Universidad Internacional del Ecuador

Escuela de Ingeniería Mecánica Automotriz

Trabajo de Integración Curricular

Artículo Investigación para la obtención del Título de Ingeniería en Mecánica Automotriz

DISEÑO DE ESTRATEGIA DE CONTROL ELECTRONICO PARA INYECTORES GDI MEDIANTE DOBLE SOLENOIDE.

Sergio Andrés Buitrón Basantes

Directora: Msc. F. Paulina Vizcaino I.

Co-Director: Msc. Gorky G. Reyes C.

Quito, 2021

CERTIFICACIÓN

Yo, Sergio Andrés Buitrón Basantes, declaro bajo juramento, que el trabajo aquí descrito es de mi autoría; que no ha sido presentado anteriormente para ningún grado o calificación profesional y que se ha consultado la bibliografía detallada.

Cedo mis derechos de propiedad intelectual a la Universidad Internacional del Ecuador, para que sea publicado y divulgado en internet, según lo establecido en la Ley de Propiedad Intelectual, reglamento y leyes.

Sergio Andrés Buitrón Basantes

Nosotros, Fernanda Paulina Vizcaino Imacaña, Guillermo Gorky Reyes Campaña, certificamos que conocemos al autor del presente trabajo siendo él responsable exclusivo tanto de su originalidad y autenticidad, como de su contenido.
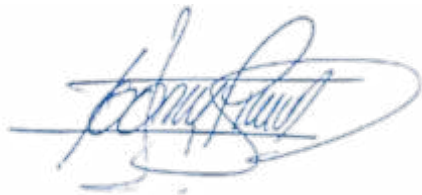
Fernanda Paulina Vizcaino Imacaña   Guillermo Gorky Reyes Campaña
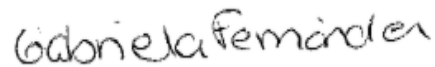
ACUERDO DE CONFIDENCIALIDAD

La Biblioteca de la Universidad Internacional del Ecuador se compromete a:

1. No divulgar, utilizar ni revelar a otros la **información confidencial** obtenida en el presente trabajo, ya sea intencionalmente o por falta de cuidado en su manejo, en forma personal o bien a través de sus empleados.

2. Manejar la **información confidencial** de la misma manera en que se maneja la información propia de carácter confidencial, la cual bajo ninguna circunstancia podrá estar por debajo de los estándares aceptables de debida diligencia y prudencia.

_____

Guillermo Gorky Reyes Campaña

Coordinador Titulación

Escuela de Ingeniería Automotriz

_____

Gabriela Fernández

Gestora Cultural

DEDICATORIA

Este trabajo nace de una pregunta, quizás la más profunda de todas: ¿Por qué? Mi vida entera se basa en responder una y otra vez esta pregunta, cuando he resuelto una no pasa mucho hasta encontrarme nuevamente con la duda. El hombre consciente de su existencia ha impuesto la técnica a la naturaleza olvidando que germinó de ella. El hombre ha olvidado que Dios creo la naturaleza. El hombre que domina la técnica solo entiende fenómenos, pero no conoce la naturaleza. Responder la pregunta ¿Por qué? No significa comprender a Dios, por el contrario, significa entender la existencia. Por eso este trabajo va dedicado a Él, Dios de toda esencia y a quien nunca se podrá preguntar ¿Por qué? Dedico este trabajo a todos los hombres que dedican sus vidas a dilucidar de entre todos los fenómenos la esencia de las cosas, quienes desean de todo corazón saber: ¿Por qué? A todos quienes intentan conocer, a ellos. A esta universidad que ha despertado mi pasión por lo que hago. Por, directa o indirectamente, formar mi carácter, mi personalidad y mis objetivos. A todas las personas que me han acompañado en este camino; lento de recorrer y tan paradójicamente breve ahora que acaba. Y aun cuando el hombre nunca llegue a conocer completamente, este trabajo es reflejo de ese intento. Indudablemente quien fui ha construido quien soy y lo que conozco hoy es el fundamento de lo que seré mañana, este trabajo lleva implícito mi vida entera.

-Sergio Andrés Buitrón Basantes

AGRADECIMIENTO

No hay palabras que expresen cuan agradecido estoy con mis papás. Todo el esfuerzo económico que debieron hacer para permitirme lograr esta meta. Gracias mi amado viejo. Gracias mi amada vieja. Entiendo que quien soy ahora es fruto de un pedazo de su vida que han dedicado a mí. Gracias UIDE por la beca económica que ustedes me han confiado. No hubiera sido posible este logro sin tal privilegio. Gracias a Dios por cada don que me ha dado. A todas las personas, sin excepción alguna, que fueron parte de este proceso gracias por formar parte de mi vida. Quiero insistir que quien redactó este trabajo, mi persona, es la suma de todas las cosas. Por esto, gracias nuevamente a todos ustedes.

Sergio Andrés Buitrón Basantes

Índice de Contenido

Índice de Tablas

Índice de Figuras

Índice de Ecuaciones

Diseño de Estrategia de Control Electrónico para Inyectores GDI Mediante Doble Solenoide.

Sergio Buitrón[1], Paulina Vizcaino[2], Gorky Reyes[3]

[1]*Estudiante Ing. Mecánica Automotriz Universidad Internacional del Ecuador*

[2-3]*Profesor Ing. Mecánica Automotriz Universidad Internacional del Ecuador*

Av. Simón Bolívar y Av. Jorge Fernández, Quito, Ecuador

[1]sebuitronba@uide.edu.ec

[2]pvizcaino@uide.edu.ec

[3]gureyesca@uide.edu.ec

Resumen-INTRODUCCIÓN: Las nuevas demandas por más estrategias de inyección que mejoren los procesos de formación de mezcla en el motor Otto se plantean desde una necesidad de potencia, eficiencia y emisiones bajas. Modelos matemáticos que describen el proceso de activación del inyector GDI han sido desarrollados. Junto con estos, nuevas estrategias que buscan reducir los retrasos entre la activación y la apertura o cierre del elemento. Tales estrategias buscan reducir tiempos de activación y la no linearidad del inyector para pulsos de inyección cortos (menores a 1,2 mS). MATERIALES Y MÉTODOS: El análisis del sistema estático de la aguja del inyector permite evaluar una nueva estrategia de activación. Los procesos físicos de este sistema se encuentran implícitos en la onda de corriente generada por la activación. Anomalías en la onda de control permitieron relacionar el sistema mecánico con el electro-magnético. RESULTADOS: Mediante un doble solenoide se disminuyó el tiempo de ascenso de corriente pico que genera la fuerza magnética para levantar la aguja del inyector. Esta estrategia prescinde de un voltaje Boost y disminuye la complejidad del circuito de control. La tasa de $\frac{\mathrm{d}i}{\mathrm{d}t}$ aumentó lo que indica una concatenación de flujo magnético mayor y más veloz. Estos resultados iniciales permiten aplicar esta estrategia en un inyector operativo.

Palabras clave- corriente pico, tiempo de activación, campo magnético, solenoide, estrategia de control.

Abstract- INTRODUCTION: The new demands for more injection strategies that improve the mixture formation processes in the Otto engine arise from a need for power, efficiency and low emissions. Mathematical models that describe the GDI injector activation process have been developed. Along with these, new strategies that seek to reduce the delays between the activation and the opening or closing of the element. Such strategies seek to reduce activation times and injector non-linearities for short injection pulses (less than 1.2 mS). MATERIALS AND METHODS: The analysis of the static system of the injector needle allows to evaluate a new activation strategy. The physical processes of this system are implicit in the current wave generated by the activation. Anomalies in the control wave allowed to relate the mechanical system with the electromagnetic one. RESULTS: By means of a double solenoid, the peak current rise time which generate the magnetic force to lift the injector needle was decreased. This strategy dispenses a Boost voltage and reduces the complexity of the control circuit. The $\frac{\mathrm{d}i}{\mathrm{d}t}$ rate increased indicating a faster and higher magnetic flux concatenation. These initial results allow this strategy to be applied to an operational injector.

Keywords- peak current, activation time, magnetic field, solenoid, control strategy**.**

1. INTRODUCCIÓN

La fase inicial de la combustión dice Heywood, está enteramente determinada por la mezcla alrededor de la bujía [1]. Muchos años antes Jovaj ya identificó que mientras más rápido se evapore el combustible mejor será la combustión [2]. Una combustión rápida y completa es sinónimo de eficiencia y bajas emisiones. La estratificación de la mezcla y la operación enteramente cualitativa del motor Otto son técnicas actuales que buscan cumplir las dos premisas anteriores. Ambas técnicas son conseguidas exclusivamente en sistemas de inyección directa. El control electrónico de los inyectores GDI ha devenido en varias estrategias de inyección en busca de aprovechar las ventajas termodinámicas de la estratificación. Pero debido a la ventana de aplicación tan estrecha que tiene esta estrategia existen muchos intentos por mejorar el comportamiento del inyector en cantidades cada vez menores que permitan expandir la formación de mezcla estratificada sin expandir el tiempo requerido. El dinamismo de operación presenta al control electrónico una carga inmensa. El estado transitorio de los motores con todas sus variables tiempo-dependientes siguen impidiendo un grado ultra refinado de operación.

El objetivo de esta investigación fue el de proponer una nueva estrategia de control electrónico para inyectores GDI a través de tres etapas. La investigación se dividió en tres etapas: un estudio preliminar de los inyectores GDI usados a nivel nacional y la identificación del control usado. Un análisis de la relación entre el sistema electromagnético y el mecánico del inyector con el que finalmente se propuso un control del actuador mediante un doble solenoide. Tal estrategia se evaluó en términos de rapidez de elevación de corriente di/dt e indirectamente en parámetros como potencia requerida y picos de voltaje. En la década de los noventa se creó el primer sistema de inyección directa a Gasolina [44] y a partir de este se desarrolló todo el fundamento y técnica que dió como resultado las tecnologías actuales [45]; todo como consecuencia de una demanda por motores más eficientes y limpios propia de la época. Los motores actuales buscan perfeccionar ambas condiciones para poder ser todavía una solución de movilidad en esta década llena de restricciones ecológicas.

Kekedjia, H., *et al* [7] trabajó en el modelado de un inyector de GLP. Afirmó lo siguiente: la fuerza magnética es inversamente proporcional a la distancia de la aguja, la aguja está en equilibrio absoluto cuando está abierta, la aguja y la armadura son equivalentes a dos masas unidas por un muelle, todo inyector es no-lineal en aperturas incompletas y una corriente baja de mantenimiento disminuye el tiempo de cierre. Yin, Z., *et al* [8] describió a la perfección las condiciones electromagnéticas que se deben cumplir para abrir un inyector GDI. Llevo a cabo esta experimentación con cámaras de alta velocidad y osciloscopios. Kumar, A. *et al* [10] comparó transistores para las demandas de alta frecuencia de conmutación y bajo consumo de potencia en los inyectores, la tecnología MOSFET fue la mejor se acopló a los requerimientos. Estos estudios respaldan el conocimiento base sobre el cual parte el estudio. Di Gaeta *et al* [12] demostró a través de un modelado matemático que es posible detectar electricamente las pulsaciones de presión no solo del inyector sino también de la bomba. Parotto, M., *et al* [5] experimentó con dos métodos para corregir variaciones de dosificación entre inyectores de mínimo 0,5 mg; una estrategia detectaba la apertura por disminución de la presión en el riel y otra detectaba el cierre a partir de la zona balística del inyector. Estos estudios demuestran que los sensores básicos detectan implícitamente todos los procesos físicos en el conjunto inyector-riel. Además, Lu, H., *et al* [13] evaluó las estrategias usadas para activar inyectores GDI y resalta sobre todo que de todos los métodos el que mejor resultados dio es el de pre-energización. Logró disminuir 26% el tiempo de apertura. Bianchi, G., M., *et al* [9] buscó disminuir el tiempo entre dos activaciones consecutivas de un inyector, a través de un modelo matemático disminuyó el intervalo en 30%. Descubrió que una parte los retrasos se originan en la elasticidad inherente del material cuando la aguja impacta. Yamakado, M., *et al* [11] propuso un inyector con dos solenoides y encontró ventajas de implementación, pero no en tiempos de apertura. Estos tres estudios demuestran los esfuerzos constantes por evitar la zona balística de los inyectores y disminuir la energía requerida para elevar o descender la aguja. Una menor cantidad de energía se traduce en una disminución de los retrasos de apertura-cierre. Este estudio reconoce las condiciones generales eléctricas de trabajo de los

inyectores para luego verificar estas en el modelo matemático y así derivar las ventajas que propone la nueva estrategia.

2. FUNDAMENTO TEÓRICO

Hay muchos enfoques posibles para estudiar los inyectores GDI y de cada uno se derivan conclusiones distintas. Si bien el componente mecánico determina, entre otras cosas, la formación del chorro [3]. Este trabajo estudia la dosificación de combustible como producto del control electrónico (Pulso eléctrico) que gestiona el sistema mecánico del inyector. Por esta razón en esta sección se expone el desarrollo cronológico de la tecnología y el estado actual de su constitución y control.

2.1 Sistemas GDI

En definitiva, hay dos momentos en la producción de sistemas GDI. Peter, H. organiza todos los sistemas existentes en dos grupos: aquellos regulados por presión y aquellos regulados por cantidad [1].

Tabla 1: versión de sistemas GDI

	Melco	Siemens 1era G.	Bosch HDP1	Bosch HDP2 DECOS	Bosch HDP5
Presión	45-53 bar	Hasta 120 bar	50-100 bar	40-120 bar	50-250 bar
Regulación LP	Mecánica	Mecánica	Mecánica	Bomba PWM	Bomba PWM
Regulación HP	Por presión	Por presión	Por presión	Por cantidad	Por cantidad
Aplicación	Mitsubishi	Renault, Opel, Mercedes Benz	VW FSI	Volvo, GM, Audi, Ford	Sistemas Turbo actuales

Fuente: Autores

El suministro de combustible de baja presión es mecánico o electrónico; este último es conocido como sistema DECOS y regula la velocidad de la bomba de combustible mediante PWM. Los sistemas regulados por presión difieren entre sí por la presencia de un retorno. Para los sistemas más actuales una válvula limita el combustible que ingresa antes que descartar un exceso. En la figura 1 se muestra la última generación de los dos principales elementos en el sistema GDI de Bosch GmbH (principal referente en tecnología).

Figura 1: Bomba HDP6 e inyector HDEV5.
Fuente: [3]

2.2 Inyección

El control electrónico de los inyectores tiene tres funciones: dosificación, adecuación y sincronización. [4]. El control debe determinar el ancho de pulso según demanda del conductor. Restringir la dosificación de acuerdo con el aire disponible y ejecutar la activación oportunamente para lograr una mezcla adecuada. Esta libertad de sincronización permite que aire y combustible interactúen entre ellos y con el entorno en periodos de tiempo diferentes. Los motores GDI usan varios factores para inducir esta interacción aire-combustible. Según el mecanismo primordial de interacción se consiguen grados diferentes de estratificación; definida como la mezcla irregular de combustible en la cámara. [5].

Tabla 2: tipos de formación de mezcla

	Dirigida por pistón	Dirigida por aire	Dirigida por chorro
Ubicación inyector	Lateral	Lateral	Central (Transversal o Longitudinal
Movimiento de la mezcla por	Cabeza de pistón	Corriente turbulenta de aire	Geometría del inyector
Desventajas	Cilindro mojado. Aumento de Hcs	Geometría compleja de admisión. Comportamiento poco predecible del aire	Depósitos de carbón en bujía. Fallas de encendido. Depósitos en inyector.

Fuente: Autores.

Solo las primeras generaciones (hasta 2005) del sistema ocupó mecanismos de mezcla por pistón o por aire [5]. Ambas estrategias producían altos niveles de HC

y nada de beneficios en cuanto a NOx y potencia. En la siguiente figura se aprecia gráficamente las diferencias entre las estrategias de formación de mezcla.
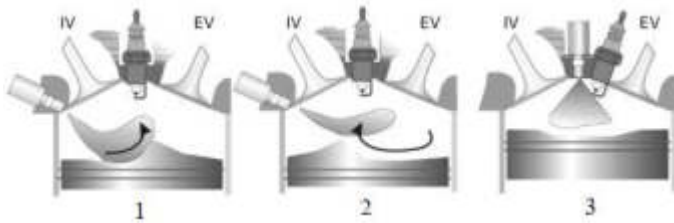

Figura 2: tipos de formación de mezcla
Fuente: [6]
1. Inyección guiada por pistón 2. Inyección guiada por aire 3. Inyección guiada por spray

2.3 Activación del Inyector

Un inyector GDI usa un circuito con doble riel de alimentación. Un voltaje boost que asegura la rápida saturación de la bobina asistido después por un voltaje nominal de batería que mantiene las condiciones estables de la bobina. Ya que se debe superar rápidamente la zona balística del inyector correspondiente al trayecto de apertura total es necesario usar el voltaje boost. El tiempo que se aplica este voltaje es constante pero monitoreado por otro factor, la corriente. Para los inyectores GDI se determinan umbrales máximos y mínimos de amperaje con los cuales está definida la apertura o cierre de la aguja. Estos valores de corriente son producto de las restricciones físicas opuesta a la aguja. En otras palabras, estos límites representan la fuerza que se necesita para mantener abierta la aguja [7]. La siguiente figura muestra un tipo de control por corriente que no usa inversión de voltaje, se regula la condición estable mediante una onda pulsada de voltaje nominal y presenta un fuerte pico de voltaje.
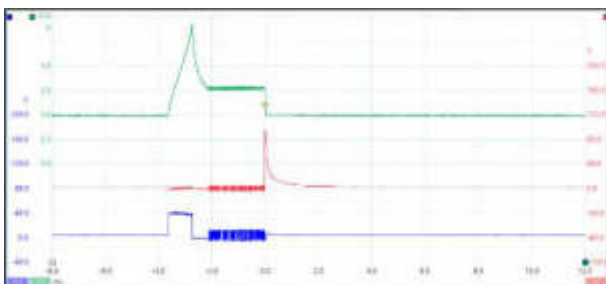

Figura 3: Onda de activación del inyector.
Fuente: [8]

2.4 Circuito de Control

En la figura 4 se observan los componentes y conexiones internas entre los elementos del circuito. Los rieles de alimentación son HV (voltaje boost) y voltaje LV (nominal). Se utilizan tres transistores mosfet para alimentar la bobina del inyector y dos para conmutar el circuito a tierra. Los transistores H1 y H2 se protegen de una polarización inversa mediante diodos. Los umbrales de corriente son monitoreados por los circuitos diferenciales ubicados en el Drain de los transistores L2 y L1. Aunque este circuito presente posibilidades de inversión de polaridad no todos los circuitos optan por estas características, tal es el caso de la Figura 3. De todas formas estos circuitos son los que menos retrasos presentan por activación y su complejidad dificulta su implementación.
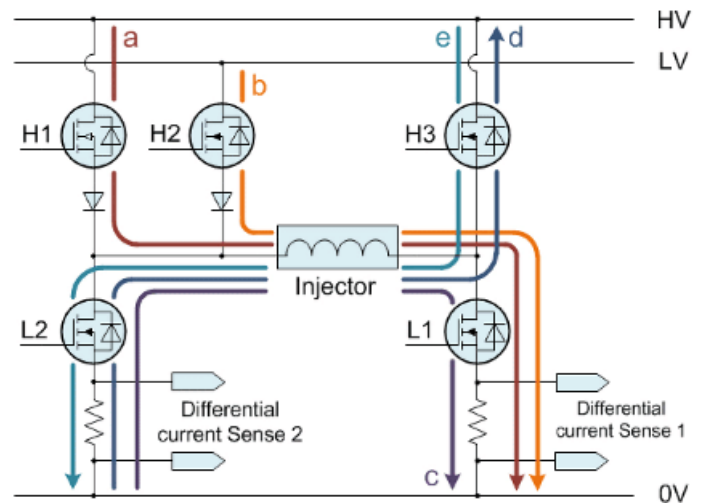

Figura 4: circuito de control inyector.
Fuente: [13]

2.5 Principios Físicos del Solenoide

El movimiento de la aguja del inyector es producto de una sistemática transformación de energía. La energía eléctrica es transformada a energía magnética y la energía magnética es transformada en energía cinética. Según la Ley de Ampere dice que un campo magnético es producto de la corriente que define una superficie [11]:

$$\oint_C H * dI = I_{contenida} \qquad \text{[Ec. 1]}$$

La intensidad de campo se entiende como la extensión de la influencia de las líneas de acción en la

zona de influencia. Su densidad de flujo es la cantidad de influencia y está relacionada por la ecuación [9]:

$$B = \mu_0 H \ [T] \qquad \text{[Ec. 2]}$$

Donde, μ_0 = permeabilidad del aire $\left[\dfrac{H}{m}\right]$

Por lo tanto un campo tiene extensión y magnitud de influencia, en otras palabras, energía. La energía magnética acumulada en el campo está definida por [11]:

$$W_m = \frac{1}{2}\int_v \ B * H \ dv \ [J] \qquad \text{[Ec. 3]}$$

Esta energía es la que generar trabajo sobre otros cuerpos ferromagnéticos como la armadura del inyector. Según Edwards, J. D., un cuerpo que se encuentre axialmente en un flujo magnético experimentará un esfuerzo de tracción [9]:

$$F = \frac{B^2}{2\mu_0} \ [N] \qquad \text{[Ec. 4]}$$

Producto de esta fuerza existe un gasto de la energía magnética que luego es reforzado cuando el cuerpo atraído concatena mayor densidad de flujo. No existe flujo de combustible hasta que los orificios se descubran y la inercia del fluido se venza (viscosidad) [25]. La densidad de campo magnético por su parte es la cantidad de flujo magnético delimitada por una sección. Defina por [11]:

$$\phi = \oint_s \ B * ds \ [Wb] \qquad \text{[Ec. 5]}$$

Dentro del solenoide, según Faraday, una fuerza electromotriz se genera en oposición al aumento o disminución del flujo. [11]:

$$\text{emf} = -\frac{d\phi}{dt} \ [V] \qquad \text{[Ec. 6]}$$

Pero este flujo es creado por la corriente que fluye en el solenoide y por lo tanto:

$$\text{emf} = -L\frac{dI}{dt} \ [V] \qquad \text{[Ec. 7]}$$

De modo que la inductancia es el termino proporcional que determina la oposición del elemento a transformar energía eléctrica en magnética y viceversa. Determinado por:

$$L = \mu_0 n^2 \pi R^2 l \ [H] \qquad \text{[Ec. 8]}$$

Donde:

$$n = \frac{N}{l}$$

R = radio

l = longitud

N = vueltas de conductor por unidad de longitud

μ_0 = permeabilidad del aire

Con este conjunto de ecuaciones se definen los procesos acaecidos en el inyector. Un voltaje aplicado EMF en el solenoide genera un flujo de corriente I que crea un campo magnético de magnitud H y con influencia B en una superficie determinada que atrae un material ferromagnético con una fuerza F.

3. MATERIALES Y MÉTODOS
3.1 Método

Después de comprender el fundamento teórico de los procesos físicos en el inyector se buscó identificar estos principios en las ondas de control de inyectores GDI usados a nivel nacional. Se tomaron en cuenta las estrategias usadas. En base a los resultados preliminares se extendió el fundamento teórico para unificar los sucesos electromagnéticos con los mecánicos. En base a estos resultados se experimentó teórica y prácticamente con la idea propuesta para observar los resultados. La posibilidad de usar un doble solenoide que permite concatenar de forma más rápida el campo magnético que levanta la aguja del inyector; siendo los dos solenoides idénticos, mientras uno permanece activado el segundo es la única razón de concatenar un campo magnético mayor y por lo tanto abre ele inyector. Dada la naturaleza de la propuesta ningún mecanismo fue usado, sino que solamente se estudió sobre una base puramente eléctrica-electrónica y puesto que los resultados son medibles, pero no observables se evaluó la eficacia de la propuesta bajo los siguientes parámetros.

La fuerza magnética generada por el solenoide tanto en la prueba base (estudio preliminar de inyectores GDI) como en la prueba de simulación, es la misma y garantiza la apertura del inyector. La variable dependiente fue la rapidez de elevación de la corriente $\frac{dI}{dt}$ influenciada por la inductancia. Se buscó disminuir el tiempo de formación del campo magnético que levanta la aguja. Esta formación es restringida por la pendiente de crecimiento de la corriente $\frac{dI}{dt}$. Por último, se evaluaron ventajas secundarias por usar un doble bobinado bajo parámetros como: potencia disipada requerida, picos de voltaje autoinducido, reducción de voltaje de alimentación y simplificación de circuito de activación.

En el estudio preliminar se midieron variables de control en inyectores GDI de tres vehículos presentes en el mercado nacional. Estas variables fueron medidas con equipo especializado. La idea principal era modificar el control electrónico únicamente por lo que se mantuvieron lineamientos o normativas aplicadas en la construcción de solenoides para inyectores con respecto a su extensión espacial, material de constitución o su caracterización. Para el banco de pruebas se usó una plataforma de software abierto y el mismo equipo de medición usado en el estudio preliminar.

3.2.1 Vehículos de Estudio

Según un pronóstico publicado en la web de Allied Market Research [26] el mercado de sistemas de inyección directa de gasolina en 2022 alcanzará un valor de $9 712 millones de dólares. Segmento de mercado que lleva creciendo desde el 2015 a un ritmo de 20,7% anual. Las empresas productoras de automóviles han encontrado en los sistemas GDI una excelente solución para el aumento de eficiencia y potencia específica de nuevos motores. Cada vez más se extiende el portafolio de vehículos con este sistema en el mercado ecuatoriano. Se escogieron tres vehículos de distinto año de fabricación. La siguiente tabla muestra las unidades estudiadas.

Tabla 3: vehículos estudiados.

	Cruze	CX3	Haval H6
Fabricante	Chevrolet	Mazda	Great Wall
Año de fabricación	2017	2018	2020
Motor	LE2	Skyactive G	GW4C20
Proveedor sistema inyección	Delphi	Denso	Delphi
Inyector usado	69746NH	PE0113250B	28483708
Imagen del inyector			

Fuente: Autores

3.2.2 Equipo de medición

Se usó un osciloscopio automotriz Hantek 1008C. El equipo tiene 2.4 MSa/s de tasa de muestreo y una precisión de tiempo base de +/- 50 ppm. Los canales tienen una sensibilidad de 10mV/div. Este equipo cuenta con las siguientes certificaciones: EN 55022:2010, EN 61000-3-2:2006+A1:2009+A2:2009, EN 61000-3-3:2008, EN 55024:2010 [27]. Este equipo se usó para medir el voltaje aplicad al inyector y comprender la estrategia de control.

Figura 5: Osciloscopio Hantek 1008C
Fuente: [27]

Corriente fue medida con una probeta amperimétrica HANTEK CC-65. La precisión de la pinza se agrega a la precisión del osciloscopio. La siguiente tabla indica la precisión según rangos de trabajo. En la figura siguiente se puede observar la ilustración del equipo. Esta se usó para medir la corriente que fluye por los solenoides y mediante la onda producida encontrar la pendiente de creación de la corriente $\frac{dI}{dt}$.

Tabla 4: precisión pinza amperimétrica.

	Atenuación		
	1mV/10mA	*1mV/100mA*	*1mV/100mA*
Rango	10mA-20A	100mA-40A	40A-65A
Precisión	+/- (1,5% +/- 5mA)	+/- (2% +/- 20mA)	+/- (4% +/- 0,3mA)

Fuente: [28]

Figura 6: Pinza amperimétrica CC-65
Fuente: [27]

Conforme a la descripción de la normativa AS40401 se asumió que los solenoides de los inyectores estudiados cumplen con ella. Contiene lineamientos generales sobre estándares de materiales o construcción de solenoides[29]. Con respecto a la caracterización de los inyectores de inyectores GDI, SAE establece métodos uniformes de prueba en banco bajo la normativa J2713 [30].

3.2.4 Equipo de control

Actualmente existen varias opciones de plataformas Open Source. Se evaluaron tres propuestas comercialmente disponibles y en la siguiente tabla se justifica la elección realizada.

Tabla 5: Opciones de plataforma Open Source

	Raspberry Pi	BeagleBone	Arduino UNO
Precio	+	+++	+
Input análogo	No	No	Sí
Output PWM	Sí	No	Sí
Interfaz propia	No	No	Sí
Aplicación usual	Comunicación serial	IoT	Control electrónico

Fuente: Autores.

Por todas las ventajas y el excelente costo el hardware usado para generar los pulsos de control según la nueva estrategia fue Arduino UNO. Este controlador usa un integrado ATmega328P. Posee seis entradas analógicas y seis entradas PWM. El código fuente es abierto y existe basta información de su programación [31]. Se usó dos entradas analógicas como parámetros de entrada para los pulsos de activación de los MOSFET que alimentan los solenoides. Para controlar estos MOSFET se usó dos salidas PWM. Estas salidas simularon un trabajo del solenoide con varias pre y post inyecciones. La siguiente figura muestra una representación del hardware descrito.
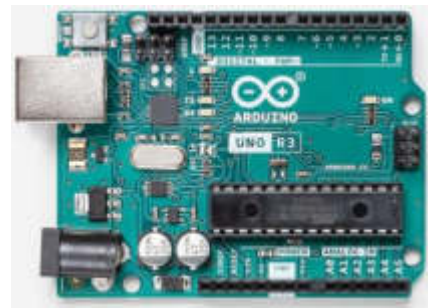

Figura 7: Arduino UNO
Fuente: [31]

4. RESULTADOS Y DISCUSIÓN
4.1 Análisis preliminar

La figura 8 muestra la onda de control del inyector 69746NH. La estrategia aplicada es una de tres etapas [33, 25]. También conocida como una estrategia de "Boost-hold" está acompañada de un control pulsado que disminuye la potencia disipada en el solenoide porque una vez levantada la aguja la mayoría de las fuerzas oponiéndose al movimiento han sido canceladas [34]. La alimentación es por doble riel y control PWM. Gracias a la inversión de polarización la energía acumulada en el solenoide se descarga rápidamente, a esto se conoce como Freewheeling.
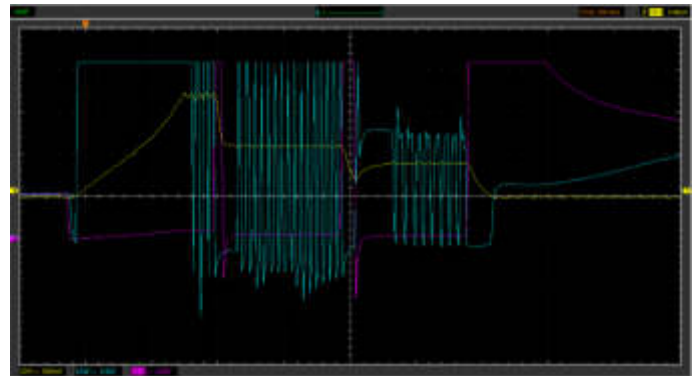

Figura 8: Control inyector 69746NH
Fuente: Autores.

Ya no se observa el uso de voltaje nominal a 2000 rpm para crear la corriente de cierre porque el Freewheeling es suficiente para cerrar el inyector y porque la mayor presión disminuye el pulso necesario, en la tabla 6 se muestra el tiempo del pulso calculado y la presión en el riel para todas las ondas analizadas. La figura 9 expone esta condición.
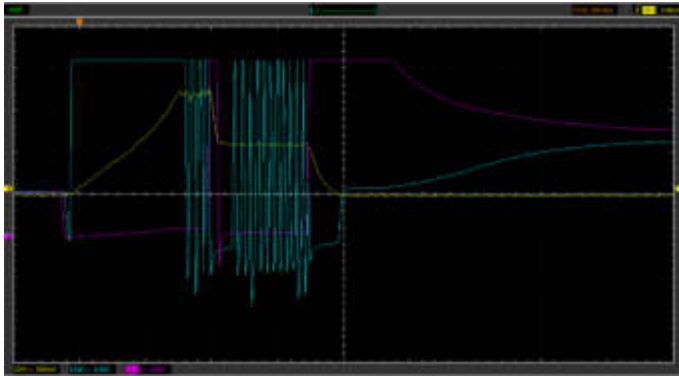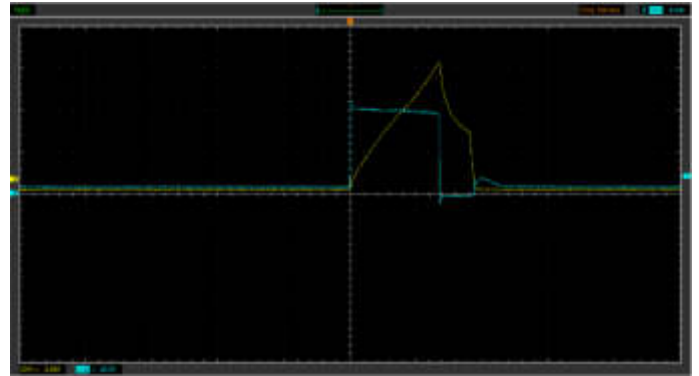
Figura 9: Control inyector 69746NH 2000rpm
Fuente: Autores.

La figura 10 muestra la onda de control del inyector PE0113250B. Esta onda de control contempla una activación sencilla sin control de apertura e integra un control anti-rebote que aumenta la precisión de cierre del actuador [35]. Al cortar la alimentación del solenoide un circuito de control de corriente calcula el momento adecuado para alimentar por un breve tiempo el solenoide con voltaje nominal. Esto crea una fuerza magnética que desacelera la aguja, disminuye la fuerza elástica de retorno ejercida por el muelle y reduce la energía del impacto contra el asiento del inyector.
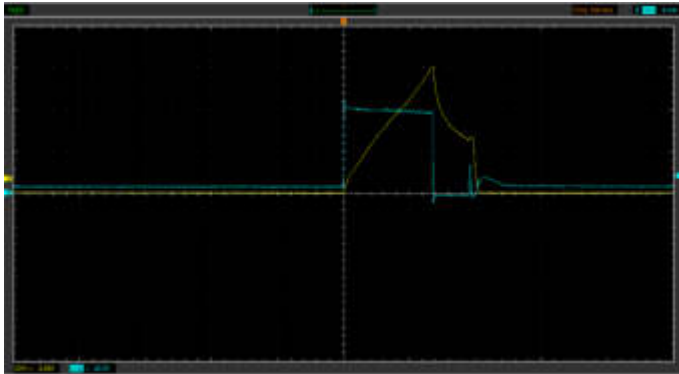

Figura 10: Control inyector PE0113250B
Fuente: Autores.

A 2000 rpm esta medida es omitida, pero en este caso particular la presión medida en el riel se ha mantenido estable como se observa en la tabla 6. Este hecho es explicado por una mayor tolerancia en combustible inyectado en regímenes de mayor carga y obsérvese también que el pulso de inyección tampoco cambió.

Figura 11: Control inyector PE0113250B 2000rpm
Fuente: Autores.

El último sistema estudiado fue el de un inyector 28483708. Esta estrategia de control tiene tres fases, pero no tan refinada como la del inyector 69746NH. No se aplica Freewheeling y por lo tanto el campo magnético se desvanece naturalmente. Para asegurar un cierre rápido asegurando un lapso de inyección fijo al final se usa un control PWM. Medida que apoya a disminuir la energía del cierre del inyector. La figura 12 muestra el proceso descrito.
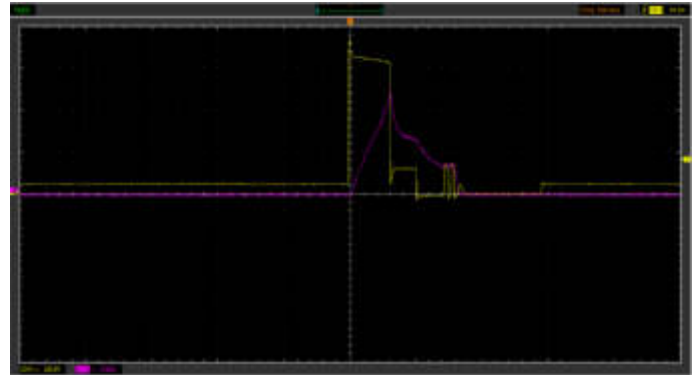

Figura 12: Control inyector 28483708
Fuente: Autores.

Y nuevamente a 2000 rpm el control pulsado el final del tiempo de inyección es omitido. La figura 13 muestra esta condición.
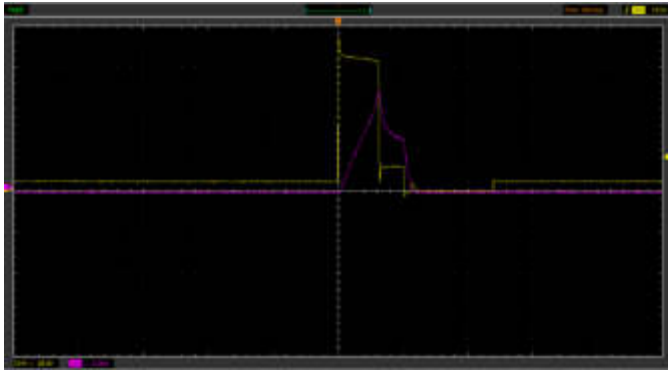
Figura 13: Control inyector 28483708 2000rpm.
Fuente: Autores.

Debido a la atenuación 20:1 usada para capturar el alto voltaje en los dos últimos vehículos (PE0113250B y 28483708) se observa ligeramente overshoot y undershoot en la señal. La tabla muestra parámetros importantes dentro de los cuales ocurrieron los pulsos de inyección.

Tabla 6: Parámetros inyección identificados.

	69746NH	PE0113250B	28483708
Método de control	Boost-hold	Simple (anti-rebote)	Boost-hold
t_{ii}	1,2 ms	1 ms	0,8 ms
P_i	30 bar	30 bar	30 bar
$t_{i2000rpm}$	0,8 ms	0,95 ms	0,6 ms
$P_{2000rpm}$	44 bar	30 bar	50 bar
I_{max}	12 A	6 A	12 A
di/dt	40 A/ms	8,57 A/ms	40 A/ms
V_{max}	64,8 v	40 v	65 v

ti:pulso inyección ralentí
P_i: Presión riel ralentí

Fuente: Autores.

El estudio estático del inyector se realizó para los inyectores de mayor consumo de corriente y específicamente se tomó en cuenta el inyector 69746NH debido a una característica presentada en la onda de control que simplificó el modelo matemático y también porque la onda de control es la que mayor complejidad electrónica presenta. Esta característica es explicada a continuación.

4.2 Doble solenoide

Además de ser evidencia del control aplicado, las ondas de control medidas en los inyectores llevan grabadas los sucesos físicos que ocurren durante su transcurso [36, 37]. Se ha visto que la inductancia media la transformación de energía eléctrica en mecánica y viceversa. Cuando la aguja del inyector se comienza a levantar, la armadura de la aguja permite condensar un mayor flujo magnético lo que permite que el solenoide deje fluir más corriente. Este suceso es estudiado en la onda con una resolución de tiempo adecuada. Para Ito, A., *et al* es necesario detectar un cambio de 20 mA durante un microsegundo para definir el comienzo de elevación de la aguja del inyector [38]. Para Wieclawski *et al* ya se puede detectar el cambio de inductancia en el solenoide por la elevación de la aguja del inyector en un marco de tiempo de 50 microsegundos [39]. Wieclawski afirma que "Aunque la disminución en corriente es ligera puede ser observada como un aplanamiento o curvatura de la corriente". La corriente disminuye porque parte de la energía eléctrica que es transformada en energía magnética es gastada en forma de energía mecánica (elevación de la aguja). Es por esta particularidad que se ha escogido la onda de control del inyector 69746NH.
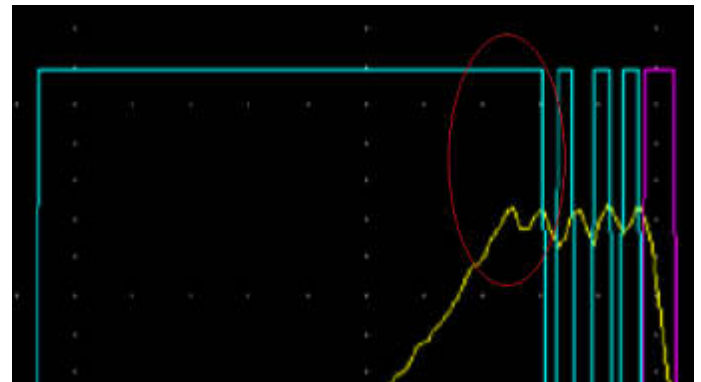

Figura 14: anomalía en control inyector 69746NH
Fuente: Autores.

En la figura 14 se observa una disminución breve de la corriente que no coincide con ningún pulso de control ni limitación de corriente del inyector 69746NH, de hecho, poco después de este suceso el control cambia a estado transitorio. La amplitud de esta desviación es de aproximadamente 50 mA y ocurre en un lapso de 20 microsegundos. Tal anomalía concuerda con las descripciones hechas en [38] y [39]. Esto quiere decir que el inyector después de 0,3 ms ha creado la fuerza magnética necesaria para vencer todas las oposiciones del sistema con una corriente de 12 amperios. A partir de esta

información se modeló la estrategia de control con dos solenoides.

De la Ecuación 7 se deduce que mientras el solenoide crea el campo magnético no hay corriente estable y por lo tanto el solenoide actúa como un cortocircuito. Al resolver la ecuación se obtiene un valor de inductancia de:

$$L = 1,62 \, m\text{H}$$

Al medir la resistencia del solenoide se encontró un valor de:

$$R = 0,6 \, \Omega$$

El solenoide presenta una longitud de 25mm y un diámetro de 20mm. El calibre del solenoide es AWG 27. La tabla 7 muestra las características de este cable.

Tabla 7: Características cable esmaltado AWG 27

Diámetro	R @ 1000 m	I_{max} continua
0,36 mm	168, 82 Ω	0,288 A

Fuente: Autores

Aunque el producto vectorial de la ley de Ampere (Ecuación 1) requiere un contorno cerrado la tangencialidad del campo concatenado en el centro del solenoide define la integral y se obtiene el siguiente resultado:

$$\text{H} \int_a^b dl_{ab} = \text{H}l_{ab} = nl_{ab}\text{I}$$

$$\text{H} = n\text{I}$$

Donde:

$n = vueltas \, por \, metro$

$$\text{H} = 33\,336 \, \frac{\text{A}}{\text{m}}$$

No se realizó un análisis metalográfico del núcleo ya que no se pudo obtener una muestra del núcleo, aún sin núcleo ferromagnético se obtuvo excelentes resultados. Remplazando los datos en la Ecuación 2, se obtuvo:

$$\text{B} = 41,9 \, m\text{T}$$

Según la Ecuación 4 la armadura que está axialmente ubicada en el solenoide experimenta una fuerza de:

$$\text{F} = 698,33 \, \text{N}$$

Esta fuerza es contraria a la fuerza elástica, la fricción y la presión del. Puesto que la posición de la perturbación en la corriente de activación del inyector siempre ocurrió en el mismo lugar se descubrió que a pesar del aumento de la presión sobre la aguja del inyector siempre actuaban las mismas fuerzas con las mismas magnitudes y sentidos. En consecuencia, los dos solenoides se calcularon para concatenar la misma intensidad de campo de 33 336 $\frac{\text{A}}{\text{m}}$. El cable usado para construir los nuevos solenoides fue AWG 24. La siguiente tabla muestra las características del conductor.

Tabla 8: Características cable esmaltado AWG 24

Diámetro	R @ 1000 m	I_{max} continua
0,51 mm	84,19 Ω	0,577 A

Fuente: Autores.

El calibre usado en el solenoide del inyector modelo soporta una corriente continua máxima de 0,288 A pero la corriente se eleva hasta 12 A. Esto es posible ya que los periodos de flujo de la corriente son muy cortos y el efecto Joule en el solenoide es mínimo. Bajo esta misma consideración se trataron los nuevos solenoides. Se mantuvo la geometría del solenoide en el inyector modelo. Una longitud de 20mm y un diámetro de 20mm. En esa distancia se alcanzan 39,17 vueltas lo que da una relación $n = 1958,5$ vueltas/m. Se usaron 2,46 m de conductor lo que dio una resistencia R = 0,21Ω. La inductancia se calculó con Ecuación 8, dio como resultado una inductancia L = 30µH. La tabla 9 resume las características del solenoide.

Tabla 9: Características solenoide construido

$\emptyset$	20mm
l	20mm
η	1958,5
L	30µH
R	0,21Ω

Fuente: Autores.

La nueva estrategia de control consistió en lo siguiente. Según la ley de Ampere el campo magnético creado por un conductor es proporcional a la superficie que la corriente delimite con su circulación. Esto quiere decir que la sumatoria de las corrientes del mismo sentido en un contorno definido serán las que determinen la intensidad de campo magnético, sin importar su origen. Esta condición permite distribuir la corriente total en varios solenoides que generan la misma intensidad de

campo calculado H en el inyector modelo. En la figura 8 se observa una corriente de mantenimiento de 6 A. Luego de ser levantada la aguja es necesaria esta magnitud de corriente para mantener el movimiento de la aguja. Esto quiere decir que sin el campo magnético generado por esta corriente el inyector se cierra. La intensidad de campo en esta condición es de $H = 16\,668\,\frac{A}{m}$. El primer solenoide crea esta intensidad de campo que no abre el inyector. El segundo solenoide crea la intensidad de campo restante de $H = 33\,336\,\frac{A}{m}$ que vence todas las fuerzas sobre la aguja en reposo. Los dos solenoides son idénticos en constitución y tienen las siguientes características. La constante de tiempo τ es igual a:

$$\tau = \frac{L}{R} = 142{,}86\ \mu S$$

Con esta constante se calcula la corriente máxima que maneja el solenoide y se determina la duración de los pulsos de activación. En la Ecuación 7 se aumentó la resistencia para completar el modelo físico de Kirchoff para un inductor. De este modelo se obtiene la solución a la ecuación diferencial de la forma:

$$I_{(t)} = \frac{V}{R}\left(1 - e^{-t\left(\frac{R}{L}\right)}\right) \qquad \text{[Ec. 9]}$$

De la Ecuación 9 se obtiene una corriente máxima de $I_{142{,}86\ \mu S} = 36$ A. Esta corriente es enorme para ser soportada continuamente como ya se explicó anteriormente. Para que el primer solenoide genere una intensidad de campo magnético $H = 16\,668\,\frac{A}{m}$ necesita una corriente de:

$$I = \frac{H}{\eta} = 8{,}51\ A$$

Para calcular la corriente necesaria se despeja t de la Ecuación 9, como resultado la siguiente ecuación:

$$t = \frac{L}{R}\left(\ln\frac{V}{V - I_{(t)}*R}\right) \qquad \text{[Ec. 10]}$$

Entonces el pulso de activación para lograr 8,51 A es de:

$$t = 23{,}03\mu S$$

El resto de intensidad de campo lo desarrolla el segundo solenoide. Esta intensidad es de la misma magnitud y por lo tanto requiere de los mismos parámetros de control, $I = 8{,}51\ A$ y $t = 23{,}03\mu S$. El segundo solenoide se activa mientras ocurre el pulso del primero. El tiempo total teórico para alcanzar la corriente máxima es de:

$$t = 46{,}06\mu S$$

En este tiempo la aguja del inyector en teoría ya se levanta. Después de este suceso hay que mantener la corriente de 8,5 A tanto como sea requerido por el comando de inyección. Ya que este trabajo no contempla el estado transitorio del inyector mayor investigación se necesita para modelar el tiempo que tarda la aguja en lograr su carrera máxima y el tiempo que tarda en comenzar el movimiento del fluido (inyección de combustible). Por este motivo se escogió arbitrariamente un tiempo máximo de 0,1ms. Luego de que el primer solenoide es activado se usa una señal PWM de 50% a 50kH para mantener la corriente entre 8-8,5 A. El segundo solenoide ya no está activado.

4.3 Programación de control

Arduino se programó para controlar dos MOSFET de potencia de acuerdo con el siguiente circuito mostrado en la figura 15.
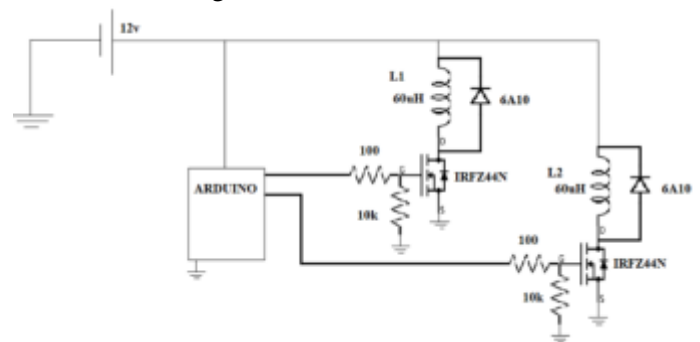

Figura 15: Circuito de control
Fuente: Autores.

Arduino activa L1 durante $46\mu S$. A los $23\mu S$ L2 es activado por los $23\mu S$ restantes. L1 y L2 son concéntricos y comparten el mismo medio permeable. Después de los $46\mu S$ L2 es desactivado y Arduino espera $10\mu S$ para activa L1 con la señal pulsada de 50kH al 50% de ciclo de trabajo. Se produce esta señal hasta completar $100\mu S$.

4.4 Resultados

La siguiente figura muestra la onda de control producida por Arduino con la programación planteada.
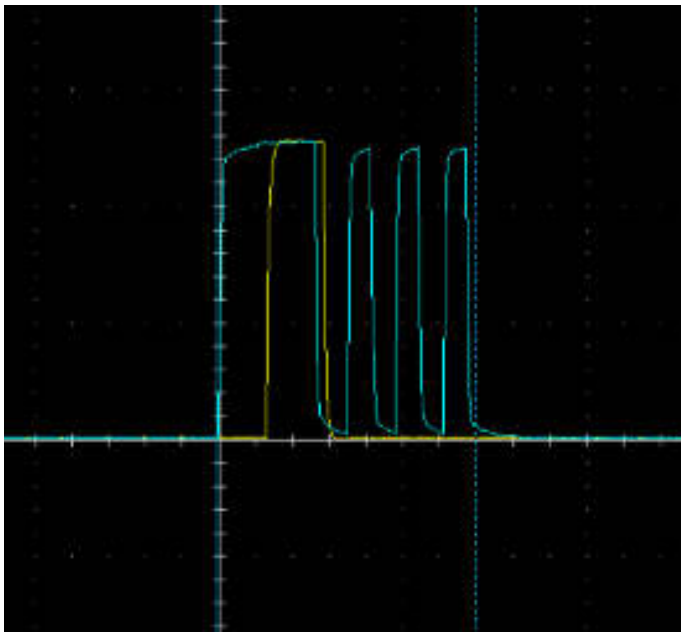
Figura 16: Control de arduino.
Fuente: Autores.

En Arduino se usaron los pines 9 y 10 (Onda azul y onda amarilla respectivamente). Se observa claramente el periodo durante el cual ambas ondas se superponen, en este momento se genera la activación de los dos solenoides. Estas ondas de control producen el siguiente comportamiento en el lado de potencia del transistor. La figura 17 muestra los resultados finales.
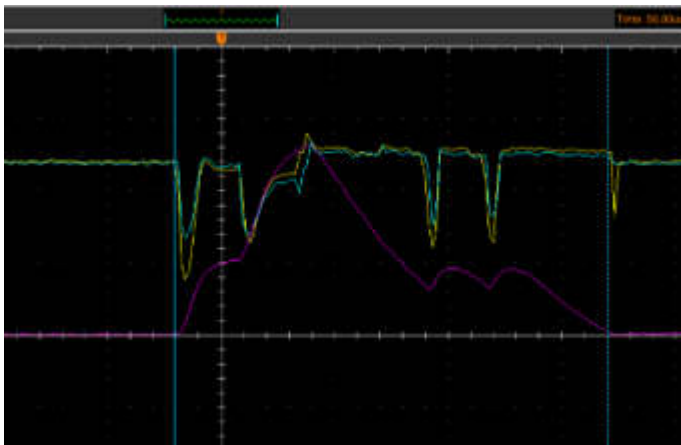

Figura 17: Control de doble solenoide.
Fuente: Autores.

La línea violeta indica la onda de corriente generada. Se observa un pico de 5,2 A. A continuación, una zona de mantenimiento y de baja intensidad. Los picos de voltaje sin necesidad de inversión de la alimentación son de +/- 33% del voltaje de alimentación. La línea amarilla es el voltaje medido en el Drain del transistor T1. La línea azul es el voltaje medido en el Drain del transistor T2. El pico de corriente no es el calculado debido a un método de conmutación muy básico y por variación en la inductancia de los solenoides.

4.5 Análisis de resultados

Los resultados de corriente son distintos a los calculados como resultado de que: la baja corriente manejada por Arduino (40 mA máx.) no alcanza a abrir totalmente el canal de conducción en el MOSFET y se produce una caída de voltaje máxima de 8v que es en efecto el voltaje real aplicado al solenoide. Se requiere un driver de MOSFET y mayor refinación del circuito para aumentar la rapidez de conmutación (corriente suministrada al GATE del transistor entre 2 y 2,5 A) [40, 41, 42, 43]. Como el voltaje es menor al calculado existe una saturación anticipada del solenoide y por ende las zonas planas de voltaje observadas antes de que termine el pulso de control. Por otra parte, existe una mayor conducción de corriente en el segundo solenoide debido al cambio de reluctancia en el circuito magnético producto del campo magnético del primero y esto permite variar la inductancia del segundo. De lo contrario se esperaría una curva similar a la primera, importante notar que se ha logrado esto sin un material permeable como núcleo de ambos solenoides. Se espera que un material ferromagnético refuerce mucho más este comportamiento. Los dos solenoides al ser de iguales características presentan las mismas magnitudes de oposición a la creación de un campo magnético. Se observa esto en la imitación del voltaje en el Drain de T2 (línea azul) tanto cuando se abre el canal de conducción en T1 y cuando el solenoide crea una diferencia de potencial al aumentar la corriente. El pulso total de activación es de 176 μS y la corriente máxima se consigue en 60 μS. Tomando una línea recta entre el origen y el valor máximo de corriente se obtiene un $\frac{di}{dt} = 86{,}67 \frac{\text{A}}{\text{ms}}$.

En la siguiente tabla se resumen los cambios conseguidos gracias a la estrategia y se compara con el inyector modelo escogido.

Tabla 10: Comparativa de solenoides.

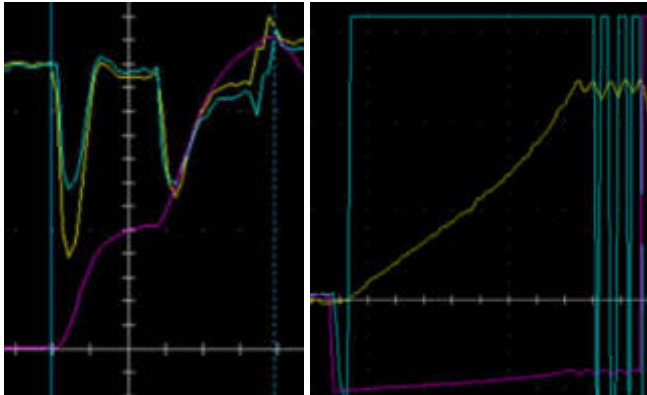|	Solenoide inyector: 69746NH	Doble solenoide
V alimentación	64,8 v	12 v
t para Imáx	0,3 mS	0,06 mS
Imáx	12 A	5,2 A
di/dt	40 A/mS	86,67 A/mS

Fuente: Autores.

Figura 18: comparativa *di/dt* entre doble solenoide y solenoide inyector 69746NH.
Fuente: Autores.

En la Figura 18 a la izquierda con línea violeta la corriente producida por la estrategia de doble solenoide a la derecha con línea amarilla la corriente del solenoide en el inyector 69746NH. Esta comparación visual del aumento inicial de corriente de activación con los valores de la tabla 10 confirma que, aunque no se hayan obtenido los valores de corriente calculados por imperfecciones en el circuito de activación o un cálculo erróneo de inductancia y por lo tanto una constante de tiempo τ equivocada, la estrategia sí permite conseguir un aumento de corriente rápido sin necesidad de un voltaje Boost. Esta rapidez de aumento permite concatenar más rápido el campo magnético que abre el inyector. Lo que permitiría disminuir el retraso entre la alimentación del inyector y el levantamiento de la aguja. En consecuencia, el comportamiento no-lineal del inyector en tiempos de inyección menores a 1,2 mS puede ser mejorado.

5. CONCLUSIONES

Varias ondas de control para inyectores GDI presentes a nivel nacional se analizaron. Se reconoció métodos propuestos en literatura que aumentan la precisión de cierre o rapidez de apertura. Dentro de la onda de corriente se encuentran implícitos los procesos mecánicos que ocurren en el inyector. Se encontró que no existe variación de la fuerza sobre la aguja del inyector proporcional al aumento de presión. El tiempo de retraso de elevación de la aguja en un inyector siempre es el mismo. Una anomalía en una onda de control compleja del inyector 69746NH permitió evaluar las condiciones eléctricas-magnéticas que son requeridas para elevar la aguja. A través de un doble solenoide se disminuyó el tiempo requerido para aumentar la corriente hasta un valor pico. Los resultados fueron buenos, aunque no corroboran los cálculos teóricos: un primer solenoide que genere un campo de mantención no solo elimina la necesidad de alimentación con doble riel y voltaje elevado, sino que también permite una mayor concatenación de campo magnético debido a la disminución de reluctancia en el circuito magnético. El segundo solenoide es capaz de conducir un mayor amperaje debido a una mayor inductancia producto de la mayor reluctancia. Refinación en el circuito de control es necesaria para alcanzar las cotas de corriente que cumplen las condiciones eléctricas-magnéticas definidas. Sin embargo, el tiempo de creación del pico de corriente teóricamente indica una disminución del 200% esto significa que sin un circuito de control complejo y una estrategia de control de corriente que requiera de un voltaje Boost sí se ha conseguido mejorar las características de activación eléctrica. La aplicación de este método en un inyector operativo es esencial para corroborar este resultado.
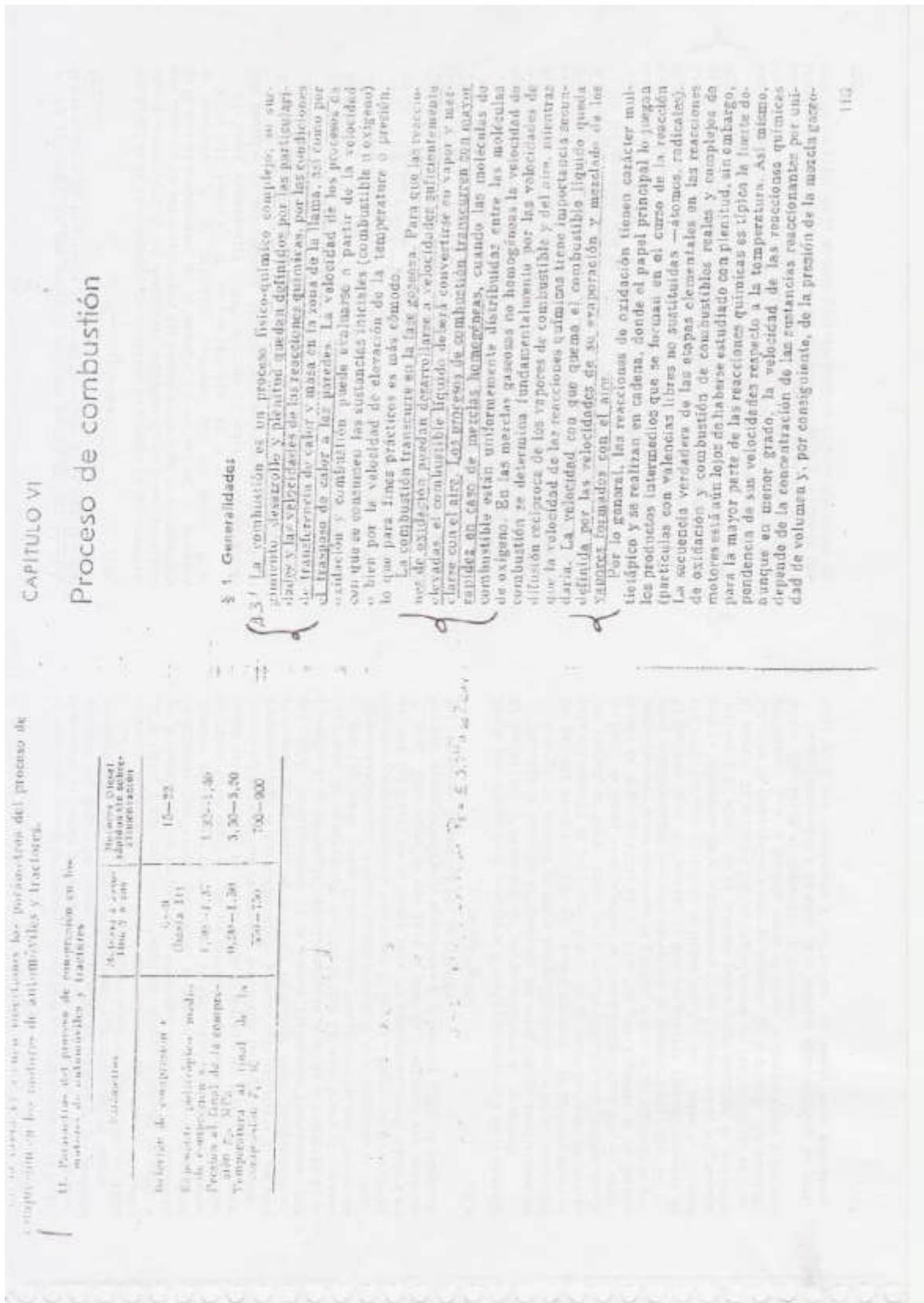
REFERENCIAS

[1] Heywood, J., B., *Internal Combustion Engines Fundamentals*, 1era ed., Duffy, A., Morris, J., New York, Chicago, San Francisco, Atenas, Londres, Madrid, México, Milán, Nueva Delhi, Singapur, Sydney, Toronto: McGraw-Hill Education, 1988.

[2] Jovaj, M., S., *Motores de automóvil*, 1era ed., traducción de Ludeña, A., Moscú, URSS: editorial MIR, 1982.

[3] Ando, H., Arcoumanis, C., *Flow and Combustion in Reciprocating Engines*, Arcoumanis, C.,

Kamimoto, T., Berlin, Heidelberg: Springer/Verlag, 2008.

[4] Denton, T., *Advanced Automotive Fault Diagnostics*, 4ta ed., Institute of the Motor Industry, New York: Routledge, 2017.

[5] Parotto, M., Sgatti, S., Sensi, F., "Advanced GDI Injector Control with Extended Dynamic Range," SAE International, doi: 10.4271/2013-01-0258, 2013.

[6] Subic, A., Cvetkovic, D., "Virtual Design and Development of Compact Fast-acting Fuel Injector Solenoid Actuator," Int. J. Vehicle Design, Vol. 46, No. 3, pp. 309-327, 2008.

[7] Kekedjian, H., Krepec, T., "Further Development of Solenoid Operated Gas Injectors with Fast Opening and Closing," SAE Technical Paper 940450, https://doi.org/10.4271/940450, 1994.

[8] Yin, Z., Jing, X., Fu, T., Zhou, H. et al., "Study on Dynamic Characteristics of High-Speed Solenoid Injector by Means of Contactless Measurement," SAE Technical Paper 2017-01-2113, doi: 10.4217/2017-01-2313, 2013.

[9] Bianchi, G., Pelloni, P., Filicori, F., Vannini, G., "Optimization of the Solenoid Valve Behavior in Common-Rail Injection Systems," SAE Technical Paper 2000-01-2042, doi: 10.4271/2000-01-2042, 2000

[10] Kumar, A., Banerjee, N., Mohan, N., Ashok B., "Study of the injector drive circuit for a high pressure GDI injector," 2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE), pp. 1-6, doi: 10.1109/ICRAIE.2016.7939562, 2016.

[11] Yamakado, M., Kadomukai, Y., Abe, M., Kubo, H., Hamada, Y., "Design and Development of a Battery-Voltage-Driven Fuel Injector for Direct-Injection Gasoline Engines," 10.4271/2001-01-1348, 2001.

[12] di Gaeta, A., Fiengo, G., Palladino, A., Giglio, V., "A control model of a Common-Rail System for Gasoline Direct Injection Engine," Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference, pp. 6614-6619, doi: 10.1109/CDC.2009.5400211, 2009.

[13] Lu, H., Deng, J., Hu, Z., Wu, Z. et al., "Impact of Control Methods on Dynamic Characteristics of high-Speed Solenoid Injector," SAE Int. J. Engines, Vol. 7, No. 3, pp. 1155-1164, doi: 10.4271/2014-01-1445, 2014.

[14] Peter, H., Benzin-Direkt-Einspritzsysteme: Komponenten, Funktionen, Diagnose., 1era ed., Krafthand Medien GmbH, Schätzl Druck & Medien, Donauwörth, Alemania, 2015.

[15] Pauer, T., Yilmaz, H., Zumbrägel, J. et al., "New Generation Bosch Gasoline Direct-injection Systems." MTZ Worldwide, vol. 78, pp. 16–23, https://doi.org/10.1007/s38313-017-0053-6, (2017).

[16] (2021) Bosch Mobility Solutions, [Online], disponible en: https://www.bosch-mobility-solutions.com/en/

[17] Reif, K., Automobilelektronik: eine Einfürungfür Ingenieure, 3ra ed., Vieweg &Teuber, GWV Fachverlage GmbH, Wiesbaden, Alemania, 2009.

[18] Reif, K., Grundlagen Fahrzeug-und Motortechnik, Bosch Fachinformation Automobil, Springer Vieweg, Friedrichschafen, Alemania, 2017.

[19] U. Spicher, T. Heidenreich, 2-Stratified-charge Combustion in Direct Injection Gasoline Engines, pp. 20-44, Hua Zhao, Advanced Direct Injection Combustion Engine Technologies and Development, Vol. 1, Woodhead Publishing, 2010.

[20] Denk, F., Elektromagnetische Einspritzventile für Ottomotoren, Universität Ilmenau, Münster, Alemania, 2018.

[21] (2021) Pico Technology [Online], disponible en: https://www.picoauto.com/library/automotive-guided-tests/gasoline-direct-injection-injector-voltage-and-current/

[22] Edwards, J., D., *Electrical Machines and Drive,* 1era ed., Macmillan Education, Londres, Inglaterra, 1991.

[23] Feyman, R., Leighton, R., Sands, M., *"The Feyman's Lecture on Physics: Mainly Electromagnetism and Matter,"* 1era ed., Wesley Publishing, Palo Alto, USA, 1964.

[24] Ida, N., *"Engineering Electromagnetics,"* 1era ed., Springer Science & Business Media, New York, USA, 2000.

[25] K., Christ, "Kalibrierung von Magnet-Injektoren für Benzin-Direkteinspritzsysteme mittels Körperschall." Dr. Ing. Thesis, Karlsruher Instituts für Technologie (KIT), Karlsruher, Alemania, 2011.

[26] (2007) Singh Pundir, A. [online]. Disponible: https://www.alliedmarketresearch.com/gasoline-direct-injection-system-gdi-market

[27] (2021) Hantek Electronic Co., Ltd [online]. http://www.hantek.com/products/detail/13170

[28] "Hantek CC-65 current clamp data sheet," Hantek Electronic Co., Ltd, Qingdao, China.

[29] SAE International. (2013). [Online]. Disponible: https://doi.org/10.4271/AS40401

[30] SAE International. (2018). [Online]. Disponible: https://doi.org/10.4271/J2713_201808

[31] (2021) Arduino Official Store. [Online]. Disponible: https://store.arduino.cc/usa/arduino-uno-rev3

[32] Stewart, M., Surface Production Operations, 4 - Reciprocating pumps, Pg. 311-414, Gulf Professional Publishing, 2019.

[33] Zhang, X., Palazzolo, A., Kweon, C., Thomas, E. et al., "Direct Fuel Injector Power Drive System Optimization," SAE *Int. J. Engines* 7(3): 2014, doi: 10.4271/2014-01-1442.

[34] Tutaj, J., Fijalkowski, B., "A new Fuel-Injection Mechatronic Control Method for Direct-Injection Interal Combustion Engines" Sciendo, doi: 10.2478/ama-2018-0042, 2018.

[35] Glasmachers H., Melbert, J., Koch, A., "Sensorless Movement Control of Solenoid Fuel Injectors," SAE Technical Paper Series, *New SI Engine and Component Design 2006* No. 2006-01-0407, 2006.

[36] Tagliatela Scafati, F., Pirozzi, F., Cannavacciuolo, S., Alloca, L. et al., "Real Time Control of GDI Fuel Injection During Ballistic Operation Mode," SAE Technical Paper 2015-24-2428, doi: 10.4217/2015-24-2428, 2015.

[37] Cvetkovic, D., Cosic, I., Subic, A., "Improved Performance of Electromagnetic Fuel Injector Solenoid Actuator Using a Modelling Approach," *International Journal of Applied Electromagnetism and Mechanic*, Vol. 27, pg. 251-273, IOS Press, Australia, 2008.

[39] Ito, A., Kawano, M., Fujita, S., "ECU Structure Strategy to Detect Lift Timing of GDI Solenoid Injectors with High Precision," SAE Technical Paper 2017-01-1628, doi: 10:4217/2017-01-1628, 2017.

[39] Wieclawski, K., Maczak, J., Szczurowski, K., "Electric Current Waveform of the Injector as A Source of Diagnostic Information," *Journal of Sensors*, Vol. 20, 4151, doi: 10.3390/s20154151, MDPI, Polonia, 2020.

[40] Barkhordarian, V., "Power MOSFET Basics," International Rectifier, Application note, El Segundo, California, S.F.

[41] Havanur, S., "Power MOSFET Basics: understanding the Turn/On Process," Vishay Siliconix, Application note AN850, 2015.

[42] "IRFZ44N data sheet," International Rectifier, California, USA.

[43] "IR2110 data sheet," International Rectifier, California, USA.

[44] Iwamoto, Y., Noma, K., Nakayama, O., Yamauchi, T., Ando, H., "Development of Gasoline Direct Injection Engine," SAE Technical Paper 970541, DOI: https://doi.org/10.4271/970541

[45] Fu-Quan, Z., Ming-Chia, L., "A Review of Mixture Preparation and Combustion Control Strategies for Spark-Ignited Direct-Injection Gasoline Engines," SAE Technical Paper 970627, DOI: https://doi.org/10.4271/970627

6. ANEXOS

[1]   Jovaj, M., S., *Motores de automóvil*, 1era ed., traducción de Ludeña, A., Moscú, URSS: editorial MIR, 1982.

La llama, que se está propagando desde el sitio de la chispa en la bujía, representa la fuente del surgimiento de la detonación y de otras alteraciones del desarrollo normal del proceso de trabajo en los motores de encendido por chispa que funcionan con combustible ligero.

La autoinflamación es el resultado final del desarrollo de toda una serie de reacciones en cadena que tienen lugar delante del frente de la llama y que van acompañadas del desprendimiento de calor, y de la formación de productos intermedios debidos a la oxidación incompleta (peróxidos, aldehídos, alcoholes, etc.). Este fenómeno aparece en aquellos casos cuando la velocidad de desprendimiento de calor, como resultado de las reacciones químicas, supera la velocidad del calor que se extrae desde el sistema en reacción. Cuanto mayor sea la velocidad de reacción y su efecto térmico tanto más rápidamente surge la autoinflamación y tanto más corto será el periodo de retardo τ_i, es decir, el intervalo de tiempo desde la culminación del rápido calentamiento de la mezcla carburante homogénea, por ejemplo, al comprimirla adiabáticamente, o desde el comienzo de la inyección del combustible líquido al aire caliente hasta el instante en que aparece la llama. Manteniendo invariables las condiciones de termotransferencia, la duración del retraso de la inflamación es inversamente proporcional a la velocidad de las reacciones que transcurren delante del frente de la llama para los correspondientes valores de T_z y p_z. Esto significa que para el caso de reacciones con cinética normal, cuyas velocidades se someten a la función (166), la duración del retraso se determina por la relación

$$\tau_i = B p^{-n} e^{E/RT},$$ (167)

donde los valores de la constante B, del exponente n y de la energía de activación E dependen de las propiedades de la mezcla.

Durante la llamada inflamación monoetápica en condiciones de $T = const.$ la presión se eleva insignificantemente en el transcurso del período de retraso (fig. 49, d). Sin embargo, en muchos casos la autoinflamación de los combustibles de hidrocarburos transcurre en dos etapas, es decir, antes de la propia inflamación (formación de los focos de llama viva) aparece la denominada llama fría. El incremento de la temperatura en la llama fría puede oscilar en los límites desde decenas hasta centenas de grados, a lo que corresponde la primera relativamente pequeña elevación de presión Δp_{fr} (fig. 49, b). El período total del retraso τ_i está constituido respectivamente de dos períodos del retraso de la llama fría τ_1 y del retraso τ_2. En este caso la influencia de la temperatura sobre los retrasos τ_1 y τ_2 resulta inversa y el efecto de la presión sobre ellos también es distinto. El retraso de la llama fría τ_1 rápidamente decrece al elevar la temperatura de compresión y casi no depende de la presión. La duración del retraso τ_2, a la inversa, crece al elevar la temperatura y bruscamente disminuye al aumentar la presión.

Como consecuencia de esto, para muchos combustibles, en particular para los parafínicos y nafténicos, en cierta zona suficiente-

---

56. La velocidad de una reacción química simple en función de la concentración y de la presión puede ser expresada de la siguiente manera:

$$\frac{dc_i}{d\tau} = A p^{m_i} e^{-E_i/RT},$$ (166)

donde $dc_i/d\tau$ es la velocidad de variación en tiempo de la concentración de cualquiera de las sustancias iniciales; A, un factor constante que depende de las propiedades del combustible y de la composición de la mezcla carburante; E, la energía de activación; m_i, el orden de reacción; R, la constante universal de los gases; $R = 8,314$ kJ/(kmol·K), T, la temperatura en K.

El factor exponencial $e^{-E/RT}$ expresa la fracción de moléculas con energías de movimiento térmico que sobrepasan cierto valor de la energía de activación E indispensable para vencer la barrera energética de una reacción elemental (rotura de los enlaces moleculares internos existentes sustituyéndolos por otros nuevos).

En caso de reacciones químicas complejas, la energía de activación es una magnitud convencional, determinada por las etapas de transformación más lenta.

57. Variación relativa de la velocidad de reacción en función de la temperatura

Energías de activación E (kJ/kmol)	Temperatura K				
	300	1000	1500	2000	
84 000	1	$6 \cdot 10^5$	$4,56 \cdot 10^6$	$4 \cdot 10^6$	$2 \cdot 10^6$
168 000	1	$3 \cdot 10^{11}$	10^{13}	$8 \cdot 10^{12}$	$2 \cdot 10^{12}$

Nota. Las velocidades de reacción convencionalmente se han adoptado igual a la unidad para $T = 300$ K.

que son las que limitan la velocidad de transformación de las reacciones de oxidación y combustión de conjunto. Para la mayoría de las reacciones de oxidación convencional (efectivo) $E_{ef} = (8,4 ... 16,8) \cdot 10^4$ kJ/kmol. Esto conduce a que las velocidades de reacción sean extremadamente pequeñas a bajas temperaturas pero con el crecimiento de éstas las velocidades de reacción con rapidez aumentan, además, tanto más aceleradamente, cuanto mayor sea la energía de activación y menor la temperatura (tabla 12).

En los procesos que se efectúan en los motores de combustión interna las velocidades de reacción que se desarrollan en la mezcla carburante, calentada al ser comprimida hacia temperaturas relativamente bajas (menos de 1000 K), son las que determinan las condiciones de la autoinflamación o del surgimiento de la detonación.

La velocidad de propagación de la llama depende de las velocidades de las reacciones químicas a elevadas temperaturas (1500 K y más), cercanas a las temperaturas finales de la combustión.

Autoinflamación. La ignición de los combustibles líquidos pulverizados que se inyectan en el aire caliente tiene importancia decisiva para los motores Diesel, pues determina no sólo el comienzo de la combustión, sino que también las particularidades del desarrollo ulterior del proceso de trabajo. La autoinflamación de la parte que todavía no arde de la mezcla aire-combustible delante del frente de

[Página 124]

turbulenta) hasta velocidades supersónicas, cuando la autoinflamación se propaga conjuntamente con el frente de la onda de choque, es decir, tienen lugar fenómenos similares a la propagación de la onda detonante.

Las ondas de choque surgen en la cámara de combustión como resultado del desarrollo rápido de las etapas finales de las reacciones que transcurren delante del frente de la llama en los focos de autoinflamación. La compresión adicional y el calentamiento del gas, en el frente de las ondas de choque que surgieron, estimula la aparición del autoencendido de aquellas porciones de la mezcla en las cuales las reacciones mencionadas también están cerca de su culminación, es decir, están a su vez lo suficientemente preparadas para la autoinflamación. Dicho proceso de culminación de la combustión, co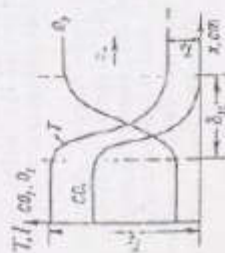mo resultado de la autoinflamación de una parte de la mezcla de trabajo, que va acompañado de la aparición de ondas de choque que a su vez estimulan la autoinflamación del resto de la mezcla aún no quemada, es lo que constituye la esencia fundamental de la combustión detonante en los motores.

§22 Propagación de la llama. A diferencia de los procesos de autoinflamación, que se caracterizan por la aceleración de las reacciones químicas a causa del incremento de la concentración de los centros activos y de la acumulación de calor en el sistema reaccionante, la llama se propaga debido a la transmisión de calor y a la difusión de los centros activos de la mezcla quemada hacia la mezcla fresca situada al lado. El frente de la llama que se propaga en una mezcla carburante inmóvil o en movimiento laminar, representa una estrecha zona que separa la mezcla fresca, cuya temperatura es T_0, de los productos de combustión, que tienen la temperatura T. En el comienzo de esta zona (fig. 50) la mezcla se calienta parcialmente por conducción, pero en mayor medida como resultado de la difusión de los productos calientes en ella desde la zona de la llama. En vista de que las reacciones químicas pueden alcanzar elevadas velocidades sólo cuando tienen suficientemente altas temperaturas, la zona de las transformaciones químicas en el frente de la llama ocupa solamente una pequeña parte del grosor total δ_f, allí donde las temperaturas son ya cercanas a las finales de la combustión T. En esta zona arde la mezcla altamente diluida (a cuenta del intercambio difusivo) con los productos de combustión completa o parcial que poseen elevada actividad química, lo que contribuye al rápido desarrollo de las reacciones químicas en cadena en medida no menor que la elevación de la temperatura. Semejante mecanismo cadenario-térmico de la

Fig. 50. Variación aproximada de la temperatura y de la concentración de las sustancias iniciales y de los productos de combustión en el frente de una llama laminar

[Página 125]

transformaciones químicas es característico para todos los procesos de combustión.

La velocidad con que el frente de combustión laminar (por capas) se desplaza con respecto a la mezcla (hecha en dirección perpendicular (normal) a su superficie se denomina *velocidad normal de la llama* u_n o *velocidad de combustión laminar* u_n. En vista de que este velocidad se determina por dos factores: la velocidad de los procesos de intercambio (transmisión de calor y difusión en el frente de llama) y la velocidad normal de la llama en función de los dos factores indicados puede aproximadamente expresarse por la fórmula

$$u_n \sim \sqrt{\chi \cdot w_r}, \qquad (158)$$

donde $\chi = \dfrac{\lambda}{c_p \rho}$ es el coeficiente de combustión de la temperatura de la mezcla carburante, que se adopta igual al coeficiente de difusión; w_r, la velocidad de reacción a la temperatura cercana a la máxima temperatura de la llama T.

El grosor del frente laminar de la llama δ_1 es fino. Por ejemplo, a la presión atmosférica en una mezcla aire-metano de composición estequiométrica $\delta_1 \approx 1$ mm, además la mayor parte (aproximadamente el 80%) corresponde a la zona de calentamiento y sólo el 20% a la zona de las propias transformaciones químicas. En la mezcla que arde más rápidamente $2H_2 + O_2$, $\delta_1 \approx 0,1$ mm. La magnitud δ_1 queda definida por los dos mismos factores que determinan la velocidad de propagación de la llama, aunque la influencia relativa de las velocidades de intercambio de calor y masa y de las velocidades de reacción en este caso resulta opuesta, a saber:

$$\delta_1 = \text{const} \, \frac{\chi}{u_n} \approx \sqrt{\frac{\chi}{w_r}}. \qquad (169)$$

Comparando las expresiones (168) y (169), observamos que al aumentar la velocidad de reacción w_r y, simultáneamente disminuye el grosor del frente de la llama. Al aumentar los coeficientes de conducción de la temperatura o de la difusión también crece u_n, pero el frente de llama se hace más extenso.

En caso de combustibles de hidrocarburos las velocidades máximas de propagación de la llama las tienen las mezclas un tanto enriquecidas ($\alpha = 0,85 \ldots 0,90$), para las cuales se consiguen las mayores velocidades de reacción y valores más elevados (en $3-3,5\%$), que siendo $\alpha = 1$ del coeficiente de variación molecular en el proceso de combustión. Al empobrecer o enriquecer aún más la mezcla, las velocidades de propagación de la llama suavemente disminuyen (fig. 51).

Los máximos valores de u_n, a presión atmosférica y $t_0 = 20\,°C$, para las mezclas de combustibles de hidrocarburos con el aire, que generalmente son las utilizadas en los motores, oscilan entre límites relativamente pequeños ($35-55$ cm/s). Las menores velocidades u_n

125

a que la influencia de la presión sobre la velocidad de las reacciones químicas —relativamente pequeña, mientras que las velocidades de reacciones de los productos de difusión y transmisión de calor disminuyen inversamente proporcional a la presión. Con precisión suficiente para los cálculos prácticos, puede admitirse que en las mezclas airegasolina las velocidades normales de propagación de la llama están relacionadas con las temperaturas inicial de la mezcla y con la presión por la siguiente expresión:

$$u_n = f(p) \sim p^{-x} \quad (17a)$$

Combustión turbulenta. En la mayoría de casos, la masa principal de la mezcla de trabajo, en el cilindro del motor se encuentra en movimiento turbulento bastante intenso, constituido por vórtices dirigidos y pulsaciones desordenadas de las velocidades de los flujos gaseosos. Bajo la acción de la turbulencia la velocidad de propagación de la llama considerablemente se acelera.

Los torbellinos grandes o las pulsaciones fuertes retuercen el frente de llama rompiéndolo en focos ardientes separados (moles), lo que puede conllevar a que la superficie real de combustión aumente varias veces (fig. 52). Las pulsaciones turbulentas de tamaños comparables con el grosor del frente de llama laminar no originan su retorcimiento, pero reforzando los procesos de transmisión de calor y difusión en el propio frente incrementan la velocidad de combustión en la superficie de los moles grandes. Bajo la acción de los dos factores mencionados las velocidades de propagación de la llama turbulenta crecen prácticamente en forma proporcional a la intensidad de la turbulencia (a la magnitud media cuadrática de las velocidades de las pulsaciones turbulentas u'). En este caso las dependencias de las velocidades de la combustión turbulenta respecto a la composición de la mezcla, de su temperatura inicial y de su presión, resultan un poco distintas a las funciones de la velocidad normal de propagación de la llama u_n respecto de los mismos factores.

Según resultados de las últimas investigaciones, la velocidad de propagación de la llama turbulenta se expresa por

$$u_t \approx u_n + Bu'$$

Fig. 52 Carácter aproximado de las incurvaciones y discontinuidades del frente de una llama turbulenta por efecto de turbulencia.

---

en todos los regímenes tales condiciones, en particular el motorac, las mayores.

Al empobrecer o enriquecer la mezcla las velocidades normales de la llama ha disminuido, sin límite, sino que tiene cierto mínimo y después se apaga. Esto está relacionado con que a bajas u_n crecen las pérdidas relativas de calor desde la zona de combustión. El límite de enriquecimiento máximo de la mezcla por encima del cual la propagación de la llama resulta imposible, se denomina límite de concentración superior α_{min}, mientras que el inferior α_{max} límite de concentración —límite inferior— de propagación de la llama.

Para la mayoría de los combustibles de motores, a presión atmosférica y $t_0 = 20\,°C$, el límite superior de propagación de la llama limitar $\alpha_{min} = 0.25\ldots 0.55$, el inferior $\alpha_{max} = 1.55 \ldots 1.8$. Precisamente los límites de variación de la composición de la mezcla.
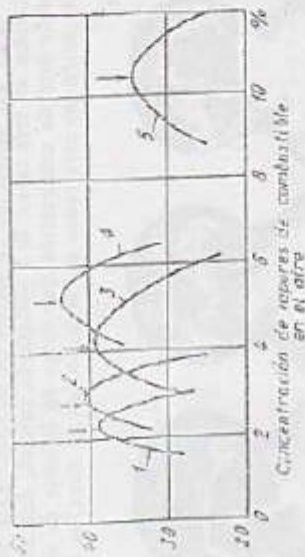
Fig. 51 Variación de los combustibles de motores, en función de la composición de la mezcla para una serie de hidrocarburos individuales siendo $p = 0.1$ MPa, $t_0 = 20\,°C$: 1 — hidrógeno; 2 — benceno; 3 — propano; 4 — propano; 5 — metano; con fluctuaciones del valor de $\alpha = \alpha_1$

para los cuales la combustión aún continúa manteniéndose lo suficientemente estable, son bastante más estrechos. Para los motores a gasolina, por lo general, son de $0.6 < \alpha < 1.3$. Si los motores funcionan con combustibles gaseosos los límites del posible empobrecimiento de la mezcla son un poco mayores.

Al aumentar la temperatura inicial de la mezcla carburante, crecen las velocidades normales de propagación de la llama y se amplían un tanto los límites de concentración, lo que se origina por el aumento de las velocidades de reacción y de los coeficientes de conductividad de la temperatura de la mezcla. dicho aumento está vinculado con las temperaturas más elevadas de la llama T_z. La elevación de la presión, por el contrario, conduce a cierta disminución de las velocidades normales de propagación de la llama. Esto se debe

mente se sobreenriquecidas ($\alpha < 0,3 \ldots 0,4$). En estas zonas tiene lugar la descomposición de las moléculas de los hidrocarburos sin suficiente suministro de aire que va acompañada del desprendimiento de carbono sólido en forma de partículas muy diminutas de carbonilla. Precisamente ésta es la zona que origina la aparición de humos en los motores Diesel al incrementar el suministro de combustible por ciclo, es decir, la carga por encima de ciertos límites.

§ 2. Combustión en los motores de encendido por chispa

Siendo normal el proceso de funcionamiento de los motores de encendido por chispa, la mezcla de combustible evaporado, aire y gases residuales lo suficiente homogénea se enciende por la chispa eléctrica y se quema en el proceso de propagación a lo largo de toda la carga del frente de la llama turbulenta. Este proceso se puede dividir en tres fases: la I —llamada inicial, durante la cual un pequeño foco de combustión que surge en la zona de altas temperaturas entre los electrodos de la bujía (en la vena de la chispa

Fig. 53. Carácter de la llama turbulenta en la cámara de combustión con válvulas laterales en un motor de carburador:

a — frente de llama; b — zona encendida; c — mezcla no quemada

la temperatura sobrepasa los 10 000 °C) progresivamente se transforma en un amplio frente de la llama turbulenta; la II —o fase principal— en la que ocurre la rápida propagación de la llama turbulenta a través de la mayor parte de la cámara de combustión, siendo prácticamente constante el volumen de esta última, puesto que el émbolo se encuentra cerca del P.M.S.; la III —fase de combustión residual de la mezcla detrás del frente de la llama, en las capas cercanas a las paredes y en los huelgos entre la culata y la cara del pistón, esta fase abarca parte de la carrera de expansión.

Para tener una clara idea sobre el carácter de la propagación de la llama se muestran algunas fotografías tomadas durante una filmación a alta velocidad a través de una mirilla transparente (de cuarzo) (fig. 53) instalada en la culata.

9—1293

---

donde B es un coeficiente directamente proporcional a la velocidad de la reacción a la temperatura de la llama T_{ll}, es decir, $B \sim \rho^{m} e^{-E/RT_{ll}}$.

La elevación de la presión influye positivamente sobre u_{t}, (para el caso de mezclas con gasolina $u_{t} \sim p^{0,05}$). La temperatura inicial de la mezcla y su composición influyen más sobre u_{t} que sobre u_{n}.

Simultáneamente con el crecimiento de u_{t} aumenta la extensión de la zona de combustión δ_{t} (véase la fig. 52). A causa de esto, la velocidad de propagación de las fronteras delanteras de la zona de la llama turbulenta no caracteriza unívocamente la velocidad real de desprendimiento de calor, que responde a la terminación completa de la combustión. En los motores de carburador el ancho de la zona de combustión en el frente de la llama turbulenta puede alcanzar 20—25 mm, lo que deberá tenerse en cuenta al comparar los diagramas indicados de presión con registro de la propagación de la llama.

Combustión difusiva. La combustión de mezclas heterogéneas, por ejemplo de los chorros de gas que penetran en el aire o de gotas suspendidas en él de combustible líquido pulverizado, sustancialmente difiere de la combustión de las mezclas homogéneas. La velocidad con que arde una mezcla gaseosa heterogénea se determina prácticamente por la velocidad de la agitación difusiva entre el combustible y el comburente, puesto que las reacciones químicas de combustión transcurren con velocidades considerablemente más altas que los procesos de mezclado. Como consecuencia, esta combustión se denomina difusiva.

La velocidad con que se queman las gotas de combustible líquido se determina fundamentalmente por la velocidad de su evaporación (los vapores de combustible que se forman sobre la superficie de la gota, se difunden en el aire que rodea la gota, formando a cierta distancia de ella la mezcla carburante, además este proceso transcurre con mayor rapidez que la vaporización del combustible).

Las diminutas gotas ($\varnothing < 40 \ \mu m$), uniformemente distribuidas en el aire, queman aproximadamente con la misma velocidad que una mezcla de vapores de combustible y aire homogénea; al mismo tiempo los límites del posible empobrecimiento general de las mezclas heterogéneas (tanto de gotas como de gases), para los cuales se conserva todavía la combustión estable, resultan considerablemente más amplios que en las homogéneas. En las mezclas heterogéneas se forman siempre zonas donde $\alpha = 0,85 \ldots 0,90$, correspondientes a las mayores velocidades de reacción y temperaturas de los productos de la combustión. Estas zonas sirven de centros de inflamación de la mezcla pobre que las rodea.

Precisamente esto explica la posibilidad de que los motores Diesel funcionen a bajas cargas con mezclas extremadamente pobres ($\alpha > 4$).

Debido a esas mismas razones la combustión de las mezclas heterogéneas, siendo bajo el coeficiente total de exceso de aire ($\alpha < 1,4 \ldots 1,5$), por lo general, conduce a la formación de humo y carbonilla, ya que existen zonas con mezclas locales considerable-

Al atacarse el frente de llama a las paredes la velocidad de su propagación disminuye, lo que se explica por la menor intensidad de la turbulencia en las capas cercanas a dichas paredes. Cuando la llama alcanza las paredes en la mayor parte de la cámara, la velocidad de combustión decrece también como consecuencia de la disminución de la superficie del frente de la llama, pero la combustión sin embargo, aquí no termina. Durante un tiempo aún suficiente mente prolongado continúan los procesos de combustión residual en las capas cercanas a las paredes y detrás del frente de la llama turbulenta. La velocidad de la combustión residual de la misma manera que el proceso en su fase inicial, en mayor medida depende de las propiedades físico-químicas de la mezcla de trabajo que la velocidad en la fase principal.
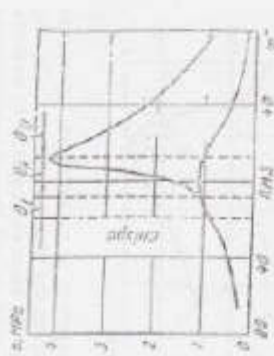
Trazar límites bien definidos entre las diferentes fases del proceso de combustión en los motores no es posible, ya que el carácter y las velocidades de combustión varían gradualmente. Como instante de culminación de la primera fase de combustión y el comienzo de la fase principal en los motores de encendido por chispa generalmente se adopta el «punto de separación» de la línea de combustión respecto a la línea de compresión en el diagrama indicado, es decir el instante en que comienza la combustión (fig. 55). La fase inicial dura desde el instante en que salta la chispa entre los electrodos de la bujía hasta el «punto de separación». Por analogía con los motores Diesel a veces esta lapso convencionalmente se denomina período de retraso de la inflamación, o período de inducción, lo que en principio no es cierto. Siendo el encendido por chispa, el retraso de la inflamación prácticamente no existe, cerca de los electrodos de la bujía inmediatamente surge el foco de combustión, pero hay un período, durante el cual el frente de la llama desde este foco se propaga en forma relativamente lenta y la fracción de mezcla quemada es todavía tan pequeña que el incremento de la presión en el diagrama indicado no se registra.

Así por ejemplo, si Δp constituye el 1% de la presión máxima de combustión, lo que aproximadamente corresponde a la combustión del 1.5% de la mezcla de trabajo, entonces el volumen que ocupan los productos de la combustión será ya cerca del 3% del volumen de la cámara (adoptamos que la relación de expansión de los gases quemados en la combustión es $\lambda = 3.3$). Comparando los gráficos (fig. 54, c y 55) se concluye que hacia el instante en que comienza la fase principal de la combustión —el «punto de separación» 62°.

Fig. 55. Diagrama indicado de un motor de carburador que corresponde al esquema de propagación de la llama ilustrado en la fig. 54, a

---

En la fig. 54, a, se ilustran los cambios de propagación de la llama que acompañan al cambio al variar el carácter del movimiento turbulento de la mezcla de trabajo en la cámara de combustión. Conociendo los intervalos de tiempo entre las posiciones consecutivas del frente de la llama puede determinarse la velocidad de su propagación en cualquier punto. Al principio, mientras el foco de combustión se encuentra la velocidad de la llama es relativamente baja, ya que sobre ella influyen solamente pulsaciones turbulentas de pequeñas

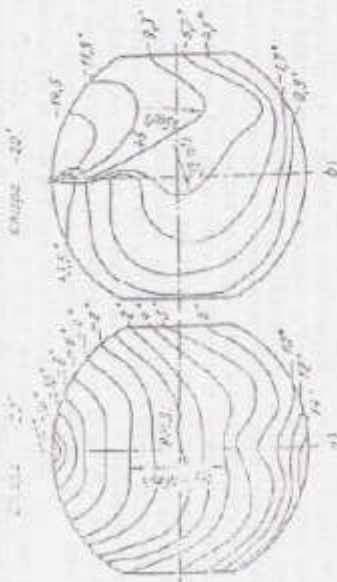

Fig. 54. Esquemas de las posiciones instantáneas consecutivas del frente de llama en la cámara de combustión de un motor que funciona con combustible ligero, a iguales intervalos de tiempo caracterizados por el ángulo φ de rotación del cigüeñal desde el P.M.S. (n = 600 rpm; α = 1):

dimensiones. En este período la velocidad de combustión preponderantemente se determina por las propiedades físico-químicas de la mezcla carburante.

A medida que crecen las dimensiones del foco inicial de la llama, con mayor grado se reflejará la acción positiva de las pulsaciones turbulentas mayores sobre la velocidad de combustión. En la fase principal la velocidad de propagación de la llama es aproximadamente proporcional a la intensidad del que, a su vez, crece proporcionalmente con la frecuencia de rotación del cigüeñal. En consecuencia, la duración de la fase principal de la combustión, expresada en grados de rotación del cigüeñal, a igualdad de las demás condiciones, casi no depende del régimen de velocidad del motor.

En la fig. 54 se indican los valores medios de las velocidades visibles de propagación de la llama u_t en algunos sectores. Si la frecuencia de rotación es más elevada, estas velocidades consideradablemente se incrementan y en los motores rápidos modernos a gasolina ellas alcanzan en la parte central de la cámara de combustión de 60 a 80 m/s.

antes del P.M.S.)— el radio medio de la semiesfera de la llama constituye cerca del 30% del radio de la cámara de combustión. Semejantes dimensiones del foco inicial de la combustión son ya suficientes para que sea posible el ulterior desarrollo rápido de la velocidad de propagación de la llama bajo la acción de las pulsaciones turbulentas de grandes dimensiones.

El límite de separación entre las fases principal y final de la combustión en los motores que funcionan con combustible ligero, convencionalmente, por lo general, se adopta el instante en que se alcanza la presión máxima en el diagrama indicado (punto l de la fig. 56). La combustión, sin embargo, aún no acaba y la temperatura media de los gases en el cilindro continúa creciendo cierto tiempo, alcanzando su máximo valor en el punto 2. Sin embargo, la disminución de la velocidad de desprendimiento de calor, debido a que el frente de la llama en la mayor parte de la cámara de combustión alcanza las paredes, conduce a que el incremento de presión, por efecto de la expansión de los gases, originada por el movimiento del pistón y la transmisión de calor a las paredes.

La efectividad del proceso de trabajo en el cilindro del motor se determina tanto por la plenitud total de la combustión, como por su velocidad, es decir, depende de la oportunidad con que se realiza el desprendimiento del calor, para asegurar una relación de expansión de los productos de combustión cercana a la relación de compresión geométrica. El trabajo máximo del ciclo, y por lo tanto las máximas pot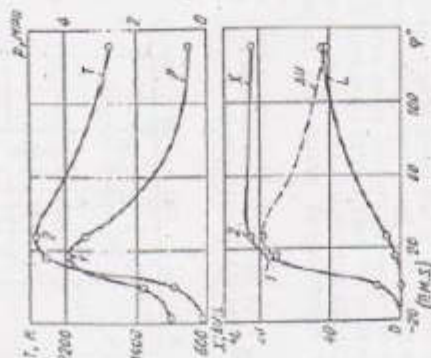encia y economía del motor que funciona con combustible ligero, a igualdad de las demás condiciones, se logra organizando el proceso de combustión de tal manera que los puntos del comienzo y final de la fase principal están aproximada y simétricamente dispuestos con respecto al P.M.S. Esto se puede obtener instalando en el lugar correspondiente el instante de la ignición. El ángulo, en grados de rotación del cigüeñal, desde el instante en que salta la chispa en la bujía hasta el P.M.S. se denomina ángulo de avance al encendido φ_a.

Fig. 56. Variación de la presión p, temperatura de los gases T, cambio de energía interna ΔU, trabajo externo de los gases L, proceso de combustión en el motor GAZ-2x ($\varepsilon = 1,02$; $n = 2000$ rpm; $p_1 = 1$ MPa; $\eta_1 = 0,33$?)

El ángulo de avance al encendido debe ser tanto mayor, cuanto mayor sea la duración de la fase inicial de la combustión θ_1 (fig. 55), así como también cuanto más lentamente se desarrolle la combustión en la fase principal.

En los motores rápidos modernos de combustible ligero con relaciones de compresión $\varepsilon = 8 \ldots 9$, la potencia máxima generalmente corresponde al instante cuando la máxima presión se alcanza a los 12—15° después del P.M.S.; además, la duración de la fase principal de la combustión es $\theta_{II} = 25 \ldots 30°$.

La velocidad del desprendimiento de calor en la fase principal es la que determina la rapidez de incremento de la presión en función del ángulo de rotación del cigüeñal ($dp/d\varphi$), y correspondientemente, la dinámica de acción del esfuerzos de los gases sobre las piezas del mecanismo biela-manivela, de la cual depende la denominada rigidez de funcionamiento del motor. La magnitud $dp/d\varphi$ depende de la velocidad con que transcurre el proceso en la fase de combustión rápida. Para iguales valores de $dp/d\varphi$, pero siendo más suave el desarrollo del proceso, el motor funciona con menos ruido. En los motores con moderadas relaciones de compresión ($6—7$) los valores máximos de $dp/d\varphi$ generalmente constituyen $0,1—0,12$ MPa.° Para valores más elevados de la relación de compresión $\varepsilon = 9 \ldots 10$, la velocidad de elevación de la presión alcanza $0,15—0,25$ MPa.°

La relación entre los procesos de combustión y la utilización del calor desprendido en los motores se determina por las curvas características de desprendimiento activo de calor, que en sí representan la variación, en función del ángulo φ, de la fracción de calor, utilizada en elevar la temperatura del fluido operante (de su energía interna) y en realizar el trabajo externo, con respecto al calor total introducido al ciclo:

$$\chi = \frac{\Delta U + \int p\,dV}{Q} \qquad (17)$$

En la fig. 56 se muestran, en calidad de ejemplo, una parte del diagrama indicado, desde el instante de la ignición hasta el comienzo de apertura de la válvula de escape, así como las curvas de la temperatura media T de los gases en el cilindro, del trabajo exterior de los gases, sus energías internas ΔU y del coeficiente de desprendimiento activo de calor χ, en función del ángulo φ para el motor GAZ-21 cuando funciona con la mariposa de gases completamente abierta. En el punto l, correspondiente a p_{max}, la cantidad de calor desprendido activo abarca el 73%; en la fase de combustión residual la misma alcanza el 85% del calor de combustión disponible. El punto 2, que corresponde a la T_{max}, está situado a 11° del punto l. El 15% restante lo conforman las pérdidas de calor a las paredes y las debidas a la combustión incompleta. Al funcionar los motores en regímenes de media carga o con mezclas empobrecidas, los valores de χ en el punto p_{max} generalmente no sobrepasan el 50% y el proceso

y en los expulsores, es decir, se incrementa la fracción de la mezcla que quema en la tercera fase. Todo esto es su conjunto conduce a que en los motores con elevados e disminuyen los ángulos óptimos de avance al encendido (correspondientes a la máxima potencia), se reduce la duración de la combustión hasta el instante en que se alcanza la máxima presión (el punto p_{max} se acerca al P.M.S.), pero simultáneamente con esto disminuye el coeficiente de desprendimiento de calor activo en los puntos p_{max} y T_{max}, crece el valor relativo de los procesos de combustión residual en la tercera fase.

Frecuencia de rotación. Al aumentar la frecuencia de rotación del cigüeñal del motor, el tiempo reservado al desarrollo del proceso de combustión disminuye y al mismo tiempo se incrementa la intensidad de la turbulencia de la carga de trabajo. Por lo expuesto, como ya se anotó, la velocidad con que se propaga el frente de la llama en la fase principal de combustión crece aproximadamente proporcional a la frecuencia de rotación, mientras que la duración de la fase principal θ_{II}, expresada en grados de rotación del cigüeñal, por lo general, permanece constante. La duración de la fase inicial de la combustión θ_I se incrementa con el crecimiento de la frecuencia de rotación.

Si aumentamos la frecuencia de rotación n, manteniendo constante la composición de la mezcla y el ángulo φ_a, entonces podrá observarse un desarrollo cada vez más tardío del proceso de combustión por ciclo (fig. 58, a). Con el correspondiente aumento del ángulo φ_a puede conseguirse que las líneas de crecimiento de la presión en la fase principal de combustión para diversos valores de n prácticamente coincidirán (fig. 58, b). Al elevar la frecuencia de rotación se incrementa un tanto la duración de la fase de combustión residual

por pero cierta disminución de la efectividad del desprendimiento de calor vinculada con esto se compensa por la disminución de la transferencia de calor a las paredes, debido a la reducción del tiempo que los gases con elevadas temperaturas se encuentran en el cilindro.

Forma de la cámara de combustión. La turbulencia de la carga de trabajo en el cilindro, suscitada por el ingreso de la mezcla a través de secciones de paso relativamente estrechas en los órganos de la
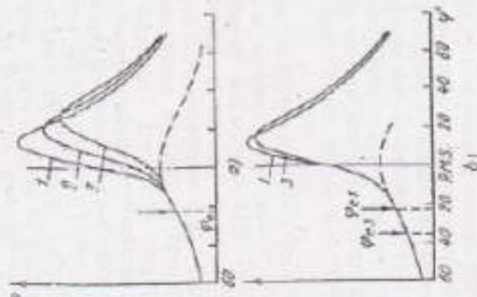
Fig. 58. Influencia de la velocidad de rotación n en el carácter de los diagramas indicados:
$1 - n = 1500\ rpm;\ 2 - n = 2500\ rpm;\ 3 - n = 3500\ rpm.$

distribución de los gases (válvulas, conductos de admisión), puede ser adicionalmente reforzada gracias al traspaso de la mezcla desde el cilindro a la cámara de combustión al final de la carrera de compresión. Esto se logra proporcionando a las cámaras de combustión la forma más adecuada, para la cual en cierta parte de ella se forman huelgos relativamente angostos entre la superficie inferior de la culata y la cara del pistón, que se disminuyen explosores (véase la fig. 32).

Los expulsores se disponen de tal manera que permiten crear turbulencia adicional a la carga en aquellas zonas hasta las cuales el frente de llama llega desde la bujía en último turno. Con esto se consigue la aceleración de la combustión residual. La bujía se coloca de manera que no origine demasiada turbulencia cerca de ella y, al mismo tiempo, que se garantice buen barrido respecto a los gases residuales en la zona de la bujía, dirigiendo hacia ella parte del flujo de la mezcla que entra a través de la válvula de admisión.

Cuanto más cerca del centro de la cámara de combustión se instale la bujía, tanto más corto será el camino recorrido por el frente de llama hasta los puntos más alejados. Al colocar la bujía en la parte central se obtiene la mayor superficie del frente de la llama, en consecuencia, la velocidad de crecimiento de la presión resultan más altas que cuando la bujía está situada al costado. Es indispensable, sin embargo, considerar que esto se refiere solamente a las cámaras de combustión que tienen forma simétrica. La creación de expulsores puede influir más sobre la velocidad de combustión que la disposición de la bujía.

Sin embargo, siendo la superficie de los expulsores grande, la parte relativa de la mezcla no quemada que se enfría intensamente y está encerrada en angostas ranuras entre la culata y la cara del pistón, resulta ser bastante considerable, lo que puede conducir a una notable disminución del coeficiente activo de desprendimiento de calor en el punto p_{max} y, por lo tanto, a la disminución del trabajo útil del ciclo. Es por esta razón que la superficie de los expulsores generalmente no sobrepasa el 30-40% de la superficie del pistón.

Entre las principales medidas empleadas para mejorar el proceso de combustión en los motores de combustión por chispa emplean circuitos de toxicidad, figuran las siguientes:

1. La intensificación del encendido por chispa empleando circuitos a transistores o tiristores, lo que permite ampliar un poco los límites del empobrecimiento efectivo de la mezcla al funcionar a pequeñas cargas y en los regímenes transitorios, disminuir el consumo de combustible y la expulsión de los gases CO y C_nH_m.

2. La creación de turbulencia de la carga de trabajo en los cilindros, utilizando estructuras de conductos de admisión que aseguren la dirección tangencial o en espiral del movimiento de la mezcla, lo que reduce la duración de la combustión y mejora el grado de uniformidad de ciclos consecutivos.

3. La estratificación de la carga de trabajo de tal manera que en la zona de la bujía se concentra la mezcla más rica, en tanto que a

§ 4. Principales alteraciones de la combustion normal en los motores de encendido por chispa

Detonación. Externamente la detonación se revela durante el funcionamiento del motor a grandes cargas en la aparición del golpeteo metálico estridente, que es resultado de repercusiones periódicas múltiples sobre las paredes de la cámara de combustión de ondas de choque que se forman en los gases. Al suceder esto al final de la combustión se registran vibraciones de la presión en forma de toda una serie de agudos picos que van gradualmente amortiguándose y que pueden observarse en los diagramas indicados (fig. 60). La frecuencia de estas vibraciones de la presión es igual a la frecuen-
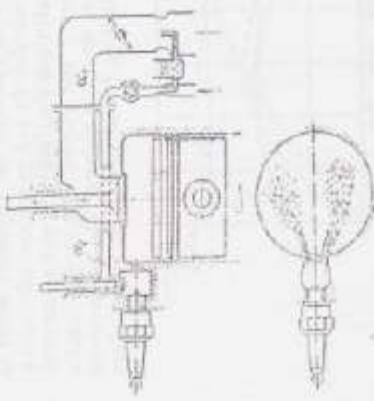
Fig. 59. Estructura esquemática de un mo... ... cámara de encendido por dardo.

Fig. 60. Diagramas indicados típicos de un motor de encendido por chispa cuando funciona con detonación.
a — detonación súbit; b — detonación intensa

cia principal del golpeteo audible. Ella depende de las velocidades de propagación de las ondas de choque y del camino recorrido entre las reflexiones consecutivas desde las paredes que está determinado por las dimensiones del cilindro (principalmente por su diámetro).

Cuando la detonación es débil la velocidad... ... en cada ciclo de trabajo, la amplitud de las vibraciones de presión es pequeña... ... de las ondas de choque en los productos de la combustión oscilan entre los límites de 1000–1200 m s. Siendo el diámetro del cilindro de 100 mm la frecuencia de las vibraciones es aproximadamente igual a 5000 Hz. Cuando la detonación es intensa, el fuerte golpeteo con frecuencia un poco mayor surge en cada ciclo, disminuyendo al mismo tiempo la potencia del motor y apareciendo humo negro en los gases de escape.

El trabajo prolongado del motor con detonación es estrictamente intolerable, puesto que cuando surgen las ondas de choque crece bruscamente la emisión de calor desde los gases quemados hacia las paredes, lo que no puede conducir al recalentamiento del motor y a la destrucción de algunas piezas en las cámaras de combustión (se pueden quemar las aristas de los pistones, las juntas entre los cilindros y la culata, los electrodos de las bujías). El carácter vibratorio de la carga sobre el trabajo cuando hay detonación puede suscitar la destrucción de la capa antifricción de los cojinetes de las bielas. El desgaste de la parte superior de la camisa del cilindro también aumenta, ya que las ondas de choque destruyen la película de aceite que cubre la superficie metálica, como resultado de lo cual surge fricción seca y simultáneamente se refuerza el desgaste corrosivo bajo la acción de los químicamente activos que contienen los productos de combustión, en particular los óxidos de nitrógeno.

La detonación se produce por el autoencendido de la carga de trabajo más alejada de la bujía, hasta la cual el frente de llama llega...

desde la misma en último turno. Las últimas porciones de la mezcla no quemada se comprimen hasta la presión p_x y, si esta compresión es adiabática, entonces su temperatura puede alcanzar valores de T_x. $T_a(p_x/p_a)^{\frac{k-1}{k}}$, que sobrepasan en mucho la temperatura de autoencendido, incluso de los combustibles de elevada octanaje. Así por ejemplo, pero $T_a = 740$ K, $p_a = 0,085$ MPa, $p_x = 4$ MPa y el valor medio del exponente adiabático de la mezcla no quemada $k = 1,34$, su temperatura alcanza 900 K.

A la aparición de la detonación contribuyen todos aquellos factores que incrementan la velocidad de las reacciones que transcurren delante del frente de la llama en la última parte de la carga, a saber:

a) elevada capacidad de reacción del combustible, cuanto mayor sea esta capacidad tanto menor será el número de octano;

b) aumento de la relación de compresión que origina la elevación de la presión y temperatura en la última parte de la carga. Una función media estadística aproximada entre los valores tolerables de ε, para los cuales aparece la detonación audible, y los números de octano de la gasolina se ilustra en la fig. 51;

c) aumento del ángulo de avance del encendido, en caso del cual la máxima presión p_z se obtiene cuando el pistón está más cerca del P.M.S. y correspondientemente se elevan los valores de p_z y T_z (véase la fig. 54);

d) composición de la mezcla ($\alpha \approx 0,9$) correspondiente a las presiones y temperaturas de combustión más elevadas, así como a las máximas velocidades en que se desarrollan las reacciones que tienen lugar delante del frente de la llama en las mezclas calentadas por compresión;

e) malas condiciones de enfriamiento de las últimas partes de la carga y el diseño inapropiado de la cámara de combustión que contribuye a la lenta combustión residual.

Los factores que impiden el surgimiento de la detonación son aquellos que aceleran la combustión de la última parte de la carga en el frente de llama o que mediante cualquier otro procedimiento dificultan la aparición de la autoinflamación explosiva.

Entre estos factores figuran: e) enérgica turbulencia de la carga de trabajo; b) disminución del camino recorrido por el frente de llama desde la bujía hasta los puntos más alejados de la cámara de combustión; c) la existencia de expulsores en la zona donde se encuentra la última
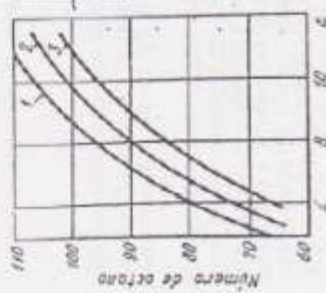
Fig. 51. Funciones estadísticas medias del número de octano de los combustibles para motores con diferentes diámetros de los cilindros con respecto a la relación de compresión:
1 — 112 mm; 2 — 90 mm; 3 — 66 mm

140

parte de la carga, que permiten su mejor enfriamiento y dificultan el surgimiento de focos suficientemente grandes de autoinflamación explosiva, capaces de originar la formación de ondas de choque.

La influencia de la frecuencia de rotación del cigüeñal sobre la detonación depende de las propiedades de las gasolinas utilizadas. Para el caso de hidrocarburos parafínicos y nafténicos, que se caracterizan por su tendencia biotópica y por poseer baja sensibilidad a las temperaturas, la tendencia del motor a detonar paulatinamente disminuye a medida de incrementar la frecuencia de rotación. Pero
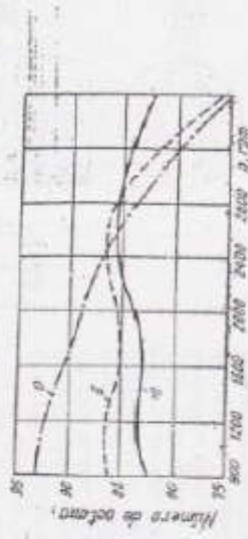
Fig. 52. Variación de los números de octano requeridos en función de la frecuencia de rotación del cigüeñal para gasolinas de diferentes sensibilidades de temperatura:
los cifras de las curvas indican los valores de la sensibilidad de temperatura

para el caso de las gasolinas modernas de alto octanaje para automóviles, que se destacan por su alto contenido de hidrocarburos aromáticos y, por lo tanto, poseen elevada sensibilidad a las temperaturas, la tendencia que tienen los motores a la detonación prácticamente no depende de la frecuencia de rotación (fig. 52).

La tendencia a la detonación de los motores, siendo iguales la frecuencia de rotación y duración de la combustión, es considerablemente más débil en aquellos que tienen menores dimensiones de los cilindros. Esto se explica por ser menores los volúmenes de las porciones residuales de la mezcla no quemada, en las cuales no es tan probable el surgimiento de la autoinflamación del tipo explosivo.

Autoencendido prematuro. Las piezas que están fuertemente recalentadas en la cámara de combustión del motor (electrodos centrales y aisladores de las bujías, cabezas de las válvulas de escape), si sus temperaturas sobrepasan ciertos límites (700—800° C), pueden provocar el autoencendido prematuro de la mezcla de trabajo o el denominado encendido por incandescencia antes de que salte la chispa. Como fuentes de dicho fenómeno pueden servir también las partículas grandes incandescentes (ardientes) de carbonilla.

Desde las superficies incandescentes el frente de la llama se propaga de manera análoga al de la llama proveniente de la chispa de la bujía, pero el instante de la ignición resulta en este caso incontro-

141

... ensimismo ... entonces la potencia del motor disminuye ... y del efecto del consumo adicional de trabajo de compresión ... ya que desde el efecto del consumo adicional de trabajo de la emisión ... los diagramas indicados, para el caso de este fenómeno, ... representados en la fig. 63.

... es la anomalía de la combustión más ... de encendido por chispa, ya que origina el ... muy intenso del traspaso de calor a las paredes, ... en la relación de p_z y T_z, puesto que los gases que ya tuvieren ... de apretarse son adiabáticamente comprimidos por el ...

Fig. 63. Diagramas indicados de un motor de encendido por chispa con encendido anticipado

... culolo y el tiempo de contacto de los gases quemados con las paredes aumenta considerablemente. Esto conduce al rápido autorreforzamiento del encendido por incandescencia, es decir, este aparece cada vez con mayor anticipación durante la carrera de compresión, a causa de lo cual pueden quemarse (fundirse) los pistones. El peligro del encendido anticipado se acrecienta porque extremamente éste por lo general se manifiesta sólo en forma de golpes secos, que no siempre se logra descubrir entre el conjunto de ruidos al funcionar el motor a elevadas cargas. Puesto que el encendido por incandescencia surge generalmente sólo en uno de los cilindros, la disminución total de la potencia de un motor multicilíndrico es insignificante.

Las fuentes del encendido anticipado que con mayor frecuencia se presentan son los electrodos centrales recalentados y las campanas de los aisladores de las bujías. Por eso es indispensable seleccionar las bujías en estricta correspondencia con las particularidades del motor. Dichas bujías deberán poseer suficientemente elevado grado térmico, que caracteriza la resistencia de la bujía contra el recalentamiento y, al mismo tiempo, no deberán ser excesivamente «frías», o sea, la temperatura de los electrodos centrales de las bujías cuando el motor funciona en vacío no tendrá que ser menor de 400 °C, para evitar la formación de residuos carbónicos propios de la combustión.

Autoencendido retrasado. En los motores a gasolina de automóviles con altas relaciones de compresión, especialmente al funcionar con combustibles que tienen aditivos antidetonantes, en la parte no elemental de la carga puede en régimen incorrecto de encendido por incandescencia, después de realizar la propagación del frente de la llama son las partículas de carbonilla incandescente (ordinaria) que se desprenden de las paredes y se encuentran suspendidas en la carga de trabajo. Dichas partículas se depositan en las cámaras de combustión cuando el motor trabaja durante prolongado tiempo en regímenes de cargas muy pequeñas y en vacío, pero se separan de las paredes al incrementar la carga. Desde estas partículas incandescentes comienzan a propagarse frentes de llama adicionales y la velocidad de combustión bruscamente aumenta al final de la fase principal. Los términos de los diagramas indicados correspondientes a la culminación de la combustión, tienen forma de picos agudos (fig. 64), se elevan las presiones máximas p_z y especialmente los valores de $dp/d\varphi$, que pueden alcanzar 1,0 kPa/°, mientras que para la combustión normal constituyen 0,2 MPa.°

Semejante alteración de la combustión en los motores a gasolina, que en la literatura extranjera han obtenido la denominación de «rumble» (estrépito, estruendo), es característica para los automóviles con motores de gran cilindrada y, por lo tanto, con gran reserva de potencia, los cuales en condiciones de ciudades considerable parte de su tiempo funcionan con cargas muy pequeñas.

Autoencendido por compresión desconectado. Semejante inflamación aparece con cierta frecuencia en los motores modernos para automóvil y consiste en que éstos continúan a veces durante un tiempo suficientemente largo trabajando en vacío después de desconectar el sistema de ignición. Este fenómeno se explica a veces equivocadamente como si fuera encendido por incandescencia, a pesar de que no tiene nada en común con este último.

Siendo la relación de compresión lo suficiente elevada ($\varepsilon > 6$), al girar el cigüeñal de un motor normalmente calentado, cuya marcha en vacío está casi cerrada (correspondiente al régimen de funcionamiento en vacío) la temperatura de la mezcla carburante al final del proceso de compresión alcanza valores lo suficiente altos como para que pueda producirse la autoinflamación de la mezcla a baja frecuencia de rotación ($n = 300 \cdots 400$ rpm).
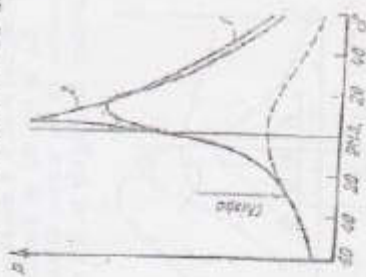
Fig. 64. Diagramas indicados superpuestos de un motor de encendido por chispa
1 — con combustión normal; 2 — con autoencendido retrasado

[2]   Heywood, J., B., *Internal Combustion Engines Fundamentals*, 1era ed., Duffy, A., Morris, J., New York, Chicago, San Francisco, Atenas, Londres, Madrid, México, Milán, Nueva Delhi, Singapur, Sydney, Toronto: McGraw-Hill Education, 1988.

29. Borgnakke, C., Davis, G. C., and Tabaczynski, R. J.: "Predictions of In-Cylinder Swirl Velocity and Turbulence Intensity for an Open Chamber Cup in Piston Engine," SAE paper 810224, SAE Trans., vol. 90, 1981.

30. Arcoumanis, C., Bicen, A. F., and Whitelaw, J. H.: "Squish and Swirl-Squish Interaction in Motored Model Engines," Trans. ASME, J. Fluids Engng, vol. 105, pp. 105–112, 1983.

31. Ikegami, M., Mitsuda, T., Kawauchi, K., and Fujikawa, T.: "Air Motion and Combustion in Direct Injection Diesel Engines," JARI technical memorandum no. 2, pp. 231–245, 1971.

32. Liou, T.-M., and Santavicca, D. A.: "Cycle Resolved Turbulence Measurements in a Ported Engine With and Without Swirl," SAE paper 830419, SAE Trans., vol. 92, 1983.

33. Fitzgeorge, D., and Allison, J. L.: "Air Swirl in a Road-Vehicle Diesel Engine," Proc. Instn Mech. Engrs (A.D.), no. 4, pp. 151–168, 1962–1963.

34. Lichty, L. C.: Combustion Engine Processes, McGraw-Hill, 1967.

35. Shimamoto, Y., and Akiyama, K.: "A Study of Squish in Open Combustion Chambers of a Diesel Engine," Bull. JSME, vol. 13, no. 63, pp. 1096–1103, 1970.

36. Dent, J. C., and Derham, J. A.: "Air Motion in a Four-Stroke Direct Injection Diesel Engine," Proc. Instn Mech. Engrs, vol. 188, 21/74, pp. 269–280, 1974.

37. Asanuma, T., and Obokata, T.: "Gas Velocity Measurements of a Motored and Firing Engine by Laser Anemometry," SAE paper 790096, SAE Trans., vol. 88, 1979.

38. Asanuma, T., Babu, M. K. G., and Yagi, S.: "Simulation of Thermodynamic Cycle of Three-Valve Stratified Charge Engine," SAE paper 780319, SAE Trans., vol. 87, 1978.

39. Hiroyasu, H., Elchian, A., Heywood, J. B., Tabaczynski, R. J., and Wall, J. C.: "Performance and NO$_x$ Emissions Modelling of a Jet Ignition Prechamber Stratified Charge Engine," SAE paper 760161, SAE Trans., vol. 85, 1976.

40. Zimmerman, D. R.: "Laser Anemometer Measurements of the Air Motion in the Prechamber of an Automotive Diesel Engine," SAE paper 830452, 1983.

41. Meintjes, K., and Alkidas, A. C.: "An Experimental and Computational Investigation of the Flow in Diesel Prechambers," SAE paper 820275, SAE Trans., vol. 91, in Diesel Engine Combustion, Emissions, and Particulates, P-107, SAE, 1982.

42. Namazian, M., and Heywood, J. B.: "Flow in the Piston-Cylinder-Ring Crevices of a Spark-Ignition Engine: Effect on Hydrocarbon Emissions, Efficiency and Power," SAE paper 820088, SAE Trans., vol. 91, 1982.

43. Wentworth, J. T.: "Piston and Ring Variables Affect Exhaust Hydrocarbon Emissions," SAE paper 680109, SAE Trans., vol. 77, 1968.

44. Tabaczynski, R. J., Hoult, D. P., and Keck, J. C.: "High Reynolds Number Flow in a Moving Corner," J. Fluid Mech., vol. 42, pp. 249–255, 1970.

45. Daneshyar, H. F., Fuller, D. E., and Deckker, B. E. L.: "Vortex Motion Induced by the Piston of an Internal Combustion Engine," Int. J. Mech. Sci, vol. 15, pp. 381–390, 1973.

CHAPTER 9

COMBUSTION IN SPARK-IGNITION ENGINES

9.1 ESSENTIAL FEATURES OF PROCESS

In a conventional spark-ignition engine the fuel and air are mixed together in the intake system, inducted through the intake valve into the cylinder, where mixing with residual gas takes place, and then compressed. Under normal operating conditions, combustion is initiated towards the end of the compression stroke at the spark plug by an electric discharge. Following inflammation, a turbulent flame develops, propagates through this essentially premixed fuel, air, burned gas mixture until it reaches the combustion chamber walls, and then extinguishes. Photographs of this process taken in operating engines illustrate its essential features. Figure 9-1 (color plate) shows a sequence of frames from a high-speed color movie of the combustion process in a special single-cylinder engine with a glass piston crown.[1] The spark discharge is at about $-30°$. The flame first becomes visible in the photos at about $-24°$. The flame, approximately circular in outline in this

FIGURE 9-1 (On color plate opposite p. 498)
Color photographs from high-speed movie of spark-ignition engine combustion process, taken through glass piston crown. Ignition timing 30° BTC, light load, 1430 rev/min, (A/F) = 19.[1]

FIGURE 9-2
Cylinder pressure, mass fraction burned, and volume fraction enflamed for five consecutive cycles in a spark-ignition engine as a function of crank angle. Ignition timing 30° BTC, wide-open throttle, 1044 rev/min, $\phi = 0.98$.[4]

view through the piston, then propagates outward from the spark plug location. The blue light from the flame is emitted most strongly from the front. The irregular shape of the turbulent flame front is apparent. At TC the flame diameter is about two-thirds of the cylinder bore. The flame reaches the cylinder wall farthest from the spark plug about 15° ATC, but combustion continues around parts of the chamber periphery for another 10°. At about 10° ATC, additional radiation—initially white, turning to pinky-orange—centered at the spark plug location is evident. This afterglow comes from the gases behind the flame which burned earlier in the combustion process, as these are compressed to the highest temperatures attained within the cylinder (at about 15° ATC) while the rest of the charge burns.[2,3]

Additional features of the combustion process are evident from the data in Fig. 9-2, taken from several consecutive cycles of an operating spark-ignition engine. The cylinder pressure, fraction of the charge mass which has burned (determined from the pressure data, see Sec. 9.2), and fraction of the cylinder volume enflamed by the front (determined from photographs like Fig. 9-1) are shown, all as a function of crank angle.[4] Following spark discharge, there is a period during which the energy release from the developing flame is too small for the pressure rise due to combustion to be discerned. As the flame continues to grow and propagate across the combustion chamber, the pressure then steadily rises above the value it would have in the absence of combustion. The pressure reaches a maximum after TC but before the cylinder charge is fully burned, and then decreases as the cylinder volume continues to increase during the remainder of the expansion stroke.

The flame development and subsequent propagation obviously vary, cycle-by-cycle, since the shape of the pressure, volume fraction enflamed, and mass fraction burned curves for each cycle differ significantly. This is because flame growth depends on local mixture motion and composition. These quantities vary in successive cycles in any given cylinder and may vary cylinder-to-cylinder. Especially significant are mixture motion and composition in the vicinity of the spark plug at the time of spark discharge since these govern the early stages of flame development. Cycle-by-cycle and cylinder-to-cylinder variations in combustion are important because the extreme cycles limit the operating regime of the engine (see Sec. 9.4.1).

Note that the volume fraction enflamed curves rise more steeply than the mass fraction burned curves. In large part, this is because the density of the unburned mixture ahead of the flame is about four times the density of the burned gases behind the flame. Also, there is some unburned mixture behind the visible front to the flame: even when the entire combustion chamber is fully enflamed, some 25 percent of the mass has still to burn. From this description it is plausible to divide the combustion process into four distinct phases: (1) spark ignition; (2) early flame development; (3) flame propagation; and (4) flame termination. Our understanding of each of these phases will be developed in the remainder of this chapter.

The combustion event must be properly located relative to top-center to
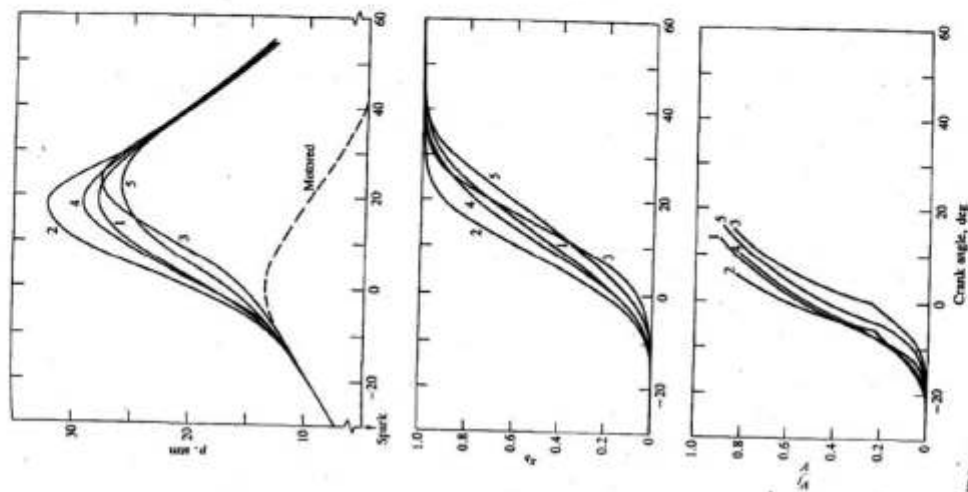
obtain the maximum power or torque. The combined duration of the flame development and propagation process is typically between 30 and 90 crank angle degrees. Combustion starts before the end of the compression stroke, continues through the early part of the expansion stroke, and ends after the point in the cycle at which the peak cylinder pressure occurs. The pressure versus crank angle curves shown in Fig. 9-3a allow us to understand why engine torque (at given engine speed and intake manifold conditions) varies as spark timing is varied relative to TC. If the start of the combustion process is progressively advanced before TC, the compression stroke work transfer (which is *from the piston to the cylinder gases*) increases. If the end of the combustion process is progressively delayed by retarding the spark timing, the peak cylinder pressure occurs later in the expansion stroke and is reduced in magnitude. These changes reduce the expansion stroke work transfer *from the cylinder gases to the piston*. The optimum timing which gives the maximum brake torque—called *maximum brake torque*, or *MBT*, *timing*—occurs when the magnitudes of these two opposing trends just offset each other. Timing which is advanced or retarded from this optimum gives lower torque. The optimum spark setting will depend on the rate of flame development and propagation, the length of the flame travel path across the combustion chamber, and the details of the flame termination process after it reaches the wall. These depend on engine design and operating conditions, and the properties of the fuel, air, burned gas mixture. Figure 9-3b shows the effect of variations in spark timing on brake torque for a typical spark-ignition engine. The maximum is quite flat.
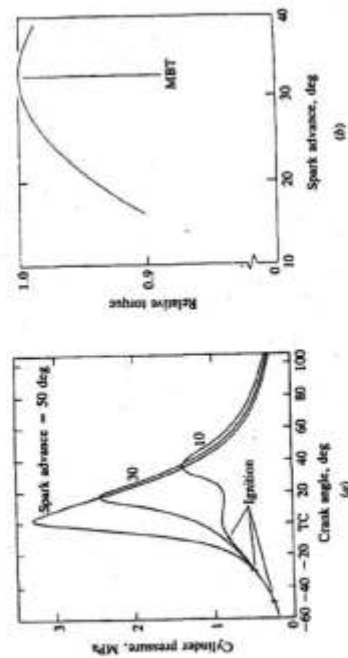
Empirical rules for relating the mass burning profile and maximum cylinder pressure to crank angle at MBT timing are often used. For example, with optimum spark timing: (1) the maximum pressure occurs at about 16° after TC; (2) half the charge is burned at about 10° after TC. In practice, the spark is often retarded to give a 1 or 2 percent reduction in brake torque from the maximum value, to permit a more precise definition of timing relative to the optimum.

So far we have described normal combustion in which the spark-ignited flame moves *steadily* across the combustion chamber until the charge is fully consumed. However, several factors—e.g., fuel composition, certain engine design and operating parameters, and combustion chamber deposits—may prevent this normal combustion process from occurring. Two types of abnormal combustion have been identified: knock and surface ignition.

Knock is the most important abnormal combustion phenomenon. Its name comes from the noise that results from the autoignition of a portion of the fuel, air, residual gas mixture ahead of the advancing flame. As the flame propagates across the combustion chamber, the unburned mixture ahead of the flame—called the *end gas*—is compressed, causing its pressure, temperature, and density to increase. Some of the end-gas fuel-air mixture may undergo chemical reactions prior to normal combustion. The products of these reactions may then autoignite: i.e., spontaneously and rapidly release a large part or all of their chemical energy. When this happens, the end gas burns very rapidly, releasing its energy at a rate 5 to 25 times that characteristic of normal combustion. This causes high-frequency pressure oscillations inside the cylinder that produce the sharp metallic noise called knock.

The presence or absence of knock reflects the outcome of a race between the advancing flame front and the precombustion reactions in the unburned end gas. Knock will not occur if the flame front consumes the end gas before these reactions have time to cause the fuel-air mixture to autoignite. Knock will occur if the precombustion reactions produce autoignition before the flame front arrives.

The other important abnormal combustion phenomenon is *surface ignition*. Surface ignition is ignition of the fuel-air charge by overheated valves or spark plugs, by glowing combustion-chamber deposits, or by any other hot spot in the engine combustion chamber: it is ignition by any source other than normal spark ignition. It may occur before the spark plug ignites the charge (preignition) or after normal ignition (postignition). It may produce a single flame or many flames. Uncontrolled combustion is most evident and its effects most severe when it results from preignition. However, even when surface ignition occurs after the spark plug fires (postignition), the spark discharge no longer has complete control of the combustion process.

Surface ignition may result in knock. Knock which occurs following normal spark ignition is called spark knock. Knock which is preceded by surface ignition to distinguish it from knock which has been preceded by surface ignition. Abnormal combustion phenomena are reviewed in more detail in Sec. 9.6.

FIGURE 9-3
(a) Cylinder pressure versus crank angle for overadvanced spark timing (50°), MBT timing (30°), and retarded timing (10°). (b) Effect of spark advance on brake torque at constant speed and (A/F) at wide-open throttle. MBT is maximum brake torque timing.[5]

tions in the burned and unburned gas are then determined by conservation of mass:

$$\frac{V}{m} = \int_0^{x_b} v_b \, dx + \int_{x_b}^1 v_u \, dx \qquad (9.1)$$

and conservation of energy:

$$\frac{U_0 - W - Q}{m} = \int_0^{x_b} u_b \, dx + \int_{x_b}^1 u_u \, dx \qquad (9.2)$$

where V is the cylinder volume, m is the mass of the cylinder contents, v is the specific volume, x_b is the mass fraction burned, U_0 is the internal energy of the cylinder contents at some reference point θ_0, u is the specific internal energy, W is the work done on the piston, and Q is the heat transfer to the walls. The subscripts u and b denote unburned and burned gas properties, respectively. The work and heat transfers are

$$W = \int_{V_0}^V p \, dV'' \qquad Q = \int_{\theta_0}^{\theta} \left(\frac{\dot{Q}}{360N}\right) d\theta \qquad (9.3)$$

where $\dot{Q}$ is the instantaneous heat-transfer rate to the chamber walls.

To proceed further, models for the thermodynamic properties of the burned and unburned gases are required. Several categories of models are described in Chap. 4. Accurate calculations of the state of the cylinder gases require an equilibrium model (or good approximation to it) for the burned gas and an ideal gas mixture model (of frozen composition) for the unburned gas (see Table 4.2). However, useful illustrative results can be obtained by assuming that the burned and unburned gases are different ideal gases, each with constant specific heats; ie,

$$pv_b = R_b T_b \qquad u_b = c_{v,b} T_b + h_{f,b} \qquad (9.4)$$

$$pv_u = R_u T_u \qquad u_u = c_{v,u} T_u + h_{f,u} \qquad (9.5)$$

Combining Eqs. (9.1) to (9.5) gives

$$\frac{pV}{m} = x_b R_b T_b + (1 - x_b) R_u T_u \qquad (9.6)$$

and

$$\frac{U_0 - W - Q}{m} = x_b(c_{v,b} T_b + h_{f,b}) + (1 - x_b)(c_{v,u} T_u + h_{f,u}) \qquad (9.7)$$

where

$$T_b = \frac{1}{x_b} \int_0^{x_b} T_b \, dx \qquad T_u = \frac{1}{1 - x_b} \int_{x_b}^1 T_u \, dx$$

9.2 THERMODYNAMIC ANALYSIS OF SI ENGINE COMBUSTION

9.2.1 Burned and Unburned Mixture States

Because combustion occurs through a flame propagation process, the changes in state and the motion of the unburned and burned gas are much more complex than the ideal cycle analysis in Chapter 5 suggests. The gas pressure, temperature, and density change as a result of changes in volume due to piston motion. During combustion, the cylinder pressure increases due to the release of the fuel's chemical energy. As each element of fuel-air mixture burns, its density decreases by about a factor of four. This combustion-produced gas expansion compresses the unburned mixture ahead of the flame and displaces it toward the combustion chamber walls. The combustion-produced gas expansion also compresses those parts of the charge which have already burned, and displaces them back toward the spark plug. During the combustion process, the unburned gas elements move away from the spark plug; following combustion, individual gas elements move back toward the spark plug. Further, elements of the unburned mixture which burn at different times have different pressures and temperatures just prior to combustion, and therefore end up at different states after combustion. The thermodynamic state and composition of the burned gas is, therefore, non-uniform. A first law analysis of the spark-ignition engine combustion process enables us to quantify these gas states.

Consider the schematic of the engine cylinder while combustion is in progress, shown in Fig. 9-4. Work transfer occurs between the cylinder gases and the piston (to the gas before TC; to the piston after TC). Heat transfer occurs to the chamber walls, primarily from the burned gases. At the temperatures and pressures typical of spark-ignition engines it is a reasonable approximation to assume that the volume of the reaction zone where combustion is actually occurring is a negligible fraction of the chamber volume even though the thickness of-the turbulent flame may not be negligible compared with the chamber dimensions (see Sec. 9.3.2). With normal engine operation, at any point in time or crank angle, the pressure throughout the cylinder is close to uniform. The condi-
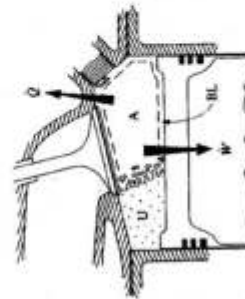
FIGURE 9-4
Schematic of flame in the engine cylinder during combustion: unburned gas (U) to left of flame, burned gas to right. A denotes adiabatic burned-gas core, BL denotes thermal boundary layer in burned gas, W is work-transfer rate to piston, Q is heat-transfer rate to chamber walls.

are the mean temperatures of the burned and unburned gases. Equations (9.6) and (9.7) may now be solved to obtain

$$x_b = \frac{pV - p_0 V_0 + (\gamma_b - 1)(W + Q) + (\gamma_b - \gamma_u)mc_{v,u}(\bar{T}_u - T_0)}{m[(\gamma_b - 1)(h_{f,u} - h_{f,b}) + (\gamma_b - \gamma_u)c_{v,u}\bar{T}_u]} \quad (9.8)$$

and

$$\bar{T}_b = \frac{R_u}{R_b}\bar{T}_u + \frac{pV - mR_u\bar{T}_u x_b}{mR_b x_b} \quad (9.9)$$

If we now assume the unburned gas is initially uniform and undergoes isentropic compression, then

$$\frac{\bar{T}_u}{T_0} = \left(\frac{p}{p_0}\right)^{(\gamma_u - 1)/\gamma_u} \quad (9.10)$$

This equation, with Eqs. (9.8) and (9.9) enables determination of both x_b and $\bar{T}_b$ from the thermodynamic properties of the burned and unburned gases, and known values of p, V, m, and Q. Alternatively, if x_b is known then p can be determined. Mass fraction burned and cylinder gas pressure are uniquely related.

While Eq. (9.9) defines a mean burned gas temperature, the burned gas is not uniform. Mixture which burns early in the combustion process is further compressed after combustion as the remainder of the charge is burned. Mixture which burns late in the combustion process is compressed prior to combustion and, therefore, ends up at a different final state. A temperature gradient exists across the burned gas with the earlier burning portions at the higher temperature.[7,8] Two limiting models bracket what occurs in practice: (1) a *fully mixed* model, where it is assumed that each element of mixture which burns mixes instantaneously with the already burned gases (which therefore have a uniform temperature), and (2) an *unmixed* model, where it is assumed that no mixing occurs between gas elements which burn at different times.

In the fully mixed model the burned gas is uniform, $T_b = \bar{T}_b$, and the equations given above fully define the state of the cylinder contents. In the unmixed model, the assumption is made that no mixing occurs between gas elements that burn at different times, and each burned gas element is therefore isentropically compressed (and eventually expanded) after combustion.† Thus:

$$\frac{T_b(x_b', x_b)}{T_b(x_b')} = \left[\frac{p(x_b)}{p(x_b')}\right]^{(\gamma_b - 1)/\gamma_b} \quad (9.11)$$

† This model applies to burned gas regions of the chamber away from the walls. Heat transfer to the walls results in a thermal boundary layer on the walls which grows with time. The gas in the boundary layer is not isentropically compressed and expanded.
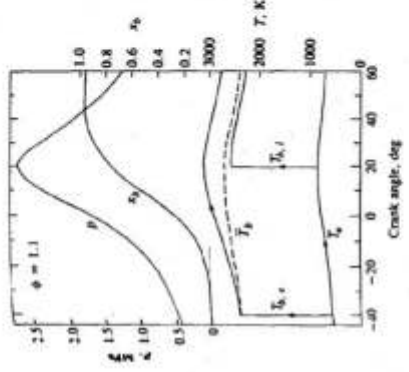
FIGURE 9-5
Cylinder pressure, mass fraction burned, and gas temperatures as functions of crank angle during combustion. T_u is unburned gas temperature, T_b is burned gas temperature, the subscripts e and l denote early and late burning gas elements, and $\bar{T}_b$ is the mean burned gas temperature.[8] (*Reprinted with permission. Copyright 1973, American Chemical Society.*)

where $T_b(x_b', x_b)$ is the temperature of the element which burned at the pressure $p(x_b')$ when the pressure is $p(x_b)$, and

$$T_b(x_b') = \frac{h_{f,u} - h_{f,b} + c_{p,u}T_u(x_b')}{c_{p,b}} \quad (9.12)$$

is the temperature resulting from isenthalpic combustion of the unburned gas at $T_u(x_b')$, $p(x_b')$. An example of the temperature distribution computed with this model is shown in Fig. 9-5. A mixture element that burns right at the start of the combustion process reaches, in the absence of mixing, a peak temperature after combustion about 400 K higher than an element that burns toward the end of the combustion process. The mean burned gas temperature is closer to the lower of these temperatures. These two models approximate respectively to situations where the time scale that characterizes the turbulent mixing process in the burned gases is (1) much less than the overall burning time (for the fully mixed model) or (2) much longer than the overall burning time (for the unmixed model). The real situation lies in between.

Measurements of burned gas temperatures have been made in engines using spectroscopic techniques through quartz windows in the cylinder head. Examples of measured temperatures are shown in Fig. 9-6. The solid lines marked A, B, and C are the burned gas temperatures measured by Rassweiler and Withrow[7] using the sodium line reversal technique in an L-head engine, for the spark plug end (A), the middle (B), and the opposite end (C) of the chamber, respectively. Curves labeled W_2 and W_3 were measured by Lavoie[8] through two different windows, W_2 and W_3 (with W_2 closer to the spark), again in an L-head engine. Each set of experimental temperatures shows a temperature gradient across the burned gas comparable to that predicted, and the two sets have similar shapes.

flame with a constant burning velocity propagates outward from the center of a spherical container. Applying this gas motion model to an engine, it can be concluded that a window in the cylinder head initially views earlier burned gas (of higher temperature and entropy) and that as more of the charge burns, the window views later burned gas of progressively lower entropy. The experimental curves fit this description: they cross the constant entropy lines toward lower entropy. Note that the gradient in temperature persists well into the expansion stroke, indicating that the "unmixed" model is closer to reality than the "fully mixed" model.

More accurate calculations relating the mass fraction burned, gas pressure, and gas temperature distribution are often required. Note that the accuracy of such calculations depends on the accuracy with which the time-varying heat loss to the chamber walls can be estimated (see Sec. 12.4.3) and whether flows into and out of crevice regions are significant (see Sec. 8.6), as well as the accuracy of the models used to describe the thermodynamic properties of the gases. Appropriate more accurate models for the thermodynamic properties are: an equilibrium model for the burned gas, and specific heat models which vary with temperature for each of the components of the unburned mixture (see Secs. 4.1 and 4.7). In the absence of significant crevice effects, Eqs. (9.1) and (9.2) can be written as

$$\frac{V}{m} = \bar{v}_b x_b + \bar{v}_u(1 - x_b) \tag{9.13}$$

$$\frac{U_0 - W - Q}{m} = \bar{u}_b x_b + \bar{u}_u(1 - x_b) \tag{9.14}$$

where

$$\bar{v}_b = \frac{1}{x_b}\int_0^{x_b} v_b\, dx \quad \text{and} \quad \bar{v}_u = \frac{1}{1 - x_b}\int_{x_b}^1 v_u\, dx$$

and similar definitions hold for $\bar{u}_b$ and $\bar{u}_u$. For a given equivalence ratio, fuel and burned gas fraction:

$$h_u = h_u(T) \qquad\qquad h_b = h_b(T_b, p) \tag{9.15a, b}$$

$$pv_u = \left(\frac{R}{M_u}\right)T_u \qquad pv_b = \left(\frac{R}{M_b}\right)T_b \tag{9.16a, b}$$

and

$$\bar{u}_u = \bar{h}_u - p\bar{v}_u \qquad \bar{u}_b = \bar{h}_b - p\bar{v}_b \tag{9.17a, b}$$

To simplify the calculations, it is convenient to assume that, for the burned gas, $\bar{u}_b = u_b(T_b, p)$ and $\bar{v}_b = v_b(T_b, p)$. This corresponds to the fully mixed assumption described above. The effect of neglecting the temperature distribution in the calculation of mass fraction burned is small. In addition, the heat losses from the unburned gas can usually be neglected; the unburned gas is then compressed isentropically. T_u is specified for some initial state of the unburned gas (where

---

FIGURE 9-6
Burned gas temperatures measured using spectroscopic techniques through windows in the cylinder head, as a function of cylinder pressure. Temperatures measured closer to spark plug have higher values. Dashed lines show isentropic behavior.[7,8]

In the unmixed model, the temperature of each burned gas element follows a different isentropic line as it is first compressed as p increases to p_{max} and then expanded as the pressure falls after p_{max}. The measured temperature curves in Fig. 9-6 do not follow the calculated isentropes because of gas motion past the observation ports. As has already been mentioned, the expansion of a gas element which occurs during combustion compresses the gas ahead of the flame and moves it away from the spark plug. At the same time, previously burned gas is compressed and moved back toward the spark plug. Defining this motion in an engine requires sophisticated flow models, because the combustion chamber shape is rarely symmetrical, the spark plug is not usually centrally located, and often there is a bulk gas motion at the time combustion is initiated. However, the gas motion in a spherical or cylindrical combustion bomb with central ignition which can readily be computed illustrates the features of the combustion-induced motion in an engine. Figure 9-7 shows calculated particle trajectories for a stoichiometric methane-air mixture, initially at ambient conditions, as a laminar

FIGURE 9-7
Particle trajectories in unburned and burned gas as flame propagates outward at constant velocity from the center of a spherical combustion bomb. Stoichiometric methane-air mixture initially at 1 atm and 300 K.

FIGURE 9-9
Calculated temperature distribution in the adiabatic core of the burned gas zone for the unmixed model assuming thermodynamic equilibrium. $\phi = 1.0$. Dashed line is temperature of each element just after it burns.

divided into an adiabatic core and a boundary layer that grows in thickness with time. In the adiabatic core, in the absence of mixing between gas elements that burn at different times, burned gas is compressed and then expanded isentropically. The burned gas temperature distribution can be calculated as follows. Given the pressure versus crank angle data, the unburned mixture state can be determined using Eq. (9.18) above. Each small element of unburned mixture burns in a constant-enthalpy constant-pressure process. So the burned state of an element of unburned charge, which burns at $p = p_i$, can be obtained from the relation

$$h_b(T_{b,i}, p_i) = h_u(T_{u,i}, p_i)$$

After combustion, this element which burned at $p = p_i$ is compressed and expanded along the isentropic:

$$s_b(T_b, p) = s_b(T_{b,i}, p_i)$$

An example of the temperature distribution computed in this manner for this unmixed model in the burned gas adiabatic core is shown in Fig. 9-9. The element ignited by the spark is compressed to the highest peak temperature at T_{max}. The temperature difference across the bulk of the charge ($0.05 < x_b < 0.95$) is about 200 K.

9.2.2 Analysis of Cylinder Pressure Data

Cylinder pressure changes with crank angle as a result of cylinder volume change, combustion, heat transfer to the chamber walls, flow into and out of crevice regions, and leakage. The first two of these effects are the largest. The effect of

---

$x_b = 0$) by p_0, V_0, M_u, and the mass of charge m. Then, since for any isentropic process

$$\left(\frac{\partial T}{\partial p}\right)_s = \frac{v - (\partial h/\partial p)_T}{(\partial h/\partial T)_p}$$ (9.18)

T_u can be determined.

Equations (9.13) to (9.18) constitute a set of nine equations for the nine unknowns $\bar{v}_u$, $\bar{v}_b$, $\bar{u}_u$, $\bar{u}_b$, $\bar{s}_u$, $\bar{s}_b$, T_u, T_b, and x_b or p. One convenient solution method is to eliminate x_b from Eqs. (9.13) and (9.14) to obtain

$$\frac{(V/m) - \bar{v}_u}{\bar{v}_b - \bar{v}_u} - \frac{(U/m) - \bar{u}_u}{\bar{u}_b - \bar{u}_u} = f(T_b, T_u) = 0$$ (9.19)

where $U = U_0 - W - Q$. T_u can be determined from Eq. (9.18). Equation (9.19) can then be solved using an appropriate iterative technique for T_b, and x_b can be obtained from Eq. (9.13). An alternative formulation based on the rate of change of pressure $dp/d\theta$ and equations for $dT_u/d\theta$, $dT_b/d\theta$, $dm_b/d\theta$, and $dV_b/d\theta$ can be found in Ref. 10. Some examples of mass fraction burned curves obtained from measured pressure data, with gasoline and methanol fuels, are shown in Fig. 9-8. With accurate pressure versus crank angle records, values of final mass fraction burned should be close to but lower than unity, usually in the range 0.93 to 0.96; the difference from unity is the combustion inefficiency for lean mixtures (see Figs. 3-9) and incomplete oxygen utilization for rich mixtures (see Fig. 4-20).

More accurate burned gas temperature calculations need to account for the presence of a thermal boundary layer (of order 1 mm thick) around the combustion chamber walls (see Sec. 12.6.5). The burned gas region in Fig. 9-4 can be

FIGURE 9-8
Mass fraction burned curves determined from measured cylinder pressure data using two-zone combustion model: (a) gasoline; (b) methanol. ϕ = fuel/air equivalence ratio.[11]

volume change on the pressure can readily be accounted for; thus, combustion rate information can be obtained from accurate pressure data provided models for the remaining phenomena can be developed at an appropriate level of approximation. The previous section has developed the fundamental basis for such calculations.

Cylinder pressure is usually measured with piezoelectric pressure transducers. This type of transducer contains a quartz crystal. One end of the crystal is exposed through a diaphragm to the cylinder pressure; as the cylinder pressure increases, the crystal is compressed and generates an electric charge which is proportional to the pressure. A charge amplifier is then used to produce an output voltage proportional to this charge. Accurate cylinder pressure versus crank angle data can be obtained with these systems provided the following steps are carried out: (1) the correct reference pressure is used to convert the measured pressure signals to absolute pressures; (2) the pressure versus crank angle (or volume) phasing is accurate to within about 0.2°; (3) the clearance volume is estimated with sufficient accuracy; (4) transducer temperature variations (which can change the transducer calibration factor) due to the variation in wall heat flux during the engine cycle are held to a minimum. Log p versus log V plots can be used to check the quality of cylinder pressure data. The first three of the above requirements can be validated using log p–log V diagrams for a motored engine. If the effects of thermal cycling are significant, the expansion stroke on the log p–log V plot for a firing engine shows excessive curvature.[12]

Figure 9-10 shows pressure-volume data from a firing spark-ignition engine on both a linear p–V and a log p–log V diagram. On the log p–log V diagram the compression process is a straight line of slope 1.3. The start of combustion can be identified by the departure of the curve from the straight line. The end of com-

FIGURE 9-10
(a) Pressure-volume diagram; (b) log p–log (V/V_{max}) plot. 1500 rev/min, MBT timing, imep = 513 kPa. $\phi = 0.8$, $r_c = 8.72$, propane fuel.[12]

bustion can be located approximately in similar fashion; the expansion stroke following combustion is essentially linear with slope 1.33. Since both the compression of the unburned mixture prior to combustion and the expansion of the burned gases following the end of combustion are close to adiabatic isentropic processes (for which pV^γ = constant; $\gamma = c_p/c_v$), the observed behavior is as expected. More extensive studies[12,13] show that the compression and expansion processes are well fitted by a polytropic relation:

$$pV^n = \text{constant} \quad (9.20)$$

The exponent n for the compression and expansion processes is 1.3 (±0.05) for conventional fuels. It is comparable to the average value of γ, for the unburned mixture over the compression process, but is larger than γ_b for the burned gas mixture during expansion due to heat loss to the combustion chamber walls (see Figs. 4-13 and 4-16).

Log p–log V plots such as Fig. 9-10 approximately define the start and end of combustion, but do not provide a mass fraction burned profile. One well-established technique for estimating the mass fraction burned profile from the pressure and volume data is that developed by Rassweiler and Withrow.[3] They correlated cylinder pressure data with flame photographs, and showed how Eq. (9.20) could be used to account for the effect of cylinder volume change on the pressure during combustion. Assuming that the unburned gas filling the volume V_u ahead of the flame at any crank angle during combustion has been compressed polytropically by the advancing flame front, then the volume $V_{u,0}$ it occupied at time of spark is

$$V_{u,0} = V_u \left(\frac{p}{p_0}\right)^{1/n} \quad (9.21)$$

Similarly, the burned gas behind the flame filling the volume V_b would, at the end of combustion, fill a volume $V_{b,f}$ given by

$$V_{b,f} = V_b \left(\frac{p}{p_f}\right)^{1/n} \quad (9.22)$$

The mass fraction burned x_b is equal to $1 - (V_{u,0}/V_0)$ and to $V_{b,f}/V_f$, where V_0 and V_f are the total cylinder volumes at time of spark and at the end of combustion, respectively. Since $V = V_u + V_b$, Eqs. (9.21) and (9.22) then give:

$$x_b = \frac{p^{1/n}V - p_0^{1/n}V_0}{p_f^{1/n}V_f - p_0^{1/n}V_0} \quad (9.23)$$

This method is widely used, though it contains several approximations. Heat-transfer effects are included only to the extent that the polytropic exponent n in Eq. (9.22) differs appropriately from γ. The pressure rise due to combustion is proportional to the amount of fuel chemical energy released rather than the mass of mixture burned. Also, the polytropic exponent n is not constant during combustion. Selecting an appropriate value for n (whether n is assumed to be con-

stant or to vary through the combustion process) is the major difficulty in applying this pressure data analysis procedure.
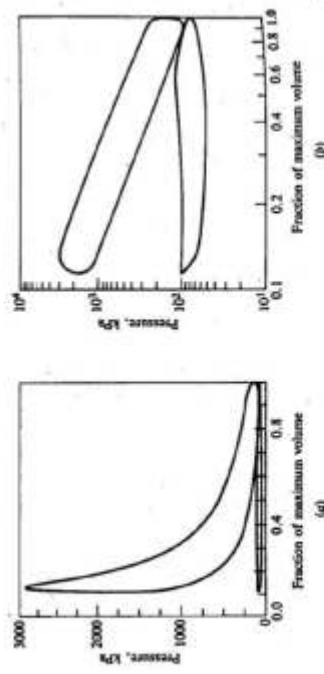
The effects of heat transfer, crevices, and leakage can be explicitly incorporated into cylinder pressure data analysis by using a "heat release" approach based on the first law of thermodynamics. A major advantage of such an approach is that the pressure changes can be related directly to the amount of fuel chemical energy released by combustion, while retaining the simplicity of treating the combustion chamber contents as a single zone. Figure 9-11 shows the appropriate open-system boundary for the combustion chamber.[14] The first law for this open system is

$$\delta Q_{ch} = dU_s + \delta Q_{ht} + \delta W + \sum h_i\, dm_i \quad (9.24)$$

The change in sensible energy of the charge dU_s is separated from that due to change in composition: the term δQ_{ch} represents the "chemical energy" released by combustion. The work is piston work and equal to $p\, dV$. δQ_{ht} is heat transfer to the chamber walls. The mass flux term represents flow across the system boundary. In the absence of fuel injection, it represents flow into and out of crevice regions (see Sec. 8.6).

The accuracy with which this energy balance can be made depends on how adequately each term in the above equation can be quantified. Assuming that U_s is given by $mu_s(T)$, where T is the mean charge temperature and m is the mass within the system boundary, then

$$dU_s = mc_v(T)\, dT + u(T)\, dm$$

Note that this mean temperature determined from the ideal gas law is close to the mass-averaged cylinder temperature during combustion since the molecular weights of the burned and unburned gases are essentially identical. Crevice effects can usually be modeled adequately by flow into and out of a single volume at the cylinder pressure, with the gas in the crevice at a substantially lower temperature. Leakage to the crankcase can usually be neglected. Then Eq. (9.24), on substituting for dU_s and dm_i ($=dm_{cr} = -dm$), becomes

$$\delta Q_{ch} = mc_v\, dT + (h' - u)dm_{cr} + p\, dV + \delta Q_{ht} \quad (9.25)$$
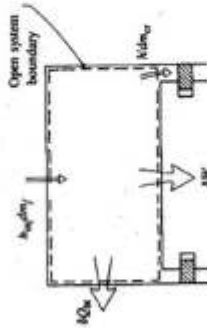
FIGURE 9-11
Open system boundary for combustion chamber for heat-release analysis.

where

$dm_{cr} > 0$ when flow is out of the cylinder into the crevice
$dm_{cr} < 0$ when flow is from the crevice to the cylinder
h' is evaluated at cylinder conditions when $dm_{cr} > 0$ and at crevice conditions when $dm_{cr} < 0$

Use of the ideal gas law (neglecting the change in gas constant R) with Eq. (9.25) then gives

$$\delta Q_{ch} = \left(\frac{c_v}{R}\right)V\, dp + \left(\frac{c_v}{R} + 1\right)p\, dV + (h' - u + c_v T)\, dm_{cr} + \delta Q_{ht} \quad (9.26)$$

This equation can be used in several ways. When the heat or energy release term, δQ_{ch}, is combined with the heat-transfer and crevice terms, the combination is termed net heat release—the combustion energy release less heat lost to the walls. It is equal to the first two terms on the right-hand side of Eq. (9.26) which, together, represent the sensible energy change and work transfer to the piston. While heat losses during combustion are a small fraction of the fuel energy (10 to 15 percent), the distributions of heat release and heat transfer with crank angle are different; heat transfer becomes more important as the combustion process ends and average gas temperatures peak. The net heat-release profile obtained from integrating the first two terms on the right-hand side of Eq. (9.26), normalized to give unity at its maximum value, is often interpreted as the burned mass fraction (or, more correctly, the energy-release fraction) versus crank angle profile.

Use of Eq. (9.26) requires a value for c_v/R [$= 1/(\gamma - 1)$]. The ratio of specific heats γ for both unburned and burned gases decreases with increasing temperature and varies with composition (see Figs. 4-13, 4-16, and 4-18). As the mean charge temperature increases during compression and combustion and then decreases during expansion, γ should vary. An approximate approach, modeling $\gamma(T)$ with a linear function of temperature fitted to the appropriate curves in Figs. 4-13, 4-16, and 4-18 and with γ constant during combustion, has been shown to give adequate results.[15]

The convective heat-transfer rate to the combustion chamber walls can be calculated from the relation

$$\frac{dQ_{ht}}{dt} = Ah_c(T - T_w)$$

where A is the chamber surface area, T is the mean gas temperature, T_w is the mean wall temperature, and h_c is the heat-transfer coefficient (averaged over the chamber surface area). h_c can be estimated from engine heat-transfer correlations (see Sec. 12.4.3). Since crevice effects are usually small, a sufficiently accurate model for their overall effect is to consider a single aggregate crevice volume where the gas is at the same pressure as the combustion chamber, but at a different temperature. Since these crevice regions are narrow, an appropriate assumption is that the crevice gas is at the wall temperature. Inserting this crevice model

into Eq. (9.26), with $\gamma(T) = a + bT$, gives the *chemical energy-* or *gross heat-release rate*:

$$\frac{dQ_{ch}}{d\theta} = \frac{\gamma}{\gamma-1} p \frac{dV}{d\theta} + \frac{1}{\gamma-1} V \frac{dp}{d\theta}$$
$$+ V_{cr}\left[\frac{T'}{T} + \frac{T}{T_w(\gamma-1)} + \frac{1}{bT_w}\ln\left(\frac{\gamma-1}{\gamma'-1}\right)\right]\frac{dp}{d\theta} + \frac{dQ_{ht}}{d\theta} \qquad (9.27)$$

An example of the use of Eq. (9.27) to analyze an experimental pressure versus crank angle curve for a conventional spark-ignition engine is shown in Fig. 9-12. The integrated heat release is plotted against crank angle. The lowest curve shown is the net heat release. The curve at the top of the figure is the mass of fuel within the combustion chamber times its lower heating value. It decreases slightly as p_{max} is approached due to flow into crevices. The difference between the final value of Q_{ch} and $(m_f Q_{LHV})$ should equal the combustion inefficiency (which is a few percent of $m_f Q_{LHV}$). The combustion inefficiency can be determined from the exhaust gas composition (see Sec. 4.9.4). Inaccuracies in the cylinder pressure data and the heat-release calculation will also contribute to this difference. An important advantage of a heat-release analysis that relates the pressure changes to the amount of fuel chemical energy within the cylinder is that this error can be determined. In the example in Fig. 9-12, the measured combustion inefficiency was close to the amount shown.

Two-zone models (one zone representing the unburned mixture ahead of the flame and one the burned mixture behind the flame) are used to calculate the mass fraction burned profile from measured cylinder pressure data.[10] Figure 9-8 shows results from such an analysis, using the methodology described in Sec. 9.2.1. The advantage of a two-zone analysis is that the thermodynamic properties of the cylinder contents can be quantified more accurately. The disadvantages are that the unburned and the burned zone heat-transfer areas must both now be estimated, and a model for the composition of the gas flowing into the crevice
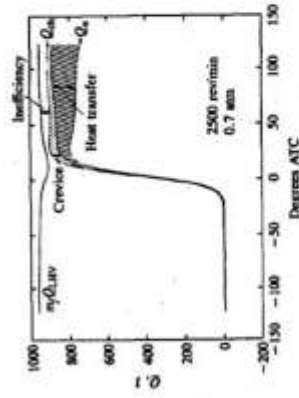
FIGURE 9-12
Results of heat-release analysis showing the effects of heat transfer, crevices, and combustion inefficiency.[14]

region must be developed. Due to this complexity, crevice models are usually omitted despite the fact that their impact can be significant.

9.2.3 Combustion Process Characterization

The mass fraction burned profiles as a function of crank angle in each individual cycle shown in Fig. 9-2 and the chemical energy- or gross heat-release curve in Fig. 9-12 have a characteristic S-shape. The rate at which fuel-air mixture burns increases from a low value immediately following the spark discharge to a maximum about halfway through the burning process and then decreases to close to zero as the combustion process ends. It proves convenient to use these mass fraction burned or energy-release fraction curves to characterize different stages of the spark-ignition engine combustion process by their duration in crank angles, thereby defining the fraction of the engine cycle that they occupy. The flame development process, from the spark discharge which initiates the combustion process to the point where a small but measurable fraction of the charge has burned, is one such stage. It is influenced primarily by the mixture state, composition, and motion in the vicinity of the spark plug (see Sec. 9.3). The stage during which the major portion of the charge burns as the flame propagates to the chamber walls is next. This stage is obviously influenced by conditions throughout the combustion chamber. The final stage, where the remainder of the charge burns to completion, cannot as easily be quantified because energy-release rates are comparable to other energy-transfer processes that are occurring.

The following definitions are most commonly used to characterize the energy-release aspects of combustion:

Flame-development angle $\Delta\theta_d$. The crank angle interval between the spark discharge and the time when a small but significant fraction of the cylinder mass has burned or fuel chemical energy has been released. Usually this fraction is 10 percent, though other fractions such as 1 and 5 percent have been used.†

Rapid-burning angle $\Delta\theta_b$. The crank angle interval required to burn the bulk of the charge. It is defined as the interval between the end of the flame-development stage (usually mass fraction burned or energy-release fraction of 10 percent) and the end of the flame-propagation process (usually mass fraction burned or energy-release fraction of 90 percent).‡

Overall burning angle $\Delta\theta_o$. The duration of the overall burning process. It is the sum of $\Delta\theta_d$ and $\Delta\theta_b$.

---

† This angle is sometimes called the *ignition delay*. Since the flame starts to propagate outward immediately following the spark discharge there is no delay, and the terminology used here is preferred (see Sec. 9.3).

‡ An alternative definition for $\Delta\theta_b$ uses the maximum burning rate to define an angle or time characteristic of the bulk charge burning process[4] (see Fig. 9-13). $\Delta\theta_d$ and $\Delta\theta_b$ are usually closely comparable.

flame can be obtained from photographs taken with techniques that are sensitive to density changes in the flow field, such as schlieren and shadowgraph. With these techniques, a parallel light beam is passed through the combustion chamber. Portions of the beam which pass through regions where density gradients normal to the beam exist are deflected, due to the refractive index gradients that result from the density gradients. In the schlieren technique, the beam is focused on a knife edge; the deflected parts of the beam are displaced relative to the knife edge and produce lighter or darker regions when subsequently refocused onto film. With the shadowgraph technique, the parallel beam emerging from the combustion chamber is photographed directly; deflected parts of the beam produce lighter and darker regions on the film. With these techniques, details of flame structure can be discerned.

Figure 9-14 shows a set of photographs from one engine cycle, from a high-speed schlieren movie taken in a special visualization spark-ignition engine operating at 1400 rev/min and 0.5 atm inlet pressure. Also shown are the cylinder pressure versus crank angle data, and the mass fraction burned profile calculated from the pressure data using the method of Rassweiler and Withrow[2] (see Sec. 9.2.2). This engine had a square-cross-section cylinder with two quartz walls to permit easy optical access, but otherwise operated normally.[17] Visualization of

FIGURE 9-14
Sequence of movie frames from one engine cycle in a square-cross-section cylinder, single-cylinder, engine with two glass walls, and corresponding pressure and mass fraction burned curves. 1400 rev/ min, 0.5 atm inlet pressure.

FIGURE 9-13
Definition of flame-development angle, $\Delta\theta_d$, rapid-burning angle, $\Delta\theta_b$, on mass fraction burned versus crank angle curve.

Figure 9-13 illustrates these definitions on a mass fraction burned, or fraction of fuel energy released, versus crank angle plot. While the selection of the 10 and 90 percent points is arbitrary, such a choice avoids the difficulties involved in determining accurately the shape of the curve at the start and end of combustion. These angles can be converted to times (in seconds) by dividing by $6N$ (with N in revolutions per minute).

A functional form often used to represent the mass fraction burned versus crank angle curve is the Wiebe function:

$$x_b = 1 - \exp\left[-a\left(\frac{\theta - \theta_0}{\Delta\theta}\right)^{m+1}\right] \qquad (9.28)$$

where θ is the crank angle, θ_0 is the start of combustion, $\Delta\theta$ is the total combustion duration ($x_b = 0$ to $x_b = 1$), and a and m are adjustable parameters. Varying a and m changes the shape of the curve significantly. Actual mass fraction burned curves have been fitted with $a = 5$ and $m = 2$.[16]

9.3 FLAME STRUCTURE AND SPEED

9.3.1 Experimental Observations

The combustion process in the spark-ignition engine takes place in a turbulent flow field. This flow field is produced by the high shear flows set up during the intake process and modified during compression, as described in Chap. 8. The importance of the turbulence to the engine combustion process was recognized long ago through experiments where the intake event, and the turbulence it generates, was eliminated and the rate of flame propagation decreased substantially. Understanding the structure of this engine flame as it develops from the spark discharge and the speed at which it propagates across the combustion chamber, and how that structure and speed depend on charge motion, charge composition, and chamber geometry, are critical to engine optimization. This section reviews experimental evidence that describes the essential features of the flame development and propagation processes.

Direct flame photographs such as those in Fig. 9-1 indicate the location and shape of the actual reaction zone which radiates in the blue region of the visible spectrum. An irregular front is apparent. Further insight into the structure of the

the flame is especially important during the early stages of flame development when the pressure rise due to combustion is too small to be detected.

These photographs show how the flame "ball," roughly spherical in shape, grows steadily from the time of spark discharge. The effect of turbulence is already visible in the convoluted flame surface in Fig. 9-14a. The volume enflamed behind the front continues to grow in a roughly spherical manner, except where intersected by the chamber walls, as seen in Fig. 9-14b and c. The mass fraction burned and the associated pressure rise due to combustion become significant by the time the flame front has traversed two-thirds to three-quarters of the field of view. Note that the fraction of the cylinder filled with enflamed charge is less than is suggested by the photos because the front of the flame is approximately spherical and the cylinder has a square cross section. Maximum cylinder pressure occurs close to the time the flame makes contact with the far wall, as seen in Fig. 9-14e. Finally, the unburned mixture ahead of and within the front burns out and the density gradients associated with the flame reaction zone disappear, clearing the field of view as shown in Fig. 9-14e and f.

A useful relationship between the mass fraction burned, $x_b (= m_b/m)$, and the volume fraction occupied by the burned gas, $y_b (= V_b/V)$, can be obtained from the identities

$$m = m_u + m_b \qquad V = V_u + V_b$$

and the ideal gas law:

$$x_b = \left[1 + \frac{\rho_b}{\rho_u}\left(\frac{1}{y_b} - 1\right)\right]^{-1} \qquad (9.29)$$

While the density ratio (ρ_u/ρ_b) does depend on the equivalence ratio, burned gas fraction in the unburned mixture, gas temperature, and pressure, its value is close to 4 for most spark-ignition engine operating conditions. Thus, the plot of x_b against y_b has a universal form,[10] as shown in Fig. 9-15. This curve is an important aid in interpreting flame geometry information.
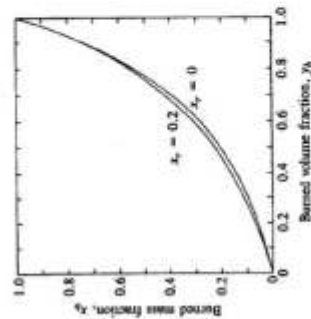
FIGURE 9-15
Relation between mass fraction burned x_b and volume fraction burned y_b. x_r is residual mass fraction.[4]

Crank angle, degrees from ignition
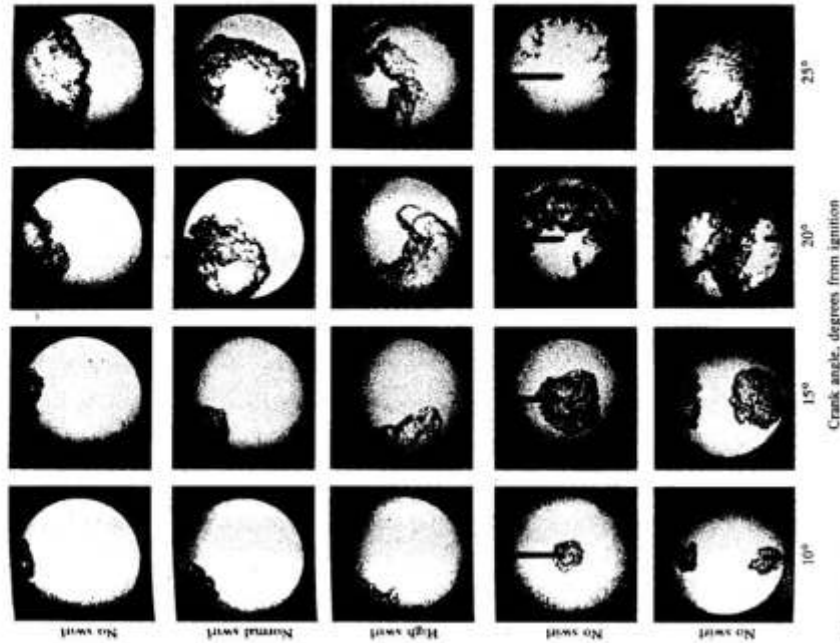
FIGURE 9-16
Laser shadowgraph photographs of engine combustion process taken in single-cylinder engine with transparent cylinder head. From top to bottom: side plug without swirl; side plug with normal swirl; side plug with high swirl; central plug without swirl; two plugs without swirl.[18]

The above-described features of the developing and propagating flame are common to almost all engine geometries and operating conditions. Figure 9-16 shows shadowgraph photographs of the flame at fixed crank angle intervals after ignition, taken through a transparent cylinder head with different geometric and flow configurations.[18] The approximately spherical development of the flame

from the vicinity of the spark plug, except where it intercepts the chamber walls, is evident for side and center ignition with one plug, and for ignition with two plugs in the absence of any intake generated swirl. With normal levels of swirl, the flame center is convected with the swirling flow, but the flame front as it grows is still approximately spherical in shape. Only with unusually high levels of swirl and aerodynamic stabilization of the flame at the spark plug location does the flame become stretched out and distorted by the flow in a major way.[18]

At any given flame radius, the geometry of the combustion chamber and the spark plug location govern the flame front surface area—the area of the approximately spherical surface corresponding to the leading edge of the flame contained by the piston, cylinder head, and cylinder wall. The larger this surface area, the greater the mass of fresh charge that can cross this surface and enter the flame zone. The photos in Fig. 9-16 illustrate the importance of flame area. The center plug location gives approximately twice the flame area of the side plug geometry at a given flame radius, and burns about twice as fast, at a fixed crank angle interval after spark). The arrangement with two spark plugs at opposite sides of the chamber is not significantly different in enflamed volume from the single center plug because, once the flame fronts are intersected by the cylinder wall, the flame front areas are comparable.

Mixture burning rate is strongly influenced by engine speed. It is well established that the duration of combustion in crank angle degrees only increases slowly with increasing engine speed.[19] Figure 9-17 shows how the interval between the spark discharge and 10 percent mass fraction burned, the flame development angle $\Delta\theta_d$, and the interval between the spark and 90 percent mass fraction burned, the overall burning angle $\Delta\theta_d + \Delta\theta_b$ (see Sec. 9.2.3), vary with engine speed.[20] Both intervals increase by a factor of about 1.6 for a factor of 4 increase in engine speed; i.e., the burning rate throughout the combustion process increases almost, though not quite, as rapidly as engine speed. Additionally, at a given engine speed, increasing in-cylinder gas velocities (e.g., with intake generated swirl) increases the burning rate: the flame size for the swirling flows in Fig. 9-16 is larger than for the quiescent case with the same plug location at the crank angle intervals after spark shown. Increasing engine speed and introducing swirl both increase the levels of turbulence in the engine cylinder at the time of combustion (see Sec. 8.2.2). Increased turbulence increases the rate of development and propagation of the turbulent premixed engine flame.
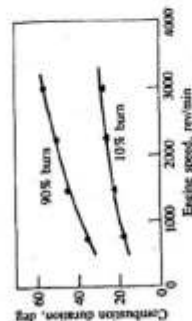
FIGURE 9-17
Effect of engine speed on flame-development angle (0 to 10 percent burned) and overall burning angle (0 to 90 percent burned). $\phi = 1$, intake pressure 0.54 atm, spark 30° BTC.[20]

It is also well established that unburned mixture composition and state affect the burning rate. Reducing the inlet pressure (and maintaining the ratio of exhaust to inlet pressure fixed to hold the residual gas fraction constant) increases both the flame development and rapid burning angles.[19] The fuel/air equivalence ratio affects the burning rate. Both flame development and burning angles show a minimum for slightly rich mixtures ($\phi \approx 1.2$) and increase significantly as the mixture becomes substantially leaner than stoichiometric.[19,20] The burned gas fraction in the unburned mixture, due to the residual gas fraction *and* any recycled exhaust gases, affects the burning rate: increasing the burned gas fraction slows down both flame development and propagation.[20] Fuel composition changes can be significant also. While mixtures of isooctane or conventional gasolines with air and burned gases (at identical conditions) have closely comparable burning rates, propane, methane, methanol, and ethanol mixtures exhibit modest differences in burning rate and hydrogen-air mixtures substantial differences. The basic combustion chemistry of the fuel, air, burned gas mixture influences the combustion process. However, the relative importance of combustion chemistry effects depends on combustion chamber design and burn rate. Faster burning engines (which have higher turbulence) are less sensitive to changes in mixture composition, pressure, and temperature than are slower burning engines (which have lower turbulence). The effects of chamber geometry, gas motion, and gas composition and state are interrelated.[21]

9.3.2 Flame Structure

Laminar flames in premixed fuel, air, residual gas mixtures are characterized by a laminar flame speed S_L and a laminar flame thickness δ_L (see Sec. 9.3.3). The laminar flame speed is the velocity at which the flame propagates into quiescent premixed unburned mixture ahead of the flame. There are several ways to define the thickness of a laminar flame.[22] Given the molecular diffusivity D_L (see Sec. 4.8), dimensional arguments give the most commonly used definition: $\delta_L = D_L/S_L$. Turbulent flames are also characterized by the root mean square velocity fluctuation, the turbulence intensity u' [Eq. (8.3)], and the various length scales of the turbulent flow ahead of the flame. The integral length scale l_I [Eq. (8.8)] is a measure of the size of the large energy-containing structures of the flow. The Kolmogorov scale l_K [Eq. (8.11)] defines the smallest structures of the flow where small-scale kinetic energy is dissipated via molecular viscosity.

Several dimensionless parameters are used to characterize turbulent premixed flames. The dimensionless parameter used to define the turbulence is the turbulent Reynolds number, $Re_T = u'l_I/\nu$. For homogeneous and isotropic (no preferred direction) turbulence, the integral and Kolmogorov scales are related by Eq. (8.14): $l_K/l_I = Re_T^{-3/4}$. A characteristic turbulent eddy turnover time τ_T can be defined as

$$\tau_T = \frac{l_I}{u'}$$

A characteristic chemical reaction time is the residence time in a laminar flame:

$$\tau_L = \frac{\delta_L}{S_L}$$

The ratio of the characteristic eddy turnover time to the laminar burning time is called the *Damköhler number*:

$$Da = \frac{\tau_T}{\tau_L} = \left(\frac{l_I}{\delta_L}\right)\left(\frac{S_L}{u'}\right) \quad (9.30)$$

It is an inverse measure of the influence of the turbulent flow on the chemical processes occurring in the flame. Other ratios are of interest. The ratio δ_L/l_K is a measure of the stretch or local distortion to which a laminar flame is subjected by the turbulent flow. Unless $l_I/\delta_L \gg 1$ the concept of a localized flame region has little significance. The ratio u'/S_L is a measure of the relative strength of the turbulence.[22]

Different regimes of turbulent flames are apparent in the plot of Damköhler number versus turbulent Reynolds number in Fig. 9-18.[22] It has been assumed that $D_L \approx v$ and that the relationships for homogeneous isotropic turbulence are valid. Two regimes—distributed reactions and reaction sheets—are normally identified. In the distributed reaction regime, chemical reactions proceed in distributed reaction zones and thin-sheet flames do not occur. A sufficient condition for this regime is $l_I \ll \delta_L$. In the reaction sheet regime, propagating reaction fronts are wrinkled and convoluted by the turbulence. A sufficient condition for the existence of reaction sheets is $l_K > \delta_L$. For $Re_T > 1$, there is a region in Fig. 9-18 where $l_I > \delta_L > l_K$: the characteristics of flames in this regime are unclear.
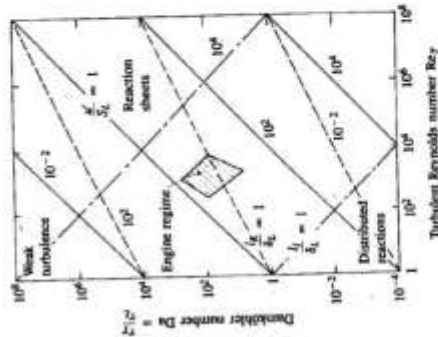
FIGURE 9-18
Different turbulent flame regimes on plot of Damköhler number versus turbulence Reynolds number; u' is turbulence intensity; S_L is laminar flame speed; l_I, l_K, and δ_L are integral scale, Kolmogorov scale, and laminar flame thickness. (From Abraham et al.[22])

Values of Da and Re_T for a typical spark-ignition engine (the cross-hatched region in Fig. 9-18) lie predominantly in the reaction sheet flame regime. Engine operation at high speed (the lower right boundary) and low load (the lower left boundary) gives values of Da and Re_T which fall below the $l_K/\delta_L = 1$ line. This is largely due to the low values of laminar flame speed that result from the high amounts of residual gas and EGR under these conditions (see Sec. 9.3.4 and Fig. 6-19). Whether the flame structure under these conditions is significantly different is not known. Observations of engine flames to date, described below, lie above the $l_K/\delta_L = 1$ line, within the reaction sheet regime. One would expect, then, the structure of the flame in a spark-ignition engine, once developed, to be that of a thin reaction sheet wrinkled and convoluted by the turbulent flow.

Detailed observations have been made of flame structure from ignition to flame extinguishing at the far cylinder wall. A flame develops from the spark discharge which causes ignition as follows. In the initial breakdown phase of ignition, a cylindrical discharge between the spark plug electrodes is established.[23] As electrical energy is fed into the discharge, the arc expands and exothermic chemical reactions capable of sustaining a propagating flame develop. Figure 9-19 shows how this development of a flame kernel occurs, with a set of shadowgraph photographs taken at 40-μs intervals of the spark plug electrode gap in one cylinder of a 2-liter conventional engine. The first photograph is between 20 and 50 μs after the spark breakdown occurred. The complete sequence (~200 μs) corresponds to 1.3 crank angle degrees. The outer boundary of this developing flame kernel is approximately spherical and is smooth with modest irregularities, corresponding to a thin reaction zone with high-temperature gases inside. As this developing sheetlike flame grows it interacts with the turbulent flow field in the vicinity of the spark plug: the flame outer surface becomes increasingly convoluted and the flame center can be convected away from the plug in a direction and with a velocity that can vary substantially cycle-by-cycle, as seen in Fig. 9-20.[4,17]

The structure of the flame continues to develop as it propagates across the chamber. Evidence, largely from schlieren photographs and studies of flame structure with laser diagnostics, shows that early in the burning process the flame

FIGURE 9-19
Shadowgraph photographs of spark-generated kernel between the spark plug electrodes. First photograph on left, 20 to 50 μs after breakdown; 40 μs between photos. Stoichiometric mixture, 1100 rev/min. (Courtesy A. Douad, Institut Français du Petrole.)
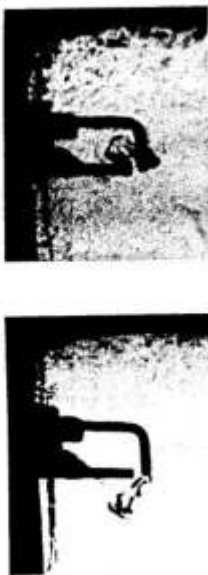
FIGURE 9-20
Schlieren photographs of developing flame, 5° after spark discharge, showing different convection processes in two different cycles. Spark plug wires 0.8 mm diameter. 1400 rev/min; 0.5 atm inlet pressure; propane fuel; $\phi = 0.9$; spark advance 45° BTC.[17]

is a thin, moderately wrinkled but simply connected, front or reaction sheet between unburned and burned gas. The thickness of the front is about 0.1 mm which is comparable to the thickness of a laminar flame under the prevailing conditions. The scale of the wrinkles is typically about 2 mm at engine speeds of 1000 to 2000 rev/min. As the flame propagates across the chamber, the thickness of the reaction sheet front remains roughly constant, the flame front becomes more convoluted, and the scale of the wrinkles tends to decrease with time.[24]

Further evidence that the thin reaction sheet front becomes highly wrinkled and convoluted by the turbulent flow field into a thick turbulent flame "brush" is provided by the schlieren photographs in Fig. 9-21. These show a flame propagating across the combustion chamber of a square-cross-section single-cylinder engine with a special transparent piston crown containing a 13-mm wide channel to isolate a small section of the flame.[25] The flame on the left is for a propane-air mixture; that on the right is hydrogen-air. The energy density per unit volume of mixture and the flow field are comparable for each fuel. The effective thickness of this turbulent flame—the average distance between the region ahead of the flame, where only unburned mixture exists, and the region behind the flame, where only

FIGURE 9-21
Schlieren photographs of flame in square-cross-section cylinder engine, with narrow channel in piston crown (at bottom of pictures) which permits observation of 13-mm wide section of flame. (a) Propane fuel, spark timing 36° BTC, photograph at 14° ATC, flame thickness ≈ 4.6 mm. (b) Hydrogen fuel, spark timing 3° BTC, photograph at 10° ATC, flame thickness ≈ 1.5 mm. Stoichiometric mixture; 1380 rev/min, 0.5 atm inlet pressure.[25]
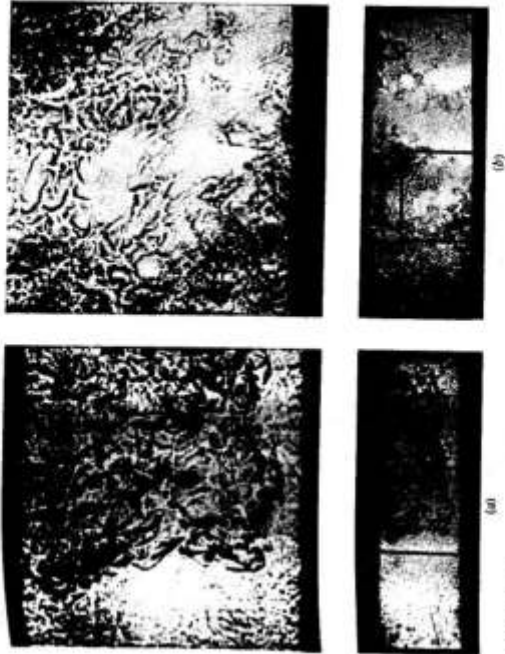
FIGURE 9-22
Enlarged schlieren photographs of (a) flame front and (b) flame back in square-cross-section cylinder engine with two glass side walls. 1400 rev/min, 0.5 atm inlet pressure, propane fuel, $\phi = 0.9$.[17]

burned mixture exists—is apparent.† It was 4 to 5 mm for propane and 1.5 mm for hydrogen. The difference is due to the substantially higher laminar flame speed for the hydrogen, air, burned gas mixture (see Sec. 9.3.4) which increases the Damköhler number and shifts the flame toward the weak-turbulence (i.e., less wrinkled) flame regime in Fig. 9-18.

Additional insight into the structure of the developed engine flame can be obtained by enlarging photographs of the leading and trailing edges of the flame, obtained with the schlieren or shadowgraph technique. Figure 9-22a shows the front of the flame 40° after the spark, when it has propagated about halfway across the chamber. It shows the irregular but smoothly curved surfaces which comprise the leading edge of the flame. Figure 9-22b shows the back of the flame 70° after the spark, when the front of the flame has just reached the wall of the combustion chamber farthest from the spark plug. It shows large clear regions of burned gas behind the flame and smaller clear regions connected by a lacelike

† These photographs were selected from a large number to give the minimum flame thickness corresponding to the flame front perpendicular to the channel length.
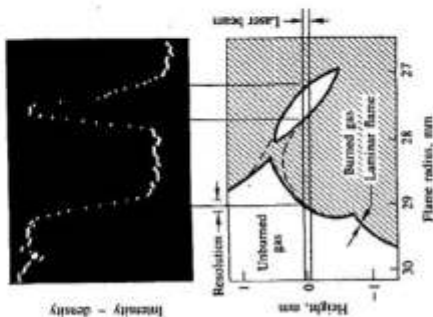
FIGURE 9-23
Upper picture: oscillogram of output of optical multichannel analyzer showing the intensity of light scattered from a narrow laser beam as a function of distance through the flame. Intensity is a measure of gas density. Lower picture: schematic of flame structure corresponding to this signal. 600 rev/min, 1 atm inlet pressure, propane fuel, $\phi = 1.0$.[24]

FIGURE 9-24
Shadowgraph photographs, with light beam normal to flame front, of 10-mm diameter section of flame. Photographs, arranged in order of increasing turbulent Reynolds number, suggest this is an appropriate scaling parameter (see Table 9.1).[24]

structure within which the remaining regions of unburned mixture are being consumed.[11] The analogy with a crumpled sheet of paper is appropriate.

Laser scattering experiments, where "snapshots" of the density profile along a laser beam passed through the flame were obtained using Rayleigh scattered light from the gas molecules, provide explicit evidence of this structure.[26] Figure 9-23 shows an oscillogram of the output of the optical analyzer. Each dot represents the intensity of the scattered light which is a measure of the gas density. The signal on the left corresponds to unburned gas. The flame is propagating from right to left. The oscillogram shows a thin transition zone of width 0.25 mm between unburned and burned gas. The fraction of oscillograms showing such an "island" of unburned gas. The fraction of oscillograms showing such "islands" varied from 0 at 300 rev/min to 20 percent at 1800 rev/min. An interpretation of this signal consistent with the available photographic evidence is shown underneath.[24]

The above results suggest that increasing engine speed, which increases turbulence levels in the unburned charge, increasingly convolutes and probably multiply-connects the thin reaction sheet flame front. Enlarged schlieren photographs of a 9-mm diameter section of the developed engine flame, viewed normal to the flame surface, indicate that increasing turbulence intensity and decreasing turbulence scales result in increasingly finely wrinkled flame structures. Figure 9-24 shows a set of such photographs, arranged in order of increasing turbulent Reynolds number. Increases in turbulence intensity, which was measured, were achieved by increases in engine speed and by modifying the inlet valve. Relevant parameters for each photograph are given in Table 9.1.

The above theoretical discussion and experimental evidence indicates that
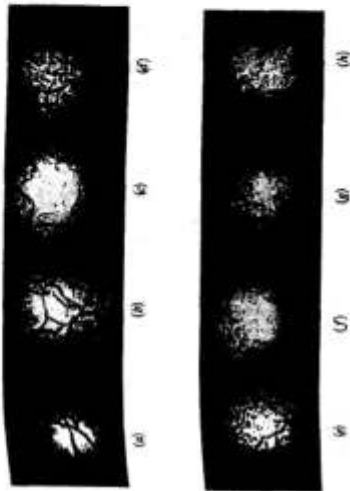
developed turbulent flames in spark-ignition engines, under normal operating conditions, are highly wrinkled and probably multiply-connected thin reaction sheets. The overall thickness of the turbulent flame "brush," front to back, is of order 1 cm. The thickness of the thin reaction sheet is comparable to estimates of the laminar flame thickness under the prevailing unburned mixture conditions which are of order 0.1 mm. The scale of the wrinkles is of order 1 mm.[24] Direct evidence to date is limited to the low to mid engine speed and low to high engine

TABLE 9.1
Parameters for shadowgraph photographs in Fig. 9-24

Photograph	Engine speed, rev/min	Turbulence intensity, m/s	Valve	Re_M	Re_T
a	300	0.44	S	106	229
b	600	0.88	S	157	503
c	300	1.07	US	173	611
d	900	1.33	S	193	760
e	1200	1.80	S	224	1024
f	600	1.95	US	234	1117
g	900	2.90	US	285	1658
h	1200	4.0	US	333	2263

Note: Valve: S, shrouded; US, unshrouded (produced higher turbulence due to less ordered flow). $Re_M = l_M V/\nu$; $Re_T = l_T V/\nu$ (see Sec. 8.2.1).[7]

load ranges. Whether the structure becomes significantly different at high engine speed is not known. Models of this turbulent flame development and propagation process are reviewed in Chap. 14.

9.3.3 Laminar Burning Speeds

An important intrinsic property of a combustible fuel, air, burned gas mixture is its laminar burning velocity. This burning velocity is defined as the velocity, relative to and normal to the flame front, with which unburned gas moves into the front and is transformed to products under laminar flow conditions. Some details of flame structure help explain the significance of this quantity. A flame is the result of a self-sustaining chemical reaction occurring within a region of space called the flame front where unburned mixture is heated and converted into products. The flame front consists of two regions: a preheat zone and a reaction zone. In the preheat zone, the temperature of the unburned mixture is raised mainly by heat conduction from the reaction zone: no significant reaction or energy release occurs and the temperature gradient is concave upward ($\partial^2 T / \partial x^2 > 0$). Upon reaching a critical temperature, exothermic chemical reaction begins. The release of chemical energy as heat results in a zone where the temperature gradient is concave downward ($\partial^2 T / \partial x^2 < 0$). The region between the temperature where exothermic chemical reaction begins and the hot boundary at the downstream equilibrium burned gas temperature is called the reaction zone. The thicknesses of the preheat and reaction zones can be calculated for one-dimensional flames from conservation equations of mass and energy. The thickness of the preheat zone $\delta_{L,\text{ph}}$ is

$$\delta_{L,\text{ph}} = \frac{4.6\bar{k}}{\bar{c}_p \rho_u S_L} \qquad (9.31)$$

where $\bar{k}$ and $\bar{c}_p$ are the mean thermal conductivity and specific heat at constant pressure in the preheat zone and S_L is the laminar burning velocity. Thus, the factors which govern the laminar burning velocity of a specific unburned mixture—the velocity at which this flame structure propagates relative to the unburned gas ahead of it—are the temperature and species concentration gradients within the flame and the mixture transport and thermodynamic properties.

Laminar burning velocities in engines are usually measured in spherical closed vessels by propagating a laminar flame radially outward from the vessel center. The laminar burning velocity is then given by

$$S_L = \frac{dm_b/dt}{A_f \rho_u} \qquad (9.32)$$

where the mass burning rate is determined from the rate of pressure rise in the vessel and A_f is the flame area. Because the laminar flame thickness [e.g., given by Eq. (9.31)] under engine conditions is of order 0.2 mm[24] and is therefore much less than characteristic vessel dimensions, in applying Eq. (9.32) the flame can be

treated as negligibly thin. Laminar burning velocities for methane, propane, isooctane, methanol, gasoline, and hydrogen—premixed with air—at pressures, temperatures, and equivalence ratios which occur in engines have been measured using this technique.[29-33] Also, the effect of a burned gas diluent on laminar burning velocity with gasoline-air mixtures has been determined.[32] Correlations derived from these data are the most accurate means available for estimating laminar burning velocities for mixtures and conditions relevant to spark-ignition engines.

The effect of the mixture fuel/air equivalence ratio on laminar burning velocity for several hydrocarbon fuels and methanol is shown in Fig. 9-25. The burning velocity peaks slightly rich of stoichiometric for all the fuels shown. The values for isooctane and gasoline are closely comparable. Data at higher pressures and temperatures have been fitted to a power law of the form:

$$S_L = S_{L,0}\left(\frac{T_u}{T_0}\right)^{\alpha}\left(\frac{p}{p_0}\right)^{\beta} \qquad (9.33)$$

where $T_0 = 298$ K and $p_0 = 1$ atm are the reference temperature and pressure, and $S_{L,0}$, α, and β are constants for a given fuel, equivalence ratio, and burned gas diluent fraction. For propane, isooctane, and methanol, these constants can be represented by

$$\alpha = 2.18 - 0.8(\phi - 1) \qquad (9.34a)$$

$$\beta = -0.16 + 0.22(\phi - 1) \qquad (9.34b)$$

and

$$S_{L,0} = B_m + B_\phi(\phi - \phi_m)^2 \qquad (9.35)$$

where ϕ_m is the equivalence ratio at which $S_{L,0}$ is a maximum with value B_m.
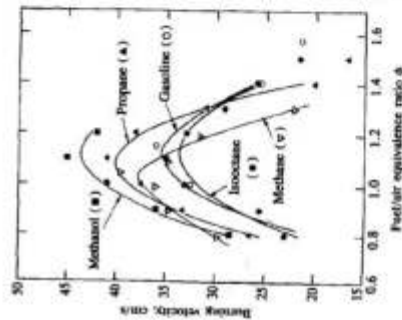
FIGURE 9-25
Laminar burning velocity for several fuels as a function of equivalence ratio, at 1 atm and 300 K. Lines are least-squares polynomial fits to data.[29,30]

TABLE 9.2
Parameters ϕ_m, B_m, and B_ϕ for Eq. (9.35)

Fuel	ϕ_m	B_m cm/s	B_ϕ cm/s	Ref.
Methanol	1.11	36.9	−140.5	30
Propane	1.08	34.2	−138.7	30
Isooctane	1.13	26.3	−84.7	30
Gasoline	1.21	30.5	−54.9	32

Note: Values of $S_{L,0}$ given by Eq. (9.35) are obtained from least-square fits of Eq. (9.35) to data over the range $p = 1$–8 atm, $T_u = 300$–700 K. They do not correspond exactly to the laminar flame speed data at 1 atm and 298 K in Fig. 9-25.

Values of ϕ_m, B_m, and B_ϕ are given in Table 9.2.[30] For gasoline (a reference gasoline with average molecular weight of 107 and an H/C ratio of 1.69) additional data were available and were correlated by:[32]

$$\alpha_\phi = 2.4 - 0.271\phi^{3.51} \qquad (9.36a)$$

$$\beta_\phi = -0.357 + 0.14\phi^{2.77} \qquad (9.36b)$$

For methane, simple equations such as (9.34a, b) do not adequately correlate the data over the range of p and T_u relevant to engines. However, laminar burning velocity data from a spherical constant-volume bomb experiment have been obtained along an unburned gas isentropic path, as the pressure in the bomb rises during combustion. Variation in laminar burning velocity along such unburned gas isentropes does correlate with a power law:

$$S_{L,i} = S_{L,0}\left(\frac{p_u}{p_{u0}}\right)_s^\varepsilon \qquad (9.37)$$

Values for $S_{L,0}$ and ε from the literature are summarized in Table 9.3.

TABLE 9.3
Parameters for methane–air laminar burning velocity correlation [Eq. (9.37)]

ϕ	p, atm	$S_{L,0}$,† cm/s	ε	Ref.
1.0	0.5	49	0.51	31
1.0	1.0	35	0.2	31
0.8–1.2	1–8	‡	0.17–0.19	33

† At 298 K initial temperature.
‡ See Fig. 9-25.
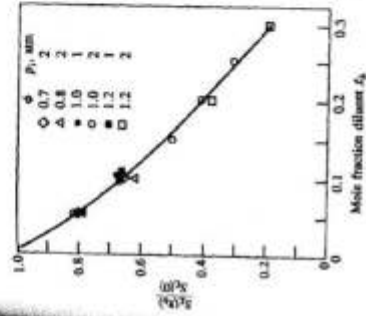
FIGURE 9-26
Effect of burned gas mole fraction $\tilde{x}_b$ in unburned mixture on laminar burning velocity. Fuel: gasoline.[31]

The presence of burned gas in the unburned cylinder charge due to residual gases and any recycled exhaust gases causes a substantial reduction in the laminar burning velocity. Any burned gas in the unburned mixture reduces the heating value per unit mass of mixture and, thus, reduces the adiabatic flame temperature. It acts as a diluent. The effect of increasing burned gas or diluent fraction on laminar flame speed is shown in Fig. 9-26. The diluent used was a mixture of CO_2 and N_2, chosen to match the heat capacity of actual gasoline-air combustion products.† The proportional reduction in laminar burning velocity is essentially independent of the unburned mixture equivalence ratio, pressure, and temperature over the range of interest in engines. The data in Fig. 9-26 are correlated by the relation:

$$S_L(\tilde{x}_b) = S_L(\tilde{x}_b = 0)(1 - 2.06\tilde{x}_b^{0.77}) \qquad (9.38)$$

where $\tilde{x}_b$ is the mole fraction of burned gas diluent. Other studies corroborate the magnitude of this burned gas effect.[31]

Note that for equal heat capacity added to the unburned mixture, burned gases have a much larger effect on laminar burning velocity than does excess air. For example, the laminar burning velocity of a stoichiometric mixture as it is leaned to $\phi = 0.8$ is reduced by 23 percent. The excess air required has a heat capacity of about 0.2 times that of the combustion products of the undiluted mixture. Adding the same heat capacity by adding stoichiometric burned gases (which requires a burned gas mole fraction of 0.175) reduces the laminar burning

_____

† The water in actual residual and exhaust gas was omitted. A mixture of 80 percent N_2 and 20 percent CO_2, by volume, was used.

velocity by 55 percent.[32] Proper allowance for the burned gas fraction in estimating laminar burning velocities for spark-ignition engines is most important.

The above correlations define the laminar burning velocity as a function of unburned mixture thermodynamic properties and composition, only. It has been assumed that flame thickness and curvature effects are negligible.[28] Our interest in laminar burning velocity is twofold: first, it is used to define the characteristic chemical reaction time of the mixture in Eq. (9.30); second, a presumed consequence of the wrinkled thin-reaction-sheet turbulent-flame structure is that, locally, the sheet propagates at the laminar burning velocity. The above correlations adequately characterize a quiescent burning process. However, laminar flame propagation can be influenced by the local flow field in the unburned gas. If the flame thickness is less than the Kolmogorov scale, the primary effect is one of straining which affects both the flame area (usually referred to as flame stretching for an area increase) and the local (laminar) burning velocity. While this problem is not yet well understood, it is known that straining can affect the laminar burning velocity and can cause flame extinction. The laminar burning velocity decreases with increasing strain rate, and the Lewis number of the unburned mixture has a significant influence on this rate of decrease. The Lewis number is the ratio of diffusivities of heat and mass. For stoichiometric mixtures it is close to one; it increases above about unity as the unburned fuel-air mixture is leaned out. Thus the local flow field may have a discernable effect on the local burning velocity of the thin laminarlike reaction-sheet flame, especially for lean or dilute mixtures.[34]

9.3.4 Flame Propagation Relations

If the heat-release or mass burning rate analysis of Sec. 9.2.2 is coupled with an analysis of flame geometry data, substantial additional insight into the behavior of spark-ignition engine flames is obtained. Flame photographs (such as those in Figs. 9-1, 9-14, and 9-16 and Refs. 4 and 35) effectively define the position of the front or leading edge of the turbulent engine flame. The "shadow" of the enflamed zone, under normal engine conditions, is close to circle: only in the presence of very high swirl does substantial distortion of the flame shape occur.[18] Thus, to a good approximation, the surface which defines the leading edge of the turbulent flame (ahead of which only unburned mixture exists) is a portion of the surface of a sphere. Figure 9-27 indicates the geometrical parameters which define this flame surface: r_c, α_c, z_c, the coordinates of the flame center; r_f, the radius of the best-fit circle to the flame front silhouette; and the geometry of the combustion chamber walls. The flame is initiated at the spark plug; however, it may move away from the plug during the early stages of its development as shown. We define the *flame front area* A_f as the spherical surface of radius r_f coinciding with the leading edge of the flame contained within the combustion chamber, and the *enflamed volume* V_f as the volume within the chamber behind this flame front.

The thermodynamic analysis of cylinder pressure data allows us to define additional geometrical parameters. The *burned gas radius* r_b is the radius of the
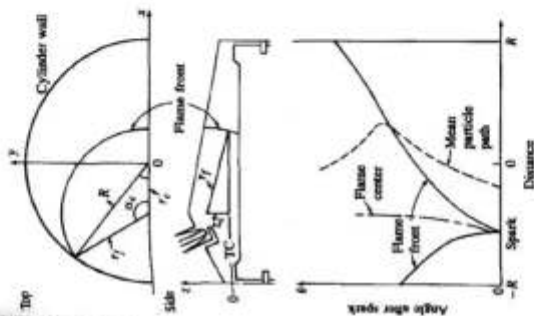
FIGURE 9-27
Schematic of spherical flame front in engine combustion chamber identifying parameters which define flame geometry. (*From Beretta et al.*[4])

spherical surface within the combustion chamber which would contain all the burned gas behind it; i.e.,

$$V_b(r_b, r_c, \alpha_c, z_c) = V_b(\rho, \theta) \qquad (9.39)$$

The *spherical burning area* A_b is the area of this spherical surface; i.e.,

$$A_b = \frac{\partial V_b(r_b, r_c, \alpha_c, z_c)}{\partial r_b} \qquad (9.40)$$

The *laminar burning area* A_L is the surface area the flame would have if it burned at the laminar flame speed; i.e.,

$$A_L = \frac{dm_b/dt}{\rho_u S_L} \qquad (9.41)$$

where S_L is the laminar flame speed in the unburned mixture ahead of the flame (see Sec. 9.3.3).

Several velocities can be defined. The *mean expansion speed* of the front u_f is given by

$$u_f = \frac{dA_b/dt}{L_b} \qquad (9.42)$$

where A_S is the "shadow" area enclosed by the "best-fit" circle through the leading edge of the flame and

$$L_S = \frac{\partial A_S}{\partial r_f} \tag{9.43}$$

is the arc length within the chamber of this "best-fit" circle. The mean expansion speed of the burned gas u_b is

$$u_b = \frac{\partial V_b/\partial t}{A_b} \tag{9.44}$$

This derivative is taken with the piston position fixed since only burned volume changes due to combustion are of interest. The burning speed S_b is defined by

$$S_b = \frac{dm_b/dt}{\rho_u A_b} \tag{9.45}$$

The mean gas speed just ahead of the flame front u_g is

$$u_g = u_b - S_b \tag{9.46}$$

Note that combining Eqs. (9.41) and (9.44) gives the relation

$$S_b A_b = S_L A_L$$

Also, it follows from Eqs. (9.29), (9.43), and (9.44) that

$$\frac{u_b}{S_b} = \frac{\rho_u}{\rho_b}(1 - y_b) + y_b = \frac{\rho_u/\rho_b}{[(\rho_u/\rho_b) - 1]x_b + 1} \tag{9.47}$$

As x_b and $y_b \to 0$, u_b/S_b approaches the expansion ratio ρ_u/ρ_b. As x_b and $y_b \to 1$, u_b/S_b approaches unity.

The variation of the above quantities during the engine combustion process, coupled with the photographs and discussion in Sec. 9.3.2, provide substantial insight into the flame development and propagation process. Figure 9-28 shows results from an analysis of cylinder pressure data and the corresponding flame front location information (determined from high-speed movies through a window in the piston) of several individual engine operating cycles. The combustion chamber was a typical wedge design with a bore of 102 mm and a compression ratio of 7.86. The flame radius initially grows at a rate that increases with time and exhibits substantial cycle-by-cycle variation in its early development (Fig. 9-28a). Later ($r_f \gtrsim 30$ mm) the growth rate, which approximates the expansion speed u_b, reaches an essentially constant value. The flame radius r_f is initially equal to the burned gas radius r_b; it increases above r_b as the flame grows and becomes increasingly distorted by the turbulent flow field (Fig. 9-28b). Eventually $r_f - r_b$ goes to an essentially constant value of about 6 mm for $r_b \gtrsim 30$ mm. This difference, $r_f - r_b$, is approximately half the thickness of the turbulent flame brush.
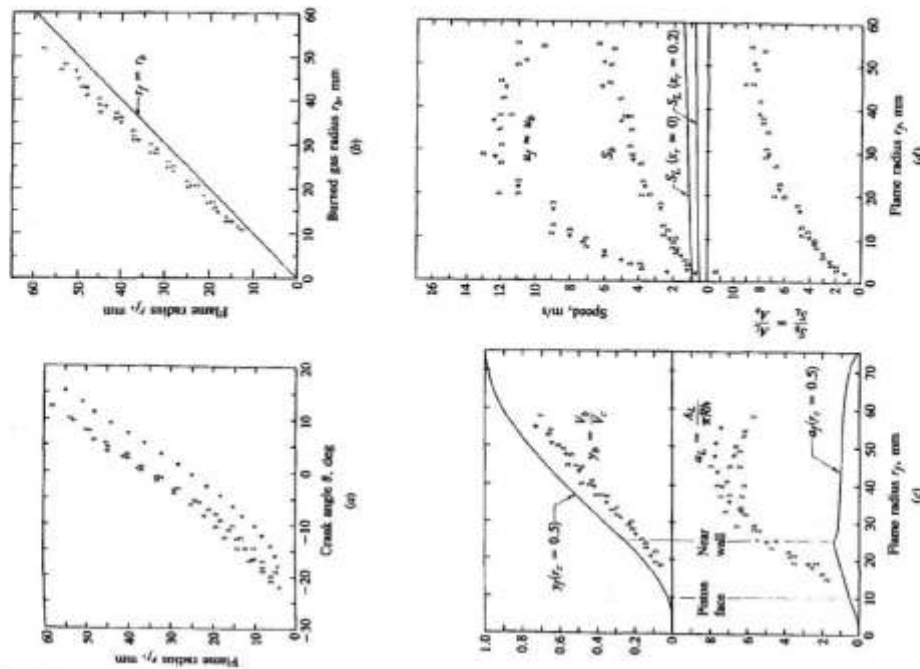
FIGURE 9-28
Variation of flame geometry and velocity parameters during four individual combustion cycles at 1044 rev/min, $\phi = 0.98$, 1 atm inlet pressure: (a) flame radius r_f versus crank angle; (b) flame radius r_f versus burned gas radius r_b; (c) normalized enflamed volume y_f, burned volume y_b, normalized flame front area a_f, and laminar area a_L versus flame radius; (d) front expansion speed u_b, burning speed S_b, and laminar flame speed S_L versus flame radius. (From Beretta et al.[a])

Normalized enflamed and burned volumes, and flame front area and laminar burning area, are shown in Fig. 9-28c. Volumes are normalized by the cylinder volume, and areas by πRh, where h is the average clearance height and R the cylinder radius. Discontinuities occur in the flame area a_f at the points where the flame front contacts first the piston face and then the near cylinder wall. The laminar area A_L is initially close to the flame area A_f and then increases rapidly as the flame grows beyond 10 mm in radius. During the rapid burning combustion phase ($y_f \gtrsim 0.2$) the value of y_f is significantly greater than y_b. During this phase, the laminar area exceeds the flame area by almost an order of magnitude. These observations indicate the existence of substantial pockets of unburned mixture behind the leading edge of the flame.[4]

The ratio of the volume of the unburned mixture within the turbulent flame zone ($V_f - V_b$) to the reaction-sheet area within the flame zone ($A_L - A_f$) defines a characteristic length

$$I_T = \frac{V_f - V_b}{A_L - A_f} \tag{9.48}$$

which can be thought of as the scale of the pockets of unburned mixture within the flame. For the data set of Fig. 9-28, I_T is approximately constant and of order 1 mm.[24]

These flame geometry results would be expected from the previous photographic observations of how the flame grows from a small approximately spherical smooth-surfaced kernal shortly after ignition to a highly wrinkled reaction-sheet turbulent flame of substantial overall thickness. Initially, the amount of unburned gas within the enflamed volume is small. During the rapid burning phase of the combustion process, however, a significant fraction (some 25 percent; see Fig. 9-2) of the gas entrained into the flame zone is unburned.

The front expansion speed u_f, burning speed S_b, and laminar flame speed S_L are shown in Fig. 9-28d. The expansion speed increases as the flame develops to a maximum value that is several times the mean piston speed of 3.1 m/s and is comparable to the mean flow velocity through the inlet valve of 18 m/s.[24] The burning speed increases steadily from a value close to the laminar flame speed at early times to almost an order of magnitude greater than S_L during the rapid burning phase. During this rapid burning phase, since ($r_f - r_b$) is approximately constant, the flame front expansion speed and the mean burned gas expansion speed are essentially equal. The difference between $u_b \approx u_f$ and S_b is the unburned gas speed u_e just ahead of the flame front. Note that the ratio u_f/S_b, ($\approx u_b/S_b$) at spark to unity as the flame approaches the far wall, as required by Eq. (9.47).

The effect of flame propagation on the flow field in the unburned mixture ahead of the flame is important because it is the turbulence just ahead of the flame that determines the local burning velocity. Measurements of mean velocities, rms fluctuation velocities, and turbulence intensities have been made using laser doppler anemometry (see Sec. 8.2.2) at different locations within engine combustion chambers (e.g., Refs. 36 and 37). Such data are difficult to interpret
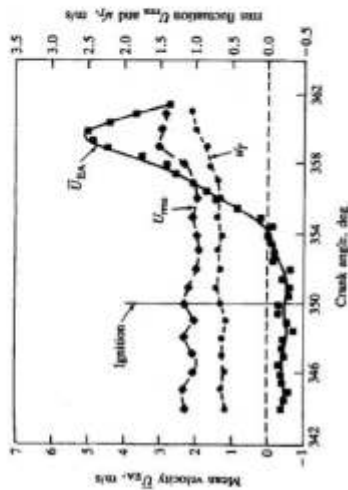
FIGURE 9-29
Laser doppler anemometer measurements of ensemble-averaged mean velocity U_{EA} [Eq. (8.20)], rms fluctuation in individual-cycle mean velocity U_{rms} [Eq. (8.21)] and turbulence intensity u_T' [Eq. (8.22)], close to the spark plug and close to the cylinder wall. Disc-shaped chamber, spark plug in cylinder wall. Measurement at $x/B = 0.57$, $B = 76$ mm, 300 rev/min, $S_p = 0.83$ m/s.[24]

because the mean flow varies cycle-by-cycle, the turbulence is not homogeneous, and the flame motion and shape show substantial cyclic variations. The results in Fig. 9-29 were taken in a special single-cylinder engine with a disc-shaped combustion chamber where the spark plug was located in the cylinder liner. Shown are the ensemble-averaged mean velocity, cyclic variation in mean velocity, and the turbulence intensity, normal to the front, during the major portion of the combustion process, close to the chamber center.

The mean velocity normal to the front increases steadily from shortly after ignition, as the combustion-produced gas expansion displaces unburned mixture toward the wall. It peaks as the flame arrives at the measurement location. The cyclic variation in mean velocity and the turbulence intensity normal to the front remain essentially constant until a few degrees before the flame arrival. These two quantities are comparable in magnitude; thus the turbulence intensity is lower than the rms fluctuation velocity (in this case by about a factor of 2) (see Sec. 8.2.2). Whether the increase in turbulence as the flame approaches is due to rapid distortion resulting from the compression of the unburned mixture which occurs during combustion or is the result of inadequate resolution of cycle-to-cycle flow variations is unclear. Rapidly imposed distortions of a turbulent flow field, such as those imposed by combustion-produced gas expansion, would lead to an increase in vorticity and turbulence intensity. Other studies, e.g., Ref. 37, indicate there is little or no increase in turbulence intensity ahead of the flame.

The variation of burning speed with engine speed has also been carefully examined in a study where flame position was determined from high-speed movies, mass burning rates from cylinder pressure, and turbulence information

from experiments at equivalent motored conditions.[35] The turbulence quantity obtained during motoring experiments was the ensemble-averaged root mean square velocity fluctuation defined by Eq. (8.18). Values of S_b/S_L and u_f'/S_L were determined at two points in the combustion process: at a flame radius of 30 mm (the end of the flame development process) and at mass fraction burned equal to 0.5 (halfway through the rapid burning phase). To correct the motored turbulence data for the higher pressure levels corresponding to engine firing conditions, a simple rapid distortion model (see Sec. 14.4.2) based on conservation of angular momentum in turbulent eddies was used. A linear correlation between S_b and u_f' results, as shown in Fig. 9-30, for the rapid burning combustion phase. Note that as u_f'/S_L goes to zero, S_b/S_L approaches a value close to unity.

Once the flame front reaches the far cylinder wall (see Fig. 9-27) the front can no longer propagate; however, combustion continues behind the front until all the unburned mixture entrained into the enflamed region is consumed. This final burning or termination phase of the combustion process can be approximated by an exponential decay in the mass burning rate with a characteristic time constant τ_b of order 1 ms. Since these "islands" or "pockets" of unburned mixture behind the leading edge of the flame have a characteristic scale l_T based on the laminar flame area [Eqs. (9.41) and (9.48)], it follows that

$$l_T = \tau_b S_L \tag{9.49}$$

In summary, the above flame data analysis procedures show that the relationships between r_f and r_b, V_f and V_b, A_L, u_f', S_b and S_L are distinctly different in the three phases of combustion: (1) the development phase, where a highly wrinkled reaction-sheet "thick"-overall turbulent flame evolves from the essentially spherical flame kernal established by the spark discharge; (2) the rapid-burning phase, where this thick "developed" turbulent flame propagates across the combustion chamber to the far wall, during which most of the mass is burned; and (3) the termination phase after the flame front has reached the far wall and propagation of the front is no longer possible, when the remaining unburned mixture within the flame burns up. The burning velocity, in the rapid-burning phase of the combustion process, scales with turbulence intensity, which in turn scales with engine speed.
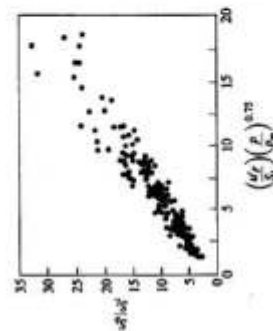
FIGURE 9-30
Variation of burning speed with turbulence intensity. The ensemble-averaged rms velocity fluctuation was measured during motoring engine operation. The ratio p/p_m (firing pressure/ motoring pressure) corrects for the effect of additional compression on the turbulence intensity. Range of engine speeds and spark timings.[35]

9.4 CYCLIC VARIATIONS IN COMBUSTION, PARTIAL BURNING, AND MISFIRE

9.4.1 Observations and Definitions

Observation of cylinder pressure versus time measurements from a spark-ignition engine, for successive operating cycles, shows that substantial variations on a cycle-by-cycle basis exist. Since the pressure development is uniquely related to the combustion process, substantial variations in the combustion process on a cycle-by-cycle basis are occurring. In addition to these variations in each individual cylinder, there can be significant differences in the combustion process and pressure development between the cylinders in a multicylinder engine. Cyclic variations in the combustion process are caused by variations in mixture motion within the cylinder at the time of spark cycle-by-cycle, variations in the amounts of air and fuel fed to the cylinder each cycle, and variations in the mixing of fresh mixture and residual gases within the cylinder each cycle, especially in the vicinity of the spark plug. Variations between cylinders are caused by differences in these same phenomena, cylinder-to-cylinder.

Cycle-by-cycle variations in the combustion process are important for two reasons. First, since the optimum spark timing is set for the "average" cycle, faster-than-average cycles have effectively overadvanced spark-timing and slower-than-average cycles have effectively overretarded spark-timing and result. Second, it is the extremes of the cyclic variations that limit engine operation. The fastest burning cycles with their overadvanced spark timing are most likely to knock. Thus, the fastest burning cycles determine the engine's fuel octane requirement and limit its compression ratio (see Sec. 9.6.3). The slowest burning cycles, which are retarded relative to optimum timing, are most likely to burn incompletely. Thus these cycles set the practical lean operating limit of the engine or limit the amount of exhaust gas recycle (used for NO emissions control) which the engine will tolerate. Due to cycle-by-cycle variations, the spark timing and average air/fuel ratio must always be compromises, which are not necessarily the optimum for the average cylinder combustion process. Variations in cylinder pressure have been shown to correlate with variations in brake torque which directly relate to vehicle driveability.

An example of the cycle-by-cycle variations in cylinder pressure and the variations in mixture burning rate that cause them are shown in Fig. 9-31. Pressure and gross heat-release rate [calculated from the cylinder pressure using Eq. (9.27)] for several successive cycles at a mid-load, mid-speed point are shown as a function of crank angle. The maximum heat-release rate and the duration of the heat release or burning process vary by a factor of two from the slowest to the fastest burning cycle shown. The peak cylinder pressure varies accordingly. The faster burning cycles have substantially higher values of maximum pressure than do the slower burning cycles; with the faster burning cycles peak pressure occurs closer to TC.

The heat-release rate data in Fig. 9-31 show that there are cycle-by-cycle variations in the early stages of flame development (from zero to a few percent of

the point of efficiency, hydrocarbon emissions, torque variations, and roughness. The partial-burn and misfire regimes are discussed in Sec. 9.4.3.

Various measures of cycle-by-cycle combustion variability are used. It can be defined in terms of variations in the cylinder pressure between different cycles, or in terms of variations in the details of the burning process which cause the differences in pressure. The following quantities have been used:

1. *Pressure-related parameters.* The maximum cylinder pressure p_{max}; the crank angle at which this maximum pressure occurs θ_{pmax}; the maximum rate of pressure rise $(dp/d\theta)_{max}$; the crank angle at which $(dp/d\theta)_{max}$ occurs; the indicated mean effective pressure [which equals $\int p\, dV/V_d$, see Eqs. (2.14), (2.15), and (2.19)].

2. *Burn-rate-related parameters.* The maximum heat-release rate (net or gross, see Sec. 9.2.2); the maximum mass burning rate; the flame development angle $\Delta\theta_d$ and the rapid burning angle, $\Delta\theta_b$ (see Sec. 9.2.3).

3. *Flame front position parameters.* Flame radius, flame front area, enflamed or burned volume, all at given times; flame arrival time at given locations.

Pressure-related quantities are easiest to determine; however, the relation between variations in combustion rate and variations in cylinder pressure is complex.[38] Equation (9.26) defines the factors that govern this relationship. Because the rate of change of pressure is substantially affected by the rate of change of cylinder volume as well as rate of burning, changes in the phasing of the combustion process relative to TC (e.g., which result from changes in flame development angle) as well as changes in the shape and magnitude of the heat-release rate profile affect the pressure. Figure 9-32 illustrates how the magnitude of the maximum cylinder pressure p_{max} and the crank angle at which it occurs θ_{pmax} vary as the crank angle at which combustion effectively starts (e.g., θ at which 1 percent of the cylinder mass has burned) and the burning rate are varied. Curve *CABE* shows how p_{max} and θ_{pmax} vary for a fixed fast-burning heat-release profile (the duration of the heat-release process and its maximum value are held
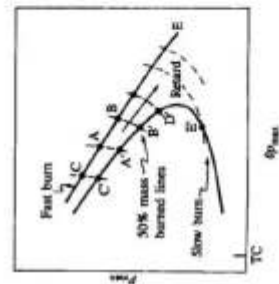
FIGURE 9-32
Schematic of variation in maximum cylinder pressure and crank angle at which it occurs, in individual cycles. *CABE* typical of fast heat-release process; *CABDE* typical of slow heat-release process. (*From Matekunas.*[34])

FIGURE 9-31
Measured cylinder pressure and calculated gross heat-release rate for ten cycles in single-cylinder spark-ignition engine operating at 1500 rev/min, $\phi = 1.0$, $p_{inlet} = 0.7$ atm, MBT timing 25° BTC.[14]

the total heat release) and in the major portion of the combustion process—the rapid-burning phase—indicated by the variations in the maximum burning rate.

As the mixture becomes leaner with excess air or more dilute with a higher burned gas fraction from residual gases or exhaust gas recycle, the magnitude of cycle-by-cycle combustion variations increases. Eventually, some cycles become sufficiently slow burning that combustion is not completed by the time the exhaust opens: a regime where *partial burning* occurs in a fraction of the cycles is encountered. For even leaner or more dilute mixtures, the *misfire* limit is reached. At this point, the mixture in a fraction of the cycles fails to ignite. While spark-ignition engines will continue to operate with a small percentage of the cycles in the partial-burn or misfire regimes, such operation is obviously undesirable from

constant), as the phasing of this combustion process relative to TC is varied, A corresponds to MBT timing where the start of combustion is phased to give maximum brake torque, B corresponds to retarded timing, and C to over-advanced timing. $C'A'B'D'E'$ is a similar curve for a slow-burning heat-release profile. A' corresponds to MBT timing, and B' and D' to increasingly retarded timing. Note that with a sufficiently slow-burning heat-release profile, beyond D', θ_{pmax} decreases as the burn process is increasingly retarded. This occurs when the rate of increase of pressure due to combustion becomes so low that it is more than offset by the pressure decrease due to volume increase: eventually for extremely slow and late burning, the maximum pressure approaches the motored pressure at TC. The dashed lines show the constant start-of-combustion timing relative to MBT for each burn rate curve. Note that θ_{pmax} for constant relative timing varies little as the heat-release profile or burn rate varies.[38]

We can now explain the effects of variations in the heat-release profile (both in the development stage of the burning process, which effectively changes the location of the start of combustion, and in the rate of burning throughout the process) on p_{max} and θ_{pmax} when the spark timing occurs at a fixed crank angle. For a fixed burning rate profile (duration of burn and maximum burning rate) as the start of combustion is delayed to be closer to TC, p_{max} decreases and θ_{pmax} initially increases (A to B or A' to B'). This is the effect of a change in relative timing or phase of the burning process due to a slower initial rate of flame development with fixed spark timing. If, in addition to the flame development being slower, the heat-release rate throughout the burning process is lower, then that combustion process is even more retarded from the optimum and p_{max} decreases and θ_{pmax} increases further, to their values at D'. The effect of a faster initial flame development and faster burning rate, with fixed spark time, is the opposite. The magnitudes of the changes in p_{max} and θ_{pmax} depend, obviously, on the extent of the cyclic variations; they also depend on whether the average burn process is fast or slow. For fast-burning engines, a larger fraction of the heat release occurs near TC when the chamber volume is changing relatively slowly. Thus pressure variations are mainly due to combustion variations. With slow-burning engines, where a significant fraction of the energy release occurs well after TC, the effect of volume change also becomes significant and augments the effect of combustion variations. For large variations and a slow average burning process, p_{max} can fall below E', and θ_{pmax} then decreases. A fast-burning combustion process significantly reduces the impact of cyclic combustion variations on engine performance.[39]

We can now evaluate the various measures of combustion variability. The maximum pressure variation has been shown to depend on both changes in phasing and burning rate. The magnitude of this variation depends on whether the combustion chamber is faster or slower burning, on average. It also depends on whether the burning process is substantially retarded relative to MBT. It depends, too, on cyclic cylinder fuel and air charging variations. Thus the interpretation of variations in p_{max} [or in the maximum rate of pressure rise $(dp/d\theta)_{max}$] in terms of variations in the rate and phasing of the burning process

must be done with care. The location of maximum pressure θ_{pmax} also depends on relative phasing of combustion and on the burn rate profile. In addition, for slow-burning chambers and retarded timing (around E') variations produce little change in θ_{pmax}. However, for fast-burning chambers, with MBT or only slightly retarded timing, the location of peak pressure depends essentially on the phasing of each combustion process relative to its MBT phasing, and is independent of charging variations. For these reasons, θ_{pmax} is a useful measure of variability in combustion event phasing.[38]

One important measure of cyclic variability, derived from pressure data, is the *coefficient of variation in indicated mean effective pressure*. It is the standard deviation in imep divided by the mean imep, and is usually expressed in percent:

$$COV_{imep} = \frac{\sigma_{imep}}{imep} \times 100 \qquad (9.50)$$

It defines the cyclic variability in indicated work per cycle, and it has been found that vehicle driveability problems usually result when COV_{imep} exceeds about 10 percent.

Figure 9-33 illustrates the relationships between p_{max}, θ_{pmax}, and imep for 120 cycles of an engine cylinder at fixed operating conditions and three different spark timings.[38] The MBT timing data show a spread in imep at a fixed value of θ_{pmax}. This imep data band is relatively flat and is centered around $\theta_{pmax} \approx 16°$; only at later values of θ_{pmax} does imep fall off significantly. The vertical spread in imep around $\theta_{pmax} = 16°$ is due to variations in the amount of fuel entering the cylinder each cycle; normal variations in the burn profile under these conditions, which effectively change the phasing of the combustion process, produce only modest reductions in imep. For early θ_{pmax} (the extreme upper left of Fig. 9-33a), the variations in p_{max} are also due mainly to these fuel-charging variations, cycle-by-cycle; these are the fastest burning cycles with the most advanced phasing. As
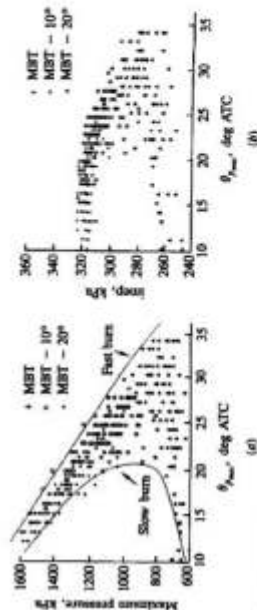
FIGURE 9-33
(a) Individual-cycle maximum pressure versus crank angle at which p_{max} occurs. (b) Individual-cycle indicated mean effective pressure versus θ_{pmax}.[38]

θ_{max} increases, the dispersion increases as cyclic variations in phasing and burning rate have increasing impact.

An important issue is whether variations in the early stages of flame development and variations in subsequent portions of the burning process are independent of each other or are correlated. Plots of early flame development angle (spark to 1 percent mass burned) against the burning angle (1 to 90 percent burned) from individual cycles for several different combustion chambers indicate the following. There is a trend with increasing flame development angle for the burning angle to increase (or the burning rate to decrease); however, there is much scatter about this trend (for a given value of flame development angle different cycles show a substantial range in burning angle), and the quantitative aspects of the trend depend on operating conditions and on combustion chamber design. In addition, as the mean rapid burning angle increases (due to changing operating conditions or a slower-burning chamber design) the mean flame development angle, the cyclic variation in the flame development angle, and the cyclic variation in the rapid burning angle all increase.[40] This topic is discussed more fully in the following section.

The shapes of the frequency distributions in individual-cycle pressure data (e.g. in p_{max}, $(dp/d\theta)_{max}$, θ_{max}) and in burn rate data such as $\Delta\theta_i$, $\Delta\theta_b$, $(dQ/d\theta)_{max}$ depend on whether the combustion process is fast and "robust" (e.g., with close-to-stoichiometric mixtures at higher loads at optimum timing—well away from the lean operating limit of the engine) or slower and less repeatable, closer to the lean or dilute-mixture operating limit. Under robust combustion conditions these distributions are close to normal distributions.[41-43] When the combustion process is much slower, the cyclic variability becomes large and the distribution becomes skewed toward the slower burning cycles which have low imep (due to the substantial retard of these slower cycles). When partial burning and then misfire occur, the low-pressure tail of the distribution approaches the motored pressure value at TC.[44] Examples of the frequency distributions of imep in these two combustion variability regimes are shown later in Fig. 9-36.

Cylinder pressure data are often averaged over many cycles to obtain the mean cylinder pressure at each crank angle. The primary use of this average pressure versus crank angle data is in calculating the average indicated mean effective pressure (which is a linear function of p). Since combustion parameters are not linearly related to the cylinder pressure [see Eq. (9-27)], analysis of the average pressure data will not necessarily yield accurate values of average combustion parameters. The error will be most significant when the combustion variability is largest. It is best to determine mean combustion parameters by averaging their values obtained from a substantial number of individual cycle analysis results. The number of cycles which must be averaged to obtain the desired accuracy depends on the extent of the combustion variability. For example, while 40 to 100 cycles may define imep to within a few percent when combustion is highly repeatable, several hundred cycles of data may be required when cyclic combustion variations are large.[13]

9.4.2 Causes of Cycle-by-Cycle and Cylinder-to-Cylinder Variations

Cycle-by-cycle combustion variations are evident from the beginning of the combustion process. Analysis of flame photographs from many engine cycles taken in special research engines with windows in the combustion chamber has shown that dispersion in the fraction of the combustion chamber volume inflamed is present from the start of combustion (e.g., see Refs. 3 and 23). Dispersion in burning rate is also evident throughout the combustion process (see Figs. 9-2 and 9-31). Three factors have been found to influence this dispersion.[45]

1. The variation in gas motion in the cylinder during combustion, cycle-by-cycle

2. The variation in the amounts of fuel, air, and recycled exhaust gas supplied to a given cylinder each cycle

3. Variations in mixture composition within the cylinder each cycle—especially near the spark plug—due to variations in mixing between air, fuel, recycled exhaust gas, and residual gas

The relative importance of these factors is not yet fully defined, and depends on engine design and operating variables. The variation in the velocity field within the engine cylinder throughout the cycle, and from one cycle to the next, has been reviewed in Sec. 8.2.2. Toward the end of the compression stroke, the ensemble-averaged rms velocity fluctuation is of comparable magnitude to the mean piston speed, and may be larger than the mean flow velocity if there is no strongly directed local mean flow pattern (see Figs. 8-8 and 8-9). This ensemble-averaged velocity fluctuation combines both cycle-by-cycle variation in the mean flow and the turbulent velocity fluctuations. During compression, these two components are of comparable magnitude (see Figs. 8-9 and 9.29). While this data base is limited, it indicates that substantial variations in the mean flow exist, cycle-by-cycle, both in the vicinity of the spark plug and throughout the combustion chamber. Velocity variations contribute in a major way to variations in the initial motion of the flame center as it grows from the kernel established by the spark, and in the initial rate of growth of the flame; they can also affect the burning rate once the flame has developed to fill a substantial fraction of the combustion chamber. Variations in gas motion near the spark plug convect the flame in its early stages in different directions and at different velocities, cycle-by-cycle. This affects the flame's interaction with the cylinder walls, changing the flame area development with time. Variations in the turbulent velocity fluctuations near the spark plug will result in variations in the rate at which the small initially laminarlike flame kernel develops into a turbulent flame. Variations in the mean flow throughout the chamber will produce differences in flame front shape; also, they may produce differences in turbulence which affect the propagation velocity of the front (see Fig. 9-30 for the relation between mean burning speed and turbulence intensity).

It is well known that, on a time-averaged basis, the fuel, air, and recycled exhaust gas flows into each cylinder of a multicylinder engine are not identical. These flow rate differences are typically a few percent (see Secs. 7.6.2 and 7.6.3). It is also known that the flow patterns within the different cylinders are not necessarily identical due to differences between the individual intake manifold runner and port geometries in many production engines. All these factors contribute to cylinder-to-cylinder variations in the combustion process: there can be significant differences in the mean burn rate parameters as well as in the cyclic variations in these parameters.[41] Also, the limited data available on the variation in mixture composition within each cylinder for each cycle indicates that cyclic charging variations in individual cylinders are comparable in magnitude to cylinder-to-cylinder differences (i.e., of order ±5 percent[44]). Whether the amount of residual gas left in the cylinder varies significantly, cycle-by-cycle, is not known. At higher loads, where the combustion process is more repeatable (and always completed relatively early in the expansion stroke) and the residual gas fraction smaller (see Sec. 6.4), variations in the total amount of residual are not expected to be significant. At light loads (particularly at idle), where combustion variability is much higher and partial-burning cycles may occur, and especially with high valve overlap engine designs, variations in the residual gas mass and its composition may become important.

In addition, mixing of fuel, air, recycled exhaust, and residual is not complete: nonuniformities in composition exist within the cylinder at the start of combustion. Composition variations, cycle-by-cycle, in the vicinity of the spark plug electrode gap will affect the early stages of flame development, especially as the flame grows through the laminarlike burning phase following the creation of a small flame kernel by the spark discharge (see Sec. 9.5.1). Figure 9-34 indicates the extent of these composition nonuniformities. The available data comes from experiments where a small rapid-acting sampling valve located in the spark plug center electrode was used to extract gas from the vicinity of the electrode gap, close to the start of combustion, for individual cycle composition analysis. Figure 9-34a shows the cycle-by-cycle air/fuel ratio fluctuations in the burned gases sampled from one cylinder of a four-cylinder gasoline-fueled carbureted engine just after combustion has started. The standard deviation was typically 2 to 6 percent of the mean (A/F).[46-48] Figure 9-34b shows the relationship between total hydrocarbon and CO_2 concentrations in unburned mixture, sampled just before spark discharge. The CO_2 concentration is a measure of the burned gas fraction in the sampled unburned mixture; hence on average it correlates inversely with the total hydrocarbon concentration. However, there is substantial fluctuation in CO_2 concentration about the mean value, at a given fuel fraction, indicating significant fluctuations, cycle-by-cycle, in the mixing of fresh mixture with residual gas. Nonuniformities in EGR distribution between cylinders and EGR mixing within the cylinder would also increase the variations in burned gas fraction locally at the spark gap, cycle-by-cycle.[48]

Experiments in a multicylinder production SI engine, where the fuel/air ratio nonuniformities and the nonuniform mixing of fresh mixture with residual
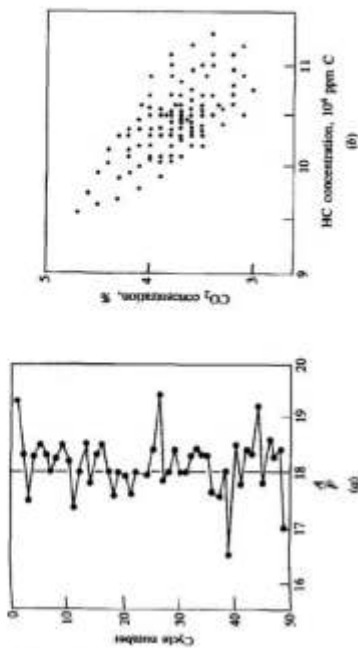
FIGURE 9-34
(a) Air/fuel ratio in 50 consecutive cycles, in vicinity of spark plug, measured just after ignition with a rapid-acting sampling valve located in the plug center electrode. Engine operated at 1400 rev/min, MBT timing, imep = 314 kPa.[47] (b) CO_2 and unburned HC concentrations in gas sampled in individual cycles from the vicinity of the spark plug just prior to ignition. Engine operated at 1200 rev/min, $\phi = 0.98$, $P_{atm} = 0.5$ atm, gasoline fuel.[48]

gas were removed in turn as contributors to cycle-by-cycle variations (by comparing premixed propane operation with conventional carbureted operation with gasoline, and by removing residual gas by purging with nonfiring cycles), showed that the three contributing factors to cyclic combustion variations—velocity variations, fuel/air ratio variations, and residual gas mixing variations—are of comparable importance at road-load conditions.[45]

An explanation for cycle-by-cycle variations can be developed from the description of the turbulent flame propagation process in Sec. 9.3. Conditions in the vicinity of the spark plug will influence the initial stages of the flame propagation process—establishing a stable kernel and its development into a turbulent flame. During the developed flame propagation phase, the average conditions in the bulk gas within the combustion chamber will be the determining factors since the flame front spans the chamber, effectively averaging out local nonuniformities. By conditions are meant the turbulent velocity fluctuations and length scales in the flow, proportions of fuel, air, and burned gas in the mixture, and the mixture state.

Using the turbulent combustion model described in Sec. 14.4.2 (which is based on the description of the flame development and propagation process in Sec. 9.3), the flame development angle $\Delta\theta_d$ (the time to burn a few large eddies and establish a developed turbulent flame) can be expressed as[20]

$$\Delta\theta_d = C\left(\frac{l_i}{u'}\right)^{1/3}\left(\frac{l_M}{S_L}\right)^{2/3} \tag{9.51}$$

The turbulent flow field influences $\Delta\theta_d$ through l_t, u', and l_M, the integral scale, the turbulence intensity, and the microscale, respectively. The mixture composition influences $\Delta\theta_d$ through the laminar flame speed S_L. There is therefore a variability in the flame development period, since all of these quantities can vary in the vicinity of the spark plug on a cycle-by-cycle basis.

The pressure development during the rapid-burning developed turbulent flame propagation phase, when the flame spans the combustion chamber, depends on the average rate of burning in the flame. Thus, variations in the turbulent flow field and mixture composition across the gas entering the flame front are averaged out and are not important. However, variation in chamber-average quantities are significant. The combustion model in Sec. 14.4.2 leads to the following expression for the maximum burning rate:

$$\left(\frac{dm_b}{dt}\right)_{max} = \frac{Cm_f(A^*/B)(\rho_u^*/\rho_u)^{1.0/0.9}[(\bar{u}'^*S_L^2)/\nu_b]^{2/3}}{\nu^{*1/3}} \qquad (9.52)$$

Here, m_f is the mass of fuel in the chamber, h is the instantaneous (mean) clearance height, B is the bore, ρ the density, $\bar{u}'$ the average turbulence intensity across the flame front, and ν the kinematic viscosity; the * denotes the value at the time of the maximum burning rate and the subscript i the value at spark; C is a constant depending on engine design. It can be seen from Eq. (9.52) that cycle-by-cycle variations in the maximum burning rate can result from variations in the overall flow pattern within the combustion chamber (which vary $\bar{u}'^*$) and from variations in the amount of fuel (m_f) that enters the cylinder each cycle.† Also, it can be seen that variations in the flame development process will result in variations in the maximum burning rate because the crank angle at which the maximum burning rate occurs is shifted, and all the starred parameters in Eq. (9.52) will have different values.

From the discussion of flame development and structure in Sec. 9.3, and Eqs. (9.51) and (9.52) above, we would expect that mixture conditions and motion leading to slower flame development rates (longer flame development angles, $\Delta\theta_d$)—lower turbulence intensities and more dilute mixtures—would also give lower burning rates (longer rapid burning angles, $\Delta\theta_b$). Data from many different engines and a wide range of operating cases show that this is the case, on average, though there is substantial variation about the mean trend. Figure 9-35 shows these trends; it also shows that the standard deviation of $\Delta\theta_d$ and the standard deviation of $\Delta\theta_b$ for a given chamber and operating condition generally increase as the average burning process becomes slower.[40]

One final factor of importance is how variations in flame development and burning rate affect engine torque. With fixed spark timing, such variations in the combustion process cycle-by-cycle result in slower developing and/or burning

† Variations in the total amount of air, recycled exhaust gas, and residual in the chamber could also, for some operating regimes, be significant.

cycles being retarded and faster developing and/or burning cycles being over-advanced. The curve of torque versus combustion timing (relative to optimum timing), Fig. 9-3b, is almost independent of the burning rate; i.e., a given magnitude retard (of say 10°) relative to optimum timing gives almost the same reduction in torque for a very fast burn as it does for a very slow burn. This is because the burning process, for optimum timing, is centered at about 10° ATC independent of the burn rate, and retard or advance shifts this "center" by equal amounts for all burn rates.[16] One of the major advantages of fast-burn engines is now apparent. The magnitude of the variations in the flame development process and subsequent flame propagation rate are decreased as the burning rate is increased (see Fig. 9-35): the ratio of standard deviation in $\Delta\theta_d$ and $\Delta\theta_b$ to the mean values remains approximately constant. Thus, these smaller combustion variations in fast-burn engines, which correspond to modest retard and over-advance in nonaverage burn rate cycles, have little effect on torque. In contrast, the larger combustion variations of slow-burning engines result in significant cyclic torque variations.
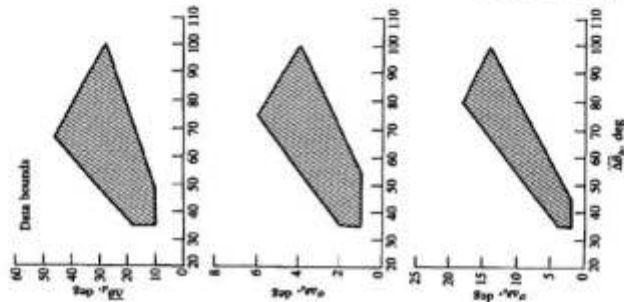
FIGURE 9-35
Variation in mean value of flame development angle $\Delta\theta_d$ (spark to 1 percent mass burned) and standard deviation of flame development angle and rapid burning angle $\Delta\theta_b$ (1 to 90 percent mass burned) with mean rapid burning angle $\Delta\theta_r$. Range of combustion chamber geometries and engine operating conditions.[40]

9.4.3 Partial Burning, Misfire, and Engine Stability

As the unburned mixture in a spark-ignition engine is leaned out with excess air or is diluted with increasing amounts of burned residual gas and exhaust gas recycle, the flame development period, the duration of the rapid burning phase, and the cycle-by-cycle fluctuations in the combustion process all increase. Eventually a point is reached where engine operation becomes rough and unstable, and hydrocarbon emissions increase rapidly. The point at which these phenomena occur effectively defines the engine's *stable operating limit*.† These phenomena result from the lengthening of all stages of the combustion process as the unburned mixture is diluted. With increasing dilution, first a fraction of the cycles burns so slowly that combustion is only just completed prior to exhaust valve opening. Then as burning lengthens further, in some cycles there is insufficient time to complete combustion within the cylinder; also, flame extinguishment before the exhaust valve opens and before the flame has propagated across the chamber may start to occur in some cycles. Finally, misfiring cycles where the mixture never ignites may start to occur. The proportion of partial burning or nonburning cycles increases rapidly if the mixture is made even more lean or dilute, and the point is soon reached where the engine will not run at all.

The impact on engine stability of increasing combustion variability, due to increased exhaust gas recycle at part-load, is shown in Fig. 9-36. Figure 9-36a shows the distributions of individual-cycle indicated mean effective pressure values for 0, 20, and 28 percent EGR. Without EGR at these conditions, the spread in imep is narrow. Increasing EGR widens the distribution significantly and cycles with low imep, and eventually zero imep, occur. Figure 9-36b shows how the coefficient of variation of imep and hydrocarbon emissions increase as EGR is increased. Slow burn, then partial burn, and then misfire cycles occur with increasing frequency. In the slow-burn cycles, combustion is complete but ends after 80° ATC and the indicated mean effective pressure is low (between 85 and 46 percent of the mean value). Imep in partial-burn cycles was less than 46 percent of the mean. In misfiring cycles, imep < 0. Empirically, it has been found that COV$_{imep}$ [see Eq. (9.50)] is about 10 percent at the engine's stable operating limit, which here occurs just before the onset of partial-burning cycles⁴⁹.

An explanation of combustion phenomenon at the engine stable operating limit has been developed by Quader.⁵⁰ It involves the following terms:

Ignition-limited spark timing or the **ignition limit**. The spark timing [advanced from maximum brake torque (MBT) timing] at which *misfire* (i.e., failure of flame initiation) first occurs at a given mixture composition, in a given small but arbitrary fraction of cycles (e.g., 0.5 to 1 percent).

---

† This limit has often been called the *lean operating limit*. Since what limits engine operation in practice is excessive torque fluctuations, cycle-by-cycle, and high hydrocarbon emissions, resulting from the use of mixtures made overly dilute with either air or burned gases (or with both), *stable operating limit* is a more appropriate term.
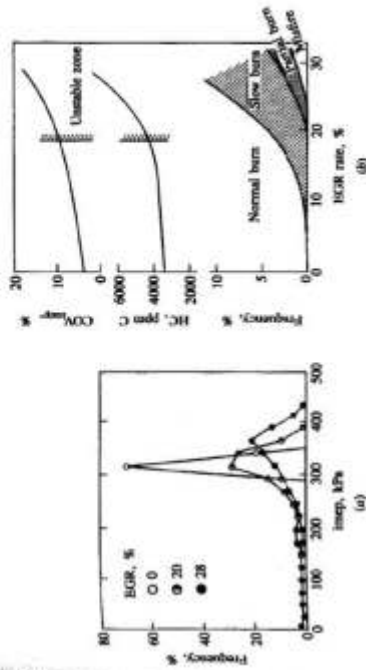
FIGURE 9-36
(a) Frequency distributions in indicated mean effective pressure at different EGR rates; 0 percent gave excellent engine stability, 20 percent acceptable stability, and 28 percent poor stability. (b) Coefficient of variation in imep, HC emissions, and percentage of normal, slow, partial-burn, and misfire cycles. Engine conditions: 1400 rev/min, $\phi = 1.0$, MBT timing, imep = 324 kPa.⁴⁹

Partial-burn-limited spark timing or the **partial-burn limit**. The spark timing (retarded from MBT) at which incomplete flame propagation occurs at a given mixture composition in a given small percent of the cycles (again, this frequency is selected arbitrarily for experimental convenience).

Lean misfire limit at MBT spark. The leanest mixture stoichiometry at which the engine could be stabilized to operate at MBT spark timing with a misfire frequency below a specified value (again, this frequency, usually a percent or less of the cycles, is selected arbitrarily for convenience). A *dilute misfire limit*, the maximum amount of exhaust gas recycle that can be absorbed at a given stoichiometry for stable engine operation, can be similarly defined.

Engine experiments have defined the locations of the ignition limit line and the partial-burn limit line; they are shown qualitatively in Fig. 9-37. At a given spark timing, on this spark timing versus equivalence ratio plot, progressive leaning of the mixture fed to the engine will lead to the onset of misfire or to the onset of partial burning, depending on the location of the lines and the spark timing selected. The individual figures show the possible interactions of the maximum brake torque (MBT) timing line—leaner mixtures require greater advance—with the ignition limit and partial-burn limit lines. At MBT timing, the partial-burn limit may or may not be reached prior to misfire or the ignition limit. It will depend on the engine and ignition system design and operating conditions. For spark timings retarded relative to MBT, partial burning and not failure of flame initiation is the primary cause of unstable engine operation.
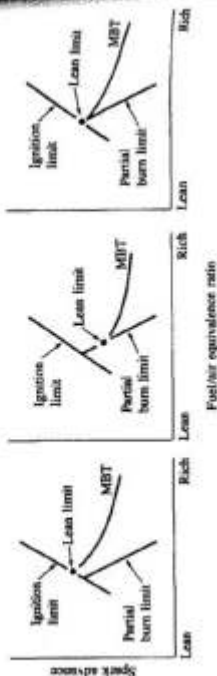
FIGURE 9-37
Schematics of three possible combinations of ignition limit, partial-burn limit, and MBT timing curves as function of fuel/air equivalence ratio. (From Quader.[50])

An example of these limiting combustion regimes for lean engine operation is shown in Fig. 9-38. With MBT timing, as the mixture is leaned out (at constant air flow rate), combustion in all cycles changes to partial burning in some cycles which changes to no ignition in some cycles at the ignition limit. In the partial-burning regime, the most common type of incomplete-combustion cycle was a slow-burning cycle which required more time to complete burning than was available: flame extinguishment during expansion was much less common. Engine performance measurements showed that the engine stability limit—evidenced by minimum fuel consumption and onset of rapid increase in HC emissions—occurred at $\phi = 0.65$, just before the partial-burn limit line where some slow-burning cycles occur but combustion is still complete in all cycles.[44]

From the above it is clear that flame initiation is a necessary but not sufficient condition for complete combustion. Too-slow flame development and prop-

FIGURE 9-38
Actual timing combustion regimes for lean-operating engine. 1200 rev/min, volumetric efficiency ≈ 60 percent, methane fuel, 40 mJ spark energy, 2.5 ms spark duration.[44]

gnition following successful ignition is usually the factor which limits engine operation with dilute mixtures. Experiments have shown that a limited interval, of order 80 crank angle degrees (depending on engine geometry and spark plug location), is available during the engine cycle when conditions are favorable for complete flame propagation.[49,51] Outside of this interval, the mixture pressure and temperature are too low, and the turbulence intensity is too low to sustain a sufficiently rapid rate of combustion. Thus, it is factors which increase the flame development and propagation rates which primarily extend the partial-burn limit.

9.5 SPARK IGNITION

In spark-ignition engines, the electrical discharge produced between the spark plug electrodes by the ignition system starts the combustion process close to the end of the compression stroke. The high-temperature plasma kernel created by the spark develops into a self-sustaining and propagating flame front—a thin reaction sheet where the exothermic combustion chemical reactions occur. The function of the ignition system is to initiate this flame propagation process, in a repeatable manner cycle-by-cycle, over the full load and speed range of the engine at the appropriate point in the engine cycle. Shadowgraph and schlieren photographs of the kernel created by the discharge between the plug electrodes, the growth of that kernel, and its transition to a propagating flame have already been presented in Figs. 9-19 and 9-20, and described in the accompanying text. A spark can arc from one electrode to another when a sufficiently high voltage is applied. Ignition systems commonly used to provide this spark are: battery ignition systems where the high voltage is obtained with an ignition coil (coil ignition systems); battery systems where the spark energy is stored in a capacitor and transferred as a high-voltage pulse to the spark plug by means of a special transformer (capacitive-discharge ignition systems); and magneto ignition systems where the magneto—a rotating magnet or armature—generates the current used to produce a high-voltage pulse.

This section reviews our basic understanding of electrical discharges in inflammable gas mixtures relevant to engine ignition (Sec. 9.5.1), the major design and operating characteristics of conventional engine ignition systems (Sec. 9.5.2), and, briefly, some alternative approaches to generating a propagating flame (Sec. 9.5.3).

9.5.1 Ignition Fundamentals

A spark can arc from one plug electrode to the other only if a sufficiently high voltage is applied. In a typical spark discharge, the electrical potential across the electrode gap is increased until breakdown of the intervening mixture occurs. Ionizing streamers then propagate from one electrode to the other. The impedance of the gap decreases drastically when a streamer reaches the opposite electrode, and the current through the gap increases rapidly. This stage of the
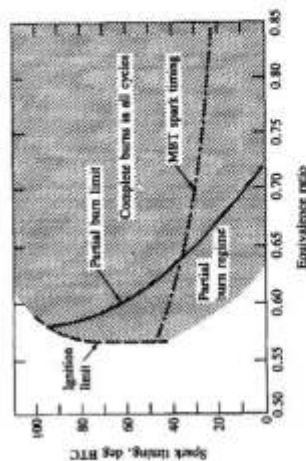
[3]   Ando, H., Arcoumanis, C., *Flow and Combustion in Reciprocating Engines*, Arcoumanis, C., Kamimoto, T., Berlin, Heidelberg: Springer/Verlag, 2008.

Chapter 3
Flow, Mixture Preparation and Combustion in Four-Stroke Direct-Injection Gasoline Engines

Hiromitsu Ando and Constantine (Dinos) Arcoumanis

3.1 Introduction

Since the launch of the first direct-injection (DI) gasoline engine in 1996, many Japanese and European car manufacturers have introduced this technology into the market with the expectation that it would spread worldwide. There was a commonly-shared view that it represented one of the most effective fuel economy improvement technologies for CO_2 reduction. Unfortunately, this technology hasn't achieved significant penetration into the mass production market even after 10 years from its launch, despite its obvious scientific and technological advantages.

In this chapter, the basic features of direct injection gasoline engines and the technologies adopted to realize them will be described. Then the technological evolutions over the last 10 years, through worldwide efforts by researchers in both industry and academia, will be reviewed and the prospects for the future of the direct-injection gasoline engine technology summarized.

Finally, the latest developments, i.e. the spray-guided combustion system representing the most promising concept for meeting the future CO_2 targets, are presented and examples are provided of the new generation of fuel injection systems that may allow the mass production of stratified direct injection gasoline engines for automotive applications.

3.2 Significance and Limits of the Direct-Injection Gasoline Engine Technology

It is generally accepted that the fuel economy improvements of DI gasoline engines are due to their higher compression ratio, higher specific heat ratio, pumping loss reduction (lean burn, EGR) and cooling loss reduction. The effects of these factors

Hiromitsu Ando
University of Fukui, Fukui, Japan, e-mail:

Constantine (Dinos) Arcoumanis
City University London, London, UK, e-mail: c.arcoumanis@city.ac.uk
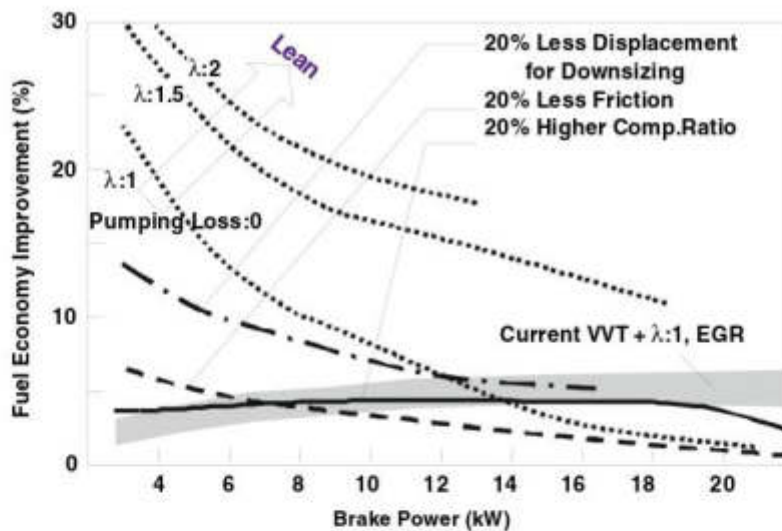
Fig. 3.1 Factors of fuel economy improvements

on fuel economy improvement are shown schematically in Fig. 3.1 and can be summarized as follows:

- Although there are large effects in the low load region, their influence diminishes at higher loads
- The effect of lean burn is large mainly due to the higher specific heat ratio rather than reduction of the pumping loss.
- The effect of higher specific heat ratio is maintained even at higher loads.
- The reason the DI gasoline engine technology is so efficient lies in its ability to take advantage of all the factors shown in Fig. 3.1, that is, the higher specific heat ratio realized by the stable lean burn, the higher compression ratio permitted by various anti-knock measures, the pumping loss reduction through lean burn and the cooling loss reduction through lowering of the burned gas temperature and mixture stratification effects.

Although the fuel economy benefits of direct-injection gasoline engines compared with their conventional port fuel injection counterparts were large at the timing of their launch in 1996, they have gradually lost some of their advantages over the last 10 years or so. The main factors are the following:

- Port fuel injection engines have improved significantly
- Breakthrough transmission technologies have emerged
- Operating conditions for the lean operation of DI gasoline engines could not be extended in order to meet the increasingly more stringent emission control requirements.

In addition, diesel engines have entered the European market with superior fuel efficiency and outstanding performance which reduced the momentum for research into DI gasoline engines.

The most distinctive evolution of port fuel injection engines has been the introduction of Variable Valve Timing (VVT) technologies which are now quite standard. Through the VVT, pumping loss reduction can be realized by using the Atkinson cycle effect where valve overlap is reduced under low load conditions to reduce the amount of residual gases. As a result of the improved combustion stability, engine idle speed can be reduced to the same level as in DI gasoline engines. In addition, by adjusting the timing of the valve overlap, hot exhaust gas is evacuated selectively which results in a slight improvement of the anti-knock characteristics; further improvement can be achieved through optimization of the thickness of the coolant flow wall between the combustion chamber and the water gallery. This can be realized through advances in the casting techniques which also contributes to the improvement of the anti-knock characteristics.
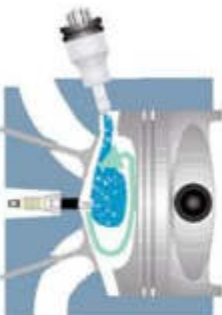
Almost all of the fuel economy improvement technologies are aiming at the low load conditions where thermal efficiency is rather poor. New transmissions such as CVT, multi-stage AT, multi-stage MT and AMT (Automated Manual Transmission) are technologies able to improve the powertrain efficiency by minimizing low load engine operation. The end result has been that the significant effects on fuel consumption reduction introduced by the DI gasoline engine have been gradually depreciated by the introduction of these transmission technologies.

Finally, although the evolution of the lean NOx catalyst technology has been remarkable, as will be described later, and its efficiency almost exceeds 90% under relatively wide ranging conditions, the efficiency of its competitor, the three-way catalyst used by port fuel injection engines, frequently exceeds 99%. It can therefore be argued that the efficiency of DI gasoline engines has been sacrificed in order to achieve emission levels comparable to those of port fuel injection engines. However, at present fuel consumption is attracting more attention than local exhaust emissions due to the emphasis on global warming, which is responsible for the renewed interest on DI Diesels and spray-guided stratified gasoline engines.

3.3 Technologies for Direct-Injection Gasoline Engines

It is well known that there are three basic concepts of direct injection gasoline engines: wall-guided, air-guided and spray-guided (Table 3.1). The concept investigated since the 1950 s had been the spray-guided concept in a configuration known as narrow/close spacing which locates closely the fuel spray and the spark plug in the central region of the cylinder head. [1, 2, 3, 4, 5, 6, 7]. In this concept, the spark plug is projected into a specific region of the fuel spray in order to achieve stable spark ignition and, as such, is it exposed to fuel droplets and/or a rich fuel-air mixture which encourages soot formation and spark plug fouling leading to misfires. It is only recently that the first encouraging signs have emerged for a successful

Table 3.1 Characteristics of gasoline direct-injection concepts

	Wall-guided	Air-guided	Spray-guided
Fuel economy	0	+	++
HC	0	+	++
Smoke	0	+	++
Power	0	−	0 to +
Robustness	0	−	−
Mixture preparation			

implementation of the spray-guided concept in a production engine; more details are provided later.

In 1996 a new concept called wide spacing was proposed [8]. This concept locates the fuel spray and the spark plug at some distance from each other. The fuel is injected towards the hot piston surface and the formed fuel vapour is directed towards the spark plug through a combination of piston-bowl geometry and strong in-cylinder bulk flow. The main advantage of this concept is that the spark plug fouling problem is resolved since the fuel approaches the spark plug after vaporization and mixing with air. The wall-guided engine concept used the following four main technologies:

- An in-cylinder flow called the reverse tumble generated by the upright straight intake ports.
- A swirl injector nozzle preparing well dispersed and atomized fuel spray at relatively low injection pressures.
- A piston with a cavity designed to optimize fuel spray impingement and redirection towards the spark plug
- Mixing control mechanism combining a homogeneous mixture realized by injection during the early stages of the intake stroke and a stratified mixture realized by injection during the later stages of the compression stroke.

After the launch of the first mass production DI gasoline engine in 1996, many alternative systems have been introduced into the market, employing swirl or tumble in place of the reverse tumble, injection systems using a swirl nozzle with an asymmetric spray shape named casting net spray and achieved through a novel nozzle hole design, and a fan-spray nozzle producing a two-dimensional thin shaped spray through a slit located at the outlet of the nozzle hole [9, 10, 11, 12]. Both systems are classified into the wide spacing concept, and have the common feature of using an intense in-cylinder flow and a purpose-designed piston cavity. Although some systems have been referred to as wall-guided and others as flow-guided, this classification seems to have little meaning since all first-generation DI gasoline engines in production have employed guidance by flow and wall simultaneously and experienced increased heat losses as a result of the complicated piston cavity with its larger surface area. Flow-guided is probably a term introduced to highlight the necessity for simplification of the piston crown geometry which more recently led to the introduction of the air-guided concept. Existing wall-guided combustion systems in production today are illustrated in Fig. 3.2 in terms of their corresponding piston and spray configurations.

The spray-guided concept under intense development today offers the best potential partly because of the possibility of using a simpler piston geometry. In order to realize the spray-guided concept, it is important to prevent impingement of fuel droplets onto the spark plug. To achieve this, many ideas have been proposed such as using an air-assisted injector, a spray collision injector, an outwards opening nozzle and a multi-hole nozzle, the latter two combined with higher pressure injection [13, 14, 15, 16, 17]. The targets are to dilute the fuel through air entrainment and
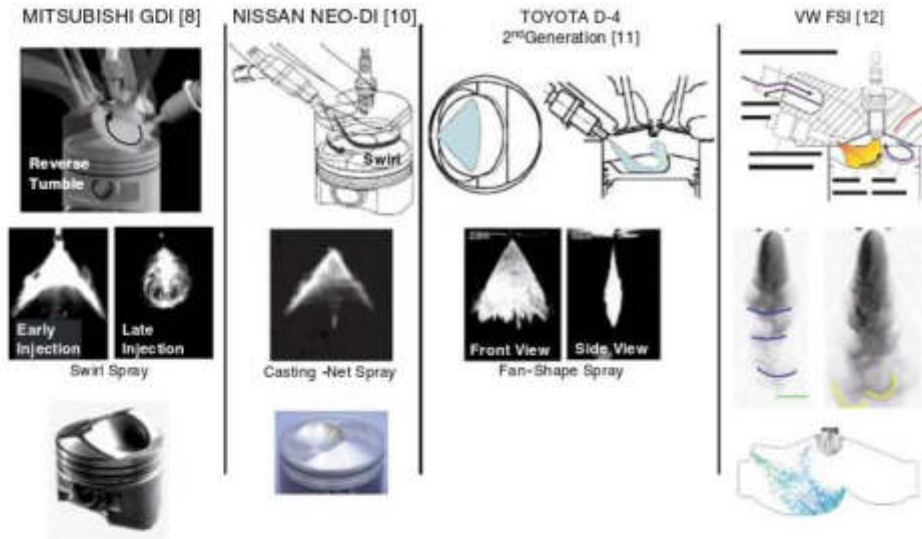
Fig. 3.2 Wall-guided gasoline direct-injection engines

improved atomization and to prevent the concentration of large fuel droplets and rich fuel vapour around the spark plug at the time of ignition through control of the spray development as a function of the thermodynamic in-cylinder conditions. Out of the emerging advanced fuel injection technologies, the multi-hole and outward-opening nozzles shown in Figs. 3.3 and 3.4, have real potential for becoming a

$$\frac{L}{D_1} < \frac{L}{D_2} \Longrightarrow \varphi_1 > \varphi_2$$
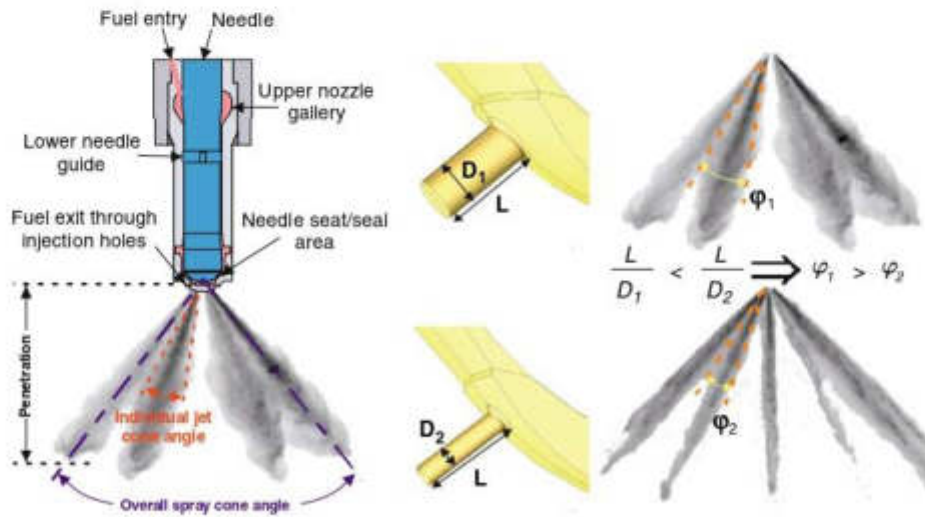
Fig. 3.3 Multi-hole injector for spray-guided gasoline direct injection engines and its key geometric parameters (spray and jet cone angle, hole diameter D and length L)
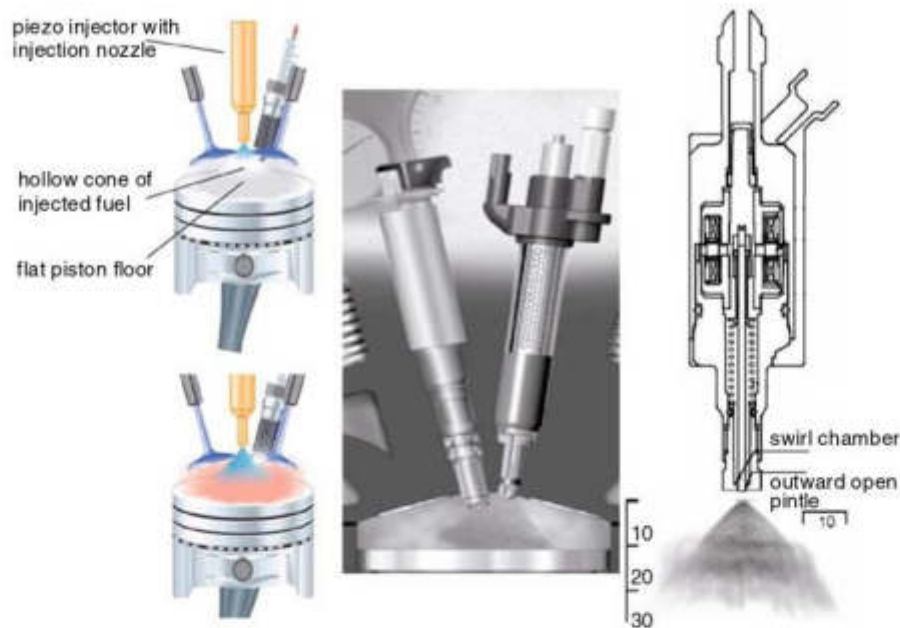
Fig. 3.4 Outward-opening injectors for spray-guided gasoline direct injection engines

mass production technology. In case of the multi-hole nozzle, the spark plug can be located between two of the sprays or over one of the sprays with a small aperture.

On the other hand, an outward-opening nozzle allows intense interaction between fuel droplets and the surrounding air, with atomized fuel concentrated on the outer periphery of the spray and the spark plug located in the recirculation zone formed at the leading edge of the spray. Both nozzle types are able to prepare a fuel-air mixture with near stoichiometric mixture strength around the spark plug while avoiding direct fuel spray impingement on the spark electrodes. More details about these two emerging fuel injection technologies and the generated sprays are given in subsequent sections. Overall, it can be argued that the potential of the spray-guided concept is much higher than the other two, despite its questionable robustness relative to the wall-guided concept (Table 3.1) due to the fact that its combustion development is controlled exclusively by the stability of the fuel spray characteristics.

3.4 General Features of Direct-Injection Gasoline Engines

The distinctive features of the DI gasoline engine are centred on its potential to achieve stable lean burn, to reduce and even eliminate engine-out unburned hydrocarbon emissions, to improve the anti-knock characteristics and to control torque precisely and quickly.

3.4.1 Lean Burn

Combustion of a lean fuel/air mixture should be discussed from two viewpoints: thermal efficiency and NOx formation. With respect to thermal efficiency, reduction of pumping losses, reduction of heat losses and higher specific heat ratio are the main advantages of lean burn. These are realized by increasing the total amount of working fluid, lowering the mean burned-gas temperature and increasing the percentage of diatomic molecules in the working fluid. All these effects can be achieved when the fuel-air mixture is overall lean. The DI gasoline engine which can maintain stable combustion under overall lean mixture conditions has the best potential to fully explore the thermodynamic advantages of lean burning.

With respect to NOx reduction, it is necessary to control the local equivalence ratio to remain always lean and to suppress flame propagation. The first requirement is to lower the local temperature or oxygen concentration in the burned gas region to prevent or reduce thermal NOx formation. The local temperature of the burned gas continues to increase by means of adiabatic compression realized in the region ahead of the propagating flame. In particular, when combustion is controlled by flame propagation, a non-uniform temperature distribution takes place in the cylinder which results in substantial NOx formation in the high temperature region. This is the reason for the second prerequisite for NOx reduction which is the elimination of flame propagation. The developing consensus on this point is that the ultimate solution for satisfying the requirement for NOx reduction is HCCI combustion, where simultaneous ignition of the mixture takes place across the cylinder volume and any temperature increases maintain mixture homogeneity.

It can be argued from the above that, from the viewpoint of NOx reduction, the direct-injection gasoline engine is not the optimal solution since the local combustion field is generated by spark ignition and the resulting flame traverses the combustion chamber tracing the fuel/air mixture with moderate local equivalence ratio. The temperature of the burned gas zone continues to rise and generates large amounts of thermal NOx. Therefore, the engine-out NOx emissions of direct injection gasoline engines operating under lean overall conditions remain at relatively high levels of about 0.2–0.35 of those in a conventional port fuel injection engine operating at stoichiometric conditions which necessitates a sophisticated aftertreatment system for reducing NOx.

When high output is desirable, it is necessary to operate the engine under stoichiometric or slightly rich conditions, because it is the fuel quantity that determines the engine output under conditions of fixed air mass. Therefore, lean engine operation with higher efficiency is quite restricted under realistic driving conditions. This represents the essential limit to the advantages that can be offered by direct injection gasoline engines.

3.4.2 Minimization of Fuel Escaping Power Generating Combustion

In port fuel injection engines, a lot of fuel is captured in the oil film formed on the cylinder liner surface and/or in the piston crevices. This fuel is scraped and released

from the walls and cavities by the ascending piston during the exhaust stroke and is burned rather quickly in the cylinder when the intake valves open, due to the intense turbulence generated and the large amount of fresh air introduced into the cylinder during the opening period of the intake valves. Although this fuel is not strictly counted as unburned hydrocarbons, the fuel burned during the exhaust and intake strokes which does not contribute to the work output of the engine should be treated as escaped fuel. It is sometimes overlooked that escaping fuel is one of the most important factors penalizing the thermal efficiency of port fuel injection engines. It is thus very fortunate that DI gasoline engines, where the air around the cylinder liner does not contain fuel, offer the possibility of minimizing the fuel escape and enhancing further their thermal efficiency.

3.4.3 Improvement of Anti-Knock Characteristics

At higher loads, the DI gasoline engine operates by adopting the early injection strategy where the combustion characteristics are identical to those of the premixed port injection engine. However, improved engine performance can be realized by the higher volumetric efficiency and the improved anti-knock characteristics.

In the case of port fuel injection engines, the latent heat of evaporation is supplied by the surface of the intake port(s), the intake valves or the cylinder liner. In the case of early direct injection, the fuel spray follows the piston and the impingement of the liquid fuel on its crown can be carefully minimized. Therefore, the latent heat is supplied to the fuel by the intake air and this causes efficient charge air cooling; it is estimated that the charge air is cooled by about 15 K. This implies that the gas temperature at the end of the compression stroke can be reduced by about 30 K, which enhances the volumetric efficiency and suppresses knocking.

In the case of port fuel injection engines, significant transient knock can take place during several cycles at the start of vehicle acceleration. Transient knock is caused by selective transport of low boiling point gasoline components with their lower octane numbers [18]. The direct injection engine, however, is not affected by such transient knock because all of the gasoline components are transported into the cylinder. A knock suppression period, caused by delay in the surface heating of the combustion chamber, follows the transient knock period. As a result, in the case of direct-injection gasoline engines that are not sensitive to transient knock, ignition timing can be advanced by 10 s or more at the start of the acceleration period which, in general, is sufficient to complete the vehicle acceleration.

For further improvements of the full-load performance, an alternative knock suppression method named 'two-stage mixing' has been proposed [19]. Since there is consensus that the most distinctive feature of a direct injection engine is the 'freedom of mixing', 'two-stage mixing' takes advantage of this freedom for knock suppression; Fig. 3.5 illustrates the relevant procedure. Fuel is injected into the engine cylinder twice; the first injection takes place during the early stages of the intake stroke to prepare a premixed very lean mixture while the second injection takes place during the later stages of the compression stroke to prepare a stratified mixture for ignition.
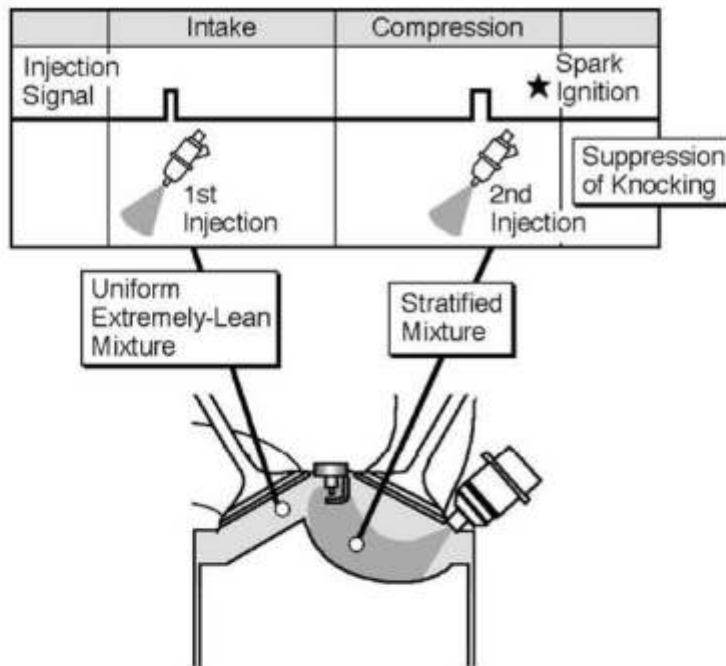
Fig. 3.5 Characteristics of two-stage mixing [18]

This knock-prevention process can be explained as follows:

- The premixed mixture prepared by the first injection is too lean to cause knock.
- The stratified mixture prepared by the second injection may form a stoichiometric mixture at some locations, however, the life time of that mixture prior to ignition is not long enough for knock precursor reactions to proceed.

Although in the case of a single late injection, a large amount of soot can be emitted when the average mixture strength becomes rich, in the case of two-stage mixing soot emission is not a problem except at very low engine speeds when the average air/fuel ratio is 12. The process of soot formation and oxidation can be explained as follows:

- When spark ignition takes place, only the stratified charge is ignited because the surrounding premixed mixture is too lean to be ignited.
- Combustion products generated in the rich stratified mixture containing CO and soot are transported by means of flame propagation into the surrounding lean mixture.
- Soot acts as ignition site and ignites the very lean mixture beyond its flammability limit; in this process CO may assist ignition and soot is oxidized in the combustion zone of the lean mixture.

3.4.4 Precise and Rapid Torque Management

In terms of engine management, the direct injection gasoline engine technology offers another advantage, the 'freedom of control', that is the potential of precise and quick management of torque with the best possible robustness.

In the case of port fuel injection engines, engine torque is controlled by the intake-air volume and, thus, is dependent on throttle valve operation. Even with an electronic throttling system, a change in air intake is somewhat delayed because the intake manifold with its large volume is located downstream of the throttle valve. In the case of direct-injection gasoline engines, however, torque can be easily controlled by just changing the injected fuel quantity while retaining the air intake constant; to be more specific, the torque in each cylinder and at each engine revolution can be freely controlled. This freedom can be used in the idle stop application [20] of a hybrid vehicle with a small capacity motor and in the advanced engine-transmission integrated control system. To gain customer acceptance, the idle-stop system needs to be able to achieve fast start and acceleration of the vehicle under the standstill mode.

Figure 3.6 compares the starting process in a port fuel injection engine with that of a direct injection engine. In the port fuel injection engine, the procedure for the earliest re-starting process is as follows: (i) detect the cylinder under exhaust stroke at the engine cranking, (ii) inject fuel into the port of the cylinder, (iii) the injected fuel is carried into the cylinder during the intake stroke and is compressed and burned. In the case of the direct injection gasoline engine, the cylinder under compression is detected at the beginning of the engine cranking and fuel is injected
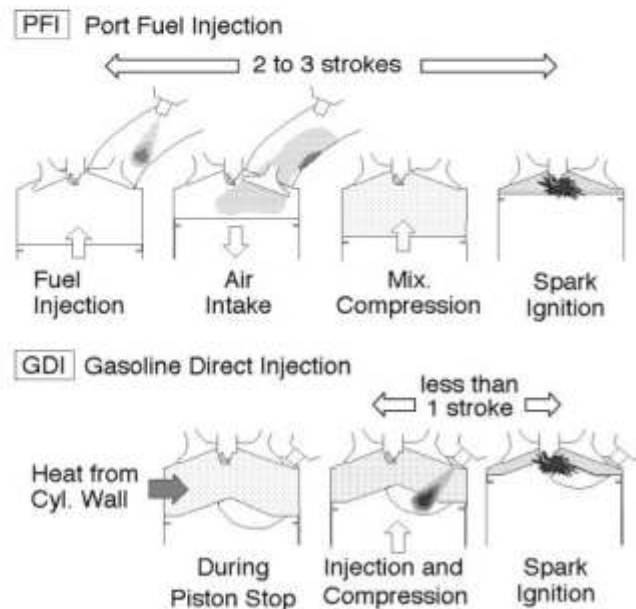

Fig. 3.6 Starting process of port injection engine vs. direct injection engine [19]

into this cylinder. The air temperature in that cylinder is already high enough due to the heated cylinder walls even at the suspended engine operation, and the fuel injected at the later stage of the compression stroke becomes ready for ignition, with complete combustion taking place once the mixture is ignited. Consequently, the delay in torque generation is minimized to one-fifth or one-tenth of that in the port fuel injection engine. Accordingly, the vehicle can execute a desirable idle-stop without giving the driver any uncomfortable feeling. Recently, the possibility to start the engine only by fuel injection, that is, without using the starter motor has been considered. The applicability of this idea will depend on the simplicity of the device to control the piston phase at the starting point of the engine.

The concept of 'freedom of control' can be applied to hybrid electric vehicles which require starting the engine when the vehicle is driven by the motor. Smooth and gradual torque increase from zero is desirable to suppress the shock caused by the torque gap. In the case of port fuel injection engines, however, it is difficult to maintain stable combustion under extremely low air and fuel conditions and some torque gap is inevitable. On the contrary, in a direct injection gasoline engine where stable combustion can be realized by a very small fuel quantity while keeping an overall very lean mixture, very low torque can be generated and the shock associated with engine starting can be eliminated completely. This effect is quite distinctive in a hybrid system using a small motor where frequent switching between motor and engine takes place.

The advantage of the DI gasoline engine in controlling the cycle- and cylinder-resolved torque by tuning only the injected fuel quantity is used in the integrated control of engine and transmission. In the case of CVT, for example, it is used to compensate the harsh and tough trade-offs of the efficiency, drivability and durability. It is also used to solve the problem of the tensional resonance caused by the connection of the shaft with poor rigidity to the pulleys with large inertia through rapid torque management synchronized with the natural vibration of the system. It can thus be argued that the strategies for resolving the trade-off problems using the 'freedom of torque' management approach may allow advances in various fields where technological breakthroughs are urgently needed.

3.5 Possibility for Further Reduction in NOx Emissions

The worldwide enforcement of exhaust emission reductions and the trend towards more stringent targets is creating difficulties for production direct injection gasoline engines to meet future regulations. Although it cannot be denied that they are handicapped relative to the mature three-way catalyst technology for port fuel injection engines, the catalyst and the catalyst reaction control technologies for direct injection gasoline engines have also advanced. At present, the developed exhaust after treatment system has the potential to meet the most stringent worldwide emission regulations of Japan, Europe, U.S.A. and California, in particular.

3.5.1 Advances in Lean NOx Catalyst Technologies

The first production direct injection gasoline engine adopted for NOx emission control a selective reduction catalyst which is less sensitive to sulphur poisoning. This has later been replaced by a NOx trap catalyst of higher performance as a result of the worldwide trend in gasoline sulphur reduction. A NOx trap catalyst operates by capturing NOx under lean operating conditions and converting it into nitrate of alkali metals and alkaline-earth metals. When the engine operating condition changes from lean to rich during short acceleration periods, the nitrates are converted into carbonates by means of the CO contained in the exhaust gas and the adsorbed NOx is released.

The first NOx trap used Ba as a NOx adsorbing metal. Although it was known that K shows better performance than Ba, K is lower in electronegativity and is thermally unstable. It is, therefore, for the following reasons that the K adsorbent has been unpopular in its use for automotive NOx control:

- The higher the catalyst temperature is, the more K adsorbent outpours from the wash coat.
- K penetrates into the catalyst substrate.

Despite all its problems, recent advances of NOx trap technologies have been remarkable and some solutions were found [21, 22]. The first solution is to mix K adsorbent with zeolite for K stabilization, while the second solution is to coat the substrate surface with Si which, due to its stronger chemical affinity to K, prevents K penetration into the substrate. It was thus found that, by adding silica compound to the wash coat, it is possible to have high-storage performance up to 550°C and extended lean-operating region, while the heat resistance can be improved to 850°C.

Overall, progress of NOx trap technologies has steadily continued. New methods, such as a method to stabilize K by adding MgO, a method to use Na in place of K and a pore distribution optimization method [23, 24], were proposed. The most effective breakthrough technology has been the introduction of metal substrates. Because the problem of the K penetration can now be resolved, higher efficiency can be achieved by the increase of doped K. Besides these NOx trap technologies, new direct NOx decomposition catalysts using new materials such as La-Ba-Mn-In-Cu perovskite or the oxides of Ba and CO were also proposed [25, 26, 27]. It is expected that one of these advanced concepts for lean NOx catalysts will reach sooner or later mass production.

3.5.2 Catalyst Reaction Control by Mixture Preparation and Combustion Control

Various catalyst reaction control technologies that make use of the distinctive feature of direct injection gasoline engines, the 'freedom of mixing', have been proposed [28, 29]. One of them is the method of 'two-stage combustion', illustrated in
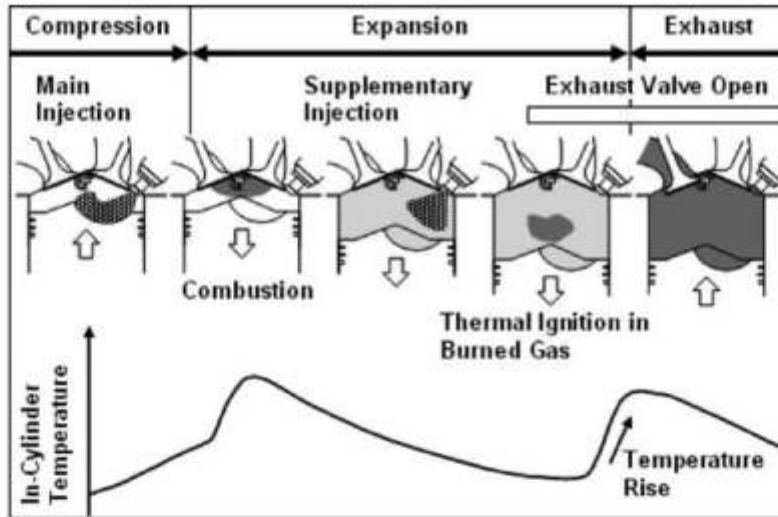
Fig. 3.7 Characteristics of two-stage combustion [28]

Fig. 3.7. The main fuel injection takes place at the later stages of the compression stroke for achieving stratified charge combustion and combustion is completed at an early stage of the expansion stroke; a second injection of a smaller amount of fuel takes place at the later stage of the expansion stroke. Since a large amount of excess air remains in the burned gas and its temperature is sufficiently high, this supplementary fuel is easily burned and the exhaust gas temperature rises rapidly; the end result is that the catalyst is warmed up quickly. Unfortunately, two-stage combustion consumes approximately twice the fuel of conventional combustion because this supplementary fuel is converted to heat rather than to work. Therefore, increased fuel consumption has accompanied the concept of 'two stage combustion'.

To solve this problem, an additional method named 'Stratified Slightly Lean Combustion' was proposed where the fuel is injected at an early stage of the compression stroke to realize a stratified slightly lean mixture. Figure 3.8 illustrates the situation of the exhaust gas which contains substantial amounts of CO generated in the slightly rich zone and O_2 remaining in the slightly lean zone; CO and O_2 are carried over the catalyst surface where the light-off temperature of CO is as low as 150°C. The catalyst is heated for several seconds up to this temperature in order to start the catalytic oxidation of CO on its surface by the two-stage combustion. Thereafter, the combustion mode is switched over to the stratified slightly lean combustion where CO oxidation instantly begins on the catalyst surface and HC oxidation is also induced because the catalyst surface is heated selectively to a temperature higher than the light-off temperature of HC. In such a manner, the stratified slightly lean combustion substantially reduces the time duration of the two-stage combustion. The end result is that better fuel economy and faster catalyst light-off can be simultaneously achieved.
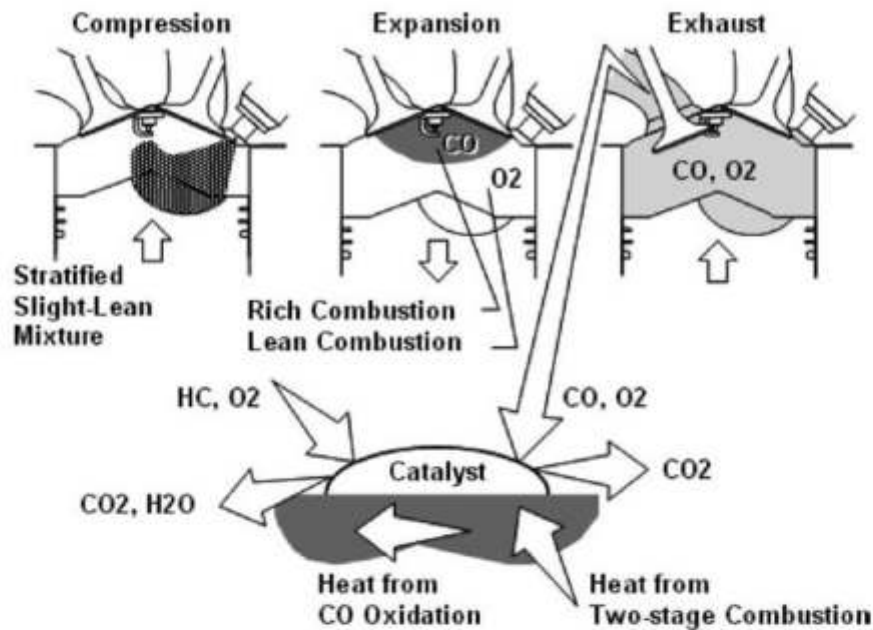
Fig. 3.8 Characteristics of stratified slight-lean combustion [28]

The NOx reduction method using a NOx-trap catalyst is basically a reversible reaction of nitrate and carbonate. NOx is adsorbed by the formulation of nitrate and desorbed by the formulation of carbonate. The desorbed NOx is then converted into O_2 and N_2 in the three-way catalyst located downstream of the NOx trap by a reducing agent such as the CO present in the exhaust gas. During this process, part of the CO in the exhaust is consumed for carbonate formulation which leads to a shortage of the reducing agent to purify the desorbed NOx; the result is that substantial NOx is temporarily emitted from the tail-pipe. The increase of engine-out CO leads to more residual CO in the exhaust gas downstream of the NOx trap catalyst, but it is not sufficient to reduce subsequent NOx levels since the desorbed NOx also increases with CO emissions. Thus the catalyst needs to be provided with a reducing agent that is less effective on NOx desorption. For such purpose, a method of introducing an additional fuel injection for a very limited period during the expansion stroke, to supply hydrocarbons to the three-way catalyst, was invented. A large amount of K loaded onto the NOx-trap catalyst as adsorbent interacts with π electrons in the HC, and then the HC adsorption into the noble metal is disturbed. Consequently, the HC emissions are not reduced by the NOx trap catalyst and flow into the three-way catalyst downstream; this HC is decomposed into CO and H_2 which achieve the reduction of the desorbed NOx. The end result of this control technology is that NOx emissions at NOx purging are minimized.

In summary, these are the methods realized by the inherent advantage of direct injection gasoline engines, the 'freedom of mixing'. In the relevant research of cata- lyst reactions, it is important to remember this important advantage and pursue any

innovative ideas in an effort to solve the remaining problem of the quite complicated NOx trap catalyst system which remains a prerequisite for the future market success of direct injection gasoline engines.

3.6 Present and Future Prospects

More than 10 years have already elapsed since the launch of the first mass production direct injection gasoline engine. In the meantime, some non-engine breakthrough technologies have emerged, other combustion concepts such as the Homogeneous Charge Compression Ignition (HCCI) have received worldwide attention, diesel engines are now well established and recently the spray-guided concept using either an outward opening nozzle or a multi-hole nozzle is emerging as the most serious technology offering hope for the mass production of the direct-ignition gasoline engine. Needless to say that continuation of the evolution of catalyst technologies and catalyst reaction control technologies is also very important to the success of the spray-guided concept. It can also be argued that the outstanding torque control capability of DI engines will be very useful in various applications and may contribute to their popularity.

3.6.1 Stoichiometric DI Gasoline Engines

As research into the spray-guided concept has intensified and the expectations from the wall-guided concept have diminished, a compromise mid-way solution, the stoichiometric direct injection gasoline engine has started attracting a lot of attention [30, 31, 32]. This is the concept which aims to achieve relatively modest improvements in thermal efficiency through a higher compression ratio. In particular, the fuel is injected into the cylinder during the induction stroke and a homogeneous and stoichiometric mixture is available for ignition in a similar way to PFI engines. This approach has the advantage of simultaneously reducing hydrocarbon emissions, improving power and fuel economy while taking advantage of the 99% or so efficiency of catalytic converters. These simultaneous benefits are the result of:

- High pressure (5–20–MPa) fuel injection and careful timing of injection which prevent impingement of the fuel on the piston and cylinder walls leading to low HC emissions.
- Charge cooling by the evaporating spray (~ 15 K) which allows higher compression ratios ($\sim 12:1$) to be used, leading to increased power (up to $\sim 15\%$) and a very modest fuel economy advantage (3–5%) relative to PFI engines [33].

Further benefits can be achieved with cylinder deactivation or stratified-charge-start options [34]. Although the overall fuel economy improvement is not significant when it is applied to naturally aspirated engines, as shown in Table 3.2, the homogeneous and stoichiometric concept offers higher potential when it is applied

Table 3.2 Feasibility of gasoline direct-injection applications

Features	Sources	Stoichiometric		Stoichiometric and lean	
		NA	TC	NA	TC
Fuel economy	Lean burn	None	None	++	+++
	High EGR tolerance	+	+	+	+
	High copm. ratio	+	++	+	++
Power	Charge air cooling	+	++	+	++
	2-Stage mixing	+	+++	+	+++
Exhaust gas	Catalyst warm-up by 2-stage combustion	+	++	+	++
	Catalyst warm-up by stratified slight-lean	+	++	+	++
	Necessity of NOx trap	None	None	–	—
Response	Torque control by fuel quantity	None	None	++	++
	Quick start-up	+ for HEV or idle stop application			

to turbocharged and/or supercharged engines. It has been argued for quite some time that engine downsizing with turbocharging is one of the candidates for fuel economy improvements, but the idea has not spread widely. To increase the knock limited torque, it is required to lower the compression ratio resulting in lower thermal efficiency. The improved anti-knock characteristics of the direct injection gasoline engine minimize the requirement for compression ratio reduction. Therefore, the combination of turbocharging and direct injection gasoline engine technology makes a lot of scientific sense since it takes advantage of all the positive aspects of turbocharging but with no negatives as direct injection counterbalances these. When combined with the downsizing concept, it offers the possibility that a certain percentage of the thermal efficiency loss caused by 'giving up' lean burn may be recovered. In the case of a direct injection gasoline, engine turbo-lag which represents one of the generic weaknesses of turbocharging, can be reduced. Because the engine is operated in the lean mode before the acceleration, the mass of exhaust gas flowing into the turbine is larger than that of a port fuel injection turbocharged engine. The rotational speed of the turbine and compressor before acceleration is maintained at two or three times higher level than that of a port fuel injection engine, thus reducing turbo-lag. Inferior low-end torque has been considered to be another disadvantage of turbocharging; however, this problem has been improved by advanced transmission. Through the rapid shift of the gear position or the pulley ratio to switch the engine speed to high, the requirement for low-end torque has been reduced. What remains an issue is the customers' degree of acceptability of the turbocharging concept. However, through turbocharging the driver and passengers can experience the superb acceleration feeling during overtaking from high speed cruising since, at the same gear position, the vehicle speed increases linearly with increasing boost pressure. Although this is a big advantage for European drivers, it is less so for Japanese drivers who cannot experience a similar feeling since the speed limit in Japanese highways is only 100 km/h.

3.6.2 Stratified Spray-Guided DI Gasoline Engines

As already mentioned earlier, more than 10 years after the introduction into the mass market of the GDI Mitsubishi engine the expectations are that the wall-guided concept (1st generation DI engine) will soon be replaced by the more promising spray-guided concept where the injector and the spark plug are closely spaced in the central part of the cylinder head (Fig. 3.9). The main reason for the introduction of the spray-guided concept, despite its obvious technical difficulties, is that it is capable of substantially expanding the speed-load range of the stratified-charge operation, thus taking better advantage of its superior fuel economy relative to the first generation DI engines. The expansion of the stratified-charge operating range is due to the different mechanism for achieving fuel stratification which does not depend on either the shape of the piston cavity or the in-cylinder flow. The advantages of the spray-guided concept in terms of fuel-economy, HC and smoke emissions are summarized in Fig. 3.2 relative to the two alternative configurations of the wall-guided and air-guided systems. The close proximity of the injector to the spark-plug results in short time separation between injection and ignition which necessitates very accurate positioning of a vaporized fuel cloud of the right mixture stoichiometry at the spark plug gap at the time of ignition. Although conceptually this seems rather straightforward, in practice it is very difficult to have high
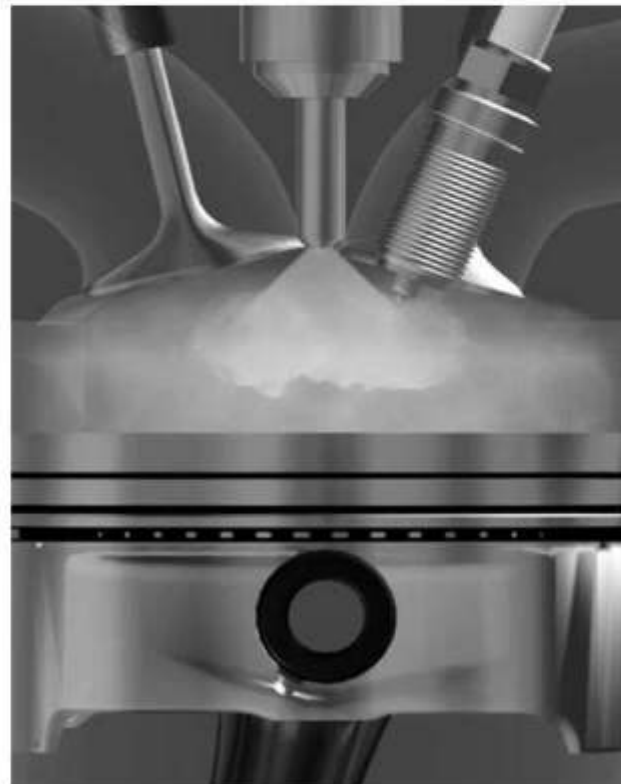

Fig. 3.9 Cylinder head configuration for spray-guided concept [35]

repeatability in the spray structure, mixture formation and ignition for every injection that takes place in the engine cylinder. It should be stressed that the local conditions are quite unfavourable for robust ignition due to:

- high gas-and liquid-phase velocities in the vicinity of the spark gap
- the presence occasionally of large droplets
- fluctuations in the spray cone angle
- variations in the spark duration and its stretching during ignition
- variations in the local pre-injection velocity field and air/fuel ratio.

It can thus be argued that the characteristics of the fuel injection system play the dominant role in the repeatability of the combustion process which will determine the chances of success of the spray-guided stratified DI gasoline engine concept as a mass-production technology for automotive applications.

Two types of fuel injectors have been developed over the last 2 years specifically for DI gasoline engines: the solenoid-actuated multi-hole injectors and the piezo-electric outward-opening pintle injectors. Both offer significant advantages relative to the swirl injectors, used extensively in the first-generation DI gasoline engines based on the wall-guided concept (Table 3.3), and deserve further attention and discussion.

3.6.2.1 Piezoelectric Outward-Opening Pintle Injectors

Contrary to the multi-hole injectors which show close similarities with their diesel counterparts, the outward-opening pintle injector represents a relatively new design purpose-built for the second-generation DI gasoline engines. Figure 3.10 gives detailed information about their standard design as well as some of its numerous variants which have been tested as means of understanding the all-important link between needle internal design and spray characteristics. The fuel from the rail and high-pressure connecting pipe is entering into the nozzle gallery and is directed towards the nozzle exit through three or four flow passages located at the space between the lower needle guide and the nozzle body. Below the lower guide there is a dead volume where the four high speed jets entering from the flow passages are mixing before the fuel exits the injector in the form of a hollow-cone spray; the direction of the spray is mainly determined by the seat angle of the pintle-type needle. More details about the internal flow and its link with the spray characteristics are provided in [37, 38, 39, 40] representing experiments and CFD calculations in both real-size and large scale models of outward-opening pintle injectors. High speed spray images taken just at the nozzle exit at atmospheric back pressure (Fig. 3.11) have revealed the formation of strings/ligaments originating at the annular area of the opening pintle whose location is not fixed during the injection period and their spacing varies as a function of the fuel flow rate. Although the origin of these ligaments/strings has been attributed to the formation of a two-phase flow inside the nozzle, associated with either cavitation or air entrainment, there are still uncertainties in the cause-and-effect relationship between internal flow and spray

Table 3.3 Comparison of the three dominant high-pressure injector nozzles [36]

Nozzle configurations	Swirl	Outward opening	Multi-hole
Spray stability/tolerance	+	++	+
Flexibility of spray pattern	+	0	++
Resistance against backpressure influence	–	++	++
Multi-injection capability	0	+	0
Costs	0	–	+
Robustness against fouling	+	+	+

Fuel entry

Upper nozzle gallery

Lower needle guide

Flow passage

Dead volume below
lower needle guide

Needle seat/seal area

Fuel exit through 360°
circumference

(a)

Standard design

3-passage guide

Ring below
the guides

20° slope
with long stem

Inward seal band

(b)

(c)
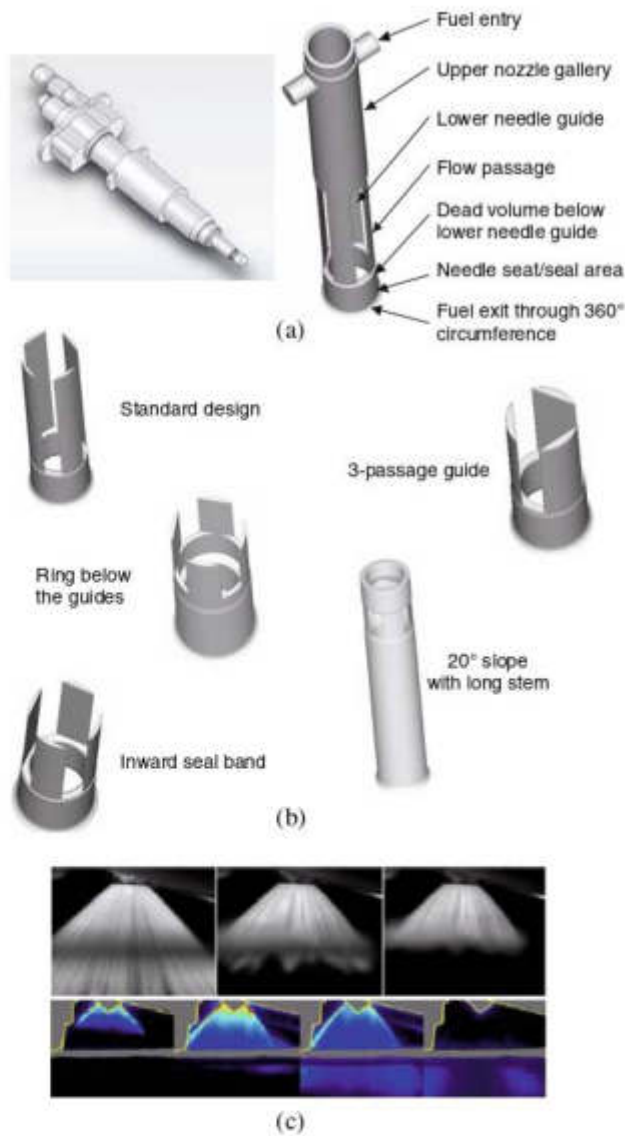
Fig. 3.10 Outward-opening pintle injector. (**a**) Standard design details. (**b**) Nozzle alternative designs (**c**) in-cylinder Mie and LIF images

structure. Independent of the mechanism of formation of the hollow-cone spray and its emergence as clearly separated liquid ligaments rather than as a continuous film, the penetration of this structure into the engine cylinder is a function of the prevailing thermodynamic conditions (pressure/density/temperature). As Fig. 3.12 shows, the 'streaky' structure of the spray under nearly atmospheric conditions gradually diminishes with increasing back pressure, giving rise to a more compact fuel cloud with clearly identifiable leading edge vortices: Furthermore, the position of
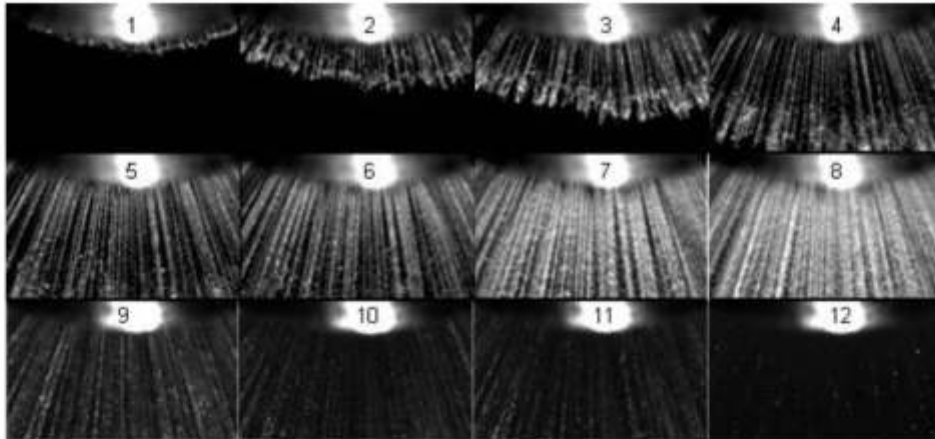
Fig. 3.11 High-speed sequence of spray images taken at the nozzle exit vicinity showing the initial development of the strings (1–4), during the main injection period (5–8) and at the end of injection (9–12) [57]

the leading edge vortex varies with cylinder pressure and, fortunately, higher back pressure forces this vortex to remain close to the spark plug [41]; injection timing thus offers some mechanism for controlling the position of this critical for ignition vortex.
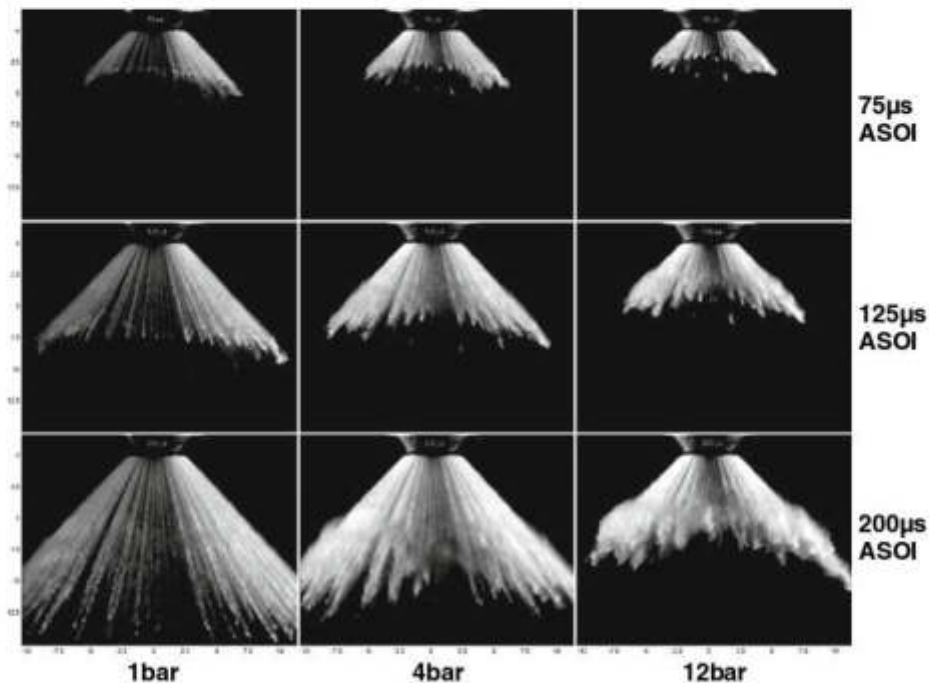
Fig. 3.12 High-speed sequence of spray images from a single injection event, showing the spatial distribution of fuel at different chamber pressures [40]

The creation of vortices at the spray edge and their stability in space relative to the spark plug represent jointly the key to the success of the spray-guided concept. In addition, the leading vortex grows in size as more fuel is injected [42], thus offering some flexibility in the positioning of the spark plug for both stratified and stoichiometric operating conditions. Due to the hollow spray configuration, air is entrained from underneath, which creates an upwards air motion and sets up a second, clockwise rotating vortex that balances the main anti-clockwise leading edge vortex formed earlier (Fig. 3.13a). The two vortices together are converting the kinetic energy of the spray into rotational motion, thus effectively reducing the spray penetration [42] and containing it near the spark plug. Interestingly enough, if the same amount of fuel is divided into two separate injections separated by a small dwell time, a third vortex is formed at the outer edge of the spray, which assists in bringing the fuel cloud closer to the spark plug (Fig. 3.13b). Confirmation of the degree of atomization and the 'ignitability' of the spray has been provided by PDA measurements of the droplets' size and velocity near the spark plug (Fig. 3.14) which confirmed that the droplet velocities are lower and their size smaller in the recirculation zone near the spark plug than in the leading edge zone. However, in the case of split injection the opposite trend has been observed [41].

The relative positioning of the spray's 'ignitable' recirculation zone with respect to the spark gap is illustrated in Fig. 3.15 for two extreme cases representing unsuccessful ignition and associated with large injection-to-injection, variations of the spray cone angle. In one case (Fig. 3.15a) the spray impinges on the electrodes as a result of air trapped at the needle head which gives rise to a spray cone angle larger than the nominal and, in the other (Fig. 3.15b) the cone angle is smaller than the nominal, representing the well-known phenomenon of hydraulic flip where air is trapped this time on the cartridge side.
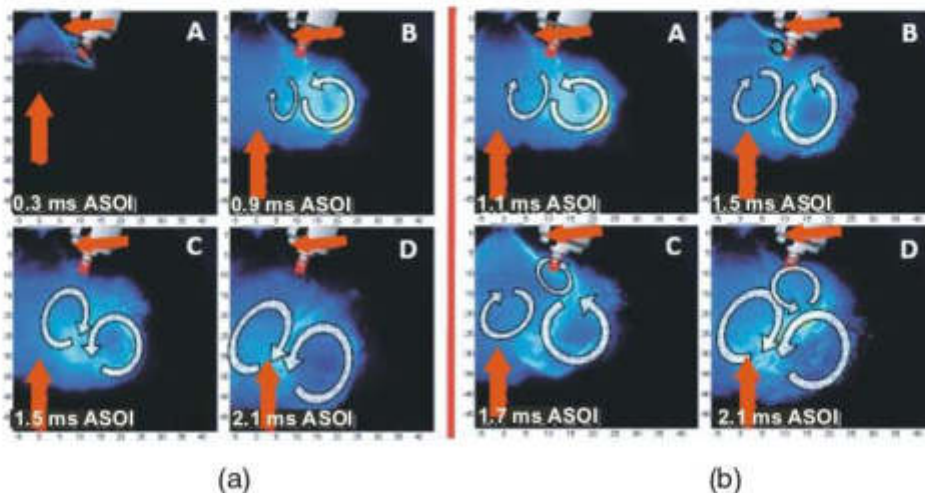

Fig. 3.13 Planar Mie images showing the liquid fuel distribution at four time steps from a (**a**) single and (**b**) split injection event [42]
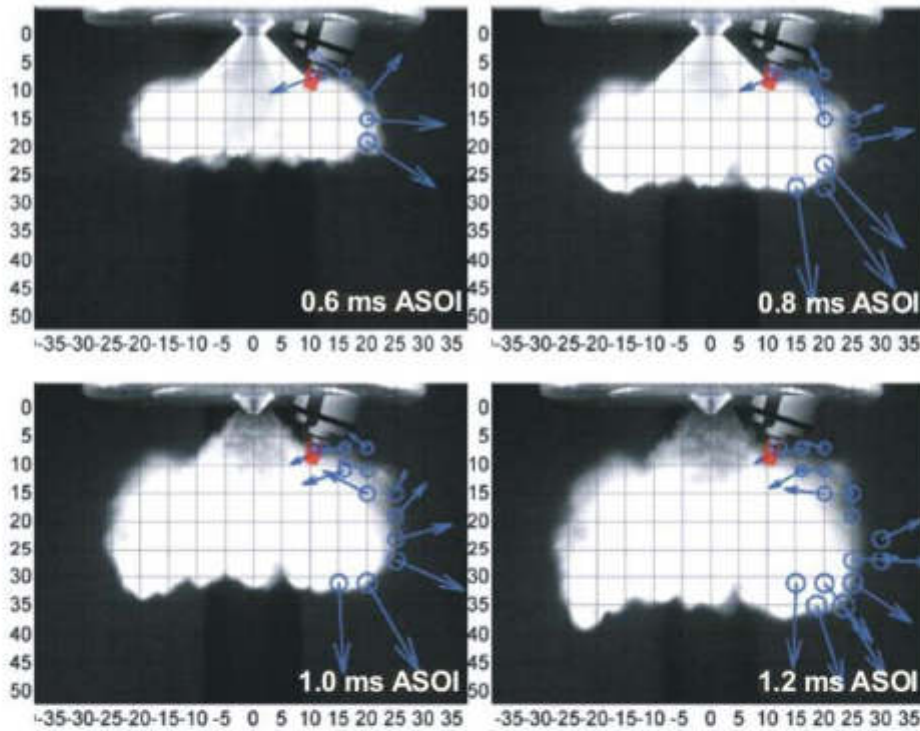
Fig. 3.14 High-speed images taken at four time steps during a single injection event with superimposed averaged PDA data (D_{10} droplet diameters and velocities); fuel temperature is at 90°C [42]

The combustion stability and robustness of an outward-opening pintle injector were examined in an optical engine under stratified charge conditions [42]. A stable operating window for misfire-free combustion was identified, extending over 4–6 CAD, with partial burns and misfires increasing as the rich and the lean burn zones were approached (Fig. 3.16a); this operating window of ±2.5 CAD is defined in terms of phase shifting the ignition from the end of injection at the given operating conditions (Fig. 3.16b).

Finally, the effect of flash-boiling on the spray generated by an outward-opening pintle injector has been investigated in a spray chamber [43]; flash-boiling occurs when the fuel temperature is high and the cylinder pressure is low as in the case of a warmed-up engine starting with nearly closed throttle. The results shown in Fig. 3.17 have revealed that at low fuel temperatures and back-pressures (non-flashing) the recirculation zone at the spray edge almost diminishes but increasing fuel temperature initiates flash boiling (superheated state) which increases the spray frontal area and restores the leading edge vortex. Nevertheless, the spray formed by the outward-opening pintle injector maintained its characteristic spray throughout these fuel temperature variations, with a small reduction in spray penetration occurring at flash boiling.
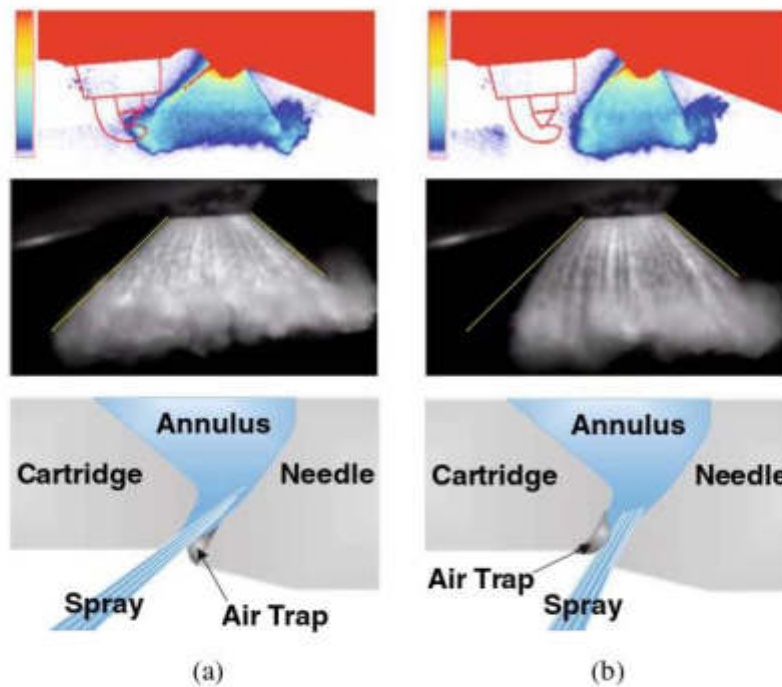
Fig. 3.15 Outward-opening pintle injector spray instability. (a) Wide angle spray and (b) narrow angle spray [57]

In concluding, it has to be stressed that, in addition to the previously identified advantages of the sprays formed by pintle injectors, the piezoelectric actuation of the needle available in all production outward-opening pintle injectors adds significantly to its 'market value'. Despite their much higher cost than multi-hole injectors, the emerging piezoelectric outward-opening pintle injectors allow very rapid opening and closing times and, therefore, short injection durations, multiple injections per engine cycle, and control of needle lift through voltage adjustment (thus the ability to change the fuel mass flow independent of injection duration). Furthermore, there is consensus that these injectors exhibit reduced tendency for fouling and spark-plug wetting. It is, therefore, not surprising that Mercedes-Benz and BMW have introduced into the market the first spray-guided stratified DI gasoline engines for medium-size passenger cars equipped with piezoelectric outward-opening pintle injectors [35, 44, 45].

3.6.2.2 Solenoid-Driven Multi-Hole Injectors

This type of injector is geometrically very familiar to the automotive industry due to its popularity or, better, dominance of the direct-injection diesel engine market. It has many advantages, as already indicated in Table 3.3, which have encouraged its early use in research and development of various spray-guided configurations and

(a)

(b)
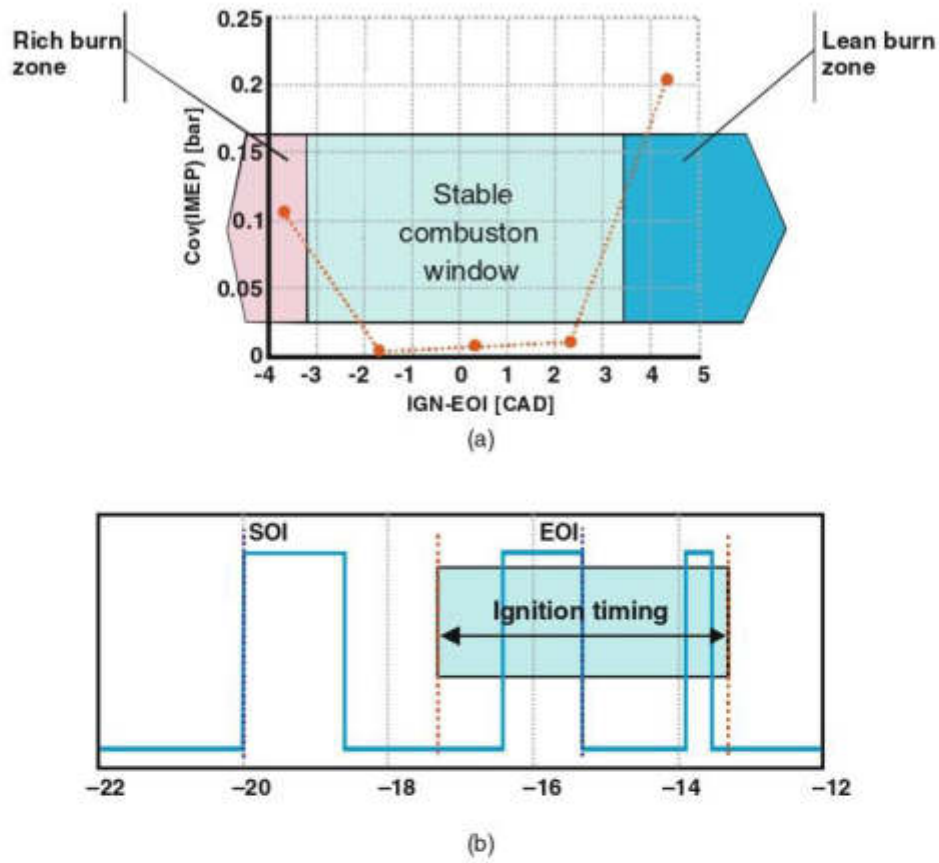
Fig. 3.16 (**a**) Stable combustion window for stratified charge operation. (**b**) Example of injection timing and the stable combustion window (shaded area)

(a)   (b)
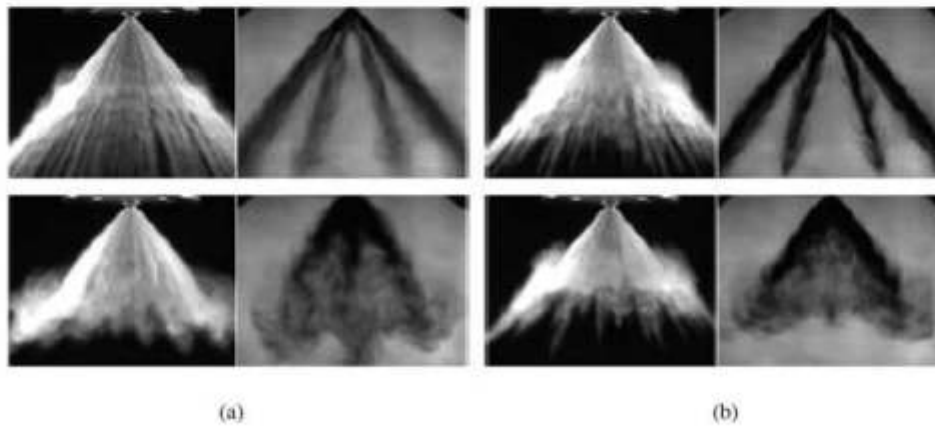
Fig. 3.17 Spray chamber measurements of multi-hole (*black background*) and pintle outward opening nozzle (*white background*) sprays in non-flashing (*top row*) and flashing (*bottom row*) conditions at chamber pressures of (**a**) 0.3 bar and (**b**) 0.8 bar [42]

concepts, but also some disadvantages such as a tendency for injector fouling and spark-plug wetting.

The main advantage of multi-hole injectors has been the enhanced flexibility in the geometry and spatial orientation of the nozzle holes (see Fig. 3.18) which
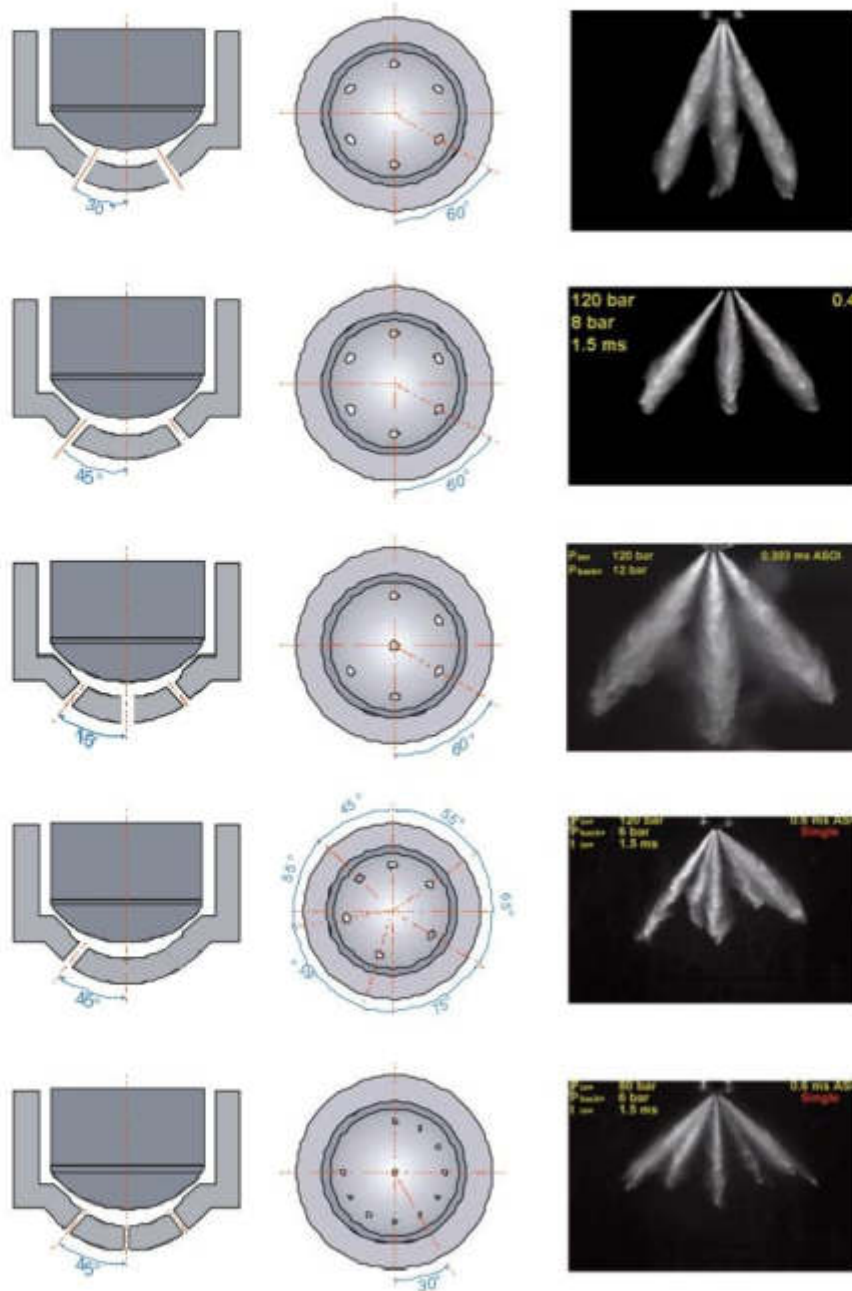
Fig. 3.18 Possible configurations of multi-hole nozzle injectors [47]
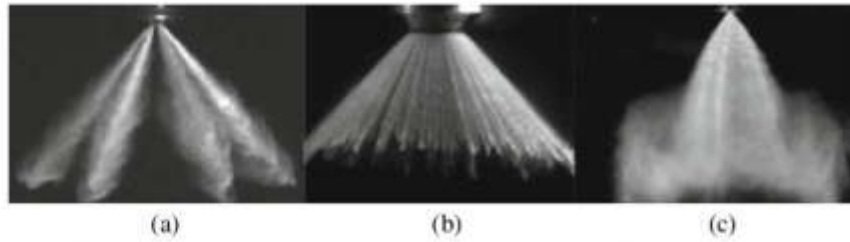
Fig. 3.19 Comparison of sprays generated by (**a**) multi-hole, (**b**) outward opening and (**c**) swirl pressure atomizers

allows a wide variety of hole numbers (6–12), hole sizes, hole pattern (symmetric and asymmetric) and length-to-diameter ratios (L/D). The generated sprays from each of these holes have very different structure than those generated by pintle and swirl-type injectors, as shown in Fig. 3.19; their fishbone structure, which has been for many years the subject of intense investigations by researchers in industry and academia involved with diesel engines, is closely linked not only with the internal nozzle flow but also with the air-entrainment pattern (Fig. 3.20a). Sprays from multi-hole nozzles are characterized in global terms by the overall spray cone angle, the individual jet cone angle and its tip penetration (Fig. 3.20b and also Fig. 3.3). Concerning the local spray characteristics, these are quantified by phase Doppler anemometry (PDA) in terms of droplet velocities and sizes (diameter) visually represented as in Fig. 3.21 by spheres of various size (μm) and vectors of various length (m/s). Under atmospheric conditions and typical injection pressures, sprays from multi-hole injectors [46, 47, 48] can exhibit typical droplet sizes in the range 10–20 μm and droplet velocities well exceeding 100 m/s which leads to steep velocity gradients near the spark plug. According to [49], an increase of L/D, through reduction of the hole diameter, gives rise to smaller droplet sizes while the spray angle increases with decreasing L/D.

As mentioned earlier, multi-hole injectors are considered to be more susceptible to fouling (Fig. 3.22) due to their proximity to the spark plug which facilitates
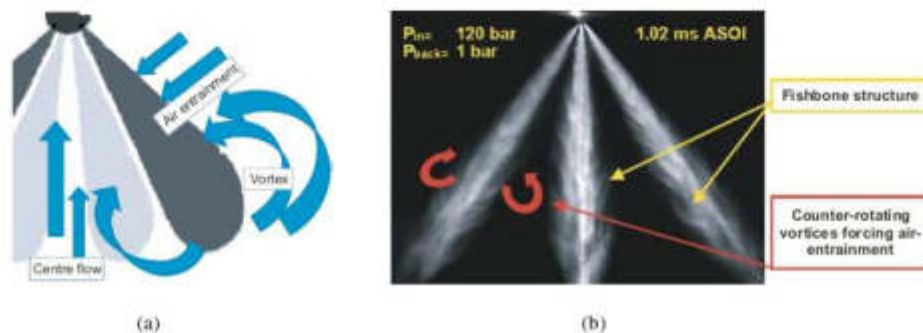

Fig. 3.20 Example of air-entrainment in multi-hole nozzle sprays. (**a**) Schematic representation [42], (**b**) Mie spray image [58]
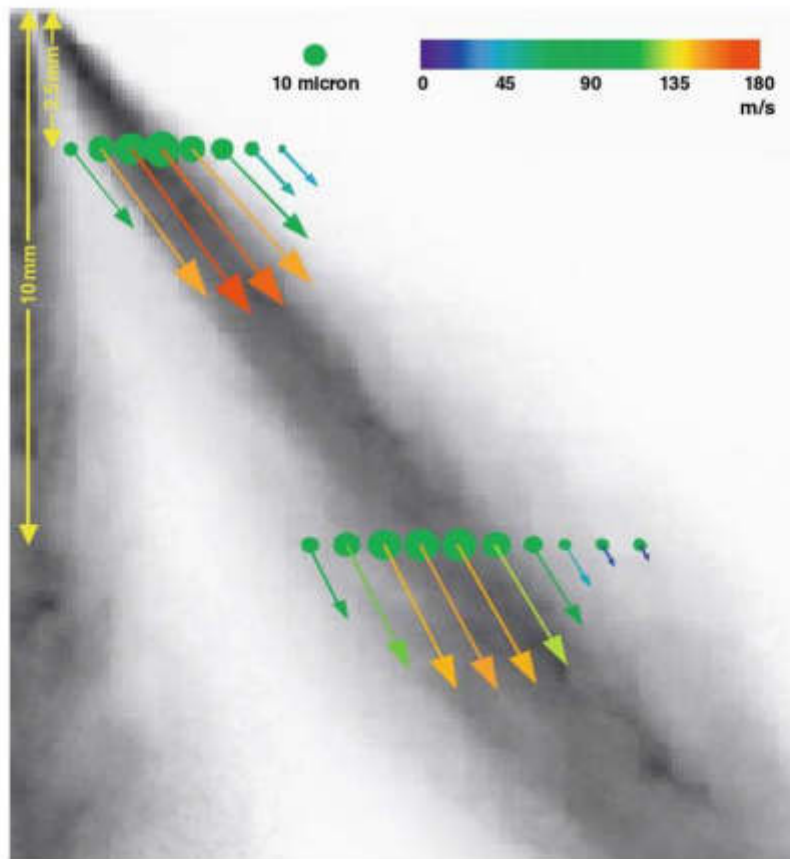
Fig. 3.21 Droplet sizes and velocities from a 6-hole nozzle superimposed on a shadowgraphy spray image [58]
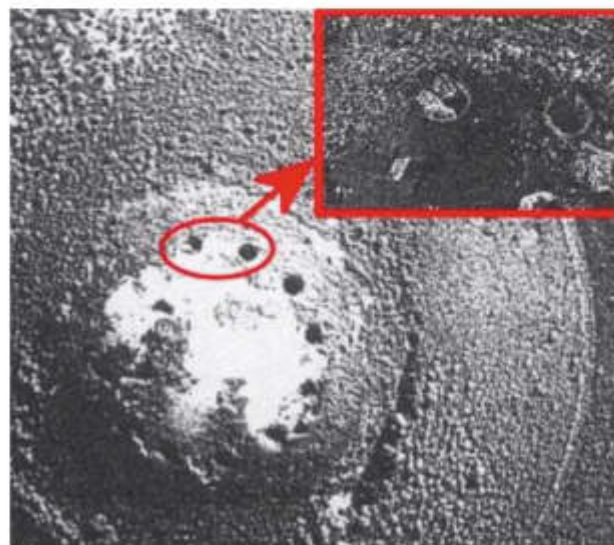

Fig. 3.22 Multi-hole injector nozzle fouling due to carbon deposits

contact of the propagating flame with the external nozzle surface in a high temperature gas environment. At the same time, the need for spark ignition of the mixture in the recirculation zone at the spray edges, in a region of steep velocity gradients, can easily lead to spark electrode wetting. As Fig. 3.23 shows, one of the jets/sprays from the multi-hole injector strikes the ground electrode producing a stream of large droplets; this could be the result of a small variation in the individual spray cone angle leading to poor spray targeting of the spark gap area. As discussed in [50], concerning the same figure, although the orientation of the ground electrode facing the oncoming spray shields the spark plug from the spray plume and reduces the local fuel velocities at the time of spark, the formed wake behind the electrode causes strong fluctuations and degrades ignition stability. In the worst case, misfire takes place which is totally unacceptable in production engines: conditions which are unfavourable to robust ignition and early flame-kernel growth include [33]:

- high gas-and-liquid-phase velocities
- large number of droplets
- steep temporal gradients in air/fuel ratio
- poor design and orientation of spark plug electrodes.

Another factor which has recently attracted attention is nozzle cavitation as a mechanism that contributes to spray instability and can cause significant cone angle fluctuations. Although most of the understanding about the onset and development of cavitation comes from diesel injectors [51, 52, 53], confirmation about the various forms of in-nozzle cavitation has been provided by experiments, supported in some cases by CFD calculations, in real-size and enlarged multi-hole injectors [38, 54]; further evidence was provided by experiments in a real-size, single-hole injector [55]. In particular, there are at least three types of cavitation in the nozzle of multi-hole gasoline injectors:

- geometric cavitation in the holes
- string or vortex-type cavitation in the sac volume
- needle cavitation originating in the vicinity of the needle and extending to the opposite hole.
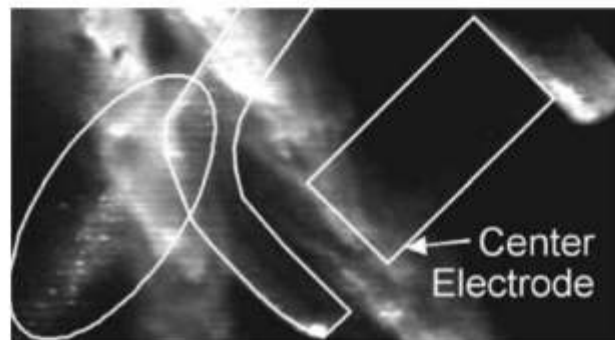

Fig. 3.23 Example of spray/spark plug interaction in a firing direct-injection spark-ignition engine [50]

It is generally accepted that all these cavitation types are affected and controlled by needle lift, radius of curvature of hole inlet, sac volume and hole orientation and geometry. In addition, string cavitation is the most likely cause of spray cone fluctuations but the detailed link with the spray structure is still relatively unknown in multi-hole injectors.

Another thermodynamic effect which can change significantly the fuel distribution at the nozzle exit and the resulting spray structure is flash boiling, a phenomenon occurring when the fuel temperature is high and already discussed in connection with the sprays generated by pintle injectors. As Fig. 3.24 clearly shows, at high fuel temperatures and low ambient pressures, the individual sprays are rapidly atomized to such a degree that it becomes impossible to distinguish the spray plumes; although flash boiling is highly desirable from the atomization point of view, it can only take place within a very narrow window of engine operating conditions and it is not considered important as an atomization mechanism. However, in relative terms, it can be argued that the effect of flash boiling is much more pronounced in the sprays generated by multi-hole rather than pintle injectors.

Overall, it can be argued that the solenoid-driven multi-hole injector is a simpler and less expensive system than piezoelectric outward-opening pintle injectors offering significant advantages in terms of the flexibility for adjusting the spray configuration to the engine geometry, the narrow cone angle of individual sprays, and the control of tip penetration and atomization through injection pressure and timing of injection (back pressure). Although experiments in single and multi-cylinder engines employing multi-hole injectors have been promising, it is widely accepted that further refinement of their design is needed to ensure stable ignition and combustion with acceptable gaseous and particulate emissions. Faster opening
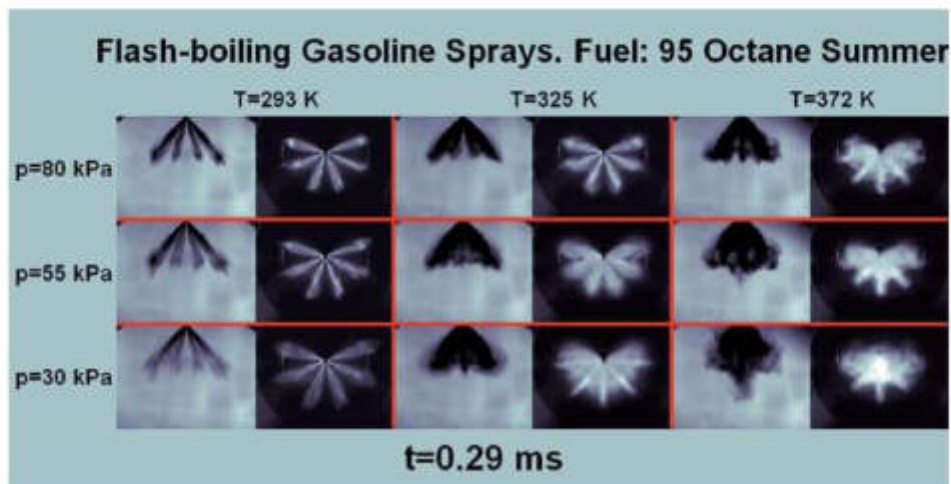

Fig. 3.24 The effect of pressure/temperature on the flash boiling of gasoline sprays from multi-hole nozzle injectors [56]

and closing times combined with much higher than today's injection pressures may make multi-hole injectors directly comparable to the piezoelectric outward-opening pintle-type injectors that, at present, seem to be the preferred fuel injection system.

3.7 Conclusions

Following nearly 10 years of refinement of the first-generation direct-injection gasoline engines, based on the wall-guided concept, the second-generation of stratified DI engines are now entering the mass market equipped with the highly sophisticated piezoelectric outward-opening pintle injectors. In the meantime, stoichiometric DI engines have reached production, taking advantage of the outstanding efficiency of three-way catalysts. All three initiatives highlight in a very consistent way that the concept of direct in-cylinder fuel injection is here to stay due to its unquestionable advantages relative to port fuel injection. Further developments in the technology of NOx catalysts and high pressure multi-hole injection systems are expected to assist the spray-guided combustion mode in establishing itself as the preferred gasoline engine technology concept of the future. Nevertheless, it can be argued that unless the DI gasoline engine is either turbocharged and/or supercharged, it won't achieve its fuel potential and may never reach 'mass production' status at a time of intense competition from advanced diesel and hybrid gasoline/diesel engine systems. It thus remains to be seen whether the many years of research into two-and-four stroke direct-injection gasoline engines will have the desired outcome for the global automotive industry in producing a second-generation supercharged/turbocharged DI gasoline engine which will compete directly with the most advanced diesels in terms of fuel efficiency, performance and exhaust emissions and be superior to the most advanced port-injection engines.

References

1. Bishop, I.N., Simko, A., A New Concept of Stratified Charge Combustion – The Ford Combustion Process (FCP), SAE Transactions, Paper 680041, Vol.77, pp. 93–117, 1968.
2. Mitchell, E., Cobb, J.M., Frost, R.A., Design and Evaluation of a Stratified Charge Multi-fuel Military Engine, SAE Transactions, Paper 680042, Vol.77, pp.118–128, 1968.
3. Alperstein, M., Schafer, G.H., Villforth, F.J., Texaco's Stratified Charge Engine – Multi-fuel, Efficient, Clean, and Practical, SAE Paper 740563, 1974.
4. Urlaub, A.G., Chmela, F.G., High-Speed, Multi-fuel Engine: L9204 FMV, SAE Paper 740122, 1974.
5. Scussel, A.J., Simko, A.O., Wade, W.R., The Ford PROCO Engine Update, SAE Transactions, Paper 780699, Vol. 87, pp. 2706–2725, 1978.
6. Wood, C.D., Unthrottled Open-Chamber Stratified Charge Engines, SAE Paper 780341, 1978.

7. Schapertons, H., Emmenthal, K.-D., Grabe, H.-J., Opperman, W., VW's Gasoline Direct Injection (GDI) Research Engine, SAE Paper 910054, 1991.

8. Iwamoto, Y., Noma, K., Nakayama, O., Yamauchi, T., Ando, H., Development of Gasoline Direct Injection Engine, SAE Transaction, Paper 970541, 1997.

9. Harada, J., Tomita, T., Mizuno, H., Mashiki, Z., Ito, Y., Development of Direct Injection Gasoline Engine, SAE Paper 970540, 1997.

10. Tagaki, Y., Itoh, T., Muranaka, S., Iiyama, A., Iwakiri, Y., Urushihara, T., Naitoh, K., Simultaneous Attainment of Low Fuel Consumption, High Output Power and Low Exhaust Emissions in Direct Injection SI Engines, SAE Paper 980149, 1998.

11. Kanda, M., Baika, T., Kato, S., Iwamuro, M., Koike, M., Sito, A., Application of a New Combustion Concept to direct injection gasoline engine, SAE Paper 2000-01-0531, 2000.

12. Hentschel, W., Block, B., Hovestadt, T., Meyer, H., Ohmstede, G., Richter, V., Stiebels, B., Winkler, A., Optical Diagnostics and CFD-Simulations to Support the Combustion Process Development of the Volkswagen FSI Direct-Injection Gasoline Engine, SAE Paper 2001-01-3648, 2001.

13. Cathcart, G., Railton, D., Improving Robustness of Spray Guided DI Combustion Systems: The Air-Assisted Approach, JSAE Spring Convention 20015360, 2001.

14. Befrui, B., Kneer, R., Breuer, S., Reckers, W., Robart, D., Wanlin, H., Weiten, C., Investigation of a DISI Fuel Injector for a Close-Arranged Spray-Guided Combustion System, SAE Paper 2002-01-1133, 2002.

15. Ortmann, R., Arndt, S., Raimann, J., Wuerfel, G., Grzeszik, R., Methods and Analysis of Fuel Injection, Mixture Preparation and Charge Stratification in Different Direct-Injection SI Engines, SAE Paper 2001-01-0970, 2001.

16. Achleitner, E., Berger, S., Frenzel, H., Klepatsch, M., Warnecke, V., Gasoline Direct Injection System with Piezo Injectors for Spray-Guided Combustion Processes, MTZ worldwide 5/2004, Vol. 65, p. 2, 2004.

17. Green Car Congress Mercedes-Benz Premiers New Gasoline Direct Injection System for More Power and Lower Fuel Consumption, http://www.greencarcongress.com/2006/02/mercedesbenz_pr.html.

18. Ando. H., Takemura, J., Koujina, E., A Knock Anticipating Strategy Basing on the Real-Time Combustion Mode Analysis, SAE Paper.890882, 1989.

19. Kuwahara, K., Ueda, K., Ando, H., Mixing Control Strategy for Engine Performance Improvement in a Gasoline Direct Injection Engine, SAE Paper 980158, 1998.

20. Ueda, K., Kaiahara, K., Krose, K., Ando, H., Idling Stop System Coupled With Quick Start Features of Gasoline Direct Injection, SAE Paper 2001-01-0545, 2001.

21. Hori, M., Taniguch, S., Noda, N., Abe, F., Iwachido, K., Tanada, H., Watanabe, T., Yamada, N., Ando, H., Development of the NOx Adsorber Catalyst for Use with High-Temperature Condition, SAE Paper 2001-01-1298, 2001.

22. Tamura, Y., Kikuchi, S., Okada, K., Koga, K., Dogahara, T., Nakayama, O., Ando, H., Development of Advanced Emission-Control Technologies for Gasoline Direct-Injection Engines, SAE Paper 2001-01-0254, 2001.

23. Kojima, S., Jimbo, T., Katoh, K., Miyashita, S., Watanabe, M., Analysis and Simplification of Thermal Endurance Tests of NOx Storage-Reduction Catalysts, SAE Paper 2004-01-1496, 2004.

24. Iizuka, H., Kaneeda, M., Higashiyama, K., Kuroda, O., Shinotsuka, N., Watanabe, H., Improvement of Heat Resistance for Lean NOx Catalyst, SAE Paper 2004-01-1495, 2004.

25. Iwamoto, S., Kouno, Y., Inoue, M., Direct Decomposition of NO by Ba-Loaded Catalyst", 90th CATSJ A-Page 239, 2002.

26. Hamada, H., Evolution of NOx removal catalyst 90th CATSJ A-Page 244, 2002.

27. Ishihara, T., Ando, M., Sada, K., Takiishi, K., Nishiguchi, N., Takita, Y., Direct Decomposition of NO into N_2 and O_2 over La(Ba)Mn(In)O_3 Perovskite Oxide, J. Catal., 220, 104–114, 2003.

28. Noma, K., Iwamoto, Y., Murakami, N., Iida, N., Nakayama, O., Ando, H., Optimized Gasoline Direct-Injection Engine for the European Market, SAE Paper 980150, 1998.

29. Yamamoto, S., Tanaka, D., Takemura, J., Ando, H., Mixing Control and Combustion in Gasoline Direct Injection Engines for Reducing Cold-Start Emissions, SAE Paper 2001-01-0550, 2001.

30. Sadakane, S., Sugiyama, M., Kishi, H., Abe, S. Harada, J., Sonoda, Y., Development of a New V-6 High Performance Stoichiometric Direct Injection Gasoline Engine, SAE Paper 2005-01-1152, 2005.

31. Nishida, M., Isobe, R., Goto, T., Hanzawa, H., Aiga, S., The New 2.3L Direct Injection Turbo Gasoline Engine from Mazda, 14th Aachen Automotive Colloquium, Aachen, pp. 939–960, 2005.

32. Middendorf, H., Krebs, R., Szengel, R., Pott, E., Fleis, M., Hagelstein, D., Volkswagen Introduces the Worlds First Double Charge Air Direct Injection Petrol Engine, 14th Aachen Automotive Colloquium, Aachen, pp. 961–986, 2005.

33. Drake, M.C., Haworth, D.C., Advanced Gasoline Engine Development using Optical Diagnostics and Numerical Modeling, Invited Plenary Lecture, 31st International Symposium on Combustion, Combustion Institution, Heidelberg, Germany, August 6–11, 2006.

34. Ladenfield, T., Kufferath, A., Gerhardt, J., Gasoline Direct Injection – SULEV Emission Concept, SAE Paper 2004-01-0041, 2004.

35. Altenschmidt, F., Bertsch, D., Bezner, M., Laudenbach, N., Zahn, M., Schaupp, U., Kaden, A., The Analysis of the Ignition Process on SI-Engines with Direct Injection in Stratified Mode, International Symposium on Internal Combustion Diagnostics, Baden-Baden, Germany, pp. 395–411, 2006.

36. Stach, T., Schlerfer, J., Vorbach, M., New Generation Multi-Hole Full Injector for Direct-Injection SI Engines – Optimization of Spray Characteristics by Means of Adapted Injector Layout and Multiple Injection", SAE Paper 2007-01-1404, 2007.

37. Marchi, A., Nouri, J.M., Yan, Y., Arcoumanis, C., Internal Flow and Spray Characteristics of Pintle-type Outwards Opening Piezo Injectors for Gasoline Direct-Injection Engines, SAE Paper 2007-01-1406, 2007.

38. Papoulias, D., Giannadakis, E., Mitroglou, N., Gavaises, M., Cavitation in Fuel Injection Systems for Spray-Guided Direct Injection Gasoline Engines, SAE Paper 2007-01-1418, 2007.

39. Nouri, J.M., Hamid, M.A., Abo-Serie, E., Marchi, A., Mitroglou, N., Arcoumanis, C., Internal and Near Nozzle Flow Characteristics in an Enlarged Model of an Outwards Opening Pintle-Type Gasoline Injector, Proceedings of Third International Conference on Optical and Laser Diagnostics (ICOLAD 2007), City University, London, 22–25 May 2007, 2007.

40. Nouri, J.M., Hamid, M.A., Yan, Y., Arcoumanis, C., Spray Characterization of a Piezo Pintle-Type Injector for Gasoline Direct Injection Engines, Proceedings of Third International Conference on Optical and Laser Diagnostics (ICOLAD 2007), City University, London, 22–25 May 2007, 2007.

41. Skogsberg, M., Dahlander, P., Denbratt, I., Spray Shape and Atomization Quality of an Outward-Opening Piezo Gasoline DI Injector, SAE Paper 2007-01-1409, 2007.

42. Skogsberg, M., A Study on Spray-Guided Stratified Charge Systems for Gasoline DI Engines, PhD Thesis, Department of Applied Mechanics, Chalmers University of Technology, Gothenburg, Sweden, 2007.

43. Dahlander, P., Lindgren, R., Denbratt, I., High-Speed Photography and Phase Doppler Anemometry Measurements of Flash-Boiling Multihole Injector Sprays for Spray-Guided Gasoline Direct Injection, Paper ICLASS06-0112, ICLASS, Kyoto, Japan, 2006.

44. Schwarz, C., Schünemann, E., Durst, D., Fischer, J., Witt, A., Potentials of the Spray-Guided BMW DI Combustion System, SAE Paper 2006-01-1265, 2006.

45. Fischer, J., Kern, W., Unterweger, G., Witt, A., Durst, B., Schünemann, E., Schwarz, C., Methods for the Development of the Spray-Guided BMW DI Combustion System, International Symposium on Internal Combustion Diagnostics, Baden-Baden, pp. 413–423, 2006.

46. Yan, Y., Gashi, S., Nouri, J.M., Lockett, R.D., Arcoumanis, C., Investigation of Spray Characteristics in a Spray-Guided DISI Engine Using PLIF and LDV, Proceedings of Third International Conference on Optical and Laser Diagnostics (ICOLAD 2007), City University, London, 22–25 May 2007, 2007.

47. Mitroglou, N., Nouri, J.M., Yan, Y., Gavaises, M., Arcoumanis, C., Spray Structure Generated by Multi-Hole Injectors for Gasoline Direct-Injection Engines, SAE Technical Paper Series, 2007-01-1417, 2007.

48. Mitroglou, N., Nouri, J.M., Gavaises, M., Arcoumanis, C., Spray Characteristics of a Multi-Hole Injector for Direct-Injection Gasoline Engines, Int. J. Eng. Res., 7, No.3, 255–270, 2006.

49. Skogsberg, M., et al., Effects of Injector Parameters on Mixture Formation for Multi-Hole Nozzles in a Spray-Guided Gasoline DI Engine, SAE Technical Papers, 2005-01-0097, 2005.

50. Fansler, T.D., Drake, M.C., Düwel, I., Zimmerman, F.P., Fuel-Spray and Spark-Plug Interactions in a Spray-Guided Direct-Injection Gasoline Engine, International Symposium on Internal Combustion Diagnostics, Baden-Baden, pp. 81–97, 2006.

51. Arcoumanis, C., Gavaises, M., Nouri, J.M., Abdul-Wahab, E. Horrocks, R., Analysis of the Flow in the Nozzle of a Vertical Multi-Hole Diesel Engine Injector, SAE Paper 980811, 1998.

52. Arcoumanis, C., Badani, M., Flora, H. Gavaises, M., Cavitation in Real-Size Multi-Hole Diesel Injector Nozzles, SAE Paper 2000-01-1249, 2000.

53. Roth, H., Gavaises, M., Arcoumanis, C., Cavitation Initiative, its Development and Link with Flow Turbulence in Diesel Injector Nozzles, SAE Paper 2002-01-0214, 2002.

54. Nouri, J.M., Mitroglou, N., Yan, Y., Arcoumanis, C., Internal Flow and Cavitation in a Multi-Hole Injector for Gasoline Direct-Injection Engines, SAE Technical Paper Series, 2007-01-1405, 2007.

55. Birth, I.G., Recks, M., Spicher, U., Experimental Investigation of the In-Nozzle Flow of Valve Covered Orifice Nozzle for Gasoline Direct Injection Engines, International Symposium Internal Combustion Diagnostics, Baden-Baden, pp. 59–78, 2006.

56. Dahlander, P., Annual Report, Combustion Engine Research Centre, Chalmers University of Technology, Gothenburg, 2006.

57. Marchi, A., Internal Flow and Spray Characteristics of the Pintle Type Piezo Injector, PhD Thesis in preparation, School of Engineering and Mathematical Sciences, City University, London, 2008.

58. Mitroglou, N., Multihole Injectors for Direct-Injection Gasoline Engines, PhD Thesis, School of Engineering and Mathematical Sciences, City University, London, 2006.

[4]   Denton, T., *Advanced Automotive Fault Diagnostics*, 4ta ed., Institute of the Motor Industry, New York: Routledge, 2017.
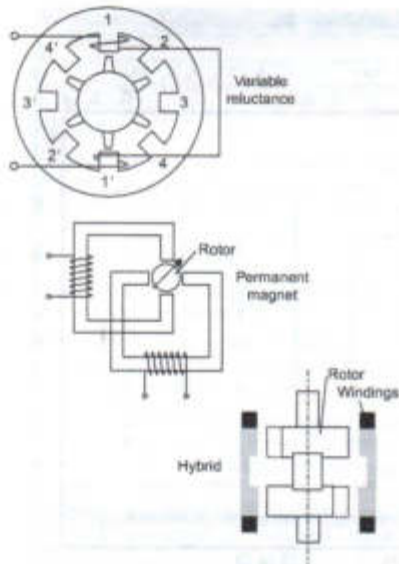
Figure 4.42 Stepper motor principle

This is mainly because of the ease with which they can be controlled by electronic systems. Stepper motors fall into the following three distinct groups, the basic principles of which are shown in Figure 4.42:

▶ variable reluctance motors;
▶ permanent magnet (PM) motors;
▶ hybrid motors.

The underlying principle is the same for each type. All of them have been and are being used in various vehicle applications. The basic design for a permanent magnet stepper motor comprises two double stators. The rotor is often made of barium-ferrite in the form of a sintered annular magnet. As the windings are energised in one direction then the other, the motor will rotate in 90° steps. Half step can be achieved by switching on two windings. This will cause the rotor to line up with the two stator poles and implement a half step of 45°. The direction of rotation is determined by the order in which the windings are switched on or off or reversed. The main advantages of a stepper motor are that feedback of position is not required. This is because the motor can be indexed to a known starting point and then a calculated number of steps will move the motor to any suitable position.

Figure 4.43 Stepper motor and throttle potentiometer on a throttle body

The stepper motor, when used to control idle speed, is a small electro-mechanical device that allows either an air bypass circuit or a throttle opening to alter in position depending on the amounts that the stepper is indexed (moved in known steps) (Figure 4.43).

Stepper motors are used to control the idle speed when an ISCV is not employed. The stepper may have four or five connections back to the ECU. These enable the control unit to move the motor in a series of 'steps' as the circuits are earthed to ground. These devices may also be used to control the position of control flaps, for example, as part of a heating and ventilation system.

The individual earth paths can be checked using the oscilloscope. The waveforms should be similar on each path. Variations to the example shown here may be seen between different systems (Figure 4.44 and Figure 4.45).

4.3.4 Solenoid actuators

The basic operation of solenoid actuators is very simple. The term 'solenoid' actually means 'many coils of wire wound onto a hollow tube'. This is often misused but has become so entrenched that terms like 'starter solenoid', when really it is a starter actuator or relay, are in common use. A good example of a solenoid actuator is a fuel injector.

When the windings are energised, the armature is attracted due to magnetism and compresses the spring. In the case of a fuel injector, the movement is restricted to approximately 0.1 mm. The period that an injector remains open is very small; under various operating conditions, between 1.5 and 10 ms being

87
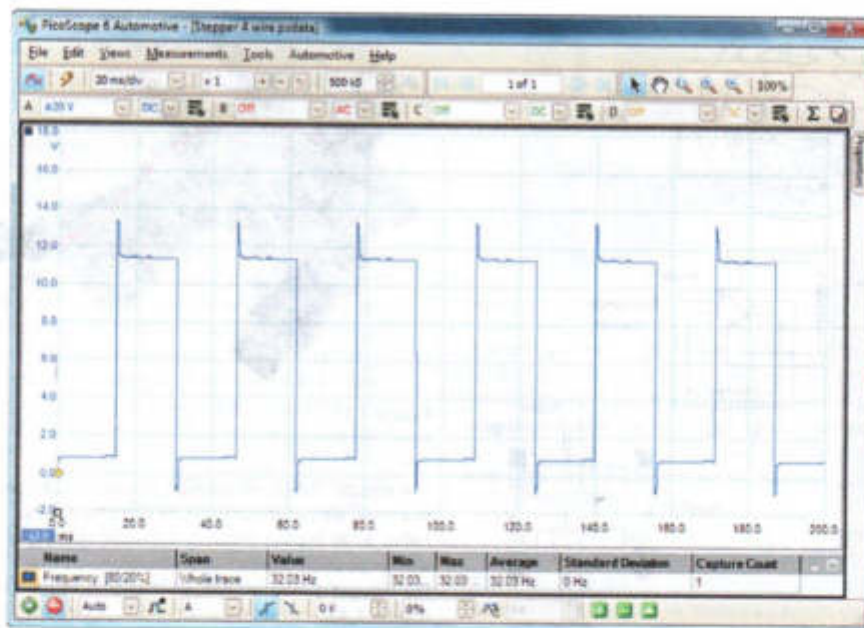
Figure 4.44 Stepper motor signals
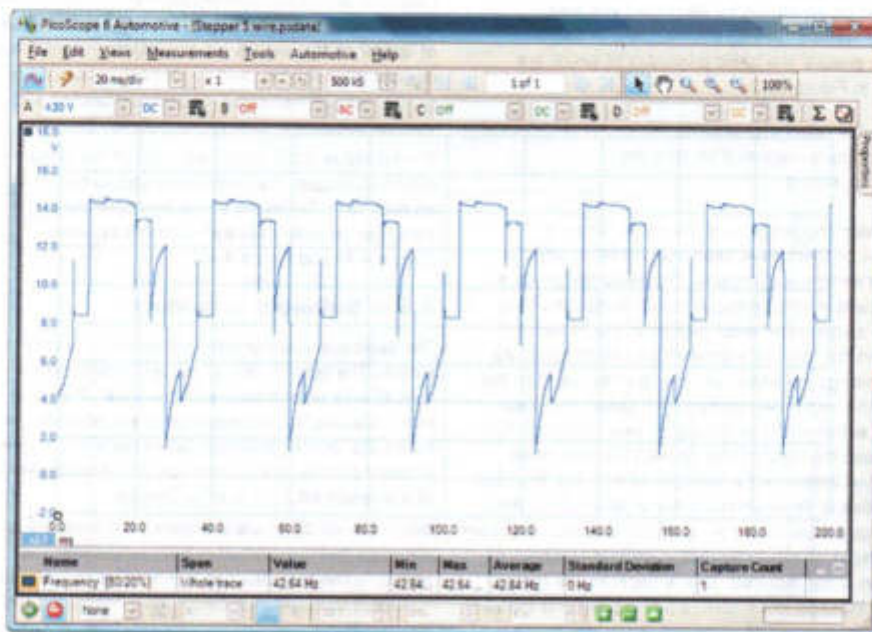
Figure 4.45 Alternative stepper motor signal

88

typical. The time it takes an injector to open and close is also critical for accurate fuel metering. Some systems use ballast resistors in series with the fuel injectors. This allows lower inductance and resistance operating windings to be used, thus speeding up reaction time.

Other types of solenoid actuators, for example door lock actuators, have less critical reaction times. However, the basic principle remains the same.

4.3.4.1 Single-point injector

Single-point injection is also sometimes referred to as throttle body injection (Figure 4.46).

A single injector is used (on larger engines two injectors can be used) in what may have the outward appearance to be a carburettor housing.

The resultant waveform from the single-point system shows an initial injection period followed by voltage-pulsing of the injector in the remainder of the trace. This 'current limiting' section of the waveform is called the supplementary duration and is the part of the injection trace that expands to increase fuel quantity.

Figure 4.46 Throttle body with a single injector

This shows better in a current rather than voltage waveform (Figures 4.47 and 4.48).

4.3.4.2 Multi-point injector

This injector is an electro-mechanical device which is fed by a 12 V supply. The voltage will only be present

4

Figure 4.47 Single-point injector voltage waveform

Figure 4.48 Single-point injector current waveform

when the engine is cranking or running because it is controlled by a relay that operates only when a speed signal is availabl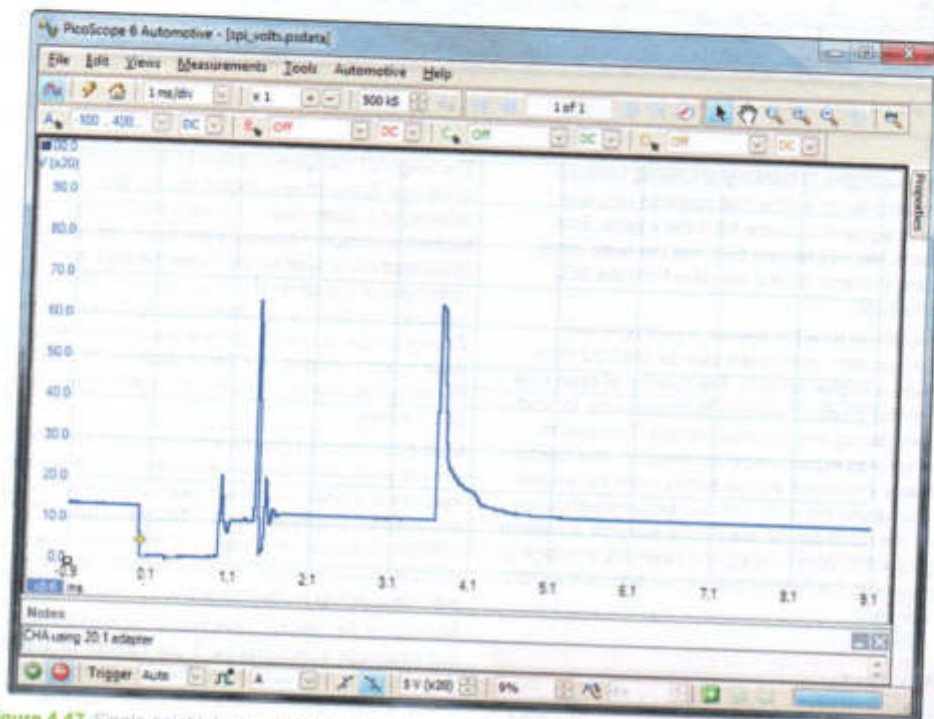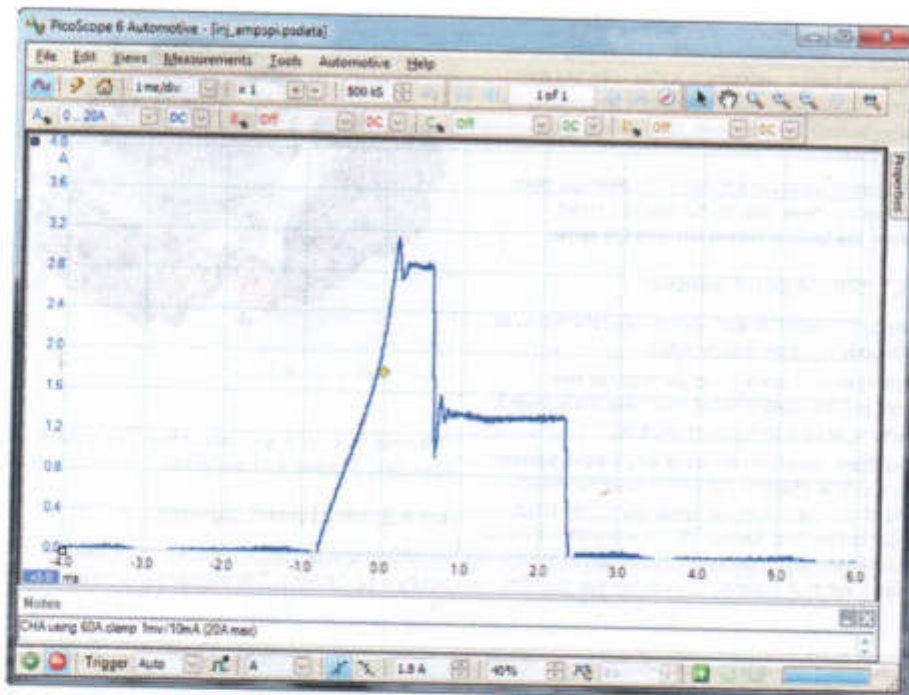e from the engine. Early systems had this feature built into the relay; most modern systems control the relay from the ECU (Figure 4.49).

The length of time the injector is held open will depend on the input signals seen by the ECU from its various engine sensors. The duration of open time or 'injector duration' will vary to compensate for cold engine starting and warm-up periods. The duration time will also expand under acceleration. The injector will have a constant voltage supply while the engine is running and the earth path will be switched via the ECU; the result can be seen in the example waveform (Figure 4.50). When the earth is removed, a voltage is induced into the injector and a spike approaching 60V is recorded.

Key fact

The length of time an injector is held open depends on the sensor input signals to the ECU.

The height of the spike will vary from vehicle to vehicle. If the value is approximately 35V, it is because a Zener diode is used in the ECU to clamp the voltage. Make sure the top of the spike is squared off, indicating the Zener dumped the remainder of the spike. If it is not squared, this indicates the spike is not strong enough to make the Zener fully dump, meaning there is a problem with a weak injector winding. If a Zener diode is not used in the computer, the spike from a good injector will be 60V or more.

Multi-point injection may be either sequential or simultaneous. A simultaneous system will fire all four injectors at the same time with each cylinder receiving two injection pulses per cycle (720° crankshaft rotation). A sequential system will receive just one injection pulse per cycle, which is timed to coincide with the opening of the inlet valve.

Monitoring the injector waveform using both voltage and amperage allows display of the 'correct' time that the injector is physically open. The current waveform (the one starting on the zero line) shows that the waveform is 'split' into two defined areas.

90

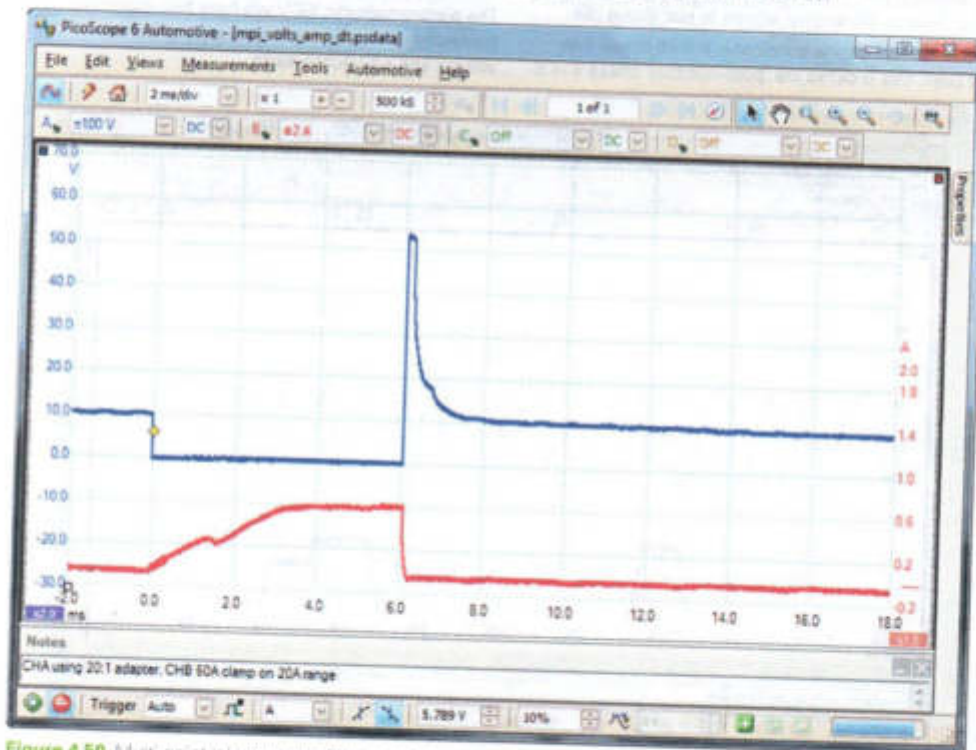Figure 4.49 Multi-point injectors on the rail. Also shown are the pressure regulator and sensor

4

Figure 4.50 Multi-point injector waveform, red shows the current and blue the voltage signal

[5]   Parotto, M., Sgatti, S., Sensi, F., "Advanced GDI Injector Control with Extended Dynamic Range," SAE International, doi: 10.4271/2013-01-0258, 2013.

SAE International®

Advanced GDI Injector Control with Extended Dynamic Range	2013-01-0258 Published 04/08/2013
Marco Parotto, Stefano Sgatti and Fabio Sensi Magneti Marelli SpA Powertrain Division	

ABSTRACT

Considering the world-wide market for GDI engines, the introduction of tighter polluting emission legislation, additional costs, vehicle fuel economy and pollutants reduction become substantial drivers. Focusing on particulate reduction, direct injection gasoline engines require advanced combustion strategies. The main levers used are injection splitting in order to reduce wall impingement (due to lower penetration) and higher rail pressure level to reduce droplet size. To reach this target it is necessary to improve precision in term of injected quantity in the small quantity region with high fuel rail pressure and during the actuation of multiple injections. As a consequence of the requirements of high quantities at full load, known GDI solenoid injectors show an unacceptable spread in terms of small injected quantity when the energizing time is small such that the injector works in the ballistic zone. Following these premises an electronic approach is needed. New hardware and software solutions should be integrated into the ECU in order to match the new injectors performance requirements. This paper presents two different strategies developed by Magneti Marelli: the first one is based on rail pressure measurement, the second one is based on injector electric signal analysis. Results in terms of correction of injected quantity using Magneti Marelli solenoid injectors are presented. The developed strategies demonstrate the possibility to work with multiple small injections at high pressure without changing injector hardware.

INTRODUCTION

As a consequence of the introduction of EU6 emission legislation a particulate number (PN) limit will be defined; the main variables influencing the formation of such pollutant are combustion and fuel metering. This paper will focus on the second one, that is fuel metering. At first we must consider using multiple injections: splitting the total amount of fuel requested by each stroke allows smaller (and shorter in time) single injections; this leads to a lower jet penetration into the combustion chamber [1] thus reducing wall wetting and decreasing particulate formation. Parameters like droplet size, spray targeting, injection pressure and injectors holes geometry help the reduction of pollutants in addition to the accuracy of the small injections focused on this paper. However the management of small injections forces the use of the injectors in their ballistic mode. A new mechatronic approach is needed in order to fulfill the new performances requested in this operating range and to overcome the limitations deriving from tolerances of mass production injectors. According to these considerations Magneti Marelli is currently developing solutions to control injectors in ballistic area with increased precision. The challenge for new systems is the availability of feedback information to allow observation and control of the injected quantity.

Magneti Marelli is now evaluating two possibilities. The first one considers the elaboration of injectors control signal, read by a specific hardware with a specific software on-board to detect needle closing time; electromagnetic actuator energizing time is then adjusted in order to exactly achieve the correct injection target. The second one evaluates the signal coming from the pressure sensor, and after a specific calculation the fuel quantity is modified to exactly achieve the correct injection target. Magneti Marelli aims to enable operations range down to 0.5-1 mg with minimal deviations. Magneti Marelli development of both strategies is presented in this paper and the results of injectors fuel metering before and after electronic corrections are compared.

MAGNETI MARELLI DRE

In the range of small injections the injector needle does not reach its full lift.
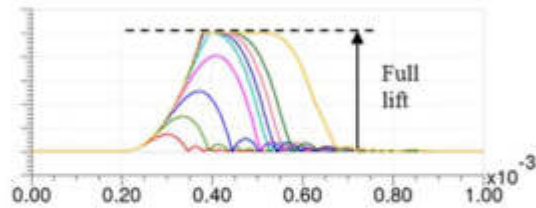

Figure 1. Simulation of needle lift behavior at small pulse-width

When injection times are very short the needle motion follows a parabolic trajectory without reaching the upper position; this is usually known as the ballistic range. Figure 1 shows a simulation of the needle movement using different electric commands; it is clearly evident that ballistic motion corresponds to the smaller energizing times. Many key factors such as spring forces, electro-magnetic forces, needle strokes, friction variations and squish are known as needle lift dynamic modifiers due to manufacturing tolerances and ageing effects on solenoid injectors. Up to now this range has not been used, because of lower requirements and most of all because of components tolerances. The availability of the ballistic area of operations will open the way to further advances in combustion concept and particularly in the emission optimization field. Magneti Marelli approach is to develop different algorithms to exactly control the injected fuel quantity during ballistic operations also. These models simulate components aging and the effects of modifications in environmental conditions as system pressure and fuel temperature.

The starting point was the analysis of different alternatives in term of electronic control unit (ECU) strategies, with the intent that guarantee small injections with low tolerance. At the beginning it was not possible to predict which was the best strategy and most of all which strategy would be able to work correctly on the real engine. Two separate projects were started following two different physical approaches, principally for the learning phase:

1. "Pressure DRE": measure of the rail pressure with specific methods and elaboration of the data to find out the whole injections characteristic. This function is called **P-DRE**

2. "Closing time DRE": measure of the electric signals interchanged between ECU and injectors and elaboration of the data to find out the closing time of each injector as a function of fuel pressure and command pulse width. This function will be called **Tc-DRE**

Both approaches allow the ECU to manage the small injections corrections. Each of the strategies presented in this paper is split in two parallel phases:

1. Learning phase. During this phase the algorithm measures and analyzes the data and then maps them in its memory.

2. Correction phase. Using data collected during learning phase this algorithm can correct the pulse-width of each injector command to exactly achieve the injected quantity requested by the engine.

DRE (Dynamic Range Extension) is the name given to Magneti Marelli solutions.

1. P - DRE

P- DRE Description

The physical concept behind this strategy is the following: in a closed system there is a direct correlation between the fuel quantity that leaves the system and the reduction of fuel pressure in the same system. Typical GDI systems show this characteristic, being composed by an high pressure pump, high pressure piping, a rail and a defined number of GDI injectors; fuel leakage is generally negligible and GDI injectors, unlike diesel injectors do not have any internal leakage or back flow. The high pressure pump, equipped with demand controlled electro-mechanic valve, is a mandatory component; this feature, together with injections control, allows the ECU to seal the hydraulic circuit.

The main challenge is the possibility to have precise measures of rail pressure using the standard automotive fuel pressure sensor. Two different software algorithms were developed in order to measure the injected fuel and reach the highest possible precision allowed by the available hardware. Both solutions show advantages and disadvantages, see table 1.

Table 1. Advantages and disadvantages of P-DRE algorithm developed

	Advantages	Disadvantages
Algorithm #1	Fast learning	One shot learning
Algorithm #2	Continuous learning	Slow learning

The two algorithms can work together thus enhancing the advantages and minimizing the disadvantages.

P- DRE Observations

Pressure pulsations inside fuel rail are taken into account for both algorithms through the introduction of specific models.

P-DRE strategy assumes injectors shot to shot behavior to be repeatable with small deviations. It is important to underline

108

the independency of this solutions from the injectors (but very low fuel leakage is a mandatory characteristic).

P - DRE Methodology

Both algorithms were developed starting from simulations; after this initial phase, a simplified laboratory control circuitry was built; encouraging results followed long sessions of laboratory tests and finally a dedicated software algorithm was embedded in a Magneti Marelli electronic engine control unit reaching the performances described in the next paragraphs.

P- DRE Hardware

One of the major points of interest of the pressure strategies is that they don't need any special hardware. It means that, for each of the performed tests, a Magneti Marelli standard GDI ECU and a standard automotive fuel pressure sensor were used. The strategy measures the pressure drop, inside the rail, produced by single or multiple fuel injections. The ECU samples the signal from the pressure sensor to know the fuel rail pressure value. Fuel pump and injectors commands follow each a specific pattern and the measured pressure drop is automatically converted into an injected quantity value: this is stored into maps indexed by corresponding pulse-width and injections number. Separate maps are available for each injector. The pressure sensor choice needs a further explanation: even though standard sensors can be used, it is also true that the measurement precision is proportional to the full scale of the sensor. The concept to keep in mind is that increasing GDI systems fuel pressure the precision of the measure may decrease. This paper presents results found using a Magneti Marelli GDI ECU and a standard automotive pressure sensor with a full scale of 260 bar.
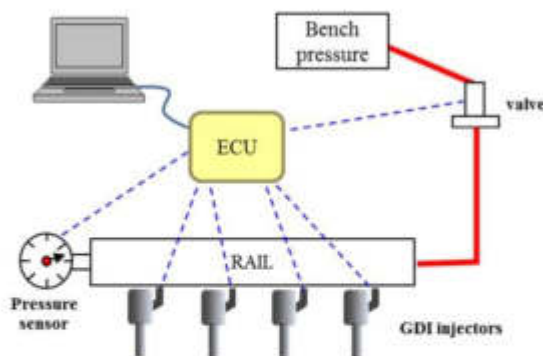

Figure 2. Laboratory hydraulic layout

Figure 2 shows the laboratory hydraulic layout used for test execution. Every injected fuel quantity was measured using an analytical balance. The test bench contains: four Magneti Marelli GDI injectors from IHP3 family, a fuel rail, a fuel pressure sensor, a pipeline to feed the rail, a pressure regulation valve and an ECU. The electronic control unit receives the pressure signal and sends commands to the valve and to the four injectors. The fuel rail pressure target can be set to a fixed value between 0 and 250 bar. The inlet fluid temperature is kept constant. Every software variable is monitored by a connected laptop. When calculating the fuel injected quantity from measured delta pressure, fuel temperature may have a relevant influence. The calculation physically depends on the volume of the hydraulic circuit and on the bulk module. This last physical entity, that is the bulk module, depends on the fuel temperature. In P-DRE strategies the transformation (delta pressure to fuel quantity) and the temperature relationship are taken into account with specific models.

P- DRE Laboratory Results

Many tests were performed to calibrate the algorithms and to analyze the system, the results are presented in this paper: they are related to the four GDI IHP3 injectors, that are representative samples of production spread. Using an analytical balance, and scanning the entire pulse width region for each injector, a characteristic curve was found precisely; to increase the precision of the measure a large number of injections for each point was repeated. The focus was on the average behavior, because this strategies were not thought to correct the shot to shot dispersion. The whole characteristic curve, measured at a controlled level of rail pressure, is shown on figure 3.
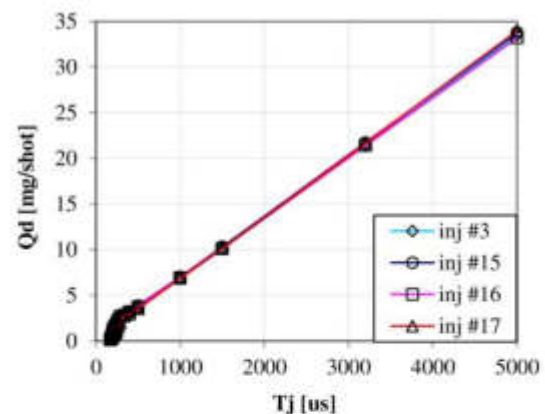
Figure 3. Selected injectors characteristic curve

Figure 4 shows a detail of the injector characteristic (figure 3). Inside linear operating range (from 4mg up to over 30mg) the fuel quantity dispersion is less than 10%, same as solenoid injectors from other manufactures [2]. Under 4mg the relationship between pulse duration of injector command and fuel injected quantity becomes non-linear. It is generally possible to find different behaviors recognizable in three different areas.
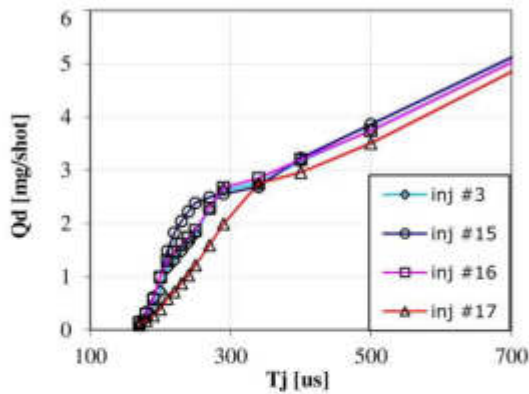
Figure 4. Injector characteristic curve detailed on small quantities

In the first one, below 2mg, the valve needle moves in ballistic mode; then, in the second area, from 2 mg to 3 mg, there is a transition zone where relationship can invert and so fuel quantity may even decrease when pulse-width increases. The third area, over 3 mg, is completely linear. It is possible to analyze the dispersion among injectors belonging to the same family: the spread of smallest fuel injected quantities may be over 50% as illustrated in figure 5. The precision of injections of small fuel quantities in ballistic zone must be increased in order to have the capability to manage multiple fuel injections at low engine load.
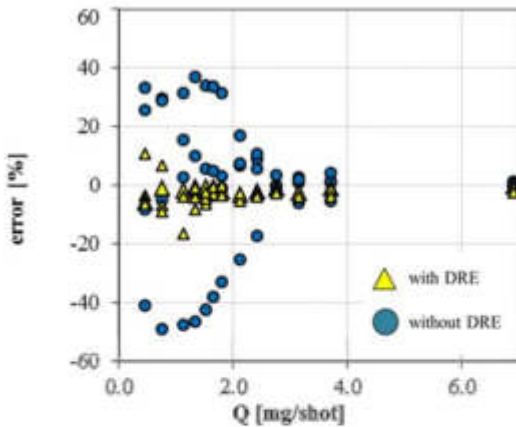
Figure 5. Dispersion of 4 injectors with and without DRE correction using algorithm #1

When the strategy runs the ECU uses the learning phase to build the injectors characteristic curve from few selected points. Data are collected into dedicated maps (one map for each injector), then correction phase may start. The test of correction data shows that errors are well below 20% for both algorithms, see figure 5 for algorithm #1 and see figure 6 for algorithm #2. These two figures show the dispersion with and

without the correction with a focus on the injection of small fuel quantities.
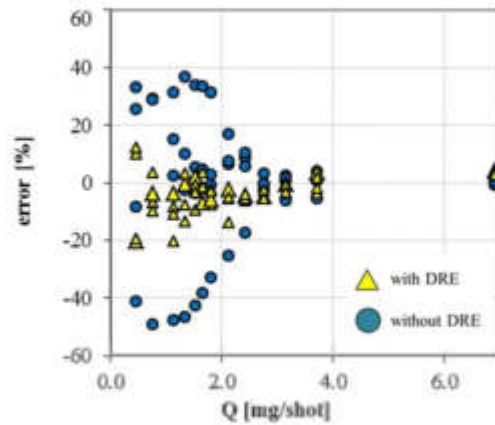
Figure 6. Dispersion of 4 injectors with and without DRE correction using algorithm #2

Even the spread of the ballistic range is very small compared with the uncorrected value of the original injectors.

The target of Magneti Marelli P-DRE strategy is to reach a maximum dispersion of 20% when injected fuel quantity goes down to 0.5mg.

P- DRE Vehicle Results

The experience and the practice developed and carried out in the hydraulic laboratory for many months were transferred on the final environment of the DRE strategies that is in a vehicle with a GDI engine. Having tried and tested the algorithms in their ideal workplace they were moved to the real one. Working on the vehicle it was necessary the check that the new boundary conditions could not affect negatively the functionality of the software. This transfer phase was implemented for both pressure based DRE strategies.
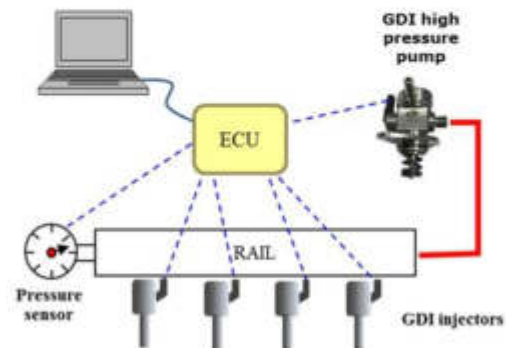
Figure 7. Vehicle hydraulic layout

The test was performed on a roller test bench; the hydraulic layout is shown in figure 7: the internal combustion engine is equipped with a high pressure fuel pump and four Magneti Marelli GDI injectors from IHP1 family. In this case it is possible to evaluate the results only in an indirect way due to the impossibility to measure the fuel injected quantity of each injector directly on the engine. During the learning phase the system works with constant fuel rail pressure and the algorithm inside the ECU collects the information. The injected fuel quantities corresponding to the electrical commands are measured and their values are stored into the permanent memory of the ECU. Once the learning phase is finished the system uses the collected data to correct the injected quantities for the engine working point in use.
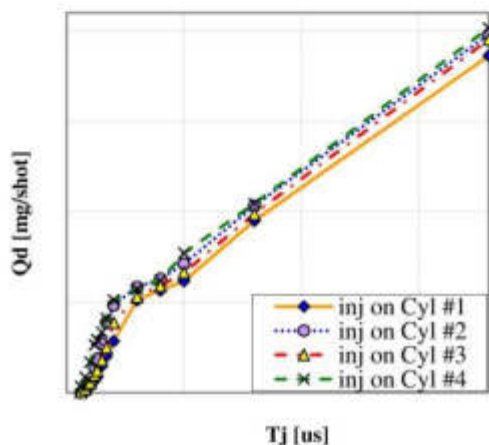
Figure 8. Hydraulic characteristic of the four selected injector for engine tests

Four injectors were selected, exactly as in laboratory environment, as a representative sample of production dispersion, see figure 8. But in this case a comparison of indirect results was necessary in order to evaluate the effects of the strategy. This paper describes the results of the tests made with three different injection conditions with idle engine speed:

1. Single fuel injections (figure 9).

2. Split fuel injections, up to 3 pulses per stroke with a defined DT (figure 10): injectors work in ballistic mode when engine works at engine low load conditions. As already described this behavior normally shows a big dispersion. Correction is switched off.

3. Split fuel injections, up to 3 pulses per stroke with a defined DT(figure 10). Same as before, but correction is switched on.

Time
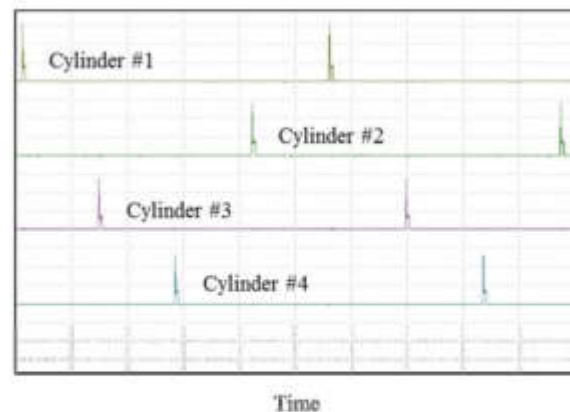
Figure 9. Current signal of four injectors driven with single injection

Time
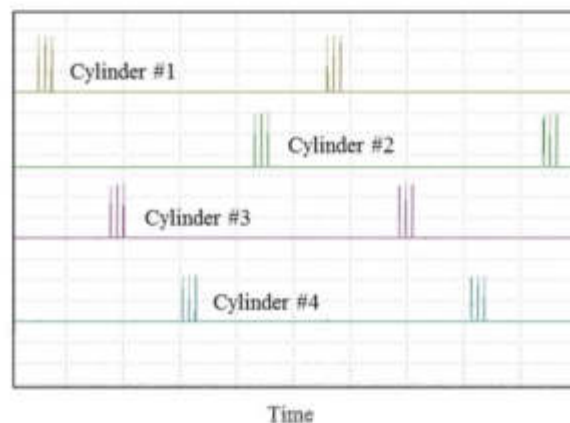
Figure 10. Current signal of four injectors driven with multiple injections

The output variables are: engine speed, fuel rail pressure and the four gas temperatures from exhaust manifold. The engine management system (EMS) in the vehicle was originally programmed to control fuel injectors only in the linear part of their characteristic. When the split of fuel injection is active, the single fuel injection pattern is transformed in a multiple fuel injection pattern. The total fuel quantity requested by the engine (i.e. 4-6 mg in engine idle speed conditions) is divided in three small fuel injections, each of them contributes to the combustion process with a fuel quantity of 1,5-2 mg. This fuel quantity is small enough to shift the injectors commands into the ballistic zone, without changing any other combustion parameter.

The engine management system was not programmed for the management of multiple injections; for this reason the performances of the engine get worse immediately after splitting. Cylinder #1 temperature drops, because injector #1

111

is too lean when it is used in ballistic area; engine speed and fuel rail pressure become highly irregular. The EMS is not programmed for multiple injections, so it is not correct to compare result when DRE correction is OFF.
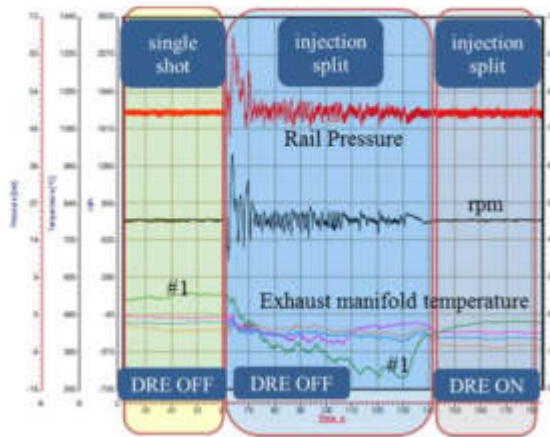
Figure 11. Vehicle results

Figure 11 shows that, after switching ON the correction of DRE, it is possible to have similar engine performances (exhaust gas temperature, rpm and rail pressure) even in fuel injection splitting conditions. This means that DRE correction works as previewed.

2. Tc -DRE

Tc- DRE Description

The closing time strategy works differently. The strategy starts from the following concept: the knowledge of the real hydraulic opening time of each injector depends on the knowledge of injector real closing time.
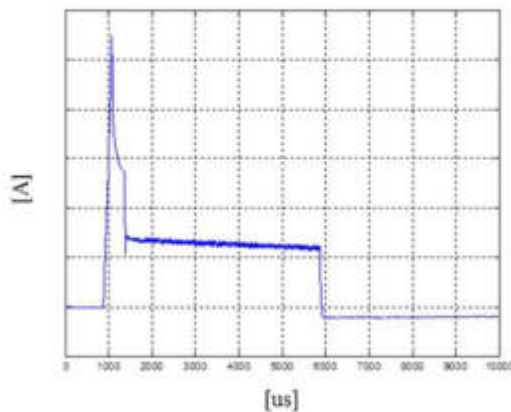
Figure 12. Current profile at a generic pulse width

GDI injectors command has a current profile like in figure 12, and a voltage profile as shown in figure 13. A very high voltage value is applied to its terminals, the current rises abruptly generating a large electromagnetic force that attracts the needle opening the fuel passage. When the electrical signal stops, the electromagnetic force drops to zero because the coil is quickly de-energized by a large value of negative voltage and the needle is pushed against the valve seat by the action of the spring load closing the fuel passage.
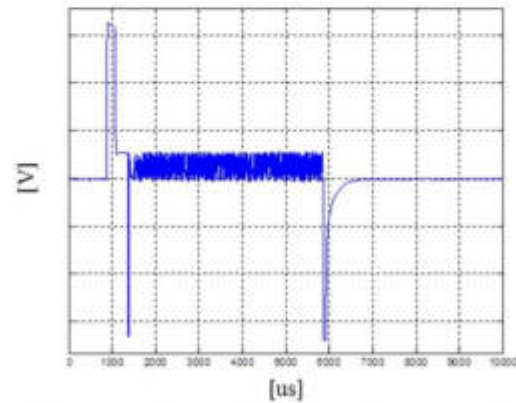
Figure 13. Voltage profile at a generic pulse width

The injected fuel quantity is a function of the injector flow rate and depends directly on the opening and closing time of the needle. Past experience says that opening response time of series production of MM injectors is quite similar from one injector to another. It can be assumed that there is a direct correlation between fuel injected quantity and injector closing time. Hence the knowledge of the closing time allows the precise estimation of the real hydraulic opening period. There is an offset linked to the delay in the repeatable behavior during the opening time. The results of a simulation, in figure 14, show the injector flow rate with increasing energizing time in the range of small fuel injected quantities. The assumption of this strategy is that the opening dynamic can be considered to be repeatable for every injector.
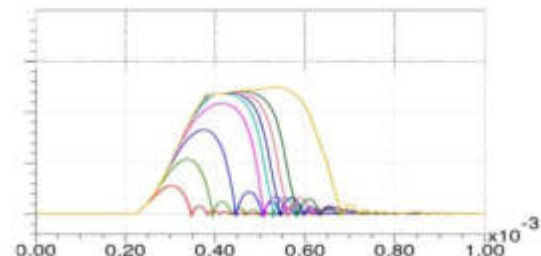
Figure 14. Simulation of injector flow rate with increasing pulse-width in small injections range

112

Tc - DRE Observations

The closing time strategy analyses the needle dynamic; fuel injected quantity is a result of mechanical movements and static injector flow rate. Figure 14 shows, in a CAE simulation of injector behavior, the slope boundaries of the flow rate curves during opening and closing phase of the needle. The transitory phase of the flow rate may be considered as negligible. It is useful to have the highest precision in the series production of fuel injectors in term of static flow rate, because it is a parameter that the strategy of closing time does not have the possibility to correct. It must highlighted that pressure strategy (P-DRE) can also correct differences of static flow rate.

Tc-DRE strategy assumes that injector shot to shot behavior is repeatable with small tolerances. The strategy works only with specific injectors, like Magneti Marelli injectors of IHP3 family.

Tc - DRE Methodology

The ECU measures automatically the closing time extracting the necessary information from the electric signal (figure 12 and figure 13). During the early phases of the development Magneti Marelli technicians compared the electrical signal with a knock sensor in order to detect which was the signal perturbation that carry the closing time information. The results of the validation process showed that the information had been chosen correctly. The second challenge was the definition of the method to extract the information automatically using only special algorithms and circuits. This was achieved using an appropriately designed hardware.
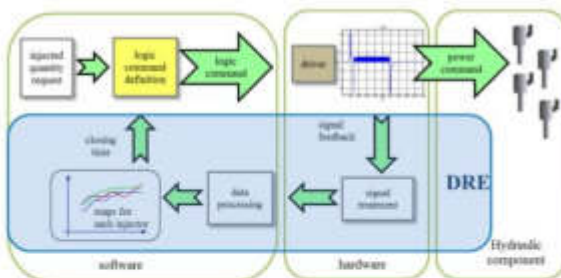

Figure 15. Tc-DRE design

Figure 15 shows the Tc-DRE design: three main vertical groups are presented:

1. Hydraulic component. The injectors are driven by the power signal received by the ECU power stage.

2. Hardware. It is the ECU hardware made by the power stage and all the electronic circuitry in charge of signal treatment.

3. Software. It is the main element, the software elaborates the data received by the hardware extracting the closing time

information saved on maps. The calculation of injectors command is performed considering the engine request and the closing time.

The grouping of signal treatment, data processing and closing time management forms the Tc-DRE perimeter.

Tc - DRE ECU Hardware

The ECU measures the closing time with a mathematical elaboration of electrical signals. This requires the presence of a special hardware circuit inside the ECU for the signal conditioning (amplifier and filter) and for the signal processing (fast acquisition at least at 300KHz). Closing times must be measured with a precision of about 1-2 microsecond, especially in ballistic mode.

Tc - DRE Laboratory Hardware

Tests were performed in hydraulic laboratory. The layout of the bench is the same that was described in the P-DRE strategy and was shown in figure 2.

Tc - DRE Laboratory Hardware

Having a look at figure 16 is it possible to appreciate how much the closing time changes during the ballistic use of the injectors (fuel injection times below 300us). The function requires the highest precision in the closing time measurement due to the high slope to detect the right information. As it was expected, when the electric command of the injector is quite large its behavior is out of the ballistic range and the closing time becomes constant. When the needle reaches the fully open position (figure 1), the closing flight becomes independent from the pulse-width.
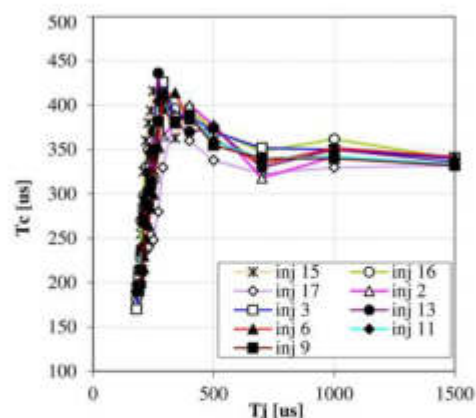
Figure 16. Closing time function of pulse-width

Figure 17 and figure 18 show a detail of the characteristic curve of some of the injectors from Magneti Marelli IHP3 family that were selected to test the strategy. The characteristic curve is displayed with the command values corresponding to the small fuel injections area. The

dispersion of fuel injected quantities for each injector is clearly visible; moreover it is possible to evaluate its value, in terms of relative deviation, in figure 19. This family of injectors has a very high dispersion under 3mg per stroke. The target of this strategy is to have a maximum dispersion of 20% with injected fuel quantity down to 0.5mg.

The strategy is tested on one injector at a time in laboratory environment. The injector is driven and corrected to obtain the requested fuel injected quantity. When the closing time is known, using a closed loop control, the strategy defines the correct command to reach the requested target. Figure 20 shows the results that have been obtained in term of relative deviation compared to the requested target. The figures show how the strategy provides good results: the most part of the deviation points are well under the predicted target at 20% with the exception of one injector.
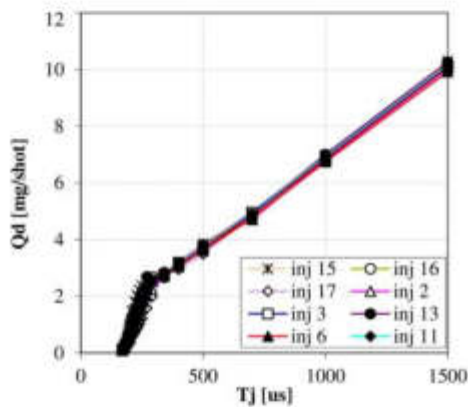
Figure 17. injector characteristic of 9 injectors analyzed with closing time strategy
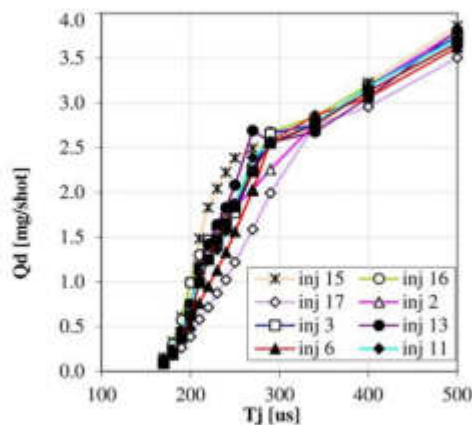
Figure 18. injector characteristic of 9 injectors analyzed with closing time strategy detailed on small quantities
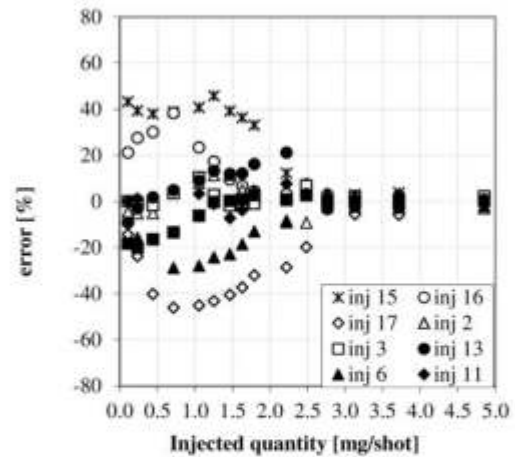
Figure 19. Injected quantity dispersion versus quantity target of injectors without DRE activation at small quantities
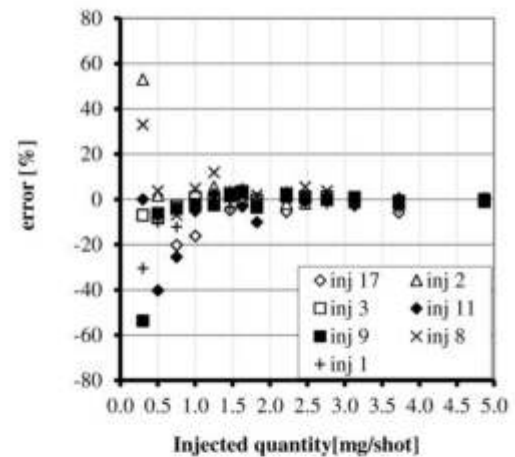
Figure 20. Injected quantity dispersion versus quantity target with DRE activated at small quantities

SUMMARY/CONCLUSIONS

The study presented in this paper describes Magneti Marelli strategies to control GDI injection events at small quantities. To comply with Euro 6 legislation, precise fuel metering is a key factor to optimize mixture formation, injection strategies and combustion. Small quantity injection events correspond to operation in the ballistic mode. In this range production injectors have an unacceptable spread. To comply with the new challenging targets an electronic aid approach is necessary to improve the performance of mechanical injectors. Two different families of strategies were described. One is pressure based, using information contained in the rail pressure signal. With specific algorithms the ECU is able to

114

learn the characteristic curves of the individual injectors and to adjust the drive signal to obtain the desired injection quantity. The second strategy described in the paper works by analyzing the signals used for driving the injectors themselves. Measuring the injector closing time the ECU derives, by software elaboration, a signal adjustment for each injector and each pulse width, thereby obtaining the targeted injected quantity. Performance, results and comparisons with normal operation were described for each strategy. The results lead to an extension of the linear range down to 0.5mg with acceptable tolerances in terms of individual injection event quantity errors.

REFERENCES

1. Bonandrini, G., Di Gioia, R., Papaleo, D., and Venturoli, L., "Numerical Study on Multiple Injection Strategies in DISI Engines for Particulate Emission Control," SAE Technical Paper 2012-01-0400, 2012, doi: 10.4271/2012-01-0400.

2. Skiba, S. and Melbert, J., "Dosing Performance of Piezo Injectors and Sensorless Closed-Loop Controlled Solenoid Injectors for Gasoline Direct Injection," *SAE Int. J. Engines* 5(2):330-335, 2012, doi: 10.4271/2012-01-0394.

CONTACT INFORMATION

Author:

Ing. Marco Parotto
Magneti Marelli S.p.A. - Powertrain Division, Italy
marco.parotto@magnetimarelli.com

Co-author:

Ing. Stefano Sgatti
Magneti Marelli S.p.A. - Powertrain Division, Italy
stefano.sgatti@magnetimarelli.com

Co-author:

Ing. Fabio Sensi
Magneti Marelli S.p.A. - Powertrain Division, Italy
fabio.sensi@magnetimarelli.com

DEFINITIONS/ABBREVIATIONS

DT - Dwell Time; time between 2 consecutive injections

CAE - Computer-aided engineering

ECU - Electronic control unit

Tc - Closing time of injection needle

Tj - Pulse-width of injector command

EMS - Engine management system

MM - Magneti Marelli

DRE - Dynamic range extender

IHP1 & HP3 - Names of Magneti Marelli injector families

Qd - Dynamic injected quantity

SAE *International*

115

[6]   Subic, A., Cvetkovic, D., "Virtual Design and Development of Compact Fast-acting Fuel Injector Solenoid Actuator," Int. J. Vehicle Design, Vol. 46, No. 3, pp. 309-327, 2008.

Virtual design and development of compact fast-acting fuel injector solenoid actuator

Aleksandar Subic*

School of Aerospace, Mechanical and Manufacturing Engineering,
RMIT University, Bundoora East Campus,
P.O. Box 71, 3083 Bundoorn, Victoria, Australia
E-mail: aleksandar.subic@rmit.edu.au
*Corresponding author

Dean Cvetkovic

School of Electrical and Computer Engineering,
RMIT University, City Campus,
G.P.O Box 2476V, 3001 Melbourne, Victoria, Australia
E-mail: dean.cvetkovic@rmit.edu.au

Abstract: Modern petrol engines demand lightweight, compact, fast-acting fuel-injectors. There has been an increasing focus in recent years on the development of automotive fuel injector solenoids that would allow reduction in size while not compromising their performance. The main objective of this research is to develop a specific virtual design approach capable of reducing the size of the fuel injector solenoid while improving its response time and attraction force. The simplicity and effectiveness of the developed virtual design methods allow for quick and accurate evaluation of alternative fuel injector designs early in the design and development stage.

Keywords: electro-mechanical; modelling; control; fuel injector; solenoid; actuating force; response time.

Biographical notes: Aleksandar Subic is the Head of School of Aerospace, Mechanical and Manufacturing Engineering at RMIT University, Australia and Chair in Mechanical Design. He is the Director of SAE-A and member of Editorial Board of the *International Journal of Vehicle Design*. He leads the research group in Automotive Design and has published over 200 internationally peer reviewed publications. He is internationally renowned for his work in sustainable engineering design.

Dean Cvetkovic is a Lecturer in the School of Electrical and Computer Engineering, RMIT University, Australia. He graduated with a BEng in Electrical Engineering, MEng and PhD from RMIT University. His main research focus over the past ten years has been electro-mechanical and

electro-magnetic modelling and simulation of a range of physical phenomena and devices, including fuel-injector solenoids, using the finite element method.

1 Introduction

The application of ordinary circuit equations to describe the behaviour of fast-acting automotive fuel injector solenoids is often inaccurate and time-consuming (Cvetkovic and Subic, 2000; Subic and Cvetkovic, 2001). Solenoids consist of magnetic components and other materials that are characterised by non-linear behaviour. As solenoid dimensions get smaller, non-linear effects have a more profound impact on its operational performance. Unfortunately, traditional tools for simulating circuits are clumsy at best in predicting the behaviour of fuel injector solenoid components. However, special software simulators combine circuit theory for analysing currents and voltages, equations of motion to compute mechanical interactions, and Finite Element Analysis (FEA) to compute magnetic fields and compensate for potential non-linearities. Design engineers who consider such effects can confidently typify designs and avoid unforeseen difficulties that previously would have shown up only in hardware.

The main objective of this research is to improve the operational performance of a fuel injector solenoid by increasing its attraction force and reducing its response time (Cvetkovic, 2001). This objective has been achieved by applying several modelling and simulation procedures that resulted in a new intelligent process for miniaturisation of fuel injector solenoid designs and the reduction of fuel injector response time.

Miniaturisation of the fuel injector solenoid, consists of four specific steps. Firstly, several fuel injector solenoid models, each characterised by a particular plunger pole shape, are modelled and simulated in order to miniaturise the plunger (moving component), plugnut (stationary component) and determine the maximum attraction force possible. The second step involves determining the plunger length characteristic using a range of fuel injector solenoid 2D models. The third step focuses on miniaturising the yoke (case) and introducing the permanent magnet in a 3D modelling environment for a range of fuel injector solenoid models. In the last step permanent magnet principles has been adopted to finally yield a miniaturised fuel injector solenoid model within a 2D environment. The outcomes of all four steps are verified using appropriate computational simulations.

In this research, modelling and simulation of fast acting fuel injector solenoids is undertaken using Ansoft's (1997a, 1997b) Maxwell 2D/3D and Electro-mechanical System Simulator (EMSS) software packages. The general procedure for designing the fast acting fuel injector solenoid commences with the preparation of a first-pass model for FEA. This includes selecting preliminary geometry for the fuel injector solenoid, such as the coil, and associated mechanical parts, particularly those that moved. In this particular study, a computational model of the basic physical fuel injector solenoid is constructed. This model is in the form of a 2D and 3D cross-section. In order to progress with the FEA, it is essential to make several assumptions, such as the range of drive currents, materials and their attributes, and boundary conditions, which dictate how the magnetic fields behave at solenoid component interfaces and at the edges (Alotto et al., 1996). The design of fuel injector solenoids involves frequent calculations

of attraction force and response time. Simulation of both the electromechanical component and circuitry of fuel injector solenoid system using EMSS is required in order to answer the following questions:

- Does the plunger closes quickly enough?
- How much energy the device needs?
- What waveshape typifies the transient current running through the coil after closure?

2 Brief theoretical background

Solenoid actuators are electro-mechanical devices that convert electrical energy into mechanical energy associated with linear motion. Generally, solenoid devices are used in a wide range of mechanical applications, such as fuel-injection systems, starting motors, Anti-lock Breaking Systems (ABS) etc. This research focuses primarily on a solenoid of a fast acting automotive fuel injection system for petrol engines, as shown in Figure 1. The purpose of a solenoid in this application is to control how long the injector is open and therefore how much fuel goes into the system.

Figure 1 Basic fuel injection system used in modern vehicles

Electronic Control Unit (ECU)

Fuel Injector   Common Rail   Fuel Supply Pump
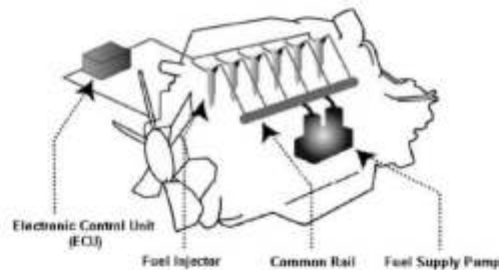
Today, every new model vehicle has a fuel injector. There is an increasing focus now on simple devices such as solenoid actuators to make the fuel injectors more efficient and smaller, which would result in smaller fuel injectors, and eventually smaller engines. However, beside miniaturising the solenoid actuator it is also of interest to improve its operational performance. The operational performance of the fuel injector solenoid depends on two major factors, such as the response time and attraction force of the plunger. The fast actuation of fuel injector solenoid is influenced by the response time or the elapsed time from the initial switch on, until the plunger is in the fully attracted position. The energising time for a solenoid to complete a given stroke, measured from the beginning of the initial pulse to the seated or energised position, is influenced by the attraction force of the plunger. An increase in the attraction force would lead to a reduction in the response time or the energised time. From this relation the fuel consumption can be minimised and the overall engine operation can be improved.

The fuel injector is an electromechanical device that sprays and atomises the fuel. The principle is similar to the way a perfume spray atomises perfume. With direct fuel injection, the fuel is sprayed directly into the throat of the inlet port. In addition, the

injector must vary the quantity of fuel to suit the engine operating conditions. This is achieved by varying the time that the injector is open. The required conical spray pattern is obtained by pumping the fuel through a pintle-type nozzle (Hillier, 1996). Fuel flow takes place when the nozzle valve is opened by a solenoid.

3   Computational modelling of fuel injector solenoids

Computational modelling and simulation of fast acting fuel injector solenoids involves the development of a suitable computer model of the actual physical system and an analysis using this model for the purpose of either understanding the behaviour of the system or evaluating various scenarios (within the limits imposed by the set of design criteria). It is necessary to identify the relationships between the design parameters of the solenoid device and their effects on solenoid's performance. This is achieved by deriving a set of mathematical equations relating the structural, functional and performance parameters of the solenoid design. There are several factors that affect the switching time of the fuel injector solenoid, including: weight, attraction force and plunger cross sectional area, applied voltage, stroke and magnetic path length, number of coil turns, and other factors (Takashi and Yoshihisa, 1995).

3.1   Overview of 2D and 3D models

The original fuel injector solenoid actuator has been accurately measured prior to modelling in order to determine the essential component dimensions, such as length and width of plunger, coil, yoke etc. From this point, the collected data is used to model a virtual 2D and 3D fuel injector solenoid model. The dimensions of the virtual 2D model are identical to the original fuel injector, as shown in Figure 2. Also, the input parameters, such as the weight of the plunger and the number of coil turns had to be determined. The purpose of modelling the original fuel injector solenoid actuator is to build a base model that will be used then as a basis throughout the miniaturisation process (Cvetkovic and Subic, 2000; Cvetkovic, 2001).

The main objective of this process is to develop a miniature fuel injector, compared to the actual fuel injector. Four specific miniaturisation procedures are further developed and used to determine the maximum attraction force characteristic for a range of fuel injector solenoids. These procedures will be evaluated before a miniaturiased fuel injector solenoid is finalised as follows:

- Modelling and simulation of various plunger pole shapes. This will be done in order to determine the relationship between the pole shape and operational performance for a range of solenoid actuators in terms of maximising the attraction force.

- Modelling and simulation of various plunger lengths. This will be undertaken in order to determine the relationship between the plunger length and attraction force at a particular stroke.

- Modelling and simulation of 3D solenoid actuator. The main design differences include a rectangular shape of the plunger, compared to the cylindrical shape plunger used for 2D solenoid models. This particular procedure consists of seven models. The modelling principle is similar to the plunger pole shape and length procedures.

119

For the modelling and simulation of a miniaturised 3D solenoid, it is important to introduce as few components as possible, miniaturise the yoke component, and assign a permanent magnet to a solenoid actuator in order to obtain the maximum attraction force.

- Modelling and simulation of 2D permanent magnet solenoid actuator. Several 2D solenoid models that include permanent magnets will be investigated in order to determine the relationship between the various permanent magnet properties and the attraction force.

Figure 2 2D cylindrical cross-sectional (left) and physical (right) fuel injector solenoid model
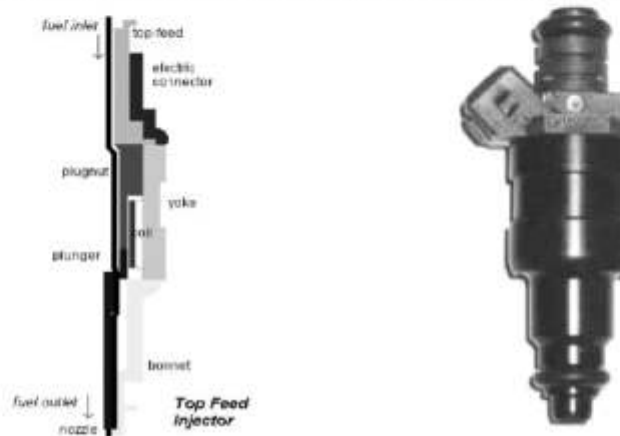
The reduction of fuel injector solenoid response time has gain increased importance in fuel injection systems for petrol engines. To control such engines precisely, it is important for the fuel injector solenoid to achieve a fast response operation time of 1 ms, which is more than 20 times faster than the conventional fuel injector solenoid. The response time is the elapsed time from the initial switch on until the plunger is in the fully attracted position. In order to analyse the reduction of the response time of fuel injector solenoid actuators, an appropriate mathematical model is required, capable of simulating the electronic, mechanical and magnetic system behaviour (Subic and Cvetkovic, 2001; Cvetkovic, 2001). The main focus of this research is modelling of the mechanical system behaviour. The reduction of fuel injector solenoid response time will be investigated here using the following three main modelling approaches:

- standard EMSS fuel injector solenoid model
- spring characteristic EMSS fuel injector solenoid model
- plunger rebound decay EMSS fuel injector solenoid model (bounce effect).

3.2   Modelling, simulation and control of solenoid actuation

3.2.1   Miniaturisation procedures: modelling and simulation of plunger pole shapes

The main objective of plunger pole shape procedure is to examine the influence of various plunger pole shapes in relation of the attraction force characteristics. In order to improve the thrust or attraction force of solenoid actuators, there are several features that need to be evaluated accurately. From the theoretical point of view, the reluctance needs to be decreased and the flux density on the plunger increased. However, from the modelling and simulation point of view, it is necessary to simulate the 2D Maxwell axisymmetric plunger pole shape solenoid model in order to determine the virtual attraction force (Cvetkovic and Subic, 2000; Cvetkovic 2001).

Solenoids can provide different force vs. displacement profiles by varying the angle or the shapes of the plunger pole or other surfaces between the two members, such as plugnut. In general, a flat surface provides high forces at small gaps, while a sharper angle can provide higher forces at large gaps, as shown in Figure 3. In order to examine a more uniform force/air gap relation, several pole and plugnut contact areas have been geometrically modelled using the standard 2D Maxwell axisymmetric solenoid. The geometry is systematically reduced in relation to the original fuel injector solenoid as required. The affects of internal and external changes in the magnetic circuit, affect force characteristics.

Figure 3   Different plunger pole shapes of the solenoid actuator
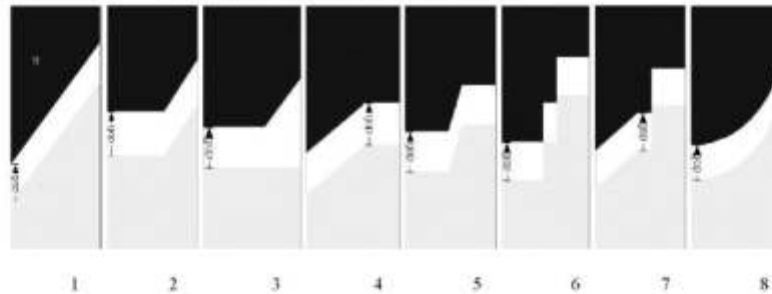
3.2.2   Miniaturisation procedures: modelling and simulation of various plunger-length models

The development of plunger length solenoid models is similar in principle to the plunger pole shape solenoid models. In order to determine the force-stroke characteristics for various plunger lengths, a particular method is used here to determine the highest attraction force at a specific stroke for a set plunger length, as shown in Figure 4. From the theoretical background and the mathematical models developed to date, it can be seen that the strength of the magnetic field is different at each position inside the coil (Cvetkovic and Subic, 2000; Subic and Cvetkovic, 2001; Cvetkovic, 2001). The most effective actuation or the strongest attraction force acting on the plunger would most probably occur in the centre of the coil. The modelling and simulation of solenoid

actuators is therefore conducted in order to determine the strongest attraction force possible. As in the case of the pole shape model, the following objectives are being targeted:

- miniaturisation of the plunger
- determining the highest attraction force possible in the solenoid actuator.

Figure 4   Plunger length miniaturisation approach for solenoid actuator model
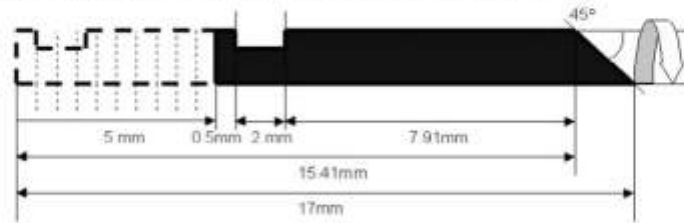
3.2.3   Miniaturisation procedures: modelling and simulation of 3D solenoid actuator model

The 3D solenoid actuator model is developed using a similar principle as the one used to develop the plunger pole shape and plunger length models in previous sections. A wide range of models have been explored and ultimately used to miniaturise the size of key solenoid components, such as yoke. The developed models are based on the following main design intents:

- introduce as few components as possible in the solenoid actuator model
- miniaturise the yoke component
- assign a permanent magnet to the solenoid actuator.

The initial 3D model aims at creating simple geometry and introducing as few components as possible in the solenoid actuator system. This model does not have top and bottom sleeves (plugnut). Also the coil and plunger are shifted in (-z) direction by 1 mm (by the thickness of the bottom sleeve). The next model investigated has the yoke length reduced by 8.4%. Due to the reduction in yoke length by 8.4%, the attraction force of this model is reduced to 1.65 N. Considering that a gap is created when the top and bottom sleeve are excluded from the solenoid actuator system, it is possible to insert a permanent magnet in this space to replace the top sleeve. The purpose of inserting a permanent magnet into the solenoid actuator is to determine if the attraction force has improved. The size of the solenoid actuator is reduced when the top and bottom sleeves are excluded and the attraction force is increased to 2.15 N at 0 mm stroke. The permanent magnet solenoids (also known as magnetic latching or self-holding solenoids) utilise the attraction force advantage of a high performance permanent magnet. When the coil is energised, the plunger moves toward the pole piece. Once energised, and the plunger movement has occurred, the plunger will remain in the energised (holding) position. The coils magnetic field and the permanent magnetic field are additive. To release the solenoid from the 'hold' position, the coil field must cancel the permanent

magnet field or reverse its direction, and the current must flow in the opposite direction. Figure 5 shows a miniaturised permanent magnet component displaced at various positions between the yoke and coil.

Figure 5   3D permanent magnet solenoid models
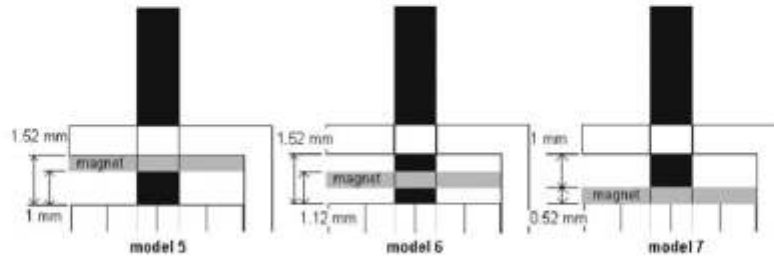
The magnet shifted in (-z) direction by 0.52 mm, (or by the thickness of the magnet). Considering that the miniaturised dimensions of the solenoid have been established, it is necessary next to determine the maximum attraction force developed by the associated three models. The computational results indicate the highest attraction force of 2.49 N at stationary plunger position. Therefore, it can be concluded that the location of the magnet component definitely affected the attraction force of the solenoid actuator.

3.2.4   Miniaturisation procedures: modelling and simulation of 2D solenoid actuator with permanent magnet

The miniaturised 3D solenoid actuator model is constructed using a permanent magnet in order to generate the maximum possible attraction force. Further miniaturisation of the permanent magnet solenoid actuator will be conducted within the 2D Maxwell environment. The attraction force is investigated here in relation to the 2D solenoid model size. The plunger type used in this model is of the flat pole shape, as shown in Figure 6. The permanent magnet, positioned between the coil, yoke and plunger, is identical to the 3D permanent magnet solenoid model. The type of material assigned is Neo35. The overall stroke of the solenoid actuator is 3 mm, with the air gap of 0.165 mm. In order to define the vector properties for magnetisation and polarisation, the vector's angle of magnetisation from the global x-axis is set to 0° (pointing to the right hand side). When developing a miniaturised permanent magnet model, the feature of magnet orientation is of great importance.

Four different vector directions have been investigated, including 0°, 90°, 180° and 270°. This investigation has two main objectives:

- to determine the highest attraction force that can be developed in the four vector directions

- to use the particular direction to model four additional models with different permanent magnet properties assigned to the geometrical magnet object.

123

Figure 6   Direction of the permanent magnet solenoid models
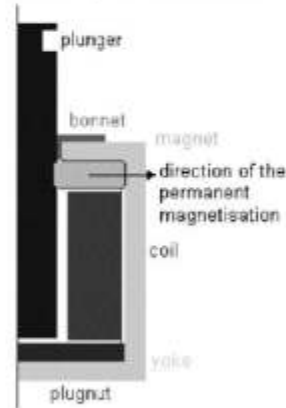
Table 1 provides a summary of all fuel injector solenoid models and parameters used in the simulations. Figure 7 shows the schematic of the so-called evolution process used in this research to achieve a miniaturised fuelled injector solenoid.

Table 1   Summary of all solenoid parameters and material properties

Solenoid models/parameters	2D plunger pole shape and 2D plunger length solenoid	3D solenoid	2D permanent magnet solenoid	Unit
Voltage	12	12	12	(V)
Coil resistance	104.3	271	629.3	(Ω)
Mass of plunger	0.75	0.47	0.4	(g)
Stroke	5	2.47	3	(mm)
Coil-current	552	207	134	(AT)
Coil turns	4800	4679	7035	(T)
Airgap	0.5	0.25	0.165	(mm)
Plunger	SS430	SS430	SS430	
Yoke	Cold rolled steel	Cold rolled steel	SS430	
Bonnet	Cold rolled steel	–	Brass	
Plugnut	SS430	SS430	SS430	
Coil	Copper	Copper	Copper	
Sleeve (tube)	Brass	Brass	Brass	
Magnet	–	Neo35	Neo35, NdFe35, NdFe30, SmCo24, SmCo28	

124

Figure 7   Evolution process of fuel injector solenoid
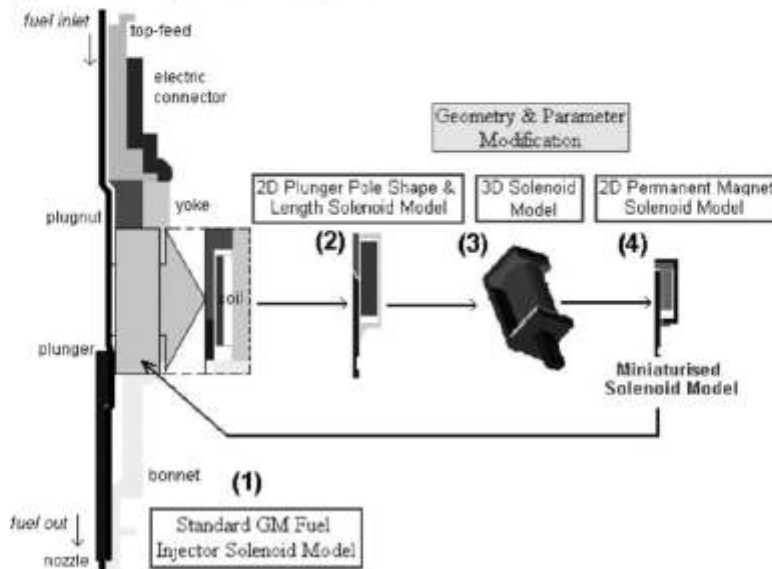
3.2.5   Reduction of response time: standard EMSS fuel injector solenoid

The electromechanical fuel injector solenoid system consists of magnetic, electrical and mechanical subsystems. The electrical subsystem consists of DC voltage pulse, 12 V, coil resistor element, 14.5 Ω, number of coil turns, 315 and a coil current of 827 mA (Subic and Cvetkovic, 2001). The mechanical subsystem consists of loads on mechanical device, such as externally applied force (−0.01813 N), gravitational force acting on a mass and the mass of the plunger. The force source value is determined from plunger mass (1.85 g) times the acceleration due to gravity 9.8 kg-m/sec^2 or −0.01813 N. The Standard EMSS model is investigated here on the plunger weight basis. Mechanical subsystem is set at four different weights, 1.55, 1.70, 1.85 and 2.00 g. The magnetic subsystem (fuel injector solenoid model) is connected to both of these subsystems via four ports. The main objective of system simulation is to calculate the minimum response time for the solenoid to close after the coil is energised.

3.2.6   Reduction of response time: spring characteristic EMSS fuel injector solenoid

The main objective of the spring model is to create a more realistic virtual model. Ideal spring parameters exhibit negligible mass and mechanical losses. A linear ideal spring is included to produce a force directly proportional to the relative displacement of the two ends of the spring. Since the force at one end of the spring must be equal and opposite to the force at the other end, the force is transmitted through the spring in a manner similar to the way current passes through an inductor. The model consists of an externally

applied force, gravitational force acting on a mass, and two springs with different relaxed lengths, as shown in Figure 8. The long spring's relaxed length equals the length of the plunger stroke of 0.3 mm.

Figure 8 Spring model of the solenoid fuel injector
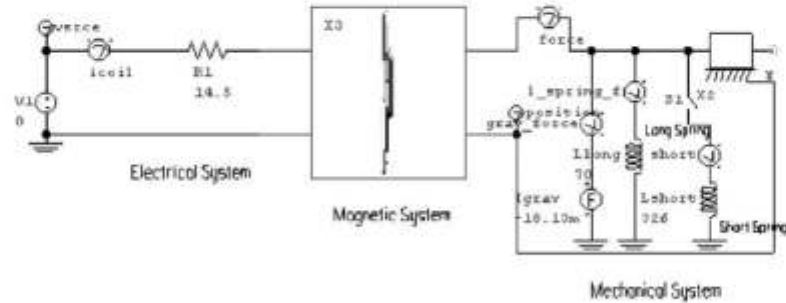
The principle of switch operation is relatively simple. By energising the solenoid coil a magnetic force is generated, which pulls the plunger in a closed position. The long spring's relaxed length equals the length of the plunger stroke of 0.3 mm. The short spring (at rest and out of contact when the plunger is in the fully open position) prevents the plunger from 'hammering' down against the plugnut (plunger stop). An electrical switch in series with the short spring models causes this effect by closing when the plunger reaches the short spring position. When the plunger contacts the short spring, it is near the minimum gap position, and the switch closes in order to activate the short spring force. Switching occures automatically as a result of setting the positive and negative control terminals of the switch to monitor plunger position. De-energising the coil allowes the long spring to push the plunger back open. At this stage of the spring EMSS fuel injector solenoid modelling, the plunger position moves from 0.3 mm to 0 mm. The plunger contacts the short spring at a gap of 0.08 mm. From the initial plunger position of 0.3 mm to a position of 0.08 mm, the short spring is out of the circuit. Long spring operates only at that position. Current probes placed in the circuit branches provide data to characterise the interaction of the forces. As a load, the long spring with the setting force of −1.71 N and a spring constant of 7.8 N/m is used. The short spring force is 6.53 N and spring constant 81.6 N/m. The position threshold is at 0.08 mm.

3.2.7 *Reduction of response time: plunger rebound decay model*

The plunger rebound decay model is developed due to constant plunger opening and closing movement in the nozzle of the injector. There are several factors that influence the motion of the plunger, such as the effects of liquid-induced forces, spring preload forces, and most importantly the rebound forces (Subic and Cvetkovic, 2001; Cvetkovic, 2001). The rebound force represents both the elastic and inelastic collision forces between the plunger and plugnut. These particular factors are included in order to improve the realism of the developed virtual approach to reducing the response time of fuel injector solenoids. In adition, it is also important to stabilise the plunger motion at its closing time. Considering that the plunger represents an integral part of the injection

nozzle, its smooth and rapid operation plays an important part in the overall design of the fuel injector.

In most of the mechanical system applications, the mechanical stops that are included in mass and inertia models are represented by very high damping values or damping coefficients. These are switched on when the mass position exceeds one of the position limits. Most of the time, they behave as ideal stops. However, a real stop could cause a rebound effect or have different cushioning behaviour. These 'bounce effects' are modelled in Maxwell Spice as shown in Figure 9.

Figure 9   EMSS rebound delay circuit model
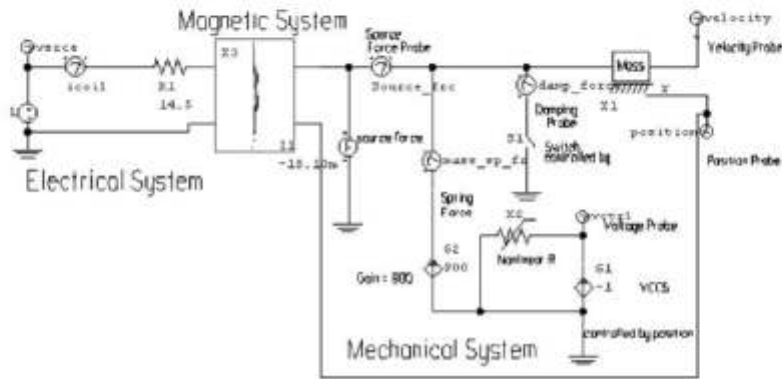
This circuit operates on the principle that when the stop material compresses, there is a spring-like restoring force proportional to the amount of compression. Also, a linear damping coefficient is included representing energy loss during material compression. The starting position or the stroke is 0.3 mm. The circuit senses when the mass travels beyond 0 mm, and develops a signal proportional to the amount of negative displacement. The Voltage Controlled Current Source (VCCS) *G1* in Figure 9 uses position as a control voltage, with a gain of −1. The current represented mass position is forced through a nonlinear resistor to produce the control voltage *VCTRL*. The VCCS *G2* injects the spring force into the mass. Its gain of 900 represents the stop's spring constant *K*, N/m. The voltage-controlled switch simulates rebound damping. Without this damping, the position would show multiple bounces. After the system reaches steady state, the source and spring forces are in balance. The investigation of such models has been used in this research for validation of general procedures and techniques developed. In the case of EMSS modelling, the manipulation of physical parameters of the fuel injector solenoid (such as mass and the starting position of the moving plunger, minimum and maximum positions, and damping coefficients) have been included in order to improve the fast acting fuel injector solenoid.

4 Optimisation of fuel injector solenoid design

4.1 Miniaturisation of a fuel injector solenoid

The first section of this paper assesed modelling and simulation approaches involving a range of 2D and 3D models, which have been essential in the process of solenoid miniaturisation. Throughout the miniaturisation process, each step has been evaluated separately, so that every feature of the fuel injector solenoid could be miniaturised as much as possible until the main condition is satisfied. The second section assessed the response time of the original fuel injector solenoid. The virtual results were compared to determine the minimum possible response time operation. The process of miniaturisation of fuel injector solenoids has been undertaken by modelling standard cylindrical axisymmetrical solenoids with a conical type plunger. From this model, simulations were done for various plunger pole shapes. The simulations of force vs. displacement in the solenoid enabled application of the plunger pole shape solenoids for either long or short strokes.

The essential feature of the plunger pole shape design is in fact the surface area manipulation. Since the pole face area and flux density are related to the attraction force, the Magneto-Motive Force (MMF) is also related to magnetic flux density. The attraction force is proportional to the square of the magnetic flux density in the air gap. Therefore, it could be determined that the attraction force is dependent on the square of MMF, on the effective cross-section of the plunger, and on the inverse square of the air gap. The initial computation involves the flux line plots of plunger pole shapes. These computational plots are important in establishing the shortest magnetic flux path possible across the air gap. Eight-plunger pole shape solenoid models are analysed. From the resultant plots of linear flux paths across the air gap for shapes (5)–(7) it is possible to predict that these shapes will have the highest attraction force at 0 mm stroke. The strength of the flux density must be maintained across the air gap in order to achieve the highest attraction force (the three plots have shown this characteristic). Next, the parametric solution is generated by sweeping the plunger displacement from 0 mm to 5 mm. Considering that the standard 2D solenoid model used in this case, operates at 552 ampere-turns AT, the closest force results can be found by correlating 600 AT plot, as shown in Figure 10.

Figure 10 Parametric force vs. stroke computations at 600 (AT)
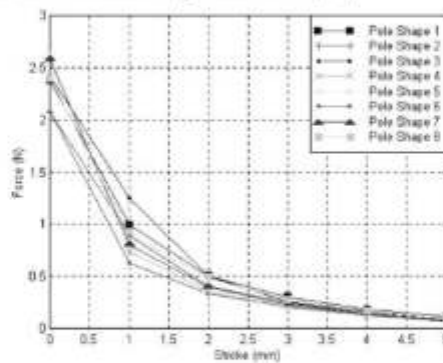
From this plot, it can be concluded that at the maximum MMF, and stationary plunger position, shapes (4), (5) and (7) show the strongest attraction force. As for middle range plunger displacements, shape (3) is the most effective. Considering the surface area of all the pole shapes, shape (1) is the largest, and as a result, it is most suitable for longer plunger displacements. However, shape (3) has the smallest surface area, and it's best operation in terms of attraction force is at stationary plunger displacement.

The design of plunger length solenoid represents the next step in the miniaturisation process. The method aims to determine the highest attraction force at a particular stroke for a particular plunger length. The plunger will exhibit the maximum attraction force when the maximum magnetic field is in the centre of the coil. Accordingly, the air-gap is positioned at the middle of the coil where the maximum force can occur. The force vs. stroke relationship is generated for (11) different solenoid models at 600 AT. At the stationary plunger position, models (5), (6) and (7) develop the maximum force of around 2.77 N. Plunger lengths in this case are 19, 19.5 and 20 mm. As plunger displacement is increased by increments of 1 mm, models (8) and (9) will develop the maximum force of 2 N. At 2 mm stroke, the maximum force is developed by models (10) and (11), 21.5 and 22 mm plunger length respectively. In comparison with the plunger pole shape approach, the results here indicate that the attraction force has increased. This increase in force is due primarily to an increase in the plunger length. From this approach, it has been determined that the attraction force is not necessarily increased as the plunger length increases. The plunger length needs to be optimal in order to achieve the maximum attraction force. Therefore, it is necessary to improve the miniaturisation process in order to achieve a maximum attraction force by developing an appropriate 3D solenoid model.

The 3D solenoid actuator presented here is based on a completely new solenoid design. The main design differences include the rectangular shape of the plunger, compared to the cylindrical shape used in previous 2D solenoid models. By incorporating three important approaches, seven models have been developed. For the modelling and simulation of the miniaturised 3D solenoid, it is important to introduce as few components as possible, miniaturise the yoke component, and assign a permanent magnet to the solenoid actuator. Models (5)–(7) have a permanent magnet of smaller thickness, and the magnet component is gradually displaced at three different positions between the yoke and top of the coil. Model (6) has the highest force of 2.5 N at 0 mm stroke. This proves that inclusion of permanent magnets in solenoids has the potential to increase the attraction force. Overall, the fuel injector solenoid has been miniaturised by now, and the force level has reached the approximate constant level in comparison to 2D solenoid models. Now, the permanent magnet concept is adopted back into 2D modelling and simulation of the final miniaturised solenoid device.

The fuel injector solenoid size is reduced even further in this way, whereby the miniaturised model is now capable of developing the highest possible attraction force. The plunger type used in this model is of the flat pole shape, with the permanent magnet stationed in between the coil, yoke, and plunger, similar to the 3D permanent magnet solenoid actuator. Since the MMF of the system is 134 AT and the overall stroke of the solenoid actuator is 3 mm, with the air gap of 0.165 mm, it is clear that all key parameters have been minimised to a certain extent. The dynamic magnetic flux distribution of the simulated permanent magnet solenoid represents the best indicator of the suitability and quality of developed solenoid designs. The strongest attraction force has been achieved in $0°$ direction. At $0°$, 200 AT, the maximum force is −3.04 N. Considering that $0°$ is the preferred permanent magnet direction, the following investigation involves four

additional magnet properties, such as NdFe30, NdFe35, SmCo24, and SmCo28. It has been determined that the model with Neo35 permanent magnet achieved the highest attraction force, but at the expense of highest energy consumption. Figure 11 shows that Neo35 permanent magnet has the highest attraction force compared to other properties.

Figure 11   Force vs. stroke characteristic for various permanent magnet properties
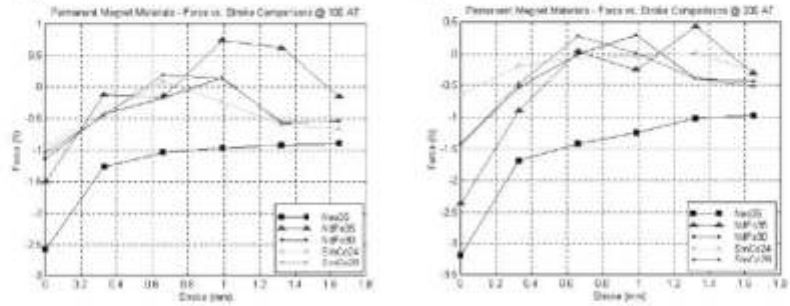
At the stationary plunger position, 100 AT, Neo35 property has developed a force of −2.6 N, and at longer strokes, it reached saturated level, around −1 N until its operation was completed. The other properties exhibit irregularities in attraction force over the complete plunger displacement. At 200 AT, the maximum attraction force is 3.2 N. The dynamic characteristics (magnetic flux density) of the permanent magnet solenoids from 0 mm to 3 mm stroke are analysed with respect to achieving the maximum attraction force, as shown in Figure 12. Due to an increase in plunger displacement, the magnetic flux density is mainly concentrated in the upper part of yoke.

Figure 12   Magnetic flux density of 2D permanent magnet solenoid dynamic characteristics
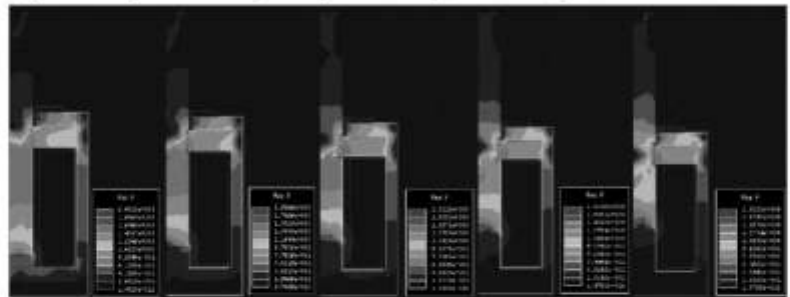
In this approach, the miniaturisation process was flexible, accurate, and the CPU time could be significantly reduced. For the developed miniaturisation process the dimensions of essential solenoid components, (such as coil, yoke, plunger lengths and widths) have been compared and illustrated in Figure 13 (left).

As a result, the computation of the respective attraction force for the models in question has improved, as shown in Figure 13 (right). The development of the 2D

permanent magnet solenoid led to a new miniaturised fuel injector solenoid with the maximum attraction force.

Figure 13 Miniaturisation process (left) and force vs. stroke comparison (right) for various 2D and 3D solenoid models
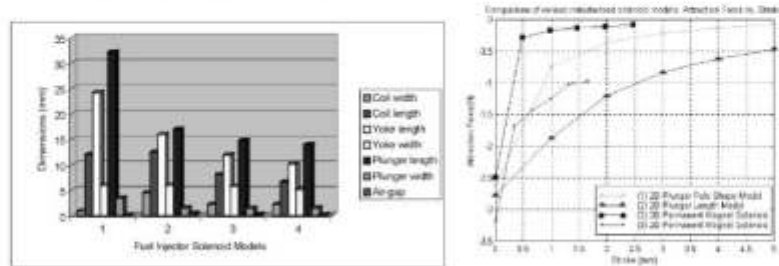
4.2   Reduction of fuel injector solenoid response time

The initial step in finding the fast response solution involves the effects of the plunger weight of solenoid actuator. The weights range from 1.55 g to 2.00 g. At final stroke displacement, 0.3 mm, the smaller weights of the moving plunger, 1.55 g, give the longest response time. However, for shorter strokes, from 0.2 mm to 0 mm, the response time is shorter. In case of a plunger of 2.00 g, at 0.3 mm stroke, the response time is the shortest. But at shorter strokes, from 0.2 mm to 0 mm, the response time is the longest. The weight of the moving plunger has a significant effect on the response time of the solenoid.

In the spring characteristic model, the plunger position moves from 0.3 mm to 0 mm. The plunger contacts the short spring at a gap of 0.08 mm. From the initial plunger position of 0.3 mm to a position of 0.08 mm, the short spring is out of the circuit. At this position only long spring operates. At the initial position, the response time was 1.194 m sec and at 0.25 mm, the response was 0.717 m sec. In comparison to the standard EMSS models, the response time is reduced by 0.050 m sec.

In the rebound delay model, a 1.665 N step force is applied to a plunger mass of 1.85 g. Considering the operation principle of the rebound delay circuit, there is a spring-like restoring force proportional to the amount of compression. Due to this operation, there is a linear damping coefficient representing energy loss during material compression. Similar to the spring model, the starting position or the stroke is 0.3 mm. This type of control operates on the feedback principle, and in terms of response time, the results seem to be effective. Initially, the response time was 1.121 msec, and at 0.25 mm, the response time was 0.404 m sec. Overall, these results prove the effectiveness of the 'bounce model' in terms of rebound delay, which resulted in shorter response time. Figure 14 illustrates the response time minimisation process by comparing the three principle models developed in this research.

The rebound EMSS model is more complex than the standard and spring models previously considered. It resulted in minimum response time. A comparative analysis of the response times achieved by the three models is shown in Figure 15.

Figure 14 Process of reducing the response time of the fuel injector solenoid
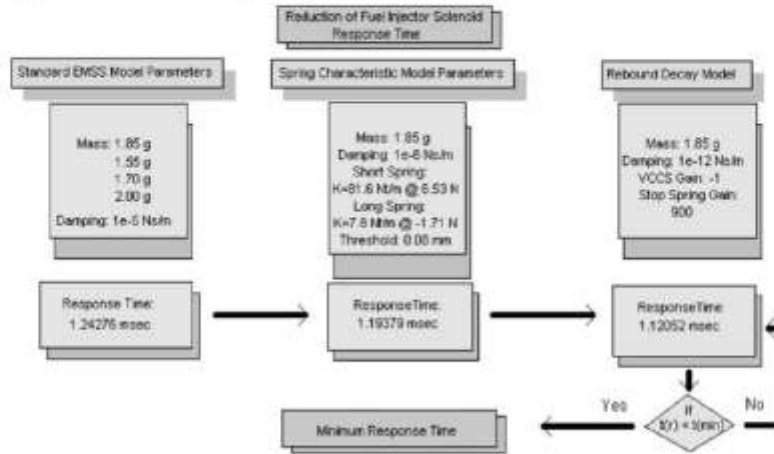
Figure 15 Comparison of the response times achieved using the three models
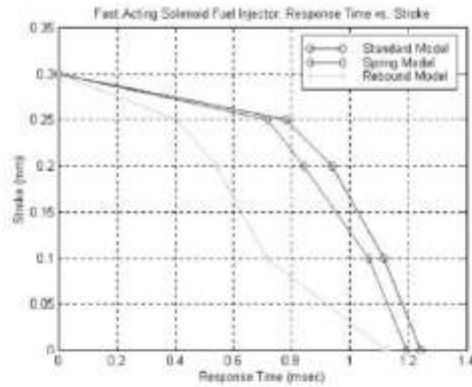
5 Discussion of results

The initial process of miniaturisation of the fuel injector solenoids has been undertaken by modelling the essential surface area features of the plunger pole shape design. Our initial modelling and simulations revealed that solenoids can provide different force vs. displacement profiles by varying the angle or the shapes of the plunger pole or other surfaces between the two members, such as plugnut (Dyck et al., 1999; Yoon et al., 1997). In general, a flat surface provides high forces at small gaps, while a sharper angle can provide higher forces at large gaps, as shown in Figure 3. The design of plunger length solenoid was the next step in the miniaturisation process. The method was

132

established to determine the highest attraction force at a particular stroke for a particular plunger length. The results showed that the increase in attraction force was due to an increase in plunger length. From this approach, it has been determined that the attraction force was not necessarily increased as the plunger length was increased. In order to improve our miniaturisation design, a 3D solenoid actuator was developed. The computational results proved that inclusion of permanent magnets in the solenoid increased the attraction force despite the optimisation of yoke geometry. At this stage the permanent magnet (Neo35) concept was adopted back into 2D 'final' miniaturised solenoid model which generated the maximum attraction force when compared to other permanent magnet properties.

Reduction of fuel injector response time has been achieved through virtual modelling and simulation. The virtual fuel injector solenoid modelling and simulation consisted of three EMSS models. The three EMSS models included the standard EMSS fuel injector solenoid model, spring characteristic EMSS model, and plunger rebound decay EMSS model (bounce effect). The reduction of response time was certainly achieved by the three EMSS models. Finally, the minimum virtual response time was 1.12 msec, developed by EMSS Rebound Delay Model. Throughout this research, a new approach for miniaturisation, force improvement and respose time reduction was flexible, accurate and allowed for simple adaptability to other solenoid actuator applications.

6   Conclusion

The presented research focused on the modelling and simulation of fast acting fuel injector solenoids. The main objective of the presented design approach was to improve the operational performance of existing fuel injector solenoids by reducing their size and response time, while increasing the attraction force required for effective actuation performance. Using the developed approach, the initial size of a typical automotive fuel injector has been reduced by 35%, the attraction force increased by 26% and the response time reduced by 76%. The modelling and simulation procedures have been successful in reducing the response time that is typically difficult and tedious to achieve through experimentation or tiral-an-error. In this research, the minimum response time was achieved by using the virtual solenoid rebound delay model. The simplicity and effectiveness of the developed methods allow for quick and accurate design and evaluation. The developed modelling and simulation approaches have greatly improved the operational performance of the fuel injector solenoid as confirmed through correlation of computational results.

References

Alotto, P., Girdinio, P. and Nervi, M. (1996) 'Mesh adaptation in finite element analysis of 2D steady state time harmonic eddy current problems', *IEEE Transactions on Magnetics*, Vol. 32, No. 3, pp.1361–1364.

Ansoft (1997a) *Getting Started: A 2D Magnetostatic Problem, Ansoft Helping Manual*, Ansoft Corporation, Pittsburgh, USA.

Ansoft (1997b) *Getting Started: A 2D Parametric Problem, Ansoft Helping Manual*, Ansoft Corporation, Pittsburgh, USA.

Cvetkovic, D. (2001) *Modelling and Simulation of Fast Acting Fuel Injector Solenoids*, Master of Engineering Thesis, RMIT University, Australia.

Cvetkovic, D. and Subic, A. (2000) 'Modelling and simulation of a miniaturised actuating solenoid device', Paper presented to *IASTED International Conference 2000 Applied Simulation and Modelling*, Banff, Alberta, Canada, pp.70–76.

Dyck, D., Lowther, D.A., Malik, Z., Spence, R. and Nelder, J. (1999) 'Response surface models of electromagnetic devices and their application to design', *IEEE Transactions on Magnetics*, Vol. 35, No. 3, pp.1821–1824.

Hillier, V.A.W. (1996) *Hillier's Fundamentals of Automotive Electronics*, Stanley Thornes Ltd., UK.

Subic, A. and Cvetkovic, D. (2001) 'Miniaturisation of electro-mechanical actuators for fuel injection systems', Paper presented to *NAFEMS World Congress, The Evolution of Product Simulation*, Lake Como, Italy, April, pp.415–426.

Takashi, K. and Yoshihisa, K. (1995) 'Development of high-speed solenoid valve: investigation of solenoids', *IEEE Transactions on Industrial Electronics*, Vol. 42, No. 1, pp.1–7.

Yoon, S-B., Hur, J., Chun, Y-D. and Hyun, D-S. (1997) 'Shape optimization of solenoid actuator using the finite element method and numerical optimization technique', *IEEE Transactions on Magnetics*, Vol. 33, No. 5, pp.4140–4142.

[7]   Kekedjian, H., Krepec, T., "Further Development of Solenoid Operated Gas Injectors with Fast Opening and Closing," SAE Technical Paper 940450, https://doi.org/10.4271/940450, 1994.

SAE TECHNICAL
PAPER SERIES

940450

Further Development of Solenoid Operated Gas Injectors with Fast Opening and Closing

Harry Kekedjian and Tadeusz Krepec
Concordia Univ.

SAE *The Engineering Society
For Advancing Mobility
Land Sea Air and Space*®
I N T E R N A T I O N A L°

**International Congress & Exposition
Detroit, Michigan
February 28-March 3, 1994**

400 Commonwealth Drive, Warrendale, PA 15096-0001 U.S.A. Tel: (412)776-4841 Fax:(412)776-5760

GLOBAL MOBILITY DATABASE

All SAE papers, standards, and selected books are abstracted and indexed in the Global Mobility Database.

Printed in USA   90-1203C/PG

940450

Further Development of Solenoid Operated Gas Injectors with Fast Opening and Closing

Harry Kekedjian and Tadeusz Krepec
Concordia Univ.

ABSTRACT

The existing state of the art in solenoid operated injectors technology promotes their use mainly for gasoline manifold injection in spark ignition engines. The attempts to use such injectors for direct fuel injection are hampered by slower and less repeatable dynamic response, as compared with hydraulically or mechanically operated diesel injectors. Also, the impact of the solenoid overheating and the resulting loss of power, is considered detrimental for such injectors when installed in the cylinder head of an engine. The next (third) generation of solenoid operated injectors is dealing with all these disadvantages. First, it is using a special inductive driving circuit for the solenoid to boost the opening of the injector needle by amplifying the supply voltage, but only for a very short instant of time. Second, it is utilizing the differential angle in the conical needle seat to reduce the unbalanced pressure force acting on the injector needle when closed; this allows the injector to open at higher gas pressure or to use a stronger return spring to accelerate its closing. The paper explains the principles of these improvements and simulates the injector's behaviour on a computer.

INTRODUCTION

Recently, there is a trend towards the use of gaseous fuels, such as natural gas and hydrogen, in automotive engines. This work is stimulated by the expected shortage of conventional fuels in the near future and by pollution regulations. However, several specific problems, such as on-board storage difficulties, backfiring, knock and pre-ignition, as well as engine power decrease, limit the efficient use of gaseous fuels in Otto-cycle engines. Therefore, research has started recently to use natural gas and hydrogen in Diesel-cycle engines, which allow for higher thermal efficiency than the lower compression engines with fuel supply through the intake manifold.

This paper describes the follow-up development in solenoid operated injectors for direct gas injection into the cylinder of a high compression engine, using microprocessor technology. This concept was first proposed by Krepec (Ref. 1) and realized by Giannacopoulos and Miele, using commercially available diesel injectors modified towards solenoid use (Ref. 2 & 3); next, a second generation of the gas injectors was developed, investigated and optimized by Kekedjian and Hong (Ref. 4). Finally, the third generation of redesigned and improved gas injectors is being developed, as described in this paper.

PROBLEMS ASSOCIATED WITH SOLENOID OPERATED HIGH PRESSURE GAS INJECTORS

In the R & D work on direct injection of gaseous fuels into the cylinder of a diesel engine, there are several problems which have to be overcome. These are:

1. **Control of nozzle opening**. In some earlier cases (Ref 5 & 6), the researchers used the existing diesel fuel injection pump and a modified diesel injector with a hydraulic actuator made to move the injector needle and to open the nozzle when the pressure wave from the pump arrives. In other cases, a solenoid actuator was used (Ref. 2,7 & 8) to move the injector needle according to the signal from an electronic control unit. While the first method was rather acceptable for a

development work, due to the high additional cost of a fuel injection pump, the second method, more feasible for the use of modern electronic injection control, is still not developed enough to be implemented. Some of the intermediate solutions, to use mechanical actuation with solenoid operated timing control, are already in use, but are only limited to special engine designs, like 2-stroke Detroit Diesel (Ref. 9).

2. **Control of nozzle leakage**. In the case of gaseous fuel injection, the leakage problems are more severe than with liquid diesel fuel injectors, due to much lower gas viscosity. While the gas pressure does not have to be as high as the liquid fuel pressure (gas does not need to be atomized) the gaseous jet penetration requires substantial pressure at the range of 100 bar; then, this pressure must be maintained inside the injector during 90 % of the engine cycle time with minimum leakage through the nozzle valve seat into the cylinder. The problem is aggreviated by the fact that no soft material can be used for the nozzle seat sealing, because the nozzle is protruding into the combustion chamber and such material should resist the high temperature.

3. **Nozzle valve seat wear**. It can be expected in a gaseous fuel injecting nozzle that the valve seat in the nozzle body will wear more rapidly than in a liquid fuel injecting nozzle. There are two reasons: one is the smaller damping of the nozzle needle movement when it approaches the seat at high velocity, as compared with diesel fuel which is providing a better cushion for the seat; second is the less effective cooling of the needle seat, as compared with the liquid fuels which are absorbing the heat from the nozzle at higher rate.

4. **Injector needle bouncing and post injection**. The lack of liquid fuel between the injector needle and the nozzle body, as well as the reduced tight fit surface between these two parts in an injector with pressure balanced needle (Ref. 4), increases the tendency of the needle to bounce on both ends: at the needle upper stop and at the seat. While in a properly designed injector, the upper stop bouncing should not affect the injector discharge characteristic, in such injectors where the needle lift is very small (Ref. 7), this type of bouncing can alter the injected fuel discharge rate and, consequently, the injected fuel dose. Also, the injection repeatability can be affected by bouncing at the upper stop of the needle when a small fuel dose is being injected (example: at idle). The bouncing at the needle seat is more harmful. It can lead, not only to the post-injection, but also to faster nozzle seat wear.

5. **Solenoid force, duty cycle and cooling**. One of the reasons that solenoids are not recommended as actuators

for diesel injectors, is their sensitivity to the high temperature which affects their force output. Being installed directly in the cylinder head of an engine, they are heated much more than the solenoid injectors of gasoline engines installed in an intake manifold. At the same time, the requirements regarding application of their force in diesel engines are higher than in gasoline engines, due to the strictly controlled injection timing, the requirement for faster opening and the much higher injection pressures involved. In this situation, the solenoid duty cycle and cooling have a big impact on the solenoid force.

6. **Solenoid injectors optimization**. It would appear that the gas injection system consisting of a constant pressure fuel supply rail and solenoid actuated injectors operated in on-off manner, is the simplest fuel injection system. It is, to some extent, for gasoline injection in SI engines, but it is not for diesel engines. The existing injector opening and closing delays depending on the gas pressure, solenoid temperature, spring preload, nozzle opening characteristic, moving parts inertia and, to a big extent, on the solenoid driving circuit, make the design of the solenoid injector for diesel engines the subject of a complex optimization process (Ref. 4).

In the development of the third generation of solenoid operated injectors for direct injection of gaseous fuels, the above problems have been, to some extent, answered.

INJECTOR DESIGN

Fig. 1 shows the comparison of basic design configurations of the second and the third generation of gas injectors. Both, in the bottom part, conform to the "S" size of the conventional diesel injectors. The following modifications have been made to the third generation injectors to accomodate them better for effective and reliable operation (Fig. 2).

1) More powerful, but shorter solenoid was adapted.
2) Solenoid body was attached to the upper cap.
3) Electric wires have been fastened with plastic seals.
4) Incoming gas was directed around solenoid to absorb heat.
5) Larger spring was more conveniently located.
6) Needle seat was made with a reversed differential angle.
7) Injector needle rod, connecting the needle to the solenoid core, was made more elastic to reduce the needle bouncing.

In the two previous generations of injectors, the conventional CAV diesel injector body was modified and used along with the nozzle. For the third generation, a

138

special, custom made body was manufactured to match the size of a commercially available solenoid and to allow for the change in the gas inlet location for improved solenoid cooling.

In this design, the solenoid body is firmly held up against a copper spacer and the cap by two screws that bolt directly into the solenoid. The created sub-assembly can then be screwed onto the injector's body without the concern of damaging the solenoid.

A larger diameter spring and a special retainer have been placed inside the injector body, so that the bottom face of the body can be used as a stopper for the needle, as it is in the original diesel injectors. The needle lift is, thereby, accurately set and is independent on the gap between the solenoid core and the bottom of the core cavity, as long as this gap is larger than the needle lift itself.

There are more changes introduced in the injector design to address the specific problems with solenoid operated gas injectors, as mentioned above. They will be discussed in detail in the following sections.
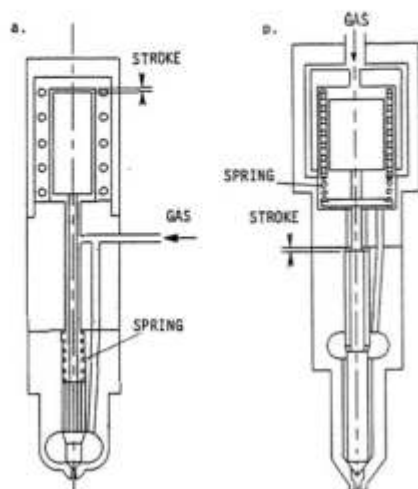
Fig. 1 Comparison of basic design configurations of second (a) and third (b) generation injectors

CHOICE OF SOLENOID

The most appropriate choice for the third generation of solenoid injectors became the Lisk L7, short 0.875 " body length, short stroke solenoid. This solenoid provides considerably more force than the Lisk L5, long 1.25", short stroke solenoid used in second generation injectors. This would guarantee opening at higher gas supply pressures and allow for the solenoid core to be further away from the solenoid backstop. Working with

a larger air gap effectively reduces the time delay between the current and the force developed by the solenoid. The delay between the voltage applied and the current developed is also reduced, since the inductance of the solenoid is inversely proportional to the air gap between the solenoid core and the backstop. Tests performed to illustrate these effects are included in solenoid model. Another important feature of the L7 short body solenoid is that the inertia added to the injector needle assembly is far outweighed by the increase in force and, therefore, does not reduce the acceleration of the needle. The excess in available force makes it possible to find the best position of the core to obtain the optimum performance.
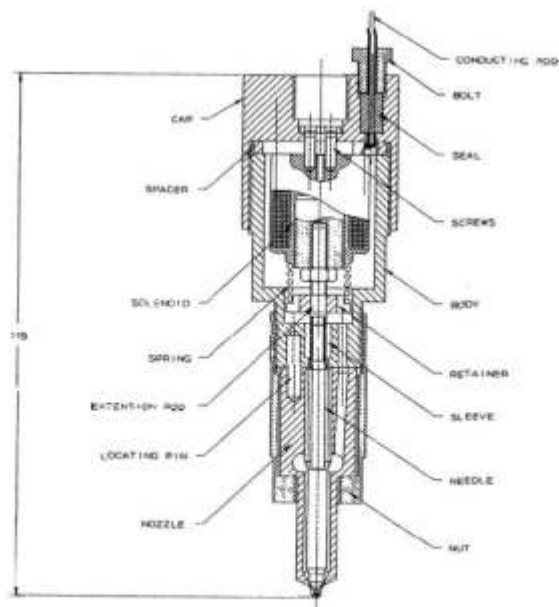
Fig 2 Third Generation of Solenoid Actuated Injectors

CONTROL OF INJECTOR OPENING FORCE

In the presented injectors, the injector needle is practically pressure balanced, after fully open. However, when closed, it is not fully balanced because of the gas pressure force not acting on the needle seat surface. In conventional diesel injectors, this surface is determined by the outside seat diameter, as shown in Fig 3a, because of the different cone angles in the nozzle seat and on the needle tip. The difference in the angles is about 50 min and is creating a circular contact line between both parts which increases the specific contact pressure and improves the sealing.

Because the unbalanced pressure force is delaying

139

the nozzle opening by the solenoid, it was found beneficial to the injector operation to change the direction of the differential angle, so that the inside seat diameter is creating the circular sealing line, as shown in Fig. 3b. This has two advantages:

1) it reduces the unbalanced needle tip surface diameter and, therefore, the solenoid force required to lift the needle, allowing the same injector to operate with higher gas pressure,
2) it reduces the length of the seat circumferencial contact line and, therefore, the gas leakage rate.

For the seat case shown in Fig. 3, the seat sealing diameter Dss will change from 3.2 mm to 2 mm, and the surface on which the unbalanced pressure force is acting will be reduced to only 40 % of the previous area. This will reduce the pressure force to only 40 % of that for the nozzle with conventional seat, allowing for the injector opening with 2.5 times higher gas pressure; or this would allow to use a higher spring preload, thus reducing the time of the injector closing.
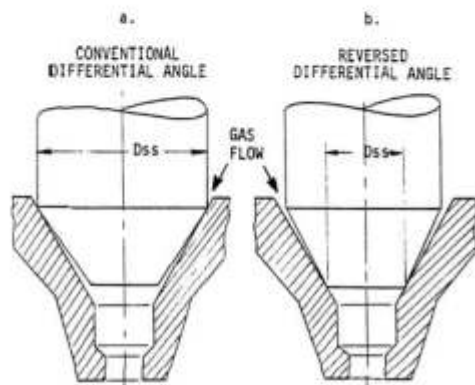
Fig. 3   Modification of the nozzle needle seat from:
   a- Conventional to
   b- With reversed cone angles

NEEDLE BOUNCING SUPPRESSION

There are two situations at which the needle bounces. The first occurs when the injector fully opens and the needle hits the stop on its way up. The second occurs when the injector is closing and the needle hits

the nozzle seat on its way down. The needle bouncing can particularly be increased in an injector for gaseous fuels, due to the fact that the gas has much lower density and viscosity than the liquid fuel. Fig. 4 shows the comparison of injector needle bouncing when there is no gas pressure and with 100 bar gas pressure; both, upper stop bouncing and seat bouncing can be well indentified.
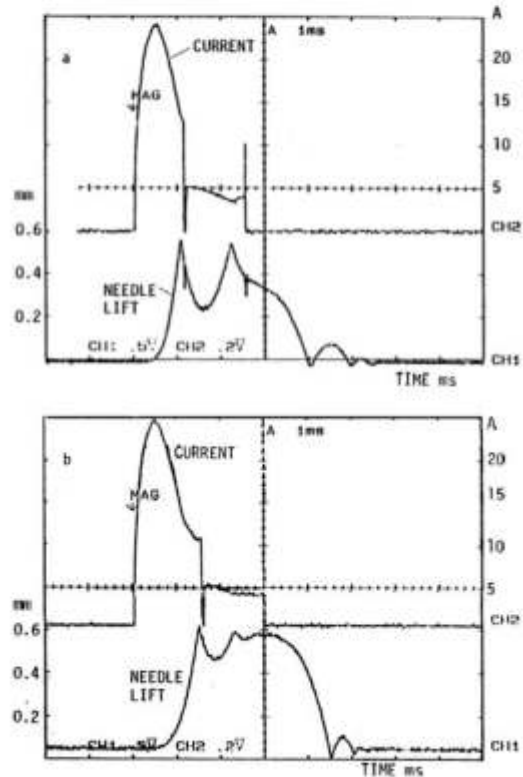
Fig. 4   Comparison of needle bouncing
   a- Without pressure
   b- With 100 bar pressure

To reduce bouncing, an approach was adapted based on the following reasoning. As can be seen in Fig. 2, the moving mass is polarized in two locations: the solenoid core at the top and the needle at the bottom. Near the middle the masses are joined together by a thin wall pipe. This section can behave similar to a stiff spring. By varying this area, the stiffness of this section can be altered. This concept of two masses connected by a spring uses the inertia of the core and the spring load of the elastic rod to absorb the kinetic energy stored in the needle after its tip hits the nozzle seat. This can be best illustrated in Fig. 5. In the first diagram, the two masses connected by the spring (elastic pipe), are moving as one unit towards the seat. As the needle hits

140

its seat, the momentum of the core mass compresses the elastic pipe, reducing it by the distance Δl, as illustrated in the second and the third diagrams. Thus, the compression of the rod exerts a force on the needle mass that holds it on its seat and suppresses the springback.
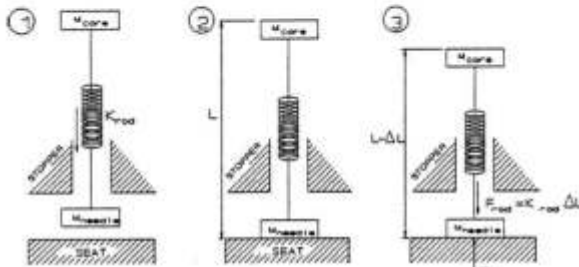
Fig. 5   Incorporated damping illustration

If the stiffness of the elastic rod is too high, the response approaches that of a single mass and the result is a large bounce. Reducing the rod stiffness has the effect of delaying the springback action to the point at which the needle is on the verge of moving away from its seat. This analysis has been confirmed by the simulation results of the injector's dynamic response, as will be shown later.

INJECTOR OPENING AND CLOSING CONTROL

Developing a strategy for controlling the current passing through the solenoid of the gas injector is as vital to its performance as its mechanical design. A microprocessor controls the fuel dose by varying the duration of the pulse sent to the solenoid driver circuit. However, problems usually arise due to the non-linearity in the case of short duration of injection and from variations in the fuel discharge characteristics resulting from the injector wear or the solenoid temperature increase. These issues should be kept in mind when planning the control strategy for the injector.

The following were the objectives for solenoid control scheme:

- minimize the opening and closing delays of the injector,
- limit the current passing through solenoid when not needed,
- ensure that the injector opens at gas pressure over 100 bar,
- boost initial solenoid voltage up to 100 V using a regular 12 V supply.

To fulfil these objectives, a new type of driving circuit was designed with an initial current boost that would minimize the opening delay of the injector. This initial current should almost instantaneously rise to a sufficiently high value. After the injector fully opens, the current should be limited to the minimum holding value for a required duration of time. This desired performance can be achieved with the circuit shown in Fig. 6.
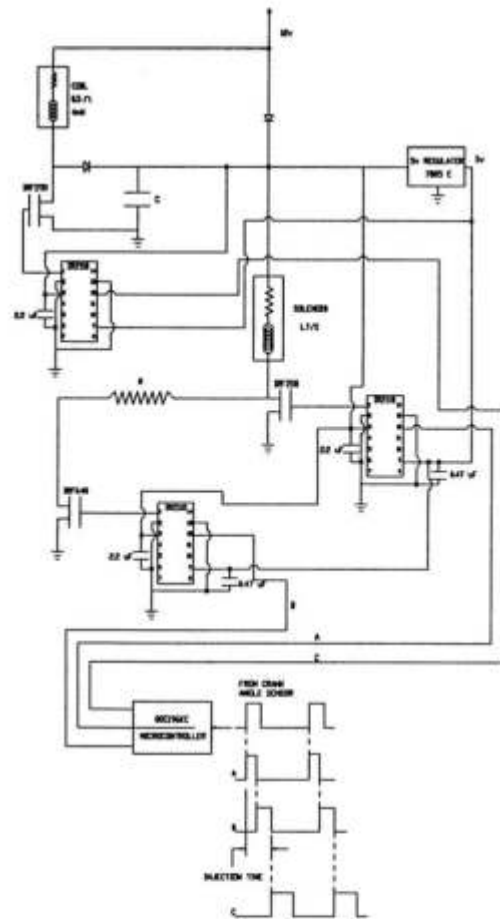
Fig. 6   Injector driver circuit

To avoid the use of a high voltage power supply, the boost to the solenoid force is provided from the capacitor which is charged up during the time when the injector is not in operation. A similar technique was proposed by Green et al (Ref. 7), where the use of a DC/DC converter was suggested for charging up the capacitor to a voltage greater than the supply voltage.

In the proposed circuit, the task of charging the capacitor is simply achieved by interrupting the current flowing through an inductive coil in order to create an instantaneous large potential difference which is next used to channel the current to the capacitor, through a diode which prevents the capacitor from discharging.

Although, boosting the solenoid current using an inductive coil, is an original approach, subsequent limiting of the current is not an uncommon technique in injector driver circuits (Ref. 10). It prevents unnecessary heating of solenoid by an excessive current and it reduces the injector closing delay, since the solenoid force does not have to decay from a higher level, to allow the injector needle to move downwards.

NOZZLE VALVE SEAT WEAR AND LEAKAGE

The work on needle bouncing limitation, did lead to the attempt to limit the needle seat wear. As discussed in the introduction chapter, direct fuel injection requires a hard steel to steal seat contact and sealing. With reduced needle damping by gaseous fuels and with shorter circumference of the seat contact line, the seat wear is becoming an important issue.

Because the seat wear is caused by the seat hammering with the needle tip during the nozzle closing, the simulation of the forces acting on the seat surface from the needle has been made with the results shown later.

Regarding the nozzle seat leakage, there is not much choice, only to accept the hard seat steal-to-steal sealing connection using the differential angle configuration, as discussed above. The present steel heat treatment and grinding technology is able to produce almost perfect stable and desired geometrical shapes, and the experience with the good quality diesel fuel injectors is showing that such a sealing in case of gaseous fuel injectors can also produce satisfactory results.

MATHEMATICAL MODEL OF SOLENOID OPERATED INJECTOR

To investigate all the above phenomena, it was necessary to simulate the solenoid operated injector's behaviour and for that reason a mathematical model of such injector has been created. It is split into four sections: the steady state and transient solenoid model, the electrical model, the mechanical model and the fuel flow model.

1. **Steady-state and transient solenoid model**. The characteristic of a solenoid is the relationship between the solenoid force as the result of the current and in function of time, as well as the inductance variations versus the solenoid core position. A simple procedure

was adopted for gathering experimental data rather than using the existing theory which only defines the steady state characteristic of the solenoid and neglects the transient. More over, the existing theory is also based on some experimental magnetization data required to relate the magnetic flux to the solenoid current and the solenoid force to the magnetic flux. The procedure used here, directly relates the force of a particular solenoid to the current in the winding for various working air-gaps, as well as the solenoid inductance, to the working air-gap.

The experimental setup consisted of a fixture which holds the solenoid in place and a piezoelectric force transducer which measures the force exerted by the solenoid core for various solenoid currents at a fixed core position, This setup is shown in Fig. 7. Current was read using an inductive clamp-on meter. The circuit used for the tests was a basic driver circuit consisting of one switching device which applies voltage to the solenoid.
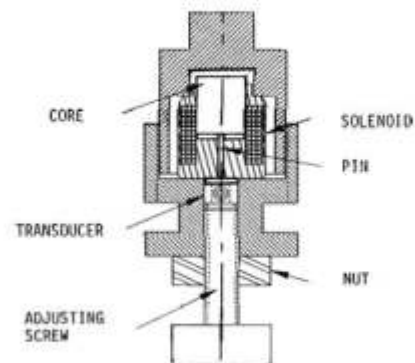
Fig. 7 Experimental fixture for solenoid testing

Fig. 8 shows the oscilloscope traces of the current and force, subject to a 12V square pulse at a constant airgap of 0.6mm. From the oscilloscope traces, the relationship between the steady state force and current could easily be read. However, the force seems to lag behind the current. This is best seen in Fig. 9, where the current was almost cut-off and the force reduced instantaneously at a certain decay rate. Approximating the force-current delay by a first order response, the force decay could be recorded as the first order time constant.

The time constant for current rise, in response to the applied voltage, was also taken, in order to evaluate the inductance of the solenoid. Note that for the force/current time constant and the inductance, there are two separate relationships that can be used in the model. While one is for lower voltage applications, the other is

142

for higher voltage applications, as in the proposed driver circuit. Both these parameters are sensitive to large current, resulting from saturation in the iron path of the solenoid, as the magnetic flux increases.

The results of the experiments are represented
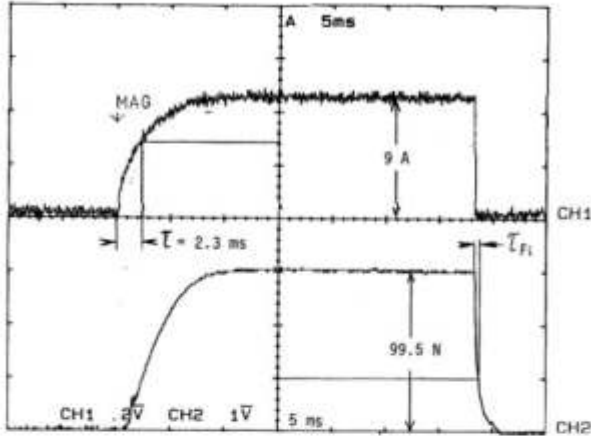
Fig. 8 Oscilloscope trace of solenoid current and force subject to a 12V pulse at an airgap of 0.6mm; τ = time constant (L/R)
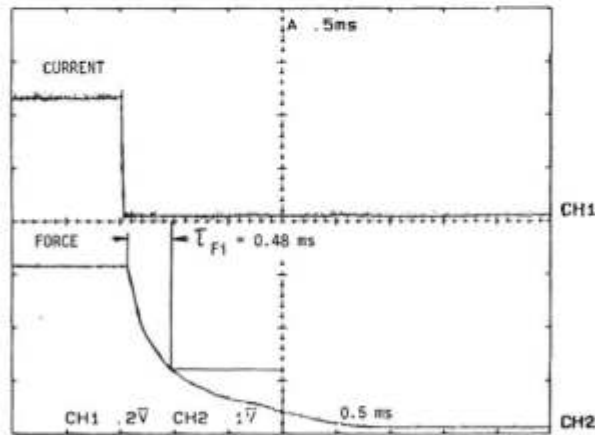
Fig. 9 Extension of oscilloscope trace of solenoid force decay at 0.6mm air-gap, from Fig. 8, with an enlarged time base

graphically in Fig. 10 through 12. Using the experimental data obtained, the electrical and mechanical models can be related by the following non-linear first order approximation:

$$F_{sol}\frac{d\tau_{Fi}}{dt} + \tau_{Fi}\frac{dF_{sol}}{dt} + F_{sol} = f(i,gap) \qquad (1)$$

where, F_{sol} = solenoid force
i = solenoid current
gap = solenoid working gap
$f(i,gap)$ = steady-state relationship between solenoid force and current from Figure 10
τ_{Fi} = solenoid force/current time constant from Figure 11
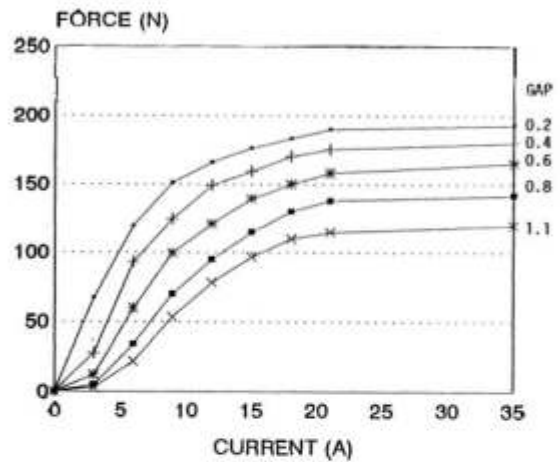
Fig. 10 Graphical representation of solenoid force vs. current for various airgaps
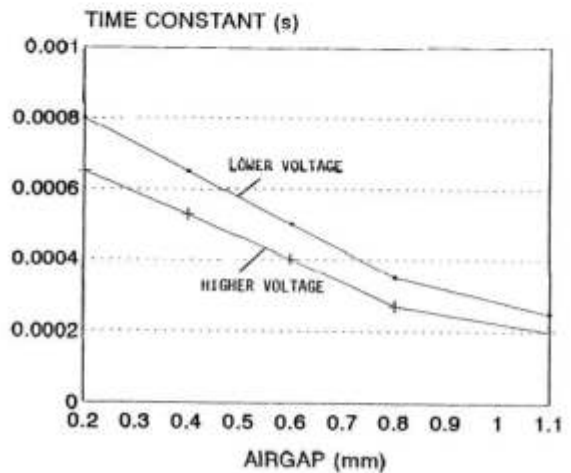
Fig. 11 Graphical representation of solenoid force/current time constant for various airgaps

143
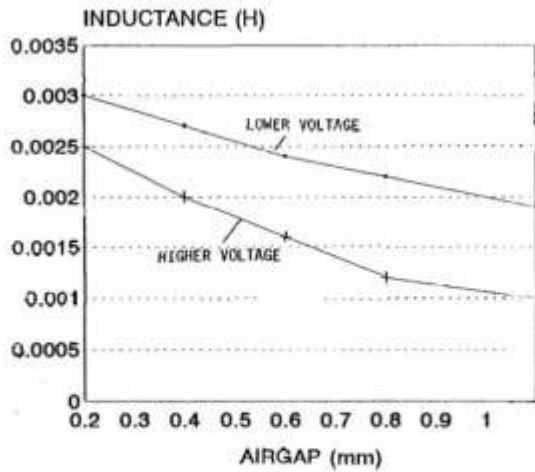
Fig. 12   Graghical representation of solenoid inductance vs. airgap

The experimental data were stored in data files and used in the program. Linear interpolations were performed for intermediate values.

2. **Electrical model.** The electrical model was used to simulate the behaviour of the current in the solenoid, when the appropriate signals are sent to the driver circuit at the time of injection.

A capacitor, which is already charged to a high voltage, discharges through the solenoid. Once the capacitor is discharged down to 12V, the diode prevents it from discharging further. This can be expressed as follows:

$$L_s \frac{di}{dt} + i \frac{dL_s}{dt} + R_s = V_c - \frac{1}{C}\int i \, dt \qquad (2)$$

$$V_c = V_{c_0} - \frac{1}{C}\int i \, dt \quad for \quad V_c > 12V \qquad (3)$$

or,

$$L_s \frac{di}{dt} + i \frac{dL_s}{dt} + Ri = V_s \quad for \quad V_c = 12V \qquad (4)$$

where, V_c = capacitor voltage
Vc_0 = initial capacitor voltage
C = capacitance

For the remainder of the injection time, the current is limited to a value that maintains a force slightly larger than the total force which tries to prevent the injector from staying open. The value of the current is found

from the steady state force/current relationship which was presented in the previous section.

3. **Mechanical model.** The mechanical representation used for the injector model is shown in Fig. 13. From this diagram, we can see the nozzle seat and the needle stopper modelled as a very stiff spring and damper. This was done in order to simulate the needle bouncing. The moving masses (needle and core) are attached by spring and damper which represents the incorporated damping system used to suppress the bouncing of the needle.
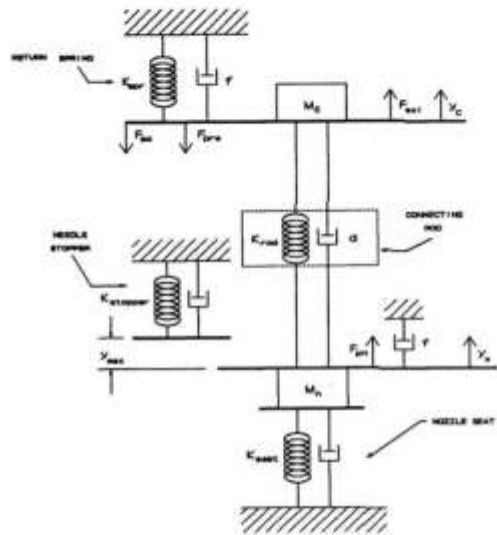
Fig. 13 Mechanical representation of injector model

The mechanical system consists of two moving masses; the needle and the solenoid core. The resultant forces on the needle and core can be expressed as follows:

$$F_n = -F_{krn} - F_{drn} + F_{pn} - F_{fn} - F_{seat} -$$
$$- F_{dnseat} - F_{stopper} - F_{dnstopper}$$
$$F_c = F_{sol} - F_{krc} - F_{drc} - F_{pc} - F_{fc} - F_{spr}$$
$$(5)$$

where, F_n = resultant force on needle
F_c = resultant force on core
F_{krn} = spring force of rod on needle
F_{drn} = damping force of rod on needle
F_{krc} = spring force of rod on core
F_{drc} = material damping force of rod on core

144

F_{pn} = pressure force on needle
F_{pc} = pressure force on core
F_{fn} = friction force on needle
F_{fc} = friction force on core
F_{spr} = force of return spring
F_{seat} = force of nozzle seat on needle
F_{dmseat} = material damping force of seat
$F_{stopper}$ = force of needle stopper on needle
$F_{dmstopper}$ = material damping force of stopper

The spring and material damping forces of the rod on the needle and core are dependant on the relative position and velocity of the two masses. The expressions for these forces are:

$$F_{krn} = K_{rod}(y_n - y_c) \qquad (6)$$

$$F_{drn} = d_{mrod}(V_n - V_c) \qquad (7)$$

$$F_{krc} = K_{rod}(y_c - y_n) \qquad (8)$$

$$F_{drc} = d_{mrod}(V_c - V_n) \qquad (9)$$

where, K_{rod} = spring constant of rod
y_{needle} = needle position
y_{core} = solenoid core position
d_{mrod} = damping value for rod
V_{needle} = needle velocity
V_{core} = core velocity

The spring constant of the rod is related to its modulus of elasticity. The damping value which is a result of internal material damping is difficult to predict. Nevertheless, a value corresponding to a typical 60% reduction of kinetic energy is assigned.

The pressure forces acting on the needle and core masses depend on the position of the needle. When the injector is closed, there is a portion of the needle tip (D_{u}) exposed to the back pressure and the corresponding area at the other end is exposed to the injector pressure. Therefore, the respective pressure forces on the needle and core are:

$$F_{pn} = A_u P_b \qquad F_{pc} = A_u P_i \qquad (10)$$

where, A_u = unbalanced area
P_b = back pressure
P_i = injector pressure

When the needle is beyond the critical lift, corresponding to the critical flow at the nozzle orifice, the pressure on the needle tip is assumed to be approximately equal to half of the supply pressure since the flow is sonic. Therefore, the pressure force is significantly lower due to the reduced unbalanced area and differential pressure across the entire axially moving

mass. The critical lift is dependant on the nozzle orifice diameter and can be found from the geometry of Figure 3, by equating the seat and orifice areas and solving for lift.

The damping caused by viscous forces and friction can best be obtained experimentally by observing the dynamics of the nozzle needle movement. Proportionality coefficients can be assigned to the velocity and the square of the velocity to account for these factors and then be incorporated into the model. As demonstrated in Fig. 4, the density of the gas is a major factor in the damping of the moving masses.

The force of the return spring is acting on the core mass and is expressed by the following:

$$F_{spr} = F_{pre} + K_{spr}y_c \qquad (11)$$

where, K_{spr} = spring constant of return spring
F_{pre} = return spring preload

The nozzle seat and needle stopper are modelled as very stiff springs with internal material damping. The spring constants and material damping values can be approximated by evaluating the compression of the seat and stopper experimentally. This can be done by assigning appropriate values that would transfer 40% of the kinetic energy to spring energy, while the rest is dissipated in the form of damping.

The seat forces act only when the needle is in contact with the seat and can be expressed as:

$$F_{seat} = K_{seat}y_n \qquad (12)$$

$$F_{dseat} = d_{mseat}V_n \qquad (13)$$

where, K_{seat} = spring constant of seat
d_{mseat} = damping value of seat

The force from the stopper acts only when the needle position exceeds the maximum needle lift. This is expressed as follows:

$$F_{stopper} = K_{stopper}(y_n - y_{max}) \qquad (14)$$

$$F_{dseat} = d_{mseat}V_n \qquad (15)$$

where, $K_{stopper}$ = spring constant of needle stopper
y_{max} = maximum lift
$d_{mstopper}$ = damping value of stopper

The accelerations of the moving masses can then be found from the following expressions:

145

$$a_n = \frac{F_n}{M_n} \qquad and \qquad a_c = \frac{F_c}{M_c} \qquad (16)$$

where: a_n = acceleration of needle
a_c = acceleration of core

The velocities of the moving masses can then be updated by integrating the accelerations and similarly the new positions can be updated by integrating the velocities.

4. Fuel flow model. The fuel flow through the injector nozzle is a function of the critical flow area which is a function of the needle position. The critical flow area begins at the nozzle seat and then switches to the nozzle orifice when the needle reaches the critical lift.

The flow through the injector nozzle is sonic and can be expressed as a function of the critical flow by using the following expression:

$$\dot{m}_f = A_{critical} C_d \sqrt{ 2 \frac{P_i^2}{RT} (\frac{k}{k+1})(\frac{2}{k+1})^{\frac{2}{k-1}} } \qquad (17)$$

where, m_f = mass flow rate of fuel

The injected fuel dose is then found by integrating the fuel flow rate.

SIMULATION

Fig. 14 shows an oscilloscope trace of the solenoid current and the needle lift when applying a 12 V pulse to the injector, without fuel pressure, for 5.5 ms. The equivalent simulated result can be seen in Fig.15.
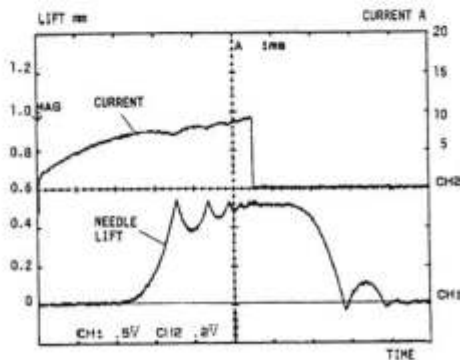
Fig. 14 Recorded injector response for a 5.5 ms, 12 V pulse, without fuel pressure
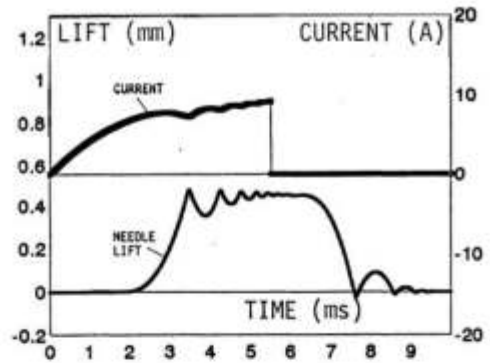
Fig. 15 Simulated injector response for a 5.5 ms, 12 V pulse, without fuel pressure

The simulation shown in Fig. 16 is the dynamic response of the injector using the proposed driver circuit at an injector pressure of 100 bar. This can be compared with the actual response that was shown in Fig. 4b.
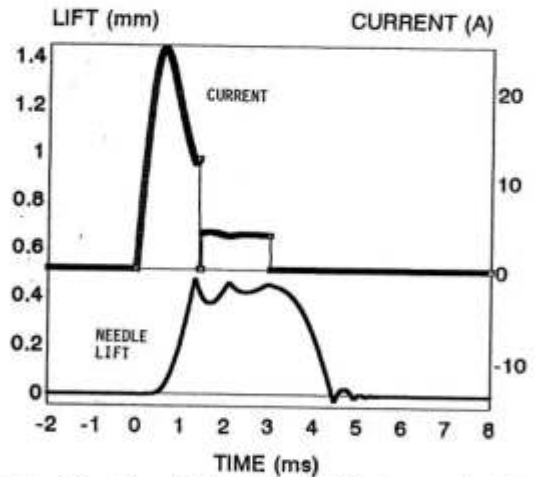
Fig. 16 Simulated response using proposed driver circuit at a fuel pressure of 100 bar

The above dynamic responses were obtained for a stiffness of 160 MN/m, of the section joining the needle and the core. The simulated response, using a 6 MN/m elastic thin wall tube can be seen in Fig. 17. Effectively, the needle bouncing amplitude has been reduced as compared with that in Fig. 16. However, only the reduction of bouncing at the upper stop is well visible; at the seat, the bouncing magnitude seems to be approximately the same. This may be due to the fact that the absolute bouncing amplitude is already low in both cases.

146
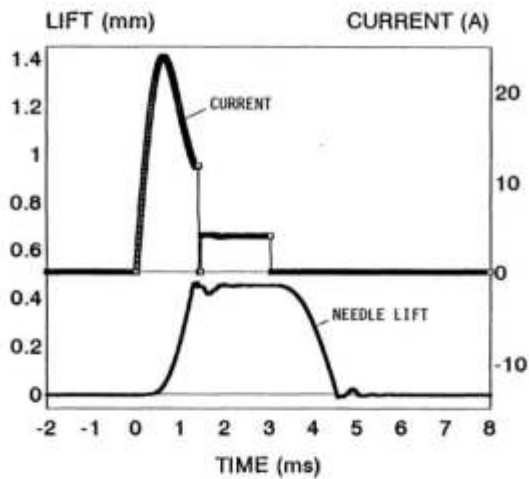
Fig. 17 Simulated injector response with suppressed needle bouncing
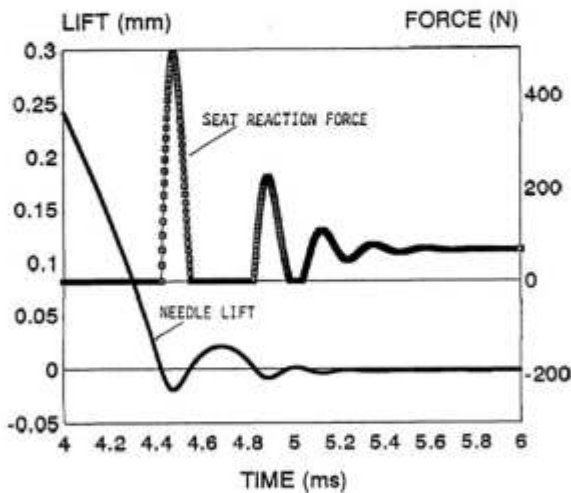
Fig. 18 Dynamic force on nozzle seat with stiff rod

This assumption was, to some extent, confirmed by the simulation results of the force acting on the needle seat which is responsible for the springback of the needle. These are shown in Fig. 18 for the stiff rod and in Fig. 19 for the elastic rod. As seen, the dynamic force, which contributes to the seat wear, is significantly lowered with the elastic rod (approximately by 35%). This creates the possibility to reduce the seat wear, particularly in the case of the fatigue impact. Another interesting feature here is the two seperate peaks of the seat force. This is a result of the delay, created by the elastic rod, between the core mass and needle mass, as

they hit the seat.

It should be noted that, in order to reduce the stiffness of the section joining the core and the needle, one can reduce the cross sectional area (smaller diameter or thin wall tube), or increase the length of the rod, or use a different material with a lower modulus of elasticity.
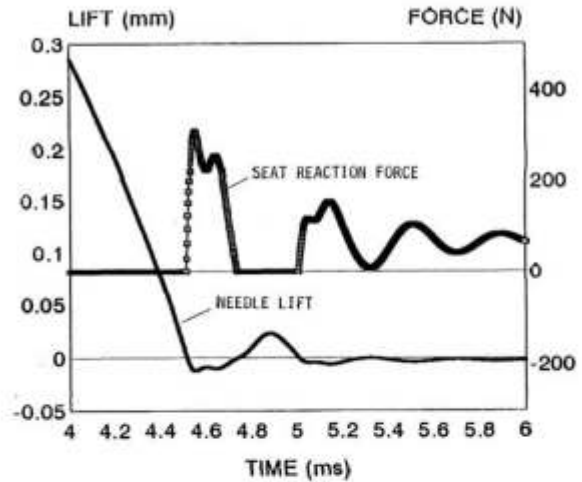
Fig. 19 Dynamic force on nozzle seat using elastic rod

The obtained results are clearly showing the potential of the elastic rod approach and are encouraging further research in this direction.

CONCLUSIONS AND RECCOMANDATIONS

The following improvements have been proposed in the design of solenoid operated injectors for direct injection of gaseous fuels in high-compressed internal combustion engines:

1. The use of inductive coil in the driving circuit, for boosting of the solenoid force, allowed for:

   a) fast opening of the injector, requiring less than 0.7 ms to reach maximum flow from the nozzle,
   b) lower heat release by the solenoid due to very short duration of high current (only 1.4 ms), i.e.to a duty cycle of only 3.5 % approximately,
   c) use of electric power supply of 12 V only to generate 100 V across the solenoid coil.

2. The use of relatively long stroke solenoid did allow for:

147

a) higher injector needle stroke, well exceeding the critical needle lift corresponding to the seat critical flow area, that is equal to the nozzle orifice flow area; this made the injector not sensitive to the needle bouncing at the upper stop,

b) larger gap between the solenoid core and the core backstop that reduces the time delay between the current and the developed force, as well as between the applied voltage and the developed current, since the inductance of solenoid is inversely related to this gap.

3. Reversed differential angle in the nozzle seat and on the needle tip cones, contributing to:

a) reduced unbalanced pressure area that allows for the use of higher gas pressure or higher spring load for the same solenoid force and reduces the injector opening delay,

b) lower gas leakage rate due to the shorter sealing contact line between the nozzle seat and the needle tip.

4. Elastic insert separating the two masses concentrated on both sides of the injector rod, that is contributing to:

a) reduced needle bouncing, mainly at the upper stop and to a lesser extent in the seat, leading to improved injection stability when fully open and to reduced post-injection tendency, when closed,

b) reduced seat wear, as implied from the reduced needle bouncing; this indicates less springback forces between the seat and the needle.

The above conclusions require more, mainly experimental investigations, and the use of a multiobjective optimization procedure because of the complexity of the factors affecting the dynamic response of a solenoid operated injector for direct gas injection.

ACKNOWLEDGEMENT

REFERENCES

1. T. Krepec, T. Tebelis and C. K. Kwok, "Fuel Control Systems for Hydrogen-Fuelled Automotive Combustion Engines--a Prognosis". Int. J. Hydrogen Energy, Vol. 9, No. 1/2, 109-114, 1984.

2. T. Krepec, T. Giannacopoulos and D. Miele, "New Electronically Controlled Hydrogen-Gas Injector, Development and Testing." Hydrogen Energy, Vol. 12, No. 12, p.p.855-861, 1987.

3. T. Krepec, G. Carrese and D. Miele. "Further Investigation on Electronically Controlled Hydrogen Storage and Direct Injection System for Automotive Applications". Int. J. Hydrogen Energy, Vol. 17, No. 12, p.p 955- 960, 1992.

4. H. Hong, T. Krepec, and R.M. Cheng, "Optimization of Electronically Controlled Injectors for Direct Injection of Natural Gas in Diesel Engines." SAE Paper No. 930928, 1993.

5. M. Takiguchi and S. Furuhama, "Combustion Improvement of Liquid Hydrogen Fueled Engine for Medium Duty Trucks." SAE Paper No. 87085, 1987.

6. J.F. Wakenell, G.G O'Neal, Q.A. Baker, C.M. Urban, "An Investigation of High Pressure/Late Cycle Injection of CNG as a Fuel for Rail Applications." Report, Southwest Research Institute, 1988.

7. C.J. Green and J.S. Wallace, "Electronically Actuated Injector for Gaseous Fuels", SAE Paper No. 892174, 1989.

8. N.Glasson and R. Green, "High Pressure Hydrogen Injection", Proceedings of the 9th World Hydrogen Energy Conference, Paris, 22-25 June, 1992.

9. R.J. Hames, R.D. Straub and R.W. Amann "DDEC Detroit Diesel Electronic Control", SAE Paper No 850542, 1985.

10. D.H. Smith and D.A. Spinweber, "A General Model for Solenoid Fuel Injector Dynamics", SAE Paper No. 800508, 1980.

[8]   Yin, Z., Jing, X., Fu, T., Zhou, H. et al., "Study on Dynamic Characteristics of High-Speed Solenoid Injector by Means of Contactless Measurement," SAE Technical Paper 2017-01-2113, doi: 10.4217/2017-01-2313, 2013.

SAE INTERNATIONAL

Study on Dynamic Characteristics of High-Speed Solenoid Injectors by Means of Contactless Measurement	2017-01-2313
	Published 10/08/2017

Zenghui Yin, Xiaojun Jing, Tieqiang Fu, and Hua Zhou
China Automotive Technology and Research Center

Chunde Yao, Jiangtao Hu, and Peilin Geng
Tianjin University

Abstract

In-cylinder direct-injected technology provides a flexible and accurate optimization for internal combustion engines to reduce emission and improve fuel efficiency. With increasingly stringent requirements for the emissions of nitrogen oxides (NO_x) and CO_2, the content of injections in an engine combustion cycle has reached 7 to 9 times in gasoline direct injection (GDI) and the diesel engine with high-pressure common rail (HPCR). Accurate control of both time and quantity of injection is critical for engine performance and emissions, while the dynamic response of injector spray characteristics is a key factor.

In this paper, a test bench was built for monitoring the dynamic response of solenoid injectors with high-speed micro-photography and synchronous current collection system. Experimental studies on the dynamic response of GDI and HPCR solenoid injectors were carried out. Comparative study showed that there was a high correlation of the delay time of opening and closing of solenoid valves with solenoid current. The solenoid valves would not open until current rate reached to a threshold value. When they closed, there was a bump in current curve. Based on the above, the dynamic characteristics of solenoid valves such as opening delay and closing delay could be obtained. As for HPCR injectors, by monitoring and analyzing the pressure of injector inlets, injection duration could be further decomposed into opening process, full-open process and closing process. Compared with GDI injectors, the change of injection pressure had a great influence on the opening delay and the closing delay of HPCR injectors, but actual fuel injection duration was different.

Introduction

In-cylinder direct-injected technology provides a flexible and accurate optimization for internal combustion engines to reduce emission and improve fuel efficiency. With increasingly stringent requirements for the emissions of NO_x and CO_2, the content of injections in an engine combustion cycle has reached 2 to 3 times in the diesel engines of GDI and HPCR. The trend requires excellently rapid and precise delay of high-speed solenoid injectors [1,2,3]. Besides nozzle parameter and injection rate scheduling, the start of injection plays a significant role, because the air in cylinders is influenced by the movement of pistons [4,5,6]. Electronically controllable solenoid-driven injectors for common rail fuel injection system have been widely used for dynamic load response and accurate fuel injection. But improvements in the major design features of these injectors are still required to reduce actual injection delay time and implement the usability of multiple injections [7,8,9,10]. Therefore, it is of great significance to thoroughly understand electronic injector dynamic response and accurate measurement, which is good for the simulation of electronic control fuel injection, the improvement of the performance of engine working process and the optimization of control strategy.

With regard to the dynamic response detection studies of solenoid valves or needles, there are many methods at home and abroad. In mechanical injectors, inductive or hall displacement sensors are often installed at the end of needle valves to measure needle valve lift. However, in electronic injectors, there is not enough space for the sensors. In addition, the changed magnetic field of electromagnetic injectors affects the measuring accuracy of the sensors, so traditional methods cannot get applied to electromagnetic injectors [11]. Shijie AN [12] builds the test bench of solenoid valve response by installing acceleration sensors on solenoid valve suction plate. At the same time, the movement process of electromagnetic valves is got through analyzing acceleration value, and dynamic response characteristics are tested. Jun Wang [13] puts forward the measure method of needle valve lift based on laser displacement sensors which can measure needle valve movement. Jinbao Zheng [14,15] designs pressure measurement system which can accurately measure the pressure of nozzle volume in electric solenoid common-rail system. On basis of the pressure curve of nozzle volume, the delay of opening and closing of needles and injection duration can be calculated. Pressure curve can also reflect needle movement. Seoksu Moon [16] uses a novel technique to monitor the movement of injector needle valves by multi-exposed X-ray phase-contrast images. R. Payri etc. [17,18]

present a methodology for modeling standard common-rail diesel injection system based on the acquisition of silicone molds, and then their visualization is found in an electronic microscope. A one-dimensional model has been obtained to analyze the dynamic behavior of solenoid-operated injectors. The previous studies are limited to the movement of a single needle or an electromagnetic valve, but little attention is paid to the corresponding relation between needle movement and original fuel injection signal. These methods can only be limited to a particular study. It is difficult for universal application. In addition, the sensors installed on injectors destroy the original basic structure of injection system, so real response characteristics cannot be got.

In this paper, a new contactless detection method was put forward on basis of the analyses of current-pressure characteristics without any change of injectors. Solenoid current was measured by high-precision current sensors, and meanwhile high-speed photography was employed to detect injection spray process. By comparing the spray picture with corresponding solenoid current and analyzing the features of current signal, especially of the moment when spray started and closed, a detection method based on magneto-resistive sensors was got. This could provide dynamic response time, opening and closing delay and actual injection duration. In addition, by monitoring and analyzing the pressure of injector inlets, injection duration could be further decomposed into opening process, full-open process and closing process.

Experimental Setup

In present work, both a GDI injector and a HPCR injector are chosen for experiments, but there are some differences with regard to their structures and using conditions. The GDI injector uses electromagnetic valves to realize fuel injection. Fuel injection finishes when electromagnetic valves cover orifices under the effect of spring force. For the HPCR injector, when solenoid valves are deactivated, spring-controlled valves will close outlet-throttles. As a result, the pressure on top of elongated rod is equal to rail pressure. The area in which high pressure acts is significantly greater than the exposed area of needle valves, offering required force to keep the needle valves closed. When injection trigger signals activate solenoid valves, control valves are open and the pressure in control chambers decreases. With the decrease in the pressure of control chambers, rail pressure acting along lower region opens needle valves. The pressure difference between control chambers and accumulation chambers controls the speeds of needle opening and closing.

GDI Injector Test System

Test system mainly consists of fuel injection system, high-speed microscopic imaging system and synchronous data acquisition system (DAQ). The fuel injection system consists of accumulator, injector and ECU. The high-speed microscopic imaging system consists of high-speed camera, microscope lens and LED cold light source with high brightness. The synchronous data acquisition system consists of NI DAQ acquisition device, high-precision current sensor and pressure sensor. The schematic diagram of GDI injector test system is shown in figure 1.
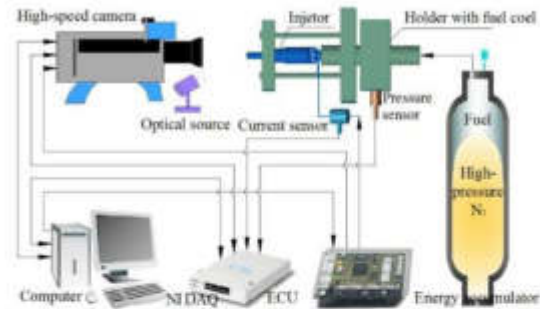
Figure 1. Schematic diagram of GDI injector test system.

In this test system, the energy accumulator is driven by high-pressure nitrogen which can provide 1-200 bars adjustable and steady pressure fuel. ECU sends injection pulse width signal and the injector completes injection according to fixed parameters. Meanwhile synchronization TTL signal sent by ECU triggers high-speed camera system and synchronous data acquisition system. Current signal and pressure signal are collected by synchronous data acquisition system. At last, images and sensor signals are matched and compared.

HPCR Injector Test System

In present work, HPCR injector test system consists of high-pressure common-rail system, temperature control system, DAQ system and injection control system. The schematic diagram of HPCR injector test system is shown in figure 2.
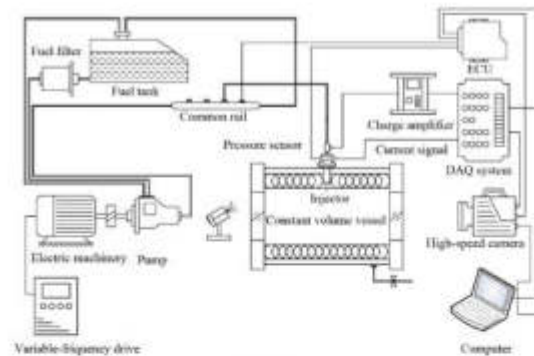
Figure 2. Schematic diagram of HPCR injector test system.

In HPCR system, common-rail pressure is established by electric machinery with the maximum pressure of 1600 bars. The injector is a p-type injector with the nozzle-hole diameter of 0.16 mm and the needle valve lift of 0.25 mm. The pressures at the inlet of the injector and the solenoid valve current are detected. Consistent with GDI injector test system, current signal and pressure signal are collected by DAQ system, and they are matched with spray images.

Experiment Method

Test equipment is connected according to the structure of test device. To ensure each spray image corresponds to the pulse width signal of fuel injection, current signal and pressure signal, fuel pressure is fit and the rate of high-speed camera and synchronous data acquisition system are set with the same frequency. The solenoid valves of

injectors are controlled by ECU. When ECU sends injection signal to fuel injectors, trigger signal is sent to high-speed camera and synchronous data acquisition system at the same time. The high-speed microscopic camera system records spray process near nozzle holes, and NI DAQ collects pulse width signal, current signal and pressure signal and then sends them to computers.

Set the start time of the pulse width signal of fuel injection as a benchmark when a full fuel injection controlled by ECU is completed. The spray process filmed by high-speed camera system is analyzed firstly. The start time of spray means opening time is determined according to nozzle local microscopic imaging. Similarly, the closing delay time of solenoid valves can be determined. Then, the key feature of current curve is analyzed when spray starts and ends. The opening delay time and the closing delay time of electronic controlled injectors can be obtained on basis of this. Finally, combined with the characteristics of injector inlet pressure during the process of injection, injection duration can be further decomposed into opening process, full-open process and closing process. Test conditions are shown in table 1, and the impact of pressure change is also taken into account.

Table 1. Test conditions.

	GDI injector	HPCR injector
Injection pulse width (ms)	3	2
Injection pressure (MPa)	6、9、12、15、18	40、60、80、100、120、140、160
Camera speed(fps)	40000	40000
Resolution	320*320	320*320

Results and Discussion

The test results will be carried out separately from the GDI injector and the HPCR injector.

GDI Injector

Adjust fuel pressure to meet the range of the working conditions of direct injection gasoline engine. Fuel injection pulse width is 3 ms. The frequency of data acquisition which collects the signals of pulse width, current and voltage is same with shutter speed.

Spray Images and the Signal of Current

The pressure of GDI injectors is so high that spray characteristics at opening and closing moment have obvious difference. Figure 3 shows the spray images at the opening and closing moment of solenoid valves. Figure 4 demonstrates that the working process of the injector obtains pulse width signal and current signal.

In figure 4, square wave signal below the signal of the injector is pulse width signal, and the top one is current signal. The two positions of A and D are current curve positions when the synchronized camera with high speed captures spray images at

opening time and closing time. The interval between point A and the opening time of electromagnetic valves is defined as the opening delay time of solenoid valves (T_{od}). The interval between point D and the closing time of electromagnetic valves is defined as the closing delay time of solenoid valves (T_{cd}). In this text, the conditions are that T_{od} is 0.350 ms and T_{cd} is 0.575 ms. Actual injection duration is 3.250 ms, longer than the pulse width.

To find the feature points of current curve which represent opening moment and closing moment, the curve is processed with derivation. Figure 5 shows current curve and current rate curve.

In figure 5, position A is the moment of solenoid valve opening when current rate reaches the maximum position. In this process, current is gradually increasing, and electromagnetic force becomes larger. When electromagnetic attraction is equal to the sum of spring preload, pressure and the gravity of ball, solenoid valves begin to move. When they begin to move, electrical energy is rapidly converted into kinetic energy, and at the same time the current rate of coils is rapidly reduced. During the closing of solenoid valves, magnetic flux becomes small. An induced electromotive force is generated in the coils, which hinders coil current from becoming further smaller. Due to self-induced EMF effect, coil current becomes large, leading to a bump (position D) in figure 4. Thus, by analyzing the characteristics of current curve, T_{od} and T_{cd} can be acquired.

325μs   350μs
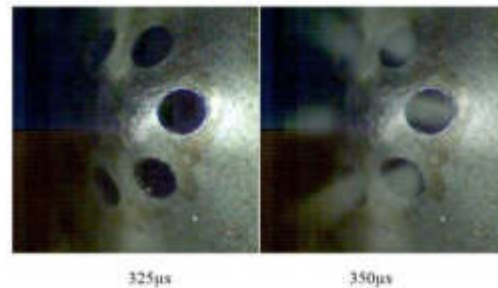
a. The moment of the start of the spray

3550μs   3575μs
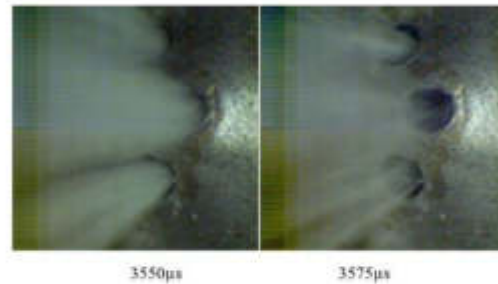
b. The moment of the end of the spray

Figure 3. Spray images when solenoid opens and closes.

There is a high correlation of the delay of opening and closing with solenoid current. What's more, the injector will not spray until current rate reaches to a threshold value. When injector spray is stopped, there will be a bump in current curve.
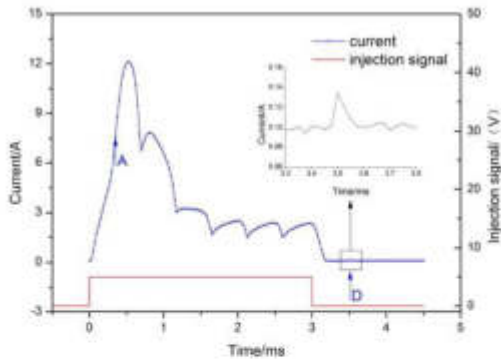
Figure 4. Detected signals in the operation process of injectors.
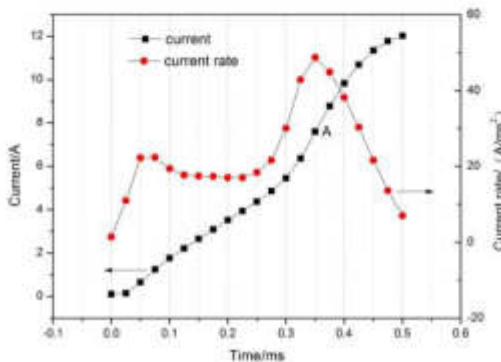
Figure 5. Detected signals in the operation process of injectors.

Compare Current Signal with Pressure Signal

Figure 6 shows the drive current of the injector and inlet pressure which are acquired by NI DAQ synchronous data acquisition system.
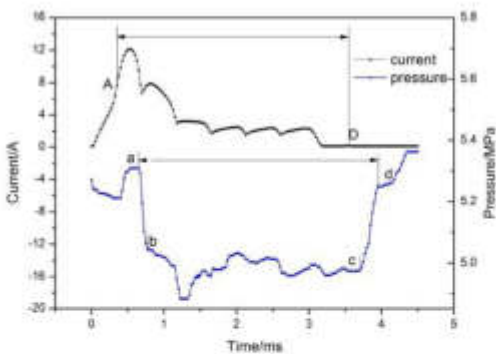
Figure 6. Drive current and pressure curve.

The upper curve shows the current change in the process of spray. The other curve is the synchronous inlet pressure of the injector. Savitzky-Gola method is used to process the pressure curve to eliminate noise. In the picture, with current incentive gradually increasing, injector inlet pressure doesn't immediately fall at the opening time of solenoid valves at A point, but presents a certain delay. The pressure gradually drops at A position. Moreover, when injector solenoid valves close, injection pressure has a delay time. The duration from A to D is equal to T_{off}. It means that the injection duration determined by current signal

is same with that by pressure signal. During the process of solenoid valves opening and closing, pressure fluctuation has less effect on pressure trends which are monotonic. Thus, it can be found that the time from a to d is the duration of solenoid valves opening; the time from b to c is solenoid valves fully opening; the time from c to d is the duration of solenoid valve closing. Corresponding fuel injection regulation can be got from the opening time and the closing time of solenoid valves determined by current signal, pressure change and injector flow characteristics.

Effect of Fuel Injection Pressure on GDI Injector Response

In the case of a certain control strategy, the characteristics of solenoid valves under different fuel injection pressures are compared. Figure 7 shows the delay time under different injection pressures.
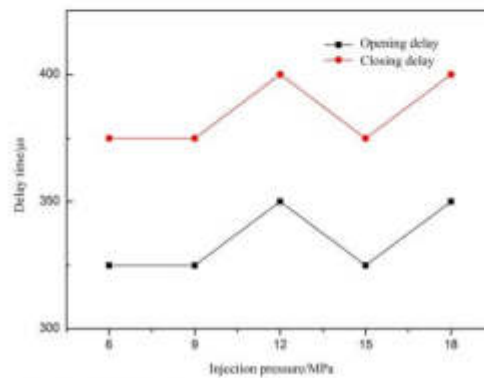
Figure 7. Delay time under different injection pressures.

In the figure, there is a fluctuation in response to the needle valves at different pressures. The fluctuation value is 25 μs, but the duration of overall injection does not change. In short, the fuel injection pressure corresponding to GDI solenoid valve injectors is very small. Compared with fuel duration, it is negligible.

HPCR Injector

Compared with the GDI injector, the HPCR injector is more complex. It includes power delay, mechanical delay and hydraulic delay. The solenoid valve and the needle value are two key components during spray process. In GDI injection system, when solenoid valves move, injection starts. However, in HPCR injection system it is just a prerequisite for spray.

Spray Images and the Corresponding Signals

High-speed solenoid valve injectors take advantage of "Peak-Hold" control mode. In the initial stage of injector pulse width signal, coil current is rapidly increased by instantaneous high voltage to improve the response speed of injector needle valves. When they open, the current in coils is reduced by PWM control to keep them open. Figure 8 shows current and current rate during injection process. Figure 9 shows pressure and its derivative at the inlet of the injector during injection process.

As shown in figure 8, the top curve means current signal; the middle one implies current rate; the square wave points pulse width signal. Position A is the start of injection signal, and position B is the end of

152

injection signal. Figure 9 shows pressure signal curve, first-order derivative and second-order derivative of pressure signal curve from top to bottom during single injection. At the moment of position A_g, the entrance pressure of the injector begins to decrease slowly. In figure 8, position A_g is the moment of solenoid valves opening when current rate reaches the maximum position. In this process, current is gradually increasing, and electromagnetic force becomes larger. When electromagnetic attraction equals to the sum of spring preload, pressure and the gravity of ball, solenoid valves begin to move. When they begin to move, electrical energy is rapidly converted into kinetic energy, and the current rate of coils is rapidly reduced. During the closing of solenoid valves, magnetic flux becomes small. At the same time, an induced electromotive force is generated in coils, which hinders coil current from becoming further smaller. Due to self-induced EMF effect, coil current becomes large, leading to a bump (position a_c) appearing in figure 3. T_{ao} is the delay of solenoid valve opening, and T_{ac} is the delay of solenoid valve closing. Thus, by analyzing the characteristics of current curve, T_{ao} and T_{ac} can be acquired. The conclusion is consistent with the research of the GDI injector.
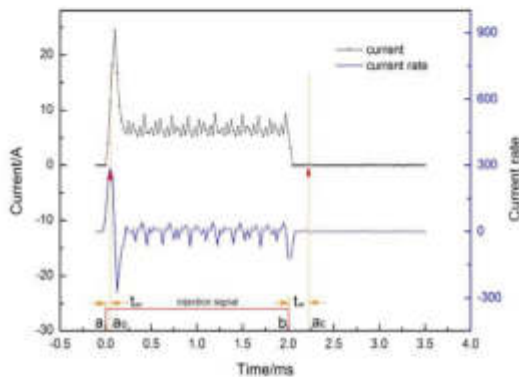
Figure 8. Current and current rate during injection process.
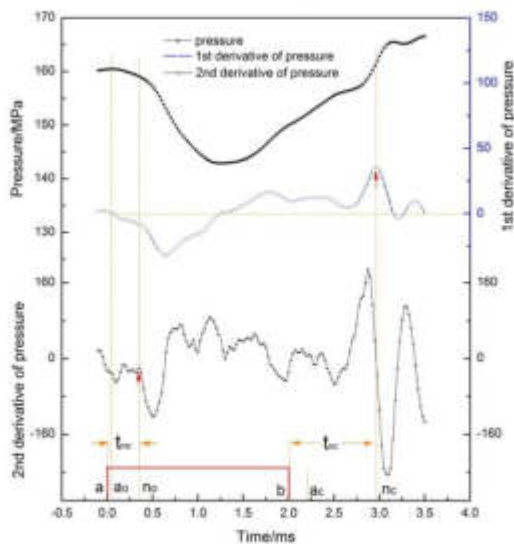
Figure 9. Pressure and its derivative during injection process.

In figure 9, the needle opens at the moment of n_o, and the entrance pressure begins to decrease fast. The second-order derivative of pressure signal curve shows that change rate begins to increase gradually at the moment of n_o when the injector begins to inject. In figure 9, the needle closes at the moment of n_c, and the first-order derivative of pressure signal curve reaches the maximum at the same time. The injector is stopped to inject at the moment of n_c. Thus, t_{no} is the delay of needle opening, and t_{nc} is the delay of needle closing. Moreover, actual injection duration is from n_o to n_c. Figure 10 shows some spray images when the needle valve opens and closes.
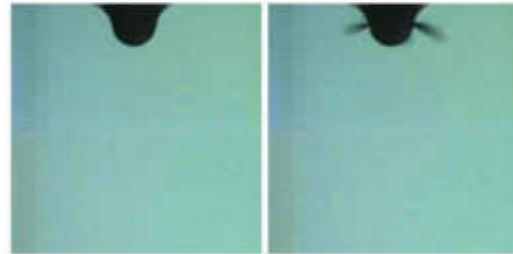
a. The moment of start of spray

b. The moment of end of spray

Figure 10. Spray images when the needle valve opens and closes.

Effect of Fuel Injection Pressure on HPCR Injector Response

Injection pressure gradually increases from 40 MPa to 160 MPa, and dynamic characteristics are studied. Figure 11 shows opening delay and closing delay under different injection pressures. Figure 12 shows actual injection duration under different injection pressures.
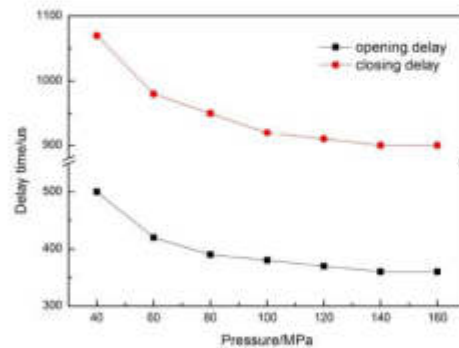
Figure 11. Pressure and its derivative during injection process.

153

As shown in figure 11, with the increase of injection pressure, the opening delay and the closing delay of needles are gradually becoming short, and tend to balance. Unlike the GDI injector, the change in fuel injection pressure has significantly impacted on the HPCR injector. When fuel injection pressure is less than 100 MPa, the opening delay and the closing delay of injector needle valves are large, and hydraulic delay is the main reason.
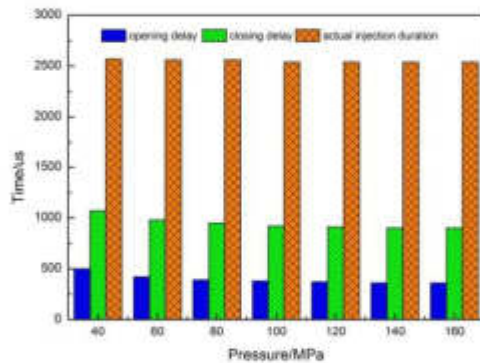
Figure 12. Pressure and its derivative during injection process.

As shown in figure 12, although fuel injection pressure gradually changes, actual injection duration keeps the same trend, and it is longer than injection pulse width. The test result is consistent with GDI injector result.

Conclusions

A new non-contact detection method is put forward to research the dynamic response of electronic injectors with high-speed micro-photography and synchronous current-pressure collection system, which solves the problems of electromagnetic interference and structure arrangement.

There is a high correlation of the delay time of the opening and the closing of solenoid valves with solenoid current. The solenoid valves will not open until current rate reaches to a threshold value. When they close, there will be a bump in current curve. On basis of this, the dynamic characteristics of solenoid valves such as opening delay and closing delay can be obtained for the GDI injector. As for the HPCR injector, by monitoring and analyzing the pressure of the injector inlet, injection duration can be further decomposed into opening process, full-open process and closing process.

For GDI and HPCR injectors, actual injection is longer than injection pulse width. With the change of injection pressure, actual injection always keeps the same trend. For dynamic response characteristics, the HPCR injector is more sensitive to pressure change.

References

1. Yao, M., Wang, H., Zheng, Z., & Yue, Y. "Experimental study of n-butanol additive and multi-injection on HD diesel engine performance and emissions," *Fuel*. 89(9), 2191-2201,2010, doi:org/10.1016/j.fuel.2010.04.008.

2. Chiodi, M., Mack, O., Bargende, M., Paule, K. et al., "Improvement of a High-Performance Diesel-Engine by means of Investigation on different Injection Strategies," SAE Technical Paper 2009-24-0008, 2009, https://doi.org/10.4271/2009-24-0008.

3. Dober, G., Tullis, S., Greeves, G., Milovanovic, N. et al., "The Impact of Injection Strategies on Emissions Reduction and Power Output of Future Diesel Engines," SAE Technical Paper 2008-01-0941, 2008, https://doi.org/10.4271/2008-01-0941.

4. Ghaffarpour, M., & Noorpoor, A. R., "NOx reduction in diesel engines using rate shaping and pilot injection," *International journal of automotive technology and management*, 7(1), 17-31,2007, dx.doi.org/10.1504/IJATM.2007.013381.

5. Vanegas, A., Won, H., Felsch, C., Gauding, M. et al., "Experimental Investigation of the Effect of Multiple Injections on Pollutant Formation in a Common-Rail DI Diesel Engine," SAE Technical Paper 2008-01-1191, 2008, https://doi.org/10.4271/2008-01-1191.

6. Han W Y, Xu S C, Zhou Y K, et al, "Start of Injection Impacts on Performance of a Di-rect-injection Spark-ignition Engine," *Journal of Tongji University (Natural Science)*, 41(4): 565-582, 2013.

7. Lee, J., Min., K., Kang, K., and Bae, C., "Hydraulic Simulation and Experimental Analysis of Needle Response and Controlled Injection Rate Shape Characteristics in a Piezo-driven Diesel Injector," SAE Technical Paper 2006-01-1119, 2006, https://doi.org/10.4271/2006-01-1119.

8. Oh, B., Oh, S., Lee, K., and Sunwoo, M., "Development of an Injector Driver for Piezo Actuated Common Rail Injectors," SAE Technical Paper 2007-01-3537, 2007, https://doi.org/10.4271/2007-01-3537.

9. Martin, D., Pischke, P., & Kneer, R., "Investigation of the influence of multiple gasoline direct injections on macroscopic spray quantities at different boundary conditions by means of visualization techniques," *International Journal of Engine Research*, 11(6), 439-454,2010, doi.org/10.1243/14680874JER525.

10. Passarini, L. C., & Nakajima, P. R., "Development of a high-speed solenoid valve: an investigation of the importance of the armature mass on the dynamic response," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 25(4), 329-335,2003, dx.doi.org/10.1590/S1678-58782003000400003.

11. Xia S F, Wang H M, Hou J L, "Research on Automatic Testing System of Injector Nozzle Body," *Tractor & Farm Transporter*, 2007,34(5):73-76.

12. An S J, Ouyang G Y, "Theory and Experimental Study for Solenoid Valve of Electronic Controlled Inector," *Transactions of CSICE*, 21(5):356-360,2003, 10.16236/j.cnki.nrjxb.2003.05.015.

13. Wang J, He L, Zhang Y T, "Automatic Measurement of Needle Lift in Electronic Controlled Fuel Injector," *Journal of Academy of Armored Force Engineering*,27(3):31-34,2013.

14. Zheng J B, Miao X L, Hong J H, "Experiment on Pressure of Nozzle Volume and Control Chamber inSolenoid Injector of Common Rail Sys-temOuyang Guangyao," *Transactions of CSICE*,31(2):171-176,2013, 10.16236/j.cnki.nrjxb.2013.02.016.

154

15. Zheng J B, Miao X L, Hong J H, "Measurement and Analysis of Pressure of Nozzle Volume in Electric Solenoid Common-Rail System," *Transactions of CSICE*,30(1):86-90,2012, 10.16236/j. cnki.nrjxb.2012.01.015.

16. Moon, S., Gao, Y., Wang, J., Fezzaa, K., & Tsujimura, T., "Near-field dynamics of high-speed diesel sprays: Effects of orifice inlet geometry and injection pressure," *Fuel*, 133, 299-309,2014, doi.org/10.1016/j.fuel.2014.05.024.

17. Payri, R., Salvador, F. J., Marti-Aldaravi, P., & Martinez-López, J., "Using one-dimensional modeling to analyse the influence of the use of biodiesels on the dynamic behavior of solenoid-operated injectors in common rail systems: detailed injection system model," Energy Conversion and Management, 54(1), 90-99,2012, doi.org/10.1016/j.enconman.2011.10.004.

18. Mulemane, A., Han, J., Lu, P., Yoon, S. et al., "Modeling Dynamic Behavior of Diesel Fuel Injection Systems," SAE Technical Paper 2004-01-0536, 2004, https://doi.org/10.4271/2004-01-0536.

Contact Information

Chunde Yao, Ph.D., Professor
Tianjin University
arcdyao@tju.edu.cn

Acknowlegements

Definitions/Abbreviations

NO_x - Nitrogen Oxides

DI - Gasoline Direct Injection

HPCR - High Pressure Common Rail

DAQ - Data Acquisition System

ECU - Electronic Control Unit

PWM - Pulse Width Modulation

TTL - Transistor Transistor Logic

DMCC - Diesel Methanol Compound Combustion

[9]   Bianchi, G., Pelloni, P., Filicori, F., Vannini, G., "Optimization of the Solenoid Valve Behavior in Common-Rail Injection Systems," SAE Technical Paper 2000-01-2042, doi: 10.4271/2000-01-2042, 2000

**SAE TECHNICAL
PAPER SERIES**

2000-01-2042

Optimization of the Solenoid Valve Behavior in Common-Rail Injection Systems

G. M. Bianchi and P. Pelloni
DIEM - University of Bologna

F. Filicori
DEIS - University of Bologna

G. Vannini
University of Ferrara

156

GLOBAL MOBILITY DATABASE

All SAE papers, standards, and selected books are abstracted and indexed in the Global Mobility Database

Optimization of the Solenoid Valve Behavior in Common-Rail Injection Systems

G. M. Bianchi and P. Pelloni
DIEM - University of Bologna

F. Filicori
DEIS - University of Bologna

G. Vannini
University of Ferrara

ABSTRACT

Injection rate profile is a powerful tool to control engine performance and emission levels. In particular, the Common Rail (C.R.) injection system has allowed flexible fuel injection in DI-diesel engines by permitting a free mapping of the start of injection, injection pressure, rate of injection and, in the near future, multiple injections,. This paper deals with improvements of stable operating condition limits of the Common Rail injector for multiple injection purposes. The focus was to optimize the behavior of the solenoid valve in order to reduce the minimum time interval between two consecutive injections required for system stability. An extensive experimental characterization of the valve has been performed in order to measure the main mechanical and electrical parameters of the assembly components. The experimental and the numerical studies have allowed optimizing the current profile and consequently the design of the anchor pin-ring assembly of the solenoid valve. A new solenoid valve electronic driving circuit has been designed for this purpose. This optimization allowed reducing the minimum valve switching time of about 300 μs. Numerical expectations have been confirmed by the first application of the new driving circuit in valve driving.

INTRODUCTION

D.I. diesel engines represent the most promising powertrain due to their high fuel conversion efficiency once their NOx and soot emissions are kept within the stringent limits imposed by regulations. Common Rail (C.R.) injection systems have offered remarkable improvements in emission and noise reduction by a free mapping of the injection profile [1-8]. To meet the further stringent requirements of EURO IV standards, the technology of the injection actuation and control must be improved. Recently, the trend toward increasing boost pressure and injection pressure through the use of C.R. injection systems allowed steps forward in meeting the emissions standards. On the other hand, today, the simultaneous reduction of NOx and soot represents a challenge for engine designers. Many emission technologies are found to reduce NOx while increasing soot or vice versa. C.R. injection systems revealed to be a effective means to achieve this goal, thanks to the independence of injection pressure on engine speed, flexibility in mapping injection timing and duration, feasibility of high injection pressure, and the possibility to implement multiple injections [2,3-4,8]. Common multiple injections are currently oriented to implement pilot, main and post-injection [8]. Nevertheless, limitations are still present in achieving stable and repeatable control of the small amount of fuel injected during pilot and post-injections, that was found to dramatically influences the combustion process and emissions [4,8]. Promising strategies in limiting the engine-out emission levels are represented by multiple injections, which split the main injection into two or more pulses, whose timing and duration are optimized depending on combustion and injection systems characteristics [9-13]. The multiple injection strategy is currently limited by the performance of the conventional electro-injectors adopted in mass production engines [7,8,14]. In the Common-Rail injection systems, the needle is actuated by a solenoid driving valve which controls the pressure on the upper and lower needle seats. The major drawback of solenoid valve actuation resides in the compromise that must be found between supplying a minimum quantity of fuel during pilot injection and guaranteeing a high enough injection rate during the main injection. The latter requirement, that is obtained by fully opening the needle, determines a minimum characteristic operating time of

1

158

the injector. This problem is compounded at high injection speeds, where the total combustion duration is too short to allow for multiple injections. In order to overcome this drawback, Renner et al. [4] have already presented promising results controlling the injection timing by means of piezo-electric actuators instead of the conventional solenoid valves. The piezo-electric actuator has the main advantage to drastically reducing the time of valve opening, as well to exert higher force levels, which can be used to increase the injection rate. Since the latter technology is still far from being feasible for mass production, efforts are currently spent in improving the conventional solenoid valve dynamics. Contribution in this area have been brought by Ganser [14] who conceived a double stop solenoid valve. The results showed that the proposed design allowed a better control of the initial shape of the injection rate.

The aim of this work is to optimize the electric control of the solenoid valve in order to reduce the minimum switching time for multiple injection purposes. The study has been performed by developing an electro-mechanical model able to simulate the valve behavior. Mechanical dynamics and electromagnetic circuits have been experimentally characterized. Model predictions in terms of predicted current profile and magnetic force at different air-gap were compared versus measurements. The application of the model has revealed that a better control of the valve opening and closing may be achieved by optimize the current profile and the anchor-pin assembly. A more flexible electronic circuit for the valve driving has been conceived and designed which allows more suitable current profiles aimed at reducing valve switching-off time. The promising numerical results were confirmed by the first application of the improved solenoid valve driving circuit.

CURRENT COMMON RAIL INJECTOR DESIGN

The injector considered in this research, shown in Figure 1, is the mass production C.R. injector for automotive D.I. diesel engines. The injector body presents an accumulation volume, in the bottom needle seat, and a control volume, in the upper part of the needle seat. Both volumes are fed at the same rail pressure. In the closed needle position, the two volumes are fed at the same pressure and the resulting force on the needle cancel each other, and hence the needle is pushed down by the spring force. The needle opening is driven by the solenoid valve placed in the upper part of the injector body and presented in Figure 2. The solenoid valve controls the pressure in the control volume and behaves as an on-off valve. By actuating the solenoid valve, the anchor is attracted by the excited solenoid coil and is pushed up. The anchor, in contact with the pin by means of the spring S2, moves the pin up, determining the opening of the ball valve and, hence, a fast pressure drop in the control volume. This results in a fast lift of the needle from its seat and, in turn, in hole opening.

Injection ends at the end of the electric excitation. From this instant the anchor-pin assembly is pushed down by the force of the spring S1 determining the ball-valve closing and therefore the raise of the pressure in the control volume. As a result the needle is pushed down by hydrodynamic forces, closing the holes.
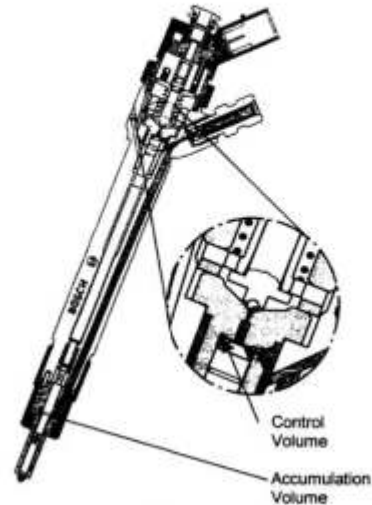
Figure 1. Common rail Injector with enlarged view on solenoid valve [8]
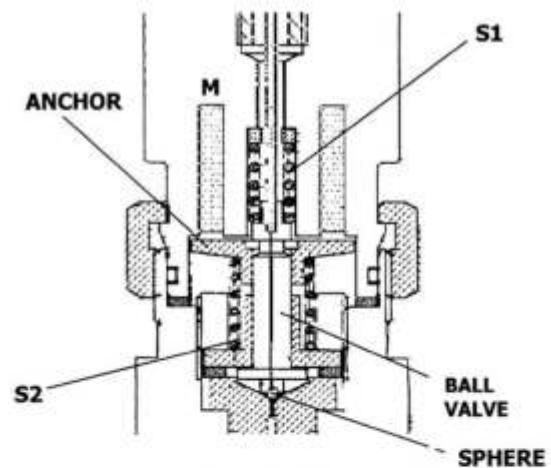
Figure 2. Schematic sketch of the solenoid valve

In the configuration illustrated, the anchor and pin are separated components. This design of the valve assembly replaced a former one where anchor and pin constituted a single component. This update was necessary to reduce the probability of the ball-valve re-opening caused by the valve bouncing back against its stop because of its non-zero impact velocity. The two-component-configuration allows reducing the impact energy since the kinetic energy of the pin can be absorbed by the spring.

2

159

The design of the mass production injector currently on the market brings the following limitations [16]:

1. Minimum elapsed time between two consecutive injections equal to 1.8 ms at fully opened valve.

2. Delay between current impulse and start of valve lift of 90 µs.

3. Minimum valve fully opening time equal to 300 µs.

CONTRIBUTION OF THE WORK

In this work the authors aimed to reduce the minimum switching time through the definition of a proper current profile. In particular the work aimed to:

1. Reduce the valve opening time

2. Eliminate the extra-displacement of the pin

3. Reduce the kinetic energy of the moving elements at the end of the lift and therefore the impact kinetic energy.

4. Design a new improved electronic control circuit.

The research has been performed by using an integrated numerical-experimental approach. In the first phase, an electrical-mechanical model has been developed and experiments have been performed in order to evaluate the electro-mechanical parameters used in the model. In the second phase, the model has been applied to the simulation of the solenoid valve solely. Numerical predictions have been compared to the experimental measurements of anchor. In a third phase a new electric driving circuit has been designed and applied to the control of the solenoid valve in order to check the feasibility of the numerical optimization in actual operating conditions. Finally, an electro-mechanical optimization of the solenoid valve dynamic has been performed. The capability of the new electronic circuit in reproducing the optimized numerical current profiles has been experimentally tested.

MEASUREMENT BENCH

In order to validate the electro-magnetic model, measurements of the inductance and magnetic force acting on the anchor at different supplying currents for fixed air-gap values were performed on a test bench. A micrometric feed screw was used to set the anchor-position and therefore the air-gap value. The test bench is presented in Figure 3. The displacement of the anchor has been brought outside the solenoid valve by using the hole used for fuel return to the tank when the ball valve opens. The force has been measured by using a strain-gage load cell with a primary temperature compensation. The cell is supplied at a constant tension of 5 V. This means that magnetic force has been measured for a fixed anchor position or, in other words, for a fixed air-gap value. In this first analysis the diesel fuel presence inside the valve was not considered since the interest was only in evaluating the magnetic force.

The inductance was measured by using an HP bridge supplied at 100 mA current at 100 Hz frequency. The value of the current has been chosen considering the compromise between avoiding saturation and neglecting the magnetic hysteresis effect. The frequency upper value was limited by the need to keep the losses in the iron small, because they are neglected in the model. Because of uncertainties below 50 µm air-gap, the measurements of the inductance and magnetic force have been carried out for air-gap values larger than 100 µm. These uncertainties were attributed to the difficult pin centering below 100 µm of displacement. Reasons for this were found in the lack of hydrodynamic forces in the absence of fuel.

Figure 3.   Test bench

SOLENOID VALVE DYNAMIC MODEL

In order to study the behavior of the solenoid valve, an electrical-mechanical numerical model has been developed. Equations have been solved using the Matlab-Simulink environment. Based on the previous description, the dynamic electro-mechanical behavior of the solenoid valve can be modeled by using a simple-forced harmonic oscillator [7,17], described by a one dimensional momentum equation:

$$m\ddot{x} + c\dot{x} + (F_o + kx) = F_T \qquad (1)$$

where m is the mass of the moving components, c the coefficient of viscous damping and k the spring elastic constant, F_o is the total spring pre-load, and F_T is the sum of the force applied to the mass along the valve axis. Since the model considers only the dynamics of the solenoid valve disconnected from the injector, the contribution of the force induced by pressure in the control volume is not considered. The friction between the anchor and pin are neglected.

In order to accurately model the anchor-pin assembly movement, the system dynamics have been split into three phases:

A. In the first one, the simulation starts from the instant of the electric excitation and ends when the anchor impacts against the stop-block, i.e., at zero air-gap

3

160

condition. The anchor and the pin are assumed to behave as one component because of spring S2, which keeps them always in contact. Hence, the velocity of the pin is assumed to be the same as that of the anchor. The mass m' of the system is the sum of the mass of both pin and anchor. The force of the spring S1 and spring S2 is accounted for by using an equivalent stiffness k'. The governing equation assumes the form

$$m'\ddot{x} + c\dot{x} + (F_o + k'x) = F \tag{2}$$

where the force F is the magnetic force that varies as a function of the air gap, i.e., as function of anchor displacement x. The variable dumping coefficient c is used as a shock absorber to simulate the impact force arising when the anchor kicks the stop-block at x=50 μm (i.e., zero air-gap condition). As it will be pointed out later, the actual current profile, visible in Figure 6, presents a second peak aimed at reducing the anchor bounce effect after impact.

B. In the second phase (i.e., for pin displacement x larger than 50 μm), the anchor is assumed to remain stopped against the block (at x= 50 μm, at zero air-gap condition), while the pin continues moving because of its inertia converting its kinetic energy into the elastic energy of spring S1. When this energy transfer is completed, the spring releases its elastic energy pushing down the pin again until it comes back in contact with the anchor (i.e, at x= 50 μm). Hence, the system consist of the pin and the spring only, and its mass is m''. The initial velocity of the pin is assumed to be the final velocity of the anchor-pin assembly at the end of phase A. Hence, the governing equation is:

$$m''\ddot{x} + c\dot{x} + (F_{o,S1} + k_{S1}x) = 0 \tag{3}$$

where x is the pin displacement, m'' is the mass of the pin, $F_{0,S1}$ and k_{S1} are the pre-load and the stiffness of the spring S1. Internal spring dumping effects are included. In this phase the second spring S2 does not play any role since it pushes the anchor against the seat.

C. The third phase begins when the pin, pushed down by the spring S1 hits the anchor at x=50 μm, pushing it down, neglecting anchor bouncing back. This hypothesis is reasonable for the modeling purposes since the second current peak visible in Figure 6 induces an increase of the magnetic force which keep the anchor at its upper stop position (i.e., x=50 μm). Then, after the contact has occurred, the anchor which is kept in contact with the pin by the spring S2. Hence, the velocity of two components, assumed to behave as only one, is evaluated from a momentum balance. The mass m' of the system is the sum of the mass of both pin and anchor. The force of the spring S1 and spring S2 is accounted for by using an equivalent stiffness k'. The model assumes the form

$$m'\ddot{x} + c\dot{x} + (F_o + k'x) = F \tag{4}$$

The variable damping coefficient c is used as shock absorber to simulate the impact force arising when the ball-valve, and therefore the pin, kicks the seat at x=0 μm. In the baseline configuration, the magnetic force F is equal to zero during valve closure operations. It will be different from zero in the optimization control profile proposed below.

EVALUATION OF THE MAGNETIC FORCE: THE ELECTRICAL MODEL

The evaluation of the magnetic force has been performed according to the following path. The magnetic force by which the anchor is driven is obtained by applying a current to a solenoid coil. This in turn induces a magnetic flux. The governing equation of the electrical circuit is:

$$v = ri + \frac{d\varphi}{dt} \tag{5}$$

where r is the electrical resistance, v the tension and φ is the total magnetic flux. Since the current i is a non-linear function of the anchor position x and magnetic flux (i.e., $i = f(x, \varphi)$), as proposed by Filicori et al. [16] for modeling the variable-reluctance motors, the current can be expressed as the sum of two terms: one, which depends on the reluctance, the other that represents a purely non-linear term:

$$i = J[x] \cdot \varphi + W[\varphi] \tag{6}$$

where $J(x)$ represents a position-dependent term associated to the air-gap reluctance and $W(\varphi)$ describes the non-linear effects in the iron part of the flux path related to magnetic saturation. The non linear term $W(\varphi)$ may be expressed by a monotonic function:

$$W[\varphi] = \sum_{n=1}^{K} \gamma_n \varphi^n \tag{7}$$

where γ_n represents the coefficients and the order of the expansion, respectively. The zero-order term has been neglected since the magnetic hysteresis is practically negligible in the case considered. Substituting Eq. (7) into Eq. (6), one obtains:

$$i = G(x) \cdot \varphi + \sum_{n=2}^{K} \gamma_n \varphi^n \tag{8}$$

where the first-order term of polynomial expansion has been included in the term G(x) [16]:

$$G(x) = J(x) + \gamma_1 \tag{9}$$

Finally, according to Filicori et al. [16], the magnetic force F may be evaluated as:

4

161

$$F = -1/2 [dG/dx] \varphi^2 \qquad (10)$$

EXPERIMENTAL EVALUATION OF THE ELECTRICAL PARAMETERS – Measurements have been performed in order to evaluate both the electrical and mechanical parameters. As regards the electrical part of the model, the linear and the non-linear terms G and W have to be measured. If small amplitude sinusoidal currents are applied, the non-linear term can be neglected since the iron saturation is almost absent. Hence, from Eq. (8) it follows that

$$i = G[x] \varphi \qquad if \; i \to 0 \qquad (11)$$

Since G(X) is the inverse of the inductance H, it can be easily evaluated from the measurement of magnetic inductance at different air-gap conditions (i.e, anchor positions). In Figure 4, the dependence of G(X) on the air-gap is presented for air-gap value above 100 μm. A linear approximation was adopted in the model to evaluate G(x) in the range of interest (i.e, 0<x<50 μm).

Since from Eq (6) one can explicitly determine the unknown non-linear term W(φ):

$$W[\varphi] = i - G[x] \varphi \qquad (12)$$

samples of the non-linear term W(φ) can be determined after the computation of G(x) from measurements at different current and flux at a fixed anchor position (i.e, W(φ) does not depend on x). The magnetic flux can be evaluated from force measurements since it is:

$$F[i] = K_f \varphi^2 [i] \qquad (13)$$

where

$$K_f = -\frac{1}{2} \frac{dG}{dx} \qquad (14)$$

From experiments the coefficients γ_n and the order K defining the polynomial expansion of W(φ) was obtained by means of the least-square method.

MODEL VALIDATION

Figure 5 compares measured and predicted magnetic forces versus current at 100 μm of air-gap. As shown in Figure 5, the model performs quite well providing an estimation of the force that is in good agreement with respect to the measured values in a wide range of currents.

Since this check might not be sufficient to bring confidence in the model, a simulation of solenoid valve dynamics was performed by applying the measured current profile shown in Figure 6. The predicted and

measured anchor lifts are compared in Figures 7, which reveals that the anchor lift profile is reproduced with reasonable accuracy.
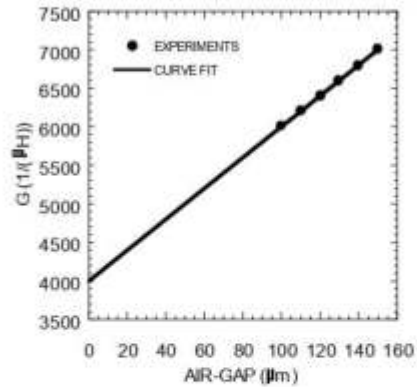
Figure 4.   Measured G(x) versus air-gap
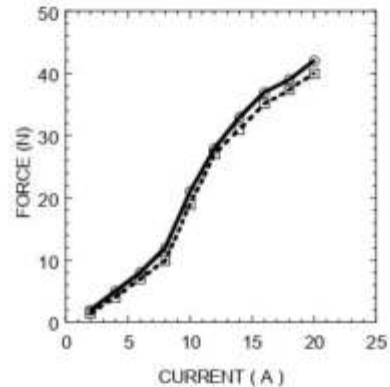
Figure 5.   Predicted and measured force versus current at 100 μm of air-gap
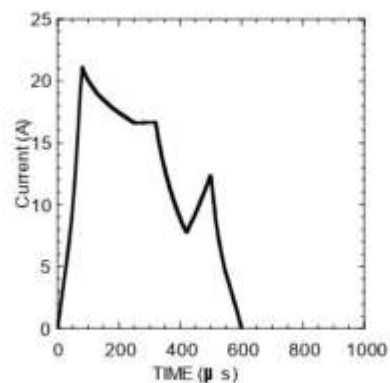
Figure 6.   Measured (actual) current profile

5

162
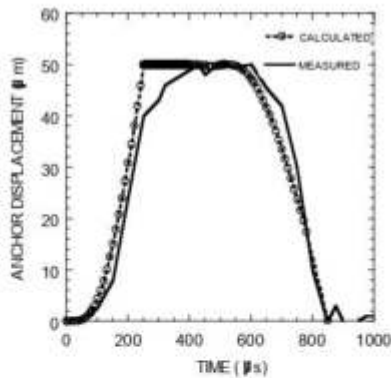
Figure 7. Comparison between predicted and measured anchor displacement
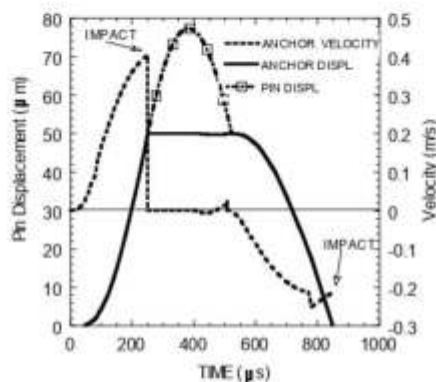
Figure 8. Predicted dynamic behavior of valve: anchor and pin displacement and anchor velocity

NUMERICAL ANALISYS OF VALVE DYNAMICS

The previous numerical analysis could be useful to investigate the valve behavior and to obtain some information for optimization purposes. Figure 8, where the velocity profile of the anchor is shown versus time, reveals that two impacts occur. The first one takes place during the opening phase, when the anchor stops at the blocks with a velocity of about 0.4 m/s. The second impact occurs during the closure operation when the anchor, and therefore the pin (i.e, we remind that in this phase the anchor and the pin are strictly in contact) end their lift with a velocity of about 0.2 m/s. Since the pin pushes down the ball-valve, this causes the ball-valve to impact with a considerable kinetic energy against its seat.

The analysis also reveals that long time is lost in compressing and relaxing spring $S1$, after the anchor has stopped against the block (i.e., x=50 μm). Beyond 50 μm, the pin continues moving up to 78 μm until its kinetic

energy is totally converted into the elastic spring energy. Then the elastic energy is released again to the pin, which is pushed down. This determines a non-active time spent when the anchor and pin are not in contact. This extra-displacement of the pin of about 56 μm (split between up- and down-lift) results in a minimum non-active time and therefore in increasing the injection duration. In fact, the return of the pin takes a displacement time of 56 % longer than that required starting from the zero air-gap condition.

Another interesting issue is that the maximum electric tension used in the baseline valve is of the order of +30 V, which is relatively smaller in comparison with that admitted by the solenoid. Moreover, in the baseline tension control, only a small negative tension of the order of −20V is used to decrease the current. These two features result in a slower raise and decrease rates of the current, respectively. This contributes to increase further the solenoid valve switching time, which is found to be of about 850 μs (Figure 8), too long for multiple-injection purposes.

The question is whether this interval can be reduced in order to comply the stringent actuation time requirements of multiple injection strategies.

OPTIMIZATION OF THE SOLENOID VALVE

The previous analysis has shown that a more rational electronic control is needed in order to shorter the characteristic operating time. Moreover also the anchor-pin assembly design may be improved once the dynamics of the moving parts are optimized, i.e., the impact energy in opening and closure is reduced if not eliminated. The improvement path has followed three main stages. First, a new driving electronic circuit was designed; second, numerical simulations were performed in order to find the proper electronic driving strategies. Finally, experimental tests were carried out in order to check the capability of the electronic circuit in reproducing the desired current profile suggested by simulations.

THE ELECTRONIC CIRCUIT FOR EFFICIENT SOLENOID VALVE CONTROL – In order to improve the dynamic behavior of the solenoid-valve system, a suitable electronic circuit for the control of the current profile in the solenoid has been developed. The circuit, shown in Figure 9, controls the solenoid current by means of a two-quadrant, DC/DC step-down switching converter, consisting of the two Insulated-Gate Bipolar Transistors (IGBT) T1, T2, which are controlled by the Injection Control Unit (ICU) through the transistor driver circuits (DR), and the two diodes D1, D2.

6

163
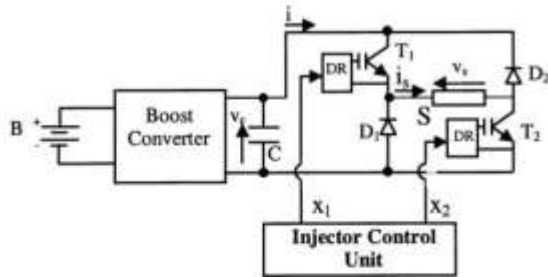
Figure 9. The electronic driving circuit for the solenoid valve actuation

The step-down converter is fed by a boost switching converter which provides an output voltage, v_c, much greater than the battery voltage. A value of v_c=100 V, which is compatible with the maximum voltage rating of the solenoid, has been chosen in order to allow for the high di/dt values required by the optimized current profile described in the following.

The capacitor C has been included not only in order to reduce the output voltage ripple of the boost converter, but also to store the electro-magnetic energy which is "recovered" by the step-down converter when a decrease in the solenoid current (i.e. a decrease in the energy stored in the magnetic field) is required, according to the current profile chosen. By choosing a suitably large capacitance value for C, the voltage v_c can be kept at a nearly constant value over the entire injection cycle.

Recovery and storage in the capacitor C of the magnetic energy stored in the solenoid, or of the mechanical energy of the moving equipment, is important in order to limit the average power consumption of the injection circuit. Energy recovery is possible since the two-quadrant step-down converter allows for bi-directional power flow, since the solenoid currents are only positive, but the solenoid voltage v_S can controlled, through Pulse-Width-Modulation (PWM) techniques, with positive or negative values in the range: $-v_c \leq v_S \leq v_c$

In particular, when both T1 and T2 are turned on (i.e. they behave almost as short circuits), the solenoid is directly connected to the voltage v_c, so that a large, positive voltage v_s=v_c is applied at the solenoid; in such conditions, according to Eq. (5) and Eq. (6), where the term ri is nearly negligible due to the small value of r, a large, positive $d\phi/dt$ (i.e. fast increase of the current and, consequently, of the solenoid force) is obtained. This means that the electric energy, taken from both the capacitor and the boost converter through the positive current i=i_s, is transformed into magnetic or mechanical force during the valve opening phase.

On the other hand, when one transistor is turned off (i.e. it behaves almost like an open circuit) the associated "free-wheeling diode" (i.e., D1 for T1 or D2 for T2) is turned on, so that the current solenoid is practically short-circuited, while no current i flows. In such conditions the magnetic flux and the corresponding current are nearly

constant, since $d\phi/dt \approx v_s \approx 0$, while no electric energy flows from the capacitor or the boost converter. Thus, the only energy transformation which can take place is the conversion of magnetic into or from mechanical energy during the opening or closing movements of the valve.

Finally, a third state can be forced on the step-down converter from the ICU, by turning off both the transistors and, consequently, by turning on both the diodes D1 and D2. This state involves a fast decrease of the solenoid current and force, owing to the large negative voltage $v_s \approx -v_c$. In such conditions, owing to the negative current $i=-i_s$, magnetic and, possibly, also mechanical energy (during the closing phase of the valve) is converted into electric energy and stored in the capacitor C.

The possibility of freely imposing at any time one of the three above-described possible states for the step-down converter allows for very flexible control of the solenoid current and force, so the optimized current profile described in the following can be directly implemented by suitably controlling through the ICU the circuit in Figure 9. Moreover, owing to the intrinsic circuit capability of recovering and storing energy, a relatively small average power consumption is involved although the peak powers required for fast valve opening and closing are quite high, due to the need for fast energy transfer.

OPTIMIZATION OF THE CURRENT PROFILE – Once a more flexible electronic circuit was designed, improvements in solenoid valve control was been conceived. The first step was aimed at eliminating the extra-displacement of the pin by controlling the approaching velocity of the anchor-pin assembly close to the zero air-gap condition. In particular, the latter goal was achieved by reducing the current rapidly and therefore reducing the magnetic force close to the end of the anchor lift. Since a reduction of the velocity might increase the valve switching time, a faster initial current raise obtained by applying higher voltages could compensate this drawback. The consequent faster initial lift might compensate the slower motion of the anchor-pin assembly approaching the zero air-gap condition. In order to increase the magnetic force, which depends on current time-derivative, a much higher tension (i.e., 100 V, as previously said) was adopted. The new current profile was found to comply to material specifications.

The braking action on the moving part of the valve is operated by a negative tension pulse of -100V close to anchor block (Figure 10). This high negative tension determines a higher negative value of the current time-derivative (Figure 11) and, in turn, a faster reduction of the magnetic force. Ending earlier the magnet excitation, the anchor-pin is subject to the breaking force of the contrast spring S1 only. By optimizing the timing and the duration of the negative tension pulse, it is possible to obtain a zero velocity at the end of the lift (50 µm). In this case, the return phase can start immediately because the spring can release immediately its elastic force without loosing time in absorbing, and then releasing, the

7

164

kinetic energy of the pin. Figures 11, 12 and 13 compare the two different control strategies in terms of current profile and pin-anchor displacement and velocity, respectively. The main result is to eliminate the impact during the opening phase and to drastically reduce, by 200 µs, the total valve switching time.

FULLY CURRENT-CONTROL CLOSING – The first benefit of this electronic control is to avoid the impact of the anchor against its block: this suggests that a new single-component design of the anchor-pin assembly is possible because the bounce back danger is eliminated. The new configuration is designed by eliminating the spring S1 whose aim was to mantain contact between anchor and pin. An overall weight reduction of 3.1 g was obtained. Despite the fact that dramatic improvements have already been obtained by optimizing the tension profile, the analysis of Figure 13 shows that the impact velocity (0.2 m/s) remains high and dangerous because the impact energy could damage the ball-valve seat. In the same fashion, it was conceived to drive the valve also during the closure phase by introducing a double tension pulse as visible in Figure 14. The second tension pulse induces a current pulse to 24 A (Figure 15), and therefore a magnetic force. The latter determines a braking action on the anchor pin-assembly opposite to the pushing spring force reducing the final velocity close to zero (0.07 m/s), as shown in Figure 17. It must be highlighted that this velocity is also the velocity of the ball-valve at impact against its seat. The use of a re-designed pin valve assembly balances the slight increase of the valve switching-off time resulting from the current-driven closure. As overall result, by combining the reduction of valve moving part weight and the fully electronic control the opening phase could be further reduced by 10 µs and the closing phase by about 30 µs. Therefore, the overall switching time for a fully opened operation was reduced by about 30%, from 850 µs to 610 µs (Figure 17). Finally, an experimental investigation was carried out in order to check the capability of the new driving circuit. Figure 18 shows the comparison between the optimized measured and predicted current profile.

It is worth nothing that the numerical optimization is feasible and that the driving circuit performance agrees with the simulations. In particular the new electronic-circuit appears to be able to reproduce with good agreement the double-pulses numerical current profile. The actual timing of the pulse matches the predicted one giving raise to encouraging expectations when experiments are be performed on actual injector supplied with diesel fuel.
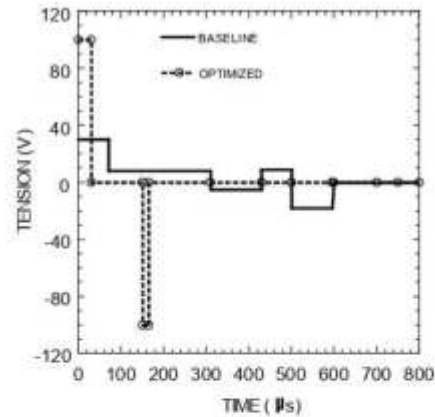
Figure 10. Baseline and optimized tension profiles for controlling anchor-pin velocity.
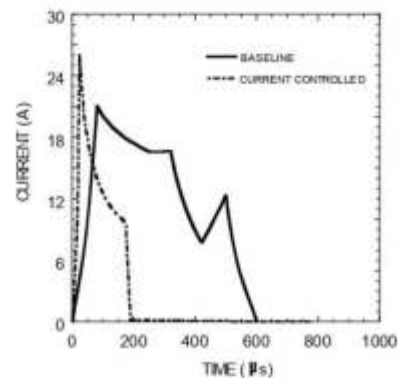
Figure 11. Baseline and optimized current profile
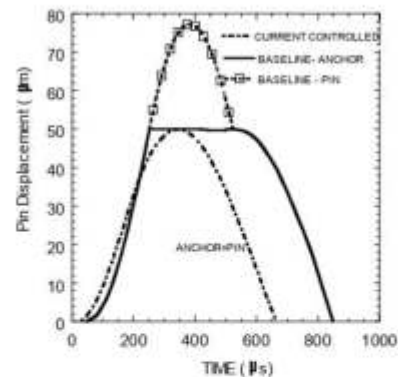
Figure 12. Baseline and current-controlled anchor-pin displacement

8

165
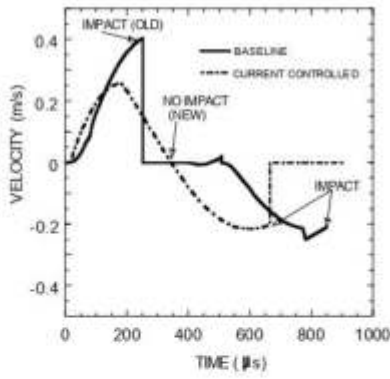
Figure 13. Baseline and current-controlled anchor-pin velocity
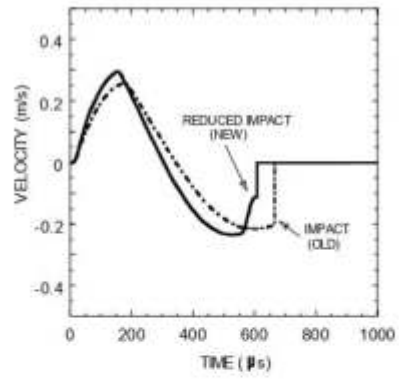
Figure 16. Current-controlled and fully optimized anchor-pin velocity
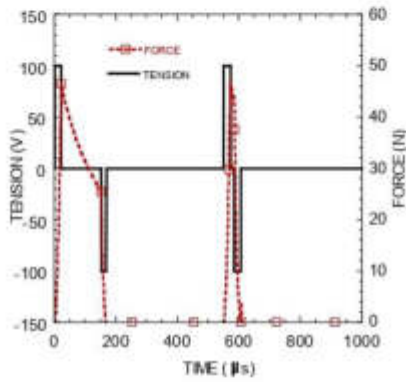
Figure 14. Fully current controlled valve: predicte tension and magnetic force
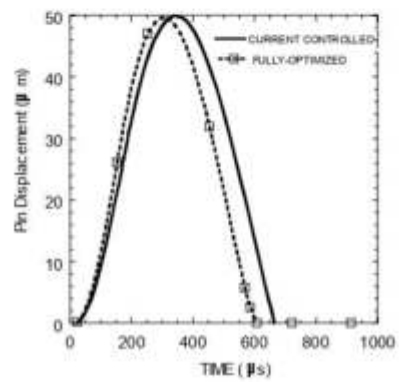
Figure 17. Current-controlled and fully optimized anchor-pin displacement
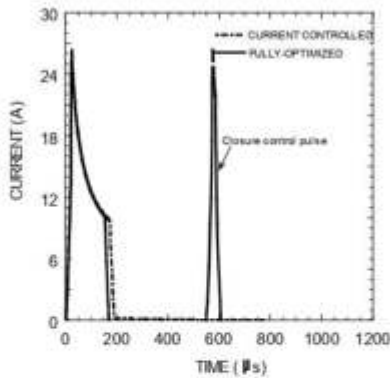
Figure 15. Single- and double-current pulse valve control
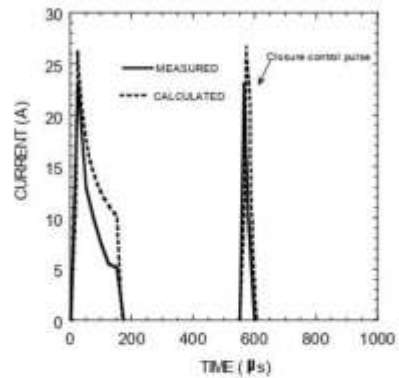
Figure 18. Comparison between predicted and measured current profiles in double-pulse valve control

9

166

CONCLUSIONS

The aim of this work was to optimize the electric control of the solenoid valve in order to reduce the minimum switching time for multiple injection purposes. The study has been performed by developing an electro-mechanical model able to simulate the valve behavior. Mechanical dynamics and electromagnetic circuit performance has between properly characterized.

The application of the model has revealed that a better control of the valve opening and closing may be achieved by optimizing the current profile and the anchor-pin assembly.

An improved more flexible electronic circuit has been developed in order to define suitable tension and current profiles for valve driving purposes.

Simulation has revealed that a proper and optimized electronic control of both the opening and closure phases allowed reduction of about 30% in the operating switching time of the solenoid valve.

This strategy eliminates the impact problems and therefore it expected to improve the long-term reliability of the components with particular emphasis on ball valve

The promising numerical results were confirmed by the first experiments where the solenoid valve was driven by means of the new electronic circuit.

REFERENCES

1. European Patent No. 0 426 205, 1993

2. Stumpp, G., Ricco, M., "Common-Rail – An Attractive Fuel Injection System for Passenger Car DI Engines", SAE Paper 960870, 1996

3. Boehner, W., and Hummel, K.,"Common Rail Injection System for Commercial Diesel Vehicles", SAE Paper 970567, 1997.

4. Renner, G., Koyannagi, K., and Maly, R.R.,"Effect of Common Rail Injector Design in the Emission Characteristics of Passenger Car DI Engines", Proceedings of "The fourth International Symposium COMODIA 98", pp.477-482, 1998

5. Guerrassi, N., P. Dupraz,"A Common-Rail Injection System For High Speed Direct Injection Engines", SAE 980803, 1998.

6. Funai, K., Yamaguchi T., Itoh, S. "Injection Rate Shaping Technology with Common-Rail Fuel System (ECD-U2)", SAE Paper 960107, 1996

7. Flaig, U., Polach, W. And Ziegler, G., "Common Rail System (CR-System) for Passenger Car DI Diesel Engines; Experience with Application for Series Production Projects", SAE Paper 1999-01-0191, 1999.

8. Ficarella, A., Laforgia, D., and Landriscina, V., "Evaluation of Instability Phenomena in a Common Rail Injection System for High Speed Diesel Engines", SAE Paper 1999-01-0192, 1999.

9. Nehmer, D. A. and Reitz, R. D., "Measurements of the Effects of Injection Rate and Split Injections on Diesel Engine Soot and NOx Emissions", SAE Paper 940668, 1994.

10. Pierpont, D. A., Montgomery, D.T., Reitz, R. D.,"Reducing Particulate and NOx Using Multiple Injections and EGR in a D.I. Diesel Engine". SAE Paper 940897, 1994

11. Tow, T., Pierpont, D. A., Reitz, R. D.,"Reducing Particulates and NOx Emissions by Using Multiple Injections in a Heavy Duty D.I. Diesel Engine", SAE Paper 940897, 1994.

12. Chan, M., Das, S. Reitz, R. D.," Modeling Multiple Injection and EGR Effects on Diesel Engine Emissions, SAE Paper 972864, 1997.

13. Han, Z., Uludogan, Ali, Hampson, G.J and Reitz, R. D.,"Mechanism and NOx Emissions Reduction Using Multiple Injection in a Diesel Engine", SAE Paper 960633, 1996.

14. Ganser. M.A.,"Common Rail Injector with Injection Rate Control", SAE Paper 981927, 1998.

15. P. Digesu, Ficarella, A., Laforgia, D., Bruni, G., and Ricco, M.,"Diesel Electro-Injector: A Numerical Simulation Code", SAE Paper 940193, 1994.

16. Filicori, F., Lo Bianco, C.G., and Tonielli, A., "Modeling and Control Strategies for a Variable Reluctance Direct-Drive Motor", IEEE Transactions on Industrial Electronics, Vol. 40, No. 1, February 1993, 1993.

167

[10]   Kumar, A., Banerjee, N., Mohan, N., Ashok B., "Study of the injector drive circuit for a high pressure GDI injector," 2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE), pp. 1-6, doi: 10.1109/ICRAIE.2016.7939562, 2016.

Study of the Injector Drive Circuit for a high pressure GDI Injector

Abhishek Kumar
SELECT
VIT University
Vellore, IN

Abhijeet Kumar
SENSE
VIT University
Vellore, IN

Neelkantha Banerjee
SELECT
VIT University
Vellore, IN

Nitin Mohan
SELECT
VIT University
Vellore, IN

Ashok B.
SMEC
VIT University
Vellore, IN

Abstract—In order to meet future emission regulations, increase the power production and reduce the fuel consumption the modern SI engines uses the electronic actuators such as solenoid injector, idle air control actuator, valve timing actuator etc. Engine performance and emissions are mainly depending on the accurate control of both the injection timing and fuel injection quantity of the injector. In order to fully exploit the advantages of modern GDI injector, the electronic system should be capable of driving a high speed solenoid injector at a very fast switching frequency efficiently and with less power consumption. The electric driving circuit is required to be designed with a fast response and precise control. The most serious problem which decreases the injector performance is the response time delay of the injector. There are many reasons for the delay in opening and closing of electromechanical injector such as voltage drop, limitation of injector driving circuit, electromagnetic interference, etc.

In this paper, an injector drive circuit has been designed and analyzed for driving a high pressure injector. Also, PWM control using NE 555 timer has been introduced for quickening the cut-off response time. In the process of designing an injector drive circuit high frequency switching component like MOSFET, IGBT and power transistor has been used. Comparison of response time between these three injector driving circuit has been given using NI Multisim. Analysis of switching characteristic among the three injector driving circuit was carried out. In the end, switching characteristic and temperature sweep of all the three IDC have been plotted.

Keywords—Injector; Injector driver circuit; power electronics; injector characteristic; PWM control

I. INTRODUCTION (*HEADING 1*)

Gasoline Direct Injection (GDI) is a type of fuel injection system in which the highly pressurized fuel is injected directly into the combustion chamber of each cylinder through a solenoid fuel injector. The fuel is supplied to the injector via a common rail fuel line. Benefits of GDI engines as compared to the contemporary PFI (Port Fuel Injection) engines are enhanced fuel economy, improved transient response and better air/fuel ratio control[1]. The solenoid fuel injector is an electromechanical device, which is operated using electric power. It mainly comprises of an induction coil, which when energized by Direct Current (DC), behaves as an electromagnet and pulls the pintle valve open. In order to further improve the fuel economy, achieve superior driving performance, and meet the requirements of stringent emission regulations, it is necessary to develop a precise and efficient control algorithm for the fuel injector. A defined control over pintle valve opening and closing time duration can be achieved by employing a suitable IDC for the fuel injector [2-4]. Zhang et al. have also tried to optimize the driving circuit. They projected a coupled simulation of injector electromagnetic, pintle (needle) rigid body motion and computational fluid dynamics (CFD) model, to generate optimal values of 1st stage current, 1st stage on-time and 2nd stage current. Two-stage current shapes were found to be the optimal power strategy for driving the fuel injector under different supply pressures. Key parameters, such as the 1^{st}stage current, the 1st stage on-time, and the 2nd stage current were obtained according to the magnetic force requirement. It helps us to get a better analysis of the performance of the driving circuits [4]. Tsai and Yu studied the design of an electrical drive for the high-pressure GDI injector for a 500cc motorbike engine. They designed a programmable injector drive circuit, and simulated using PSpice environment. They concluded that the driving parameters (such as two constant pulse time and adjustable PWM duties) can be optimized by a predetermined current control algorithm [5].

Huang et al explained the design of current controlled drive circuit. They tried to optimize the injector drive circuit by controlling the current across the solenoid, which further increased the response speed of the valve. Experimental results show that current drive circuit is feasible and reliable to execute for real world application [6]. Hardik and Shrikruishna took up rather simple approach in designing a driver circuit and ended up with quite an efficient design. They employed a 555 timer configured in the monostablemode along with an IC TLP250 in their design. The output of the circuit was used to drive the injector. They concluded that their circuit was working successfully and efficientlyand their circuit was compatible with their injector [7].Johnhas put forward a simulation methodology that helps the injector drive designer overcome the challenges involved in designing a complex injector drive circuit. The results obtained from the simulation were compared with those of prototype hardware to establish the accuracy and suitability of this simulation process [8]. Qiang et al investigated the different types of power losses associated with a solenoid injector with the help of software simulation. They concluded that there were remarkable differences in the power losses and the performances of the injector, when it worked within different driven strategies. In view of the fact that power

losses cannot be directly measured, the simulation results were validated by comparing to experimental results [9].

Lu et al studied the dynamic characteristics of a high speed servo-hydraulic solenoid injector for diesel engine with different driving circuits and control methods. They also proposed a pre-energized control strategy founded on dual power supply in order to improve the response time of the injector. They concluded that the injection delay (both opening and closing) is inversely proportional to the supply voltage. They also added that the pre-energized opening strategy significantly improved the opening response time of the injector [10]. Yamakado et al have come up with a mass producible battery driven fuel injector, with a dual coil structure, that eliminates the need of voltage step-up circuitry. They have also investigated the dynamic characteristics and design concept of their proposed fuel injector. They concluded that the injector ensures reliable fuel injection and optimizes the flow characteristics by controlling the opening coil energization time according to the changes in the battery voltage, harness resistance and fuel pressure [11].

In this paper, an injector drive circuit for a High Pressure GDI injector has been investigated. We designed and simulated various injector driving circuits with the help of NI Multisim. We employed three different switching elements, Power MOSFET, Power Transistor and IGBT, and compared their performances. We tried to make the injector circuit response quicker by decreasing the closing time of the injector using PWM control in the last stage (i.e. holding current stage). Since the induction coil of the solenoid injector draws a lot of current during its operation, a significant amount of power losses is observed in the circuit. We have also included analysis of thermal dissipation in our simulation results. As we know, increase in the heat dissipation of the driving circuit leads to extra power loss and thus decreases the efficiency of the circuit and drains the battery quickly. Therefore, it is very important to minimize the heat dissipation in the circuit in order to make it more efficient.

This paper has been organized in the following sections.

Section 2 discusses the electrical modeling of the drive circuit parameters. It points out the advantages of current drive injectors over a voltage drive injectors for controlling the fuel injector. The third section talks about the working mechanism of current driven injectors. Section 4 talks about designing an injector drive circuit using three different switching elements, namely Power MOSFET, Power transistor and IGBT. Section 5 depicts the simulation results that we obtained after analysisof the injector drive circuits.

II.   ELECTRICAL MODELLING OF INJECTOR

In early days, EFI system used to prefer voltage-drive injectors in which there was an observable lag in opening and closing of valve, which resulted in less efficient dynamic control of flow.To decrease this delay time, current driven injectors are adopted. These injectors have many advantages over voltage-driven injectors such as faster response time and also protection of solenoid from over voltage. All these advantages of current-drive injectors directed the auto industries to espouse current-drive circuit.

Fuel Injectors are designed for fast response under high pressures, which requires careful selection of the power drive circuit. The power input to the injector creates a magnetic force to open the pintle valve. Since magnetic force is directly proportional to square of supply current therefore, the pintle will move only when magnetic forces are greater than opposing force (including spring and friction forces).Thus, it can be concluded that the threshold of the pintle movement depends on the threshold level of excitation current in the coil. In other words, fast response of injector is very much dependent on the threshold current of the driving circuit. Throughout this paper, fuel injector is mostly showed as RL load because of solenoid inside it. The governing equation for the simple resistor-inductor circuit (RL Circuit) using KVL as:-

$$V_{Inj} = i_{Inj}.R_{Coil} + L_{Coil}.\frac{di_{inj}}{dt} \tag{1}$$

Where,

V_{Inj} = the voltage across the coil
R_{Coil} = the resistance of the coil
L_{Coil} = the inductance of coil.

From the above equation (1) we can find the expression of current flowing through the coil.

$$i_{Inj}.(t) = \frac{V_{Inj}}{R_{Coil}}.(1-e^{\frac{-Rt}{L}}) = \frac{V_{Inj}}{R_{Coil}}.(1-e^{\frac{-t}{\tau}}) \tag{2}$$

Where,

$\frac{L}{R}$ = electrical time constant, τ

Applying ampere's law on the electrical subsystem of the fuel injector, the following expression for Inductance value of Fuel Injector is obtained:

$$L_{Coil} = \frac{N^2.\mu_1.\frac{\pi}{4}.d^2.h}{\frac{d.w}{2}+\Delta x.h} \tag{3}$$

Where,

N = the no of turns in coil
μ_1 = the permeability in the air
h = the pintle height
d = pintle disk diameter
Δx = the air gap
w = the non-magnetic strip width.

We obtained the geometric parameters ofcommonly used GDI injectors as presented in table 1[4].

Table 1Parameters for inductance calculation [4]

Injector Parameters	Notations	Value	Unit
Pintle disk diameter	d	10	mm
Pintle height	h	2.62	mm
Non-magnetic strip width	w	0.8	mm
Permeability	μ_1	1.25e-6	-

Number of coil turns	N	160	-
Air gap	Δx	0~0.1	mm
Calculated Inductance value using eqn. 3	L_{Coil}	1.54	mH

III.   DESIGN OF INJECTOR DRIVE CIRCUITS

Three types of driving circuit have been designed and simulated considering R L as fuel injector circuit :-
- Three-stage Power Transistor driving circuit
- Single pulse IGBT driving circuit
- Three-stage Power MOSFET driving circuit.

A brief analysis of all the three driving circuit is provided based on the simulation which we performed on NI Multisim.As already discussed in previous section, GDI fuel injector can be electrically modelled by series RL load.Therefore to increase the electromagnetic force, the resistance should be decreased to its minimum possible value. As a result, current through the solenoid coil would be increased, which would result in an increase in the electromagnetic force, yielding an increase in the opening speed of the pintle valve.

In modern Fuel injector driving system based on Power MOSFET, manufacturer mostly use three stage drive current those three stage drive current are:-
- Peak current stage
- Bypass current stage
- Holding current stage

Initially, peak current is high so it can be used to quickly open the solenoid valve. It helps the fuel injector by increasing the opening speed of valve. As soon as the pintle is opened to it maximum position, drive circuit decreases the drive current to bypass current, which is lower than peak current. It protects the modern fuel injector form transient high voltage which can lead to damage of the solenoid. After stabilizing the dynamic flow of fuel, drive current changes to stage (iii) i.e. holding current stage. In this stage, drive current is in the minimum level. It helps in keeping the valve in maximum opening position and also protects the injector from high current in long term. Furthermore, the closing speed of injector is very high because drive current is very low and decrease to zero instantly. This helps in decreasing the closing time and making the system more efficient and decreasing the overall response delay.

We took a different approach to make the injector drive system more responsive. PWM control has been introduced in the last stage i.e. holding current stage of the Three-stage Power MOSFET drive circuit to increase the closing speed of injector drive circuit.

A.   Three-stage Power Transistor circuit

Firstly, a three-stage power transistor circuit to control the fuel injector has been designed. As shown in Fig.1, pulse generator is used to create three different pulse widths to drive the power transistor in three different stages current for replicating the signal from ECU.
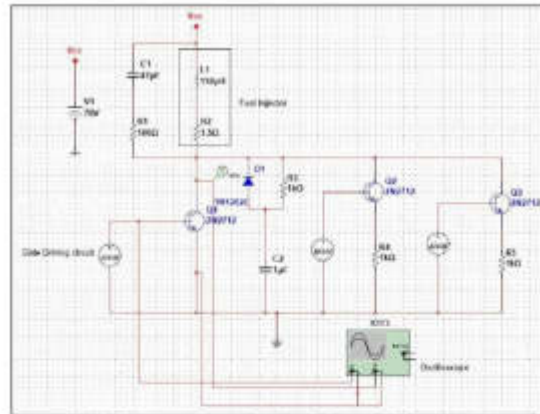
Fig.1: Three stage Power Transistor Injector Driving circuit

Resistance (R1) and capacitance (C1) across the Injector work as snubber circuit. It is added to protect the fuel Injector from abrupt voltage rise and thus prevent fuel injector from getting damaged. Oscilloscope is used to analyze gate signal and voltage across collector and emitter. Reason behind driving the circuit in three different currents (i.e. 14A, 9A, 6A) is to have better control over fuel injectors. It helps the IDC to have better and accurate control over fuel Injector.

B.   Single-Pulse IGBTs Based Injector Driving Circuit
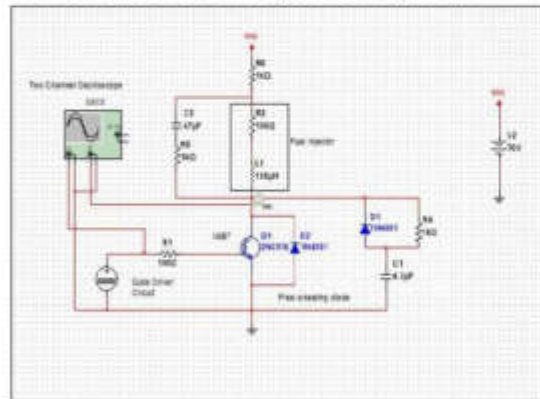
Fig.2: Single-Pulse IGBT Based Injector driving circuit

Fig.2 shows the single pulse IGBT based injector driving circuit that we designed using NI Multisim. In real time situation, a photocoupler IC 817 is also usually added with the injector driving circuit to avoid the noise interference from power electronics switching. We adopted proper measures to protect the solenoid from surge voltage (by means of snubber circuit) as well as the IGBT from over voltage(by means of a freewheeling diode).The purpose of adding diode D1 and

170

resistance R4 is to filter the signal and decrease the noise. It helps the ECU to have a better control over the IDC.

C. Three-stage Power MOSFET driving circuit

In order to have a better analysis and study on injector driving circuit for fuel injector, we decide to design one more circuit using Power MOSFET.
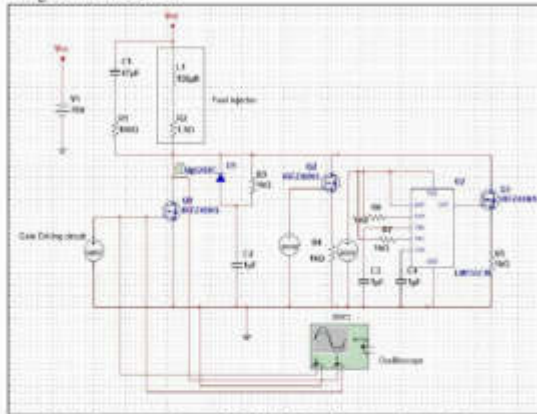
Fig.3: Three-stage Power MOSFET Based Injector Driving Circuit

We included three-stage Power MOSFET based driving circuit to have a better control on the injector as compared to other switching elements as shown in Fig.3. Three stage current for optimum performance are 14A (Peak Current), 9A (By-pass current) and 6A (Holding current). In our design, we have also included PWM control in last stage to decrease the closing speed and for making the system response faster. For our research, we are using pulse generator to simulate signal from ECU.

IV. SIMULATION AND ANALYSIS OF INJECTOR DRIVING CIRCUIT

Various driving circuit for the GDI Injector were designed and simulated. A driving circuit using three stage power transistors was designed and simulated using NI Multisim. The frequency was kept fairly high to simulate the environment of high-speed fuel injections. But the heat dissipation of the circuit was very high which lead to inefficient working of drive circuit and increase in power wastage across the circuit. Using this drive circuit can also lead to decrease in efficiency of the IC engine.

To improve the performance of the drive circuit, single pulse IGBT and three stage power MOSFET drive circuit were used to control the Fuel Injector. Abnormalities in the working of Single pulsed IGBT drive circuit were observed. This abnormal working can also lead to irregular fuel squirt from the injector or metal fatigue on the nozzle head. However, as observed from the simulation results Single pulse IGBT Drive circuits performed better as compared to the three-stage Power transistor driving circuit.

Performance of driving circuit was further improved by using power MOSFET, because of its high switching speed, and introducing PWM control in it. This controls the fuel injector precisely and more accurately as compared to the other two driving circuits. Hence, it can also help in increasing the efficiency of an IC engine. Simulation result of all three circuits is briefly analyzed below paragraph:-

A. Three-stage Power Transistor driving circuit

The circuit model of three stage power transistor circuit was built and simulated in NI Multisim platform (Fig.1). We have set different pulse conduction times to generate the three stage driving current and simulate three stage current (14A, 9A, 6A) in IDC. In this case, Pulse generator was used to create three different pulse widths to drive the injector drive circuits and control the pintle valve. In the proposed IDC, only one of the power transistors is turned on and the other two power transistors are turned off. This helps in both generating three stage driving current as well as driving the solenoid valve.
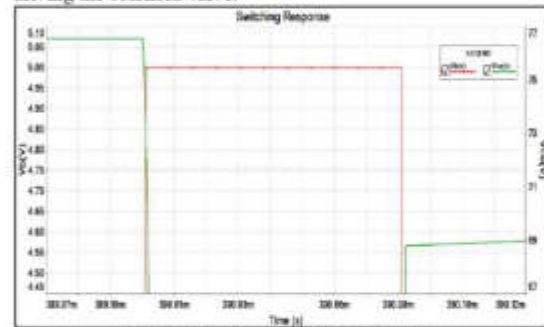
Fig.4: Switching Response of Power Transistor

As it can be easily observed in Fig.4, there is a delay while turning on and off the transistor which increase the response time of this circuit and thus decrease the efficiency of fuel injector. According to simulation result as shown in Fig.5, thermal dissipation increased abruptly and lead to power wastage and heat dissipation. The main reasons behind the increase in thermal dissipation were surge currents and voltage. Thus, the GDI injector cannot be perfectly controlled by three-stage power transistor circuit.
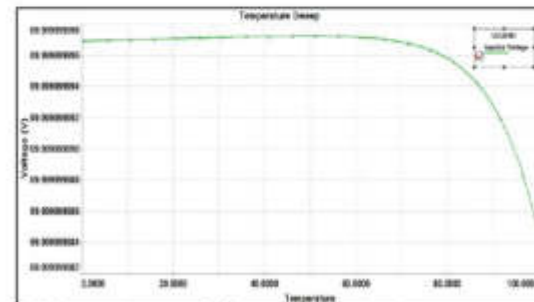
Fig.5: Temperature sweep of the three-stage Power Transistor

B. Single-Pulse IGBTs Driving Circuit

In automotive power electronics, gates are mostly control by the ECU. Based on the reading of the various sensors, ECU generates a pulse signal which turns on the injector driving circuit for a given time and thus forces the injector to maximum opening position. Remarkable effects can be seen in both turn-on and turn-off switching waveforms as shown in Fig.6. Gate voltage exhibits a step, remaining at a constant level while drain voltages rises or falls during switching.
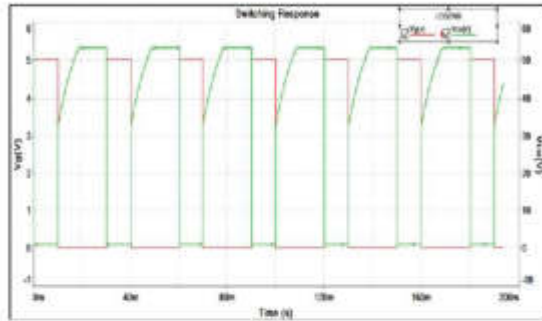
Fig 6: Switching Response of IGBT

When IGBT and MOSFET are turned on under the same condition, they exhibit almost the same kind of waveform and behave in the similar way. Therefore, it can easily be concluded from their waveform that opening time will almost be the same for single pulse-IGBT and MOSFET based injector driving circuits. However, turning-off an IGBT is very different from MOSFET. At the end of switching event, IGBTs have a tail current which is absent in case of MOSFET. This tail is caused by minority carriers trapped in the "base" of the bipolar output section of the IGBT causing the device to remain turned on and thus lead to abnormal behavior of Fuel Injector [12]. Unlike power transistors, we cannot remove these carriers to speed up switching, as there is no external connection to base output of IGBT and thus limits the switching speed. Due to this limitation, three stage Power MOSFET driving circuit is preferred over a single pulse IGBT drive circuit. It allows more precise and accurate control of fuel injector and thus helping in increasing the engine performance and fuel efficiency. It can be observed in Fig.7 temperature sweep of single pulse IGBT drive circuit varies less as compare to other three IDC in the temperature range of 0^0C-100^0C.
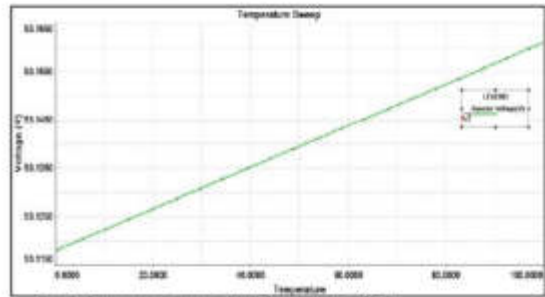
Fig.7: Temperature sweep of the Single-pulse IGBT

C. Three-stage Power MOSFET driving circuit

With the help of three stages driving circuit (different amplitude of current), we were also able to get fast response from the injector driving circuit along with better accuracy. As observed from Fig.3, introducing a PWM signal in last stage of the circuit, we were able to make injection more accurate and fast.

From Fig.8, we can see that there is a delay time of 0.016ms. This delay occurs due to the semiconductor limitations of MOSFET. There is a lag in opening and closing of injector because of this limitation. However, this time delay is less as compare to power transistor driving circuit and IGBT driving circuit.
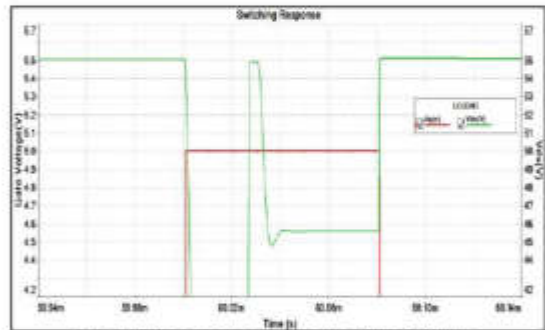
Fig 8: Switching response of three-stage Power MOSFET

As we observed from the simulations result, Performance with Power MOSFET is better than IGBT based driving circuit. Power MOSFET based driving circuit overcomes all the problems related to IGBT based driving circuit. As shown in Fig.9 thermal dissipation (39 microvolt/^{0}C) is also less as compared to all other injector driving circuit.

172
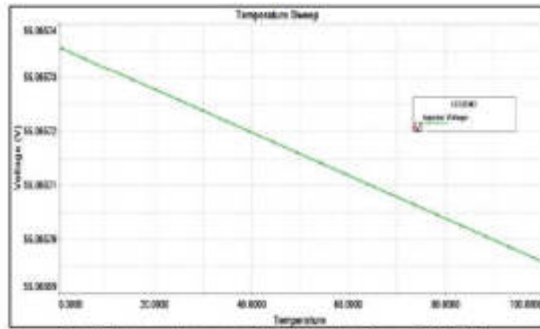
Fig. 9: Temperature sweep of the three-stage Power MOSFET

V. CONCLUSIONS

In this paper, the performances of three switching elements (Power MOSFET, IGBT and Power Transistor) in an injector drive circuit for a high pressure GDI engine were studied and compared. The results are shown in table.2,

Table II. Comparative response of the three switching elements

Parameters	Power Transistor	IGBT	Power MOSFET
Time Delay(ms)	0.8	0.06	0.016
Temperature Sweep [0-100^{0}C] (V/ ^{0}C)	$-121.008 * 10^{-12}$	$428.435 * 10^{-6}$	$-398.08 * 10^{-8}$

We obtained the following observations:

- Among the three switching elements, Power MOSFET was found to be most suitable for this application. It neither showed signs of lag due to heat dissipation nor due to trapped carriers.

- A significant amount of heat dissipation was observed in case of a power transistor due surge current and voltage. As a result there was an observable lag while switching off the power transistor.

- IGBT displayed a peculiar behavior due to minority carriers trapped in the base of the device, which cannot be cleared by external methods. As a result, we get an unwanted base current

which increases the closing time of the IGBT, thus decreasing the efficiency of the device.

REFERENCES

[1] Niall Enrich." Basic principles of operation and applications of fuel injection systems in petrol powered cars", Dept. of Mechanical and Automobile Engineering, Limerick Institute of Technology, Retrieved October 21, 2015.

[2] Zhao, Fuquan, M-C. Lai, and David L. Harrington. "Automotive spark-ignited direct-injection gasoline engines." Progress in energy and combustion science 25, no. 5 (1999): 437-562.

[3] Jiangjian, A., GaoXiyan, and Yao Chunde. "An Experimental Study on Fuel Injection System and Emission of a Small GDI Engine." In Mechatronic and Embedded Systems and Applications, Proceedings of the 2nd IEEE/ASME International Conference on, pp. 1-6. 2006.

[4] Zhang, Xiaohua, Alan Palazzolo, Chol-Bum Kweon, Erwin Thomas, Randall Tucker and Albert Kascak, Direct Fuel Injector Power Drive System Optimization. No.2014-01-1442. SAE Technical Paper, 2014.

[5] Tsai, Wen-Chang, and Peng-Cheng Yu. "Design of the electrical drive for the high-pressure GDI injector in a 500cc motorbike engine." International Journal of Engineering and Industries 2 (2011) 70-83.

[6] Huang, Du, Hongyuan Ding, Zhaowen Wang, and Ronghua Huang. "Design of drive circuit for GDI injector." In Electric Information and Control Engineering (ICEICE), 2011 International Conference on, pp. 5821-5824. IEEE, 2011.

[7] Dave, Chincholkar,"Development of Fuel Injection System for Different Fuel Injection Timing",International Journal on Mechanical Engineering and Robotics 2 (2014): 27-29.

[8] Mooney, John D. Drive circuit modeling and analysis of electronically-controlled fuel injectors for diesel engines. No. 2003-01-3361. SAE Technical Paper, 2003.

[9] Cheng, Qiang, Zhendong Zhang, and NaihuXie. "Power losses and dynamic response analysis of ultra-high speed solenoid injector within different driven strategies." Applied Thermal Engineering 91 (2015): 611-621.

[10] Lu, Haifeng, Jun Deng, Zongjie Hu, Zhijun Wu, and Liguang Li "Impact of control methods on dynamic characteristic of high speed solenoid injectors."SAE International Journal of Engines 7, no. 2014-01-1445 (2014): 1155-1164.

[11] Yamakado, Makoto, YuzoKadomukai, Motoyuki Abe, Hiromasa Kubo, and Yasunaga Hamada. Design and Development of a Battery-Voltage-Driven Fuel Injector for Direct-Injection Gasoline Engines. No. 2001-01-1348. SAE Technical Paper, 2001.

[12] Maurice, B., and L. Wuidart. "Drive Circuits for Power MOSFETs and IGBTs." ST Microelectronics, Application Note, AN524/0994 (1999).

[11]   Yamakado, M., Kadomukai, Y., Abe, M., Kubo, H., Hamada, Y., "Design and Development of a Battery-Voltage-Driven Fuel Injector for Direct-Injection Gasoline Engines," 10.4271/2001-01-1348, 2001.

SAE TECHNICAL
PAPER SERIES

2001-01-1348

Design and Development of a Battery-Voltage-Driven Fuel Injector for Direct-Injection Gasoline Engines

Makoto Yamakado, Yuzo Kadomukai and Motoyuki Abe
Mechanical Engineering Research Laboratory, Hitachi, Ltd.

Hiromasa Kubo
Automotive Products Division, Hitachi, Ltd.

Yasunaga Hamada
Hitachi Car Engineering, Ltd.

SAE The Engineering Society
For Advancing Mobility
Land Sea Air and Space®
I N T E R N A T I O N A L

174

SAE routinely stocks printed papers for a period of three years following date of publication. Direct your orders to SAE Customer Sales and Satisfaction Department.

Quantity reprint rates can be obtained from the Customer Sales and Satisfaction Department.

To request permission to reprint a technical paper or permission to use copyrighted SAE publications in other works, contact the SAE Publications Group.

GLOBAL MOBILITY DATABASE

All SAE papers, standards, and selected books are abstracted and indexed in the Global Mobility Database

Positions and opinions advanced in this paper are those of the author(s) and not necessarily those of SAE. The author is solely responsible for the content of the paper. A process is available by which discussions will be printed with the paper if it is published in SAE Transactions. For permission to publish this paper in full or in part, contact the SAE Publications Group.

Persons wishing to submit papers to be considered for presentation or publication through SAE should send the manuscript or a 300 word abstract of a proposed manuscript to: Secretary, Engineering Meetings Board, SAE.

Printed in USA

2001-01-1348

Design and Development of a Battery-Voltage-Driven Fuel Injector for Direct-Injection Gasoline Engines

Makoto Yamakado, Yuzo Kadomukai, Motoyuki Abe
Mechanical Engineering Research Laboratory, Hitachi, Ltd.

Hiromasa Kubo
Automotive Products Division, Hitachi, Ltd.

Yasunaga Hamada
Hitachi Car Engineering, Ltd.

ABSTRACT

The conventional fuel injector for direct-injection gasoline engines is driven by voltage step-up circuitry and current control circuitry. The voltage step-up circuitry boosts the battery voltage to near 100 volts. This conventional system is fairly large and a more compact system is preferable. We have developed a mass producible battery-voltage-driven fuel injector for direct-injection gasoline engines. The injector has a dual-coil structure that enables the injector to operate at the battery voltage, thus eliminating the need for either voltage step-up circuitry or current control circuitry. Deviation in battery voltage and changes in harness resistance are fully compensated through opening-coil energization-time control. In addition, with this control method, the injector can be used with a wide range of fuel pressures. In this paper, the design concept and the dynamic characteristics of the battery-voltage-driven fuel injector with a dual-coil structure and the opening-coil energized-time control method is discussed in detail.

INTRODUCTION

The Direct-Injection Spark-Ignition (DI-SI) gasoline engine is now widely used because of its low fuel consumption and high power output [1,2].

The fuel injector for the DI-SI gasoline engine injects fuel directly into the high-pressure cylinders of the engine, so the DI-SI injector needs to operate at a high-pressure. At the same time, it needs a fast valve response so that it can inject the precise amount of fuel necessary for each combustion event.

To generate a high injection pressure, the fuel injector must be capable of a strong magnetic attraction force. To obtain a fast response, it needs a fast magnetic attraction force response.

The conventional DI-SI gasoline injector is driven by voltage step-up circuitry and current control circuitry. The voltage step-up circuitry boosts the battery voltage to near 100 volts. After the valve is opened, the current control circuitry maintains the hold current to keep the valve open [3]. Figure 1 shows an example of a conventional high-voltage-driven fuel injection system for DI-SI gasoline engines. The system is basically comprised of a high-voltage injector (which we refer to as a VH injector (VH: high voltage)) and high-voltage driver circuitry.
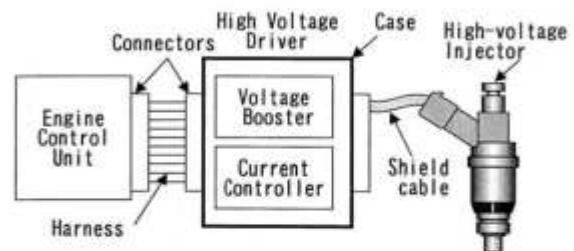
Figure 1 Conventional High-Voltage-Driven Fuel-Injection System for DI –SI Engines

Conventional injector systems are fairly large, and produce both joule heat and electromagnetic switching noise. The drive circuit should be separated from the engine control unit, and an extra case, connectors, and a shield cable is needed. The drive circuit should be installed in a cool place separate from the engine control unit, so the limited amount of space in a car is a problem. A more compact system is desirable.

We developed a mass producible battery-voltage-driven fuel injector for direct-injection gasoline engines. The injector's newly designed dual-coil structure has both an opening coil and a holding coil. The opening coil enhances valve opening, and the holding coil keeps the valve open. Each coil is designed specifically for each purpose, and the design enables the injector to be driven by battery voltage only.

However, if the battery voltage is used directly, the injector will be subjected to changes in battery voltage and harness resistance. The current flow into both coils and the driving force of the valve will vary. The flow characteristics are also susceptible to this variation.

Thus, we also developed an opening-coil energization-time control method that enables flow characteristics compensation. This control method also makes the injector compatible with a wide range of fuel pressures.

Moreover, the dimensions of the injector's core, yoke, anchor, and gaps were redesigned to raise the efficiency of the magnetic circuit compared to that of a conventional injector.

In this paper, the design and dynamic characteristics of the battery-voltage-driven fuel injector, its dual-coil structure and the opening-coil energized-time control method are described in detail.

CONSTRUCTION AND DRIVING METHOD OF THE INJECTOR

CONSTRUCTION OF THE BATTERY-VOLTAGE-DRIVEN INJECTOR

Figure 2 shows a cross-sectional view of our battery-voltage-driven fuel injector for DI-SI gasoline engines (From here, we refer to the injector as the VB injector (VB: voltage of battery)).
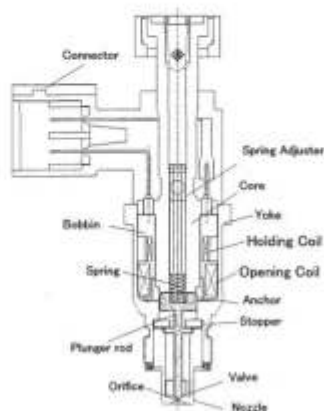
Figure 2 Battery-Voltage-Driven Fuel Injector

The injector consists of a cylindrical core with a yoke formed around the tip of the core as an integral part of the core. There is a cylindrical nozzle with a fuel injection orifice, a freely reciprocating plunger inserted in the nozzle with a valve on its tip, and a return spring to provide elastic force to the plunger to close the valve. The opening coil and holding coil provide an electromagnetic force to the plunger to open and to hold the valve using power from the battery.

The opening coil is used primarily for valve opening and has a fast response. This coil produces a strong magnetic force due to its low resistance and small number of coil turns (like a peak-hold-type injector). The holding coil holds the valve open throughout the injection event and has low power consumption due to a large number of turns (like a saturated-type injector). Each coil is wound on a magnetic core and they are magnetically coupled to each other. The connector for both coils has three contacts: VB (+), opening coil (-), and holding coil (-).

There is no difference between the configurations of the VB injector and a conventional VH injector, except for the two coils and three contacts of the VB injector. To ensure that the flow characteristics with the VH injector are comparable to those with a VB injector, we made the valve movements of the VB injector similar to those of a VH Injector.

This means the magneto-motive force (the product of the coil current and the number of coil turns) profile provided by a high-voltage drive circuit should also be obtained by applying only the battery voltage.

The current-control mechanism of the VH injector's high-voltage drive circuit is discussed in detail in the next section. The energization method for the VB injector's opening and holding coils is also explained, and the equivalent effectiveness of the magneto-motive force control enabled by the high-voltage drive circuit is shown.

DRIVING METHOD OF A CONVENTIONAL HIGH-VOLTAGE-DRIVEN INJECTOR

Figure 3 shows the high-voltage drive circuit's current control action for the VH injector, with the applied voltage, coil current, magnetic attraction force, and valve movements. The magneto-motive force has the same profile as the coil current, so it is not shown.

177
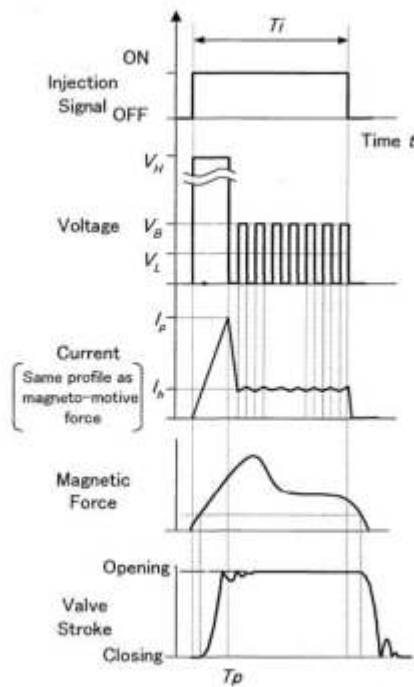
Figure 3 High-Voltage Drive Method

Where:

T_i: injection signal from the engine control unit

V_B: battery voltage V_H: boosted high voltage

V_L: average voltage during the current control period

I_p: peak current I_h: hold current

When an injection signal pulse input is received from the engine controller, the voltage booster circuit applies a high voltage V_H (which is higher than the battery voltage) to the coil.

In the injector, the voltage applied to the coil causes the coil current to be delayed by the inductance component of the coil. The coil current rises because of the high applied voltage.

The magnetic attraction force also increases, but a delay between the input magneto-motive force and the magnetic attraction force occurs because of eddy-current generation [4,5]. After the attraction force becomes larger than the return spring force and the fuel fluid dynamics force, the valve opens.

The drive circuit monitors the coil current through a current-detecting resistor. When the coil current has reached I_p, the application of V_H to the coil is suspended.

The coil current decreases, and when the current falls below the optimum current value I_h – which is adequate to hold the valve open – the current control circuitry starts to operate.

This circuitry turns the application of V_B to the coil on and off, causing the current value to remain close to I_h. The average application voltage during this on-off switching becomes V_L, and it is equivalent to the continuous V_L applied to the coil.

When the injection signal makes a high-to-low transition, the battery voltage applied to the coil is terminated.

This dampens the magnetic attraction force generated from the magneto-motive force that the coil generates. The valve is pressed against the valve seat surface by the spring force and the fuel hydraulic force.

Figure 4 shows the relationship between the applied voltage and the magnetic attraction force in terms of a simplified block diagram of a conventional high-voltage-driven injection system.
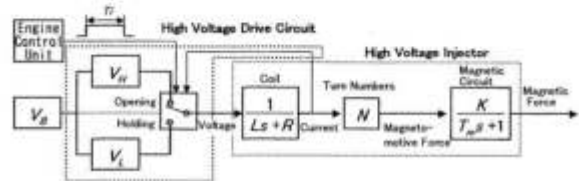
Figure 4 High-voltage-driven Injection System

Where:

L : coil inductance   R : coil resistance

N : number of coil turns   s : Laplace operator

K : constant

T_m : time constant of the magnetic circuit

The high-voltage drive circuit enables the valve opening and keeps the valve open.

This is equivalent to the circuit having a dual-voltage power supply, which allows one coil to perform two functions.

DRIVING METHOD OF THE BATTERY-VOLTAGE-DRIVEN INJECTOR

Figure 5 shows the relationships between the applied voltage and the magnetic attraction force in terms of a simplified block diagram of the battery-voltage-driven injection system.
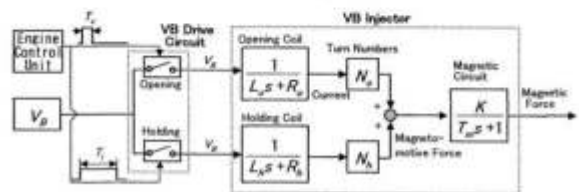
Figure 5 Battery-voltage-driven System

178

Where:

L_o : opening coil inductance

R_o : opening coil resistance

N_o : number of opening coil turns

L_h : holding coil inductance

R_h : holding coil resistance

N_h : number of holding coil turns

T_c : opening coil energization time

The battery voltage drive system does not make either V_H (> V_B) or V_L (< V_B) in the system. The valve-opening and valve-holding functions are enabled, respectively, by the fast response of the opening coil and the low power consumption of the holding coil. In fact, the holding coil current saturates at the optimum value for holding the valve open.

This means the functional parts, which have been constructed by electrical parts for the conventional system, have already installed into the VB injector.

The voltage applied to each coil is controlled through the injection signal (T_i) and the opening-coil energization-time signal from the engine-control unit.

Figure 6 shows the voltage and current changes in the opening and holding coils that correspond to an injection signal from the engine controller. The engine controller generates two signals, one for the opening coil and another for the holding coil. The pulse width of the opening-coil signal is T_c, and it is shorter than the pulse width of T_i. The pulse width of the holding-coil signal is equal to that of T_i.
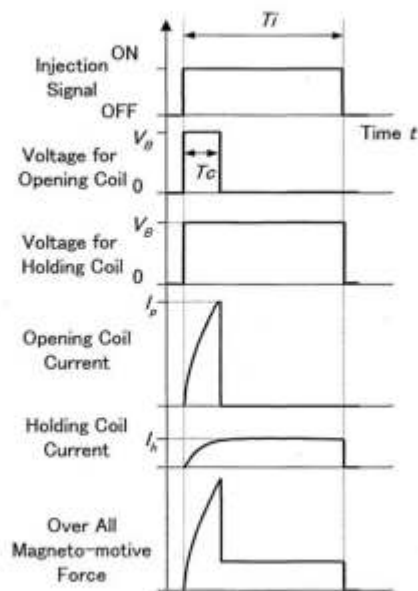
Figure 6 Battery-voltage Drive Method

The switching action is as follows.

1. When the injection signal is input to the VB injector drive circuit, battery voltage is applied to both coils. Since the opening coil has low resistance and a small number of coil turns, the current flowing into the coil rises more rapidly than the current in the holding coil. The magneto-motive force transferred to the magnetic circuit mainly comes from the opening coil.

2. After a delay equal to T_c, the control-coil current (magneto-motive force) rises to I_p, and the valve is opened. The voltage applied to the opening coil is then terminated to avoid excessive current flow in the opening coil. On the other hand, current continues to flow in the holding coil while the valve is held open. The high resistance of the holding coil means that the current flowing in it is restricted, so a current high enough to hold the valve open must be supplied.

3. After a duration of T_i and the injection signal has made a high-to-low transition, the battery voltage applied to the opening coil is terminated.

By applying voltage to both coils as described above, a magneto-motive force profile comparable to that of a high-voltage-driven injector is obtained by applying only battery voltage.

Since a fairly strong magneto-motive force is transferred to the magnetic circuit during the initial period of the valve-opening operation, the interval between the start of the injection signal and the start of the valve opening is reduced.

Since a transfer of magneto-motive force to the magnetic circuit is necessary to hold the valve open, the interval between the end of the injection signal and the start of the valve closing is reduced.

Figure 7 shows the equivalent circuit of the VB injector and the control circuit that enables the voltage application described above. The circuit is comprised of only two transistor modules and a simple ON/OFF switch circuit.
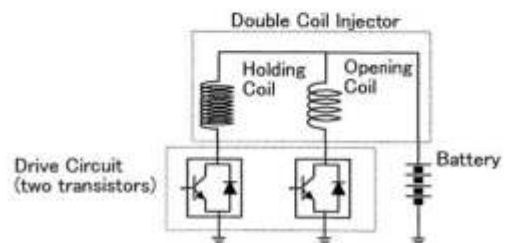
Figure 7 Battery-voltage Drive Circuit

179

EXPERIMENTAL RESULTS

We measured flow characteristics of both a VB injector and a VH injector to compare the basic performance of the two injectors.

Figure 8 shows the current profiles of the VB injector's opening and holding coils, and that of the VH injector's coil. The pulse width of the injection signal from the engine controller was 2 ms and the opening coil energization time was 0.5 ms.
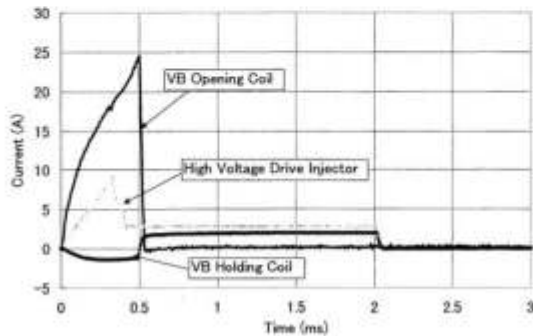
Figure 8 Coil Current

The current flowing into the VB injector's opening coil rose faster than that flowing into the holding coil and that of the HV injector. On the other hand, an opposite current occurs in the holding coil. This is because the opening and holding coils of the VB injector are wound on a magnetic core and magnetically coupled to each other.

When the battery voltage is applied to both coils, the magnetic flux generated by the opening coil increases rapidly and a counter magnetic flux is generated that causes magnetic flux-induced opposite currents to pass through the holding coil.

This counter current reduces the total flux of the magnetic circuit and weakens the magnetic attraction force for valve opening. However, when the voltage applied to the opening coil is terminated (after 0.5 ms), the current of the opening coil is cut off and the magnetic flux starts to rapidly drop. The counter magnetic flux, which enhances the holding coil current, quickly rises to its saturated value. This behavior is very convenient for stable valve holding. If the hold current changes or needs time to saturate, the interval between the end of the injection signal and the finish of the valve closing will change according to the pulse width of the injection signal from the engine controller. This strongly affects the linearity of the flow characteristics relative to the injection signal pulse width.

Figure 9 shows profiles of the VB and VH injector's magneto-motive force, and Fig. 10 shows the opening and closing response of the injectors. The response was measured by the acceleration sensor fixed on the injector's body. When the valve opens the plunger collides with the stopper, and when the valve closes the plunger collides with the nozzle. These collisions cause vibration, and the opening and closing times can be precisely identified by detecting this vibration.
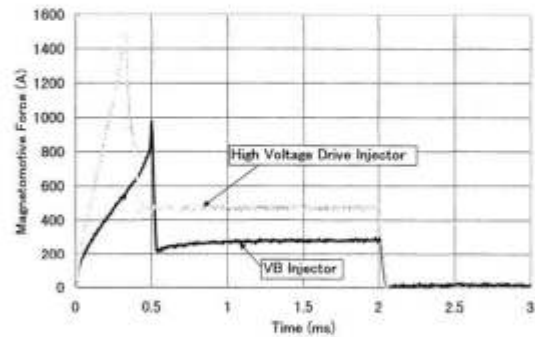
Figure 9 Magneto-motive Force
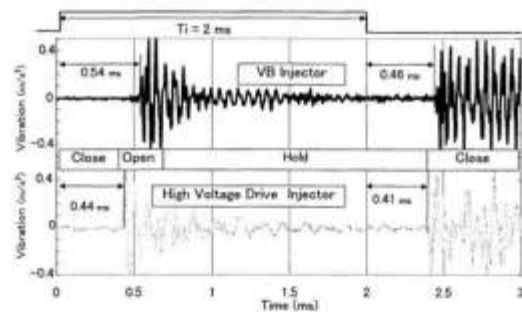
Figure 10 Opening/Closing Response

The magneto-motive force of the VB injector is weaker than that of a VH injector. On the other hand, the magnetic attraction force will saturate when the input magneto-motive force overdrives the magnetic circuit, so the effective force of the two injectors is comparable.

The valve opening response is also comparable, with a difference of only 0.1 ms. The valve operation with the VB injector is comparable to that with a VH injector.

Figure 11 shows the VB and VH injector's flow curves, created from the measured fuel flow versus the injection pulse width. These curves indicate the flow linearity and dynamic range of each injector.

The difference of the valve opening response becomes a delay towards the direction of injection signal pulse width increase. The flow linearity and dynamic range of both injectors are comparable, which demonstrates the practicality of the VB injector for DI use.

180
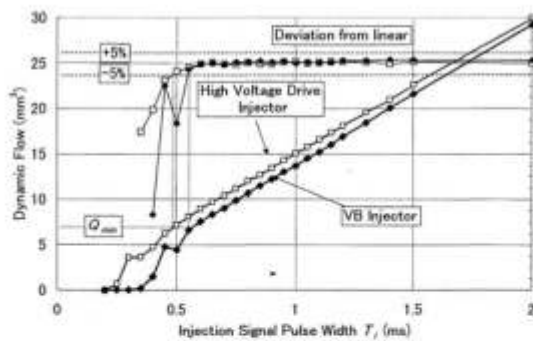
Figure 11 Flow Characteristics

THE CHALLENGE OF PRACTICAL IN-CAR APPLICATION

The VB injector can be used in DI -SI gasoline engines, as we have shown, but only with a battery voltage of 14V, and a stable fuel pressure of 7 MPa.

To apply the VB injector to DI-SI gasoline engines in the real world, counter-measures against variations in battery voltage and harness resistance are needed. These problems are solved by a circuit function in a conventional high-voltage drive system.

The VB drive system should also be adaptable to a variable fuel pressure, which will enable higher combustion efficiency in DI-SI gasoline engines. To make this possible, we developed the opening-coil energizing-time (T_c) control method for the VB injector.

These problems and counter-measures are discussed in further detail below. We analyzed the problems in detail through computer simulation and experiments.

BATTERY VOLTAGE AND HARNESS RESISTANCE VARIATION

The conventional high-voltage drive circuitry generates a stable voltage from the battery voltage by using a voltage step-up circuit and a current-control circuit.

On the other hand, the battery voltage driver for the VB injector applies the battery voltage directly to both coils (Fig. 5). The voltage of a battery installed in a car varies, though, depending on changes in the car's electrical load. In a normal driving situation, the voltage ranges from 10 to 14 V.

Moreover, causing the small resistance of the opening coil, VB injector is susceptible to the resistance changes of coil-loaded line causing the harness secular changes or coil's heat generation.

When the battery voltage drops or resistance increases, the opening-coil current (magneto-motive force) falls. The magnetic attraction force used for valve opening

decreases, which causes significant changes in the flow characteristics. For example, if the voltage drops and the resistance rises at the same time, the valve cannot open.

On the other hand, if the coil parameters are designed to enable valve opening under these conditions, an excessive current will flow into the opening coil, and the valve will be driven by too much force under the standard battery voltage and resistance. This would cause both high dissipation and valve damage.

To accommodate disturbances such as battery voltage drops or resistance increases, the valve opening force should be adjusted in accordance with the changes.

OPENING-COIL ENERGIZATION-TIME CONTROL

As shown in Fig. 5, the magnetic circuit has a time delay between the input of the magneto-motive force and the output of the magnetic attraction force. This is represented by the first-order delay with a time constant T_m and gain K. For the purposes of a frequency response analysis, it can be regarded as an integrator circuit with a cut-off frequency of $f_c = 2\pi / T_m$.

This means when the lower frequency components of the magneto-motive force are input to the magnetic circuit, the magnitude of the output magnetic attraction force corresponds to the absolute value of the magneto-motive force.

On the other hand, when the higher frequency components generated by the opening-coil current are input, the output magnetic attraction force has a magnitude that corresponds to the integrated value of the magneto-motive force.

The duration of magneto-motive force T_j is fairly smaller than the injection period that is determined by the engine speed. Therefore the magneto-motive force $U(t)$ can be described as a shock wave using frequency f.

$$U(t) = \int_{-\infty}^{+\infty} [A(f)\cos tf + B(f)\sin tf] df$$
$$= \int_{-\infty}^{+\infty} M \cos(tf - \phi) df \qquad (1)$$

Therefore the integration range can be divided as follows.

$$U(t) = \int_{-\infty}^{-f_c} M \cos(tf - \phi) df + \int_{+f_c}^{+\infty} M \cos(tf - \phi) df$$
$$+ \int_{-f_c}^{+f_c} M \cos(tf - \phi) df$$
$$= U_H(t f_H) + U_L(t f_L) \qquad (2)$$

The magneto-motive force $U(t)$ transferred to the magnetic circuit has both lower frequency components (U_L) and higher frequency components (U_H).

The magnetic attraction force is a function of magneto-motive force $U(t)$, and the higher frequency components U_H are assumed to be generated by the rapid increase in the opening-coil current, so the magnetic attraction force F_{mag} is described as

$$F_{mag}(t) = K(U_L(t f_L) + k \int_0^{Tc} U_H(t f_H) dt)$$ (3)

where K and k are constant and T_c is the opening-coil energization time.

In this way, when the battery voltage drops or the resistance rises, prolonging the opening coil's energization time (T_c) boosts the absolute value and integrated value of the magneto-motive force by prolonging the integration span.

This form of opening-coil energization-time control was incorporated in the VB injector and drive circuit. Figure 12 shows our simulation results concerning the opening-coil energization-time control. Table 1 shows the driving situation for each simulation.

Pattern (a) had a 14 V battery voltage and no additional resistance, i.e., standard conditions. Pattern (b) had a 10 V battery voltage, and pattern (c) had 0.3 Ω of additional resistance. The coil constants, spring force, and hydraulic force from the pressurized fuel were the same as for the VB injector used in the experiments discussed earlier. The time constant of the magnetic circuit was also decided on the basis of the measured value. The period of the injection signal was assumed to be 1.5 ms, and the voltage was applied to the holding coil during this period.

On the left side of Fig. 12, Case 1, the opening coil energization time (T_c) was 0.5 ms for (a), (b), and (c). On the right side, Case 2, T_c was (a) 0.5 ms, (b) 0.7 ms, and (c) 0.6 ms. The opening-coil current, holding-coil current, combined magneto-motive force, magnetic force and valve stroke are shown as a function of time.

In Case 1, where T_c was the same for all patterns, the opening-coil current of patterns (b) and (c) became lower than that of pattern (a). The opposite currents of patterns (b) and (c), caused by mutual induction, also dropped, and the drop in the magneto-motive force became smaller.

Pattern (a) can open and hold the valve open during the hold-coil energization. However, pattern (b) cannot open the valve, and pattern (c) can open the valve only with a fairly long delay and cannot hold the valve open during the energization period.
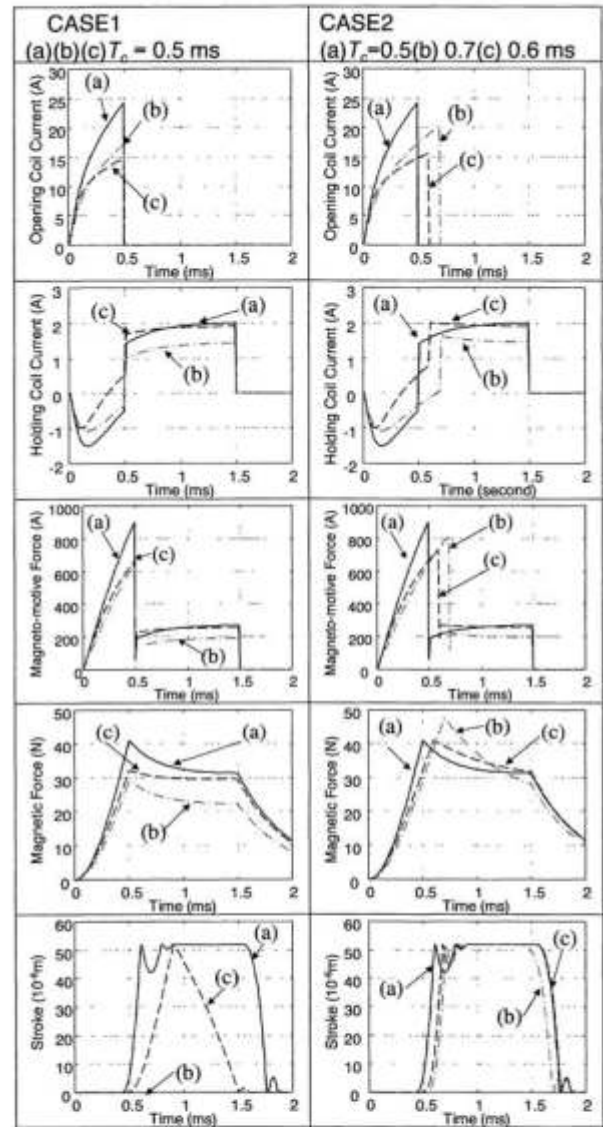
Figure 12 Effect of Opening-Coil Energized-Time Control

Table 1 Driving situation for each simulation

(a)	V_B =14 V, no additional resistance
(b)	V_B =10 V, no additional resistance
(c)	V_B =14 V, 0.3-Ω additional resistance

182

On the other hand, in Case 2, the opening-coil energization time was well controlled according to the electrical driving conditions.

The opening-coil current of patterns (b) and (c) became higher than in Case 1. The rate of change in the opening-coil current became lower as the current came closer to its saturation value. The magnetic flux-induced voltage inside of the holding coil became low, and the holding-coil current switched to a positive value even if the opening coil was still being energized.

In this way, the absolute value of the magneto-motive force input into the magnetic circuit was energized by prolonging the opening-coil energization time T_c.

Even when the peak value of the magneto-motive force of (b) and (c) were smaller than for (a), the peak value of the magnetic attraction forces became larger than that of (a). This is caused by the integrated filter simulates magnetic circuit. Patterns (b) and (c) could now open and hold open the valve. The opening-coil energization time T_c was set to a certain length for the standard electrical conditions, and when the battery voltage dropped or the resistance increased, T_c was extended. This enabled control of the valve opening force according to changes in the electrical conditions.

Figures 13 and 14 show experimental results for the injected flow characteristics of the VB injector. The electrical conditions for Fig. 13 (battery voltage of 10 V, and no additional resistance) were the same as in pattern (b) of Fig. 12, and the conditions for Fig. 14 (battery voltage of 14 V with 0.3-Ω additional resistance) were the same as in pattern (c).

As indicated by the simulation results in Fig. 12, the VB injector was able to inject fuel even if the battery voltage dropped and/or the harness resistance rose because the opening-coil energization-time (T_c) was adjusted in response to changes in the electrical conditions.

Figure14 Tc Control for a Harness Resistance Increase

Figure 15 shows the injected flow characteristics of the VB injector for different fuel pressures. The fuel pressure for each test was 7 MPa or 12 MPa.
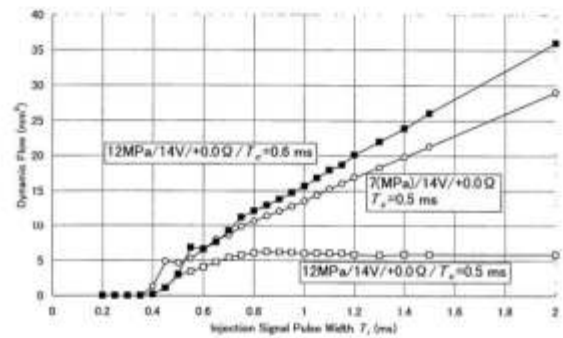
Figure15 Tc Control for Different Fuel Pressures

When the opening-coil energization-time T_c was 0.5 ms, the VB injector could operate properly at a fuel pressure of 7 MPa, but could not hold the valve at 12 MPa; this was similar to the situation depicted in Figs. 13 and 14. However, fuel injection at 12 MPa became possible when T_c was extended to 0.6 ms.

In this way, control of the opening-coil energization time T_c was also effective in enabling valve opening and fuel injection at the higher fuel pressure. This control method enables to make VB injector to be adaptable to variable fuel pressure system.

Figure 16 shows the change of flow characteristics for different values of T_c. The fuel pressure was 7 MPa, the battery voltage was 14 V, and there was no additional resistance.

Figure13 Tc Control for a Battery Voltage Drop

183
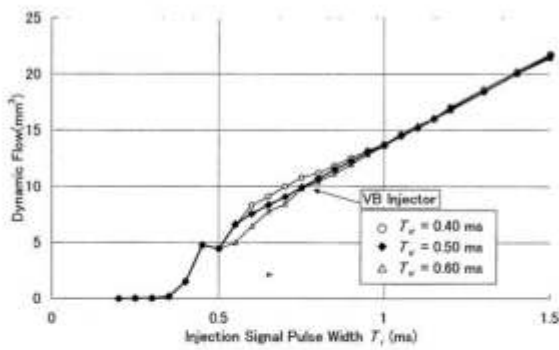
Figure16 Tc Control for Flow Characteristics

The flow characteristics just after the valve is opened can be controlled by changing the opening-coil energization-time T_c. This maximizes the flow linearity for all driving conditions.

In this way, the VB injector ensures reliable fuel injection and optimizes the flow characteristics by controlling T_c in accordance with the battery voltage, harness resistance, and fuel pressure.

FUEL INJECTION SIGNAL COMPENSATION CONTROL

Since the VB injector controls the valve-opening force by prolonging the integration span to boost the absolute value and integrated value of the magneto-motive force, it needs a certain time interval for the magneto-motive force to accumulate.

As shown by the simulation results in Fig. 13 and the experimental results in Figs. 14, 15, and 16, this time interval allows a valve-opening time lag to adapt to a drop in battery voltage, a rise in harness resistance, or a rise in fuel pressure.

Figure 17 shows the flow characteristics under different electrical conditions. The opening-coil energization time T_c was well tuned to maximize the linearity of the flow characteristics.
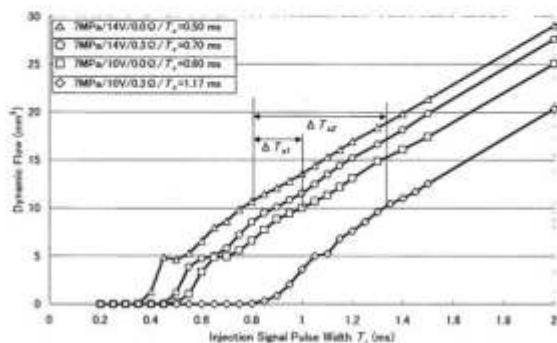
Figure 17 Flow Characteristics with Tc Control

The starting point of the flow characteristics (the intercept point on the horizontal axis) changed depending on the electrical situations.

For example, the7-MPa, 10-V case has a ΔT_{s1} delay compared with the 7-MPa, 14-V case, and the 7-MPa,10-V,+ 0.3-Ω case has a ΔT_{s2} delay.

To control the amount of injected fuel in such cases, adjusting the fuel-injection pulse width (as shown in Fig. 18) is very effective.
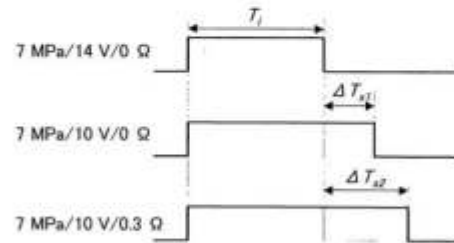
Figure 18 Injection-pulse-width Adjustment for Different Conditions

For example, when the same amount of injected fuel for injection pulse width T_i at 7 MPa and 14 V is needed, a pulse width of $T_i + \Delta T_{s1}$ is suitable for 7 MPa and 10 V, and a $T_i + \Delta T_{s2}$ pulse width is suitable for 7 MPa, 10 V, +0.3 Ω. The engine controller changes the injection pulse width in accordance with the VB injector's driving conditions.

Figure 19 shows the flow characteristics after injection pulse-width adjustment.
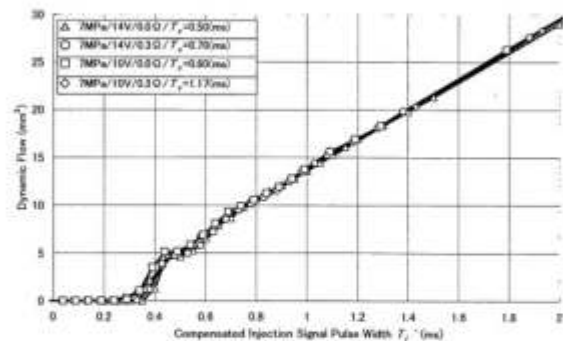
Figure 19 Flow Characteristics after Injection-pulse-width Adjustment

Thus, the impact of drops in battery voltage and increases in harness resistance are eliminated by controlling the opening-coil energization-time and the fuel-injection signal compensation. The VB injector is clearly practical for use in the real-car-world.

184

CONCLUSION

We have developed a battery-voltage-driven fuel injector for the DI-SI gasoline engine. The main features of the injector can be summarized as follows.

1. The injector has a dual-coil structure with both an opening coil and a holding coil. The opening coil enhances valve opening, and the holding coil keeps the valve open. Each coil is designed specifically for its purpose, and the injector's design makes it possible to drive it by battery voltage.

2. The driving circuit of the injector is compact and does not generate magnetic noise. This makes the driving circuit easy to install in the engine control unit.

3. The injector ensures reliable fuel injection and optimizes the flow characteristics by controlling the opening-coil energization time in accordance with changes in the battery voltage, harness resistance, and fuel pressure.

4. Control of the opening-coil energization time and the fuel-injection signal compensation provides the adaptability needed to make the injector suitable for practical use.

REFERENCES

1. Takagi, Y., Itoh, T., Muranaka, S., Iiyama, A., Iwakiri, Y., Urushihara, T., and Naitoh, K., Simultaneous Attainment of Low Fuel Consumption, High Output Power and Low Exhaust Emissions in Direct Injection SI Engines, Direct Injection SI Engine Technology, SAE 980149 (1998), pp. 195-204

2. Kume, T., Iwamoto, Y., Iida, K., Murakami, M., Akishino, K., and Ando, H., Combustion Control Technologies for Direct Injection SI Engine, SAE Paper 960600 (1996)

3. M. Pontoppidan, G .Gavuani, G. Bella and V. Rocco, Direct Fuel Injection – A study of Injector Requirements for Different Mixture Preparation Concepts, SAE Paper 970628 (1997)

4. David H. Smith, David A. Spinweber, A General Model for Solenoid Fuel Injector Dynamics, SAE Paper 8000508 (1980)

5. Harry Kekedjan, Tadeusz Krepec, Further Development of Solenoid Operated Gas Injectors with Fast Opening and Closing, SAE Paper 940450 (1994)

[12]   di Gaeta, A., Fiengo, G., Palladino, A., Giglio, V., "A control model of a Common-Rail System for Gasoline Direct Injection Engine," Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference, pp. 6614-6619, doi: 10.1109/CDC.2009.5400211, 2009.

Joint 48th IEEE Conference on Decision and Control and
28th Chinese Control Conference
Shanghai, P.R. China, December 16-18, 2009

FrA09.6

A control oriented model of a Common-Rail System for Gasoline Direct Injection Engine

Alessandro di Gaeta, Giovanni Fiengo, Angelo Palladino and Veniero Giglio

Abstract—Electronics has greatly contributed to the development of internal combustion engine. This progress has resulted in reducing environmental degradation, and yet continuing to support improvements in performance. Regarding gasoline engine, a considerable step forward has been achieved by Common Rail (CR) technology able to exactly regulate the injection pressure during whole engine speed range. As a consequence, the injection of a fixed amount of fuel is more precise and it is possible to perform multiple injections for combustion cycle.

In this paper, the authors present a mean value model aimed at the control of a CR system for a Gasoline Direct Injection (GDI) engine. The model is based on the descriptions of electro-valve, including the actuator circuit, and the fuel pressure in the rail. The performances of the proposed model are finally depicted through comparisons with experimental data collected by a CR system mounted on a 2.0 liters spark ignition engine, showing a good accuracy and reliability.

I. INTRODUCTION

Today, an huge effort is devoted by companies to preserve the earth's environment. In particular, automobile industry has to comply both the reduction of pollutant emissions enforced by international regulations and the improvement of performance required by the customers.

In this scenario, the GDI engines with High Pressure (HP) fuel injection systems, based on the CR architecture, can be considered a good solution to this aim. In fact, electronically controlled HP fuel injection system holds an important role concerning both the emission control strategy and the improvement of internal combustion engine performance [1], [2]. High pressure injection allows to finely atomize the fuel spray and to promote fuel and air mixing, resulting in significant combustion improvements [3], [4]. Following the previous consideration about the pollutant emissions and consequent role of HP system to reduce it, the prediction of CR equipment behavior is of practical relevance in design of automotive engine.

In literature, extensive research is conducted for diesel engines to reduce both NO_x and particulate (soot) emissions [5], [6], [7], or to improve the mass transfer characteristics to the catalytic surface through the analysis of a diesel oxidation catalyst [8]. In spark ignition engine, the CR system can be

useful to reduce the exhaust emission and fuel consumption and to improve the driving dynamics [9].

For example, for diesel engines, the use of CR injection systems improves performance and customers perception. In this case the modern CR systems allow multiple injections with a limited dwell time between injections, but the use of multiple injections generates a series of pressure waves, due to the closure of the injectors, which make constant injection pressure a pure theoretical assumption.

In this perspective, the CR system represents a quenched oscillating system, in which the pressure waves generated by the first injection pulse produce a variation in the injection pressure of the subsequent injections [10]. Many CR injection models have been previously proposed, based on the equations of the physics underlying the process or alternatively developed through simulation packages. The form of these models is suitable for mechanical design but is often too complex to be appropriate for control purposes.

In [11] and [12] innovative approaches in automotive control application have been considered, respectively based on mean value engine models and the discrete-continuous interactions in the fuel injection system, due to the slow time-varying frequency of the HP pump cycles and the fast sampling frequency of sensing and actuation. Whereas, a rather complete mathematical model for a common-rail injection-system dynamics numerical simulation was developed in [13] to support experimentation, layout, and control design, as well as performance optimization.

In [14] and [15] an accurate model considers the fuel injection system of diesel engines with a distributor-type pump and properly represents the unsteady flow in the pipeline connecting the jerk pump and the injector. Then the resulting hyperbolic equations are solved by using an explicit scheme of the predictor corrector type.

The authors in [16] consider the same complex fluid-dynamic phenomena and study the instability in a CR injection system for high speed diesel engines. The good accuracy of all these models is counterpoised by the complexity of the resulting equations.

In this context, the paper presents a regulation oriented model of a common-rail system for GDI engines. The proposed model is obtained exploiting the descriptions of electro-valve dynamics and of steady fuel pressure in rail. Moreover, in this work, the analysis of CR system is completed by the study of pressure alternating component, present in the HP circuit and generated by the pump in absence of injections. The robustness of the model is tested comparing some measurable variables with experimental

data.

The paper is outlined as follows. The CR system is briefly described in section II. The electro-valve dynamic model is detailed in section III, taking into account the electric circuit devoted to the actuation of the valve. Then, in section IV, the model capturing the mean pressure in the CR is illustrated and a brief experimental analysis of the pressure alternating component discussed in section V. Finally, the experimental equipment is presented and some validation results are commented in section VI. Conclusions and future activities end the paper.

II. COMMON RAIL SYSTEM

The main objective of a CR system is to supply the electro-injectors with high pressure fuel independently by the quantity of fuel. The main benefit is, therefore, to decouple the regulation of the pump by the operation of the injectors, differently from traditional injection systems where the mechanic pump generates a pressure varying with the quantity of fuel to be injected.

The CR plant for spark ignition engine, shown in Fig. 1, is composed mainly by two separated sections: a low-pressure circuit, consisting of a fuel tank (1), a fuel feed pump (2) with a preliminary filter (3) and a low-pressure pipe; an high-pressure circuit formed by a mechanical pump (4), an high-pressure pipe, a common manifold (5) equipped with a pressure sensor (7), a flow stopper, a pressure regulator valve (8) and four injectors (6). The low pressure electro-pump $(4-5$ (bar)) forces the fuel from the tank toward the HP mechanical pump, crossing the filter aimed at cleaning the fuel. The mechanical pump compresses the fuel (100-120 (bar)) and sends it into the common manifold (named common-rail) equipped with the electro-injectors. The manifold is designed so as to filter the pressure oscillations in the fuel due to the pump and the intermittent working of the injectors. Finally, the fuel pressure in the manifold is regulated through the sensor and the electro-valve that flows the excess of fuel back into the tank [17].
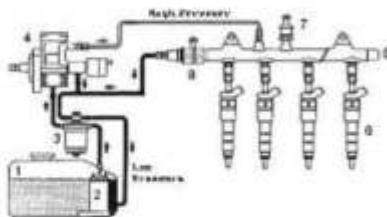
Fig. 1. Common rail injection system description: 1) fuel tank; 2) low pressure pump; 3) fuel filter; 4) high pressure pump; 5) common manifold; 6) electro-injector; 7) pressure sensor; 8) pressure regulation electro-valve.

The HP pump, shown in Fig. 2, is formed by three small pistons arranged in radial position (radial-jet) at an angular distance of 120 degrees. The pump is dragged by the engine through the camshaft and it does not require the phasing since the instant and time of injection are set by the control unit, which manages the opening of the injectors.

The alternating motion of the three small pistons is assured by a triangular cam connected to the pump's shaft and each pumping group has an intake and exhaust valve. The combined action of the three pumping groups allows to reach a pressure between 50 and 100 (bar), maintaining low level of residual pressure into the external manifolds.
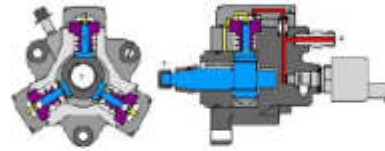
Fig. 2. High pressure pump scheme: 1) piston; 2) triangular cam; 3) shaft; 4) exit hole.

Therefore, in order to regulate the injection pressure, a mathematical model able to capture the behavior of the electro-valve and of the fuel pressure inside the CR in a wide range of engine working conditions is proposed and in the following presented.

III. PRESSURE REGULATION ELECTRO-VALVE

A. Electro-Valve Model

The valve devoted to the regulation of the pressure in the common manifold can be considered as a variable orifice without inertial effects, if the mechanical dynamics are neglected. Under this assumption, the pressure regulation electro-valve has been modeled through an electrical circuit governed by the following equation

$$L\frac{di}{dt} = -Ri + v_d, \qquad (1)$$

where also inductance variations and the counter-electromotive force have been neglected. The parameters L (H) and R (Ω) indicate, respectively, inductance and the electric resistance of coil, and $v_d(t)$ is the instantaneous voltage applied to its poles.

The values of R and L have been experimentally estimated (see Table I), forcing the valve with the battery voltage with 10 (Hz) of modulation. The comparison between dynamic model (1) and the experimental data is shown in Fig. 3, and it is possible to note how the model reproduces the electric current showing a good level of reliability and accuracy.

B. Actuation Circuit Model

Fig. 4 shows the simplified electronic circuit used to drive the electro-valve. Neglecting the voltage drop on the power transistor despite the battery $V_b(t)$, the voltage to the coil terminals can be rewritten as follows

$$v_d(t) = V_b(t)v_c(t). \qquad (2)$$

TABLE I

ESTIMATED PARAMETERS OF THE ELECTRO-VALVE MODEL

Parameter	Unit	Value
R	Ω	5.41
L	mH	18.74

187
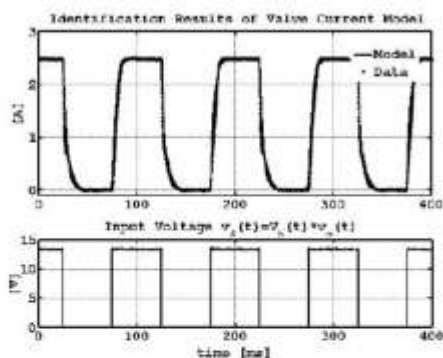
Fig. 3. The first plot depicts the comparison between the simulated (solid line) and experimental (points) electro-valve current response to a voltage input traced in the second plot.

where $v_c(t) \in \{0,1\}$ is the signal modulating in PWM the control voltage.
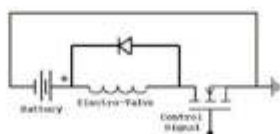
Fig. 4. A simple electric circuit aimed at actuating the electro-valve.

To this aim, the PWM frequency has been experimentally selected, $f_m = 1500$ (Hz), considered as the best compromise between reducing current ripple and a limited switching action from the components. This modulation frequency is approximately thirty times the electro-valve cutoff frequency (the electrical time constant is $\tau = R/L = 3.46$ (ms), for which $f_3 = 1/(2\pi\tau) = 46$ (Hz) and $f_m/f_3 \approx 32$), therefore ensuring a good attenuation of the harmonics introduced by the same modulation.

Coherently to what described, the voltage v_d can be approximated with its average value, V_m; that is, integrating (2) on the modulation period $T_m = 1/f_m$, it is obtained

$$V_m = \left(\frac{a\delta + b}{100} \right) V_b. \qquad (3)$$

where δ indicates the duty-cycle percentage of $v_c(t)$.

The coefficients a and b take into account the not-ideal modulation ($a = 1$ and $b = 0$ in the ideal case); their values, experimentally estimated, are listed in Table II.

Note that, since δ varies in the interval $[0; 100]$, the voltage 3 could be negative or bigger than the battery voltage, that

TABLE II

ESTIMATED PARAMETERS OF THE ACTUATION CIRCUIT MODEL

Parameter	Unit	Value
a	-	1.0573
b	%	-5.6084

is clearly unfeasible. So, to avoid this, the relation 3 must be saturated in the interval $[0; V_b]$.

Finally, replacing (3) into equation (1), the complete electro-valve model equipped with the electric circuit is obtained

$$\frac{di}{dt} = -\frac{R}{L}[i - f_i(\delta(t), V_b(t))], \qquad (4)$$

where

$$f_i(\delta, V_b) := \frac{V_b}{R} \left(\frac{a\delta + b}{100} \right) \qquad (5)$$

describes the current in stationary conditions as function of δ (control variable) and V_b (not manipulable input). The battery voltage is usually measured in an ECU (Electronic Central Unit) and taking into account to adapt the maps of the main engine actuators (injectors, spark-plugs, throttle, valves, etc.).

It is important to note that equation (4) describes now the average current behavior in the modulation period T_m. The instantaneous current, including the ripple due to the PWM, will be however governed by equation (1).

Fig. 5 shows the comparison between simulated and experimental current response of the electro-valve and actuation circuit in steady conditions. The good prediction of the steady model 5 is confirmed.
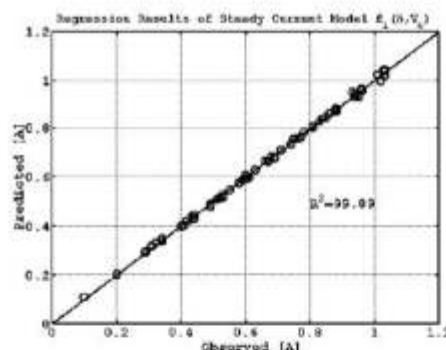
Fig. 5. Comparison of modeled $f_i(\delta, V_b)$ (predicted) and measured (observed) steady current of the electro-valve and the actuator circuit.

IV. COMMON-RAIL PRESSURE MODEL

A detailed model of the fluid-dynamics phenomena and fuel propagation in the common-rail system is not the target of this work, aimed at designing a simple phenomenological model suited for control purposes. Therefore, a mean value pressure model has been obtained through an experimental approach.

To this aim, observing pressure data in Fig. 6, obtained varying the current and the shaft speed N of the HP pump ($N := N_e/2$ where N_e is engine speed), the following model for the average pressure $\bar{p}$ in the manifold has been obtained

$$\bar{p} := f_p(i, N) = c(N)i + d(N), \qquad (6)$$

whose parameters c (bar/A) and d (bar) have been firstly estimated for each N and then fitted with the following

6616
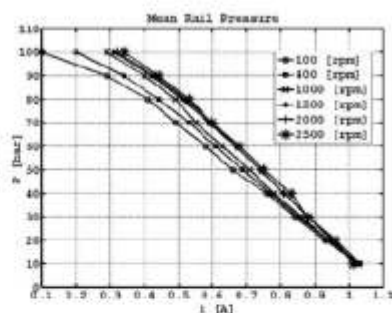
Fig. 6. Experimental data of the rail mean pressure $\bar{p}$ versus the electro-valve current i for different HP pump speeds N.

Fig. 8. Comparison of modeled $p(i, N)$ (predicted) and measured (observed) CR steady pressure.

polynomial models (plotted in Fig. 7):

$$c(N) = c_3 \left(\frac{N}{10^3}\right)^3 + c_2 \left(\frac{N}{10^3}\right)^2 + c_1 \left(\frac{N}{10^3}\right) + c_0 \quad (7)$$

$$d(N) = d_3 \left(\frac{N}{10^3}\right)^3 + d_2 \left(\frac{N}{10^3}\right)^2 + d_1 \left(\frac{N}{10^3}\right) + d_0, \quad (8)$$

whose coefficients, estimated by means of a least squares algorithm, are reported in Table III. The degree of polynomials has been chosen maximizing the well known *adjusted R-square* statistic, which is generally the best measure of fit quality when additional coefficients are added to a model.

Fig. 7. Identification results of the functions $c(N)$ (left) and $d(N)$ (right) in the pressure model.

The comparison between the identified pressure model (6), (7) and (8) and the experimental data is concisely shown in Fig. 8, whereas a 3D representation is given in Fig. 9. Observing the results, it is possible to state that the model explains the data in satisfactory way up to 80 (bar), while for higher pressure the predictive performance of model gets worse. In fact, the pressure shows a nonlinear behavior for low values of current (see Fig. 6), where the fuel flow through the regulation electro-valve is more choked.

TABLE III
ESTIMATED PARAMETERS OF THE CR PRESSURE MODEL

$c(N)$	Value	$d(N)$	Value
c_0	-97.9354	d_0	113.4366
c_1	-47.5405	d_1	41.8776
c_2	25.0493	d_2	-19.7275
c_3	-4.7962	d_3	3.5148

Fig. 9. 3D map of CR pressure model $p(i, N)$ and experimental data (dot)

We point out that the accuracy of the pressure model (6) could be further improved introducing a quadratic dependency on the current, but we preferred not to increase the complexity of a model that is mainly developed for control purposes. A feedback regulator will have to compensate model errors mainly at higher pressure.

V. ANALYSIS OF PRESSURE ALTERNATING COMPONENT

The model (6) describes the mean value, over a cycle of a HP pump, of the total rail pressure $p(t)$. During the control system design, it could be useful to have further quantitative information, even if approximated, on the alternating pressure $p_{ac}(t) := p(t) - \bar{p}(t)$ generated by the HP pump in absence of injections.

To this aim, we analyzed the amplitude of pressure oscillations at different mean pressure values $\bar{p}$ and pump speeds N. For each steady working condition we calculated the maximum and minimum percentage of overshoot and undershoot of pressure with respect to mean value, respectively $\sigma_+ = 100(p_{max} - \bar{p})/\bar{p}$ and $\sigma_- = 100(\bar{p} - p_{min})/\bar{p}$, where p_{max} and p_{min} are the maximum and minimum values of $p(t)$ observed on a enough long time interval. The results of these analysis are shown in Fig. 10. In particular, the first and second plot depict respectively the experimental values of σ_+ and σ_- got for all steady conditions. The trending curves of such distributions, averaged over N, have been

6617

189

calculated in terms of $\bar{p}$, σ_+ = $85.275\,e^{-0.0296\,\bar{p}}$ and $\sigma_- = 54.272\,e^{-0.0229\,\bar{p}}$, and their traces are superimposed in the first two plots and compared in the third.
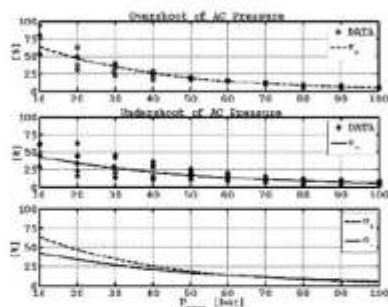
Fig. 10. Results of pressure alternating component: (first plot) maximum percentage of overshoot σ_+; (second plot) minimum percentage of undershoot σ_-; (third plot) comparison between the trending curves $\sigma_+(\bar{p})$ and $\sigma_-(\bar{p})$.

Analyzing the trends of $\sigma_+(\bar{p})$ and $\sigma_-(\bar{p})$, one can state that the percentage ripple decreases for higher mean pressure, until reaching values lower than 10% for pressure close to 100 (bar). For values higher than 50 (bar), the two curves can be considered superimposed, that is the negative and positive peaks of the alternating pressure are the same modulus values. On the contrary, for lower pressure than 50 (bar), the overshoot values are always grater than undershoot ones, until reaching the maximum difference, $\approx 20\%$, at 10 (bar).

For sake of brevity, we report for a single case the harmonic analysis of the alternating pressure measured at 60 (bar) and a pump speed of 500 (rpm). In Fig. 11, we observe a spectrum where the first four harmonics can be considered enough to describe pressure ripple, with a main harmonic of 3 (bar) of amplitude (i.e. 5% of 60 (bar)). Note that, the motion of radial pistons in the HP pump are evidently reflected in the first three harmonics of the spectrum, whose fundamental frequency is in agreement with the theoretical one, i.e. $f_P \approx 8.3$ (Hz); in fact, the HP pump completes a whole round in a time period $T_P = 1/f_P = 60/N$ (s).

We point out that in presence of injections further oscillating components would appear in $p_{ac}(t)$ at higher frequency ($f_I = nf_P$, where n is the number of injectors), as well as a continuous one (disturbance). This last determine a decreasing of the mean pressure $\bar{p}(t)$ that a feedback control action will have to compensate rapidly in order to reestablish the desired injection pressure.

VI. EXPERIMENTAL RESULTS

The performances of the proposed model are highlighted comparing the model outputs, obtained through equations (4)–(8), with experimental data collected from a CR system (see Fig. 12) mounted on a 2.0L GDI engine.

The selected duty-cycle input signal, commanded during the experiment, is reported in the second plot of Fig. 13. Figs. 14–16 show the zooms of the rail pressure and duty-cycle signals during three pressure steps. In all cases, it is
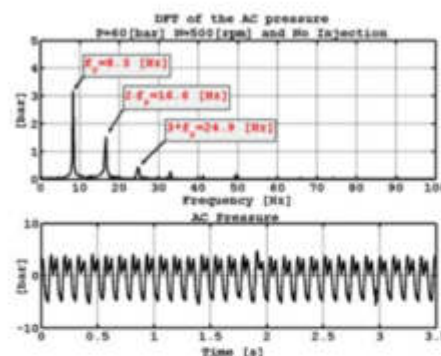
Fig. 11. DFT of the pressure alternating component at 60 (bar) of mean pressure

Fig. 12. Overall of the CR system components modeled in this work.

possible to appreciate the excellent matching of the simulated rail pressure with the measured signal. On the other hand, a presence of a constant error at 100 (bar) confirms the previous considerations on pressure model (6) at higher pressure.

VII. CONCLUSIONS

The modeling of a CR fuel injection system of a GDI engine has been developed in this paper. Furthermore a qualitative analysis of the ripple pressure generated by the HP pump in steady conditions has been presented. The synthesized model predicts satisfactorily the CR pressure both in stationary and transient conditions showing a good accuracy and reliability in all test cases. The model can be effectively used for control purposes and for real time simulation applications (e.g. hardware/software-in-the-loop). In future work, the model will be in fact adopted to design an injection pressure regulator to be used by engine control tasks.

REFERENCES

[1] M. Fry, J. King, and C. White, "A comparison of gasoline direct injection systems and discussion of development techniques," *SAE Tecnical Paper, 1999-01-0171*, 1999.

[2] P. Lino, B. Maione, and A. Rizzo, "Nonlinear modelling and control of a common rail injection system for diesel engines," *Applied mathematical modelling, Elsevier Science*, 2006.
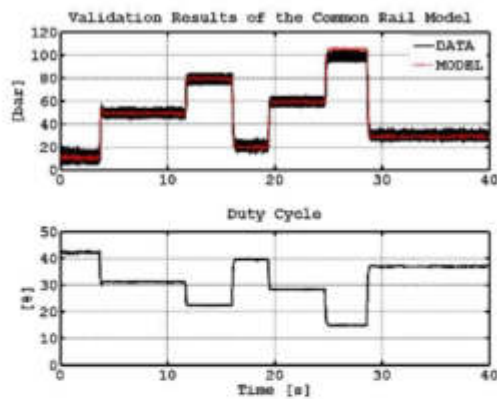
Fig. 13. Model validation: the first plot depicts the comparison between simulated (dotted line) and experimental (solid line) data of common rail pressure. The second plot depicts the duty-cycle input.
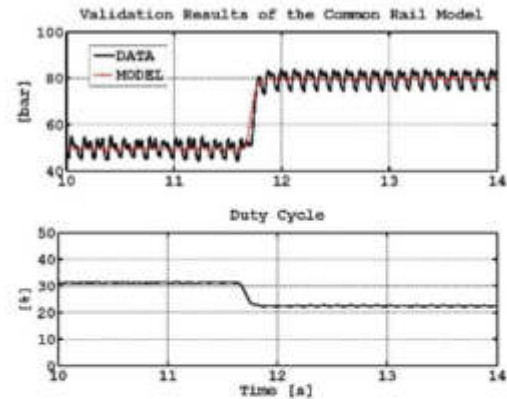
Fig. 15. Model validation: the first plot depicts the comparison between simulated (dotted line) and experimental (solid line) data of pressure step by 50 to 80 (bar). The second plot depicts the duty-cycle input.
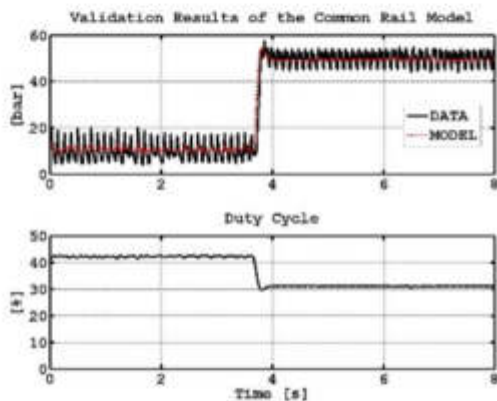
Fig. 14. Model validation: the first plot depicts the comparison between simulated (dotted line) and experimental (solid line) data of pressure step by 10 to 50 (bar). The second plot depicts the duty-cycle input.
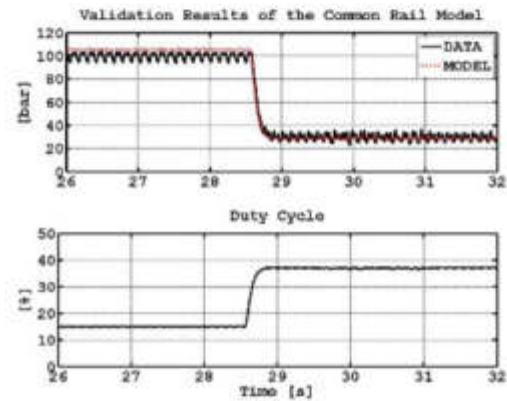
Fig. 16. Model validation: the first plot depicts the comparison between simulated (dotted line) and experimental (solid line) data of pressure drop by 100 to 30 (bar). The second plot depicts the duty-cycle input.

[3] S. HaiFeng, Z. YouTong, W. Jun, and L. LianDu, "Researches of common-rail diesel engine emission control based on cylinder pressure feedback," *IEEE Vehicle Power and Propulsion Conference*, 2008.

[4] H. Tomishima, T. Matsumoto, M. Oki, and K. Nagata, "The advanced diesel common rail system for achieving a good balance between ecology and economy," *SAE Technical Paper, 2008-28-0017*, 2008.

[5] W. Hongrong, Z. Youtong, and W. Jun, "Studies of control strategies for high pressure common rail diesel engine," *IEEE Vehicle Power and Propulsion Conference*, 2008.

[6] A. Sarvi, P. Kilpinen, and R. Zevenhoven, "Emissions from large-scale medium-speed diesel engines: 3. influence of direct water injection and common rail," *Fuel Processing Technology*, 2009.

[7] P. Eastwood, Y. Hardalupas, T. Morris, A. Taylor, K. Tufail, and T. Winstanley, "The effect of multiple fuel-injections on emissions of nox and smoke with partially-premixed diesel combustion in a common-rail diesel engine," *Institution of Mechanical Engineers: Combustion Engines and Fuels Group - Internal Combustion Engines: Performance, Fuel Economy and Emissions*, 2008.

[8] O. Lepreux, Y. Creff, and N. Petit, "Motion planning for a diesel oxidation catalyst outlet temperature," *IEEE American Control Conference*, 2008.

[9] E. Achleitner, H. Bcker, and A. Funaioli, "Direct injection systems for otto engines," *SAE Technical Paper, 2007-01-1416*, 2007.

[10] A. de Risi, F. Naccarato, and D. Laforgia, "Experimental analysis of common rail pressure wave effect on engine emissions," *SAE Tecnical Paper, 2005-01-0373*, 2005.

[11] C. Gauthier, O. Sename, L. Dugard, and G. Meissonnier, "Modelling of a diesel engine common rail injection system," *16th IFAC World Congress, Prague, CZ*, 2005.

[12] A. Baltucchi, A. B. E. Mazzi, A. S. Vincentelli, and G. Serra, "Hybrid modelling and control of the common rail injection system," *International Journal of Control*, 2007.

[13] A. E. Catania, A. Ferrari, and M. Manno, "Development and application of a complete multijet common-rail injection-system mathematical model for hydrodynamic analysis and diagnostics," *Journal of Engineering for Gas Turbines and Power*, 2008.

[14] P. Lino, B. Maione, and A. Rizzo, "A control-oriented model of a common rail injection system for diesel engines," *10th IEEE Conference on Emerging Technologies and Factory Automation*, 2005.

[15] R. Morselli, E. Corti, and G. Rizzoni, "Energy based model of a common rail injector," *Proceedings of the 2002 International Conference on Control Applications*, 2002.

[16] P. Digesu, A. Ficarella, D. Laforgia, G. Bruni, and M. Ricco, "Diesel electro-injector: A numerical simulation code," *SAE Tecnical Paper, 940193*, 1994.

[17] W. Hongrong, Z. Youtong, and W. Jun, "Studies of control strategies for high pressure common rail diesel engine," *IEEE Vehicle Power and Propulsion Conference*, 2008.

6619

191

[13]   Lu, H., Deng, J., Hu, Z., Wu, Z. et al., "Impact of Control Methods on Dynamic Characteristics of high-Speed Solenoid Injector," SAE Int. J. Engines, Vol. 7, No. 3, pp. 1155-1164, doi: 10.4271/2014-01-1445, 2014.

Impact of Control Methods on Dynamic Characteristic of High Speed Solenoid Injectors

Haifeng Lu, Jun Deng, Zongjie Hu, Zhijun Wu, and Liguang Li
Tongji Univ.

ABSTRACT

Accurate control of both the timing and quantity of injection events is critical for engine performance and emissions. The most serious problem which reduces the accuracy of the control operation in such systems is a time delay of the responsiveness for the opening and closing operation of the electromagnetic valve. Modern electronic control systems should be capable of driving high speed solenoid injectors at a very fast switch frequency with high efficiency and acceptable power requirements. In this paper, the dynamic characteristic of a high speed servo-hydraulic solenoid injector for diesel engine, with different driving circuits and control methods, is investigated. A pre-energizing control strategy based on a dual power supply is applied to speed up the opening response time of the injectors. The effect of free-wheeling circuit on the closing response time is experimentally studied and a reverse-energizing circuit is utilized for achieving a rapid extinction of the operating magnetic field. The solenoid current is measured to get the electrical delay of the injectors. Meanwhile, high-speed photography is employed to detect the total delay (electrical delay, hydraulic delay and mechanical delay) of the injection spray process. The results show that, there is a high correlation between the opening delay and the solenoid current, but the armature of the injector will not move until the current reaches a threshold value. The switching time (both the opening and the closing time) could be further optimized through the proposed driving circuits and control strategies without any change to the injectors. As a result, the interval between two injections can be significantly reduced, which can provide greater flexibility for multi-injections strategies.

INTRODUCTION

Direct injection technology has provided a highly precise and flexible method to optimize the emissions and efficiency of the piston engine, including diesel engine and gasoline engine. With increasingly stringent CO_2 and noxious emission regulations, the number of injection events in a single combustion cycle has also increased, about 2~3 times in gasoline direct injection (GDI) engines and 3~8 times in diesel engines. These trends demand a more rapid and precise response of the high speed solenoid injector. The response delay of the injector is inevitable during its opening and closing period. Because the solenoid injector is an inductive element. Solenoid valve injector along with the development of Engine electronic control technology has made great strides in the past few decades. The control methods of the solenoid valve injector have also been upgraded with the progress of the electronic component. In 1971, Motorola Inc. had used transistors to control the opening and closing of the solenoid valve [1]. The Bendix Corporation applied for a patent about inductive drive circuit in 1979, which had initially formed an idea of close-loop control on the current of the solenoid [2]. From 1981 to 1985, several companies such as Bosch, Lucas,

Denso, and Ford etc. invented different methods and apparatuses for driving the solenoid injectors, which had almost the same core idea of "Peak & Hold" current control method [3, 4, 5, 6, 7]. This idea applies a high-amperage current to solenoid injectors at the beginning of a trigger pulse, which is called "Peak" stage. Once the solenoid valve is opened, no other mechanical work needs to be performed, and therefore a smaller current is sufficient for maintaining the open position of the valve than is required for opening the valve itself, which is called "Hold" stage. "Peak & Hold" method has undoubtedly become a basic idea of driving the solenoid injector ever since, for it can not only make the current in the coil of the solenoid rise sharply, but also reduce the power consumption and keep the injector from overheating.

In addition to speeding up the opening process, engineers were also laying emphasis on minimizing the closing delay. As the solenoid injector is an inductive load, the freewheeling circuit was employed to smooth the current and eliminate the sudden voltage spike across an inductive load when its supply voltage is suddenly reduced or removed. This circuit works when the Pulse Width Modulation (PWM) is used to control the power supplied to the coil of the solenoid. While during the

1155

192

closing process, the freewheeling circuit will also slow down the extinction of the electromagnetic field in the coil, which leads to the closing delay. From 1985 to 1996, Denso, Motorola, Chrysler, Caterpillar, and Honda etc. applied for their own patents of solenoid injector driving circuits with the freewheeling control functions, which can get a fast decay path to allow the solenoid valve to return to its inactivated position [8, 9, 10, 11, 13]. After that, during 1997 to 2008, the MOSFET technology had made considerable progress, so that companies like Mitsubishi, Bosch, Toyota, Conti and Delphi applied MOSFET to replace traditional transistors to reaching higher power efficiency and the system was greatly simplified. Thus the mainstream solenoid injector driving circuit has finalized the design [14, 15, 16, 17, 18, 19].

Even though automobile manufacturers have kept refreshing their driving circuits for high speed injectors, there are rarely few elaborate investigations about the impact of control methods on the dynamic characteristic of high speed solenoid injectors, especially the switching delay. Mooney, Jrohn D (2003) establishes a process for evaluating parameters of the driving circuits via simulation [20]. Then SONG Jun (2005) BB further investigated the opening pulse and PWM pulse duty factor of the driving circuits, and found that an ideal control current can be achieved by optimized parameters [21]. Lee, Ill-Yeong (2006) designed an electronic valve driving circuit with 3 power source, which can shorten the switching delay time from 5 ms to 1.55 ms [22]. Afterwards, Ziqiang Xu simulated the driving circuit of solenoid valve employing a precise PSPICE model, which applied a method of high and low side driving. The result turned out that the model could achieve the current driving requirement of solenoid valve [23]. Huang Du (2011) designed drive circuits for a GDI injector, and verified that the drive circuits satisfied the requirements through detecting the current waveform [24]. More recent studies by Sheng-Nian Cai (2012) used soft switching technology in high speed solenoid valve driver circuit, and found it can reduce switching losses and noise. The energy consumption of the injectors could be further improved [25].

These previous researchers have provided many advanced control methods to satisfy the current requirement of solenoid injectors. However, the data of the total opening and closing delay of the injection behaviors are still scarce. Considering the fact that multi-injection strategies are becoming more and more complex and precise, the crank angle would the fuel be exactly delivered to the cylinder after the injection pulse had been given are really cared about. On the other hand, many different control methods have been proposed, and further comparison studies are still necessary.

The purpose of this paper is to investigate the impact of different control methods on the dynamic characteristic of the injection events, and the paper will only focus on the mainstream solenoid injectors, which has a better technological maturity and cost efficiency than the piezo injectors. Based on the common rail and high speed imaging system, the spray characteristics are investigated in a constant volume vessel.

The spray characteristics are compared to check the effects of different driving voltage, opening strategy, power supply mode and cut-off strategy. Meanwhile, the current of the injector is also detected and compared with the spray characteristics.

EXPERIMENTAL FACILITIES

Variable Driving System

In order to meet the test requirements, a specially designed driving circuit for the solenoid injector is employed. The Variable Driving System, showed in Figure 1, can operate in multiple modes. All the possible current paths are marked from "a" to "e". The flexible control strategies can be achieved through the combination of different paths. There are 5 MOSFETs (Metal-oxide-semiconductor field-effect transistor) in this circuit, working as switch components. Among them, H1, H2, H3 are high side switches and L1, L2 are low side switches. All the MOSFETs are integrated with body diodes and 2 discrete diodes are used in path "a" and "b", which can turn the path into "one-way streets". HV (High Voltage) is supplied by a boost circuit and the HV can vary from 40 V to 100 V. LV (Low Voltage) is a 12 V battery. All the MOESFETs are controlled by PWM (Pulse Width Modulation) signal generated from a MCU (Micro Controller Unit). It's worth noting that the MCU cannot drive the MOSFETs directly, so there are still some circuits between the MCU and the MOSFETs. To drive the high side MOSFETs, the bootstrap circuit is recommended.

a. High Voltage Supply Path
b. Low Voltage Supply Path
c. Freewheeling Path
d. Freewheeling Cut-off Path
e. Reverse Power Supply Path
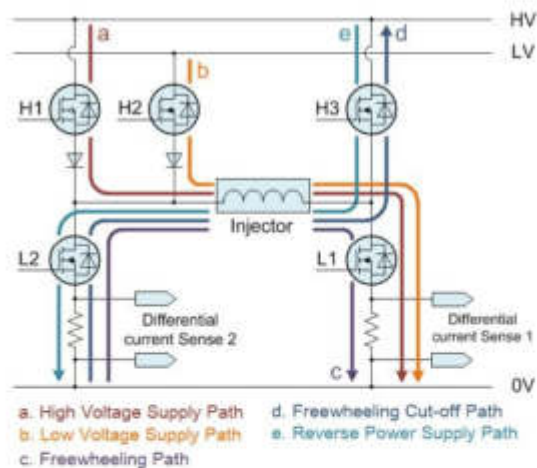
Figure 1. Schematic diagram of the Variable Driving System

As Figure 1 shows, when H1 and L1 are turned on, HV is supplied to the solenoid injector through path "a" and the coil will have a higher current. When H2 and L1 are turned on, LV is supplied to the injector through path "b", and the current will be lower than path "a". As the solenoid is an inductive load, the H1 or H2 can rapidly switch on or off, which called PWM mode, to control the current in it. During this PWM period, when H1 and H2 are both in non-conducting states, current generated

by the self-induction effect will flow through path "c", which is always called the freewheeling circle. If the L1 is suddenly cut off during the freewheeling period, the self-induction voltage of the injector will have no way to go, except the path "d". Path "d" can not only provide the discharge channel of the inductive voltage, but also can recycle some energy of the solenoid and improve the efficiency of the system. The last path "e" will not work until the L2 and H3 are turned on. The current will flow in the reverse direction, which is supposed to accelerate the extinction of the electromagnetic field in the solenoid when the injector is closing. Differential current sense 1 and 2 are used to real-time monitor the current of the injector, which can provide the conditions for the closed-loop control.

Spray Characteristics Test System

The spray injection delay characteristics test bench includes a common rail system, a constant volume, a laser generator, a high speed camera and an injection control system. The common rail system can provide injection pressure up to 150MPa. The constant volume, which is filled with nitrogen, can operate at the highest back pressure of 5MPa. The test bench configuration is illustrated in Figure 2. The high speed complementary metal oxide semiconductor (CMOS) camera (Phantomv7.3) equipped with a 105 mm focal length Nikkor lens is used in this study, and the images are collected at 33,000 fps, 28 μs exposure time, and with 192 × 360 pixel resolution.
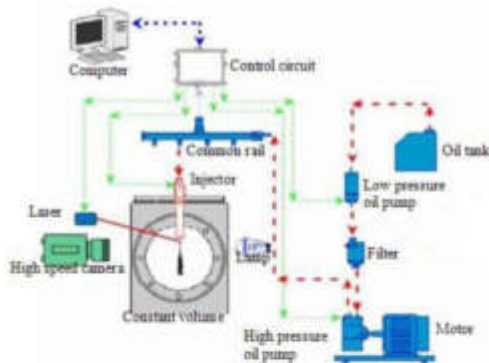
Figure 2. High pressure common rail spray test bench

The solenoid valve injector is driven by ECU. When ECU sends out the injection signal to the injector, another synchronous signal is sent out to the laser generator and then the laser generator emits a laser beam, pointing to the head of the injector. Then the high speed camera is utilized to record the laser signal and the whole spray injection process: the laser signal first appears and then the spray comes out from the nozzle tip; the laser signal disappears and then the spray departs from nozzle tip. In that case, the interval between the injection signal and the real injection can be calculated. The Figure 3 shows the whole recording process. The delay of the laser, from the given electrical signal to optical signal appears, is less than 1 us. While the delay of the injection is about hundreds of microseconds to milliseconds, so the laser delay is small enough to be ignored.
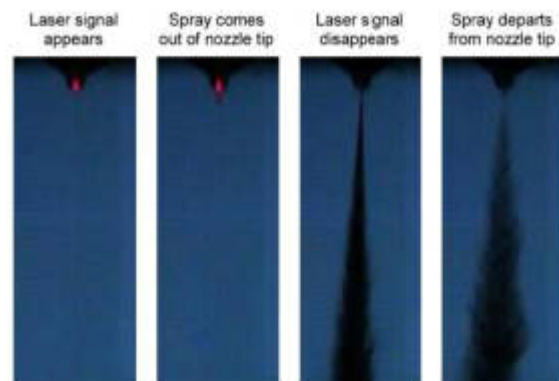
Figure 3. Laser signal and spray injection process

Parameters Definition

Injection Delay Parameters Definition

In this article, the time interval between ECU sending out a "start injection" signal and the actual injection beginning is defined as Injection Opening Delay (IOD), which in pictures reflects the time interval between the laser signal appearance and the fuel injecting out of nozzle tip. The Injection Closing Delay (ICD) is defined as the time interval between ECU sending out an "end injection" signal and the actual injection ending, which correspondingly in pictures equals to the time interval between the laser signal disappearance and the injection fuel departs from nozzle tip. The Set Injection Duration (SID) is defined as the time interval between the laser signal appearance and its disappearance; and the Real Injection Duration (RID) is the time interval between the spray coming out and departing from nozzle tip. The relationships between these parameters are shown in Figure 3.

Spray Characteristics Parameters Definition

Parameters of the spray tip penetration (S) and the spray cone angle (θ) are normally chosen by researches to present the spray features. As presented in Figure 4, the spray cone angle is the largest angle formed by the two boundaries of spray and the nozzle within the calculation range.

Figure 4. Definitions of spray core angle

194

This paper concentrates on the injection delay of the spray. Thus, we focus only on the spray cone angle. This is due to the movement of the needle of the injector can be reflected from the change of the spray cone angle, especially when the needle is closing. Figure 5 shows the spray core angle curve of a normal injection event. The angle appears later than the injection signal. Once the spray comes out, it will have a core angle. Therefore the core angle curve rises sharply at the beginning. When the spray has stabilized, the angle slowly declines and stays at a steady value. Afterwards, the closing of the needle will break the stability of the spray and the angle will rise again. At the end of injection, when the spray separates from the nozzle tip, the calculation program will judge the angle to be zero. More information about the definition of the spray cone angle is presented in Gao Y., et al (2010) [27].
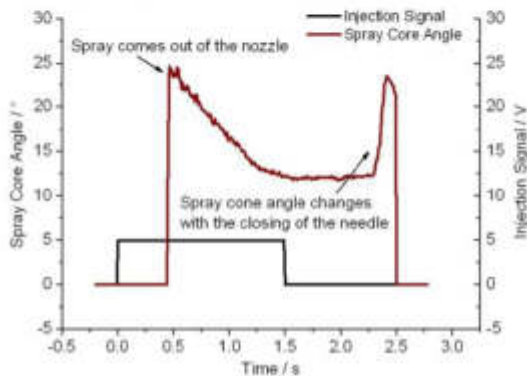
Figure 5. Spray core angle curve of the injection event

Image Processing

Data processing software is used for the image processing and the calculation of the experimental results. Figure 6 shows an example of spray image before and after the image processing. First of all, the background subtraction is performed using an image prior to the injection event. Then the Otsu segmentation algorithm based is used for the boundaries definition (Otsu N., 1979). The tests have been done and the results demonstrated that the Otsu segmentation algorithm could properly detect the estimated spray boundary even if the illumination was poor. At last, all the images were converted to binary images, on which the experimental results could be easily calculated.
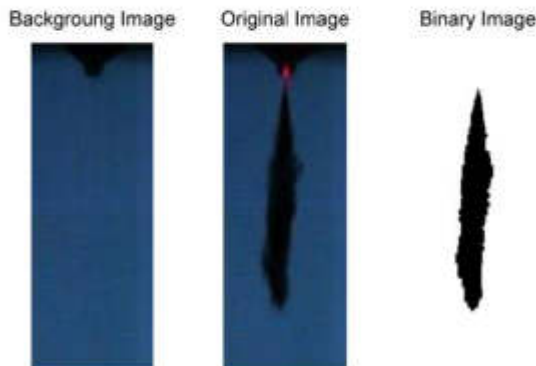
Figure 6. Image processing method

Experimental Conditions

During this study, a high speed servo-hydraulic solenoid injector for diesel engine is investigated. The ambient back pressure and the injection duration are fixed to 1 MPa and 1.5 ms respectively. Before the determination of the test conditions, the opening and closing threshold current of the solenoid injector must be checked. In Figure 7, curve "a" and "b" show the opening test of the injector. The injector is supplied by LV (Low Voltage) with a 50% duty cycle PWM, and the current gradually creeps upwards. The little bump of the curve "b" is caused by the movement of the solenoid valve.

Because the moving armature will change the electromagnetic field a little. Therefore, the opening threshold current of the injector is 4 A. Curve "c" and "d" is the closing test of the injector. The PWM duty is set to 100% at the beginning, and the injector can easily open. After that, the duty is set to 30% to let the current smoothly decline. Similarly, the little jump of the current is also caused by the moving armature, so that the closing threshold current of this injector is 2 A. All the parameters of the high speed solenoid injector are showed in Table 2.
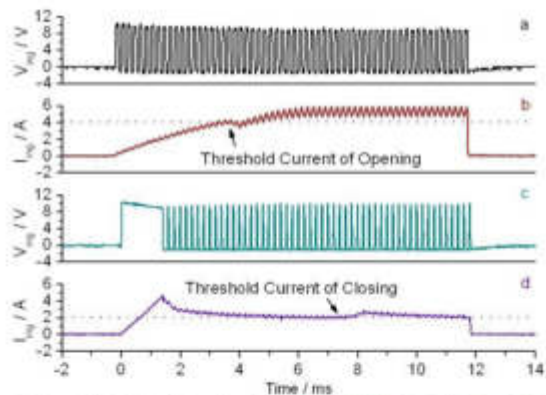
a. Solenoid voltage in the opening test   c. Solenoid voltage in the closing test
b. Solenoid current in the opening test   d. Solenoid current in the closing test

Figure 7. Test of the opening and closing current of the injector

"Peak & Hold" method, which mentioned above, is applied as a basic strategy in this study. ECU monitors the current of the injector through two differential current sense circuits, and close-loop control the current by adjusting the duty cycle of the PWM. Meanwhile, the control strategy of the main injection period (opening and closing strategies are not included) is fixed, which can improve the comparability of the experiment. In other words, no matter what opening or closing strategies are employed in the experiment, the "Peak & Hold" method is always used during the main injection period. According to the opening and closing threshold current test, the "Peak" current is set to 8 A and the "Hold" current is set to 4 A. These current values are all twice the threshold values, which can enhance the stability of the movement of the solenoid valve.

Table 2. Parameters of the high speed solenoid injector

Type	Unbalanced valve, servo-hydraulic, solenoid injector
Hole numbers of the nozzle	Single hole
Diameter of the hole	0.16 mm
Resistance of the solenoid	0.8 Ω
Inductance of the solenoid	1.9 mH
Opening threshold current of the solenoid	4 A
Closing threshold current of the solenoid	2 A

In order to thoroughly investigate the influences of control strategies on spray injection delay, there are 4 aspects of this issue have to be addressed:

1. The first issue involves the value of the HV (High Voltage) supply. It has long been recognized that the higher voltage supply can accelerate the current-rising speed of the solenoid. But how high is appropriate? For the higher the HV supply is, the lower the conversion efficiency of the boost circuit will be.

2. The second issue relates to opening strategy. There is no doubt that the "Peak & Hold" method contributes greatly to speeding up the opening process of the injector. The pre-energize strategy seems to have further optimization potential. This method will not increase the cost of the driving system. But its real effect needs to be tested.

3. The third aspect deals with the closing strategy. The reverse power supply seems to further speed up the extinction of the electromagnetic field. Also, it will increases system's complexity. Three different strategies, showed in Table 1, will be compared.

4. The fourth issue is about the power supply. Dual power supply has a higher current control precision. Meanwhile, the single power's system is more simplified. Do they have any influence on the response characteristic of the injector?

For further understanding, the spray injection delay characteristics under different injection pressures are also considered.

Table 1. Experimental conditions

Ambient back pressure	1 MPa
Injection duration	1.5 ms
Low Voltage Supply	12 V
Peak current of the close-loop control	8 A
Hold current of the close-loop control	4 A
High Voltage Supply	40 V / 60 V / 80 V / 100 V
Opening Strategy	Pre-energize / Non-pre-energize
Closing Strategy	Non Freewheeling cut off / Freewheeling cut off / Reverse power supply
Power Supply	Single power / Dual power
Injection Pressure	70 MPa / 100 MPa / 130 MPa

RESULTS AND ANALYSIS

Data Processing and Error

In this study, there are two kinds of experimental data. One is the voltage and current data of the driving circuit detected by the oscilloscope (OSC). The other is the injection delay and the spray core angle, which are calculated from the spray images. At each experimental condition, tests are repeated 10 times and results are averaged to avoid random error.

Furthermore, the injection delay is the most important indicator. The Coefficient of Variation (CV) is calculated to evaluate the dispersion of the observed data.

$$C_v = \frac{\sigma}{\mu} \times 100\%$$

(1)

where is the standard deviation and μ is the mean. The CV of each experiment condition is calculated separately. The maximum CV of the opening delay is 1.41% and the maximum CV of the closing delay is 2.67%. The highly accurate and repeatable experimental results guarantee the reliability of the following analysis.

Effects of High Voltage Value

In this test, the value of the HV (High Voltage) has been only changed. Meanwhile, the opening strategy is non-pre-energize; the closing strategy is freewheeling cut-off; dual power supply and the injection pressure is 100 MPa. Figure 8 shows the voltage and current of the solenoid at different HV (High Voltage) value. In this test, the high side MOSFET H1 and the low side MOSFET L1 are turned on at the very beginning of the injection, and the current flows through the path "a", which is mentioned in Figure 1. As a result, the current of the solenoid will sharply rise, and the H1 keeps conductive till the current reach the "peak" current value, which is set to 8 A. Then, the H1 is turned off and H2 is turned on. The solenoid is powered by the Low Voltage (LV) and the current flows through the path "b". H2 is controlled by PWM to keep the current at the "hold" current value, which is set to 4 A. After the 1.5 ms injection duration, all the MOSFET is turned off, namely freewheeling cut-off. Thus, the self-induction voltage makes the voltage of the solenoid negative. Actually, the self-induction voltage is so high that it will charge the HV through the path "d".

Obviously, a higher HV will let the current of the solenoid reach the "peak" value earlier. For the relationship of the current and voltage in the inductance is

$$u = L\frac{di}{dt}$$

(2)

where is u the external voltage, i is the current in the inductance and L is the inductance value. This equation means that the current change rate of an inductance is proportional to

196

the voltage. But, it is an interesting phenomenon that the current also falls down more quickly than the lower HV condition. This is mainly due to the higher negative self-induction voltage. When the low side MOSFET is turned off, the normal freewheeling path "c" is blocked. So the self-induction voltage begins to generate and rises sharply. When its value exceeds the HV, path "d" will become conductive. Thus, path "d" is another freewheeling path in a sense. This phenomenon indicates that if the HV is so high that even the self-induction voltage cannot exceed it, path "d" will not become conductive anymore, just like it never exists. Then the decrease of the decline of the current will be even faster and electrical delay will be further improved. But it should be noted that this method will increase the risk of breakdown the MOSFET L1.

Figure 9 shows the trend of spray core angle with different HV. The total delay (electrical delay, hydraulic delay and mechanical delay) of the injection event can be clearly observed. The opening delay and closing delay of the spray is summarized in Figure 10. From 40 V to 100 V, the opening delay declines from 550 μs to 386 μs, and the closing delay is also reduced from 1008 μs to 942 μs. Among them, from 40 V to 60 V, the opening delay is improved by 18% and the closing delay is improved by 4%. Although, in the case of 80 V and 100 V, the dynamic response is better, the increasing range of the response speed decreases. On the other hand, the effect of the HV on the closing delay is not so obviously. In consideration of the boost circuit efficiency, maybe 60 V is good enough for the injector used in this study.

Figure 8. Effects of high voltage value on the solenoid current
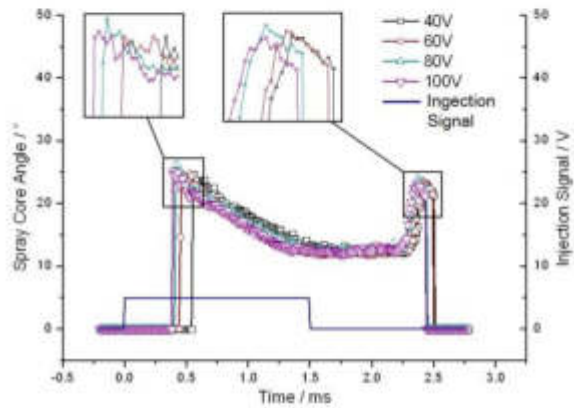
Figure 9. Spray core angle of different HV
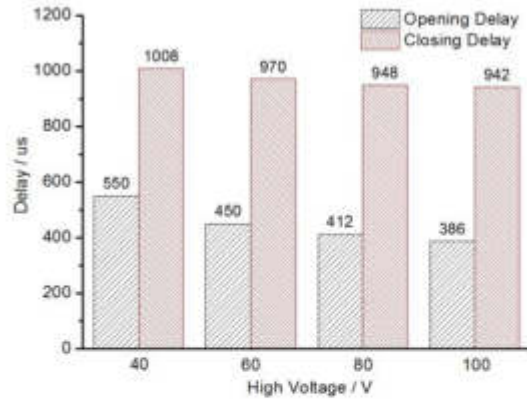
Figure 10. Injection delay of different HV

Effects of Opening Strategy

The pre-energize opening strategy supplies the power to the injector just before the injection signal has given. Meanwhile, the ECU monitors the solenoid current and modulates the duty cycle of the PWM to limit the current below the opening threshold value. In this way, the injector will be kept closed, but the electromagnetic field has been partially established, waiting

for the injection signal. This test is also carried out under the condition of 60 V High Voltage (HV) supply, freewheeling cut-off, dual power, 100 MPa injection pressure.

Figure 11 shows the solenoid voltage and current of the pre-energize strategy. Before the start of the injection signal, the current of the injector has been pulled up to 2 A, which is lower than the opening threshold current of the injector. When the injection signal is given, the "Peak & Hold" method is performed as usual.
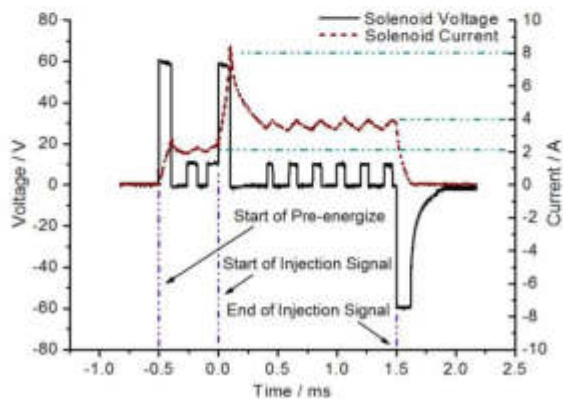
Figure 11. Solenoid voltage and current of the pre-energize strategy
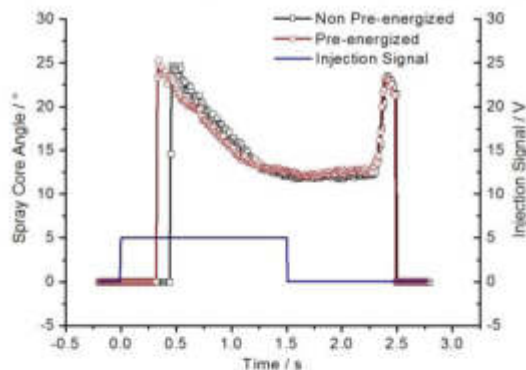
Figure 12. Spray core angle of different opening strategies

The effects of the pre and non-pre-energize opening strategy are compared in Figure 12. The spray core angle curve shows that the pre-energized strategy can really speed up the opening response of the injection. The opening delay of the pre-energize strategy is 332 μs, which is showed in Figure 13, nearly 26% improved than the non-pre-energize strategy. This opening response speed is even faster than the condition of the 100 V HV. This result indicates that the impact of the establishment of the magnetic field in the solenoid is more obvious than increasing the current change rate of the solenoid. The reason is, in case of an inductive load, like solenoid injector or motor, the current will not rise synchronously when the external electric field increases. This inductive reactance is essentially an outward manifestation that

the energy of the external electric field is converting into the energy of a magnetic field. Therefore, the pre-energize strategy is to some extent a method to establish the magnetic field in the solenoid beforehand. The inductive reactance of the solenoid is relatively reduced, so that the electrical delay of the injection can be significantly improved.
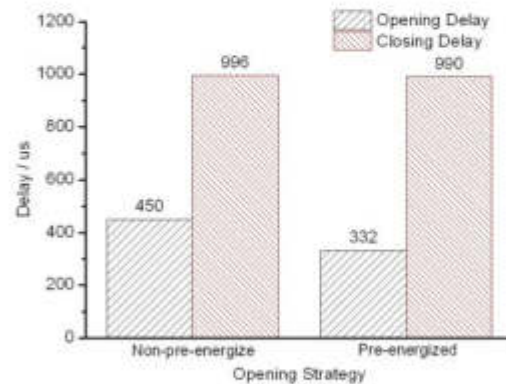
Figure 13. Injection delay of different opening strategies

Effects of Closing Strategy

This case firstly compares the difference between the freewheeling cut-off strategy and the non-freewheeling cut-off strategy. Except for the closing strategy, the other conditions of the test are the same: 60V HV, non-pre-energize, dual power and 100MPa injection pressure. The picture "a" of the Figure 14 shows the current curve of the solenoid, whose driving circuit does not cut off the low side MOSFET L1 at the end of the injection. The curve indicates that the current will slowly slip down as the freewheeling path "c" is always conductive. The little "jump" of the current curve shows the movement of the solenoid armature, when the current reaches the closing threshold value of this injector. The closing delay of this strategy, showed in Figure 16, is up to 0256 μs. This result proves the necessity of the freewheeling cut-off circuit.

The picture "c" of the Figure 14 is the voltage and current of the solenoid, which applied the reverse power supply strategy at the end of the injection. Although the voltage of picture "b" and "c" both become negative at the end of the injection, the causes of them are totally different. In the reverse power supply strategy, the high side MOSFET H3 and the low side MOSFET L2 are turned on at the end of the injection and the current will flow through path "e". Thus, after the current of the solenoid falls to zero, it will continuously decrease to a negative value. This negative current can establish a reverse magnetic flux, which can help to eliminate the electrical delay of the injection event.

The totally injection delay showed in Figure 15 and 16 indicates that the reverse power supply do improve the closing response of the injection event. But from 996 μs to 916 μs, the 8% improvement is not so encouraging relative to the increase

198

of the system complexity. The possible explanation for this phenomenon is the electrical delay of the injection is not dominant in the total injection delay during the closing process.

a. Non Freewheeling Cut-off   —— Solenoid Voltage
b. Freewheeling Cut-off   ··· Solenoid Current
c. Reverse Power Supply
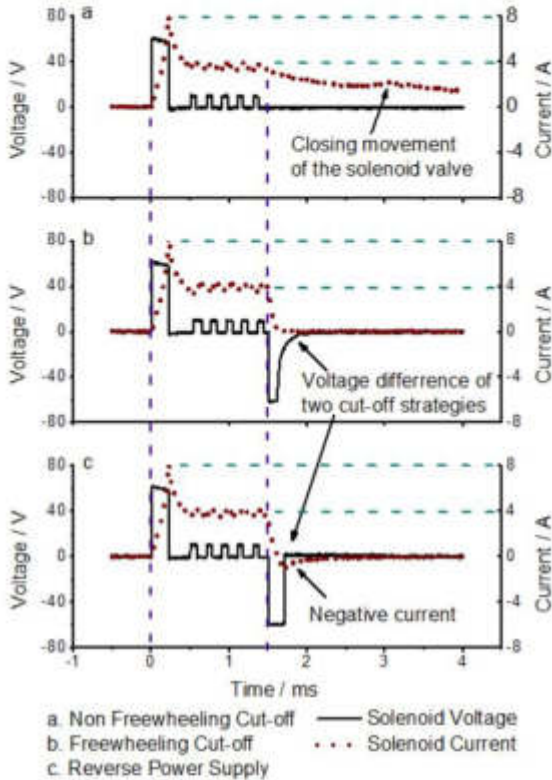
Figure 14. Solenoid voltage and current of different closing strategies

In order to explain the result, the working principle of the solenoid injector must be firstly understood. Through the configuration of the injector in Figure 17, it can be seen that the injection delay characteristic relies mostly on the needle opening and closing rate, which are determined by building up and releasing pressure in injector control chamber, which controlled by the outlet throttle hole. When the solenoid is not energized, the outlet hole is closed, and the pressure on both sides of the piston is the same. Thus, the spring will keep the needle closed. When the solenoid is energized, the outlet hole is released. As a result, the pressure of the control chamber relief, so that the pressure below the piston will push the needle upward. During the closing process, even though the electrical delay is very little and the outlet hole could be closed immediately, it still takes time to establish pressure in the control chamber. At the same time, the injection is not stopped, and oil flow supplied by the common rail is separated to two parts. Therefore, duration of the pressure establishment is longer than that of the decompression process. The hydraulic delay, which determined by the injector structure has a greater impact on the closing delay.
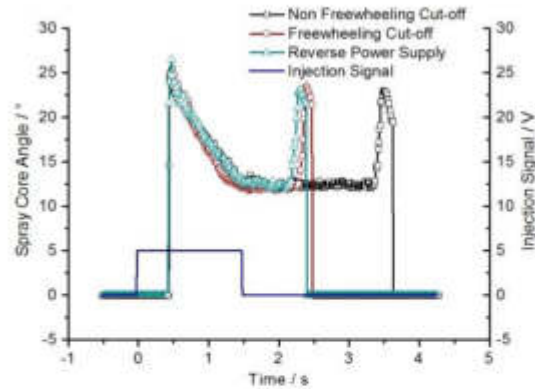
Figure 15. Spray core angle of different closing strategies
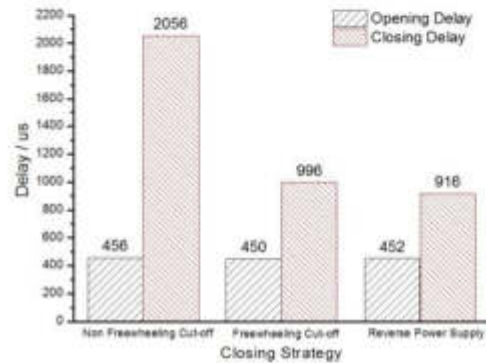
Figure 16. Injection delay of different closing strategies
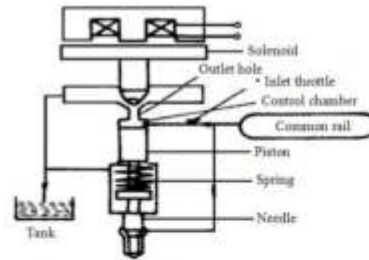
Figure 17. Working principle of the unbalanced valve servo-hydraulic solenoid injector [26]

Effects of Power Supply

The single power supply strategy is investigated in this section. It means all the power is supplied by HV alone. This strategy can simplify the complexity of the driving circuit. The comparison object is also the dual power, 60 V HV, freewheeling cut-off and 100 MPa injection pressure. The results showed in Figure 18 and 19 turn out that the single power supply strategy almost has the same injection response characteristics as the dual power supply. But the current of the "Hold" period has a larger fluctuation than the dual power

supply strategy. This is because if you only use the HV to supply the injector, the duty cycle of the PWM must be very small during the "Hold" period.
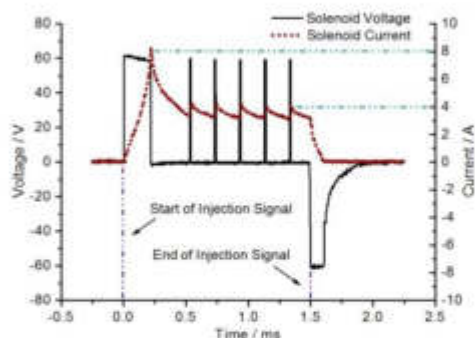
Figure 18. Solenoid voltage and current of single power supply
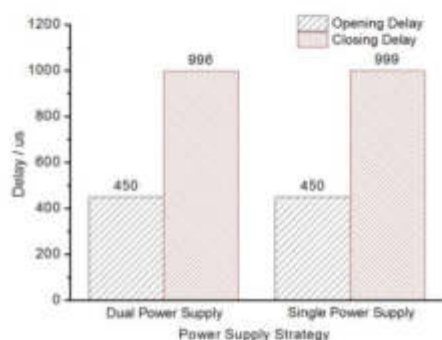
Figure 19. Injection delay of different power supply

CONCLUSIONS

In this study, the impact of different control strategies applied to driving the solenoid injector are investigated based on the high speed imaging system. The spray characteristics with different driving voltage, opening strategy, closing strategy and supply voltage are detected to evaluate the opening and closing delay of the injection event. Moreover, the solenoid voltage and current is measured to research the electrical characteristics of the injector. It can be concluded that:

1. Under the same driving strategies, the higher the supply voltage, the smaller the injection delay (both opening and closing delay) will be. But the effect of the improvement is really obvious, when HV (High Voltage) rises from 40 V to 60 V.

2. The pre-energized opening strategy can greatly speed up the opening response of the injection. Because this strategy can establish the magnetic field in the solenoid beforehand. The improvement of this strategy is better than simply increase the valve of the HV.

3. The freewheeling cut-off strategy is really necessary. It can reduce the closing delay of the injection from 2056 µs to 996 µs. The reverse power supply can also improve the closing

response of the injection event. But its effect is not so worthwhile relative to the increase of the system complexity.

4. Dual power and single power supply can both satisfy the demands of the driving system. The difference of impact on the injection delay is not obvious. But the amplitude of the current fluctuations is larger when the single power strategy is employed.

REFERENCES

1. Davis, William F., et al. "Voltage and current regulator with automatic switchover." U.S. Patent No. RE31,391. 20 Sep. 1983.

2. Harper, Patrick D. "Inductive drive circuit for setting three different levels of load current including a downshift delay." U.S. Patent No. 4,176,387. 27 Nov. 1979.

3. Streit, Klaus, Harsch Klaus, and Schulzke Peter. "Method and apparatus for operating an electromagnetic load, especially an injection valve in internal combustion engines." U.S. Patent No. 4,266,261. 5 May 1981.

4. Woodhouse, Richard Graham, and Salway Peter Hugh. "CONTROL CIRCUITS FOR SOLENOIDS." European Patent No. EP 0008509. 23 Feb. 1983.

5. Ohba, Masahiro. "Injector drive circuit." U.S. Patent No. 4,347,544. 31 Aug. 1982.

6. Takahashi, Yukiharu. "Injector driving circuit." U.S. Patent No. 4,377,144. 22 Mar. 1983.

7. Nielsen, Arnold D. "Solenoid switching driver with fast current decay from initial peak current." U.S. Patent No. 4,511,945. 16 Apr. 1985.

8. Ariyoshi, Hiromi, and Yokomori Iwao. "Current control method and apparatus for electromagnetic valves." U.S. Patent No. 4,520,420. 28 May 1985.

9. Nielsen, Arnold D. "Solenoid driver with switching during current decay from initial peak current." U.S. Patent No. 4,536,818. 20 Aug. 1985.

10. Anttila, Dennis A. "Injector driver control unit with internal overvoltage protection." U.S. Patent No. 4,618,906. 21 Oct. 1986.

11. Kissel, William R. "Electronic fuel injector driver circuit." U.S. Patent No. 4,631,628. 23 Dec. 1986.

12. Grembowicz, Conrad G., and Harrell Brad. "Driver circuit for solenoid operated fuel injectors." U.S. Patent No. 4,905,120. 27 Feb. 1990.

13. Shimanuki, Hiroshi, and Ariyoshi Toshiaki. "Fuel injection valve drive control apparatus." U.S. Patent No. 5,574,617. 12 Nov. 1996.

14. Tsutsui, Teiji. "Apparatus and method for controlling duty solenoid valves." U.S. Patent No. 5,647,387. 15 Jul. 1997.

15. Yanai, Akihiro, Kadowaki Yoshinori, and Fuwa Toshio. "DEVICE FOR DRIVING SOLENOID VALVE." European Patent No. EP 1036964. 20 Sep. 2000.

16. Hoenig, Guenter, et al. "Method and device for controlling at least one solenoid valve." U.S. Patent No. 6,250,286. 26 Jun. 2001.

17. Heinke, Friedhelm. "Method for providing current by means of an inductive component." U.S. Patent No. 6,721,158. 13 Apr. 2004.

18. Kikutani, Takashi. "Fuel injection system for internal combustion engine equipprd with injectors energized through capacitors." European Patent No. EP 1703109. 20 Sep. 2006.

19. Cheever, Gordon D.Jr. "Solenoid driver with high-voltage boost and reverse current capability." U.S. Patent No. 7,349,193. 25 Mar. 2008.

20. Mooney, J., "Drive Circuit Modeling and Analysis of Electronically-Controlled Fuel Injectors for Diesel Engines," SAE Technical Paper 2003-01-3361, 2003, doi:10.4271/2003-01-3361

21. SONG, Jun, et al. "Simulative Analysis for Drive Properties of High-Speed Solenoid Valve in Diesel Engines [J]." Vehicle Engine 5 (2005): 013.

22. Lee, Ill-Yeong. "Switching response improvement of a high speed On/Off solenoid valve by using a 3 power source type valve driving circuit." Industrial Technology, 2006. ICIT 2006. IEEE International Conference on. IEEE, 2006.

23. Ziqiang, Xu Quankui Cheng Zhu Keqing, and Lin Yang. "Design of Solenoid Driving Circuit Module for Electrically Controlled Diesel Engine Based on Circuit Simulation [J]." Transactions of the

200

Lu et al / SAE Int. J. Engines / Volume 7, Issue 3 (August 2014)

Chinese Society for Agricultural Machinery 2 (2008): 005.

24. Huang, Du, et al. "Design of drive circuit for GDI injector." Electric Information and Control Engineering (ICEICE), 2011 International Conference on. IEEE, 2011.

25. Sheng-Nian, Cai, et al. "Energy-saving Driver Circuit of High-speed Solenoid Valve Based on Soft-switch Technology." Instrumentation, Measurement, Computer, Communication and Control (IMCCC), 2012 Second International Conference on. IEEE, 2012.

26. Yang H. M., Su W. H., Wang Y., Shao S. A., "An investigation of the effects of geometric parameters of CR injectors on unevenness of response time and injection Quantity of the injectors," Transactions of CSICE, Vol.18, No.3, pp. 244-249, June 2000.

27. Gao Y., Li L. G., Gao Y., Huang W. D., Deng J., Gu T. L., "Analysis of the difference of spray angles based on various definitions calculation criterions," The 14th Annual Conference on Liquid Atomization and Spray Systems-ASIA, Jeju, Korea, 2010.

CONTACT INFORMATION

Dr. Liguang Li, Prof.
School of Automotive Studies
Tongji University, Shanghai
201804, P. R. China.
liguang@tongji.edu.cn

ACKNOWLEDGMENTS

[14]   Peter, H., Benzin-Direkt-Einspritzsysteme: Komponenten, Funktionen, Diagnose., 1era ed., Krafthand Medien GmbH, Schätzl Druck & Medien, Donauwörth, Alemania, 2015.

2. Das Kraftstoffsystem

Das Kraftstoffsystem ist in einen Niederdruckbereich und in einen Hochdruckbereich unterteilt. Die Niederdruckseite enthält die Bauteile der Vorförderung vom Kraftstoffbehälter bis zur Hochdruckpumpe und kann auf konventionelle Art mit einem Druckregler sowie mit einer bedarfsgerechten Kraftstoffpumpensteuerung ausgerüstet sein.

> ❗ Da der Kraftstoffseite bei der Benzin-Direkteinspritzung eine noch wichtigere Funktion in Bezug auf die Gemischbildung und Abgasemission zukommt, spielt die On-Board-Diagnose eine noch größere Rolle.

Auf der Hochdruckseite befinden sich die Komponenten der Hochdruckerzeugung und Regelung bis hin zum Verteilerrohr und den Injektoren. Alle Bestandteile des Systems sind auf die erhöhten Erfordernisse der Benzin-Hochdruckeinspritzung ausgelegt und daher in der Handhabung und Diagnose entsprechend umsichtig zu behandeln.

Fehlfunktionen werden schon zeitig durch die erweiterte Sensorik anhand von Rechenprozessen erkannt und durch Fehlercodes quittiert. Der Zusammenhang der Fehlermeldungen mit dem tatsächlich defekten Bauteil bedarf der präzisen Systemkenntnis.

Funktionsbeschreibung *Hochdruck*

Der Kraftstoff wird durch eine im Tank angeordnete elektrische Vorförderpumpe über den Kraftstofffilter und den Druckregler zur eigentlichen Hochdruckpumpe gefördert. Der Vorförderdruck ist zumeist von einem federbelasteten Membran-Druckregler auf den gewünschten Niederdruck begrenzt oder er wird von Kraftstoffpumpe bedarfsgerecht generiert.

> ❗ Die Herausforderung von Hochdruck-Einspritzsystemen liegt nicht nur in der Erzeugung des Hochdrucks, sondern vor allen in der Geschwindigkeit des Druckaufbaus. Denn viele Systeme machen aus Effizienz- und Emissionsgründen einen Hochdruckstart, bei dem es gilt, den notwendigen Start-Freigabedruck möglichst schnell zu erreichen.

29

die Nockenwelle : árbol de levas
erhöhen : aumentar
entnehmen : tomar, sacar de un lado.
der Fahrt : recorrido, viaje
Hitten : medio

erträglich : tolerable, aguantable.
das Rahmen : margen, marco.
Dampfblasenbildung : bloqueo de vapor.
vermeiden : evitar.
der Lauf : recorrido, marcha

2 Das Kraftstoffsystem

etlich : algunos / einige
etablieren : establecer
die Vielfalt : diversidad
die Gestaltung : conformación, estructuración, configuración.
derart : de tal manera.
weiterhin : además
Fördermodule : conjunto de alimentación
Verbauposition : posición de construcción

Die nockenwellengetriebene Hochdruck-
pumpe erhöht den Kraftstoffdruck und
fördert diesen zum Kraftstoffverteilerrohr
(Rail), von wo aus die Injektoren den Kraft-
stoff entnehmen und direkt in den Ver-
brennungsraum injizieren.

Eine Herausforderung ist die Kraftstoff-
temperatur, die nach Fahrten mit hoher
Last oder bei Heißstart mit allen Mitteln im
erträglichen Rahmen gehalten werden
muss, um eine Dampfblasenbildung zu ver-
meiden. Im Lauf der Zeit haben sich etliche
Systeme etabliert, die in ihrer technischen
Gestaltung sehr unterschiedlich sind.

aufgeladen : sobrecargado.
aufweisen : exhibir, ostentar, ofrecer.
aufgrund : debido a
Feinheit : refinamiento
Einfluss : influencia

2.1 Kraftstoffsysteme im Überblick

Die Vielfalt der Kraftstoffsysteme ist derart
groß, dass eine Unterteilung nach mehre-
ren technischen Aspekten vorgenommen
werden muss. Die Niederdrucksysteme
lassen sich in bedarfsgeregelte und nicht
bedarfsgeregelte Systeme unterteilen.
Weiterhin gibt es mehrere Variationen von
Fördermodulen und Verbaupositionen von
Kraftstofffiltern. Moderne aufgeladene
Motoren weisen aufgrund der thermischen
Einflüsse speziell bei Niederdrucksyste-
men konstruktive Feinheiten auf.

Die Hochdrucksysteme lassen sich
nach der Art der Druckregelung in hoch-
druckgeregelte und mengengeregelte Sys-
teme unterscheiden. Während die hoch-
druckgeregelten Systeme nur auf eine Art
funktionieren, gibt es bei den mengenge-
regelten Systemen eine Mehrzahl von
Möglichkeiten, die sich oft direkt an der

Mehrzahl : mayoría.
Auslegung : exposición.
Einheit : unidad, conjunto
sogar : incluso.

eingesetzt : establecida.

konstruktiven Auslegung der Hochdruck-
pumpe orientieren. Daher bildet die Art
der Regelung immer eine Einheit mit der
Hochdruckerzeugung *producción* und sogar bis auf
wenige Ausnahmen auch eine Einheit mit
der Art der Niederdruckregelung. Bleibt
noch zu erwähnen, dass die Systemausle-
gung hochgradig von dem konstruktiv be-
stimmten Hochdruckwert und dem ge-
wünschten Volumenstrom abhängig ist.
Diese zwei Faktoren werden bedingt vom
eingesetzten Brennverfahren und den da-
mit in Verbindung stehenden Injektoren
und letztlich von dem Hubraum und der
Leistung des Motors bestimmt.

mengengeregelt : regulado por volumen.
stehend : fijo
hochgradig : en alto grado
Hubraum : cilindrada.
bedingt : condicionalmente

2.1.1 System Melco im Mitsubishi GDI und bei Volvo

Das System des Mitsubishi GDI von Melco
besteht aus einer konventionellen Nieder-
druckseite mit federbelastetem Membran-
druckregler und einer Hochdruckpumpe
mit einem Druckreglergehäuse. Dort wird
ein durch einen hydraulisch-mechani-
schen Druckregler konstanter Wert von 50
bar im Verteilerrail und somit an den Hoch-
druck-Einspritzventilen eingestellt. Ein
piezoelektrischer Drucksensor im Hoch-
druckkreis überwacht den Kraftstoffdruck
und gibt das Signal zur Verarbeitung an das
Motorsteuergerät weiter. Der Hochdruck
kann vom Steuergerät nicht beeinflusst
werden. In der Hochdruckpumpe befinden
sich zudem Flatterventile, die wie ein
Bypassventil wirken. Über das Bypassven-
til wird das System mit Niederdruck wäh-

Ausnahme : excepción
erwähnen : mencionar.
Somit : por consiguiente
aufgrund : debido a

Kreis : circuito.
überwachen : controlar, supervisar.
Verarbeitung : procesamiento, manipulación
wirken : actuar.

nach folgend: subsecuente

der Anstieg: aumento
ständig: constante, permanente
die Beladung: carga.

Übersicht: esquema, vista general.

2.1

Melco GDI Kraftstoff-Systemüber-sicht (blau = Niederdruck 3,3 bar, rot = Hochdruck 45–53 bar, grün = Rücklauf).

Relais Kraftstoffpumpe

ECU

Hochdrucksensor

Verteilerleiste

Kraftstoff-filter

Nieder-Druckregler

Hoch-druckpumpe

Hoch-druckregler

Kraftstoffbehälter

Elektrische Kraftstoffpumpe

Einspritzventile

der Rücklauf: retorno.

Vorlaufphase: fase preliminar

Einschalten: conectar

Zündung: encendido.

gewährleisten: garantizar

bestehen: soportar,

Anwendung: aplicación

Dabei: como resultado

trotz: a pesar de.
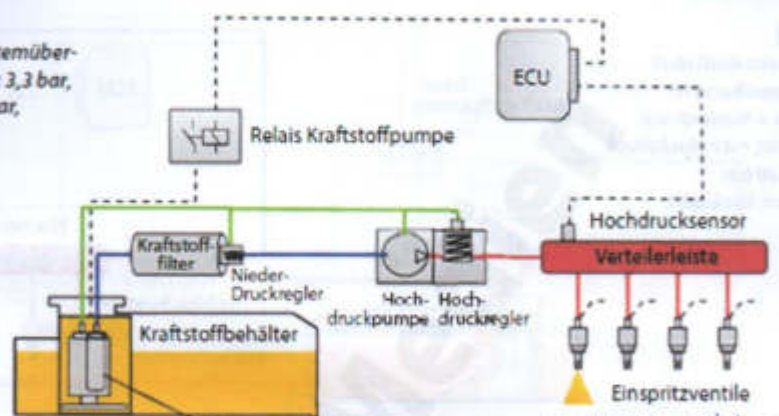
Einbauort: lugar de instalación

Verteilerleiste: riel de combustible.

einstellen: regular, graduar, ajustar.

befindlich: ubicado, situado.

rend der Vorlaufphase der Vorförderpumpe und nach Einschalten der Zündung gefüllt, um nach längeren Standzeiten einen schnellen Druckaufbau zu gewährleisten. Die Hochdruck-Einspritzventile werden mit höherer Spannung und höherem Strom angesteuert, um die neuen Betriebsbedingungen zu bestehen. Dabei kommt ein zusätzliches Düsentreibersteuergerät zur Anwendung. Auch bei diesem System ist ein Aktivkohlebehälter eingebaut, der in Intervallen gespült werden muss. Die Einbauorte liegen im Motorraum oder am Tank. Der Kraftstoff muss trotz der hochdruckseitigen Regelung nicht gekühlt werden.

2.1.2 System Siemens der 1. Generation

Das Kraftstoffsystem von Siemens ist mit einer variablen Hochdruckregelung ausgestattet. Ein Kraftstoffsystem dieser Bauart wird in der ersten Generation der Direkt-

einspritzer bei den HPI-Motoren von PSA, dem Renault IDE, dem 2,2 Direct von Opel und im M271 bei Mercedes-Benz eingesetzt.

Das Hochdruck-Regelventil wird vom Steuergerät angesteuert, um hochdruckseitig den von der Pumpe erzeugten Druck in der Verteilerleiste und somit an den Injektoren kennfeldgerecht einzustellen. Der abgeregelte Kraftstoff gelangt vom Regelventil über eine kurze Rücklaufleitung in den Niederdruckkreis auf die Saugseite der Hochdruckpumpe. Die durch die Entspannung in den Kraftstoff eingetragene Wärme wird dadurch nicht in den Kraftstoffbehälter zurückgeführt. Dort würde der nachfolgende Temperaturanstieg zur ständigen Beladung des Aktivkohlefilters führen.

Die Niederdruckregelung erfolgt zumeist am nahe dem Kraftstoffbehälter befindlichen Kraftstofffilter mit integriertem federbelasteten Membran-Druckregler. Der Niederdruck wird auf konstant 3,5 bar

ocurre (above "erfolgt")

abgeregelt: regulado

ausgestattet: equipado

Bauart: arquitectura, estructura.

zumeist: la mayoría de las veces

Rücklaufleitung: conexión de retorno.

Saugseite: lado de succión.

eingetragen: registrado

gelangen: llegar, alcanzar

dadurch: de este modo.

31

überschüssig : excedente
lieferant : proveedor.
versorgt : provisto

2 Das Kraftstoffsystem

2.2

Siemens Kraftstoff-Systemübersicht (blau = Niederdruck 3,5 bar, rot = Hochdruck bis 120 bar, grün = Rücklauf).

2.3
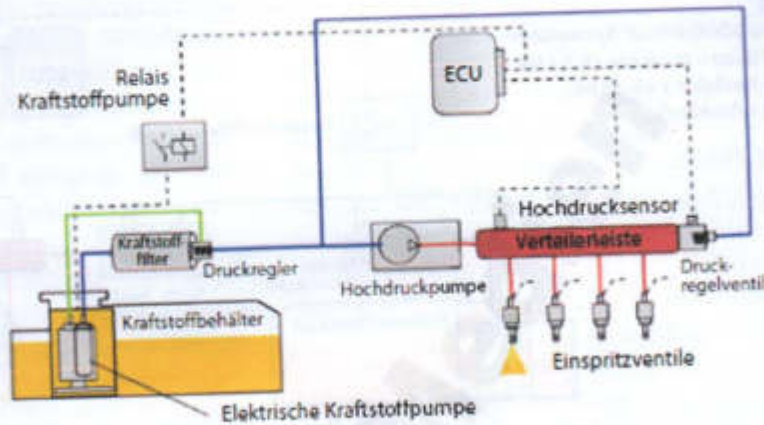
Bosch/VW Kraftstoff-Systemübersicht der ersten Generation (blau = Niederdruck 3,5 bar, rot = Hochdruck 50–100 bar, grün = Rücklauf).
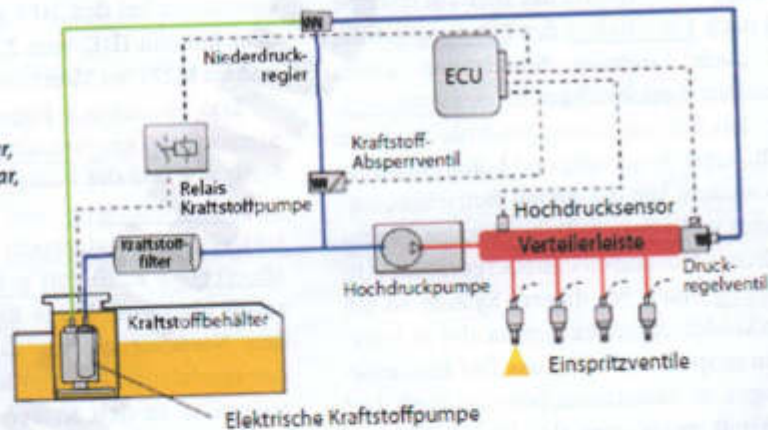
geregelt. Auf kurzem Weg gelangt der überschüssige Kraftstoff von dort aus drucklos in den Kraftstoffbehälter zurück. Die Hochdruckerzeugung ist herstellerspezifisch. So ist das System bei Mercedes und Renault mit einer Dreizylinder-Radialkolbenpumpe und bei PSA und Opel mit einer Dreizylinder-Axialkolbenpumpe ausgestattet.

2.1.3 System Bosch mit HDP 1 im VW FSI

Das Kraftstoffsystem der ersten Generation FSI bei VW vom Systemlieferanten Bosch wird niederdruckseitig von einer In-Tank-Elektro-Kraftstoffpumpe mit Kraftstoff versorgt. Der Niederdruck wird von einem im Motorraum befindlichen

2.4

System des Audi FSI
mit Bosch HDP 2
und konventioneller
Vorförderung mit
rücklaufloser Nieder-
druckregelung
direkt an der In-Tank-
Pumpeneinheit
(blau = Niederdruck
6 bar, rot = Hochdruck
40–120 bar,
grün = Rücklauf).
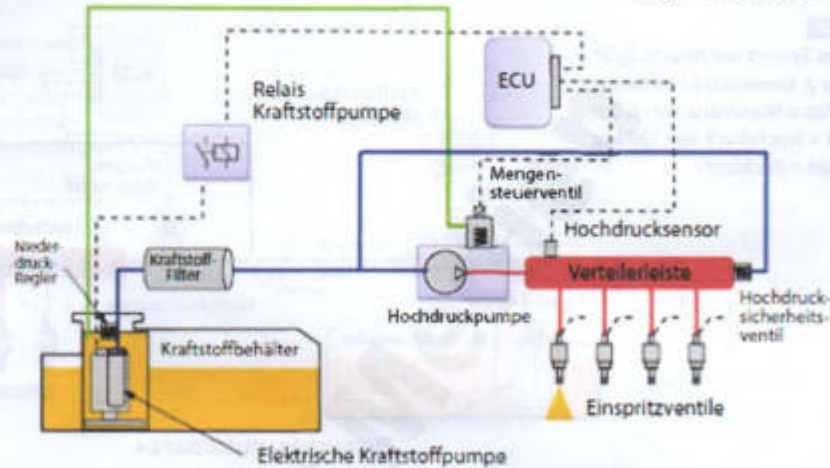
Membrandruckregler auf konstant 3 bar ge-
regelt. Bei Heißstarts allerdings verschließt
ein Absperrventil die Vorlaufleitung des
Druckreglers von der Kraftstoffpumpe.
Zur Vermeidung von Dampfblasen steigt
der Druck so bis auf den Öffnungsdruck
des Überdruckventils in der Elektro-Kraft-
stoffpumpe von etwa 6 bar an.

Die Hochdruckpumpe ist eine Dreizy-
linder-Radialkolbenpumpe und läuft im-
mer mit voller Förderleistung. Der Hoch-
druck wird je nach Betriebszustand auf 50
bar im Leerlauf bis etwa 100 bar oberhalb
von 4.000/min von einem Druckregelventil
im Rail geregelt. Das Druckregelventil
übernimmt bei der Platzierung im Hoch-
druckbereich gleichzeitig die Funktion der
Druckbegrenzung ab etwa 120 bar. Der
überschüssige Kraftstoff wird mit der
Rücklaufmenge der Hochdruckpumpe und
des Niederdruckreglers in den Kraftstoff-
behälter zurückgeleitet.

2.1.4 System Bosch mit HDP 2 im Ford SCI und Audi 2,0 FSI

Im SCI-Motor von Ford und im 2,0 Liter
FSI-Motor von Audi kommt erstmals ein Sys-
tem mit mengengeregelter Hochdruckpum-
pe HDP 2 von Bosch zum Einsatz. Das Sys-
tem von Toyota ist identisch. Es ist allerdings
ausschließlich mit eigenen Komponenten im
Hochdruckbereich ausgestattet. Die Nieder-
druckförderung findet durch eine Elektro-
Kraftstoffpumpe statt, die sich in einem
Schwalltopf im Kraftstoffbehälter befindet.

In der Einheit befindet sich gleichzeitig
der Niederdruckregler, der den überschüs-
sig geförderten Kraftstoff gleich vor Ort in
den Behälter zurückführt. Man spricht in
diesem Fall von einem rücklauflosen Nie-
derdrucksystem. Zusätzlich befindet sich
bei Ford noch das bekannte Absperrventil
(Dosierventil) zur Heißstart-Druckanhe-
bung im Niederdruckkreis. Im Gegensatz
dazu fährt Audi das Niederdrucksystem

Dampf : vapor
aufwendig : laborioso
fällen : cortar, talar
dafür : por esto

schaltung : conexión del circuito.
entsprechend : correspondiente.
Befehl : mando, orden, comando
masseseitig : de la de tierra.

2 Das Kraftstoffsystem
obliegen : corresponder.

aufnahme : admisión
leistung : capacidad

zusätzlich : adicionalmente
erhielten : conseguir, obtener

zuführen : conducir, transportar.

2.8

Das System mit Hitachi HDP
der 2. Generation und Decos:
(blau = Niederdruck 4 – 6 bar,
rot = Hochdruck 40 – 120 bar,
grün = Rücklauf).

Kraftstoffpumpen-steuergerät

ECU

Druckbegrenzungs-ventil

Kraftstoff-filter

4 – 6 bar

Niederdruck-sensor

Mengen-steuerventil

Hochdrucksensor

Verteilerleiste

Hochdruckpumpe

Hochdruck-sicherheits-ventil

Kraftstoffbehälter

Einspritzventile

Elektrische Kraftstoffpumpe
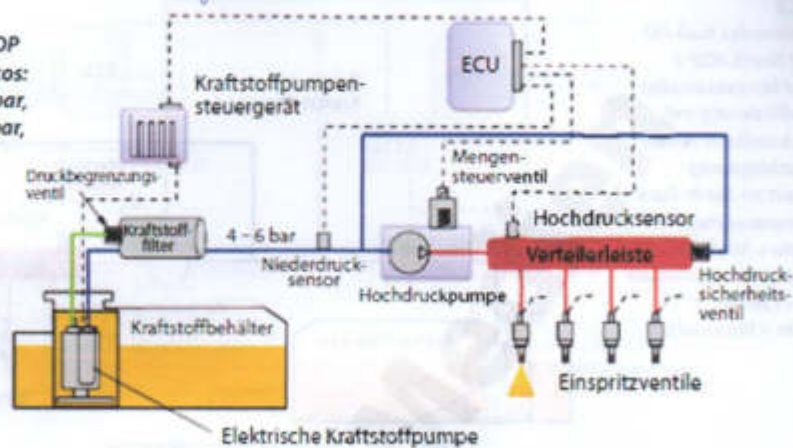
zur Dampfblasenvermeidung gleich mit aufwendigen 6 bar. Die Vordruckumschaltung fällt dafür bei diesen Motoren weg. Die Schaltung der Elektro-Kraftstoffpumpe übernimmt ein entsprechendes Relais, welches den Schaltbefehl masseseitig an der Spule des Steuerstromkreises vom Motorsteuergerät bekommt. Die Hochdruckerzeugung obliegt der Einzylinder-Hochdruckpumpe mit angebautem Mengensteuerventil, welches für eine Reduzierung der Leistungsaufnahme der Hochdruckerzeugung sorgt. Der überschüssige Kraftstoff gelangt so nicht erst in den Hochdruckkreis sondern wird über den Rücklauf zum Kraftstoffbehälter zurückgeführt. Das nun zusätzlich im Rail verbaute Hochdrucksicherheitsventil führt die Leckagemenge wieder dem Niederdruckbereich vor der Hochdruckpumpe zu. Somit steht der Rücklauf unter Vorförderdruck. Der Raildrucksensor, das Motorsteuergerät und das Mengensteuerventil bilden den Hochdruck-Regelkreis.

2.1.5 System Bosch HDP 2 und Hitachi 2. Generation mit Decos

Das System mit Bosch HDP 2 kann ebenso mit einem Niederdruckkreis mit bedarfsgeregelter Kraftstoff-Vorförderung, Decos, ausgestattet sein wie das System mit der Hitachi HDP der 2. Generation. Decos bedeutet: demand controlled fuel supply – bedarfsgeregelte Kraftstoffförderung. Die Ansteuerung der Elektro-Vorförderpumpe übernimmt ein Leistungsteil in Form eines Kraftstoffpumpensteuergerätes, welches nahe der Pumpeneinheit auf dem Kraftstoffbehälter oder auf der Abdeckung darüber verbaut ist. Die Vorförderpumpe wird drehzahlgeregelt und erzeugt damit einen vom Steuergerät angestrebten Niederdruck. Womit alle Bauteile der bisherigen Regelung entfallen. Dazu erhielt das System einen Niederdrucksensor, der in einem Regelkreis mit dem Motorsteuergerät, dem Pumpensteuergerät und der Vorförderpumpe einen Istwert als Regelungsgrund-

Abdeckung : directorio, cerramiento
angestrebt : aspirado, perseguido

womit : con lo cual.
bisherig : previo
entfallen : suprimirse

34

Istwert : valor actual

gering: pequeña, bajo, reducido, escaso
angemessen: apropiado, apropiado.
beherrschbar: manejable, dominable.
auszeichnen: distinguir
bevorzugt: preferido apropi favorecido

2.6

Das aktuelle System mit Bosch HDP 5 oder Hitachi HDP der 3. Generation und bedarfsgeregelter Vorförderung mit rücklauflosem Hochdruckkreis (blau = Niederdruck 4 – 6,5 bar, rot = Hochdruck).

Somit = por consiguiente

Senkung: reducción, descenso, disminución

Dadurch: de este modo

benötigen: necesitar, precisar

Dämpfung: amortiguación

Drossel: estrangulación

Kraftstoffpumpensteuergerät

ECU

Mengensteuerventil
Hochdrucksensor

4 – 6 bar
Niederdrucksensor
Verteilerleiste

Kraftstoffbehälter
Hochdruckpumpe mit Druck-Begrenzungsventil

Einspritzventile

Kraftstofffilter
Elektrische Kraftstoffpumpe
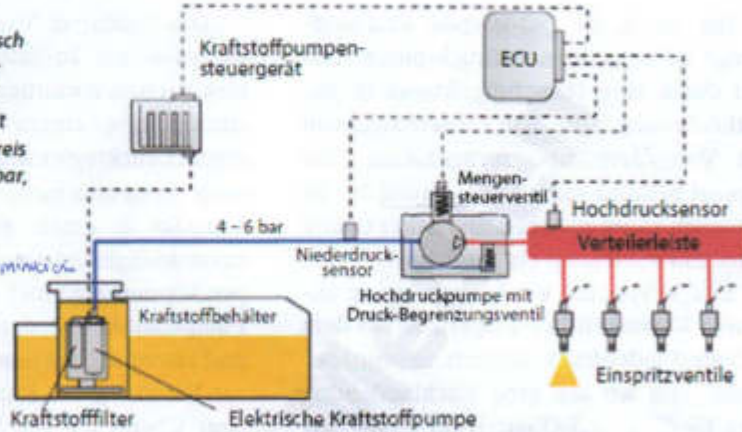
lage liefert. Somit ist eine Senkung der Leistungsaufnahme des Niederdrucksystems möglich, da der Kraftstoff nun schon auf der Vorförderseite bedarfsgerecht limitiert an die Hockdruckerzeugung gelangt. Der Hochdruckkreis ist auch bei der Hitachi HDP mit einem externen Hochdruck-Sicherheitsventil versehen. Im Unterschied zu der Bosch HDP 2 besitzt die Hochdruckpumpe keinen externen Rücklauf. Dadurch benötigt das System in der Vorförderleitung zur Hochdruckpumpe eine Dämpfungsdrossel welche die Druckpulsationen der Absteuermenge dämpfen soll. Die Kraftstoffsysteme für aufgeladene Motoren besitzen am Kraftstofffilter ein Druckbegrenzungsventil von dem aus der Rücklauf direkt in dem Kraftstoffbehälter mündet.

münden: desembocar

2.1.6 System Bosch HDP 5 und Hitachi 3. Generation mit Decos

In aktuellen Benzinhochdruck-Einspritzanlagen kommen vorwiegend Systeme

anlage: construcción

vorwiegend: predominantemente

mit bedarfsgeregeltem Niederdrucksystem und mengengeregelten Hochdruckpumpen zum Einsatz. Diese Systeme zeichnen sich durch ihre geringe Leistungsaufnahme und angemessene Pumpenbelastung im Niederdruckbereich sowie auch bei Turbomotoren beherrschbare Kraftstofftemperaturen aus.

Die Hochdruckerzeugung findet bevorzugt mit Einzylinder-Radialkolbenpumpen wie Hitachi Generation 3, Bosch HDP 5 oder auch mit saugmengengeregelten Axialkolbenpumpen statt (wie beispielsweise bei BMW, PSA oder Porsche).

Eher die Ausnahme bilden großvolumige Motoren, die zwei Pumpen benötigen, um die Volumenströme zu generieren. Die Systeme haben immer kompaktere Komponenten. Beispielsweise ist das Druck-Sicherheitsventil direkt in die Pumpe integriert und das den abgeregelten Kraftstoff intern auf die Saugseite in den Niederdruckbereich zurückführt. Somit entfällt der Rücklauf komplett.

Eher: antes, más temprano, anteriormente

Druckspitze · pico de presión
hervorgeruft: avocado
fernhalten: mantener a distancia.
Kaum: apenas, difícilmente.

einzeichnen: marcar
ebenfalls: igualmente, asimismo
festgelegt: determinado.

2 Das Kraftstoffsystem

Die durch die pulsierende Kraftstoff-menge hervorgerufenen Druckspitzen wer-den durch eine Dämpfungsdrossel in der Vorlaufleitung nahe dem Fördermodul von der Vorförderpumpe ferngehalten. Die Drossel befindet sich in der Leitung. Sie ist äußerlich kaum zu sehen und in der Grafik (Abb. 2.6) auch nicht eingezeichnet.

Einige Systeme werden mit einem ex-ternen Kraftstofffilter ausgerüstet, in dem sich ein Niederdruck-Sicherheitsventil be-findet, von wo aus eine Rücklaufleitung zum Fördermodul-Oberteil im Kraftstoff-behälter führt.

Der Niederdrucksensor ist bei einigen Systemen ebenfalls nicht mehr verbaut, da das Steuergerät mit einer Testroutine in festgelegten Zyklen einen Ersatzwert er-mitteln kann. Das äußerliche Fehlen eini-ger Komponenten soll dennoch nicht über die Komplexität dieses Systems hinweg-täuschen. Gerade im Niederdruckbereich hat sich bei aufgeladenen Motoren einiges geändert.

2.2 Das Kraftstoff-Niederdrucksystem

Der Niederdruckkreis kann bei Benzindi-rekt-Einspritzanlagen je nach Anforde-rung an das Gesamtsystem sehr unter-schiedlich ausgelegt sein. Wie bei der Saugrohreinspritzung gibt es auch hier die Varianten mit Kraftstoffrücklauf und ohne Kraftstoffrücklauf sowie mit Membran-druckregler und mit bedarfsgeregelter Vorförderung, wo der Druckregler entfällt.

Das Kraftstoff-Niederdrucksystem be-steht aus der In-Tank-Einheit mit einer elektrischen Vorförderpumpe, einer Saug-strahlpumpe, einem **Kraftstofffilter** und einem Druckregler. **Alle Bauteile** können je nach Systemvariante getrennt wie auch kompakt in einem **Förderpumpenmodul** zusammengefasst sein. Die Kraftstoffpum-pen können **auf** einer Achse mit nur einer Pumpenstufe für den Hauptförderdruck und zweistufig mit einer Strömungspumpe zur Befüllung des Pumpenraumes mit ei-nem Überdruck von 0,25 bar und einer Hauptförderpumpe in einem Gehäuse ausgeführt sein. Der Druckregler ist bis auf eine Ausnahme entweder am Kraftstofffil-ter platziert oder er befindet sich ebenfalls im Förderpumpenmodul. Das Förderpum-penmodul enthält außerdem das Grob-An-saugsieb und den Tankfüllstandsgeber.

Bei Satteltanks gibt es noch eine teles-kopartige Ansaugverlängerung mit Saug-strahlpumpe für die andere Tankhälfte. Über das Vorsieb am Boden des Förder-pumpenmoduls wird der Kraftstoff gefil-tert und durch die Saugstrahlpumpe an-gesaugt. So füllt sich der Schwalltopf. Ein Klappenventil sorgt für sofortigen Ver-schluss der Leitung, sobald der Kraftstoff-strom abbricht. So wird ein Leerlaufen des Systems verhindert. Die nach dem Venturi-Prinzip wirkende Saugstrahlpumpe wird über den Volumenstrom der Rücklaufmen-ge der elektrischen Kraftstoffpumpe be-trieben.

Die Kraftstofffilter für die Benzin-Di-rekteinspritzung müssen eine sehr feine Filtrierung des Kraftstoffs zum Schutz der

[15]   Pauer, T., Yilmaz, H., Zumbrägel, J. et al., "New Generation Bosch Gasoline Direct-injection Systems." MTZ Worldwide, vol. 78, pp. 16–23, https://doi.org/10.1007/s38313-017-0053-6, (2017).

COVER STORY MIXTURE FORMATION AND COMBUSTION

New Generation Bosch Gasoline Direct-injection Systems

Due to future exhaust gas emission legislations, vehicle manufacturers need to implement additional emissions reductions. Based on three factors – market development, legislation and end customers – Bosch has developed a new generation of gasoline direct-injection systems. According to the engineers, approximately 80 % lower PN emissions can be achieved under RDE conditions.

© Bosch

16
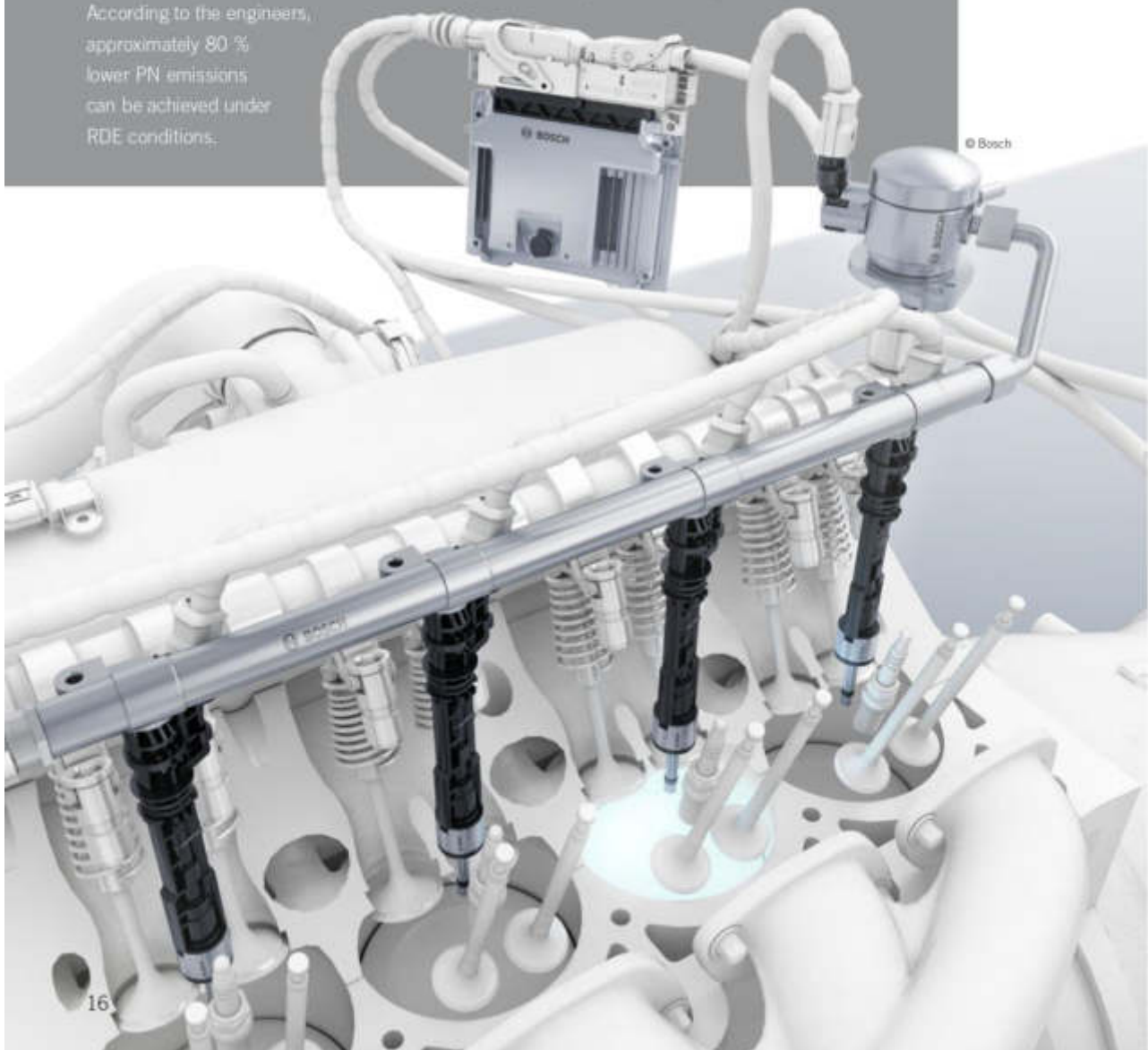
AUTHORS

Dr. Thomas Pauer
is Executive Vice President, Division
Gasoline Systems, and responsible
for Engineering and Combustion
Engines at Robert Bosch GmbH
in Schwieberdingen (Germany).

Hakan Yilmaz, M. Sc.
is Vice President System Engineering
Powertrain, Division Gasoline
Systems at Robert Bosch GmbH
in Schwieberdingen (Germany).

Dr. Joachim Zumbrägel
is Vice President Engineering
Gasoline Direct Injection of the
Business Unit Direct Injection (DI),
Division Gasoline Systems,
at Robert Bosch GmbH
in Schwieberdingen (Germany).

Dr. Erik Schünemann
is Department Manager
Engineering Combustion Systems,
Division Gasoline Systems,
at Robert Bosch GmbH
in Schwieberdingen (Germany).

REQUIREMENTS ON NEW GASOLINE DI SYSTEMS

Today, gasoline engines are the preferred solution for passenger car powertrains due to favourable unit costs, high power density, potential for low exhaust gas emissions and high robustness against worldwide fuels. Compared to 70 million gasoline engine cars in 2015, 78 million gasoline engine powered vehicles are forecasted for 2025, **FIGURE 1**. While gasoline engine volumes with port fuel injection (PFI) are expected to decline, gasoline DI engines are predicted to double during this time frame. Consequently, an estimated 39 million PFI engines as well as 39 million DI engines can be expected in 2025.

Due to future exhaust gas emission legislations, vehicle manufacturers need to implement additional reductions of gaseous and particulate emissions (particulate mass PM and particulate number PN). Exceptionally stringent requirements regarding gaseous emissions are demanded by the CARB LEVIII legislation, **FIGURE 2**. The reduction of fleet limits for NMOG and NO_x given by this legislation requires the fulfilment of SULEV 30 limits as a fleet average. Other important automotive markets, for example China (CN 6a/b and Beijing 6) or India (BS 6) are converging quickly to the European (Euro 6d) and US (LEV III/Tier 3) legislations respectively. Furthermore, Europe, China and India will ask for testing the real driving emissions (RDE) during vehicle operation on public roads. Likewise, requirements on the manufacturers regarding useful life and evaporative emissions (EVAP) will become more severe. Worldwide CO_2 legislations are calling for further reduction of tank-to-wheel CO_2 emissions. End customers expect functional benefits with every new car, for example with respect to engine and driving performance, fuel consumption and comfort or NVH.

Gasoline engines with direct-injection and turbocharging can make a contribution to the fulfilment of these demands by the manufacturers. Direct-injection is permitting both high specific power and an increased compression ratio. It provides potential for optimised catalyst heating strategies that can compensate for the additional heat sink given by the turbine. Today's gasoline DI engines predominantly apply multihole injectors with solenoid actuation that are installed into the cylinder head in side-mounted or central-mounted position and operate with system pressures of up to 25 MPa. Using laser drilling of fuel holes is allowing for large flexibility with respect to spray lay-
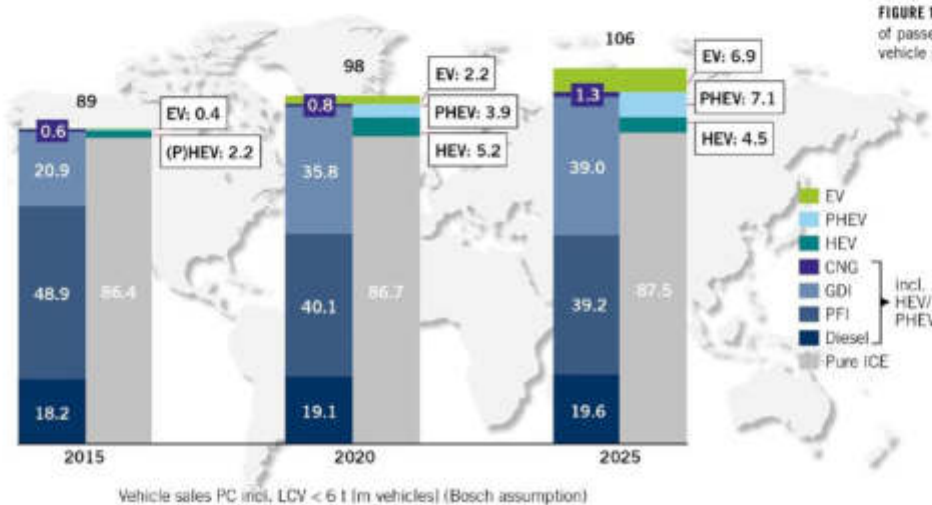
FIGURE 1 Market development of passenger cars and light-duty vehicle powertrains (© Bosch)

Vehicle sales PC incl. LCV < 6 t (m vehicles) (Bosch assumption)

	LEV III SULEV 20	EU 6d	BS 6	CN 6b national	Beijing 6
	MY 2015 (phase-in)	9/2017 (new types)	4/2020	7/2023 (all types)	12/2017
PN/PM	PM: $10-3-1$ mg/mi DI & PFI	PN: $6*10^{11}$/km DI	PN: $6*10^{11}$/km DI	PN: $6*10^{11}$/km DI und PFI	PM: 1.86 mg/km DI
NO$_x$	20 mg/mi NO$_x$+NMOG	60 mg/km	60 mg/km	35 mg/km	43.5 mg/km NO$_x$+NMHC
Cycle	FTP 75	WLTC + RDE	MIDC 90 + RDE	WLTC + RDE	FTP 75
EVAP	300 mg	2000 mg	2000 mg	700 mg	650 mg

FIGURE 2 Exhaust gas emission legislation in the US (LEV III), Europe, India and China (top); LEV III legislation and fleet limits (bottom) (© Bosch)
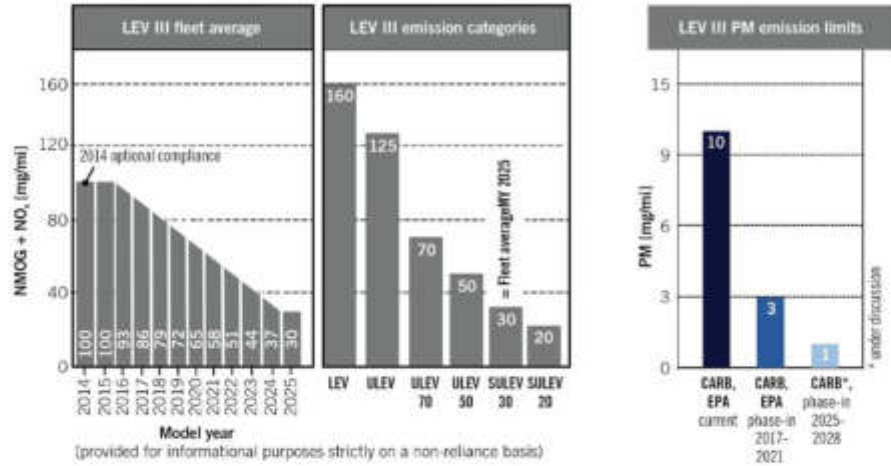
out. Central-mounted injector position close to the spark plug enables highly efficient catalyst heating operation using the injection of a small fuel quantity close to the ignition for fastest catalyst light-off and reduced exhaust gas emissions [1].

For this purpose, in addition to both appropriate spray layout and respective injection strategy, a software functionality is needed (Bosch CVO: Controlled Valve Operation) that guarantees stable injector operation with a small fuel quantity injection inside the ballistic area of the characteristic line over lifetime [2]. Incomplete fuel-mixture preparation of liquid fuel films formed at the combustion chamber walls during injection is the major source of particulate formation inside gasoline DI engines [3]. Optimising the spray layout in combination with the charge motion of the engine as well as using multi-injection strategies can already reduce piston and cylinder liner wetting considerably and hence, lower particulate emissions. Further optimisation of fuel metering and fuel-mixture preparation by an advanced DI fuel system can make a strong contribution to additional substantial improvements regarding particulate emissions [4].

18

According to the requirements of market development, worldwide legislation and end customers, the following development guidelines have been set for the new generation of Bosch DI systems:
– optimise fuel-mixture formation for potential of lowered particulate

emissions by increasing DI system pressure to 35 MPa together with improved NVH performance
– extend injector operation in combination with CVO software functionality
– advance multi-injection capabilities by reduced pause time

FIGURE 3 Variant frame and NVH performance of the high-pressure pump HDP6 (© Bosch)

Bosch NVH test bench, idle condition, 700 rpm engine, 4-lobe cam with standard flow 1.0 ccm/rev camshaft
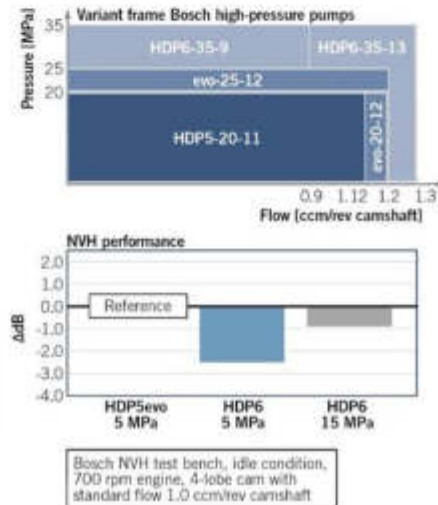
- improve injector tip geometry for contribution to reduced particulate emissions due to injector tip wetting
- provide compact system design, high modularity and flexibility for compliance with customer specific requirements
- worldwide utilisability with respect to EVAP, useful life and fuel qualities.

In the following, the technical solutions on component level including software functionalities are described and the functional results realised on system level are presented.

SYSTEM PRESSURE GENERATION WITH HIGH-PRESSURE PUMP HDP6

Increasing the DI system pressure to 35 MPa required the development of a new generation of high-pressure pumps, **FIGURE 3**. Bosch high-pressure pump HDP6 offers an enhanced variant frame with higher flexibility regarding fuel connections, angular positions and installation on the engine. The modular concept permits a choice of two piston diameters. While the small piston diameter variant minimises required driving power at volume flows of up to 0.9 ccm/rev camshaft, the large piston diameter variant allows volume flows of up to 1.3 ccm/rev camshaft. The NVH performance of the new HDP6 under idle conditions represents "best in class".
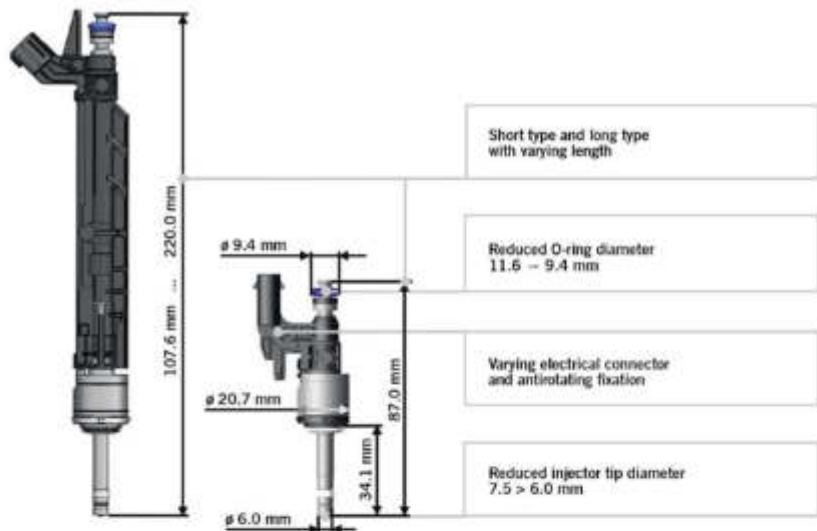
FIGURE 4 Variant frame of the multihole injector HDEV6 (© Bosch)

VARIANT FRAME OF MULTIHOLE INJECTOR HDEV6

The new HDEV6 is designed for DI system pressures of up to 35 MPa. A highly flexible variant frame, **FIGURE 4**, comprises amongst others short and long injector versions, different electrical connectors as well as antirotating fixations and permits an installation into side-mounted and central-mounted position. A compact injector design with an injector tip diameter having been decreased from 7.5 to 6.0 mm and an O-ring diameter of the fuel connection which has been lowered from 11.6 to 9.4 mm is reducing required installation space. Injector weight has been cut by 20 %.

HDEV6 FUEL METERING AND SOFTWARE FUNCTIONALITIES

The HDEV6 qdyn characteristic line shown in **FIGURE 5**, exhibits a mono-
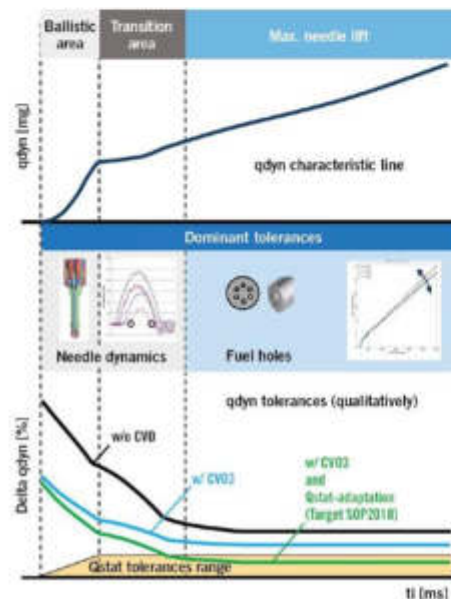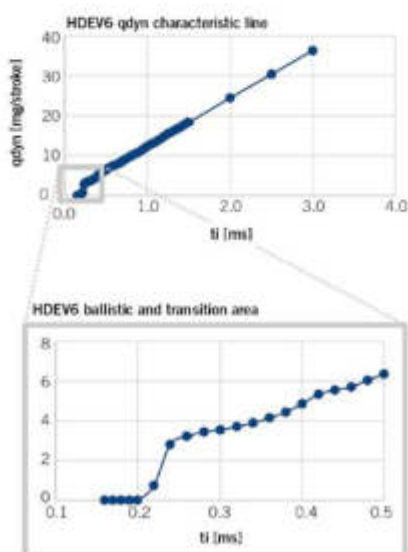
FIGURE 5 HDEV6 qdyn characteristic line with ballistic area and transition area (left); reduction of HDEV6 qdyn tolerance using Bosch CVO3 and Qstat-adaptation (right) (© Bosch)
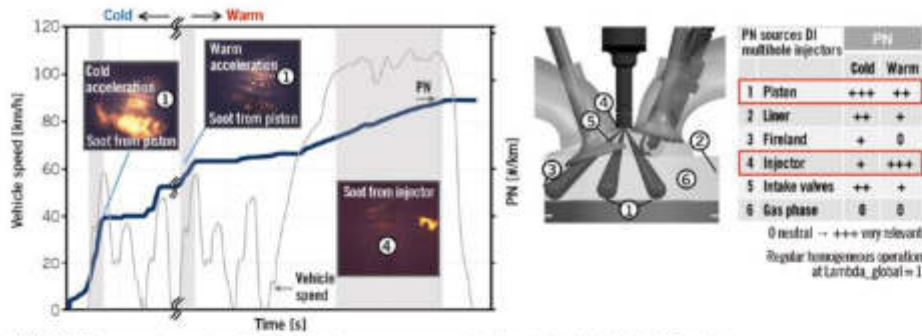
FIGURE 6 PN sources at gasoline DI engines in homogeneous operation (Lambda_global = 1) (© Bosch)

tonically increasing dependence of injected fuel mass qdyn on the injector actuation time ti over the complete operating range. Bosch CVO software functionality was further improved for operation with HDEV6 and is in the 3rd generation (CVO3) now working over the complete injector operating range (ballistic area, transition area, linear area), **FIGURE 5**.

CVO3 reduces deviations of the actual injected fuel mass from the tar-

get value over the complete injector operating area. This feature significantly contributes to fulfilment of future useful life requirements. Additional improvements can be expected by a software functionality for the adaptation of static flow addressing relevant injector parameters, which is currently under development (SOP 2018). HDEV6's multi-injection capability was considerably

improved by internal injector measures, thereby providing injector pause times smaller than 1 ms.

The reduction of valve seat leakage over injector lifetime is an important target for further improvement regarding evaporative emissions (EVAP). HDEV6 was further optimised with respect to the relevant valve seat properties and provides an improved valve seat leakage after durability run, enabling a robust
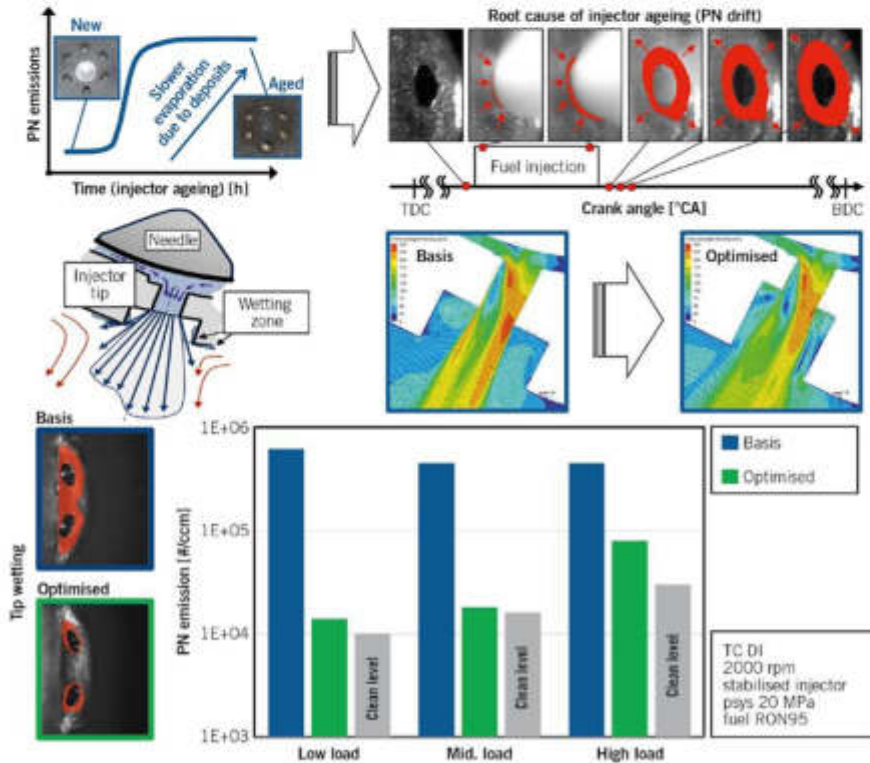
FIGURE 7 Particulate formation at the DI injector tip and optimisation measure as well as injector tip-wetting and stationary PN emission after optimisation compared to baseline injector and clean level (© Bosch)

MIXTURE PREPARATION AND REDUCTION OF PARTICULATE EMISSIONS

Soot particle formation inside combustion engines can occur if both high temperatures and very rich conditions are present simultaneously during the combustion process. For gasoline direct-injection and premixed gasoline engine combustion with globally stoichiometric mixture, this can appear, if due to the interaction of liquid fuel and combustion chamber walls a fuel wall film is formed. Regular stoichiometric premixed combustion is largely finished, when the flame front reaches the respective combustion chamber walls. Then, both high gas temperature and low oxygen content are existent. Soot particles are formed if the aforementioned fuel wall film is not evaporated and homogenised at the time when the flame front arrives at the affected combustion chamber wall. Particulate formation can be visualised easily by detection of the typical soot luminance using optical measurement techniques. Particulate formation due to wetting of intake or exhaust valves, combustion chamber roof and spark plug can be avoided by positioning DI injector and spark plug accordingly and by a thoroughly designed spray of the injector. Although contact of liquid fuel with piston and cylinder liner can be minimised by proper spray layout and injection strategy, typically it cannot be completely avoided under all boundary conditions.

Consequently, particles are formed especially during accelerations under cold engine conditions as well as with a warm engine, FIGURE 6. Here, the particulate formation at the injector tip itself is contributing to the overall particulate emission level. Due to injector internal fuel flow, spray breakup and interaction of fuel spray and fuel-hole/pre-hole as well as interaction with the surrounding air inside the cylinder during the injection process, fluctuations of the spray jets in radial direction can occur. These fluctuations can lead to a liquid fuel film build-up at the injector tip. At the arrival of the regular flame front with high gas temperatures and low oxygen content, remaining fuel at the injector tip can form soot particles. A portion of these soot particles remains at the injector tip and can lead to deposit formation. These deposits can absorb more liquid fuel and consequently are self-reinforcing the process while strongly increasing PN emissions ("PN drift") until a balance of deposit formation and deposit reduction mechanisms is realised at the injector tip ("stabilised injector"). Therefore, minimising injector tip-wetting is one of possible measures to reduce PN drift.

Internal injector flow and geometry of the injector tip can be improved in a way that injector tip-wetting is reduced considerably, resulting in a significantly increased robustness against deposit formation, FIGURE 7. For this purpose, several development tools have been used intensively, for example CFD simulation, near-field and far-field spray measurements with high temporal and spacial resolution and engine tests. By application of the optimisation measures, PN

emissions resulting from the injector tip are lowered considerably. FIGURE 7 depicts the current status compared with baseline injector and clean injector tip free of deposits at 20 MPa. A further increase of the DI system pressure to 35 MPa, which is feasible with HDEV6, results in an additional improvement of robustness against PN drift [1].

Furthermore, an injection pressure increase results in a SMD (Sauter Mean Diameter) reduction by improved primary breakup, leading to a raising liquid surface area provided for evaporation of the injected fuel, FIGURE 8 (left). At the same time, the entrainment of ambient air into the spray is elevated significantly (air-entrainment), FIGURE 8 (middle). Thus, enthalpy supporting fuel evaporation is provided. Consequently, the fuel droplets are following the air-motion much better, resulting in a strong decrease of liquid fuel film at the respective piston and cylinder liner area that is responsible for particulate formation. The mentioned effect of increased injection pressure on fuel wall film formation is shown by fundamental investigations, FIGURE 8 (right) [5].

By combining all hardware measures with an advanced calibration, PN emissions under RDE conditions – and for engine start temperatures and ambient temperatures of 20 °C – can be reduced by about 80 %, FIGURE 9. Due to delayed fuel evaporation because of cold combustion chamber walls and low in-cylinder gas temperatures, low temperature conditions present a particular challenge regarding PN emissions. Here, in addition to the available engine internal measures and an advanced calibration as well as quickest readiness of the

FIGURE 8 Influence of DI system pressure on SMD, air-entrainment and fuel film build-up (© Bosch)

215
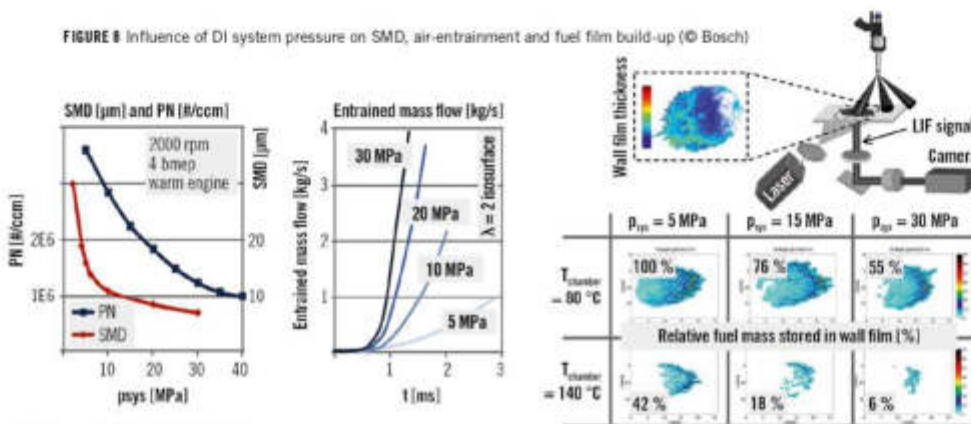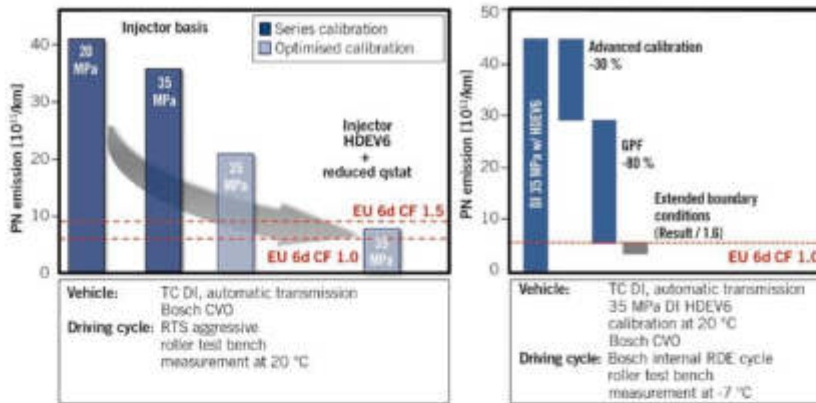
FIGURE 9 Status of engine internal PN reduction measures in the RTS aggressive cycle at 20 °C and status of PN emission reduction in Bosch internal RDE cycle at -7 °C (© Bosch)

lambda control, a gasoline particulate filter (GPF) is necessary to support manufacturers in achieving the future RDE

In order to maximise the functional benefits of the DI injector, an optimal integration into the customer-specific engine concept and combustion system is necessary. Therefore, an overarching project organisation has been established within Bosch Gasoline Systems. A combined application of development tools for injector, spray and combustion system reaching from component level to vehicle level, FIGURE 10, as well as a close cooperation with the customer starting in the early engine development phase are crucial preconditions for achieving optimal overall results.

SUMMARY

Based on the requirements of market development, legislation and end customers, Bosch has developed a new generation of gasoline direct-injection systems. Essential features are the optimised mixture formation for lowered particulate emissions by increasing the DI system pressure to 35 MPa while at the same time improving the NVH performance, an enhanced injector operating range with narrowed tolerances using software functionalities CVO3 and Qstat-adaptation, improved multi-injection capabilities. An optimised injector tip geometry contributes to the PN emission reduction due to an injector tip-wetting. The new genera-

tion is providing compactness, high modularity and flexibility for the fulfilment of customer-specific requirements. It is designed for worldwide operation regarding EVAP, useful life and fuel qualities.

Approximately 80 % lower PN emissions can be achieved under RDE conditions by optimising DI injector and calibration. For particularly challenging low temperature conditions, in addition to internal engine measures, an advanced calibration and quickest readiness of the lambda control, a gasoline particulate filter (GPF) is necessary to support vehicle manufacturers in reaching the future RDE conformity factors of 1.0. To ensure optimal integration of Bosch DI injectors into the indi-
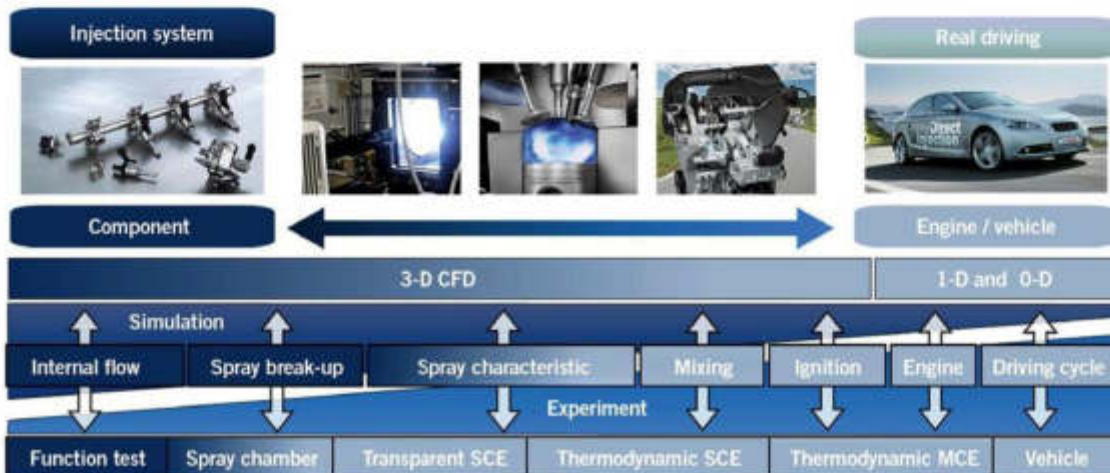
FIGURE 10 Development of injector, spray and gasoline combustion systems at Bosch (© Bosch)

vidual engine concept and combustion system together with the customer, an overarching project organisation has been established within Bosch Gasoline Systems. This organisation is applying a range of Bosch development tools starting from the component to the vehicle. The new generation of gasoline direct-injection systems is currently launched into series production and will be rolled out into the worldwide production network step by step. The introduction is supported by implementation of an integrated Industry 4.0 value stream.

REFERENCES
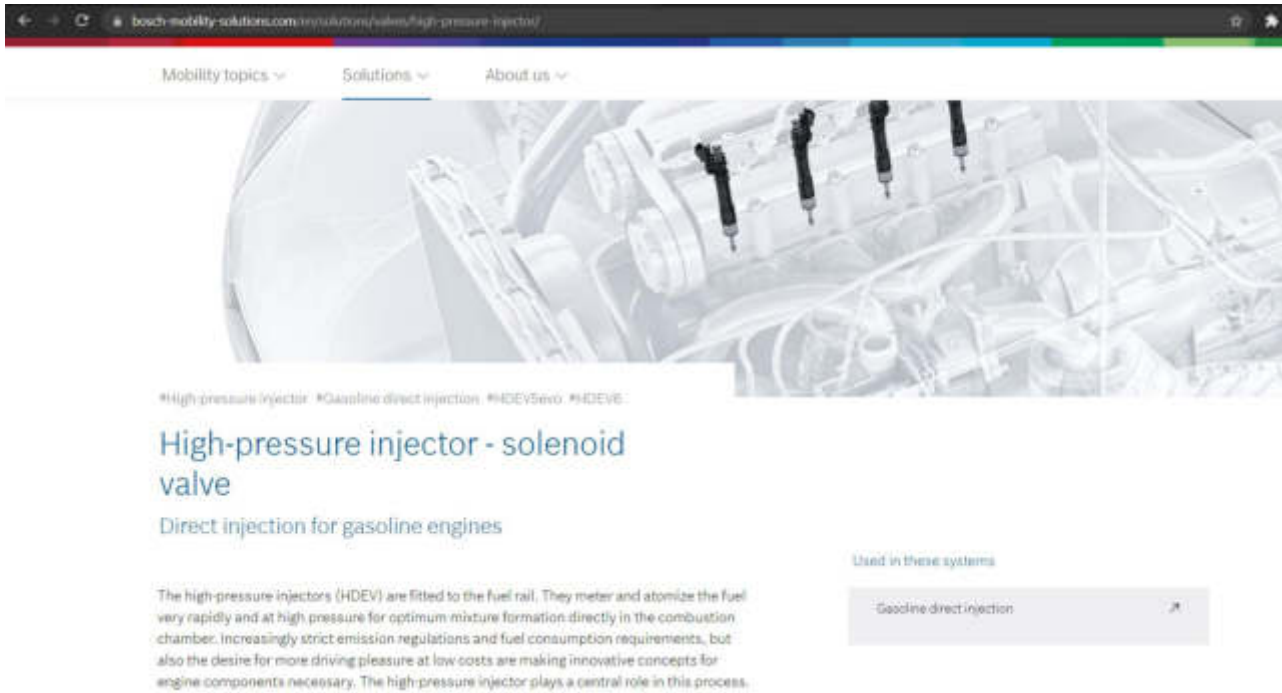
[1] Kufferath, A.; Wiese, W.; Samenfink, W.; Dageförde, H.; Knorsch, T.; Jochmann, P.: Assessment of Feasible System Solutions for Future Particle Emission Requirements. Conference IMECHE – Fuel Systems for IC Engines, London, 2015
[2] Schlüter, R.; Kümpel, J.; Okuyama, H.: Mechatronic Component Packages within Gasoline Direct Injection Systems and their Impact on OEM-Supplier-Cooperation. 7th IFAC Symposium on Advances in Automotive Control, Tokyo, 2013
[3] Kufferath, A.; Berns, S.; Hammer, J.; Busch, R.; Frank, M.; Storch, A.: EU 6 als Herausforderung für die Benzindirekteinspritzung – Eine Bewertung zukunftsfähiger Systemansätze. 33. International Vienna Motor Symposium, Vienna, 2012
[4] Wiese, W.; Kufferath, A.; Storch, A.; Rogler, P.: FIE requirements in GDI engines to meet Future emissions legislation. 2. International Engine Congress, Baden-Baden, 2015
[5] Kufferath, A.; Samenfink, W.; Hammer, J.; Schulz, F.; Könnig, M.; Schmidt, J.: Charakterisierung des Wandfilms relevanter Betriebsbedingungen für einen direkteinspritzenden Ottomotor als Grundlage zur Schadstoffminimierung. Conference 10. Tagung Motorische Verbrennung, Munich, 2011

THANKS

[16]   (2021) Bosch Mobility Solutions, [Online], disponible en: https://www.bosch-mobility-solutions.com/en/

[17]   Reif, K., Automobilelektronik: eine Einfürungfür Ingenieure, 3ra ed., Vieweg &Teuber, GWV Fachverlage GmbH, Wiesbaden, Alemania, 2009.

Die Modellierung der Füllungserfassung ist durch die Vielzahl an Variabilitäten, die Einfluss auf die Zylinderfüllung haben, äußerst komplex. Derzeit existieren unterschiedliche Modellansätze, die in der Regel teils physikalisch, teils empirisch aufgebaut sind. Insbesondere bei Motoren mit Abgasturboaufladung wird ebenfalls eine Modellierung der Prozesse im Abgassystem benötigt. Über den Brennraumdruck sind Einlass- und Auslasssystem miteinander gekoppelt, gerade bei Ventilsteuerzeiten mit großer Überschneidung der Öffnungsphasen von Einlass- und Auslassventilen (d. h. Einlass- und Auslassventile sind gleichzeitig geöffnet). Der Frischluftmassenstrom und die Restgasmasse hängen in diesen Betriebszuständen stark vom Druck im Abgassystem ab. Füllungserfassungssysteme von Turbomotoren umfassen daher gekoppelte Modelle für Einlass- und Auslasskomponenten. Die Füllungssteuerung und die Füllungserfassung sind konsistent ausgelegt und greifen auf identische Modellvorstellungen zurück, um eine robuste und stabile Laststeuerung zu gewährleisten.

5.4.2 Gemischbildungsfunktionen

Aufgabe dieser Funktionsgruppe ist es, sowohl in stationären als auch in transienten Betriebszuständen den Kraftstoff in der erforderlichen Menge und in dem gewünschten Aufbereitungszustand der Verbrennung zum richtigen Zeitpunkt zur Verfügung zu stellen. Die exakte Dosierung des Kraftstoffs ist insbesondere für die Einhaltung der Emissionsgrenzwerte von entscheidender Bedeutung. Bei direkteinspritzenden Motoren im Schichtbetrieb führt zudem bei der so genannten Qualitätsregelung eine abweichende Kraftstoffmasse direkt zu einer Momentenänderung. Eine fehlerhafte Aufbereitung des Kraftstoffs hinsichtlich des Zeitpunkts der Einspritzung oder der Einspritzstrahlgeometrie und Zerstäubung gefährdet die Entflammung des Gemischs und kann Aussetzer hervorrufen. Die wesentlichen Aufgaben der Gemischbildungsfunktionen lassen sich zusammenfassen zu: Berechnung der Grundeinspritzmasse, Regelung auf ein gewünschtes Luftverhältnis und zeitliche Koordination der Einspritzung.

Berechnung der Grundeinspritzmasse

Im homogenen Motorbetrieb wird in der Regel ein stöchiometrisches Luftverhältnis angestrebt. Dazu wird zunächst auf Basis der aktuellen Lastinformation aus dem Füllungsmodell und der aktuellen Drehzahl die erforderliche Einspritzmasse berechnet. Dieser Basiswert wird durch mehrere Korrekturen verändert (siehe Bild 5-9a). Diese berücksichtigen Instationäreffekte (Wandfilmkompensation) sowie Anpassungen der Kraftstoffmasse für den Start oder den Warmlauf. Ebenso wird die Kraftstoffmasse aus der Tankentlüftung eingerechnet. Im Schichtbetrieb bei Direkteinspritzmotoren wird die Grundeinspritzmasse aus dem Momentenwunsch ermittelt.

Regelung auf ein gewünschtes Luftverhältnis

Um in allen Betriebszuständen auch mit toleranzbehafteten Bauteilen sicherzustellen, dass das gewünschte Luftverhältnis eingestellt wird, ist der Vorsteuerung eine Regelung hinzugefügt (siehe Bild 5-9b). Dazu wird im Abgas das aktuelle Luftverhältnis mit Hilfe von λ-Sonden (auch λ-Sensoren genannt) erfasst und mit dem Sollwert verglichen. λ-Sonden messen den Sauerstoffgehalt des Abgases und erlauben so die Bestimmung des Luftverhältnisses. Der Aufbau und die Funktion von λ-Sonden werden z. B. in [Bo4] erklärt. Man unterscheidet dabei im Wesentlichen zwei Varianten der λ-Sonden. Eine *Sprungsonde* gibt abhängig vom Sauerstoffgehalt des Abgases eine Spannung U_λ aus, die beim Übergang zwischen $\lambda < 1$ und $\lambda > 1$

und umgekehrt einen charakteristischen Sprung aufweist (siehe Bild 5-10a). Der Messbereich der Sprungsonde ist daher auf den Bereich um $\lambda = 1$ beschränkt. Mit der *Breitband-λ-Sonde* kann die Sauerstoffkonzentration im Abgas über einen weiten Bereich bestimmt und damit auf das Luftverhältnis λ geschlossen werden. Das Ausgangssignal ist bei der Breitband-λ-Sonde ein Strom I_p, der vom Sauerstoffgehalt des Abgases und damit vom Luftverhältnis λ abhängt (siehe Bild 5-10b). Die optimale Effizienz des Dreiwegekatalysators hinsichtlich Oxidation von HC und CO und Reduktion von NO_x ist gegeben, wenn durch eine gezielte Schwankung des Luftverhältnisses um den Wert eins der Sauerstoffspeicher befüllt und entleert wird. Dazu erfolgt bei Systemen mit Breitband-λ-Sonden eine periodische Modulation des λ-Sollwertes mit einer Amplitude von 1 bis 3 Prozent, die als Zwangsanregung bezeichnet wird.

versorgung: suministro,

(a) *lüftung: ventilación*

(b)
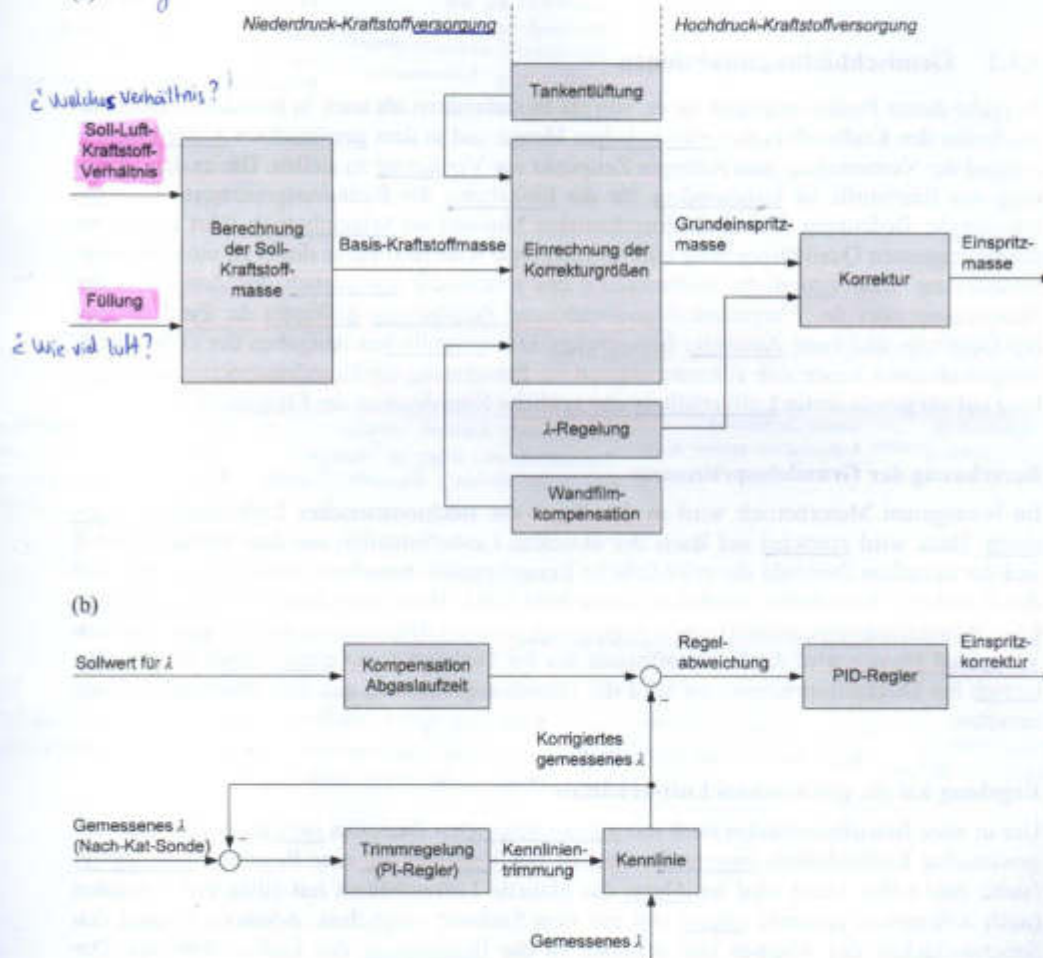
Bild 5-9 Funktionsstruktur der Gemischbildungsfunktionen: (a) Übersicht. (b) Einspritzkorrektur durch λ-Regelung
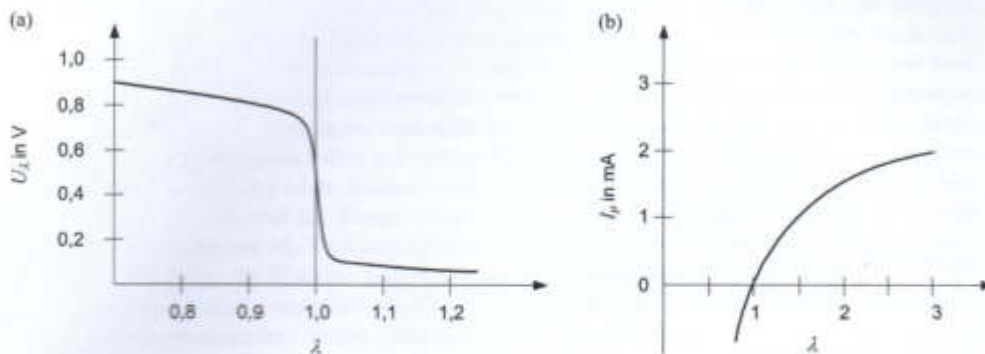
Bild 5-10 Typische Kennlinien von λ-Sonden: (a) Sprungsonde. (b) Breitband-λ-Sonde.
U_λ Spannung an der Sprungsonde, I_p Strom durch die Breitband-λ-Sonde, λ Luftverhältnis

Um im Mittelwert exakt ein stöchiometrisches Luftverhältnis einzuhalten, wird bei manchen Systemen eine zusätzliche Sprungsonde hinter dem Katalysator verwendet, die auch Nach-Kat-Sonde genannt wird. Diese λ-Sonde besitzt durch die niedrigeren Temperaturschwankungen und die geringere Querempfindlichkeit bezüglich weiterer Emissionskomponenten eine höhere Genauigkeit als die λ-Sonde vor dem Katalysator. Allerdings ist die Messung hinter dem Katalysator deutlich zeitverzögert. Der auf der Nach-Kat-Sonde basierende Regelanteil wird als Trimmregelung bezeichnet und korrigiert den λ-Sollwert oder die Sondenkennlinie der Sonde vor dem Katalysator.

Der λ-Regler ist meist als PID-Regler oder ähnlich ausgeführt. Um den Regler zu entlasten und um in den Zeitanteilen des Motorbetriebs, in denen die λ-Sonden nicht betriebsbereit sind, eine Korrektur der Grundeinspritzung zu erhalten, verfügt die λ-Regelung über eine Adaption, in der beispielsweise der I-Anteil des Reglers abgespeichert wird. Ein hoher Regelbedarf kann durch Toleranzen sowohl des Luftpfades (Luftmassenmesser, Drucksensor) als auch des Kraftstoffpfades (Einspritzventile) hervorgerufen werden.

Zeitliche Koordination der Einspritzung

Die korrigierte Einspritzmasse wird anschließend in die Ansteuersignale der Einspritzkomponenten umgerechnet. In Abhängigkeit von der Batteriespannung, den Drücken und den Temperaturen wird zunächst die benötigte Öffnungsdauer der Einspritzventile bestimmt. Besonders für direkteinspritzende Motoren ist die zeitliche Positionierung des Einspritzvorgangs entscheidend, da im Schichtbetrieb ein zündfähiges Gemisch zum Zündzeitpunkt an die Zündkerze gebracht werden muss. Daher wird in der Software sichergestellt, dass auch bei Mehrfacheinspritzung die zyklischen Einspritzimpulse zeitgerecht und konsistent zur Zündungsausgabe erfolgen. Dies ist auch für Betriebsartenumschaltungen entscheidend, bei denen zwischen Homogen- und Schichtbetrieb gewechselt wird. Eine Voraussetzung für die exakte Dosierung des Kraftstoffs ist die Einhaltung des gewünschten Kraftstoffdrucks. Dazu existiert bei Hochdrucksystemen eine Druckregelung, die in Abhängigkeit des Betriebspunktes den gewünschten Kraftstoffdruck exakt einstellt.

anschließend: seguidamente, a.la seguida, a continuación.

Daher: además

vorgang: curso, proceso, suceso.

Betriebsarten umschaltung:
cummulación / activación por tipo de operación.

Schichtbetrieb: operación estratificada / Schichtladung: estratificación.

mehrfach: múltiple, repetido.

Einhaltung: cumplimiento.

erfolgen: ocurrir, suceder.

ausgabe: distribución

[18]   Reif, K., Grundlagen Fahrzeug-und Motortechnik, Bosch Fachinformation Automobil, Springer Vieweg, Friedrichschafen, Alemania, 2017.

Reif, K., Grundlagen Fahrzeug- und Motorentechnik, Bosch Fachinformation Automobil, Springer Vieweg, Friedrichschafen, Alemania, 2017.

nicht durchströmte Rails (Returnless System, ohne Rücklauf. Bild 14). Die Einspritzventile sind direkt am Kraftstoffverteilerrohr montiert. Neben den Einspritzventilen kann bei Systemen mit Rücklauf auch ein Kraftstoffdruckregler und eventuell im Kraftstoffverteilerrohr ein Druckdämpfer integriert werden.

Die gezielte Auslegung von Abmessungen des Kraftstoffverteilerrohrs verhindert örtliche Druckänderungen durch Resonanzen beim Öffnen und Schließen der Einspritzventile. Last- und drehzahlabhängige Unregelmäßigkeiten der Einspritzmengen werden dadurch vermieden. Abhängig von den Anforderungen der verschiedenen Fahrzeugtypen besteht das Kraftstoffverteilerrohr aus Edelstahl oder Kunststoff. Zu Prüfzwecken und zum Druckabbau im Service kann ein Diagnoseventil integriert sein.

Ergänzend zu den Benzin- und Ethanol-Kraftstoffverteilerrohren gibt es auch Kraftstoffverteilerrohre für den Erdgasbetrieb (siehe Bild 16). Hierbei kommen spezielle Gasventile zum Einsatz. Die Druck- und Temperaturüberwachung erfolgt über einen entsprechenden Sensor.

Benzin-Direkteinspritzung

Einleitung

Die Benzin-Direkteinspritzung ermöglicht eine effektive Weiterentwicklung von Ottomotoren hinsichtlich Verbrauch und Abgas, bei der auch die Fahrdynamik und der Fahrkomfort nicht zu kurz kommen muss. Sie ist der Schlüssel für effektives Downsizing von Ottomotoren und ermöglicht Verbrauchseinsparungen bis zu 20 %. Durch die Synergie von Benzin-Direkteinspritzung, Abgasturboaufladung und einer variablen Nockenwellensteuerung können Drehmomente und Motorleistungen realisiert werden, die bislang nur größeren Motorhubräumen und -zylinderzahlen vorbehalten waren.

Für den Fahrer äußert sich dies z. B. im hochdynamischen Ansprechverhalten des Fahrzeugs bei Geschwindigkeitsänderungen, was im heutigen Straßenverkehr einen Komfort- und einen Sicherheitsaspekt darstellt. Überdies lässt die Benzin-Direkteinspitzung eine Gesamtoptimierung des Antriebs zu, um kostengünstige Abgasnachbehandlungskonzepte für künftige Emissionsgrenzen, wie z. B. EU6 in Europa und SULEV in USA, darzustellen.

Übersicht

Die Forderung nach leistungsfähigen Ottomotoren bei gleichzeitig niedrigem Kraftstoffverbrauch und niedrigen Emissionen führte zur Wiederentdeckung der Benzin-Direkteinspritzung. Gegenüber Saugrohr-Einspritzsystemen bietet die Benzin-Direkteinspritzung zusätzliche Freiheitsgrade aufgrund der inneren Gemischbildung. Sie bietet die Grundlage moderner und leistungsfähiger Brennverfahren wie z. B. des Schichtmagerbetriebs oder der homogenen Kompressionszündung (HCCI). Bei Turbomotoren mit stöchiometrischer Verbrennung ergeben sich Vorteile im Drehmoment

---

Handwritten notes (left column):

Weiterentwicklung: perfeccionamiento
hinsichtlich: con respecto a
zu kurz kommen: quedarse corto
einsparung: ahorro.
Aufladung: sobrecarga/sobre alimentación
Nockenwellen: árboles de levas
Drehmoment: torque
bislang: hasta ahora
vorbehalten: reservar, reservado.
verhalten: comportamiento
darstellen: constituir, representar,
heutig: moderno.
außern: manifestar, expresar
Überdies: aparte de esto.
zulassen: permitir, admitir.

Handwritten notes (right column):

Antrieb: tracción, propulsión.
Kostengünstig: barato
Künftig: futuro (adjetivo)
Forderung: exigencia
Leistungsfähig: potente, de gran potencia
Entdeckung: descubrimiento
Wiederentdeckung: redescubrimiento
gegenüber: frente a, comparado.
zusätzlich: adicional
aufgrund: debido a
ergeben: producir, arrojar

[Handschriftliche Notizen am oberen Rand:]

Überschneidung : traslape
gering : pequeño, mínimo, bajo, reducido.
Klopfneigung : tendencia a la detonación / neigung : inclinación

damals : por entonces.
aufwendig : costoso, gastoso
anforderung : exigencias, requisito
benötigt : necesitados, requerids.

Dauerhaltbarkeit : durabilidad.
verhinderten : imposibilitar, obstaculizar.
durchbruch : salida, irrupción.

[Handschriftliche Notizen am linken Rand:]
...pfung : evaporación
...ya en
...aplicación, utilización, funcionamiento
...incorporar, instalar.
...por primera vez
...sig : en serie
...funcionamiento
...zen : distinguir, caracterizar
...kt : adjunto
...formular, sugir
...dentro de
...demandado, exigido
...demandar
...estable, permanente
...acumular

[Links unten:]
...Betätigung...
...ventil balula velm...
...rze mit auf...
...der Zünd-
...ventil balula esc...
...ack-Ein-
...
...Verteiler-
...

im unteren Drehzahlbereich durch eine erhöhte Überschneidung der Ladungswechselventile und durch die geringere Klopfneigung aufgrund der Verdampfung des Kraftstoffs im Brennraum.

Das Prinzip ist nicht neu. Bereits 1937 kam ein Flugzeugmotor mit einer mechanischen Benzin-Direkteinspritzung zum Einsatz. 1951 wurde ein Zweitakt-Motor mit einer mechanischen Benzin-Direkteinspritzung erstmals serienmäßig in einem Pkw, dem Gutbrod, eingebaut. 1954 folgte der Mercedes 300 SL mit einem Viertakt-Motor und Direkteinspritzung.

Die Konstruktion eines direkteinspritzenden Motors war für die damalige Zeit sehr aufwendig. Zudem stellte diese Technik hohe Anforderungen an die benötigten Werkstoffe. Die Dauerhaltbarkeit des Motors war ein weiteres Problem. All diese Probleme verhinderten über eine lange Zeit den Durchbruch der Benzin-Direkteinspritzung.

Arbeitsweise

Benzin-Direkteinspritzsysteme sind durch eine Hochdruckeinspritzung direkt in den Brennraum gekennzeichnet (Bild 17). Das Luft-Kraftstoff-Gemisch entsteht wie beim Dieselmotor innerhalb des Brennraums (durch innere Gemischbildung). Das Kraftstoffsystem besteht aus Elektrokraftstoffpumpe, Hochdruckpumpe, Rail, Hochdrucksensor und den Einspritzventilen (Bild 18).

Hochdruckerzeugung

Die Elektrokraftstoffpumpe (Bild 18, Pos. 10) fördert den Kraftstoff mit dem Vorförderdruck von 3...5 bar zur Hochdruckpumpe (11). Diese erzeugt abhängig vom Betriebspunkt (gefordertes Drehmoment und Drehzahl) den Systemdruck. Der unter Hochdruck stehende Kraftstoff gelangt in das Rail (12) und wird dort gespeichert. Der Kraftstoffdruck wird mit dem Hochdrucksensor (13) gemessen und über das in der
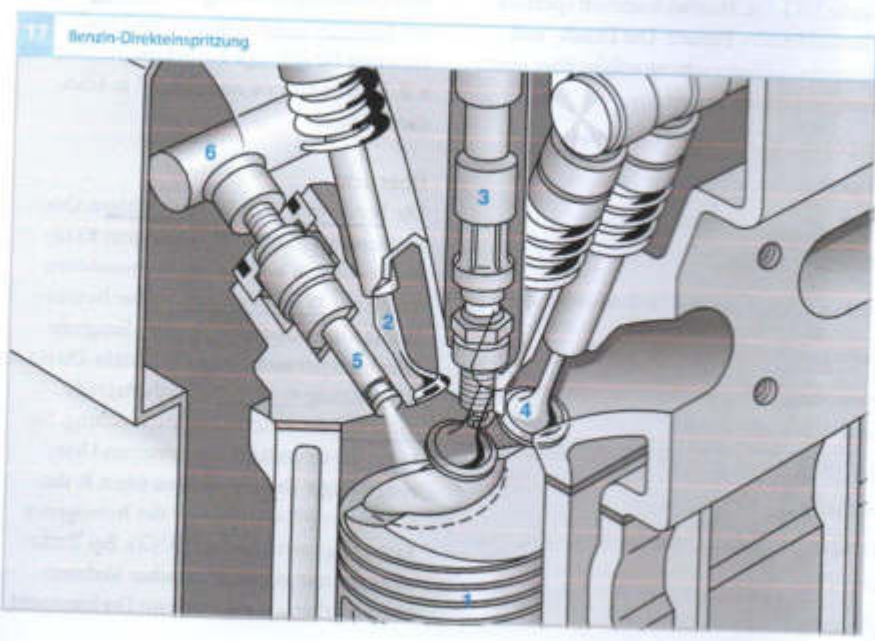
17 Benzin-Direkteinspritzung

[handwritten notes, top:]

Umgebung: antorno, ambiente
Vorrichtung: mecanismo, dispositivo
Ladungsbewegungsklappe: aleta de movimiento de carga
versteller: ajustador, regulador.
Steller: actuador, controlador.
Schub: empuje   Schub-Umluftventile: blow-off valve
Leuchte: luz (no Kocamente)

Weghfahrsperre: inmobilizador.

Strukturbild eines Ottomotors mit Benzin-Direkteinspritzung einschließlich Komponenten für die Steuerung und Regelung
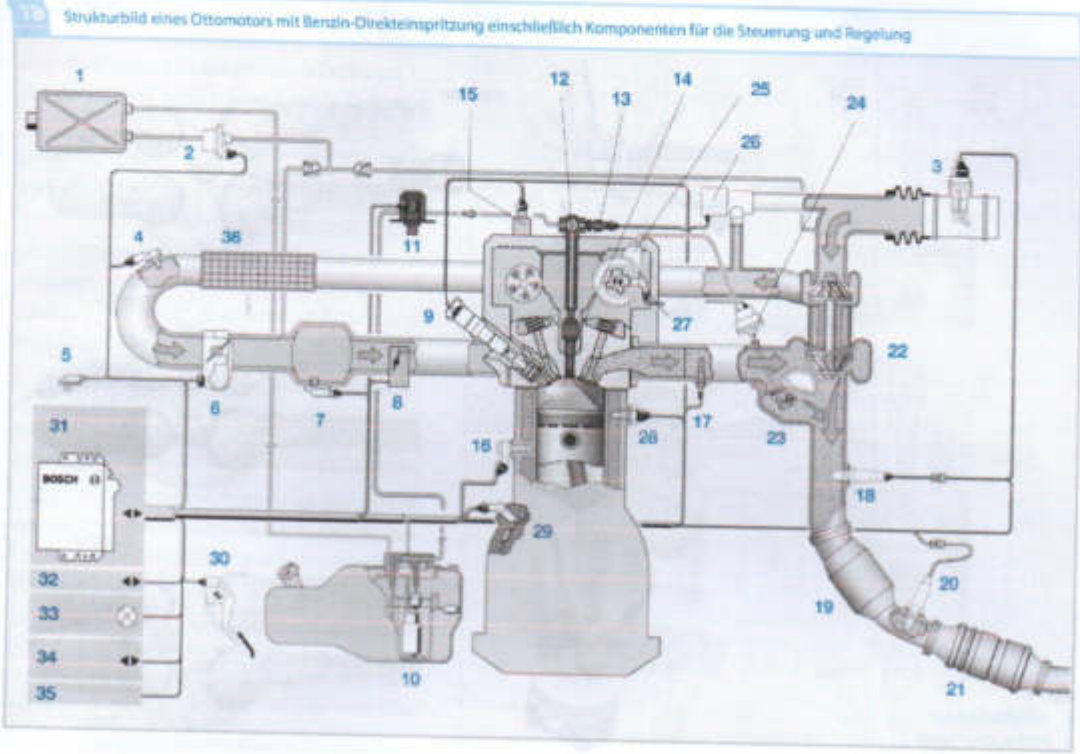
Hochdruckpumpe integrierte Mengensteu-
erventil auf Werte zwischen 50 und 200 bar
eingestellt. Am Rail, auch als „Common
Rail" bezeichnet, sind die Hochdruck-Ein-
spritzventile (14) angeordnet. Sie werden
vom Motorsteuergerät angesteuert und sprit-
zen den Kraftstoff in den Brennraum des
Zylinders ein. Die Komponenten der Bosch-
Benzin-Direkteinspritzung sind aus Edel-
stahl gefertigt und somit robust im Einsatz
mit unterschiedlichen Kraftstoffen. Die Me-
dienverträglichkeit besteht für alle gängigen
Kraftstoffe, E85 (85 % Ethanol und 15 %
Benzin) und M15 (15 % Methanol und 85 %
Benzin). Weitere Kraftstoffe können in Ab-
stimmung mit dem Fahrzeughersteller frei-
gegeben werden.

Bild 18

1 Aktivkohlebehälter
2 Tankentlüftungsventil
3 Heißfilm-Luftmassenmesser
4 kombinierter Ladedruck- und Ansaug-
   lufttemperatursensor
5 Umgebungsdrucksensor
6 Drosselvorrichtung (EGAS)
7 Saugrohrdrucksensor
8 Ladungsbewegungsklappe
9 Zündspule mit Zündkerze
10 Kraftstofffördermodul mit Elektro-
   kraftstoffpumpe
11 Hochdruckpumpe
12 Kraftstoff-Verteilerrohr
13 Hochdrucksensor
14 Hochdruck-Einspritzventil
15 Nockenwellenversteller
16 Klopfsensor
17 Abgastemperatursensor
18 λ-Sonde

19 Vorkatalysator
20 λ-Sonde
21 Hauptkatalysator
22 Abgasturbolader
23 Waste-Gate
24 Waste-Gate-Steller
25 Vakuumpumpe
26 Schub-Umluftventil
27 Nockenwellenphasensensor
28 Motortemperatursensor
29 Drehzahlsensor
30 Fahrpedalmodul
31 Motorsteuergerät
32 CAN-Schnittstelle
33 Motorkontrollleuchte
34 Diagnoseschnittstelle
35 Schnittstelle zur Wegfahrsperre
36 Ladeluftkühler

[handwritten notes, bottom:]

einstellen: regular, ajustar, colocar.
zeichnen: indicar, denominar.
ordnen: disponer, colocar.
Edelstahl: acero inoxidable

verträglichkeit: compatibilidad   freigeben: homologar.
verträglich: compatible
bestehen: soporta
gängig: corriente, común
Abstimmung: acuerdo, sintonía

224

Handwritten notes (top margin):

Weise: modo, forma
freisetzen: liberación
Hierbei: al mismo tiempo

Strömung: flujo
Ausrichtung: orientación
Größe: magnitud.
seitlich: de lado

rasch: rápido, veloz, pronto.
vollständig: completo, total.
bestehend: existente, vigente.
erleichtern: facilitar, aliviar, -

üblich: usual, habitual, convencional.
etablieren: establecer. unterhalb: debajo   Anpassung: adaptación
umstieg: paso, cambio   deutlich: claramente
durchsetzen: llevar adelante, conseguir

Bild 19
Einbaulagen des Injektors, seitlich und zentral

a
Injektor
Injektor

b
Injektor
Injektor

c
Injektor
Injektor

Bild 19
a   seitlicher Einbau
b   zentral longitudinal
c   zentral transversal

Brennverfahren und Betriebsarten

Brennverfahren

Als Brennverfahren wird die Art und Weise bezeichnet, wie das Gemisch im Brennraum gebildet und die Energie durch die Verbrennung freigesetzt wird. Hierbei werden die Abläufe durch viele Parameter beeinflusst. Wesentliche Parameter sind die Geometrie des Brennraums, die Brennraumströmung und die Ausrichtung des Kraftstoffsprays, aber auch die steuerbaren Größen wie der Einspritz- und der Zündungszeitpunkt. Die Optimierung all dieser Parameter ist die Grundvoraussetzung für ein robust ablaufendes Brennverfahren mit rascher und vollständiger Verbrennung und geringen Emissionen.

Die Kraftstoffverteilung im Brennraum wird stark durch die Einbaulage des Einspritzventils beeinflusst. Heute haben sich bei den üblichen Vierventilmotoren die beiden Einbaulagen seitlich und zentral etabliert. Bei seitlicher Einbaulage wird der Injektor unterhalb des Einlasskanals positioniert (Bild 19a) . Der Kraftstoff wird zwischen den Einlassventilen in den Brennraum eingespritzt. Ein wesentlicher Vorteil dieser Einbaulage ist die relativ einfache Anpassung eines Zylinderkopfes einer bestehenden Saugrohreinspritzung, was den Umstieg auf die Direkteinspritzung für die Motorenhersteller deutlich erleichtert.

Bei zentraler Einbaulage haben sich in Serienmotoren zwei Positionierungsmöglichkeiten durchgesetzt, der longitudinale und der transversale Einbau (Bild 19b, c). Beim longitudinalen Einbau liegen Zündkerze und Injektor im Zylinderdach zwischen den Ein-

225

vergleichweise: comparativamente
Ablagerung: deposito, sedimentación
dadurch: de este modo.
erlauben: permitir, dejar
hierzu: respecto a esto, a tal efecto.
vorhanden: disponible, existente
vorlegen: plantear, exhibir

und Auslassventilen. Dadurch kann eine
bessere Zylinderkopfkühlung erreicht wer-
den. Es bleibt auch ein größerer Freiraum
für die Einlass- und Auslasskanäle. Bei der
transversalen Einbaulage liegt der Injektor
zwischen den Einlassventilen, die Zündkerze
zwischen den Auslassventilen. Bei dieser
Positionierung bleibt die Injektorspitze ver-
gleichsweise kühl. Die Robustheit gegen
Ablagerungen an der Injektorspitze wird
dadurch verbessert. Die zentrale Einbaulage
erlaubt zudem, das volle Potential der Ver-
brauchsreduzierung durch Kraftstoffschich-
tung zu nutzen. Heutige Schichtbrennver-
fahren verwenden hierzu die transversale
Einbaulage.

Ein Brennverfahren besteht oft aus mehre-
ren verschiedenen Betriebsarten, auf die be-
triebspunktabhängig umgeschaltet wird.
Prinzipiell teilen sich die Brennverfahren in
zwei Klassen auf: in Homogen- und Schicht-
brennverfahren.

verfahren: operación, modo de obrar.

Homogenbrennverfahren

Beim Homogenbrennverfahren wird in der
Regel im gesamten Motorkennfeld ein im
Mittel stöchiometrisches Gemisch im Brenn-
raum gebildet (Bild 20). Das bedeutet, dass
immer eine Luftzahl von $\lambda = 1$ vorliegt. Da-
mit wird wie bei der Saugrohreinspritzung
die Abgasnachbehandlung durch einen Drei-
Wege-Katalysator ermöglicht. Dieses Brenn-
verfahren wird in Verbindung mit einer
Aufladung häufig beim Downsizing (Redu-
zierung des Hubraums bei gleichzeitiger Ef-
fizienzsteigerung) angewandt, um den Kraft-
stoffverbrauch zu senken.

Das Homogenbrennverfahren wird immer
im Homogenmodus betrieben, allerdings
kann es auch hier Sonderbetriebsarten ge-
ben, die motorindividuell unterschiedlich
zu bestimmten Einsatzzwecken genutzt
werden.

vorlegen: plantear.
ermöglichen: hacer posible
angewandt: aplicado.
senken: reducir, disminuir

20 Gemischverteilung im Brennraum
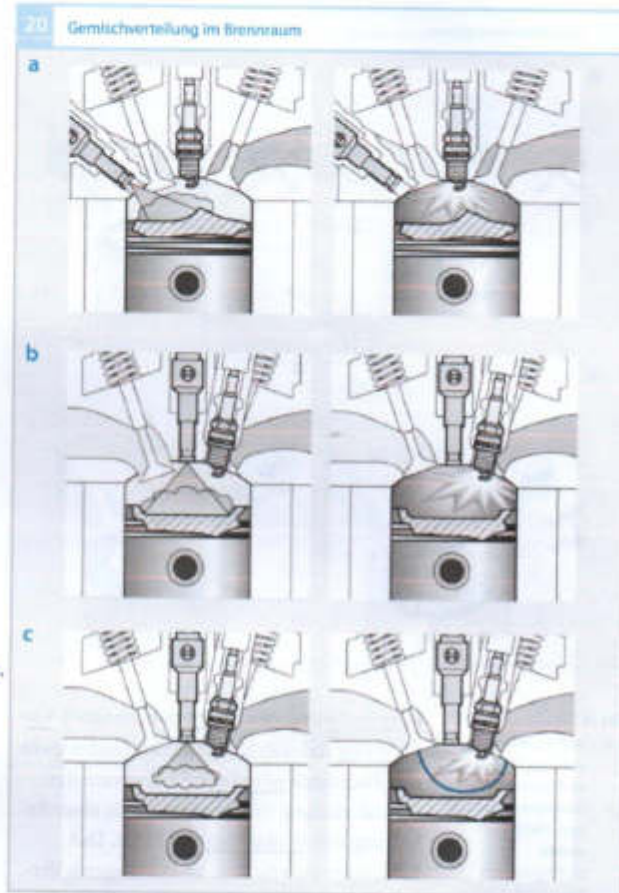
Schichtbrennverfahren

Beim Schichtbrennverfahren wird in einem
bestimmten Kennfeldbereich (kleine Last,
kleine Drehzahl) der Kraftstoff erst im Ver-
dichtungstakt in den Brennraum eingespritzt
und ggf. als Schichtwolke zur Zündkerze
transportiert (Bild 20 c). Die Wolke ist dabei
im idealen Fall von reiner Frischluft umge-
ben. Somit ist nur in der lokalen Wolke ein
zündfähiges Gemisch vorhanden. Gemittelt
über den gesamten Brennraum liegt eine
Luftzahl $\lambda > 1$ vor. Dadurch kann in größe-
ren Bereichen ungedrosselt gefahren werden,
was aufgrund der reduzierten Ladungswech-

Verdichtungstakt: carrera de compresión.

Bild 20
a seitliche Einbaulage
des Einspritzventils:
homogene Gemisch-
bildung und Ver-
brennung
b zentrale Einbaulage
des Einspritzventils:
homogene Gemisch-
bildung und Ver-
brennung
c zentrale Einbaulage
des Einspritzventils:
geschichtete Ge-
mischbildung und
Verbrennung, die
blaue Linie markiert
die Gemischwolke

verluste: pérdida
verdünnung: dilución
wegen: a consecuencia de, debido a, por.
günstig: oportuno, ventajoso

Wirkungsgrad: eficiencia
einsetzen: poner, colocar, establecer.
erfolgen: ocurre, suede
mulde: cuenca.
entweder: ya sea
nah: cercano, próximo.

luftpolster:
colchón de aire.
vereinen: unir, combinar
meist: la mayoría del tiempo.
winkel: ángulo
ca.: aprox (circa)
umsetzen: implementar

Schichtbrennverfahren

a

b

c

d
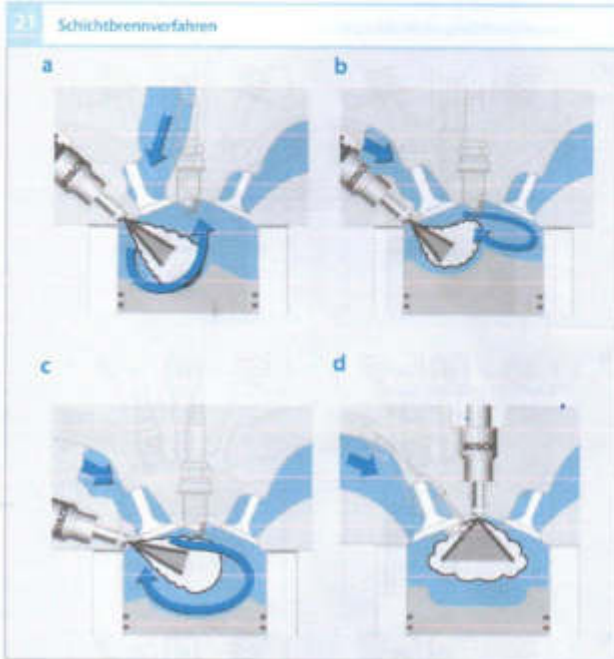
Bild 21
a–c wand- und luft-
     geführte Brenn-
     verfahren
a, b Gemischtransport
     über die Kolben-
     mulde
d   strahlgeführtes
     Brennverfahren

selverluste und der wegen der erhöhten Verdünnung reduzierten mittleren Gastemperatur, und damit günstigen Stoffwerten der Zylinderladung im Brennraum, zu einer Erhöhung des Wirkungsgrads führt. Das Schichtbrennverfahren ist ein mageres Verbrauchskonzept mit hohen Potentialen für den Ottomotor.

Heute wird in Neufahrzeugen aufgrund der hohen Kosten für das Abgassystem nur noch das Schichtkonzept mit dem größten Verbrauchspotential, das strahlgeführte Brennverfahren, eingesetzt.

Wand- und luftgeführtes Brennverfahren
Beim wand- und luftgeführten Brennverfahren sitzt der Injektor in seitlicher Einbaulage (Bild 21a–c). Der Gemischtransport erfolgt über die Kolbenmulde, die (im Falle der Wandführung) entweder direkt mit dem Kraftstoff interagiert oder die Luftströmung

Strahlgeführte
     Einspritzung
hat das größte
Verbrauchspotential

Anordnung: disposición
Umweg: desvío
Nachteilig: desfavorable
     desventajoso
Aufbereitung: preparación
Verfügung: disposición
benötigen: necesitar
Lamellenzerfall: desintegración en laminillas.
güte: calidad

im Brennraum so führt, dass (im Falle der Luftführung) der Kraftstoff auf einem Luftpolster zur Zündkerze geleitet wird. Reale geschichtete Brennverfahren mit seitlichem Injektoreinbau vereinen meist beides, abhängig vom Einbauwinkel der Injektoren, der eingespritzten Kraftstoffmenge und der Ladungsbewegung im Brennraum. Wand- und luftgeführte Schichtbrennverfahren werden seit ca. 2005 aus Kosten-Nutzen-Gründen in Serienmotoren nicht mehr umgesetzt.

Strahlgeführtes Brennverfahren
Das strahlgeführte Brennverfahren verwendet die zentrale Einbaulage. Die Zündkerze ist injektornah im Brennraumdach eingebaut (Bild 21d). Der Vorteil dieser Anordnung ist die Möglichkeit der direkten Strahlführung des Kraftstoffstrahls zur Zündkerze ohne Umwege über Kolben oder Luftströmungen. Nachteilig ist allerdings die kurze Zeit, die zur Gemischaufbereitung zur Verfügung steht. Strahlgeführte Schichtbrennverfahren benötigen daher einen Kraftstoffdruck von ca. 200 bar und eine hohe Gemischgüte. Dies wird beim Injektor für strahlgeführte Brennverfahren durch eine außenöffnende Düse mit Lamellenzerfall erreicht.

Das strahlgeführte Brennverfahren erfordert eine exakte Positionierung von Zündkerze und Einspritzventil sowie eine präzise Strahlausrichtung, um das Gemisch zum richtigen Zeitpunkt entzünden zu können. Die Wärmewechselbelastung der Zündkerze ist dabei sehr hoch, da die heiße Zündkerze unter Umständen vom relativ kalten Einspritzstrahl direkt benetzt wird. Bei guter Auslegung des Systems weist das strahlgeführte Brennverfahren einen höheren Wirkungsgrad auf als die anderen geschichteten Brennverfahren, sodass hier gegenüber dem Schichtbetrieb mit wand- und luftgeführten Brennverfahren eine noch höhere Verbrauchsersparnis erreicht werden kann.

Wärmewechselbelastung: carga de intercambio de calor
Umstand: circunstancia
da: en ese caso, entonces
benetzen: mojar

dabei : ∼ constante
Auslegung : exposición
aufweisen : exhibir, mostrar, ofrecer
sodass : de modo que
ersparnis : ahorro = Einsparung.
hub : carrera,

Außerhalb : fuera de
aufführen : mencionar, especificar, citar
Je : jamás
Je nach : dependiendo de
geeignet : apropiado, adecuado

genügend : suficiente
verbleiben : quedar, seguir
überschuss : exceso, excedente
ausnutzen : aprovechar, explotar

Außerhalb des Schichtbetriebbereichs wird auch beim Schichtbrennverfahren der Motor im Homogenmodus betrieben.

Betriebsarten

Im Folgenden sollen die unterschiedlichen Betriebsarten, die bei der Benzin-Direkteinspritzung eingesetzt werden, aufgeführt werden. Je nach Betriebspunkt des Motors wird die geeignete Betriebsart von der Motorsteuerung eingestellt (Bild 22).

Homogen

Im Homogenmodus wird die eingespritzte Kraftstoffmenge genau im stöchiometrischen Verhältnis ($\lambda = 1$), z. B. bei Super-Benzin 14,7:1, der Frischluft zugemessen. Dabei wird der Kraftstoff im Ansaughub eingespritzt, damit genügend Zeit verbleibt, um das gesamte Gemisch zu homogenisieren. Zum Bauteilschutz des Katalysators oder zur Leistungssteigerung an der Volllast wird in Teilen des Betriebskennfelds auch mit leichtem Kraftstoffüberschuss gefahren ($\lambda < 1$). Die Betriebsart „Homogen" ist bei einer hohen Drehmomentanforderung notwendig, da sie den gesamten Brennraum ausnutzt. Wegen des stöchiometrisch vorliegenden Luft-Kraftstoff-Gemischs ist in dieser Betriebsart auch die Rohemission von Schadstoffen niedrig, die zudem vom Drei-Wege-Katalysator vollständig konvertiert werden kann. Beim Homogenbetrieb entspricht die Verbrennung weitgehend der Verbrennung bei der Saugrohreinspritzung.

Schichtbetrieb

Beim Schichtbetrieb wird der Kraftstoff erst im Verdichtungstakt eingespritzt. Der Kraftstoff soll dabei nur mit einem Teil der Luft aufbereitet werden. Es entsteht eine Schichtwolke, die idealerweise von reiner Frischluft umgeben ist. Das Einspritzende ist im Schichtbetrieb sehr wichtig. Die Schicht-

22 Einspritz- und Zündzeitpunkte der verschiedenen Betriebsarten
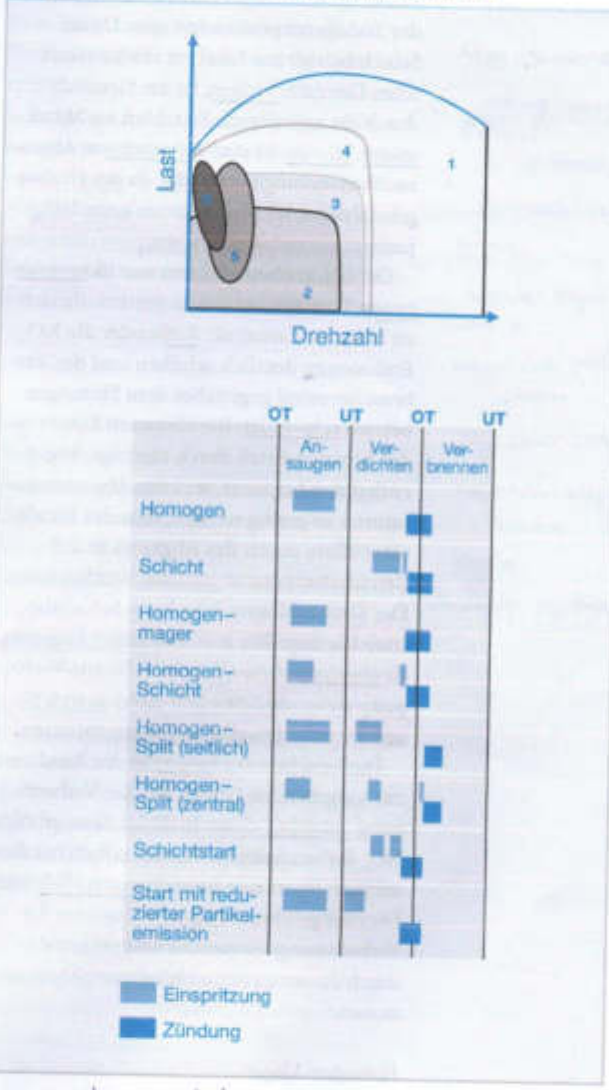
vorliegend : en cuestión
entsprechen : corresponder.
weitgehend : en gran parte
entstehen : formarse, surgir.
umgeben : rodear.

Bild 22
1 Homogen
2 Schichtbetrieb
3 Homogen-Mager
4 Homogen-Schicht
5 Homogen-Split (zum Katalysator-Heizen)
6 Schichtstart und Start mit reduzierter Partikelemission

228

Handwritten margin notes (top):

ausreichend: suficiente, bastante
vorliegen = exhibir, presentar.
umhüllend: envuelto, revestido.

weiger: pobre, lean,
hierbei . con esto
al mismo tiempo

Handwritten margin notes (left/upper):

aufwendig: costoso, gastoso.
vorgegeben: dado de antemano
Ruß: hollin
schaffen: constituir, producir

schwinden: disminuir, decrecer.
oberhalb: por encima de
abhagern: empobrecerse
Abhilfe: remedio, solución

schwelle: umbral
ausreichen: er suficiente, bastar
entstehen: formarse, originarse

stehend: estable, fijo

Handwritten margin notes (left column):

bgasrückführung:
circulación de gases.
wiederum: por otra parte
en cambio.
ähnlich: parecido, similar
im übrigen: el resto, lo demús
Aufteilung: distribución, reparto.
tragen: valer
magern: disminuir, reducir, rebajar
Anwendung: aplicación

---

wolke muss zum Zündzeitpunkt nicht nur ausreichend homogenisiert, sondern auch an der Zündkerze positioniert sein. Da im Schichtbetrieb nur lokal ein stöchiometrisches Gemisch vorliegt, ist das Gemisch durch die umhüllende Frischluft im Mittel mager. Hierbei ist eine aufwendigere Abgasnachbehandlung notwendig, da der Dreiwegekatalysator im Magerbetrieb keine NO_x-Emissionen reduzieren kann.

Der Schichtbetrieb kann nur in vorgegebenen Grenzen betrieben werden, da sich zu höheren Lasten die Ruß- oder die NO_x-Emissionen deutlich erhöhen und der Verbrauchsvorteil gegenüber dem Homogenbetrieb schwindet. Bei kleineren Lasten ist der Schichtbetrieb durch niedrige Abgasenthalpien begrenzt, weil die Abgastemperaturen so gering werden, dass der Katalysator allein durch das Abgas nicht auf Betriebstemperatur gehalten werden kann. Der Drehzahlbereich ist beim Schichtbetrieb bis ungefähr $n = 3500$ min^{-1} begrenzt, da oberhalb dieser Schwelle die zur Verfügung stehende Zeit nicht mehr ausreicht, um die Schichtwolke zu homogenisieren.

Die Schichtwolke magert in der Randzone zur umgebenden Luft ab. Bei der Verbrennung entstehen daher in dieser Zone erhöhte NO_x-Rohemissionen. Abhilfe schafft bei dieser Betriebsart eine hohe Abgasrückführrate. Die rückgeführten Abgase reduzieren die Verbrennungstemperatur und senken dadurch die temperaturabhängigen NO_x-Emissionen.

Homogen-Mager

In einem Übergangsbereich zwischen Schicht- und Homogenbetrieb kann der Motor mit Schichtbrennverfahren mit homogenem mageren Gemisch betrieben werden ($\lambda > 1$). Im Homogen-Mager-Betrieb ist der Kraftstoffverbrauch gegenüber dem Homogenbetrieb mit $\lambda = 1$ geringer, da die La-

dungswechselverluste durch die Entdrosselung geringer werden. Zu beachten sind aber die erhöhten NO_x-Emissionen, da der Dreiwegekatalysator in diesem Bereich diese Emissionen nicht reduzieren kann. Zusätzliche NO_x-Emissionen bedeuten wiederum Wirkungsgradverluste durch die Regenerierungsphasen eines hier notwendigen NO_x-Speicherkatalysators.

Homogen-Schicht

Im Homogen-Schicht-Betrieb ist der gesamte Brennraum mit einem homogen-mageren Grundgemisch gefüllt. Dieses Gemisch entsteht durch Einspritzung einer Grundmenge an Kraftstoff in den Ansaugtakt. Eine zweite Einspritzung erfolgt im Kompressionstakt. Dadurch entsteht eine fettere Zone im Bereich der Zündkerze. Diese Schichtladung ist leichter entflammbar und kann mit der Flamme – ähnlich einer Fackelzündung – das homogen-magere Gemisch im übrigen Brennraum sicher entzünden.

Der Aufteilungsfaktor zwischen den beiden Einspritzungen beträgt ungefähr 75 %. Das bedeutet, 75 % des Kraftstoffs werden bei der ersten Einspritzung, die für das homogene Grundgemisch sorgt, eingespritzt. Ein stationärer Homogen-Schicht-Betrieb bei niedrigen Drehzahlen im Übergangsbereich zwischen Schicht- und Homogenbetrieb reduziert die Rußemission gegenüber dem Schichtbetrieb und verringert den Kraftstoffverbrauch gegenüber dem Homogenbetrieb.

Homogen-Split

Der Homogen-Split-Modus ist eine spezielle Anwendung der Homogen-Schicht-Doppeleinspritzung. Er wird bei allen Motoren mit Benzindirekteinspritzung zum raschen Aufheizen des Katalysators nach dem Kaltstart eingesetzt. Durch die stabilisierend wirkende zweite Einspritzung im frühen Kompressi-

229

Handwritten notes (left column):

Anteil: porción, parte
eingehen: acabar, recibirse
einsatzbereit: disponible para trabajar
anstatt: en vez de
Vordruck: depresión
beruhen: basarse, deberse
verdunsten: evaporarse,

Handwritten notes (middle top):

sonst: caso contrario, otra vez de lo contrario
wirken: actuar, hacer efecto.
bewirken: producir, conseguir causar
wodurch: por lo cual
vermeiden: evitar, eludir
angewandt: aplicado.

Handwritten notes (right column):

Bereitstellung: Preparación
Möglichst: si es posible
Anforderung: requisito, requerimiento, exigencias
hingegen: por el contrario
volumetrico: raumlich - espacial
teilweise: parcialmente
vor allem: sobre todo
mucho que nada.

...onstakt bei seitlicher Einbaulage oder direkt vor der Zündung bei zentraler Einbaulage kann die Zündung extrem spät (bei einem Kurbelwinkel von 15 ... 30° nach ZOT) erfolgen. Ein großer Anteil der Verbrennungsenergie wird dann nicht mehr in eine Drehmomentensteigerung eingehen, sondern erhöht die Abgasenthalpie. Durch diesen hohen Abgaswärmestrom ist der Katalysator schon wenige Sekunden nach dem Start einsatzbereit.

Schichtstart

Beim Schichtstart wird die Starteinspritzmenge im Kompressionshub und unter erhöhtem Kraftstoffdruck eingespritzt, anstatt konventionell im Ansaughub bei Vordruck eingespritzt zu werden. Der Vorteil dieser Einspritzstrategie beruht darauf, dass in bereits komprimierte und damit erwärmte Luft eingespritzt wird. Dadurch verdunstet prozentual mehr Kraftstoff als bei kalten Umgebungsbedingungen, bei denen sonst ein deutlich größerer Anteil des eingespritzten Kraftstoffs als flüssiger Wandfilm im Brennraum verbleibt und nicht an der Verbrennung teilnimmt. Die einzuspritzende Kraftstoffmenge kann daher beim Schicht-Start deutlich verringert werden. Dies führt zu stark reduzierten HC-Emissionen beim Start. Da zum Startzeitpunkt der Katalysator noch nicht wirken kann, ist dies eine wichtige Betriebsart für die Entwicklung von Niedrigemissionskonzepten. Zusätzlich bewirkt diese Schichteinspritzung eine deutlich stabilere Startverbrennung, was wiederum die Startrobustheit erhöht. Um eine Aufbereitung in der kurzen, zur Verfügung stehenden Zeit zu ermöglichen, wird der Schichtstart mit einem Kraftstoffdruck von ca. 50 bar durchgeführt. Dieser Druck kann von der Hochdruckpumpe bereits durch die Umdrehungen des Starters zur Verfügung gestellt werden.

Umdrehung: revolución, giro

Start mit reduzierter Partikelemission

Aufgrund der erhöhten Anforderungen der EU6-Emissionsgesetze zur Senkung der Partikelemission werden heute im Start Einspritzstrategien mit reduzierter Partikelemission verwendet. So wird meist eine Mehrfacheinspritzung mit einer Ersteinspritzung in der Saugphase angewandt. Ein zweiter Anteil wird in die frühe Kompressionsphase einspritzt, wodurch sehr inhomogene Schichtwolken vermieden werden. Partikel werden nur in lokalen Gemischbereichen erzeugt, in denen eine Luftzahl $\lambda < 0{,}5$ besteht.

Gemischbildung, Zündung und Entflammung

Aufgabe der Gemischbildung ist die Bereitstellung eines möglichst homogenen, brennfähigen Luft-Kraftstoff-Gemischs zum Zeitpunkt der Zündung.

Anforderungen

In der Betriebsart Homogen (Homogen mit $\lambda \le 1$ und auch Homogen-Mager mit $\lambda > 1$) soll das Gemisch im gesamten Brennraum homogen sein. Im Schichtbetrieb hingegen ist das Gemisch nur innerhalb eines räumlich begrenzten Bereichs teilweise homogen, während sich im restlichen Brennraum Frischluft oder Inertgas befindet. Homogen kann eine Gas-Mischung oder eine Gas-Kraftstoffdampf-Mischung nur dann sein, wenn der gesamte Kraftstoff verdunstet ist. Einfluss auf die Verdunstung haben viele Faktoren, vor allem

- die Temperatur im Brennraum,
- die Brennraumströmung,
- die Tropfengröße des Kraftstoffs,
- die Zeit, die zur Verdunstung zur Verfügung steht.

230

Handwritten notes (margins):

maßgeblich: decisivamente
tief: hondo, bajo, profundo.
erhalten: alcanzar, conseguir
detener, lograr.
erzielen: obtener, realizar
conseguir.

d.h. = Das heißt, quiere decir   eindringtiefe: profundidad de penetración

länge: longitud   zurücklegen: recorrer.   einzeln: particular, individualmente
Abstand: distancia, diferencia, intervalo.   bezüglich: con relación a

mithilfe: con ayuda, con apoyo

einströmend:
      [wir hacia dentro

deren: cuya, cuyo

um einen: por un
          lado.

m anderen:
      por otro lado.

fördern: promover, fomentar
      favorecer

rand: deflagración,
      quemar.

weilgehend: en gran
              parte

sbildung: formación
          preparación

tscheidend: determinante

visteken: originarse
          resultar

agen: conducir

durfach: multiple

Einflussgrößen

Brennfähig ist ein Gemisch mit Ottokraftstoff mit λ im Bereich von 0,6 bis 1,6; abhängig von Temperatur, Druck und Brennraumgeometrie des Motors.

Temperatureinfluss

Die Temperatur beeinflusst maßgeblich die Verdunstung des Kraftstoffs. Bei tieferen Temperaturen verdunstet er nicht vollständig. Deshalb muss unter diesen Bedingungen mehr Kraftstoff eingespritzt werden, um ein brennfähiges Gemisch zu erhalten.

Druckeinfluss

Die Tropfengröße des eingespritzten Kraftstoffs ist abhängig vom Einspritzdruck und vom Druck im Brennraum. Mit steigendem Einspritzdruck können kleinere Tropfengrößen erzielt werden, die schneller verdunsten.

Geometrieeinfluss

Bei gleichem Brennraumdruck und steigendem Einspritzdruck erhöht sich die Eindringtiefe, d. h., die Weglänge, die der einzelne Tropfen zurücklegt, bis er vollständig verdunstet ist. Ist dieser zurückgelegte Weg länger als der Abstand vom Einspritzventil zur Brennraumwand, wird die Zylinderwand oder der Kolben benetzt. Verdunstet der so entstehende Wandfilm nicht rechtzeitig bis zur Zündung, nimmt er nicht oder nur unvollständig an der Verbrennung teil und erzeugt HC- und Partikelemissionen. Wandfilme sind bei homogenen Brennverfahren die Hauptquelle der Partikelemissionen. Die Geometrie des Motors (bezüglich Einlasskanal und Brennraum) ist auch verantwortlich für die Luftströmung und die Turbulenz im Brennraum, die wesentliche Faktoren für den Einfluss auf die Brenngeschwindigkeit sind.

Gemischbildung und Verbrennung im Homogenbetrieb

Um eine lange Zeit für die Gemischbildung zu erhalten, sollte der Kraftstoff frühzeitig eingespritzt werden. Deshalb wird im Homogenbetrieb bereits im Ansaugtakt eingespritzt und mithilfe der einströmenden Luft eine schnelle Verdunstung des Kraftstoffs und eine gute Homogenisierung des Gemischs erreicht (Bild 23a). Die Aufbereitung wird vor allem durch hohe Strömungsgeschwindigkeiten und deren aerodynamische Kräfte im Bereich des öffnenden und schließenden Einlassventils unterstützt. Bei aufgeladenen Motoren wird eine starke Tumbleströmung verwendet, die zum einen das fein aufbereitete Kraftstoffspray von der Wand fernhält, und zum anderen durch die starke Durchmischung des Kraftstoffgemisches die Verdunstung und Homogenisierung fördert. Zusätzlich erzeugt zum Zeitpunkt der Entflammung der Zerfall der Tumbleströmung in Turbulenz einen raschen Durchbrand. Die Zündungs- und Entflammungsbedingungen homogener Gemische bei der Benzin-Direkteinspritzung entsprechen weitgehend denen bei der Saugrohreinspritzung.

Gemischbildung und Verbrennung im Schichtbetrieb

Für den Schichtbetrieb ist die Ausbildung der brennfähigen Gemischwolke, die sich zum Zündzeitpunkt im Bereich der Zündkerze befindet, entscheidend. Dazu wird beim strahlgeführten Brennverfahren der Kraftstoff während der Verdichtungsphase so eingespritzt, dass eine kompakte Gemischwolke entsteht (Bild 23b). Diese wird durch den Sprayimpuls zur Zündkerze getragen. Der Einspritzzeitpunkt ist von der Drehzahl und vom geforderten Drehmoment abhängig. Bei höheren Lasten im Schichtbetrieb wird auch eine Mehrfacheinspritzung zur Homogenisierung der Ge-

Kraft: fuerza, energía.
zerfall: desintegración, descomposición, disociación
Abschluss: terminación, finalización
Erst: tan sólo, solamente
Danach: conforme a ello
abmagern: empobrecer (mecánica)

Funke: chispa.
Kern: núcleo, centro.
ausbreiten: exparer, difundir, divulgar.
abbrennen: quemar, deflagrar
Damit: por eso.   somit: por consiguiente, por tanto.
fest: firmemente   binden: atar, ligar

mischwolke eingesetzt. Die dadurch in die Gemischwolke zusätzlich eingetragene Luft ermöglicht auch eine Anpassung der Luftzahl im Gemisch auf stöchiometrische Verhältnisse.

Für eine robuste Entflammung ist das exakte Zusammenspiel zwischen Einspritzende und Zündung wichtig. Während der Einspritzung des Gemischs ist die Strömungsgeschwindigkeit der an der Zündkerze vorbeifliegenden Gemischwolke, aber auch die Kühlung des verdunstenden Kraftstoffes zu hoch für eine Entflammung (Bild 24a). Erst zum Abschluss der Einspritzung bestehen für eine sehr kurze Zeit ideale Bedingungen. In der danach folgenden Schleppe aus Brennraumluft magert das Gemisch rasch ab. In diese Schicht am Ende der Einspritzung wird der Zündfunke eingesaugt und bildet einen Flammkern aus. Dieser folgt der sich ausbreitenden Gemischwolke und brennt sie rasch ab. Damit ist der Zeitpunkt des Verbrennungsbeginns, und somit auch die Schwerpunktlage der Verbrennung, fest an das Spritzende gebunden. Der ausgebildete Zündfunke steht dagegen wesentlich länger zur Verfügung. Dieser Mechanismus der Entflammung unterscheidet sich deutlich von dem der homogenen Verbrennung, und muss auch im Motormanagement bei der Regelung der Einspritzparameter berücksichtigt werden.

Entscheidend für eine sichere Zündung und Entflammung sind unter anderem:
- die Qualität der Gemischaufbereitung,
- eine genaue Mengendosierung auch bei kleinen Einspritzmengen (Mehrfacheinspritzung),
- eine möglichst große Zündfunkenbrenndauer,
- die richtige Zuordnung von Funkenort und Kraftstoffspray,
- eine relativ genaue Einhaltung des Abstandes vom Spray zum Zündort,

23 Gemischbildungsmechanismen bei der Benzin-Direkteinspritzung

a

Tropfeninteraktion mit der Wand
Primärtropfen
Wandfilm
Tropfenverdunstung
Tropfenzerfall durch aerodynamische Kräfte
Filmverdunstung

b

Primärtropfen
Tropfenzerfall durch aerodynamische Kräfte
Tropfeninteraktion mit der Wand
Filmverdunstung
Wandfilm
Homogenisierung der Wolke, Tropfenverdunstung

Dagegen: por el contrario
berücksichtig: considerar, tener en cuenta
entscheidend: determinante, decisivo.
brenndauer: duración de encendido.
Zuordnung: coordinación
Einhaltung: cumplimiento, observancia.

Bild 23
a Homogenbetrieb davon
b Schichtbetrieb

Unveränderlich: invariable   Unveränderlichkeit:
gegenüber: frente a   invariabilidad.

einerseits: por un lado
andererseits: por otro lado
zerstäubung: pulverización
gezielt: apuntada
erzielen: lograr, obtener, conseguir

gleichmäßig: uniformemente
erforderlich: necesario, preciso
eigenschaft: característica
neigung: inclinación
hierbei: en esta ocasión

anpassen: adaptar, acomodar, adecuar
anordnung: disposición, agrupación, distribución
vor allem: sobre todo
Neben: junto a, además de
Zeitraum: lapso, espacio de tiempo.

Bild 24 Schichtbetrieb mit strahlgeführtem Brennverfahren: Kopplung der Entflammung und Verbrennung an das Einspritzende
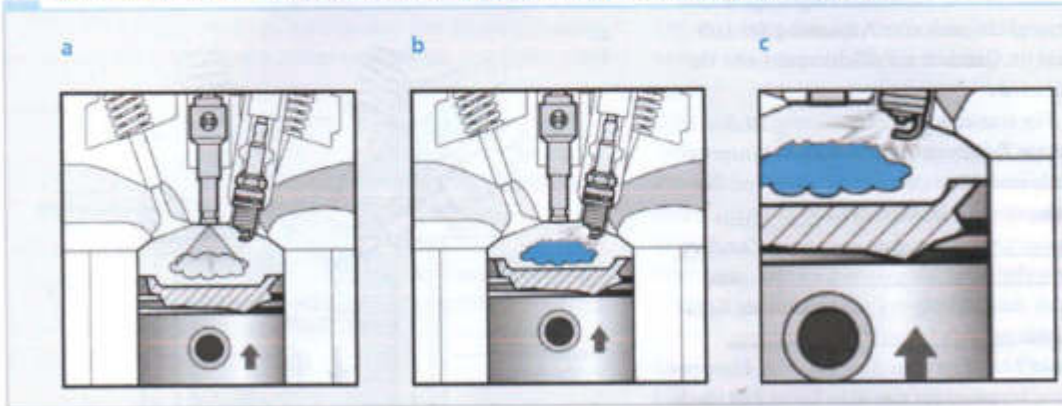
Bild 24
a Einspritzung
b Einspritzende
c vergrößerter Aus-
  schnitt aus b

- Unveränderlichkeit des Sprays gegenüber dem Brennraumdruck,
- konstante Sprayform über die gesamte Lebensdauer des Motors.

Hochdruck-Einspritzventil

Aufgabe

Aufgabe des Hochdruck-Einspritzventils (HDEV) ist es einerseits, den Kraftstoff zu dosieren und andererseits durch dessen Zerstäubung eine gezielte Durchmischung von Kraftstoff und Luft in einem bestimmten räumlichen Bereich des Brennraums zu erzielen. Abhängig vom gewünschten Betriebszustand wird der Kraftstoff im Bereich um die Zündkerze konzentriert (geschichtet) oder gleichmäßig im gesamten Brennraum zerstäubt (homogen verteilt).

Anforderungen

Spray

Für einen robusten und sauberen Verbrennungsprozess ist ein stabiles Spray erforderlich. Sprayeigenschaften, wie z. B. Spraywinkel, Sprayneigung oder Eindringtiefe sind hierbei die wesentlichen Kriterien (Bild 25). Um die Interaktion des Sprays mit der Brennraumwand oder dem Kolbenboden zu minimieren, wird die Eindringtiefe des Sprays motorspezifisch angepasst. Durch Anpassung von Kraftstoffdruck, Spritzlochanordnung und -design wird ein Optimum zwischen Zerstäubung und Eindringtiefe erzielt. Eine zusätzliche Möglichkeit zur Anpassung der Sprayausbreitung ergibt sich, indem die erforderliche Kraftstoffmenge auf mehrere Einspritzvorgänge aufgeteilt wird.

Dynamik

Neben dem Spray ist vor allem die Schaltdynamik des Hochdruck-Einspritzventiles von großer Bedeutung. Wesentlicher Unterschied der Benzin-Direkteinspritzung im Vergleich zur Saugrohreinspritzung ist ein höherer Kraftstoffdruck und eine deutlich kürzere Zeit für die Einspritzung des Kraftstoffs direkt in den Brennraum (Bild 26). Bei der Saugrohreinspritzung kann über den Zeitraum von zwei Kurbelwellenumdrehungen der Kraftstoff in das Saugrohr eingespritzt werden. Das entspricht bei einer Drehzahl von 6 000 min⁻¹ einer Einspritzdauer von 20 ms. Für den Homogenbetrieb bei der Direkteinspritzung muss der Kraftstoff im Ansaugtakt eingespritzt werden. Somit steht nur eine halbe Kurbelwellenumdre-

Erläuterung: aclaración
geneigt: inclinado / neigen: inclinar, tender
Ausbreitung: propagación
Gegebenheit: circunstancia
slank: delgado, alargado,
erforderlich: requerido, necesario, preciso.
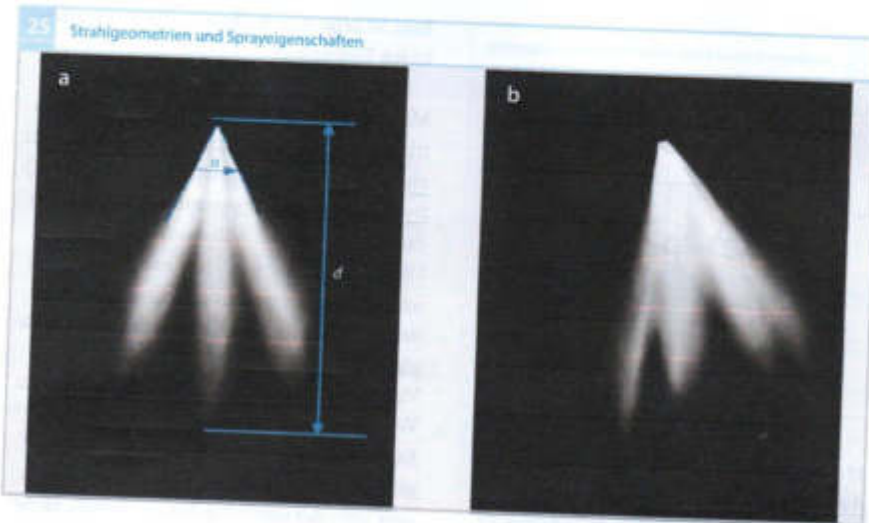
Strahlgeometrien und Sprayeigenschaften

a

b

...hung für den Einspritzvorgang zur Verfügung. Bei 6 000 min⁻¹ entspricht das einer Einspritzdauer von 5 ms. Bei der Benzin-Direkteinspritzung ist der Kraftstoffbedarf im Leerlauf (im Verhältnis zur Volllast) sehr viel geringer als bei der Saugrohreinspritzung (Faktor 1:12). Im Falle der Mehrfacheinspritzung wird die Einspritzzeit pro Teileinspritzung nochmals reduziert, was zu einer weiteren Anforderung an die Dynamik führt.

Einbau

Aus dem Brennverfahren und aus den räumlichen Gegebenheiten ergeben sich weitere, im Wesentlichen geometrische Anforderungen an das Hochdruck-Einspritzventil. Im Falle des seitlichen Einbaus (Bild 27) ist eine möglichst kleine Bauhöhe und ein schlankes Design erforderlich. Um die elektrische und hydraulische Kontaktierung realisieren zu können, wird für den zentralen Einbau (Bild 28) das Hochdruck-Einspritzventil entsprechend verlängert.
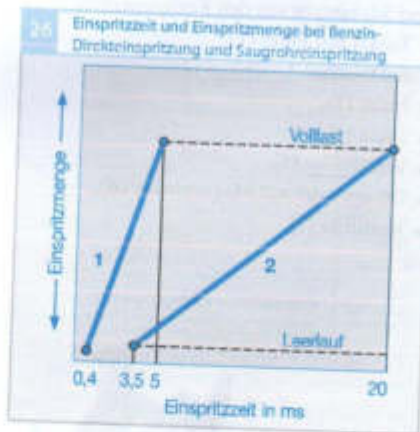
Einspritzzeit und Einspritzmenge bei Benzin-Direkteinspritzung und Saugrohreinspritzung

Einspritzmenge

Volllast

Leerlauf

0,4   3,5 5   20

Einspritzzeit in ms

1. (0,4 ms - 5 ms)
2. (3,5 ms - 20 ms)

Seitlicher Einbau eines Hochdruck-Einspritzventils

[Handwritten notes in margin:]

Aufbau: construcción
Arbeitsweise: funcionamiento, modo de trabajo
Anschluss: conexión
Hülle: róscara, cápsula, carcasa. stromdurchfloss: corriente que fluye a través
ab/heben: levantar, elevar, alzar freigeben: desbloquear drücken: empujar, presionar dabei:
dabei: como resultado de, al mismo tiempo, a la vez
Abschalten: apagar,
    desconectar, desactivar
unterbrechen: interrumpir,
    suspender, cortar.
geeignet: apropiado, adecuado
    apto.

Zentraler Einbau eines Hochdruck-Einspritzventils

Magnetinjektoren
Aufbau und *Arbeitsweise*
Das Hochdruck-Einspritzventil (Bilder 29 und 30) besteht aus den Komponenten:
- Zulauf mit Filter (1),
- elektrischer Anschluss (2),
- Feder (3),
- Spule (4),
- Ventilhülse (5),
- Düsennadel mit Magnetanker (6),
- Ventilsitz (7).
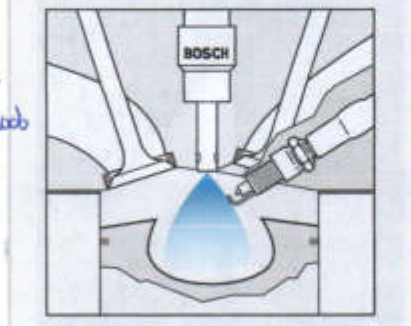
Bild 30 zeigt den Aufbau im Falle einer zentralen Einbaulage.

Bei stromdurchflossener Spule wird ein Magnetfeld erzeugt. Dadurch hebt die Ventilnadel gegen die Federkraft vom Ventilsitz ab und gibt die Ventilauslassbohrungen (8) frei. Aufgrund des Systemdrucks wird nun der Kraftstoff in den Brennraum gedrückt. Die eingespritzte Kraftstoffmenge ist dabei im Wesentlichen von der Öffnungsdauer des Ventils und dem Kraftstoffdruck abhängig. Bei Abschalten des Stroms wird die Ventilnadel aufgrund der Federkraft in den Ventilsitz gepresst und unterbricht den Kraftstofffluss. Durch eine geeignete Düsengeometrie an der Ventilspitze wird eine sehr gute Zerstäubung des Kraftstoffs erreicht.

Aufbau eines Hochdruck-Einspritzventils

Bild 29
1 Kraftstoffzulauf mit Filter
2 elektrischer Anschluss
3 Feder
4 Spule
5 Ventilhülse
6 Düsennadel mit Magnetanker
7 Ventilsitz
8 Ventilauslassbohrungen

235

Baustein: módulo, bloque
Endstufen: amplificador de potencia
Enstufenbaustein: módulo de amplificación
Wandler: convertidor, transformador

Aufbau eines Hochdruck-Einspritzventils für die zentrale Einbaulage

Ansteuerung des Einspritzventils

Um einen definierten und reproduzierbaren Einspritzvorgang zu gewährleisten, muss das Hochdruck-Einspritzventil mit einem komplexen Stromverlauf angesteuert werden (Bild 31). Der Mikrocontroller im Motorsteuergerät liefert ein digitales Ansteuersignal (a). Aus diesem Signal erzeugt ein Endstufenbaustein (ASIC) das Ansteuersignal (b) für das Einspritzventil. Ein DC/DC-Wandler im Motorsteuergerät erzeugt die Boosterspannung von 65 V. Sie wird benötigt, um den Strom in der Boosterphase möglichst rasch auf einen hohen Stromwert zu bringen. Das ist erforderlich, um die Einspritzventilnadel möglichst schnell zu beschleunigen. In der Anzugsphase (t_{an}) erreicht die Ventilnadel anschließend den maximalen Öffnungshub (c). Bei geöffnetem Einspritzventil reicht ein geringer Ansteuerstrom I_H (Haltestrom) aus, um das Ventil offen zu halten. Bei konstantem Ventilnadelhub ergibt sich eine zur Einspritzdauer proportionale Einspritzmenge (d).
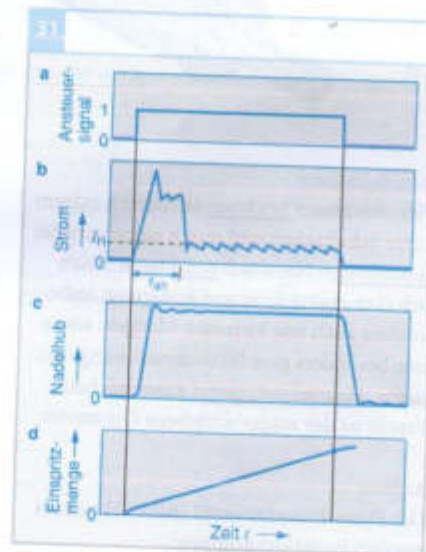
beschleunigen: acelerar, activar
Anzug: salida,
erreichen: conseguir, lograr, alcanzar.
anschließend: a continuación, seguidamente.
aus/reichen: ser suficiente, bastar.

► La cantidad completa de combustible es la integral del tiempo de activación en base a la
función de la imágen (c)

Handwritten notes (top):

aus/zeichnen: distinguirse
einstellbar: regulable, ajustable
haupteinsatzgebiet: area de aplicación principal.
erfolgen: ocurrir, suceder, realizarse
verzögerungsfrei   verzögerung: retardo, tardanza, demora.
spalt: abertura, ranura, intersticio.

vorgespannt: pretensado, pre-cargado.
Außenöffnend: abierta hacia afuera
Betätigung: acción, actuación, accionamiento.

Handwritten notes (left margin):

dünn: delgado, fino,
Schicht: capa, estrato
Weder... noch: ni... ni
erfahren: experimentar
Zugspannung: tensión
ausführen: ejecutar, desarrollar
sorgen: proporcionar, procurar, cuidar
Längenausgleich: compensación de longitud
ausgleich: compensación
ausdehnung: expansión, alargamiento
selbst: incluso, hasta
steifigkeit: rigidez
verursachen: causar, ocasionar, provocar

32 Aufbau eines Piezoinjektors

Kraftstoffzulauf
Koppler
Piezoaktormodul
Ventilgruppe
Außenöffnende Düse

Piezoinjektoren

Piezoinjektoren zeichnen sich durch extrem kurze Schaltzeiten und durch einen variabel einstellbaren Nadelhub aus. Damit lassen sich eine exakte Kraftstoffdosierung, insbesondere auch von kleinsten Mengen, sowie eine besonders gute Strahlzerstäubung realisieren. Haupteinsatzgebiet eines solchen Ventils ist der magerbetriebene Ottomotor.

Aufbau

Das Piezo-Einspritzventil (Bild 32) besteht aus drei Funktionsgruppen:

- Ventilgruppe,
- Piezo-Aktormodul,
- hydraulisches Kompensationselement.

Die Ventilgruppe besteht im Wesentlichen aus der mit einer Feder vorgespannte Ventilnadel und dem Ventilkörper. Die Nadel wird direkt über Betätigung des Piezo-Stacks bewegt. Der Öffnungs- und Schließvorgang erfolgt verzögerungsfrei. Die Nadel öffnet nach außen und gibt einen ringförmigen Spalt frei. Durch diesen tritt der Kraftstoff als dünner Film mit hoher Geschwindigkeit aus.

Das Piezo-Aktormodul ist das Stellelement. Der Piezostack besteht aus vielen piezokeramischen und elektrisch kontaktierten Schichten und ist durch eine umgebende Feder auf Druck vorgespannt. Weder im ausgelenkten noch im Ruhezustand darf der Aktor Zugspannungen erfahren.

Das Kompensationselement, auch Koppler genannt, ist als geschlossener hydraulischer Kompensator ausgeführt. Er sorgt für einen Längenausgleich zwischen Ventilgehäuse und Piezostack, der sich durch Temperatureinfluss bei unterschiedlichen Ausdehnungen einstellt. Damit ist unter allen Betriebsbedingungen, selbst in extremen Temperaturbereichen, ein konstanter Nadelhub und damit eine konstante Einspritzmenge sichergestellt. Selbst bei längeren Einspritzzeiten hat der Koppler eine ausreichende Steifigkeit, um keinen Hubverlust zu verursachen.

237

33 Eingespritzte Kraftstoffmenge als Funktion der Öffnungszeit bei festem Ventilhub für verschiedene Ansteuerparameter
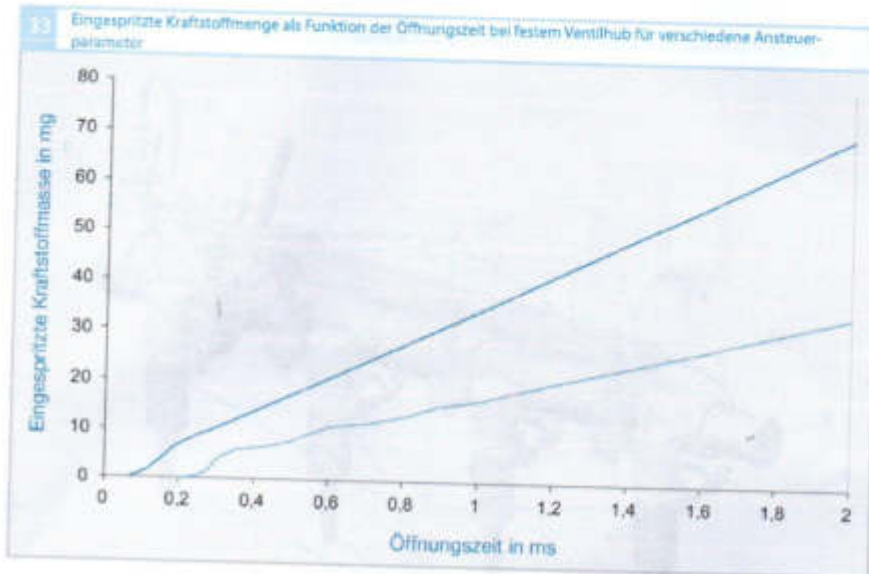
Funktion und Ansteuerung

Zur Betätigung des Piezoinjektors wird der Stack definiert elektrisch geladen. Damit öffnet das Ventil mit einer rampenförmigen Hubkurve und mit einer Schaltzeit kleiner als 0,2 ms. Umgekehrt erfolgt das Schließen des Ventils durch Entladung des Stacks. Die Schaltzeiten sind variabel. Durch die direkte Betätigung der Ventilnadel sind eine hohe Genauigkeit und Reproduzierbarkeit des Hubes von Zyklus zu Zyklus möglich, und damit eine exakte Dosierung der Einspritzmenge (Bild 33). Es lassen sich sowohl Einspritzstrategien im Teilhub- als auch im Vollhubbetrieb darstellen; auch als Kombination mit bis zu fünf Mehrfacheinspritzungen pro Arbeitstakt.

Handschriftliche Notizen:

Umgekehrt: viceversa, inversamente
Entladung: descargar.
Genauigkeit: exactitud
sowohl: ya sea.

Kraftstoffverteilerrohr für die Benzin-Direkteinspritzung aus Edelstahl
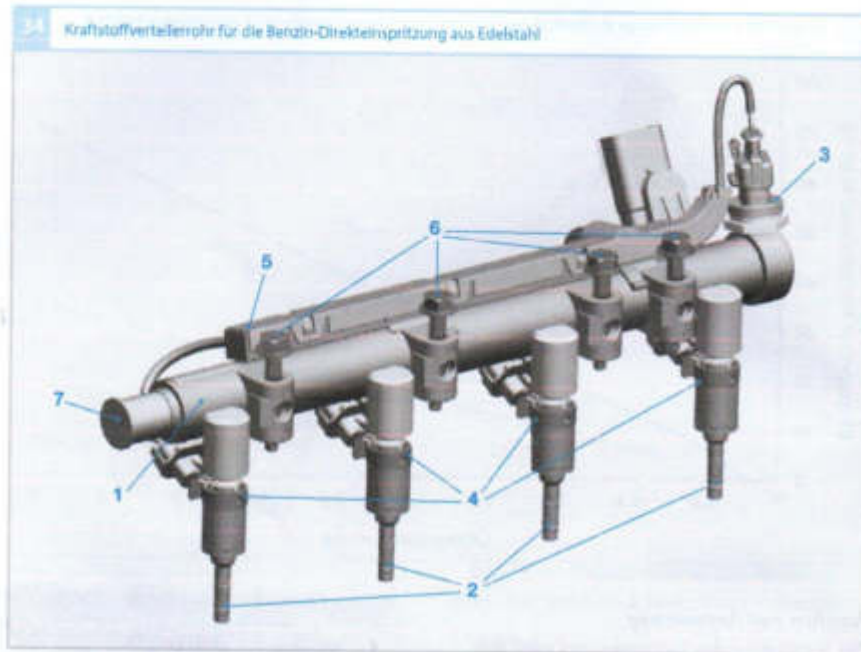
Bild 34
Systemdruck 30 MPa,
Berstdruck über 90 MPa,
Speichervolumen
50...140 cm³
1 Kraftstoffverteiler-
   rohr
2 Einspritzventil
3 Drucksensor
4 Befestigung
5 Kabelbaum
6 Schraube
7 Schutzkappe

Kraftstoffverteilerrohr

Das Kraftstoffverteilerrohr (Bild 34), auch als Rail bezeichnet, hat die Aufgabe, die für den jeweiligen Betriebspunkt erforderliche Kraftstoffmenge zu speichern und zu verteilen. Die Speicherung hängt von dem Volumen und der Kompressibilität des Kraftstoffs ab und muss für den jeweiligen Motorbedarf und Druckbereich angepasst werden. Das Volumen des Kraftstoffverteilerrohrs sorgt außerdem für eine Dämpfung im Hochdruckbereich, d. h., Druckschwankungen im Hochdruckbereich werden ausgeglichen. Am Rail sind die Anbaukomponenten für das Einspritzsystem montiert: die Hochdruckeinspritzventile (HDEV) und der Drucksensor zur Regelung des Hochdruckes.

Hochdruckpumpen für die Benzin-Direkteinspritzung

Aufgabe und Anforderungen

Die Hochdruckpumpe (HDP) hat die Aufgabe, den von der Elektrokraftstoffpumpe (EKP) mit einem Vordruck von 0,3...0,5 MPa gelieferten Kraftstoff auf das für die Hochdruckeinspritzung erforderliche Niveau von 5...20 MPa zu verdichten. Aktuelle Ausführungen sind grundsätzlich bedarfsgesteuerte Pumpen.

Aufbau und Arbeitsweise

Bild 35 zeigt eine in Öl laufende nockengetriebene Einzylinderpumpe mit integriertem niederdruckseitigen Mengensteuerventil (Zumesseinheit), hochdruckseitiger Druckbegrenzung und integriertem Druckdämpfer. Sie ist als Steckpumpe am Zylinderkopf befestigt. Der Antriebsnocken der Hochdruckpumpe sitzt auf der Motornockenwelle

239

35 Bedarfgeregelte Einzylinder-Hochdruckpumpe für Benzin-Direkteinspritzung

OUT
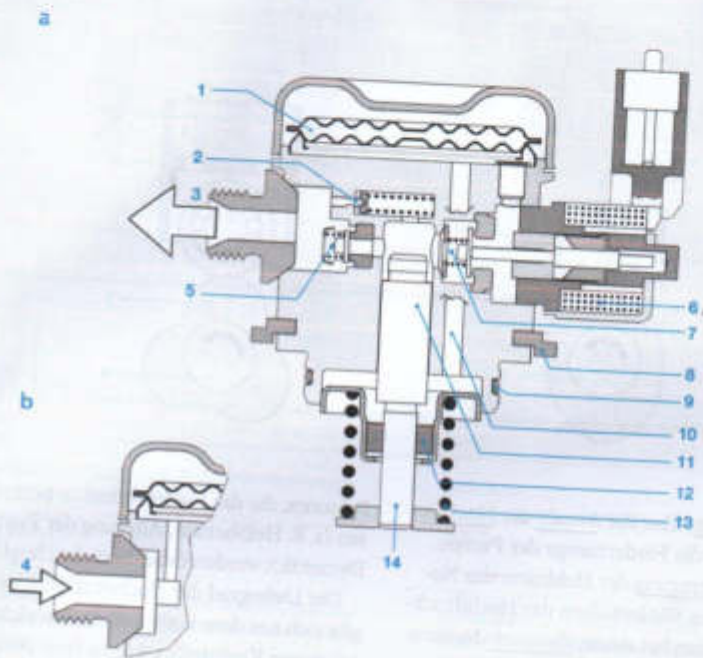
Bild 35
a Ansicht mit Hochdruckanschluss
b Detailansicht mit Niederdruckanschluss (auf gleicher Ebene winkelversetzt zum Hochdruckanschluss
c Außenansicht

1 variabler Druckdämpfer
2 Druckbegrenzungsventil
3 Hochdruckanschluss
4 Niederdruckanschluss
5 Auslassventil
6 Spule
7 Mengensteuerventil
8 Befestigungsflansch
9 Dichtring
10 Kanal zum Förderkolben (Funktion der Druckdämpfung)
11 Förderkolben
12 Kolbendichtung
13 Kolbenfeder
14 mechanischer Antrieb

Befestigung: sujeción, fijación,
Flansch: brida.
dichtring: empaque, anillo de sellado
dichtung: empaque

240

[Handwritten notes in margins:]

Drehung: rotación
abfahren: recorre

ergeben: producir, dar
anregung: estímulo, excitación, incentivo

a
b
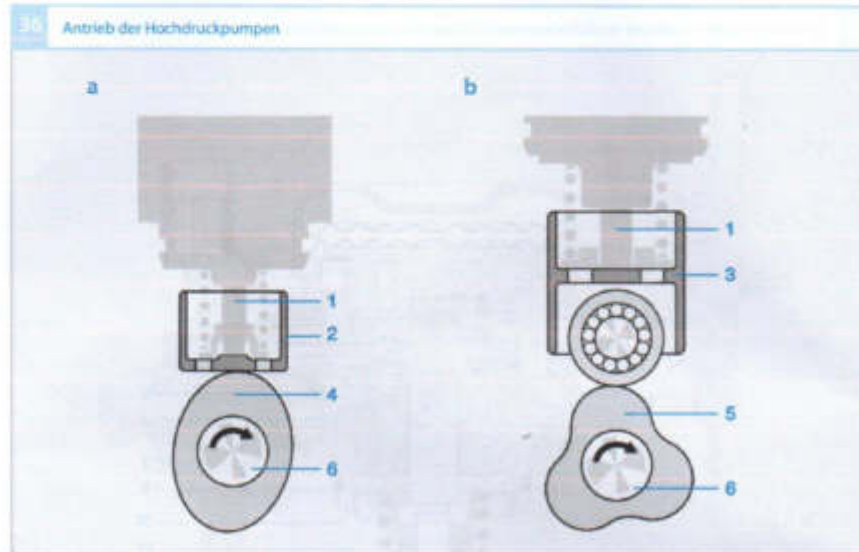
Bild 36
a Antrieb über Tassen-
   stößel
b Antrieb über Rollen-
   stößel

1 Pumpenkolben
2 Tassenstößel
3 Rollenstößel
4 Zweifachnocken
5 Dreifachnocken
6 Antriebswelle

[Handwritten left margin notes:]
Bestimmen: determina
Anzahl: cantidad
Nockenhebung: leva
Übertragung: transmisión, traspaso, trasladación
Zweifach = doble, duplicada
Tassenstößel: empujador de cubeta
Rollenstößel: empujador de rodillo
Antrieb: propulsión, mando, accionamiento
zeitlich: temporal, transitorio
auslegen: exhibir, cubrir
Trägheit: inercia.
durchmesser: calibre, diámetro.
ober: superior, alto.

und bestimmt über die Anzahl der Nocken-
erhebungen die Fördermenge der Pumpe.

Zur Übertragung der Hubkurve des No-
ckens auf den Förderkolben der Hochdruck-
pumpe werden bei einem Zweifach-Nocken
ein Tassenstößel und beim Drei- und Vier-
fach-Nocken ein Rollenstößel (Bild 36) ein-
gesetzt. Bei der Drehung der Nockenwelle
fährt der Stößel die Kontur des Nockens ab,
woraus sich die Hubbewegung des Förder-
kolbens ergibt. Im Förderhub nimmt der
Stößel die anstehenden Kräfte wie Druck-,
Massen-, Feder- und Kontaktkraft auf.

Mit dem Vierfach-Nocken ist eine zeitli-
che Synchronisierung von Förderung und
Einspritzung beim 4-Zylinder-Motor mög-
lich, d.h., bei jeder Einspritzung gibt es auch
eine Förderung. Damit wird zum einen die
Anregung des Hochdruckkreises reduziert,
zum anderen kann das Railvolumen redu-
ziert werden. Um sicherzustellen, dass bei
maximalem Kraftstoffbedarf des Motors der
Systemdruck noch ausreichend schnell vari-
iert werden kann, wird die maximale För-
dermenge auf den Maximalbedarf ausgelegt.

Faktoren, die das Förderverhalten beeinflus-
sen (z.B. Heißbenzin, Alterung der Pumpe,
Dynamik), werden dabei berücksichtigt.

Der Liefergrad der Hochdruckpumpe er-
gibt sich aus dem Verhältnis von tatsächlich
gelieferter Kraftstoffmenge zu theoretisch
möglicher Menge. Diese ist vom Kolben-
durchmesser und vom Hub abhängig. Der
Liefergrad ist über der Drehzahl nicht kons-
tant und hängt im unteren Drehzahlbereich
von Kolben- und anderen Leckagen sowie
im oberen Drehzahlbereich von Trägheit
und Öffnungsdruck des Ein- und Auslass-
ventils ab. Im gesamten Drehzahlbereich
wirkt sich das Totvolumen des Förderraums
und die Temperaturabhängigkeit der Kraft-
stoffkompressibilität aus.

Niederdruckdämpfer

Mit dem variablen Druckdämpfer (Bild 35,
Pos. 1) werden die durch die Hochdruck-
pumpe im Niederdruckkreis angeregten
Druckpulsationen gedämpft und auch bei
hohen Drehzahlen eine gute Füllung garan-
tiert. Der Druckdämpfer nimmt über die

241

vereinfacht: simplificado
querschnitt: sección transversal

Ansteuerkonzept des Mengensteuerventils für eine Hochdruckpumpe
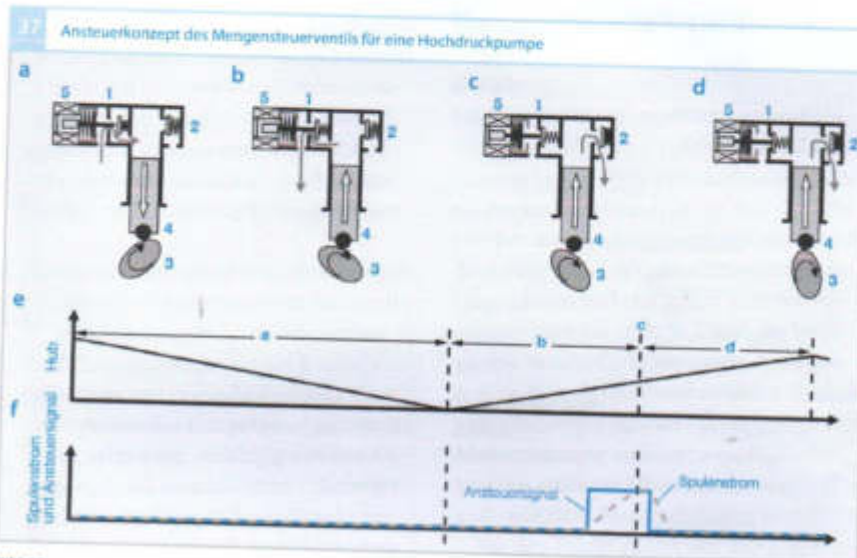
Bild 37

a–d vereinfachter Querschnitt der Hochdruckpumpe zu
      verschiedenen Zeitpunkten
a   Saughub, Mengensteuerventil geöffnet, Auslassventil
      geschlossen
b   Förderhub, Mengensteuerventil geöffnet, Auslassven-
      til geschlossen
c   Förderhub, Schließzeitpunkt des elektrisch angesteu-
      erten Mengensteuerventils, Öffnungszeitpunkt des
      Auslassventils
d   Förderhub, Mengensteuerventil bleibt auch nach
      Stromabschaltung geschlossen, Auslassventil geöffnet

e   Hubverlauf
f   Ansteuersignal und Spulenstrom des Mengensteuer-
      ventils

1   Mengensteuerventil
2   Auslassventil
3   Antriebsnocken
4   Kolben, Pfeil gibt die Bewegungsrichtung an
5   Spule

Pfeil: flecha.

Verformung seiner Membranen die im je-
weiligen Betriebspunkt abgesteuerte Kraft-
stoffmenge auf und gibt sie im Saughub zur
Füllung des Förderraums wieder frei. Dabei
ist ein Betrieb mit variablem Vordruck –
d. h. der Einsatz von bedarfsgeregelten Nie-
derdrucksystemen – möglich.

Mengensteuerventil
Mit dem Mengensteuerventil (Bild 35,
Pos. 7) wird die Bedarfssteuerung der Hoch-
druckpumpe realisiert (Bild 37). Der von
der Elektrokraftstoffpumpe gelieferte Kraft-
stoff wird über das Einlassventil des offenen
Mengensteuerventils in den Förderraum ge-
saugt. Im anschließenden Förderhub bleibt

das Mengensteuerventil nach dem unteren
Totpunkt weiterhin offen, so dass der im je-
weiligen Lastpunkt nicht benötigte Kraftstoff
unter Vordruck in den Niederdruckkreis
zurückgefördert wird. Nach Ansteuern des
Mengensteuerventils schließt das Einlass-
ventil, der Kraftstoff wird vom Pumpenkol-
ben verdichtet und in den Hochdruckkreis
gefördert. Das Motormanagement berechnet
den Zeitpunkt, ab dem das Mengensteuer-
ventil angesteuert wird in Abhängigkeit von
der Fördermenge und dem Raildruck. Der
Förderbeginn wird zur Bedarfsteuerung
variiert.

berechnen.

[19]   U. Spicher, T. Heidenreich, 2-Stratified-charge Combustion in Direct Injection Gasoline Engines, pp. 20-44, Hua Zhao, Advanced Direct Injection Combustion Engine Technologies and Development, Vol. 1, Woodhead Publishing, 2010.

Spicher, U., Heidenreich, T., stratified-charge combustion in DI gasoline engines. universität Karlsruhe.

of insulating air. Therefore, when viewed as a whole, the engine is indeed operating globally lean of stoichiometric, whereby large air ratio gradients form in the combustion chamber.

Figure 2.4 shows the range of the operating modes for gasoline direct injection engines in the engine operation map. The mixture preparation, mixture transportation to the spark plug, and reliable ignition of extremely lean mixtures are important preconditions for successful charge stratification (Wirth et al., 2003; Waltner et al., 2006). As a basic rule, three different combustion concepts for the implementation of mixture preparation and fuel guidance can be identified in the case of gasoline direct injection (see Fig. 2.5).

2.2.1   Wall-guided combustion systems

The majority of first-generation engines with stratified-charge operation were designed with wall-guided combustion systems, in which the mixture
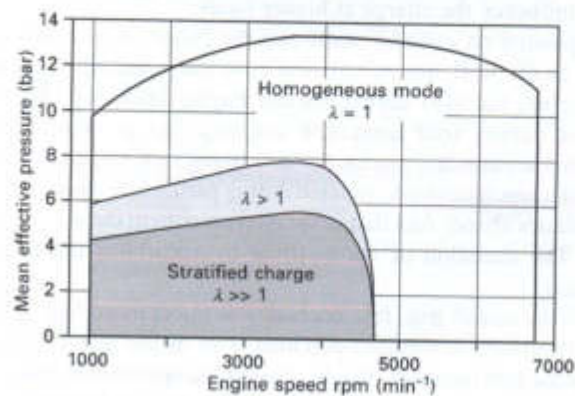
2.4 Operating modes in the engine operation map.
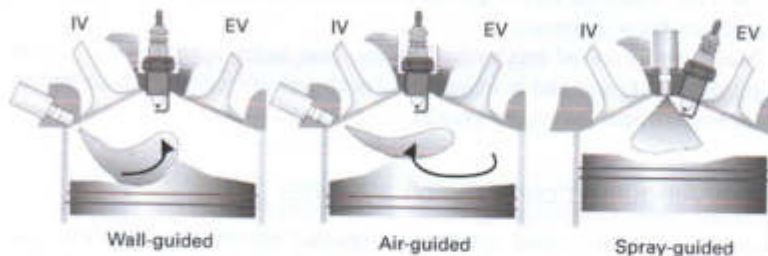
2.5 Wall-, air- and spray-guided combustion concepts (Fischer, 2004).

is prepared and transported to the spark plug via a wall in the combustion chamber. In most cases, the piston is shaped in such a way that the spray from the fuel injector is directed to the spark plug via the piston crown bowl. The mixture transportation during this process is normally supported by a swirl or tumble flow operation. As the injection jet is directed straight at the piston, this combustion procedure results in increased fuel deposits and emissions of unburned hydrocarbons. As the injection timing is linked to the piston position, it is therefore dependent on engine speed. Furthermore, desirable and stable in-cylinder flow patterns also depend on the engine speed; it is difficult to coordinate the injection and ignition timings across a wide engine speed/load range. In practice, it was not possible to realise the theoretical potential of gasoline direct injection for reducing fuel consumption in the case of wall-guided combustion operation for the following reasons:

- As mixture transport is linked to the piston position, it is difficult to properly coordinate the injection and ignition timing over a wide engine operating range.
- The tumble or swirl ports required to produce the desired intake flow adversely influence the charge at higher loads.
- Fuel is deposited on cylinder walls and the piston. A pronounced film may form on the wall, not only during cold starts and when the engine is warming up, but also during normal engine operation. Incomplete combustion causes soot formation and deposits at these locations. Deposits that accumulate cannot be fully burned off because of the low combustion temperatures encountered during partial load operation. These porous residues absorb fuel that is sprayed onto them during subsequent injection. The emission of combustible hydrocarbons increases as a consequence.
- Fuel enters the squish gap. It is necessary to inject more fuel during the transition from low- to medium-part load. This means that it is necessary to advance the fuel injection time to achieve the optimum ignition timing in terms of efficiency. However, it is possible that the injected fuel will spill over the edge of the crown bowl and penetrate the squish gap. In this case, since the fuel is not fully consumed, the emission of unburned hydrocarbons is greater.
- There are increased mechanical losses. Due to the special shape of the piston, it is larger and heavier than a conventional piston, which means the mechanical losses are greater.

2.2.2   Air-guided combustion systems

In contrast to wall-guided systems, air-guided combustion systems aim to reduce the hydrocarbon emissions that are a by-product of wall-guided systems by preventing the fuel from coming into contact with the walls of

the combustion chamber. This ideally eliminates fuel deposits on the wall of the combustion chamber. The aim in this case is to utilise change movement to effectively mix the fuel and intake air. The charge movement can also be supported by appropriately shaped piston crowns. The success of this method therefore depends on the directional orientation of the injection jet and on the generation of a specific charge movement. It is particularly important to ensure that the specific charge movement is sustained far into the compression phase in order to ensure transport of the mixture up to the spark plug. The swirl or tumble flow required to perform this operation reduces volumetric efficiency and therefore adversely affects performance.

Wall-guided operations also use highly distinctive air flows to direct the mixture which means that, in reality, clear differentiation of combustion systems is not practical. These systems can therefore be referred to as wall/air-guided combustion systems or first-generation gasoline direct injection.

2.2.3   Spray-guided combustion systems

An important feature of spray-guided combustion systems is the physical proximity of the spark plug to the fuel injector. It must be ensured, through optimum positioning of the spark plug in relation to the injection plume, that an ignitable mixture is present at the spark plug at the point of ignition for a wide range of operating conditions. Decisive factors in this case are the characteristics of the fuel injector. The injector must produce a spray pattern that is as robust and repeatable as possible, even when subjected to changes in back pressure or flow conditions. Furthermore, the system must ensure that extremely lean mixtures with steep stratification gradients at the edges of the fuel spray can be reliably ignited.

Owing to the special features of the spray-guided combustion system, particular attention must be paid to the following points:

- Coking at the spray nozzle. Coking may occur at the spray nozzle due to low combustion temperatures during part load operation and idling. These may have negative effects on the spray profile.
- Tolerances in the spray pattern. The spray pattern may change as a result of operating conditions or manufacturing tolerances and have undesirable effects on combustion performance.
- Coking at the spark plug. Low combustion temperatures during partial load operation and when idling can lead to coking at the spark plug and, as a consequence, misfiring.
- Thermal shock at the spark plug. If liquid fuel comes into contact with the hot spark plug, it cools the spark plug rapidly, therefore subjecting it to an extremely high thermal load.
- Effect of fuel injection time. As fuel injection and ignition are closely linked, fuel is injected at a thermodynamically favourable time shortly

before ignition at top dead centre (TDC). Very little time is therefore available to prepare the mixture.

- Effect of engine speed. High or low engine speeds change the speed of the intake air flow and the relative speeds of the cylinder charge and injected fuel spray. This affects mixture preparation and may interfere with air stratification in the area around the spark plug.

Owing to the problems mentioned above, a number of requirements must be satisfied in order to fully exploit the potential of a spray-guided combustion system in terms of its theoretical efficiency (Kemmler *et al.*, 2002; Hübner *et al.*, 2003):

- The spray quality of the fuel injector must be high and as reproducible as possible over all relevant engine speed and load ranges. This is the only way to achieve complete combustion with low carbon monoxide and unburned hydrocarbon emissions.
- The mixture preparation must also be optimised at retarded fuel injection times in order to achieve a thermodynamically favourable combustion process.
- To minimise hydrocarbon emissions, the geometry of the combustion chamber and spray pattern must be such that contact between liquid fuel and the combustion chamber is as low as possible.
- The mixture composition at the spark plug can be selectively modified in order to maximise the stability of combustion performance using a highly flexible injection system capable of multiple injections within a cycle.
- The sensitivity of the fuel injector to contamination must be as low as possible, and it must demonstrate considerable long-term stability in terms of the spray characteristic.
- High injection pressures reduce the mixture preparation time, which is important when fuel injection is retarded for the purposes of optimising the course of combustion. Faster evaporation of smaller fuel droplets can reduce particle emissions.
- A robust, efficient and variable ignition system that satisfies the special requirements of spray-guided operation must be employed. The spark plug should have the highest possible resistance to wear and should also be resistant to thermal shock.

In addition to the known thermodynamic benefits of direct injection, such as reduced pumping losses, the effect of modified calorific gas properties, lower temperature as a result of the internal evaporation of fuel, etc., spray-guided combustion systems also provide additional benefits in terms of efficiency. When compared with wall/air-guided systems, fuel wetting on the pistons or combustion chamber walls is reduced to the point where it is essentially eliminated. This reduces the emissions of unburned hydrocarbons. Improved

primary atomisation of the fuel spray and a more compact mixture cloud leads to faster and more comprehensive fuel conversion, which further reinforces the similarities between this process and the constant-volume process.

Direct-injected engines are less sensitive to knock due to the evaporation of fuel in the combustion chamber (internal cooling). The allowed increase in compression ratio improves the thermal efficiency of the engine throughout the entire operating range. Additionally, the higher final compression pressure also has a positive effect on the ignition of the mixture.

With wall/air-guided combustion systems, the swirl and/or tumble flow necessary to transport the mixture to the spark plug reduces the volumetric efficiency and therefore has disadvantages in terms of performance. In principle, this restriction does not apply to spray-guided systems; in this case the inlet ports can be designed as filling channels to optimise the full-load performance. As no specific charge movement is required for transportation of the mixture to the spark plug for spray-guided systems, the charge movement can be specified so that the turbulence in the combustion chamber supports the flame propagation and therefore comprehensive and rapid burn-out of the mixture.

Spray-guided systems can also benefit from efficient mixture preparation and a centrally located fuel injector and spark plug. This ensures favourable conditions that affect the efficiency of the combustion process and also unburned hydrocarbon emissions.

The positive effect on the dynamic engine response characteristics is a further consequence of direct fuel injection and unthrottled operation. When a sudden change from part-load operation to full-load operation occurs as the throttle is opened in a conventional quantity-regulated engine, the air intake system must initially be filled and a higher air intake pressure must be allowed to build up. With direct fuel injection and unthrottled stratified-charge operation, the full air intake pressure is nearly always present and the torque response is much faster.

The advantages of spray-guided combustion systems mean that stratified-charge operation can be designed not only to optimise low pollutant emissions and low fuel consumption; considerable advantages over conventional engines with port fuel injection or first-generation direct fuel injection also exist.

The intake port wall film that forms and enriches the mixture during rapid load shifts with port fuel injection is eliminated. In addition, gasoline direct injection also presents the opportunity of using special operating modes, such as stratified starting or homogeneous lean-burn operation mode.

Comparison of nozzle types

The three standard types of nozzles used nowadays in fuel injectors for gasoline direct injection are shown in Fig. 2.6. The nozzles can be differentiated

[20]   Denk, F., Elektromagnetische Einspritzventile für Ottomotoren, Universität Ilmenau, Münster, Alemania, 2018.

2.6   Elektronik

2.6.1   Konzeptübersicht

In der Arbeit von Glasmachers [Gla08] wird das zeitliche Stromprofil und Spannungsprofil der Ansteuerung eines Kraftstoffeinspritzventils untersucht. Es wird dargestellt, dass in einem Kraftfahrzeug typischerweise die Batteriespannung sowie eine aus der Batteriespannung generierte erhöhte Boostspannung zur Ansteuerung genutzt wird. Der Einfluss auf die Kraftstoffmenge durch Exemplarstreuung, Variation der Boostspannung, Variation der Batteriespannung sowie dem Kraftstoffdruck und der Temperatur werden analysiert.

Darüber hinaus wird die Möglichkeit betrachtet, in der Erfassung der elektrischen Größen Spulenstrom und Spulenspannung Rückschlüsse auf das hydraulische Verhalten des Einspritzventils zu ziehen. Die zum Einsatz kommende Elektronik lässt sich in den Teil Elektronik des Stromreglers und in den Teil Elektronik des DC-DC-Wandlers aufteilen. Der DC-DC-Wandler ist ein Bestandteil der elektrischen Versorgung des Stromreglers.

2.6.2   Prinzip des Stromreglers

Der Regler ist als Highside- und Lowside-Treiber realisiert. Elektronische Highside-Schalter, sogenannte Highside-FET's (FET = Feld Effekt Transistor), stellen Energie aus der Batterie des Systems oder aus einem DC-DC-Wandler, dem Booster, bereit. Der Stromfluss durch das Einspritzventil wird mittels eines weiteren Lowside-FET's, dem elektronischen Lowside-Schalter gegen Masse, ermöglicht.

Je nach Auslegung der Verbrennungsstrategie des Verbrennungsmotors und somit auch der Einspritzstrategie ist es möglich, mehrere Einspritzventile gemeinsam aus einem Boost-FET oder Batterie-FET zu betreiben. Dies ist der Fall, wenn gewährleistet ist, dass bei den gemeinsam zu betreibenden Einspritzventilen immer nur ein Einspritzventil einen elektrischen Einspritzvorgang leisten muss. Die Auswahl des jeweiligen Einspritzventils wird durch den jeweiligen FET gegen Masse ermöglicht.

Stellt sich die Forderung einer zeitgleichen Bestromung mehrerer Einspritzventile, so spricht man hier vom sogenannten Überlappbetrieb. In diesem Fall ist die hydraulische Einspritzung in jeder zeitlichen Zuordnung der Einspritzventile zueinander möglich. Da aus Kostengründen nur ein DC-DC-Wandler pro Motorsteuergerät implementiert wird, können Überlappprozesse Einflüsse in der nominalen Boostspannung haben.

Ein elektronisches Schaltungsbeispiel in Abb. 14 beschreibt, wie zwei Einspritzventile mit gemeinsamen Batterie- und Boost-FET elektronisch zu betreiben sind. Die farbig zugeordneten Strompfade sind schaltungstechnisch in Abb. 14 skizziert und Spulenstrom und Spulenspannung in Abb. 15 zeitlich dargestellt. Der Potentialbezug für die Masse, dem Minuspol der Batterie, wird in der Regel mit GND (im Englischen Ground) bezeichnet. Die Bezeichnung für das Einspritzventil erfolgt im Schaltungsteil, aus dem Englischen abgeleitet, mit Injector.
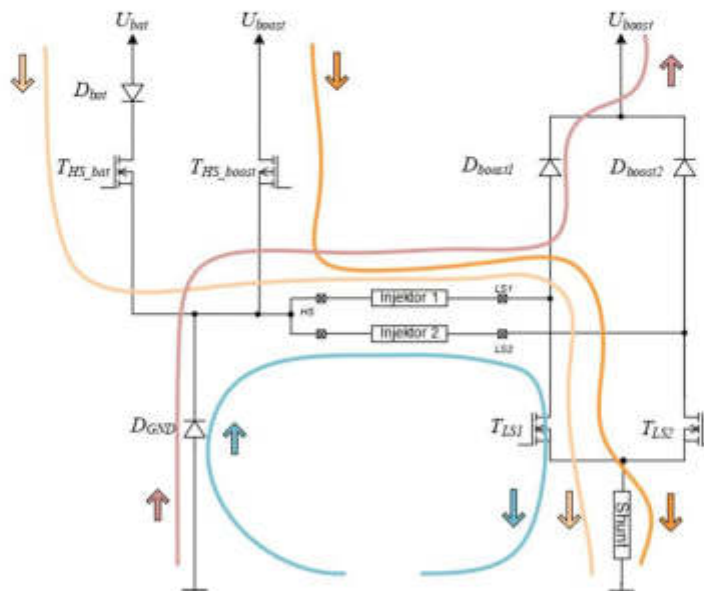
Abbildung 14: Die Variationsmöglichkeiten der Strompfade

Batterieschalter T_{HS_bat} und Boostschalter T_{HS_boost} führen dem über den Lowsideschalter T_{LS1} ausgewählten Einspritzventil 1 elektrische Energie zu. Die Verbindungsstecker am Steuergerät zum Einspritzventil entsprechen den Bezeichnungen *HS* und *LS1*.

Elektrische Energie wird vom Einspritzventil bei geöffnetem Batterie- und Boostschalter in den Booster zurückkommutiert, wenn der Lowsidesschalter offen ist, oder läuft bei geschlossenem Lowsideschalter über die Freilaufdiode D_{GND1} frei.
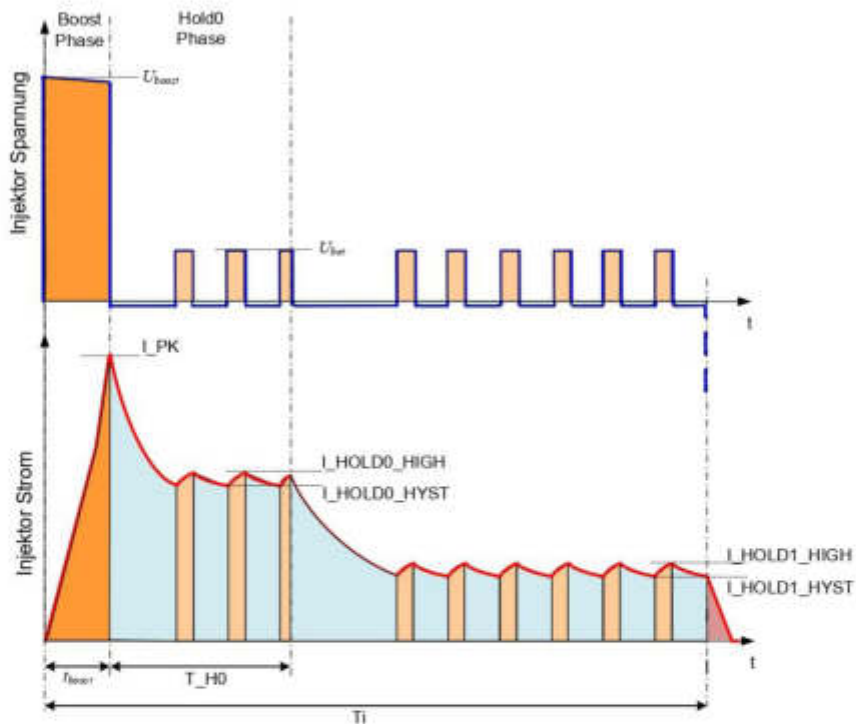
Abbildung 15: Schematisch zeitliche Darstellung von Spulenspannung und Spulenstrom

Das Konzept des Stromreglers erzeugt zeitlich je nach Einstellung der Stromsollwerte I_PK, I_HOLD0_HIGH, I_HOLD0_HYST und I_HOLD1_HIGH, I_HOLD1_HYST ein Spannungspulsmuster zeitlich unterschiedlicher Phasen. Die elektrische Beaufschlagung seitens der Motorsteuerung erfolgt aus U_{boost} oder U_{bat}. Die elektrische Beaufschlagung seitens des Einspritzventils erfolgt in der Freilaufphase mit $U_{freewheel}$ und in der Abkommutierungsphase mit U_{clamp} am Ende der zeitlichen Bestromung. Der Beginn der hydraulischen Einspritzphase erfolgt zeitlich im Bereich t_{boost} und T_H0.

2.6.3 DC-DC-Wandler

Der Wandler hat die Funktion, die zum Einspritzbeginn anliegende Boostspannung zu ge-
nerieren. Er arbeitet als elektronischer Aufwärtswandler gemäß Tietze [Tie93], der prinzi-
piell mittels einer Speicherinduktivität und eines elektronischen Schalters die gewünschte
Spulenspannung einem Zwischenkreiskondensator zur Verfügung stellt. Die Kriterien für
die Wandlerauslegung sind hier die Nominalspannung und die mittlere Leistung. Während
der Boostphase wird dem Einspritzventil eine bestimmte Boostenergie E_{boost} zugeführt.
Diese Boostenergie wird dem Zwischenkreiskondensator entnommen und kann über die
zeitliche Integration von t_{boost} der elektrischen Einspritzventilleistung, dem Produkt aus
dem Peakstrom I_{boost} und der Boostspannung U_{boost} berechnet werden. Aus der Zylinderan-
zahl Z des Motors, seiner Maximaldrehzahl n [1/min] sowie der Anzahl der Einspritzpulse
pro Verbrennungszyklus NR_PLS errechnet sich in Gl. (2.1) eine mittlere Leistung von

$$P_{DC/DC} = \frac{E_{boost}}{T_{period}} \cdot \text{NR\_PLS} = \frac{E_{boost}}{120} \cdot n \cdot Z \cdot \text{NR\_PLS} \qquad (2.1)$$

Die Auslegung des DC-DC-Wandlers ist ein maßgeblicher Kostenfaktor innerhalb einer
Motorsteuerung. Unter anderem geht die mechanische und leistungsmäßige Auslegung der
Speicherdrossel sowie der spezifizierte Arbeitsbereich über die Temperatur in die Kalku-
lation mit ein. In den Dokumentationen der Robert Bosch GmbH für das Hochdruckein-
spritzventil HDEV5 [Rob13], im Kraftfahrtechnischen Taschenbuch [Rei11a] sowie im
Nachschlagewerk für Autoelektronik [Rei11b] wird beschrieben, dass die allgemeine Aus-
legung der nominalen Boostspannung U_{boost} aktuell bei 65 V liegt.

Eine verfügbare kleinere Spannungsklasse und somit eine Reduktion der Versorgungsspan-
nung des DC-DC-Wandlers würde der Spannungsbereich mit einer Boostspannung < 60 V
darstellen. Eine Übersicht verfügbarer MOSFET-Schalter zeigt die Firma Onsemi unter
Angabe der maximalen Versorgungsspannung [Ons16]. Üblicherweise steigt die Anzahl
von Applikationen mit MOSFET-Schalter einer kleineren Spannungsklasse. Somit wirkt
sich dies direkt auf den von der Stückzahl abhängigen Komponentenpreis aus. Eine gleiche

Kostenbetrachtung kann hier bei dem jeweils im DC-DC-Wandler benötigten Speicher-
kondensator durchgeführt werden. Die gespeicherte elektrische Energie im Kondensator
ergibt sich nach Gerthsen [Ger89] zu

$$E_{cap} = \frac{1}{2} \cdot C_{cap} \cdot U_{boost}^2 \tag{2.2}$$

Die jeweils notwendige Maximalspannung besitzt auch einen direkten Einfluss auf das
Layout und die Konstruktion der elektronischen Motorsteuerung, da sich hier mit steigen-
der Spannung die einzuhaltenden Isolationsabstände spannungsführender Teile vergrö-
ßern. Somit wird im Design mehr Platz beansprucht, was sich wiederum negativ auf die
Kosten auswirkt.

2.6.4 Spulenspannungsprofil und Spulenstromverlauf

Die Messung an einem Einspritzventil ohne Ankerfreiweg veranschaulicht in Abb.13 eine
Einspritzventilansteuerung mit Aufzeichnung von Spulenstrom, Spulenspannung und
elektrischer Energie sowie das am Einspritzventil abgegriffene erzeugte Körperschallsig-
nal, welches Rückschlüsse auf Anker- und Nadelbewegung liefern kann. Die elektrische
Boostenergie wird gemäß Gl. (2.3) ermittelt.

$$E_{boost} = \int\limits_{0}^{t_{boost}} U(t) \cdot I(t)\, \mathrm{d}t \tag{2.3}$$

Die Integration beginnt mit dem Einschalten des Boosters, dem Anlegen der Boostspan-
nung. Sie endet mit dem Abschalten der Boostspannung, in der Regel, wenn der einge-
stellte Peakstrom erreicht ist.

Zu Beginn des zeitlichen Verlaufs der Bestromung zeigt das Körperschallsignal in Abb. 16
im Bereich des maximalen Spulenstroms ein ausgeprägtes negatives Minimum (1), was
den erstmaligen Jochaufschlag des stark beschleunigten Ankers und der Nadel darstellt.
Zum Ende der zeitlichen Abfolge der Bestromung ist ein ausgeprägtes Maximum (2) im
Körperschall zu erkennen. Dies ist der Punkt, an dem der Anker zusammen mit der Nadel

erstmalig seinen hydraulischen Sitz, vorab beschleunigt durch die Schließfeder, wieder erreicht.
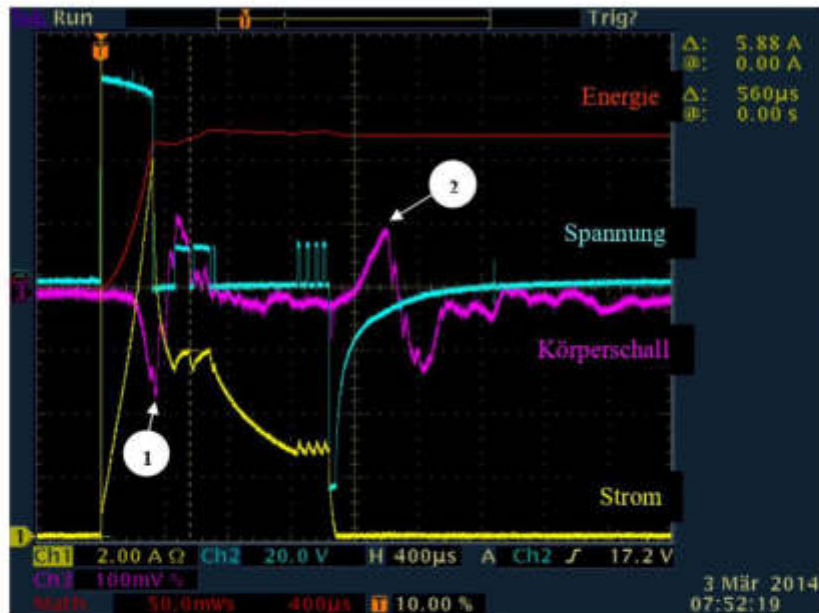
Abbildung 16: Elektrische Einspritzventilsignale und abgegriffenes
Körperschallsignal des Einspritzventils

2.7   Funktionen der Software

Die Produktionskosten für Software in Serienentwicklungen stellen sich im Vergleich zu materiell darstellbaren Bauteilkosten anders dar. Der Bereich „Einspritzsoftware" ist physikalisch gesehen Programmcode im Flashspeicher. Dieser nutzt dynamisch in der Softwarelaufzeit Variablen des RAM-Speichers (im Englischen Random-Access Memory), Register, Pipelines sowie Interrupts und Stacks eines Mikroprozessors. Die in modernen Motorsteuergeräten implementierte Software arbeitet auf speziell für den Automotive-Sektor Powertrain entwickelten Mikrocontrollern.

Von der Infineon AG wird unter dem Marktnamen „Tricore" eine Familie von Mikrocontrollern angeboten, die den Anforderungen im Automotive-Bereich genügen und durch

[21] (2021) Pico Technology [Online], disponible en: https://www.picoauto.com/library/automotive-guided-tests/gasoline-direct-injection-injector-voltage-and-current/

[22] Edwards, J., D., "*Electrical Machines and Drive,*" 1era ed., Macmillan Education, Londres, Inglaterra, 1991.

1 General Principles

1.1 Introduction

In 1820 Oersted discovered the magnetic effect of an electric current, and the first primitive electric motor was built in the following year. Faraday's discovery of electromagnetic induction in 1831 completed the foundations of electromagnetism, and the principles were vigorously exploited in the rapidly growing field of electrical engineering [1]. By 1890 the main types of rotating electrical machine had been invented, and the next forty years saw the development of many ingenious variations, along with refinement of the basic types. This was the golden age of machine development; electronics was in its infancy, and the rotating machine was king. Many machines are now obsolete which were once made in large numbers. Thus the cross-field DC machines, or rotary amplifiers, have been replaced by solid-state power amplifiers; while the Schrage motor and other ingenious variable-speed AC machines have given way to the thyristor-controlled DC motor and the inverter-fed induction motor.

Electrical machines may be broadly divided into two groups. The first group comprises the three types of classical machine which were invented in the nineteenth century: simple DC machines, AC synchronous machines and AC induction machines. In the second group are the so-called non-classical machines which are of much more recent origin: stepper motors, brushless DC motors and switched reluctance motors. The distinctive feature of the non-classical machines is their dependence on power electronic circuits to control the currents in the active coils; their development was made possible by advances in semiconductor technology. Stepper motors are complementary to the classical machines because they are designed for motion in discrete steps rather than continuous rotation. Brushless DC motors and switched reluctance motors, on the other hand, are designed for continuous rotation; they are in direct competition with classical machines and are beginning to displace them in some applications.

Classical machines still dominate the world of electrical machines and drives. Most of the motive power for industry is supplied by induction motors operating directly from the AC mains as constant-speed drives. There are many applications, however, which require an adjustable-speed drive; this can be achieved by power electronic control of a classical machine. The dominance of the classical

1

254

machines, and the extension of their domain through the use of power electronics, is reflected in the structure of this book.

As might be expected from their historical development, the classical machines share a common physical basis for their operation. This permits a progressive development of the subject from the general principles outlined in this chapter, beginning with the effects discovered by Oersted and Faraday. These two effects and their interrelation are seen most clearly when a current-carrying conductor is free to move in a magnetic field of constant intensity; this case is considered in section 1.2. But in most practical machines the conductors are not free to move; they are embedded in slots in the iron core of the machine, and forces act on the iron as well as on the conductors. It is necessary therefore to consider force production and EMF generation in more general terms, which is done in sections 1.3 and 1.4. Magnetic materials form an essential part of electrical machines; some of their properties are discussed in section 1.5, and the important concept of the magnetic circuit is introduced in section 1.6.

1.2 Conductor moving in a magnetic field

When a conductor moves in a magnetic field, an EMF is generated; when it carries a current in a magnetic field, a force is produced. Both of these effects may be deduced from one of the most fundamental principles of electromagnetism, and they provide the basis for a number of devices in which conductors move freely in a magnetic field. It has already been mentioned that most electrical machines employ a different form of construction, and the concepts developed in the next two sections are necessary for a proper understanding of their operation. Nevertheless the equations developed in this section for the force and the induced EMF remain valid for many practical machines; this important and useful result will be justified in chapter 2.

Induced EMF in a moving conductor

Consider a conductor moving with a velocity denoted by the vector u in a magnetic field B (figure 1.1). If the conductor slides along wires connected to a voltmeter, there will be a reading on the meter, showing that an EMF is being generated in the circuit. The effect may be explained in terms of the Lorentz equation for the force on a moving charge q

$$F = q(E + u \times B) \quad \text{newtons} \tag{1.1}$$

where q is the charge in coulombs, E the electric field strength in volts/metre, u the velocity in metres/second, and B the magnetic flux density in teslas. If the conductor is initially at rest, there will be no electric field E and no reading on
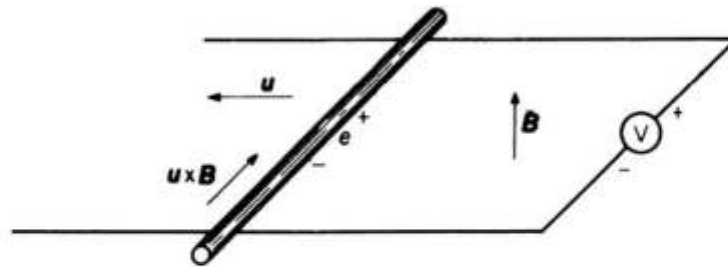
Figure 1.1 Moving conductor: induced EMF

the voltmeter. When the conductor is moving with velocity u, a force $qu \times B$ will act on any charged particle within the conductor, and the free charges (conduction electrons) will begin to move under the action of the force. There will be an accumulation of negative charge at one end of the conductor, leaving a surplus of positive charge at the other end; this will set up an electrostatic field E, and charge separation will continue until the force qE is exactly equal and opposite to $qu \times B$. The net force is then zero; there is no further motion of charge, and we have

$$E = -u \times B \tag{1.2}$$

The quantity $u \times B$ may be regarded as an induced electric field produced by the motion of the conductor, and this is opposed by an equal and opposite electrostatic field E produced by a distribution of electric charge. In virtue of the electrostatic field E there will be an electrostatic potential difference between the ends of the conductor given by the line integral of E along any path joining the ends PQ

$$v = -\int_P^Q E \cdot ds \text{ volts} \tag{1.3}$$

and this will be measured by a voltmeter connected between the wires. From eqn (1.2) this may be written as $v = e$, where

$$e = \int_P^Q u \times B \cdot ds \text{ volts} \tag{1.4}$$

and e may be regarded as the EMF induced in the conductor by its motion in the magnetic field. The integral may conveniently be taken along the axis of the conductor, which is a line of length l denoted by the vector $l = PQ$. If u, B and l are mutually perpendicular, the induced electric field $u \times B$ will be parallel to l, and its direction is determined by the right-hand screw rule of the vector product: if you look in the direction of $u \times B$, then B is displaced clockwise

from u. For a magnetic field which is uniform along the length of the conductor, we then have

$$e = \int_0^l uB \, ds = Blu \text{ volts} \tag{1.5}$$

and the sign of e is determined from the direction of $u \times B$ as shown in figure 1.1. Equation (1.5) is commonly known as the 'flux cutting rule', or more accurately as the *motional induction formula*; its relation to Faraday's law of electromagnetic induction is discussed in section 1.3.

Conductor resistance

Suppose that a resistor is connected in place of the voltmeter, so that a current i flows. If this current is distributed uniformly over the cross-section of the conductor, and the cross-sectional area is A, the magnitude of the current density is

$$J = \frac{i}{A} \text{ amperes/metre}^2 \tag{1.6}$$

and its direction is along the conductor, as shown in figure 1.2. Since the total force acting on unit charge is $E + u \times B$, Ohm's law for the moving conductor is

$$J = \sigma(E + u \times B) \tag{1.7}$$
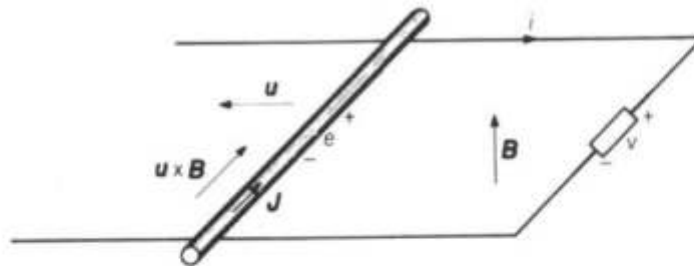
Figure 1.2   Moving conductor: induced current

where σ is the conductivity of the material. Thus the electrostatic field E must be slightly less than the induced electric field $u \times B$ when a current is flowing, the resultant force per unit charge being just sufficient to maintain the flow of current. The potential difference between the ends of the conductor is now given by

$$v = -\int E \cdot ds = \int u \times B \cdot ds - \int \frac{1}{\sigma} J \cdot ds$$

$$= \int_0^l uB \, ds - \int_0^l \frac{i}{\sigma A} \, ds = Blu - \frac{li}{\sigma A} \qquad (1.8)$$

or

$$v = e - Ri \qquad (1.9)$$

where $e = Blu$ is the induced EMF in volts and $R = l/\sigma A$ is the resistance of the conductor in ohms. The system may be represented by an equivalent circuit, as shown in figure 1.3.

Electromagnetic force on a conductor

Take the same configuration of a conductor in a magnetic field, and suppose initially that the conductor is stationary (figure 1.4). If a current i is flowing,
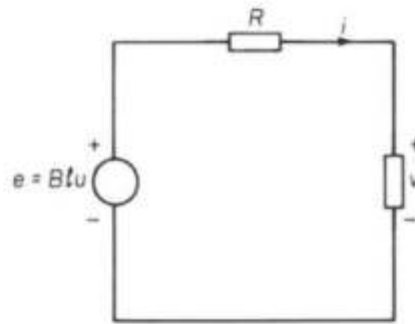
Figure 1.3   Equivalent circuit for a moving conductor
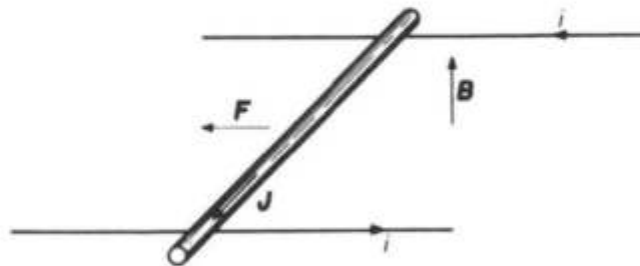
Figure 1.4   Force on a current–carrying conductor

there will be a flow of free charge along the conductor; let ρ be the charge per unit volume, and U the average drift velocity of the charge. This moving charge will experience a force in a magnetic field, and from eqn (1.1) the force per unit volume is

$$f = \rho U \times B \quad \text{newtons/metre}^3 \tag{1.10}$$

Since the free charge cannot escape from the sides, this force is transmitted to the conductor. We may express the force in terms of the current by noting that the current density J is given by

$$J = \rho U \tag{1.11}$$

The force per unit volume is therefore

$$f = J \times B \tag{1.12}$$

and the total force on the conductor is given by the volume integral

$$F = \int J \times B \, dv \quad \text{newtons} \tag{1.13}$$

In the simple case of a uniform current density given by $J = i/A$ (eqn 1.6), and a uniform magnetic flux density B perpendicular to the conductor

$$F = \int JB \, dv = \int_0^l \frac{iB}{A} A \, ds = Bli \quad \text{newtons} \tag{1.14}$$

From eqn (1.12) the direction of the force is given by the vector product $J \times B$, and it is perpendicular to both the conductor and the magnetic field.

Electromechanical energy conversion

Figure 1.5 shows the conductor connected to a voltage source v. There is a current i flowing, and the conductor is moving with a velocity u. The directions
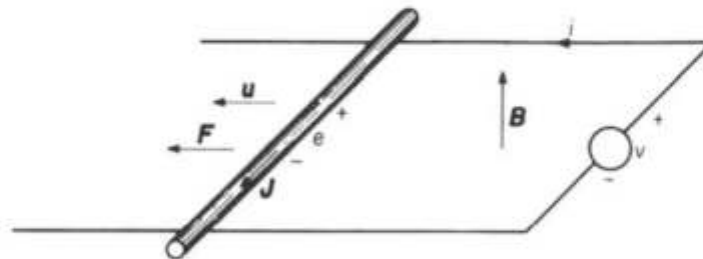
Figure 1.5 Moving current–carrying conductor

of the force F and the induced EMF e are shown in the figure. Since the force F and the velocity u are in the same direction. the conductor does mechanical work at the rate

$$P_m = F \cdot u = Fu \ \text{watts} \tag{1.15}$$

The voltage source v is driving a current i into the circuit, and it therefore does work at the rate

$$P_e = vi \ \text{watts} \tag{1.16}$$

Since the direction of current flow is the reverse of that in figure 1.2, eqn (1.9) becomes

$$v = e + Ri$$
$$= Blu + Ri \tag{1.17}$$

and we also have the force equation

$$F = Bli \tag{1.14}$$

Multiplying eqn (1.17) by i and eqn (1.14) by u gives

$$P_e = vi = Blui + Ri^2$$
$$P_m = Fu = Bliu$$

It follows that

$$P_e = P_m + Ri^2 \tag{1.18}$$

showing that the electrical input power P_e is equal to the mechanical output power P_m plus the ohmic losses in the conductors; the device is acting as a motor, converting electrical energy into mechanical energy. For the current to flow in the direction shown, the applied voltage v must exceed the induced EMF e; if v is smaller than e the direction of current flow is reversed, and the direction of the force F is reversed in consequence. The conductor then absorbs mechanical energy at the rate $P_m = Fu$; the voltage source likewise absorbs electrical energy at the rate $P_e = vi$, and

$$P_m = P_e + Ri^2 \tag{1.19}$$

The device is acting as a generator, converting mechanical energy into electrical energy plus ohmic losses. Thus the process of energy conversion is reversible, and there is no fundamental difference between generator and motor action.

Applications

Among the best-known applications of the conductor in a magnetic field are the moving-coil loudspeaker (figure 1.6) and the moving-coil meter (figure 1.7). In each of these devices the conductor, in the form of a coil, moves in the uniform radial field of a permanent magnet; motion of the coil is opposed by a spring, giving a displacement proportional to the coil current. The direct proportionality of the force or displacement to the current makes the moving-coil principle particularly useful for instrumentation. In the force-balance accelerometer, for instance, a feedback system senses the coil displacement and adjusts the current until the electromagnetic force exactly balances the acceleration reaction force. The coil current then gives a measure of the acceleration precise enough for modern inertial navigation techniques.

Another application is a special type of electrical machine. Most conventional motors and generators depend for their operation on the force between magnetised iron parts; but in the homopolar machine the force is developed directly on a conductor moving in a magnetic field of constant intensity. Faraday in 1831 made the first generator using this principle, in the form of a circular disc rotating in a magnetic field (figure 1.8). Each element of the disc at a radial distance r from the axis is moving with velocity $u = \omega r$ perpendicular to the magnetic field. The induced electric field $u \times B$ is directed along the radius, so there will be an
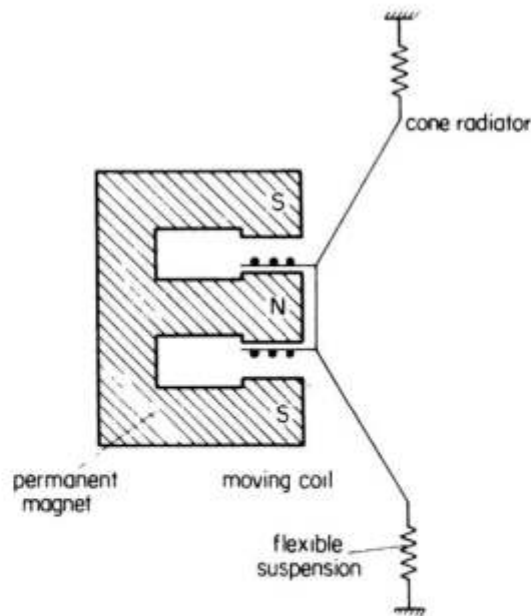
Figure 1.6 Moving-coil loudspeaker
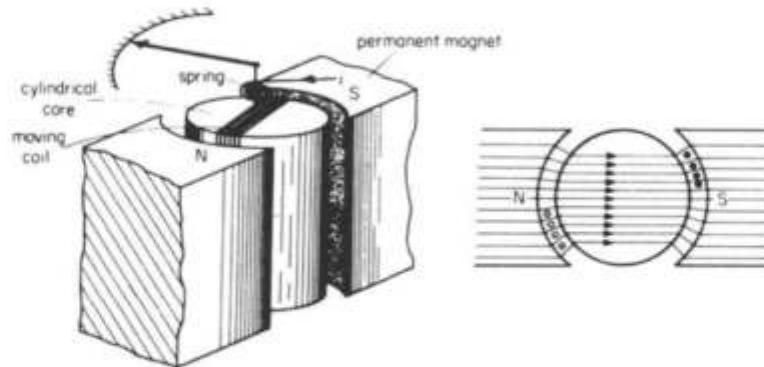
Figure 1.7 Moving–coil meter

EMF induced between the centre of the disc and the periphery. Integration along a path such as PQ gives

$$e = \int_P^Q u \times B \cdot ds = \int_0^a B\omega r \, dr = \tfrac{1}{2}B\omega a^2 \quad \text{volts} \tag{1.20}$$
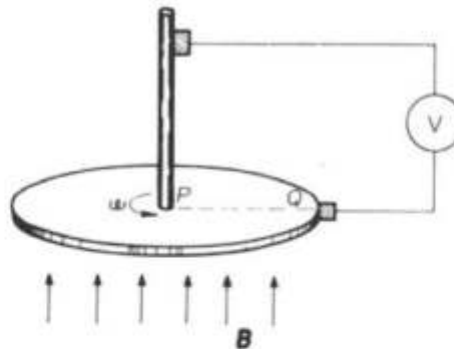
Figure 1.8 Faraday disc generator

where a is the radius of the disc in metres and ω the angular velocity in radians/second. The generated voltage is rather low for small machines at normal speeds, if the flux density is limited (as it usually is) by the saturation of iron to about 2 teslas. For example, if $a = 100$ mm, $\omega = 100$ rad/s (≈ 1000 rev/min) and $B = 2$ T, then $e = 1$ V. Large machines have been built for special low-voltage heavy-current applications, and an interesting development is the use of supercon-

ducting coils to generate very high magnetic fields [2]. Superconducting machines can be smaller, lighter and more efficient than their conventional counterparts, but their commercial application is restricted at present by the poor reliability of liquid helium refrigerators for the superconducting coils.

1.3 Electromagnetic induction

It was shown in section 1.2 that an EMF is induced in a conductor when it moves in a magnetic field. An EMF can also be induced in a stationary circuit by a time-varying magnetic field. If this magnetic field is produced by currents flowing in conductors or coils, the EMF can be induced merely by changing the current; no motion of any part of the system is required. The effect is termed transformer induction, and it appears to be physically quite distinct from motional induction. Both effects are included in Faraday's law of electromagnetic induction, which relates the induced EMF in a circuit to the rate of change of the magnetic flux linking the circuit.

Flux linkage

If a circuit consists of a conductor in the form of a simple closed curve C, the magnetic flux Φ linking the circuit is defined by the surface integral

$$\Phi = \int_S B \cdot da \quad \text{webers} \tag{1.21}$$

where S is any surface spanning the boundary C of the circuit. If the magnetic field is uniform and the circuit has an area A perpendicular to the field, this reduces to the simple expression

$$\Phi = BA \tag{1.22}$$

The concept of flux linkage arises when it is desired to calculate the flux linking a multi-turn coil. It is possible in principle to devise a twisted surface resembling an Archimedian screw, bounded by the turns of the coil, and to evaluate the integral in eqn (1.21) over this surface. But it is simpler to suppose that each turn links a certain amount of flux, so that the total flux linking the coil is the sum of the contributions from the individual turns. Thus if each turn links a flux Φ and the coil has N turns, then the total flux linking the coil, or flux linkage, is given by

$$\psi = N\Phi \quad \text{webers} \tag{1.23}$$

If the magnetic field is uniform and parallel to the axis of the coil, and each turn has an area A, then

$$\psi = NBA \qquad (1.24)$$

Usually the field is not uniform and the flux through an individual turn will depend on its position (figure 1.9). The total flux linking the coil is then given by the sum

$$\psi = \sum_{r=1}^{N} \Phi_r \text{ webers} \qquad (1.25)$$
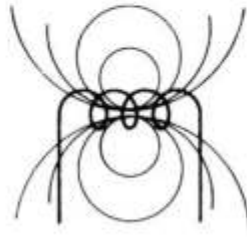
Figure 1.9 Flux linkage: non–uniform field

and an average flux per turn may be defined by the relation

$$\Phi_{av} = \psi/N \qquad (1.26)$$

Inductance

If permanent magnets are excluded, the flux linking a coil will depend on (a) the current flowing in the coil and (b) currents flowing in any adjacent coils or conductors.

Self-inductance

With a single coil carrying a current i we have

$$\psi = f(i) \qquad (1.27)$$

When there are no ferromagnetic materials present (the coil is air-cored) the relationship is linear; thus

$$\psi = Li \qquad (1.28)$$

where L is a constant known as the *self-inductance* of the coil. The unit of L is the henry when ψ is in webers and i in amperes. When the coil has an iron core the relationship between ψ and i is no longer linear, on account of the magnetic properties of the iron. The form of eqn (1.28) may still be used, but the coefficient L is no longer a constant; this can be made explicit by writing

264

$$\psi = iL(i) \tag{1.29}$$

In order to simplify the analysis it is often assumed that the inductance of an iron-cored coil is a constant; this assumption must be used with caution, for it can sometimes give completely erroneous results. This point will be discussed more fully in section 1.6.

Mutual inductance

With two coils the flux linkages are functions of the coil currents and the geometry of the system. Thus

$$\psi_1 = L_1 i_1 + M_{12} i_2$$
$$\psi_2 = L_2 i_2 + M_{21} i_1 \tag{1.30}$$

where ψ_1 and ψ_2 are the flux linkages for the two coils, and i_1 and i_2 are the corresponding currents. The coefficients M_{12} and M_{21} are known as the *mutual inductances.* When the coils do not have iron cores, it may be shown [3] that the mutual inductance coefficients are constant and equal, that is

$$M_{12} = M_{21} = M \tag{1.31}$$

If ψ_{12} is the flux linking the first coil due to a current i_2 in the second, and ψ_{21} is the flux linking the second coil due to a current i_1 in the first, then the mutual inductance is given by

$$M = \frac{\psi_{12}}{i_2} = \frac{\psi_{21}}{i_1} \tag{1.32}$$

This reciprocal property is particularly useful when the mutual inductance has to be measured or calculated, for one of the two alternative expressions in eqn (1.32) is often easier to evaluate than the other.

Faraday's law

Faraday's law of electromagnetic induction states that the EMF induced in a circuit is proportional to the rate of change of flux linkages. In SI units the constant of proportionality is unity, so that

$$e = \pm \frac{d\psi}{dt} \text{ volts} \tag{1.33}$$

The question of the sign in eqn (1.33) sometimes causes difficulty. Traditionally a negative sign is used in deference to Lenz's law, which states that any

265

current produced by the EMF tends to oppose the flux change. But this is inconsistent with the definition of inductance given in eqn (1.28) and the usual circuit conventions shown in figure 1.10. If we take the positive sign in eqn (1.33) and substitute for ψ from eqn (1.28), then

$$e = + L \frac{di}{dt} \tag{1.34}$$
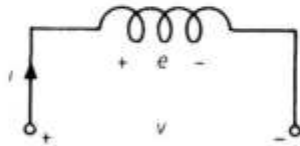
Figure 1.10 Induced EMF in a coil

For a pure inductance, with no internal resistance, Kirchhoff's voltage law applied to the circuit gives

$$v - e = 0$$

so that

$$v = e \tag{1.35}$$

Consequently the terminal voltage v is given by

$$v = + L \frac{di}{dt} \tag{1.36}$$

and this is the equation which defines the inductance element in circuit theory. The correct form of Faraday's law is therefore

$$e = + \frac{d\psi}{dt} \tag{1.37}$$

and Lenz's law may be used to resolve any uncertainty about the positive directions of e and ψ in the circuit.

Calculation of the induced EMF

Faraday's law relates the induced EMF to the rate of change of flux linkage, regardless of the way in which the change occurs. The flux linkage of a circuit may be changed in several ways: the strength of the magnetic field may be altered, either by moving the circuit relative to the source of the field or by varying the currents which create the field, or the boundary of the circuit may

be deformed while the magnetic field remains unchanged. The moving conductor in section 1.2 is an example of this last case. Consider the circuit formed by the voltmeter, the fixed rails, and the moving conductor; the area of this circuit increases steadily, and the rate of change of flux is equal to Blu. This result agrees with the previous calculation. In all cases the induced EMF may be calculated by the direct application of Faraday's law, and this is the only satisfactory method when motional and transformer effects are both present.

Particular care is needed when calculating the EMF in a moving conductor. It is tempting to use the 'flux cutting formula' $e = Blu$ in all cases, but this can give incorrect results when parts of the magnetic structure move with the conductor. The derivation of the formula given in section 1.2 is for the particular case of a conductor whose motion does not affect the source of the magnetic field in any way, and its direct application is limited to that situation. More complex problems can be treated by expressing the total magnetic field as the sum of components from different parts of the magnetic structure, and taking the sum of Blu terms with the appropriate values of u [4]. But the direct application of Faraday's law is the safest procedure in this kind of problem. Carter [3] gives a particularly good discussion of electromagnetic induction and some apparent paradoxes.

Induced EMF and inductance

The self-inductance and mutual-inductance coefficients can often be changed by relative movement of parts of the system, and Faraday's law gives the correct value for the induced EMF in these cases. For example, with the single coil shown in figure 1.10 the induced EMF is given by

$$e = \frac{d\psi}{dt} = \frac{d}{dt}(Li) = L\frac{di}{dt} + i\frac{dL}{dt} \qquad (1.38)$$

Thus if the motion of a part of the system causes L to change, there will be an EMF term additional to the normal EMF of self-induction. The voltage equation for the circuit should therefore be written as

$$v = Ri + \frac{d}{dt}(Li) \qquad (1.39)$$

Similarly, with the coupled coils shown in figure 1.11 the voltage equations are

$$v_1 = R_1 i_1 + \frac{d}{dt}(L_1 i_1 + M i_2)$$

$$v_2 = R_2 i_2 + \frac{d}{dt}(L_2 i_2 + M i_1) \qquad (1.40)$$
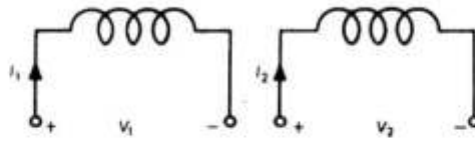
Figure 1.11 Coupled coils

When there is no motion of parts of the system the inductance coefficients are constant, and these equations reduce to the ordinary equations of coupled circuit theory.

1.4 Electromagnetic forces

In section 1.2 a method was given for calculating the force on a conductor in a magnetic field. In many practical devices, including rotating machines, magnetic forces act on the iron parts as well as on the conductors. These forces on the magnetised iron parts are often the dominant ones, and there is a need to calculate the total electromagnetic force acting on a structure made up of conductors and ferromagnetic materials. Two methods of calculation are given in this section. The first is the Maxwell stress method, which also provides a useful physical picture of the mechanism of force production. The second is an energy method, which complements the Maxwell stress method for purposes of calculations, but is less useful as a physical explanation.

The Maxwell stress concept

There is a sound scientific basis to the elementary idea that the magnetic lines of force are like rubber bands tending to draw pieces of iron together. The idea of lines or tubes of force was central to Faraday's conception of the magnetic field, but it was Maxwell who gave precise mathematical expression to this concept. Maxwell showed, as a deduction from the equations of the electromagnetic field, that magnetic forces could be considered to be transmitted through space (or a non-magnetic material) by the following system of stresses [3, 5]

(a) a tensile stress of magnitude $\frac{1}{2}BH$ newtons per square metre along the lines of force
(b) a compressive stress, also of magnitude $\frac{1}{2}BH$ newtons per square metre, at right angles to the lines of force.

Since $B = \mu_0 H$ in a non-magnetic medium, the stresses may also be written as $\frac{1}{2}\mu_0 H^2$ or $B^2/2\mu_0$.

268

If the magnetic field is perpendicular to the surface of a body (figure 1.12), there will be a tensile stress of magnitude $B^2/2\mu_0$, also perpendicular to the surface, drawing the surface into the field. If the field is parallel to the surface (figure 1.13), there will be a compressive stress of magnitude $B^2/2\mu_0$ pushing the surface out of the field. In the general case, when the flux density B makes an angle θ with the normal n, the stress t makes an angle 2θ with n (figure 1.14). The magnitude of t is still $B^2/2\mu_0$, and the three vectors n, B and t are coplanar.
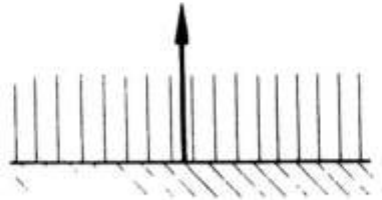
Figure 1.12   Magnetic force: field perpendicular to the surface

Figure 1.13   Magnetic force: field parallel to the surface
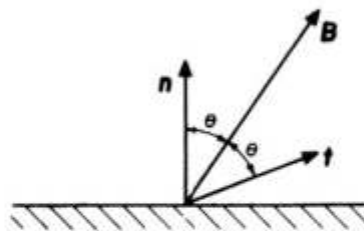
Figure 1.14   Maxwell stress vector

The force on an element of area δA is thus in the direction of t, and its magnitude is given by

$$\delta F = t\,\delta A = \frac{B^2}{2\mu_0}\,\delta A \quad \text{newtons} \qquad (1.41)$$

An interesting, and at first sight surprising, deduction is that the force will be parallel to the surface when the field is inclined at $45°$.

It is sometimes useful to relate the components of stress to the components of flux density, which may be done by resolving the vectors in directions normal and tangential to the surface. Thus if B_n and B_s are respectively the normal and tangential components of $\boldsymbol{B}$, the normal component of stress is given by

$$t_n = \frac{1}{2\mu_0}(B_n^2 - B_s^2) \qquad (1.42)$$

and the tangential component is

$$t_s = \frac{B_n B_s}{\mu_0} \qquad (1.43)$$

Although the Maxwell stress concept has been introduced in terms of magnetised iron parts, it is not restricted to this situation. The electromagnetic force acting on any combination of iron parts and conductors may be found from the Maxwell stress on a surface enclosing the bodies; the only restriction is that the surface should not pass through any magnetised parts. The concept gives an immediate qualitative picture of the way in which forces are distributed over the surface of an object in a magnetic field; it will be applied to DC machines in chapter 2 and AC machines in chapter 4.

Calculation of the force from the Maxwell stress

Equations (1.42) and (1.43) may be integrated over the surface to give the total force on an object; this presupposes that an accurate field solution is available, for example by numerical analysis [6]. When such a solution is not available, it is still possible to calculate the force approximately from the Maxwell stress. Two examples will illustrate the method.

Force of attraction

Figure 1.15 shows an electromagnet lifting an iron bar, and the problem is to calculate the force of attraction. The actual field pattern is quite complex; it would be very tedious first to solve the field equations with the boundary conditions imposed by the structure, and then to integrate the stress over the surface to find the force. A good approximation can be obtained from the following considerations. If the airgaps between the magnet poles and the bar are small, and the permeability of the magnetic material is high, the magnetic
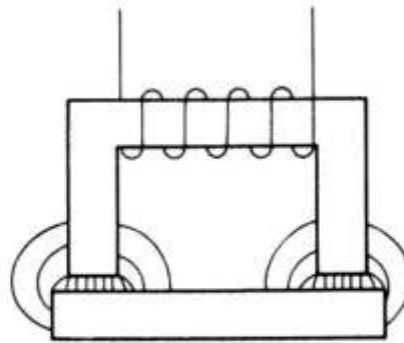
Figure 1.15   Electromagnet: actual field pattern

field in the gaps will be nearly uniform and much more intense than the fringing field outside the airgaps. (This will be justified formally in section 1.6.) Since the Maxwell stress varies as B^2, the field outside the airgaps will contribute very little to the total force, and it may be ignored. For the purpose of calculating the force, we may therefore replace the actual field distribution of figure 1.15 with the idealised distribution of figure 1.16, in which the field is uniform, confined to the airgaps and normal to the iron surfaces. If the area of each pole face is A square metres, and the magnetic flux density in each airgap is B teslas, the total force is

$$F = \frac{B^2}{2\mu_0} \cdot 2A = \frac{AB^2}{\mu_0} \text{ newtons} \tag{1.44}$$

A method of calculating the flux density from the coil current and the dimensions of the magnet will be given in section 1.6.
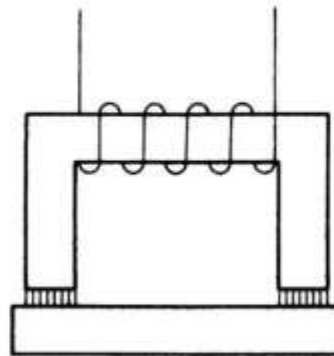
Figure 1.16   Electromagnet: idealised field pattern

[23]   Feyman, R., Leighton, R., Sands, M., *"The Feyman's Lecture on Physics: Mainly Electromagnetism and Matter,"* 1era ed., Wesley Publishing, Palo Alto, USA, 1964.

1

Electromagnetism

1-1 Electrical forces

Consider a force like gravitation which varies predominantly inversely as the square of the distance, but which is about a *billion-billion-billion-billion* times stronger. And with another difference. There are two kinds of "matter," which we can call positive and negative. Like kinds repel and unlike kinds attract—unlike gravity where there is only attraction. What would happen?

A bunch of positives would repel with an enormous force and spread out in all directions. A bunch of negatives would do the same. But an evenly mixed bunch of positives and negatives would do something completely different. The opposite pieces would be pulled together by the enormous attractions. The net result would be that the terrific forces would balance themselves out almost perfectly, by forming tight, fine mixtures of the positive and the negative, and between two separate bunches of such mixtures there would be practically no attraction or repulsion at all.

There is such a force: the electrical force. And all matter is a mixture of positive protons and negative electrons which are attracting and repelling with this great force. So perfect is the balance, however, that when you stand near someone else you don't feel any force at all. If there were even a little bit of unbalance you would know it. If you were standing at arm's length from someone and each of you had *one percent* more electrons than protons, the repelling force would be incredible. How great? Enough to lift the Empire State Building? No! To lift Mount Everest? No! The repulsion would be enough to lift a "weight" equal to that of the entire earth!

With such enormous forces so perfectly balanced in this intimate mixture, it is not hard to understand that matter, trying to keep its positive and negative charges in the finest balance, can have a great stiffness and strength. The Empire State Building, for example, swings only eight feet in the wind because the electrical forces hold every electron and proton more or less in its proper place. On the other hand, if we look at matter on a scale small enough that we see only a few atoms, any small piece will not, usually, have an equal number of positive and negative charges, and so there will be strong residual electrical forces. Even when there are equal numbers of both charges in two neighboring small pieces, there may still be large net electrical forces because the forces between individual charges vary inversely as the square of the distance. A net force can arise if a negative charge of one piece is closer to the positive than to the negative charges of the other piece. The attractive forces can then be larger than the repulsive ones and there can be a net attraction between two small pieces with no excess charges. The force that holds the atoms together, and the chemical forces that hold molecules together, are really electrical forces acting in regions where the balance of charge is not perfect, or where the distances are very small.

You know, of course, that atoms are made with positive protons in the nucleus and with electrons outside. You may ask: "If this electrical force is so terrific, why don't the protons and electrons just get on top of each other? If they want to be in an intimate mixture, why isn't it still more intimate?" The answer has to do with the quantum effects. If we try to confine our electrons in a region that is very close to the protons, then according to the uncertainty principle they must have some mean square momentum which is larger the more we try to confine them. It is this motion, required by the laws of quantum mechanics, that keeps the electrical attraction from bringing the charges any closer together.

1-1

Lower case Greek letters and commonly used capitals

α		alpha
β		beta
γ	Γ	gamma
δ	Δ	delta
ϵ		epsilon
ζ		zeta
η		eta
θ	Θ	theta
ι		iota
κ		kappa
λ	Λ	lambda
μ		mu
ν		nu
ξ	Ξ	xi (ksi)
o		omicron
π	Π	pi
ρ		rho
σ	Σ	sigma
τ		tau
υ	Υ	upsilon
ϕ	Φ	phi
χ		chi (khi)
ψ	Ψ	psi
ω	Ω	omega

There is another question: "What holds the nucleus together"? In a nucleus there are several protons, all of which are positive. Why don't they push themselves apart? It turns out that in nuclei there are, in addition to electrical forces, nonelectrical forces, called nuclear forces, which are greater than the electrical forces and which are able to hold the protons together in spite of the electrical repulsion. The nuclear forces, however, have a short range—their force falls off much more rapidly than $1/r^2$. And this has an important consequence. If a nucleus has too many protons in it, it gets too big, and it will not stay together. An example is uranium, with 92 protons. The nuclear forces act mainly between each proton (or neutron) and its nearest neighbor, while the electrical forces act over larger distances, giving a repulsion between each proton and all of the others in the nucleus. The more protons in a nucleus, the stronger is the electrical repulsion, until, as in the case of uranium, the balance is so delicate that the nucleus is almost ready to fly apart from the repulsive electrical force. If such a nucleus is just "tapped" lightly (as can be done by sending in a slow neutron), it breaks into two pieces, each with positive charge, and these pieces fly apart by electrical repulsion. The energy which is liberated is the energy of the atomic bomb. This energy is usually called "nuclear" energy, but it is really "electrical" energy released when electrical forces have overcome the attractive nuclear forces.

We may ask, finally, what holds a negatively charged electron together (since it has no nuclear forces). If an electron is all made of one kind of substance, each part should repel the other parts. Why, then, doesn't it fly apart? But does the electron have "parts"? Perhaps we should say that the electron is just a point and that electrical forces only act between *different* point charges, so that the electron does not act upon itself. Perhaps. All we can say is that the question of what holds the electron together has produced many difficulties in the attempts to form a complete theory of electromagnetism. The question has never been answered. We will entertain overselves by discussing this subject some more in later chapters.

As we have seen, we should expect that it is a combination of electrical forces and quantum-mechanical effects that will determine the detailed structure of materials in bulk, and, therefore, their properties. Some materials are hard, some are soft. Some are electrical "conductors"—because their electrons are free to move about; others are "insulators"—because their electrons are held tightly to individual atoms. We shall consider later how some of these properties come about, but that is a very complicated subject, so we will begin by looking at the electrical forces only in simple situations. We begin by treating only the laws of electricity—including magnetism, which is really a part of the same subject.

We have said that the electrical force, like a gravitational force, decreases inversely as the square of the distance between charges. This relationship is called Coulomb's law. But it is not precisely true when charges are moving—the electrical forces depend also on the motions of the charges in a complicated way. One part of the force between moving charges we call the *magnetic* force. It is really one aspect of an electrical effect. That is why we call the subject "electromagnetism."

There is an important general principle that makes it possible to treat electromagnetic forces in a relatively simple way. We find, from experiment, that the force that acts on a particular charge—no matter how many other charges there are or how they are moving—depends only on the position of that particular charge, on the velocity of the charge, and on the amount of charge. We can write the force F on a charge q moving with a velocity v as

$$F = q(E + v \times B). \tag{1.1}$$

We call E the *electric field* and B the *magnetic field* at the location of the charge. The important thing is that the electrical forces from all the other charges in the universe can be summarized by giving just these two vectors. Their values will depend on *where* the charge is, and may change with *time*. Furthermore, if we replace that charge with another charge, the force on the new charge will be just in proportion to the amount of charge so long as all the rest of the charges in the

1–2

world do not change their positions or motions. (In real situations, of course, each charge produces forces on all other charges in the neighborhood and may cause these other charges to move, and so in some cases the fields *can* change if we replace our particular charge by another.)

We know from Vol. I how to find the motion of a particle if we know the force on it. Equation (1.1) can be combined with the equation of motion to give

$$\frac{d}{dt}\left[\frac{mv}{(1 - v^2/c^2)^{1/2}}\right] = F = q(E + v \times B). \tag{1.2}$$

So if E and B are given, we can find the motions. Now we need to know how the E's and B's are produced.

One of the most important simplifying principles about the way the fields are produced is this: Suppose a number of charges moving in some manner would produce a field E_1, and another set of charges would produce E_2. If both sets of charges are in place at the same time (keeping the same locations and motions they had when considered separately), then the field produced is just the sum

$$E = E_1 + E_2. \tag{1.3}$$

This fact is called *the principle of superposition* of fields. It holds also for magnetic fields.

This principle means that if we know the law for the electric and magnetic fields produced by a *single* charge moving in an arbitrary way, then all the laws of electrodynamics are complete. If we want to know the force on charge A we need only calculate the E and B produced by each of the charges B, C, D, etc., and then add the E's and B's from all the charges to find the fields, and from them the forces acting on charge A. If it had only turned out that the field produced by a single charge was simple, this would be the neatest way to describe the laws of electrodynamics. We have already given a description of this law (Chapter 28, Vol. I) and it is, unfortunately, rather complicated.

It turns out that the form in which the laws of electrodynamics are simplest are not what you might expect. It is *not* simplest to give a formula for the force that one charge produces on another. It is true that when charges are standing still the Coulomb force law is simple, but when charges are moving about the relations are complicated by delays in time and by the effects of acceleration, among others. As a result, we do not wish to present electrodynamics only through the force laws between charges; we find it more convenient to consider another point of view—a point of view in which the laws of electrodynamics appear to be the most easily manageable.

1-2 Electric and magnetic fields

First, we must extend, somewhat, our ideas of the electric and magnetic vectors, E and B. We have defined them in terms of the forces that are felt by a charge. We wish now to speak of electric and magnetic fields *at a point* even when there is no charge present. We are saying, in effect, that since there are forces "acting on" the charge, there is still "something" there when the charge is removed. If a charge located at the point (x, y, z) at the time t feels the force F given by Eq. (1.1) we associate the vectors E and B with *the point* in space (x, y, z). We may think of $E(x, y, z, t)$ and $B(x, y, z, t)$ as giving the forces that *would be* experienced at the time t by a charge located at (x, y, z), *with the condition* that placing the charge there *did not disturb* the positions or motions of all the other charges responsible for the fields.

Following this idea, we associate with *every* point (x, y, z) in space two vectors E and B, which may be changing with time. The electric and magnetic fields are, then, viewed as *vector functions* of x, y, z, and t. Since a vector is specified by its components, each of the fields $E(x, y, z, t)$ and $B(x, y, z, t)$ represent three mathematical functions of x, y, z, and t.

1-3

274
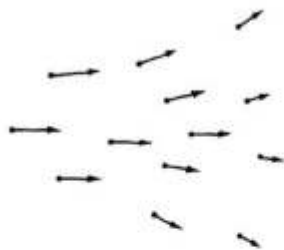
Fig. 1-1. A vector field may be represented by drawing a set of arrows whose magnitudes and directions indicate the values of the vector field at the points from which the arrows are drawn.
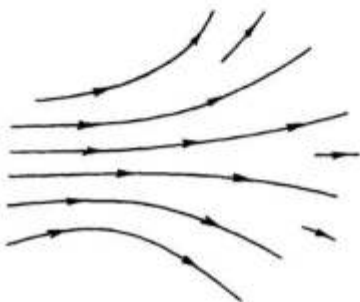
Fig. 1-2. A vector field can be represented by drawing lines which are tangent to the direction of the field vector at each point, and by drawing the density of lines proportional to the magnitude of the field vector.
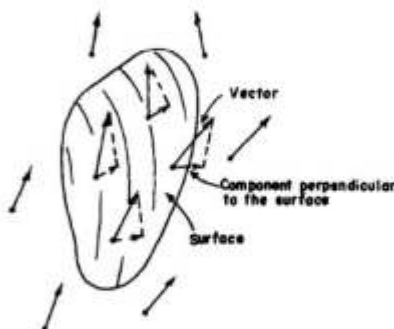
Fig. 1-3. The flux of a vector field through a surface is defined as the average value of the normal component of the vector times the area of the surface.

It is precisely because E (or B) can be specified at every point in space that it is called a "field." A "field" is any physical quantity which takes on different values at different points in space. Temperature, for example, is a field—in this case a scalar field, which we write as $T(x, y, z)$. The temperature could also vary in time, and we would say the temperature field is time-dependent, and write $T(x, y, z, t)$. Another example is the "velocity field" of a flowing liquid. We write $v(x, y, z, t)$ for the velocity of the liquid at each point in space at the time t. It is a vector field.

Returning to the electromagnetic fields—although they are produced by charges according to complicated formulas, they have the following important characteristic: the relationships between the values of the fields at *one point* and the values at a *nearby point* are very simple. With only a few such relationships in the form of differential equations we can describe the fields completely. It is in terms of such equations that the laws of electrodynamics are most simply written.

There have been various inventions to help the mind visualize the behavior of fields. The most correct is also the most abstract: we simply consider the fields as mathematical functions of position and time. We can also attempt to get a mental picture of the field by drawing vectors at many points in space, each of which gives the field strength and direction at that point. Such a representation is shown in Fig. 1-1. We can go further, however, and draw lines which are everywhere tangent to the vectors—which, so to speak, follow the arrows and keep track of the direction of the field. When we do this we lose track of the *lengths* of the vectors, but we can keep track of the strength of the field by drawing the lines far apart when the field is weak and close together when it is strong. We adopt the convention that the *number of lines per unit area* at right angles to the lines is proportional to the *field strength*. This is, of course, only an approximation, and it will require, in general, that new lines sometimes start up in order to keep the number up to the strength of the field. The field of Fig. 1-1 is represented by field lines in Fig. 1-2.

1-3 Characteristics of vector fields

There are two mathematically important properties of a vector field which we will use in our description of the laws of electricity from the field point of view. Suppose we imagine a closed surface of some kind and ask whether we are losing "something" from the inside; that is, does the field have a quality of "outflow"? For instance, for a velocity field we might ask whether the velocity is always outward on the surface or, more generally, whether more fluid flows out (per unit time) than comes in. We call the net amount of fluid going out through the surface per unit time the "flux of velocity" through the surface. The flow through an element of a surface is just equal to the component of the velocity perpendicular to the surface times the area of the surface. For an arbitrary closed surface, the *net outward flow*—or *flux*—is the average outward normal component of the velocity, times the area of the surface:

$$\text{Flux} = (\text{average normal component})\cdot(\text{surface area}). \qquad (1.4)$$

In the case of an electric field, we can mathematically define something analogous to an outflow, and we again call it the flux, but of course it is not the flow of any substance, because the electric field is not the velocity of anything. It turns out, however, that the mathematical quantity which is the average normal component of the field still has a useful significance. We speak, then, of the *electric flux*—also defined by Eq. (1.4). Finally, it is also useful to speak of the flux not only through a completely closed surface, but through any bounded surface. As before, the flux through such a surface is defined as the average normal component of a vector times the area of the surface. These ideas are illustrated in Fig. 1-3.

There is a second property of a vector field that has to do with a line, rather than a surface. Suppose again that we think of a velocity field that describes the flow of a liquid. We might ask this interesting question: Is the liquid circulating?

1-4

By that we mean: Is there a net rotational motion around some loop? Suppose that we instantaneously freeze the liquid everywhere except inside of a tube which is of uniform bore, and which goes in a loop that closes back on itself as in Fig. 1–4. Outside of the tube the liquid stops moving, but inside the tube it may keep on moving because of the momentum in the trapped liquid—that is, if there is more momentum heading one way around the tube than the other. We define a quantity called the *circulation* as the resulting speed of the liquid in the tube times its circumference. We can again extend our ideas and define the "circulation" for any vector field (even when there isn't anything moving). For any vector field the *circulation around any imagined closed curve* is defined as the average tangential component of the vector (in a consistent sense) multiplied by the circumference of the loop (Fig. 1–5).

$$\text{Circulation} = (\text{average tangential component}) \cdot (\text{distance around}). \quad (1.5)$$

You will see that this definition does indeed give a number which is proportional to the circulation velocity in the quickly frozen tube described above.

With just these two ideas—flux and circulation—we can describe all the laws of electricity and magnetism at once. You may not understand the significance of the laws right away, but they will give you some idea of the way the physics of electromagnetism will be ultimately described.

1–4 The laws of electromagnetism

The first law of electromagnetism describes the flux of the electric field:

$$\text{The flux of } E \text{ through any closed surface} = \frac{\text{the net charge inside}}{\epsilon_0}, \quad (1.6)$$

where ϵ_0 is a convenient constant. (The constant ϵ_0 is usually read as "epsilon-zero" or "epsilon-naught".) If there are no charges inside the surface, even though there are charges nearby outside the surface, the *average* normal component of E is zero, so there is no net flux through the surface. To show the power of this type of statement, we can show that Eq. (1.6) is the same as Coulomb's law, provided only that we also add the idea that the field from a single charge is spherically symmetric. For a point charge, we draw a sphere around the charge. Then the average normal component is just the value of the magnitude of E at any point, since the field must be directed radially and have the same strength for all points on the sphere. Our rule now says that the field at the surface of the sphere, times the area of the sphere—that is, the outgoing flux—is proportional to the charge inside. If we were to make the radius of the sphere bigger, the area would increase as the square of the radius. The average normal component of the electric field times that area must still be equal to the same charge inside, and so the field must decrease as the square of the distance—we get an "inverse square" field.

If we have an arbitrary curve in space and measure the circulation of the electric field around the curve, we will find that it is not, in general, zero (although it is for the Coulomb field). Rather, for electricity there is a second law that states: for any surface S (not closed) whose edge is the curve C,

$$\text{Circulation of } E \text{ around } C = \frac{d}{dt}(\text{flux of } B \text{ through } S). \quad (1.7)$$

We can complete the laws of the electromagnetic field by writing two corresponding equations for the magnetic field B.

$$\text{Flux of } B \text{ through any closed surface} = 0. \quad (1.8)$$

For a surface S bounded by the curve C,

$$c^2(\text{circulation of } B \text{ around } C) = \frac{d}{dt}(\text{flux of } E \text{ through } S)$$
$$+ \frac{\text{flux of electric current through } S}{\epsilon_0}. \quad (1.9)$$
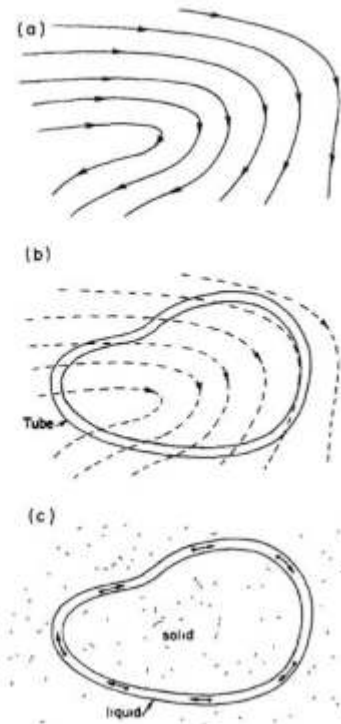
Fig. 1–4. (a) The velocity field in a liquid. Imagine a tube of uniform cross section that follows an arbitrary closed curve as in (b). If the liquid were suddenly frozen everywhere except inside the tube, the liquid in the tube would circulate as shown in (c).
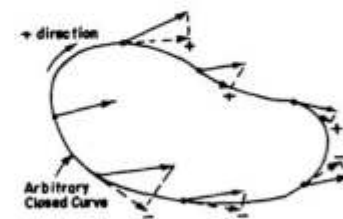
Fig. 1–5. The circulation of a vector field is the average tangential component of the vector (in a consistent sense) times the circumference of the loop.

1–5
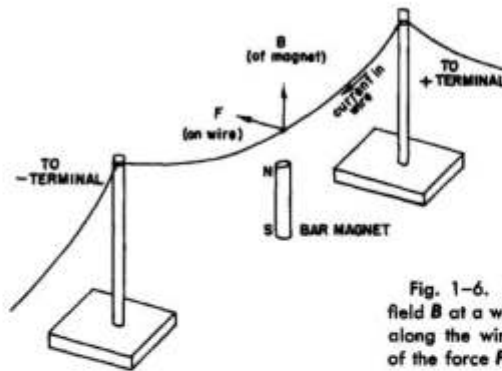
Fig. 1-6. A bar magnet gives a field B at a wire. When there is a current along the wire, the wire moves because of the force $F = qv \times B$.

The constant c^2 that appears in Eq. (1.9) is the square of the velocity of light. It appears because magnetism is in reality a relativistic effect of electricity. The constant ϵ_0 has been stuck in to make the units of electric current come out in a convenient way.

Equations (1.6) through (1.9), together with Eq. (1.1), are all the laws of electrodynamics*. As you remember, the laws of Newton were very simple to write down, but they had a lot of complicated consequences and it took us a long time to learn about them all. These laws are not nearly as simple to write down, which means that the consequences are going to be more elaborate and it will take us quite a lot of time to figure them all out.

We can illustrate some of the laws of electrodynamics by a series of small experiments which show qualitatively the interrelationships of electric and magnetic fields. You have experienced the first term of Eq. (1.1) when combing your hair, so we won't show that one. The second part of Eq. (1.1) can be demonstrated by passing a current through a wire which hangs above a bar magnet, as shown in Fig. 1-6. The wire will move when a current is turned on because of the force $F = qv \times B$. When a current exists, the charges inside the wire are moving, so they have a velocity v, and the magnetic field from the magnet exerts a force on them, which results in pushing the wire sideways.

When the wire is pushed to the left, we would expect that the magnet must feel a push to the right. (Otherwise we could put the whole thing on a wagon and have a propulsion system that didn't conserve momentum!) Although the force is too small to make movement of the bar magnet visible, a more sensitively supported magnet, like a compass needle, will show the movement.

How does the wire push on the magnet? The current in the wire produces a magnetic field of its own that exerts forces on the magnet. According to the last
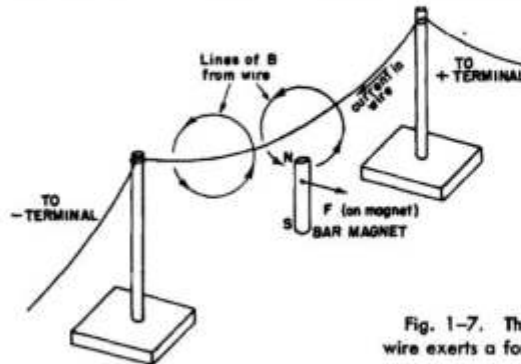
Fig. 1-7. The magnetic field of the wire exerts a force on the magnet.

* We need only to add a remark about some conventions for the *sign* of the circulation.

1-6
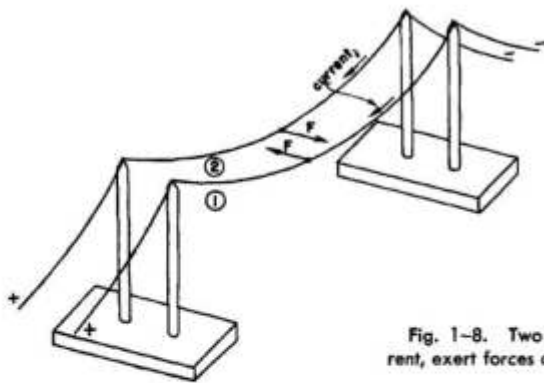
Fig. 1–8. Two wires, carrying current, exert forces on each other.

term in Eq. (1.9), a current must have a *circulation* of **B**—in this case, the lines of **B** are loops around the wire, as shown in Fig. 1–7. This **B**-field is responsible for the force on the magnet.

Equation (1.9) tells us that for a fixed current through the wire the circulation of **B** is the same for *any* curve that surrounds the wire. For curves—say circles—that are farther away from the wire, the circumference is larger, so the tangential component of **B** must decrease. You can see that we would, in fact, expect **B** to decrease linearly with the distance from a long straight wire.

Now, we have said that a current through a wire produces a magnetic field, and that when there is a magnetic field present there is a force on a wire carrying a current. Then we should also expect that if we make a magnetic field with a current in one wire, it should exert a force on another wire which also carries a current. This can be shown by using two hanging wires as shown in Fig. 1–8. When the currents are in the same direction, the two wires attract, but when the currents are opposite, they repel.

In short, electrical currents, as well as magnets, make magnetic fields. But wait, what is a magnet, anyway? If magnetic fields are produced by moving charges, is it not possible that the magnetic field from a piece of iron is really the result of currents? It appears to be so. We can replace the bar magnet of our experiment with a coil of wire, as shown in Fig. 1–9. When a current is passed through the coil—as well as through the straight wire above it—we observe a motion of the wire exactly as before, when we had a magnet instead of a coil. In other words, the current in the coil imitates a magnet. It appears, then, that a piece of iron acts as though it contains a perpetual circulating current. We can, in fact, understand magnets in terms of permanent currents in the atoms of the iron. The force on the magnet in Fig. 1–7 is due to the second term in Eq. (1.1).
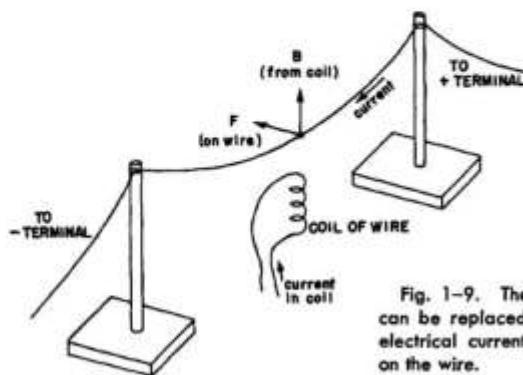
Fig. 1–9. The bar magnet of Fig. 1–6 can be replaced by a coil carrying an electrical current. A similar force acts on the wire.

1–7

Where do the currents come from? One possibility would be from the motion of the electrons in atomic orbits. Actually, that is not the case for iron, although it is for some materials. In addition to moving around in an atom, an electron also spins about on its own axis—something like the spin of the earth—and it is the current from this spin that gives the magnetic field in iron. (We say "something like the spin of the earth" because the question is so deep in quantum mechanics that the classical ideas do not really describe things too well.) In most substances, some electrons spin one way and some spin the other, so the magnetism cancels out, but in iron—for a mysterious reason which we will discuss later—many of the electrons are spinning with their axes lined up, and that is the source of the magnetism.

Since the fields of magnets are from currents, we do not have to add any extra term to Eqs. (1.8) or (1.9) to take care of magnets. We just take *all* currents, including the circulating currents of the spinning electrons, and then the law is right. You should also notice that Eq. (1.8) says that there are no magnetic "charges" analogous to the electrical charges appearing on the right side of Eq. (1.6). None has been found.

Fig. 1–10. The circulation of *B* around the curve C is given either by the current passing through the surface S_1, or by the rate of change of the flux of *E* through the surface S_2.
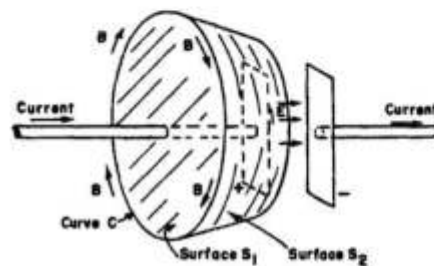
The first term on the right-hand side of Eq. (1.9) was discovered theoretically by Maxwell and is of great importance. It says that changing *electric* fields produce magnetic effects. In fact, without this term the equation would not make sense, because without it there could be no currents in circuits that are not complete loops. But such currents do exist, as we can see in the following example. Imagine a capacitor made of two flat plates. It is being charged by a current that flows toward one plate and away from the other, as shown in Fig. 1–10. We draw a curve C around one of the wires and fill it in with a surface which crosses the wire, as shown by the surface S_1 in the figure. According to Eq. (1.9), the circulation of *B* around C is given by the current in the wire (times c^2). But what if we fill in the curve with a *different* surface S_2, which is shaped like a bowl and passes between the plates of the capacitor, staying always away from the wire? There is certainly no current through this surface. But, surely, just changing the location of an imaginary surface is not going to change a real magnetic field! The circulation of *B* must be what it was before. The first term on the right-hand side of Eq. (1.9) does, indeed, combine with the second term to give the same result for the two surfaces S_1 and S_2. For S_2 the circulation of *B* is given in terms of the rate of change of the flux of *E* between the plates of the capacitor. And it works out that the changing *E* is related to the current in just the way required for Eq. (1.9) to be correct. Maxwell saw that it was needed, and he was the first to write the complete equation.

With the setup shown in Fig. 1–6 we can demonstrate another of the laws of electromagnetism. We disconnect the ends of the hanging wire from the battery and connect them to a galvanometer which tells us when there is a current through the wire. When we *push* the wire sideways through the magnetic field of the magnet, we observe a current. Such an effect is again just another consequence of Eq. (1.1)—the electrons in the wire feel the force $F = qv \times B$. The electrons have a sidewise velocity because they move with the wire. This v with a vertical *B* from the magnet results in a force on the electrons directed *along* the wire, which starts the electrons moving toward the galvanometer.

1–8

Suppose, however, that we leave the wire alone and move the magnet. We guess from relativity that it should make no difference, and indeed, we observe a similar current in the galvanometer. How does the magnetic field produce forces on charges at rest? According to Eq. (1.1) there must be an electric field. A moving magnet must make an electric field. How that happens is said quantitatively by Eq. (1.7). This equation describes many phenomena of great practical interest, such as those that occur in electric generators and transformers.

The most remarkable consequence of our equations is that the combination of Eq. (1.7) and Eq. (1.9) contains the explanation of the radiation of electromagnetic effects over large distances. The reason is roughly something like this: suppose that somewhere we have a magnetic field which is increasing because, say, a current is turned on suddenly in a wire. Then by Eq. (1.7) there must be a circulation of an electric field. As the electric field builds up to produce its circulation, then according to Eq. (1.9) a magnetic circulation will be generated. But the building up of *this* magnetic field will produce a new circulation of the electric field, and so on. In this way fields work their way through space without the need of charges or currents except at their source. That is the way we *see* each other! It is all in the equations of the electromagnetic fields.

1–5 What are the fields?

We now make a few remarks on our way of looking at this subject. You may be saying: "All this business of fluxes and circulations is pretty abstract. There are electric fields at every point in space; then there are these 'laws.' But what is *actually* happening? Why can't you explain it, for instance, by whatever it *is* that goes between the charges." Well, it depends on your prejudices. Many physicists used to say that direct action with nothing in between was inconceivable. (How could they find an idea inconceivable when it had already been conceived?) They would say: "Look, the only forces we know are the direct action of one piece of matter on another. It is impossible that there can be a force with nothing to transmit it." But what really happens when we study the "direct action" of one piece of matter right against another? We discover that it is not one piece right against the other; they are slightly separated, and there are electrical forces acting on a tiny scale. Thus we find that we are going to explain so-called direct-contact action in terms of the picture for electrical forces. It is certainly not sensible to try to insist that an electrical force has to look like the old, familiar, muscular push or pull, when it will turn out that the muscular pushes and pulls are going to be interpreted as electrical forces! The only sensible question is what is the *most convenient* way to look at electrical effects. Some people prefer to represent them as the interaction at a distance of charges, and to use a complicated law. Others love the field lines. They draw field lines all the time, and feel that writing E's and B's is too abstract. The field lines, however, are only a crude way of describing a field, and it is very difficult to give the correct, quantitative laws directly in terms of field lines. Also, the ideas of the field lines do not contain the deepest principle of electrodynamics, which is the superposition principle. Even though we know how the field lines look for one set of charges and what the field lines look like for another set of charges, we don't get any idea about what the field line patterns will look like when both sets are present together. From the mathematical standpoint, on the other hand, superposition is easy—we simply add the two vectors. The field lines have some advantage in giving a vivid picture, but they also have some disadvantages. The direct interaction way of thinking has great advantages when thinking of electrical charges at rest, but has great disadvantages when dealing with charges in rapid motion.

The best way is to use the abstract field idea. That it is abstract is unfortunate, but necessary. The attempts to try to represent the electric field as the motion of some kind of gear wheels, or in terms of lines, or of stresses in some kind of material have used up more effort of physicists than it would have taken simply to get the right answers about electrodynamics. It is interesting that the correct equations for the behavior of light in crystals were worked out by McCullough in 1843. But

people said to him: "Yes, but there is no real material whose mechanical properties could possibly satisfy those equations, and since light is an oscillation that must vibrate in *something*, we cannot believe this abstract equation business." If people had been more open-minded, they might have believed in the right equations for the behavior of light a lot earlier than they did.

In the case of the magnetic field we can make the following point: Suppose that you finally succeeded in making up a picture of the magnetic field in terms of some kind of lines or of gear wheels running through space. Then you try to explain what happens to two charges moving in space, both at the same speed and parallel to each other. Because they are moving, they will behave like two currents and will have a magnetic field associated with them (like the currents in the wires of Fig. 1–8). An observer who was riding along with the two charges, however, would see both charges as stationary, and would say that there is *no* magnetic field. The "gear wheels" or "lines" disappear when you ride along with the object! All we have done is to invent a *new* problem. How can the gear wheels disappear?! The people who draw field lines are in a similar difficulty. Not only is it not possible to say whether the field lines move or do not move with charges—they may disappear completely in certain coordinate frames

What we are saying, then, is that magnetism is really a relativistic effect. In the case of the two charges we just considered, travelling parallel to each other, we would expect to have to make relativistic corrections to their motion, with terms of order v^2/c^2. These corrections must correspond to the magnetic force. But what about the force between the two wires in our experiment (Fig. 1–8). There the magnetic force is the *whole* force. It didn't look like a "relativistic correction." Also, if we estimate the velocities of the electrons in the wire (you can do this yourself), we find that their average speed along the wire is about 0.01 centimeter per second. So v^2/c^2 is about 10^{-25}. Surely a negligible "correction." But no! Although the magnetic force is, in this case, 10^{-25} of the "normal" electrical force between the moving electrons, remember that the "normal" electrical forces have disappeared because of the almost perfect balancing out—because the wires have the same number of protons as electrons. The balance is much more precise than one part in 10^{25}, and the small relativistic term which we call the magnetic force is the only term left. It becomes the dominant term.

It is the near-perfect cancellation of electrical effects which allowed relativity effects (that is, magnetism) to be studied and the correct equations—to order v^2/c^2—to be discovered, even though physicists didn't *know* that's what was happening. And that is why, when relativity was discovered, the electromagnetic laws didn't need to be changed. They—unlike mechanics—were already correct to a precision of v^2/c^2.

1–6 Electromagnetism in science and technology

Let us end this chapter by pointing out that among the many phenomena studied by the Greeks there were two very strange ones: that if you rubbed a piece of amber you could lift up little pieces of papyrus, and that there was a strange rock from the island of Magnesia which attracted iron. It is amazing to think that these were the only phenomena known to the Greeks in which the effects of electricity or magnetism were apparent. The reason that these were the only phenomena that appeared is due primarily to the fantastic precision of the balancing of charges that we mentioned earlier. Study by scientists who came after the Greeks uncovered one new phenomena after another that were really some aspect of these amber and/or lodestone effects. Now we realize that the phenomena of chemical interaction and, ultimately, of life itself are to be understood in terms of electromagnetism.

At the same time that an understanding of the subject of electromagnetism was being developed, technical possibilities that defied the imagination of the people that came before were appearing: it became possible to signal by telegraph over long distances, and to talk to another person miles away without any connections between, and to run huge power systems—a great water wheel, connected by

1–10

281

filaments over hundreds of miles to another engine that turns in response to the master wheel—many thousands of branching filaments—ten thousand engines in ten thousand places running the machines of industries and homes—all turning because of the knowledge of the laws of electromagnetism.

Today we are applying even more subtle effects. The electrical forces, enormous as they are, can also be very tiny, and we can control them and use them in very many ways. So delicate are our instruments that we can tell what a man is doing by the way he affects the electrons in a thin metal rod hundreds of miles away. All we need to do is to use the rod as an antenna for a television receiver!

From a long view of the history of mankind—seen from, say, ten thousand years from now—there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics. The American Civil War will pale into provincial insignificance in comparison with this important scientific event of the same decade.

Magnetostatics

13–1 The magnetic field

The force on an electric charge depends not only on where it is, but also on how fast it is moving. Every point in space is characterized by two vector quantities which determine the force on any charge. First, there is the *electric force*, which gives a force component independent of the motion of the charge. We describe it by the electric field, *E*. Second, there is an additional force component, called the *magnetic force*, which depends on the velocity of the charge. This magnetic force has a strange directional character: At any particular point in space, both the *direction* of the force and its *magnitude* depend on the direction of motion of the particle: at every instant the force is always at right angles to the velocity vector; also, at any particular point, the force is always at right angles to a *fixed direction in space* (see Fig. 13–1); and finally, the magnitude of the force is proportional to the *component* of the velocity at right angles to this unique direction. It is possible to describe all of this behavior by defining the magnetic field vector *B*, which specifies both the unique direction in space and the constant of proportionality with the velocity, and to write the magnetic force as $qv \times B$. The total electromagnetic force on a charge can, then, be written as

$$F = q(E + v \times B). \tag{13.1}$$

This is called the *Lorentz force*.

The magnetic force is easily demonstrated by bringing a bar magnet close to a cathode-ray tube. The deflection of the electron beam shows that the presence of the magnet results in forces on the electrons transverse to their direction of motion, as we described in Chapter 12 of Vol. I.

The unit of magnetic field *B* is evidently one newton·second per coulomb·meter. The same unit is also one volt·second per meter². It is also called one *weber per square meter*.

13–2 Electric current; the conservation of charge

We consider first how we can understand the magnetic forces on wires carrying electric currents. In order to do this, we define what is meant by the current density. Electric currents are electrons or other charges in motion with a net drift or flow. We can represent the charge flow by a vector which gives the amount of charge passing per unit area and per unit time through a surface element at right angles to the flow (just as we did for the case of heat flow). We call this the *current density* and represent it by the vector *j*. It is directed along the motion of the charges. If we take a small area ΔS at a given place in the material, the amount of charge flowing across that area in a unit time is

$$j \cdot n \, \Delta S, \tag{13.2}$$

where *n* is the unit vector normal to ΔS.

The current density is related to the average flow velocity of the charges. Suppose that we have a distribution of charges whose average motion is a drift with the velocity *v*. As this distribution passes over a surface element ΔS, the charge Δq passing through the surface element in a time Δt is equal to the charge contained in a parallelepiped whose base is ΔS and whose height is $v \, \Delta t$, as shown in Fig. 13–2. The volume of the parallelepiped is the projection of ΔS at right angles to v times

13–1

Review: Chapter 15, Vol. 1: *The Special Theory of Relativity*
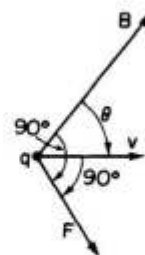
Fig. 13–1. The velocity-dependent component of the force on a moving charge is at right angles to **v** and to the direction of **B**. It is also proportional to the component of **v** at right angles to **B**, that is, to $v \sin \theta$.
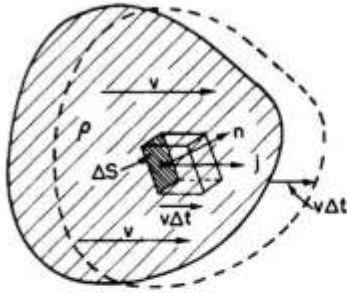
Fig. 13–2. If a charge distribution of density ρ moves with the velocity v, the charge per unit time through ΔS is $\rho v \cdot n \, \Delta S$.
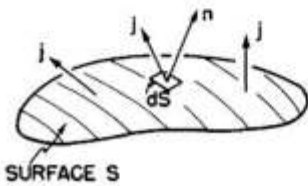
SURFACE S

Fig. 13–3. The current I through the surface S is $\int j \cdot n \, dS$.
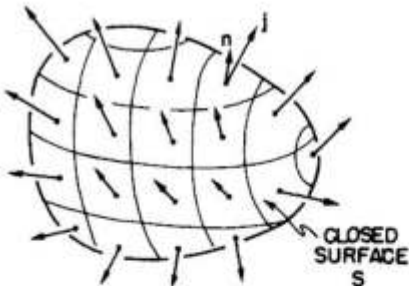
CLOSED SURFACE S

Fig. 13–4. The integral of $j \cdot n$ over a closed surface is the rate of change of the total charge Q inside.

$v \, \Delta t$, which when multiplied by the charge density ρ will give Δq. Thus

$$\Delta q = \rho v \cdot n \, \Delta S \, \Delta t.$$

The charge per unit time is then $\rho v \cdot n \, \Delta S$, from which we get

$$j = \rho v. \tag{13.3}$$

If the charge distribution consists of individual charges, say electrons, each with the charge q and moving with the mean velocity v, then the current density is

$$j = Nqv, \tag{13.4}$$

where N is the number of charges per unit volume.

The total charge passing per unit time through any surface S is called the *electric current*, I. It is equal to the integral of the normal component of the flow through all of the elements of the surface:

$$I = \int_S j \cdot n \, dS \tag{13.5}$$

(see Fig. 13–3).

The current I out of a closed surface S represents the rate at which charge leaves the volume V enclosed by S. One of the basic laws of physics is that *electric charge is indestructible;* it is never lost or created. Electric charges can move from place to place but never appear from nowhere. We say that *charge is conserved.* If there is a net current out of a closed surface, the amount of charge inside must decrease by the corresponding amount (Fig. 13–4). We can, therefore, write the law of the conservation of charge as

$$\int_{\substack{\text{any closed} \\ \text{surface}}} j \cdot n \, dS = -\frac{d}{dt}(Q_{\text{inside}}). \tag{13.6}$$

The charge inside can be written as a volume integral of the charge density:

$$Q_{\text{inside}} = \int_{\substack{V \\ \text{inside } S}} \rho \, dV. \tag{13.7}$$

If we apply (13.6) to a small volume ΔV, we know that the left-hand integral is $\nabla \cdot j \, \Delta V$. The charge inside is $\rho \, \Delta V$, so the conservation of charge can also be written as

$$\nabla \cdot j = -\frac{\partial \rho}{\partial t} \tag{13.8}$$

(Gauss' mathematics once again!).

13–3 The magnetic force on a current

Now we are ready to find the force on a current-carrying wire in a magnetic field. The current consists of charged particles moving with the velocity v along the wire. Each charge feels a transverse force

$$F = qv \times B$$

(Fig. 13–5a). If there are N such charges per unit volume, the number in a small volume ΔV of the wire is $N \, \Delta V$. The total magnetic force ΔF on the volume ΔV is the sum of the forces on the individual charges, that is,

$$\Delta F = (N \, \Delta V)(qv \times B).$$

But Nqv is just j, so

$$\Delta F = j \times B \, \Delta V \tag{13.9}$$

(Fig. 13–5b). The force per unit volume is $j \times B$.

13–2

284

If the current is uniform across a wire whose cross-sectional area is A, we may take as the volume element a cylinder with the base area A and the length ΔL. Then

$$\Delta F = j \times B A\, \Delta L. \tag{13.10}$$

Now we can call jA the vector current I in the wire. (Its magnitude is the electric current in the wire, and its direction is along the wire.) Then

$$\Delta F = I \times B\, \Delta L. \tag{13.11}$$

The force per unit length on a wire is $I \times B$.

This equation gives the important result that the magnetic force on a wire, due to the movement of charges in it, depends only on the total current, and not on the amount of charge carried by each particle—or even its sign! The magnetic force on a wire near a magnet is easily shown by observing its deflection when a current is turned on, as was described in Chapter 1 (see Fig. 1-6).

13-4 The magnetic field of steady currents; Ampere's law

We have seen that there is a force on a wire in the presence of a magnetic field, produced, say, by a magnet. From the principle that action equals reaction we might expect that there should be a force on the source of the magnetic field, i.e., on the magnet, when there is a current through the wire.* There are indeed such forces, as is seen by the deflection of a compass needle near a current-carrying wire. Now we know that magnets feel forces from other magnets, so that means that when there is a current in a wire, the wire itself generates a magnetic field. Moving charges, then, *produce* a magnetic field. We would like now to try to discover the laws that determine how such magnetic fields are created. The question is: Given a current, what magnetic field does it make? The answer to this question was determined experimentally by three critical experiments and a brilliant theoretical argument given by Ampere. We will pass over this interesting historical development and simply say that a large number of experiments have demonstrated the validity of Maxwell's equations. We take them as our starting point. If we drop the terms involving time derivatives in these equations we get the equations of *magnetostatics*:

$$\nabla \cdot B = 0 \tag{13.12}$$

and

$$c^2 \nabla \times B = \frac{j}{\epsilon_0}. \tag{13.13}$$

These equations are valid only if all electric charge densities are constant and all currents are steady, so that the electric and magnetic fields are not changing with time—all of the fields are "static."

We may remark that it is rather dangerous to think that there is such a thing as a static magnetic situation, because there must be currents in order to get a magnetic field at all—and currents can come only from moving charges. "Magnetostatics" is, therefore, an approximation. It refers to a special kind of dynamic situation with *large numbers* of charges in motion, which we can approximate by a *steady* flow of charge. Only then can we speak of a current density j which does not change with time. The subject should more accurately be called the study of steady currents. Assuming that all fields are steady, we drop all terms in $\partial E/\partial t$ and $\partial B/\partial t$ from the complete Maxwell equations, Eqs. (2.41), and obtain the two equations (13.12) and (13.13) above. Also notice that since the divergence of the curl of any vector is necessarily zero, Eq. (13.13) requires that $\nabla \cdot j = 0$. This is true, by Eq. (13.8), only if $\partial \rho/\partial t$ is zero. But that must be so if E is not changing with time, so our assumptions are consistent.

---

* We will see later, however, that such assumptions are *not* generally correct for electromagnetic forces!

(a)

(b)
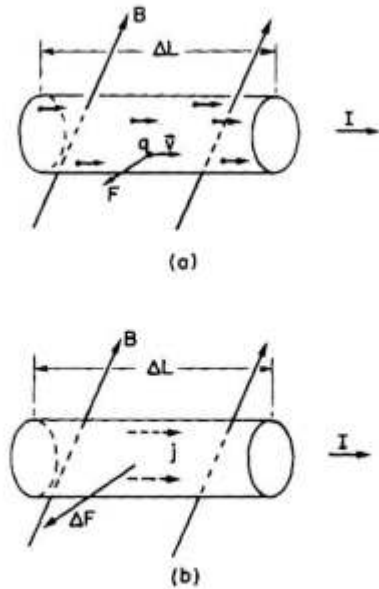
Fig. 13-5. The magnetic force on a current-carrying wire is the sum of the forces on the individual moving charges.

The requirement that $\nabla \cdot j = 0$ means that we may only have charges which flow in paths that close back on themselves. They may, for instance, flow in wires that form complete loops—called circuits. The circuits may, of course, contain generators or batteries that keep the charges flowing. But they may not include condensers which are charging or discharging. (We will, of course, extend the theory later to include dynamic fields, but we want first to take the simpler case of steady currents.)

Now let us look at Eqs. (13.12) and (13.13) to see what they mean. The first one says that the divergence of B is zero. Comparing it to the analogous equation in electrostatics, which says that $\nabla \cdot E = \rho/\epsilon_0$, we can conclude that there is no magnetic analog of an electric charge. There are *no magnetic charges* from which lines of B can emerge. If we think in terms of "lines" of the vector field B, they can never start and they never stop. Then where do they come from? Magnetic fields "appear" *in the presence of* currents; they have a *curl* proportional to the current density. Wherever there are currents, there are lines of magnetic field making loops around the currents. Since lines of B do not begin or end, they will often close back on themselves, making closed loops. But there can also be complicated situations in which the lines are not simple closed loops. But whatever they do, they never diverge from points. No magnetic charges have ever been discovered, so $\nabla \cdot B = 0$. This much is true not only for magnetostatics, it is *always* true—even for dynamic fields.

The connection between the B field and currents is contained in Eq. (13.13). Here we have a new kind of situation which is quite different from electrostatics, where we had $\nabla \times E = 0$. That equation meant that the line integral of E around any closed path is zero:

$$\oint_{\text{loop}} E \cdot ds = 0.$$

Fig. 13–6. The line integral of the tangential component of B is equal to the surface integral of the normal component of $\nabla \times B$.

We got that result from Stokes' theorem, which says that the integral around any closed path of *any* vector field is equal to the surface integral of the normal component of the curl of the vector (taken over any surface which has the closed loop as its periphery). Applying the same theorem to the magnetic field vector and using the symbols shown in Fig. 13–6, we get

$$\oint_\Gamma B \cdot ds = \int_S (\nabla \times B) \cdot n \, dS. \tag{13.14}$$

Taking the curl of B from Eq. (13.13), we have

$$\oint_\Gamma B \cdot ds = \frac{1}{\epsilon_0 c^2} \int_S j \cdot n \, dS. \tag{13.15}$$

The integral over j, according to (13.5), is the total current I through the surface S. Since for steady currents the current through S is independent of the shape of S, so long as it is bounded by the curve Γ, one usually speaks of "the current through the loop Γ." We have, then, a general law: the circulation of B around any closed curve is equal to the current I through the loop, divided by $\epsilon_0 c^2$:

$$\oint_\Gamma B \cdot ds = \frac{I_{\text{through }\Gamma}}{\epsilon_0 c^2}. \tag{13.16}$$

This law—called *Ampere's law*—plays the same role in magnetostatics that Gauss' law played in electrostatics. Ampere's law alone does not determine B from currents; we must, in general, also use $\nabla \cdot B = 0$. But, as we will see in the next section, it can be used to find the field in special circumstances which have certain simple symmetries.

13-4

13–5 The magnetic field of a straight wire and of a solenoid; atomic currents

We can illustrate the use of Ampere's law by finding the magnetic field near a wire. We ask: What is the field outside a long straight wire with a cylindrical cross section? We will assume something which may not be at all evident, but which is nevertheless true: that the field lines of B go around the wire in closed circles. If we make this assumption, then Ampere's law, Eq. (13.16), tells us how strong the field is. From the symmetry of the problem, B has the same magnitude at all points on a circle concentric with the wire (see Fig. 13–7). We can then do the line integral of $B \cdot ds$ quite easily; it is just the magnitude of B times the circumference. If r is the radius of the circle, then

$$\oint B \cdot ds = B \cdot 2\pi r.$$

The total current through the loop is merely the current I in the wire, so

$$B \cdot 2\pi r = \frac{I}{\epsilon_0 c^2},$$

or

$$B = \frac{1}{4\pi\epsilon_0 c^2} \frac{2I}{r}. \tag{13.17}$$

Fig. 13–7. The magnetic field outside of a long wire carrying the current I.

The strength of the magnetic field drops off inversely as r, the distance from the axis of the wire. We can, if we wish, write Eq. (13.17) in vector form. Remembering that B is at right angles both to I and to r, we have

$$B = \frac{1}{4\pi\epsilon_0 c^2} \frac{2I \times e_r}{r}. \tag{13.18}$$

We have separated out the factor $1/4\pi\epsilon_0 c^2$, because it appears often. It is worth remembering that it is exactly 10^{-7} (in the mks system), since an equation like (13.17) is used to *define* the unit of current, the ampere. At one meter from a current of one ampere the magnetic field is 2×10^{-7} webers per square meter.

Since a current produces a magnetic field, it will exert a force on a nearby wire which is also carrying a current. In Chapter 1 we described a simple demonstration of the forces between two current-carrying wires. If the wires are parallel, each is at right angles to the B field of the other; the wires should then be pushed either toward or away from each other. When currents are in the same direction, the wires attract; when the currents are moving in opposite directions, the wires repel.

Fig. 13–8. The magnetic field of a long solenoid.

Let's take another example that can be analyzed by Ampere's law if we add some knowledge about the field. Suppose we have a long coil of wire wound in a tight spiral, as shown by the cross sections in Fig. 13–8. Such a coil is called a *solenoid*. We observe experimentally that when a solenoid is very long compared with its diameter, the field outside is very small compared with the field inside. Using just that fact, together with Ampere's law, we can find the size of the field inside.

Since the field *stays* inside (and has zero divergence), its lines must go along parallel to the axis, as shown in Fig. 13–8. That being the case, we can use Ampere's law with the rectangular "curve" Γ shown in the figure. This loop goes the distance

13-5

287

L inside the solenoid, where the field is, say, B_0, then goes at right angles to the field, and returns along the outside, where the field is negligible. The line integral of B for this curve is just B_0L, and it must be $1/\epsilon_0 c^2$ times the total current through Γ, which is NI if there are N turns of the solenoid in the length L. We have

$$B_0 L = \frac{NI}{\epsilon_0 c^2}.$$

Or, letting n be the number of turns *per unit length* of the solenoid (that is, $n = N/L$), we get

$$B_0 = \frac{nI}{\epsilon_0 c^2}. \tag{13.19}$$
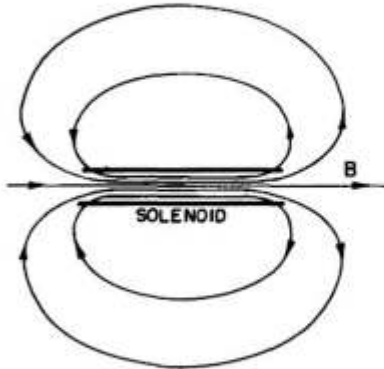
Fig. 13-9. The magnetic field outside of a solenoid.

What happens to the lines of B when they get to the end of the solenoid? Presumably, they spread out in some way and return to enter the solenoid at the other end, as sketched in Fig. 13-9. Such a field is just what is observed outside of a bar magnet. But what *is* a magnet anyway? Our equations say that B comes from the presence of currents. Yet we know that ordinary bars of iron (no batteries or generators) also produce magnetic fields. You might expect that there should be some other terms on the right-hand side of (13.12) or (13.13) to represent "the density of magnetic iron" or some such quantity. But there is no such term. Our theory says that the magnetic effects of iron come from some internal currents which are already taken care of by the j term.

Matter is very complex when looked at from a fundamental point of view—as we saw when we tried to understand dielectrics. In order not to interrupt our present discussion, we will wait until later to deal in detail with the interior mechanisms of magnetic materials like iron. You will have to accept, for the moment, that all magnetism is produced from currents, and that in a permanent magnet there are permanent internal currents. In the case of iron, these currents come from electrons spinning around their own axes. Every electron has such a spin, which corresponds to a tiny circulating current. Of course, one electron doesn't produce much magnetic field, but in an ordinary piece of matter there are billions and billions of electrons. Normally these spin and point every which way, so that there is no net effect. The miracle is that in a very few substances, like iron, a large fraction of the electrons spin with their axes in the same direction—for iron, two electrons from each atom takes part in this cooperative motion. In a bar magnet there are large numbers of electrons all spinning in the same direction and, as we will see, their total effect is equivalent to a current circulating on the surface of the bar. (This is quite analogous to what we found for dielectrics—that a uniformly polarized dielectric is equivalent to a distribution of charges on its surface.) It is, therefore, no accident that a bar magnet is equivalent to a solenoid.

13-6 The relativity of magnetic and electric fields

When we said that the magnetic force on a charge was proportional to its velocity, you may have wondered: "What velocity? With respect to which reference frame?" It is, in fact, clear from the definition of B given at the beginning of this chapter that what this vector is will depend on what we choose as a reference frame for our specification of the velocity of charges. But we have said nothing about which is the proper frame for specifying the magnetic field.

It turns out that *any* inertial frame will do. We will also see that magnetism and electricity are not independent things—that they should always be taken together as *one* complete electromagnetic field. Although in the static case Maxwell's equations separate into two distinct pairs, one pair for electricity and one pair for magnetism, with no apparent connection between the two fields, nevertheless, in nature itself there is a very intimate relationship between them that arises from the principle of relativity. Historically, the principle of relativity was discovered after Maxwell's equations. It was, in fact, the study of electricity and magnetism which led ultimately to Einstein's discovery of his principle of relativity. But let's see

13-6

288
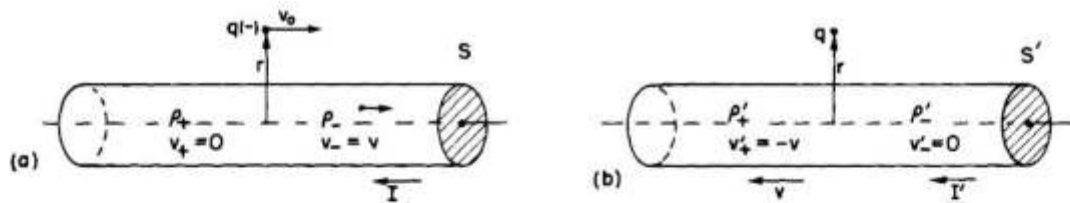
Fig. 13-10. The interaction of a current-carrying wire and a particle with the charge q as seen in two frames. In frame S (part a), the wire is at rest; in frame S' (part b), the charge is at rest.

what our knowledge of relativity would tell us about magnetic forces if we assume that the relativity principle is applicable—as it is—to electromagnetism.

Suppose we think about what happens when a negative charge moves with velocity v_0 parallel to a current-carrying wire, as in Fig. 13-10. We will try to understand what goes on in two reference frames: one fixed with respect to the wire, as in part (a) of the figure, and one fixed with respect to the particle, as in part (b). We will call the first frame S and the second S'.

In the S-frame, there is clearly a magnetic force on the particle. The force is directed toward the wire, so if the charge is moving freely we would see it curve in toward the wire. But in the S'-frame there can be no magnetic force on the particle, because its velocity is zero. Does it, therefore, stay where it is? Would we see different things happening in the two systems? The principle of relativity would say that in S' we should also see the particle move closer to the wire. We must try to understand why that would happen.

We return to our atomic description of a wire carrying a current. In a normal conductor, like copper, the electric currents come from the motion of some of the negative electrons—called the conduction electrons—while the positive nuclear charges and the remainder of the electrons stay fixed in the body of the material. We let the density of the conduction electrons be ρ_- and their velocity in S be v. The density of the charges at rest in S is ρ_+, which must be equal to the negative of ρ_-, since we are considering an uncharged wire. There is thus no electric field outside the wire, and the force on the moving particle is just

$$F = qv_0 \times B.$$

Using the result we found in Eq. (13.18) for the magnetic field at the distance r from the axis of a wire, we conclude that the force on the particle is directed toward the wire and has the magnitude

$$F = \frac{1}{4\pi\epsilon_0 c^2} \cdot \frac{2Iqv_0}{r}.$$

Using Eqs. (13.4) and (13.5), the current I can be written as $\rho_- v A$, where A is the area of a cross section of the wire. Then

$$F = \frac{1}{4\pi\epsilon_0 c^2} \cdot \frac{2q\rho_- A v v_0}{r}. \tag{13.20}$$

We could continue to treat the general case of arbitrary velocities for v and v_0, but it will be just as good to look at the special case in which the velocity v_0 of the particle is the same as the velocity v of the conduction electrons. So we write $v_0 = v$, and Eq. (13.20) becomes

$$F = \frac{q}{2\pi\epsilon_0} \frac{\rho_- A}{r} \frac{v^2}{c^2}. \tag{13.21}$$

Now we turn our attention to what happens in S', in which the particle is at rest and the wire is running past (toward the left in the figure) with the speed v. The positive charges moving with the wire will make some magnetic field B' at the particle. But the particle is now at *rest*, so there is no *magnetic* force on it! If there is any force on the particle, it must come from an electric field. It must

13-7

be that the moving wire has produced an electric field. But it can do that only if it appears *charged*—it must be that a neutral wire with a current appears to be charged when set in motion.

We must look into this. We must try to compute the charge density in the wire in S' from what we know about it in S. One might, at first, think they are the same; but we know that lengths are changed between S and S' (see Chapter 15, Vol. 1), so volumes will change also. Since the charge *densities* depend on the volume occupied by charges, the densities will change, too.

Before we can decide about the charge *densities* in S', we must know what happens to the electric *charge* of a bunch of electrons when the charges are moving. We know that the apparent mass of a particle changes by $1/\sqrt{1 - v^2/c^2}$. Does its charge do something similar? No! *Charges* are always the *same*, moving or not. Otherwise we would not always observe that the total charge is conserved.

Suppose that we take a block of material, say a conductor, which is initially uncharged. Now we heat it up. Because the electrons have a different mass than the protons, the velocities of the electrons and of the protons will change by different amounts. If the charge of a particle depended on the speed of the particle carrying it, in the heated block the charge of the electrons and protons would no longer balance. A block would become charged when heated. As we have seen earlier, a very small fractional change in the charge of all the electrons in a block would give rise to enormous electric fields. No such effect has ever been observed.

Also, we can point out that the mean speed of the electrons in matter depends on its chemical composition. If the charge on an electron changed with speed, the net charge in a piece of material would be changed in a chemical reaction. Again, a straightforward calculation shows that even a very small dependence of charge on speed would give enormous fields from the simplest chemical reactions. No such effect is observed, and we conclude that the electric charge of a single particle is independent of its state of motion.

So the charge q on a particle is an invariant scalar quantity, independent of the frame of reference. That means that in any frame the charge density of a distribution of electrons is just proportional to the number of electrons per unit volume. We need only worry about the fact that the volume *can* change because of the relativistic contraction of distances.

We now apply these ideas to our moving wire. If we take a length L_0 of the wire, in which there is a charge density ρ_0 of *stationary* charges, it will contain the total charge $Q = \rho_0 L_0 A_0$. If the same charges are observed in a different frame to be moving with velocity v, they will all be found in a piece of the material with the *shorter* length

$$L = L_0\sqrt{1 - v^2/c^2}, \qquad (13.22)$$

but with the same area A_0 (since dimensions transverse to the motion are unchanged). See Fig. 13–11.

If we call ρ the density of charges in the frame in which they are moving, the total charge Q will be $\rho L A_0$. This must also be equal to $\rho_0 L_0 A$, because charge is the same in any system, so that $\rho L = \rho_0 L_0$ or, from (13.22),

$$\rho = \frac{\rho_0}{\sqrt{1 - v^2/c^2}}. \qquad (13.23)$$
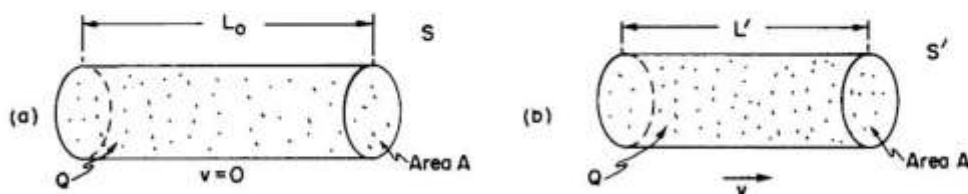
Fig. 13–11. If a distribution of charged particles at rest has the charge density ρ_0, the same charges will have the density $\rho = \rho_0/\sqrt{1 - v^2/c^2}$ when seen from a frame with the relative velocity v.

13–8

The charge *density* of a moving *distribution* of charges varies in the same way as the relativistic mass of a particle.

We now use this general result for the positive charge density ρ_+ of our wire. These charges are at rest in frame S. In S', however, where the wire moves with the speed v, the positive charge density becomes

$$\rho'_+ = \frac{\rho_+}{\sqrt{1 - v^2/c^2}}. \tag{13.24}$$

The *negative* charges are at rest in S'. So they have their "rest density" ρ_0 in this frame. In Eq. (13.23) $\rho_0 = \rho'_-$, because they have the density ρ'_- when the *wire* is at rest, i.e., in frame S, where the speed of the negative charges is v. For the conduction electrons, we then have that

$$\rho_- = \frac{\rho'_-}{\sqrt{1 - v^2/c^2}}, \tag{13.25}$$

or

$$\rho'_- = \rho_- \sqrt{1 - v^2/c^2}. \tag{13.26}$$

Now we can see why there are electric fields in S'—because in this frame the wire has the net charge density ρ' given by

$$\rho' = \rho'_+ + \rho'_-.$$

Using (13.24) and (13.26), we have

$$\rho' = \frac{\rho_+}{\sqrt{1 - v^2/c^2}} + \rho_- \sqrt{1 - v^2/c^2}.$$

Since the stationary wire is neutral, $\rho_- = -\rho_+$, and we have

$$\rho' = \rho_+ \frac{v^2/c^2}{\sqrt{1 - v^2/c^2}}. \tag{13.27}$$

Our moving wire is positively charged and will produce an electric field E' at the external stationary particle. We have already solved the electrostatic problem of a uniformly charged cylinder. The electric field at the distance r from the axis of the cylinder is

$$E' = \frac{\rho' A}{2\pi\epsilon_0 r} = \frac{\rho_+ A v^2/c^2}{2\pi\epsilon_0 r \sqrt{1 - v^2/c^2}}. \tag{13.28}$$

The force on the negatively charged particle is toward the wire. We have, at least, a force in the same direction from the two points of view; the electric force in S' has the same direction as the magnetic force in S.

The magnitude of the force in S' is

$$F' = \frac{q}{2\pi\epsilon_0} \frac{\rho_+ A}{r} \frac{v^2/c^2}{\sqrt{1 - v^2/c^2}}. \tag{13.29}$$

Comparing this result for F' with our result for F in Eq. (13.21), we see that the magnitudes of the forces are almost identical from the two points of view. In fact,

$$F' = \frac{F}{\sqrt{1 - v^2/c^2}}, \tag{13.30}$$

so for the small velocities we have been considering, the two forces are equal. We can say that for low velocities, at least, we understand that magnetism and electricity are just "two ways of looking at the same thing."

But things are even better than that. If we take into account the fact that *forces* also transform when we go from one system to the other, we find that the two ways of looking at what happens do indeed give the same *physical* result for any velocity.

13-9

One way of seeing this is to ask a question like: What transverse momentum will the particle have after the force has acted for a little while? We know from Chapter 16 of Vol. I that the transverse momentum of a particle should be the same in both the S- and S'-frames. Calling the transverse coordinate y, we want to compare Δp_y and $\Delta p'_y$. Using the relativistically correct equation of motion, $F = dp/dt$, we expect that after the time Δt our particle will have a transverse momentum Δp_y in the S-system given by

$$\Delta p_y = F\,\Delta t. \tag{13.31}$$

In the S'-system, the transverse momentum will be

$$\Delta p'_y = F'\,\Delta t'. \tag{13.32}$$

We must, of course, compare Δp_y and $\Delta p'_y$ for corresponding time intervals Δt and $\Delta t'$. We have seen in Chapter 15 of Vol. I that the time intervals referred to a *moving* particle appear to be *longer* than those in the rest system of the particle. Since our particle is initially at rest in S', we expect, for small Δt, that

$$\Delta t = \frac{\Delta t'}{\sqrt{1 - v^2/c^2}}, \tag{13.33}$$

and everything comes out O.K. From (13.31) and (13.32),

$$\frac{\Delta p'_y}{\Delta p_y} = \frac{F'\,\Delta t'}{F\,\Delta t},$$

which is just $= 1$ if we combine (13.30) and (13.33).

We have found that we get the same physical result whether we analyze the motion of a particle moving along a wire in a coordinate system at rest with respect to the wire, or in a system at rest with respect to the particle. In the first instance, the force was purely "magnetic," in the second, it was purely "electric." The two points of view are illustrated in Fig. 13–12 (although there is still a magnetic field B' in the second frame, it produces no forces on the stationary particle).

If we had chosen still another coordinate system, we would have found a different mixture of E and B fields. Electric and magnetic forces are part of *one* physical phenomenon—the electromagnetic interactions of particles. The separation of this interaction into electric and magnetic parts depends very much on the reference frame chosen for the description. But a complete electromagnetic description is invariant; electricity and magnetism taken together are consistent with Einstein's relativity.

Since electric and magnetic fields appear in different mixtures if we change our frame of reference, we must be careful about how we look at the fields E and B. For instance, if we think of "lines" of E or B, we must not attach too much reality to them. The lines may disappear if we try to observe them from a different coordinate system. For example, in system S' there are electric field lines, which we do *not* find "moving past us with velocity v in system S." In system S there are no electric field lines at all! Therefore it makes no sense to say something like: When I move a magnet, it takes its field with it, so the lines of B are also moved. There is no way to make sense, in general, out of the idea of "the speed of a moving field line." The fields are our way of describing what goes on at a point in space. In particular, E and B tell us about the forces that will act on a moving particle. The question "What is the force on a charge from a *moving* magnetic field?" doesn't mean anything precise. The force is given by the values of E and B at the charge, and the formula (13.1) is not to be altered if the *source* of E or B is moving (it is the values of E and B that will be altered by the motion). Our mathematical description deals only with the fields as a function of x, y, z, and t *with respect to some inertial frame.*

We will later be speaking of "a *wave* of electric and magnetic fields travelling through space," as, for instance, a light wave. But that is like speaking of a *wave* travelling on a string. We don't then mean that some part of the *string* is moving
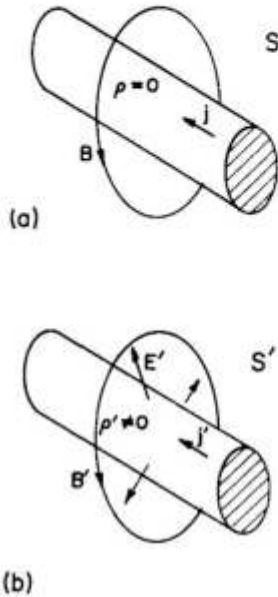
(a)

(b)

Fig. 13–12. In frame S the charge density is zero and the current density is j. There is only a magnetic field. In S', there is a charge density ρ', and a different current density j'. The magnetic field B' is different and there is an electric field E'.

in the direction of the wave, we mean that the *displacement* of the string appears first at one place and later at another. Similarly, in an electromagnetic wave, the *wave* travels, but the magnitude of the fields *change*. So in the future when we—or someone else—speaks of a "moving" field, you should think of it as just a handy, short way of describing a changing field in some circumstances.

13-7 The transformation of currents and charges

You may have worried about the simplification we made above when we took the same velocity v for the particle and for the conduction electrons in the wire. We could go back and carry through the analysis again for two different velocities, but it is easier to simply notice that charge and current density are the components of a four-vector (see Chapter 17, Vol. I).

We have seen that if ρ_0 is the density of the charges in their rest frame, then in a frame in which they have the velocity v, the density is

$$\rho = \frac{\rho_0}{\sqrt{1 - v^2/c^2}}.$$

In that frame their current density is

$$j = \rho v = \frac{\rho_0 v}{\sqrt{1 - v^2/c^2}}. \tag{13.34}$$

Now we know that the energy U and momentum p of a particle moving with velocity v are given by

$$U = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}}, \qquad p = \frac{m_0 v}{\sqrt{1 - v^2/c^2}},$$

where m_0 is its rest mass. We also know that U and p form a relativistic four-vector. Since ρ and j depend on the velocity v exactly as do U and p, we can conclude that ρ and j are *also* the components of a relativistic four-vector. This property is the key to a general analysis of the field of a wire moving with any velocity, which we would need if we want to do the problem again with the velocity v_0 of the particle different from the velocity of the conduction electrons.

If we wish to transform ρ and j to a coordinate system moving with a velocity u in the x-direction, we know that they transform just like t and (x, y, z), so that we have (see Chapter 15, Vol. I)

$$x' = \frac{x - ut}{\sqrt{1 - u^2/c^2}}, \qquad j'_x = \frac{j_x - u\rho}{\sqrt{1 - u^2/c^2}},$$

$$y' = y,\qquad\qquad j'_y = j_y,$$

$$z' = z,\qquad\qquad j'_z = j_z,$$

$$t' = \frac{t - ux/c^2}{\sqrt{1 - u^2/c^2}}, \qquad \rho' = \frac{\rho - uj_x/c^2}{\sqrt{1 - u^2/c^2}}. \tag{13.35}$$

With these equations we can relate charges and currents in one frame to those in another. Taking the charges and currents in either frame, we can solve the electromagnetic problem in that frame by using our Maxwell equations. The result we obtain *for the motions of particles* will be the same no matter which frame we choose. We will return at a later time to the relativistic transformations of the electromagnetic fields.

13-8 Superposition; the right-hand rule

We will conclude this chapter by making two further points regarding the subject of magnetostatics. First, our basic equations for the magnetic field,

$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \times \mathbf{B} = j/c^2\epsilon_0,$$

13-11

are linear in B and j. That means that the principle of superposition also applies to magnetic fields. The field produced by two different steady currents is the sum of the individual fields from each current acting alone. Our second remark concerns the right-hand rules which we have encountered (such as the right-hand rule for the magnetic field produced by a current). We have also observed that the magnetization of an iron magnet is to be understood from the spin of the electrons in the material. The direction of the magnetic field of a spinning electron is related to its spin axis by the same right-hand rule. Because B is determined by a "handed" rule—involving either a cross product or a curl—it is called an *axial* vector. (Vectors whose direction in space does not depend on a reference to a right or left hand are called *polar* vectors. Displacement, velocity, force, and E, for example, are polar vectors.)

Physically observable quantities in electromagnetism are *not*, however, right-(or left-) handed. Electromagnetic interactions are symmetrical under reflection (see Chapter 52, Vol. I). Whenever magnetic forces between two sets of currents are computed, the result is invariant with respect to a change in the hand convention. Our equations lead, independently of the right-hand convention, to the end result that parallel currents attract, or that currents in opposite directions repel. (Try working out the force using "left-hand rules.") An attraction or repulsion is a polar vector. This happens because in describing any complete interaction, we use the right-hand rule twice—once to find B from currents, again to find the force this B produces on a second current. Using the right-hand rule twice is the same as using the left-hand rule twice. If we were to change our conventions to a left-hand system all our B fields would be reversed, but all forces—or, what is perhaps more relevant, the observed accelerations of objects—would be unchanged.

Although physicists have recently found to their surprise that *all* the laws of nature are not always invariant for mirror reflections, the laws of electromagnetism do have such a basic symmetry.

13–12

294

[24]   Ida, N., *"Engineering Electromagnetics,"* 1era ed., Springer Science & Business Media, New York, USA, 2000.

C H A P T E R 10

Faraday's Law and Induction

Whether Ampere's beautiful theory were adopted, or any other, or whatever reservation were mentally made, still it appeared very extraordinary, that as every electric current was accompanied by a corresponding intensity of magnetic action at right angles to the current, good conductors of electricity, when placed within the sphere of this action, should not have any current induced through them, or some sensible effect produced equivalent in force to such a current.

—Michael Faraday; article 3 in the first series of "Experimental Researches in Electricity", Nov. 24, 1831.

10.1   INTRODUCTION

In the previous chapters, we found it useful to treat the electric and magnetic phenomena separately. In **Chapters 4** through **7**, we treated electrostatic fields by relying on the two postulates:

$$\left(\nabla \times \mathbf{E} = 0 \quad \text{and} \quad \nabla \cdot \mathbf{D} = \rho\right) \quad \text{or} \quad \left(\oint_C \mathbf{E} \cdot d\mathbf{l} = 0 \quad \text{and} \quad \oint_s \mathbf{D} \cdot d\mathbf{s} = Q\right)$$

In **Chapters 8** and **9**, some of the basic magnetic phenomena were introduced, now relying on the following two postulates:

$$\left(\nabla \times \mathbf{H} = \mathbf{J} \quad \text{and} \quad \nabla \cdot \mathbf{B} = 0\right) \quad \text{or} \quad \left(\oint_C \mathbf{H} \cdot d\mathbf{l} = I \quad \text{and} \quad \oint_s \mathbf{B} \cdot d\mathbf{s} = 0\right)$$

These postulates allowed us to treat a large number of applications and gain insight into the behavior of the electric and magnetic fields governed by the postulates.

Two aspects of this approach should be apparent by now:

(1) The discussion of electrostatic fields was independent of that for magnetostatic fields. Even though we saw in **Chapter 7** that a static field can cause a flow

629

of charge in a conductor and, therefore, a current, the electric field does not depend on this current and any of its consequences; that is, the static electric field is uniquely defined from the charge distribution in the system. The current thus generated does, in turn, generate a static magnetic field as defined by Ampere's law. However, this magnetic field, while it may coexist with the electric field, does not affect it.

(2) The discussion was limited to static applications. We have not specifically stated that the postulates of the magnetostatic field are not valid under time-varying conditions, but because the magnetostatic field relied on steady electric currents, time-dependent phenomena could not be included.

The question is: What happens if the fields are time varying? Or, perhaps, a better statement would be: How do we need to modify the fundamental postulates, to treat time-varying electric and magnetic fields?

This chapter discusses the question of time dependency in some detail. We will see that the electric and magnetic fields under time-varying conditions are interdependent. A time-varying electric field generates a time-varying magnetic field and vice versa. Thus, the time-dependent magnetic flux density $\mathbf{B}$ (and therefore $\mathbf{H}$) is dependent on the electric field intensity $\mathbf{E}$ (and therefore on $\mathbf{D}$). The above postulates will be modified to account for this dependency. Since now we must treat both the electric and magnetic fields as coupled vectors, the postulates for both fields must be included. A total of four relations are required to specify the electromagnetic field under time-dependent conditions: the curl of the electric field intensity $\mathbf{E}$, the divergence of the electric flux density $\mathbf{D}$, the curl of the magnetic field intensity $\mathbf{H}$, and the divergence of the magnetic flux density $\mathbf{B}$.

Because of the dependency between the electric and magnetic fields, we should use the term electromagnetic field when dealing with time-dependent fields. This term indicates that the two fields cannot be treated separately. The remaining part of this books deals with the electromagnetic field. In some cases, we will find it easier or more useful to treat the electric field or the magnetic field alone, but whichever field we choose to emphasize, it should be remembered that the other field can always be derived if necessary and, more importantly, it always exists.

After introducing Faraday's law as the first basic law governing time-dependent fields, we discuss some applications of Faraday's law. These include traditional power devices like transformers, motors, and generators as well as more recently developed devices like linear motors and levitating mechanisms. Other applications are acceleration of particles, electromagnetic testing of materials, electromagnetic ore prospecting, heating and melting of materials, and magnetic braking, to name a few.

The main difference at this stage, between the time-dependent and time-independent fields, is in the process of induction: A time-dependent field produces an electric field and that, in turn, induces currents in conducting materials. This induction is fundamental to the operation of many very important devices, including transformers and generators. There are two mechanisms of induction: One is due to change in the magnetic flux; the other due to motion in the magnetic field. Both are discussed since they can exist simultaneously in the same device and because both are important in design.

10.2 FARADAY'S LAW

The coupling between the electric and magnetic fields is based on an experimental relation known as Faraday's law. This law was formulated in 1831 by Michael Faraday after a series of experiments. Faraday[1] observed that if he moved a closed loop in the magnetic field of a magnet or if he moved the magnet while the loop remained stationary, a current flowed in the loop. This current was not due to external sources but rather, was induced in the loop by the change in the magnetic flux. He also found that the current was proportional to the rate of change of flux.

The current is due to an induced voltage in the loop. This voltage is called an *electromotive force* or *emf*. The electromotive force produced in this experiment can be written as

$$\text{emf} = -\frac{d\Phi}{dt} \quad [V] \tag{10.1}$$

where Φ is the flux through the loop. The physical situation is shown in Figure 10.1. In general, there may be more than one loop in the same location;
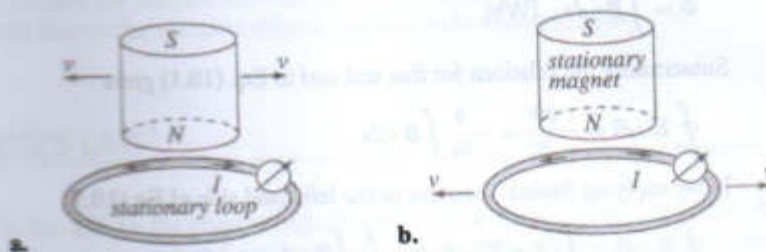
FIGURE 10.1 Faraday's law: (a) Movement of a permanent magnet in the presence of a loop generates an electromotive force and a current in the loop. (b) Movement of the loop in the presence of the stationary magnet generates an identical current.

[1]Michael Faraday (1791–1867) had little schooling, mostly as self-education. At age 22, he became assistant to Sir Humphry Davey (a well-known chemist and member of the Royal Institution) who helped his education. In 1821, Faraday demonstrated the operation of an electric motor. In 1825 he became director of the Royal Institution Laboratory and, there, in 1831 he discovered what is now known as Faraday's law. His experiment showed that the motion of a magnet near a wire loop produces an electromotive force, and therefore a current, in the closed loop. Faraday experimented in other areas as well. These included dielectrics, electrolysis, and polarization of light. The farad unit of capacitance is named after him in honor of his many and varied achievements. Faraday worked long and hard and is considered the ultimate experimentalist. His work was not all fun. In 1862, he became very ill from inhaling mercury vapors (he was using mercury for contacts) and it took him almost 2 years to recover. Faraday, more than anyone else, laid the foundation for the electromagnetic theory later developed by James Clerk Maxwell into what is today known as Maxwell's theory, and this is summarized by Maxwell's equations. Faraday's law is one of the four equations. Later, when Maxwell unified the electromagnetic theory, he gave credit to many, but mostly to Faraday.
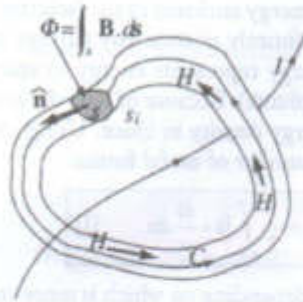
FIGURE 9.34   An arbitrary volume in the form of a general ring around the current I.

Consider a simple current-carrying conductor such as the wire shown in **Figure 9.34**. The conductor can be thin or thick. We wish to calculate the energy associated with an element of volume due to a conductor. Eq. (9.72) is now

$$W_m = \frac{1}{2}\sum_{i=1}^{1}\Phi_i I_i = \frac{1}{2}\Phi_1 I_1 \quad \text{[J]} \tag{9.85}$$

where we assumed $N = 1$. This does not diminish from the generality of the derivation. If there are multiple currents, the energy due to each current can be calculated separately. The flux generated by the current can now be calculated by first using Ampere's law to calculate the magnetic field intensity: The surface s over which the flux is evaluated is arbitrary and can be as small as necessary:

$$I = \oint_C \mathbf{H}\cdot d\mathbf{l} \quad \text{for single loops} \quad \text{or} \quad NI = \oint_C \mathbf{H}\cdot d\mathbf{l} \quad \text{for multiple loops} \tag{9.86}$$

The contour C is again arbitrary. However, we will choose a convenient contour such that $\mathbf{H}$ and $d\mathbf{l}$ are in the same direction to simplify the evaluation of the scalar product. Referring to **Figure 9.34**, we created a volume, in the form of a general tube, with cross section s in which the flux density is constant so that the flux Φ is independent of the location in the cross section along the tube. With these considerations in mind, we can write

$$W_m = \frac{1}{2}\int_s \mathbf{B}\cdot\hat{\mathbf{n}}\,ds \oint_C \mathbf{H}\cdot d\mathbf{l} = \frac{1}{2}\oint_C \left(\int_s \mathbf{B}\cdot\hat{\mathbf{n}}\,ds\right)\mathbf{H}\cdot d\mathbf{l} \quad \text{[J]} \tag{9.87}$$

Note that the surface integral representing flux was inserted into the line integral because the flux is assumed to be constant along the contour C. Since $\mathbf{B}\cdot\hat{\mathbf{n}}$ is the scalar component of $\mathbf{B}$ in the direction normal to s (see **Figure 9.34**), we can write

$$\oint_C \left(\int_s \mathbf{B}\cdot\hat{\mathbf{n}}\,ds\right)\mathbf{H}\cdot d\mathbf{l} = \oint_C \left(\int_s \mathbf{B}\cdot\mathbf{H}\,ds\right)dl = \int_v \mathbf{B}\cdot\mathbf{H}\,dv \quad \text{[J]} \tag{9.88}$$

Substituting this back into Eq. (9.87) gives

$$\boxed{W_m = \frac{1}{2}\int_v \mathbf{B}\cdot\mathbf{H}\,dv \quad \text{[J]}} \tag{9.89}$$

This magnetic energy represents the total energy enclosed in the (arbitrary) volume of the tube in **Figure 9.34**. This tube is entirely outside any current or current density distribution and, therefore, its energy represents energy in space (or any other material that may be present in the volume). Because of this, the integrand of Eq. (9.89) may be viewed as a volume energy density in space. Using the relation $B = \mu H$, we can rewrite this relation in a number of useful forms:

$$W_m = \frac{1}{2}\int_v \mathbf{B}\cdot\mathbf{H}\,dv = \frac{1}{2}\int_v \mu\mathbf{H}\cdot\mathbf{H}\,dv = \frac{1}{2}\int_v \mathbf{B}\cdot\frac{\mathbf{B}}{\mu}\,dv \quad [J] \tag{9.90}$$

Any of these relations is appropriate to use depending on which is more convenient. Note, also, that the integration may be done over any volume, including the whole of space. The result is always the energy stored in the chosen volume. If the energy is over the whole space, it represents the total energy in the system. If permeability μ is a constant, then **B** and **H** are always in the same direction and we can write the scalar products as $\mathbf{B}\cdot\mathbf{H} = BH$, $\mathbf{H}\cdot\mathbf{H} = H^2$, and $\mathbf{B}\cdot\mathbf{B} = B^2$. Viewing the integrand as an energy density, w_m, we can write

$$w_m = \frac{\mathbf{B}\cdot\mathbf{H}}{2} = \frac{BH}{2} = \frac{\mu H^2}{2} = \frac{B^2}{2\mu} \quad \left[\frac{J}{m^3}\right] \tag{9.91}$$

In practical use it is common to first calculate the magnetic energy density and then integrate over the volume of a device, or over some part of space to calculate the total magnetic energy stored in a given volume.

We conclude this section by returning to the initial result in Eq. (9.62). If the energy stored in the inductor is known, it can be used to calculate the inductance as:

$$W_m = \frac{1}{2}LI^2 \rightarrow L = \frac{2W_m}{I^2} \quad [H] \tag{9.92}$$

Since it is sometimes easier to calculate the energy due to a current-carrying conductor or a system of conductors, it is often easier to calculate the energy in the system using any of the above equations (such as Eq. (9.89)) and calculate the inductance of the system from energy. For example, using the general result for energy

$$L = \frac{1}{I^2}\int_v \mathbf{B}\cdot\mathbf{H}\,dv \quad [H] \tag{9.93}$$

However, we should be careful since Eq. (9.93) gives the total inductance and it is not always easy to differentiate between self and mutual inductances. This relation should only be used if only a self-inductance exists or if the total inductance of the system is needed.

▼ EXAMPLE 9.15 Stored magnetic energy

An infinitely long solenoid with radius a and n turns per unit length is given. The turns carry a current I. A long piece of iron, of radius b is placed in the solenoid, as shown in **Figure 9.35**. The relative permeability of iron is μ_r and that of free space is 1. Calculate the total work per unit length of the solenoid necessary to pull the iron completely out of the solenoid. Assume iron does not saturate and the B(H)

Answer.

$$L_{11} = \mu_0 n^2 \pi b^2, \quad L_{22} = \mu_0 n^2 \pi a^2, \quad L_{12} = L_{21} = \mu_0 n^2 \pi b^2 \quad \text{[H/m]}$$
$$L_t = L_{11} + L_{22} - L_{12} - L_{21} = \mu_0 n^2 \pi (a^2 - b^2) \quad \text{[H/m]}$$

9.6 MAGNETIC CIRCUITS

A useful and relatively simple tool in design of magnetic devices is the idea of the magnetic circuit. It is based on an analogy between the fundamental equations and properties of the static magnetic field and the static electric field. This analogy is based on the following relations:

Electric field   Magnetic field

$$\mathbf{E} = \frac{1}{\sigma}\mathbf{J} \qquad \mathbf{H} = \frac{1}{\mu}\mathbf{B} \tag{9.94}$$

$$V = \int_C \mathbf{E} \cdot d\mathbf{l} \qquad NI = \oint_C \mathbf{H} \cdot d\mathbf{l} \tag{9.95}$$

$$I = \int_s \mathbf{J} \cdot d\mathbf{s} \qquad \Phi = \int_s \mathbf{B} \cdot d\mathbf{s} \tag{9.96}$$

The "current" in the magnetic circuit is represented by the magnetic flux, while "voltage" is represented by the term NI. The latter, defined in Ampere's law (for example, Eq. (9.86)), is often called the **magnetomotive force**. The voltage V is an **electromotive force** supplied by a source such as a battery. We must note, however, a number of points:

(1) The magnetic circuit is only an analogy: Flux is not a "magnetic current" and the magnetomotive force is not a "magnetic voltage."

(2) This analogy between circuits only applies if the conditions for a circuit are satisfied. In particular, an electric circuit requires that current flows in closed circuits in conductors. Similarly, the flux must flow in closed circuits in "magnetic paths." These concepts must be carefully defined and understood before we can properly use the analogy between electric and magnetic circuits.

(3) The magnetomotive force in a magnetic circuit is supplied either by a coil with N turns and current I or by an equivalent permanent magnet, as shown in **Figure 9.37**.

In the equivalent circuit in **Figure 9.37c**, the meaning of the equivalent resistance R, has not yet been defined. To do so, we use the simple magnetic circuit in **Figure 9.38**. The torus is used here because it satisfies the basic condition of a magnetic circuit; namely, all flux is contained within the magnetic circuit (magnetic core). The magnetic field intensity for the torus was calculated in **Example 9.7**. Its magnitude is

$$H = \frac{NI}{2\pi r_a} \tag{9.97}$$
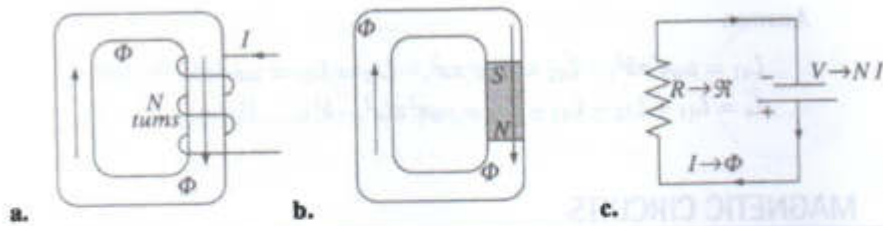
FIGURE 9.37   (a) Magnetomotive force in a magnetic circuit generated by a coil with N turns and current I. (b) A magnetomotive force generated by a permanent magnet. (c) The equivalent circuit for (a) or (b).
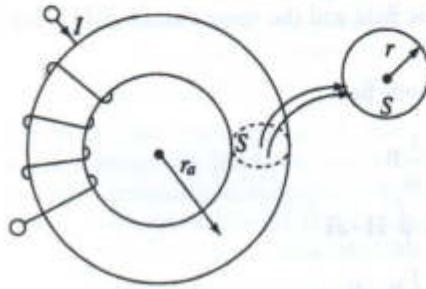
FIGURE 9.38   A closed magnetic circuit used to define reluctance.

where r_a is the average radius of the torus. The field intensity is uniform in the cross-sectional area of the torus (approximately) and, therefore, the flux is given by

$$\Phi = \int_s \mathbf{B} \cdot d\mathbf{s} = \int_s \mu \mathbf{H} \cdot d\mathbf{s} = \mu H S = \mu \frac{NI}{2\pi r} S \qquad (9.98)$$

where μ is the permeability of the material and S is the cross-sectional area of the magnetic circuit. The magnetomotive force (NI) and the flux (Φ) are known. From the equivalent circuit in **Figure 9.37c**, the equivalent "resistance" in the circuit is

$$\mathfrak{R} = \frac{NI}{\Phi} = \frac{NI}{\mu \frac{NI}{2\pi r_a} S} = \frac{2\pi r_a}{\mu S} = \frac{l}{\mu S} \qquad (9.99)$$

where $l = 2\pi r_a$ is the average length of the magnetic path. $\mathfrak{R}$ is called *magnetic reluctance* and is analogous to resistance in an electric circuit. Note that in this sense, μ is viewed as a "magnetic conductivity" and $1/\mu$ as a "magnetic resistivity." The term $1/\mu$ is also called *reluctivity* of the magnetic material in analogy to the *resistivity* of conductors.

The calculation of reluctance of a circuit is straightforward: It is equal to the length of the magnetic circuit divided by the cross-sectional area and magnetic permeability. It depends on the physical size of the device and on its permeability.

The use of a torus in the above derivation was arbitrary: Any closed magnetic circuit would do. In fact, we may consider a magnetic circuit made of a number of materials with perhaps many branches. As long as all flux is contained within the magnetic circuit, the circuit can be analyzed using this method. In particular,
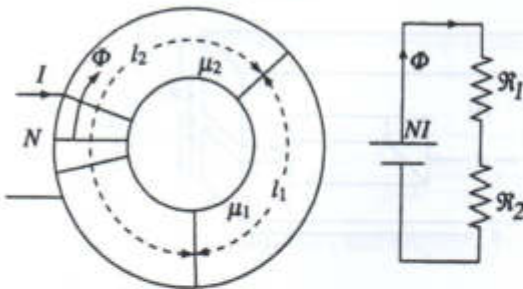
FIGURE 9.39   Magnetic circuit made of two magnetic materials and equivalent circuit.

consider the magnetic circuit in **Figure 9.39** which is made of two sections, one with permeability μ_1 the other μ_2.

The reluctance of each section is calculated from **Eq. (9.99)**. In this case, there are two reluctances, connected in series:

$$\mathfrak{R}_1 = \frac{l_1}{\mu_1 S}, \qquad \mathfrak{R}_2 = \frac{l_2}{\mu_2 S} \tag{9.100}$$

The flux in the magnetic circuit is

$$\Phi = \frac{NI}{\mathfrak{R}} = \frac{NI}{\mathfrak{R}_1 + \mathfrak{R}_2} = \frac{NI}{l_1/\mu_1 S + l_2/\mu_2 S} \tag{9.101}$$

Note also that any material can be included in the circuit, is long as the conditions of the circuit are satisfied. If, however, a very high, reluctance material is included (low μ)-such as free space, the length of this material must be kept to a minimum, otherwise the flux will spread out and the circuit is not a true circuit any more. In other words, when large air gaps are included in the magnetic circuit, the analysis of the magnetic circuit using this method may not be valid or a large error may be introduced.

The expression in **Eq. (9.101)** may be generalized for any number of magneto-motive forces and reluctances in a closed circuit as

$$\boxed{\Phi = \frac{\sum_{i=1}^{n} N_i I_i}{\sum_{j=1}^{k} \mathfrak{R}_j}} \tag{9.102}$$

▼ EXAMPLE 9.17   Magnetic circuit with a gap

The magnetic circuit in **Figure 9.40** is given. Calculate the magnetic field intensity H in the gap.

Solution.   The equivalent circuit with a source equal to $N_1 I_1$ and reluctances in each path is shown in **Figure 9.40c**. The three reluctances are calculated using the average paths shown in **Figure 9.40b**. Note that the cross-sectional area everywhere is bc except in the central leg, where it is ec. Since the flux in the gap is needed, it is best to use two loops as shown in **Figure 9.40c**, but other methods can be used.

B_1=flux density produced by loop C_1

I_1=current in loop C_1

Φ_{11} = magnetic flux, produced by loop C_1 that passes through loop C_1

Φ_{12} = magnetic flux, produced by loop C_1 that passes through loop C_2

In the discussion that follows, it is assumed that the medium in which the loops are placed is linear; that is, the permeability of the medium is independent of the currents in either loop.

The total flux in loop C_1 is

$$\Phi_{11} = \int_{s_1} \mathbf{B}_1 \cdot d\mathbf{s}_1 \quad \text{[Wb]} \tag{9.49}$$

Similarly, the flux through loop C_2 is

$$\Phi_{12} = \int_{s_2} \mathbf{B}_1 \cdot d\mathbf{s}_2 \quad \text{[Wb]} \tag{9.50}$$

The flux density $\mathbf{B}_1$ is not known unless the actual dimensions of C_1 and the value of I_1 are known. However, in very general terms, we recall that in the Biot-Savart law, if the current I is a constant, it may be taken outside the integral sign (see **Eq. (8.9)**). Therefore, $\mathbf{B}_1$ is directly proportional to the current I_1:

$$\mathbf{B}_1 = I_1 \left[\frac{\mu_0}{4\pi} \oint_{loop} \frac{d\mathbf{l}' \times \hat{\mathbf{R}}}{|\mathbf{r}_1 - \mathbf{r}'|} \right] \quad \text{[T]} \tag{9.51}$$

where $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}'$. $\mathbf{r}_1$ is the position vector of the point at which $\mathbf{B}_1$ is calculated and $\mathbf{r}'$ is the position vector of the current element $d\mathbf{l}'$. Suppose we use this equation and calculate the magnetic flux density everywhere in space due to the loop and then substitute it in **Eqs. (9.49)** and **(9.50)** and perform the integration over s_1 and s_2. The result is the magnetic fluxes Φ_{11} and Φ_{12}, which are directly proportional to I_1:

$$\Phi_{11} = L_{11} I_1 \quad \text{[Wb]} \tag{9.52}$$

$$\Phi_{12} = L_{12} I_1 \quad \text{[Wb]} \tag{9.53}$$

The terms L_{11} and L_{12} are constants, independent of current; only the geometry of the configuration plays a role in the values of L_{11} and L_{12}. L_{11} is defined as

$$L_{11} = \frac{\Phi_{11}}{I_1} \quad \left[\frac{\text{Wb}}{\text{A}} \right] = \text{[H]} \tag{9.54}$$

This is called the *self-inductance* of loop C_1. It is a self-inductance because all terms involved in its calculation relate only to itself. Loop C_2 has no effect on the self-inductance of loop C_1. Similarly,

$$L_{12} = \frac{\Phi_{12}}{I_1} \quad \text{[H]} \tag{9.55}$$

L_{12} is called the *mutual inductance* between loop C_1 and loop C_2. It is a mutual inductance because the flux Φ_{12} depends both on loop C_1 and on loop C_2, as is evident from **Eq. (9.50)**.

Before proceeding we note the following:
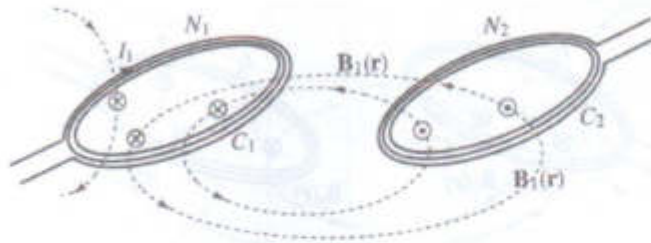
FIGURE 9.21 Inductance of multiple loop circuits. Circuit (1) produces a magnetic flux part of which passes through circuit (2).

(1) Inductance (self or mutual) has units of [Wb/A]. This unit is designated as the henry [H].

(2) Inductance depends only on the geometrical configuration of the circuits and is independent of current. Although the current produces the flux, once we divide by the current, the result contains only geometrical terms (and permeability).

(3) Any device that has inductance may be called an *inductor*. In practice, an inductor is usually a coil or a solenoid designed specifically for its inductance and often used as a circuit element.

The above results can be generalized in two ways: One is to ask ourselves what happens if instead of a single turn, we have two or more turns either in C_1 or C_2 or both. This is shown in **Figure 9.21**, where N_1 loops are placed together to form a circuit C_1 and N_2 loops are placed to form a circuit C_2. We argue as follows: If the flux due to a single loop carrying a current I is equal to Φ, the flux due to N loops carrying the same current I is $N\Phi$ from the principle of superposition, provided that the same flux passes through all loops. Thus, the total flux in C_1 is $N_1\Phi_{11}$, and the total flux passing through C_2 is $N_2\Phi_{12}$. The equations for self- and mutual inductance can now be written as

$$L_{11} = \frac{N_1\Phi_{11}}{I_1} \quad \text{and} \quad L_{12} = \frac{N_2\Phi_{12}}{I_1} \quad \text{[H]} \tag{9.56}$$

The term $N_1\Phi_{11}$ is the total flux that links circuit C_1 with itself or the *flux linkage* of C_1 with itself. The term $N_2\Phi_{12}$ is the total flux that links circuit C_1 with circuit C_2 or the flux linkage between circuit C_1 and C_2. These terms are denoted as

$$\Lambda_{11} = N_1\Phi_{11} \quad \text{and} \quad \Lambda_{12} = N_2\Phi_{12} \quad \text{[weber · turns]} \tag{9.57}$$

The unit of flux linkage is the weber · turns [Wb · t]. With this notation, the self- and mutual inductances are defined as

$$L_{11} = \frac{\Lambda_{11}}{I_1} \quad \text{and} \quad L_{12} = \frac{\Lambda_{12}}{I_1} \quad \text{[H]} \tag{9.58}$$

This definition is more general and includes in it the definition in Eqs. **(9.54)** and **(9.55)**.

C H A P T E R

8

The Static Magnetic Field

"There's a South Pole," said Christoper Robin, "and I expect there's an East Pole and a West Pole, though people don't like talking about them."

Winnie-The-Pooh

8.1 INTRODUCTION

After discussing the static electric field and steady currents, we are now ready to take another significant step in the study of electromagnetics; the study of the static magnetic field. But what exactly is a magnetic field? This question will be answered gradually, but, for a simple description, we may say it is a new type of force field in the same sense that the electric field is a force field. Take, for example, a magnet. It attracts or repels other magnets and generates a "magnetic field" around itself. The permanent magnet generates a static (time independent) magnetic field. A direct current can also generate a static magnetic field. How do we know that? As with many other aspects of electromagnetics, we know by experiment.

The various properties of the magnetic field will be discussed primarily from the point of view of currents. The importance of this for engineering design is twofold; first, it indicates that a static magnetic field can be generated to suit design purposes. Second, it provides a link between the electric field and the magnetic field. Thus, we can view the electric field as a source of current, which, in turn, is the source for the magnetic field and, therefore, at least a partial explanation for the term "electromagnetics." The reason for the qualification is that the opposite does not happen in the case of static fields: A static magnetic field does not generate an electric field. We will see in **Chapters 10** and **11** that a time-dependent magnetic field does generate a time-dependent electric field, and at that point the link between the electric and magnetic field will be complete.

For many years, electric and magnetic fields were thought to be separate phenomena even though both were known since antiquity. It was not until 1819 when

471

Hans Christian Oersted[1] found that the needle of a compass moved in the presence of a current-carrying wire that a link between the two fields was found. He concluded that the only way this can happen is if the current generates a magnetic field around the wire. He used a compass to "map" the behavior of the magnetic field around the wire. Following his initial discovery, Andre Marie Ampere[2] quickly established the correct relation between current and magnetic field in what is now known as Ampere's law. With all that, as late as the beginning of the 20th century, it was still common to use separate units for electric and magnetic quantities without any connection between the two systems.

In this chapter, we discuss the relationships between the steady electric current and the static magnetic field. These will be in the form of basic postulates which, as in the case of the static electric field, are experimental in nature. Two important relations, the Biot-Savart and Ampere's laws, will allow us to calculate the magnetic field due to electric currents. The equivalency between current and magnetization will explain the rather poorly understood properties of the permanent magnet.

8.2 THE MAGNETIC FIELD, MAGNETIC FIELD INTENSITY, AND MAGNETIC FLUX DENSITY

What, then, is the magnetic field? We might start to answer this question by playing with two magnets. The first effect to notice is that there is a force between the two magnets; the magnets either attract or repel each other as shown in **Figure 8.1**. Since attraction happens at a distance, each magnet must have a domain in which it attracts the other magnet. This is exactly what we called a field. That the field is a vector field we can establish using a compass as a measuring device. The direction of the force is established by the direction of the compass needle in space. Placing the compass at as many positions as we wish, a complete map of the vector field is established. This simple measurement establishes the following:

(1) A field exists throughout space.

(2) The field is stronger closer to the magnet.

(3) The two ends of the magnet have different properties; one attracts the north pole of the compass and is labeled the south pole; the other attracts the south

---

[1] Hans Christian Oersted (1777–1851), Danish scientist and professor of physics. He tried for many years to establish the link between electricity and magnetism, a link which was suspected to exist by him and many other scientists of the same period. He finally managed to do so in his now famous experiment of 1819 in which he showed that a current in a wire affects a magnetic needle (compass needle). He disclosed his experiments, all made in the presence of distinguished witnesses, in 1820. Oersted was very careful to ensure that what he saw was, in fact, a magnetic phenomenon by repeating the experiments many times and with various "needles," in addition to the magnetic needle (to show that the effect does not exist in conducting materials such as copper or insulating materials such as glass - only in magnetized materials). Intervening materials between the wire and needle were also tested. As was the custom of the day, his work was written and communicated in Latin in a pamphlet titled: "Experimenta circa efficaciam, conflictus eletrici in acum magneticam."

[2] See Footnote 1 in Chapter 7.
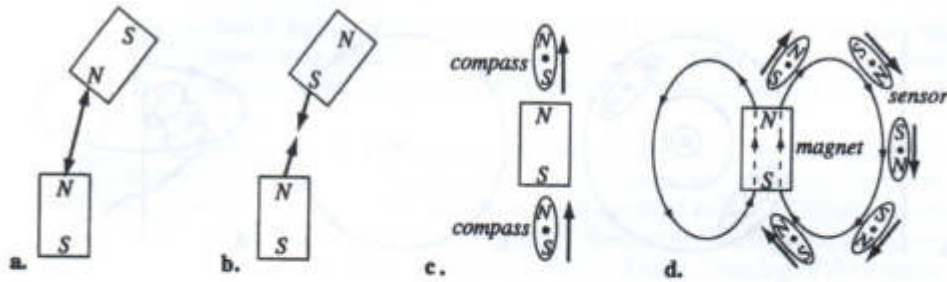
FIGURE 8.1 The permanent magnet. (a) Two permanent magnets repel each other. (b) Two permanent magnets attract each other. (c) Identification of the poles of a magnet. (d) Mapping the magnetic field of a permanent magnet using a compass.

pole of the compass and is labeled the north pole of the magnet (**Figure 8.1c**). This arbitrary identification is convenient because of its relation with the Earth's magnetic field.

(4) By placing the compass at different locations in space, we can map the magnetic field. One field line is shown in **Figure 8.1d.**

This simple experiment is all that is necessary to establish the existence of the magnetic field of a magnet. What we need to do now is to show that the same occurs due to steady electric currents and, more importantly, to find the exact relationship between the electric current and the magnetic field. For now, we will say that the magnetic field is a *field of force* acting on a magnet; the only qualification is that the magnet must be small (e.g., the needle of a compass) to render the measurement valid.

The first of these, showing that a force exists due to the magnetic field produced by a current, is easily performed using the above idea but for a steady current. This is, in fact, what Oersted did in his historical experiment. Consider a straight wire, carrying a steady (direct) current I as shown in **Figure 8.2a.** Positioning the compass at various locations in space, we note that a force exists on the compass. Since this force cannot be an electrostatic force (no charges on the compass) and it certainly cannot be gravitational (since disconnecting the current will cancel the force), we must conclude that the force is similar to that between two magnets. Thus, the current in the wire has generated a magnetic field with properties identical to those of an equivalent magnet.

Experiment shows that the direction of the compass now is tangential to any circle centered at the current and depends on its direction. In **Figure 8.2b**, the direction is counterclockwise, while in **Figure 8.2c**, it is clockwise. The direction of the compass (arrow pointing from S to N) is taken as the direction of the magnetic field. From this experiment, we establish two very general and important properties:

(1) The direction of the magnetic field due to a current is defined by the right-hand rule shown in **Figure 8.2d.** The rule states that

"if the thumb of the right-hand shows the direction of current, the fingers show the direction of the magnetic field."
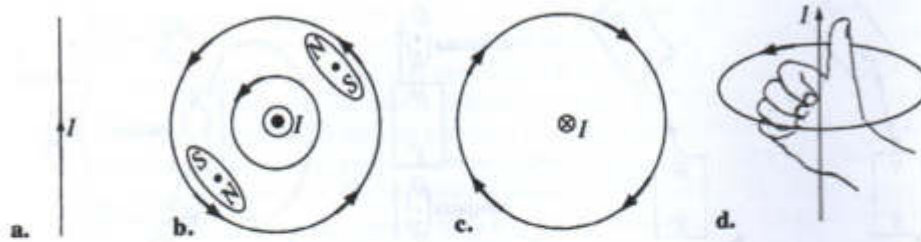
FIGURE 8.2   Oersted's experiment. (a) Current in a wire. (b) Direction of the compass needle for a current out of the page. (c) Direction of the compass for a current into the page. (d) The right-hand rule establishes the relation between the direction of current and the resulting magnetic field.

(2) The lines of magnetic field (that is, the lines to which the compass is tangential) are always closed lines. We will return to this aspect of the field shortly, after we have a better definition of the magnetic field.

The above experiment is lacking in one respect: Nothing has been established as far as the strength of the force exerted on the compass. We can intuitively say that the further away from the current we are, the lower the force, but exact values of force, or its relationship with the current, cannot be established from this experiment.

Since a magnetic field is generated by a steady current but not by a static charge, we conclude that the magnetic field and, therefore, the force in the magnetic field are related to the motion of charges (current) or, more specifically, to the velocity of charges. We recall that in the case of a static electric charge, a force acted on the charge as $F = q\mathbf{E}$. Similarly, there is a force acting on a moving charge due to the magnetic field. As with the electrostatic force, we can establish the force in a magnetic field by performing a series of experiments from which both the magnitude and direction of the force are found. These experiments lead to the following relation between the magnetic force $\mathbf{F}_m$, electric charge q, charge velocity $\mathbf{v}$, and a new quantity $\mathbf{B}$:

$$\mathbf{F}_m = q\mathbf{v} \times \mathbf{B} \quad [\text{N}] \tag{8.1}$$

where $\mathbf{B}$ is called the ***magnetic flux density***. This relation is known as the Lorentz[3] force equation for a moving charge in the magnetic field, whereas $\mathbf{F}_e = q\mathbf{E}$ is the Coulomb force for a charge q in the electric field. For a general field, which includes both electric and magnetic components, the force on a charge is the sum of the two forces and may be written as

$$\mathbf{F}_{total} = \mathbf{F}_e + \mathbf{F}_m = q\mathbf{E} + q\mathbf{v} \times \mathbf{B} \quad [\text{N}] \tag{8.2}$$

[3] Hendrik Antoon Lorentz, (1853–1928). Dutch physicist who is best known for his work on the effects of magnetism on radiation. For this he received the Nobel Prize in 1902. In his attempt to explain electricity, magnetism, and light, he arrived at the Lorentz transformation, which helped Albert Einstein in formulating his theory of relativity.

This relation is known as the Lorentz-Coulomb force equation (sometimes only as the Lorentz force equation). Note that a stationary charge ($v = 0$) only experiences an electric force, whereas a moving charge experiences both an electric and a magnetic force.

The relation $q\mathbf{v} \times \mathbf{B}$ may be viewed as the defining relation for the magnetic flux density $\mathbf{B}$ since all other quantities are known or measurable. This is similar to using the relation $\mathbf{F} = q\mathbf{E}$ to define the electric field intensity as $\mathbf{E} = \mathbf{F}/q$ (force per unit charge). Although we will define the magnetic flux density in a different way shortly, the discussion here is useful in that it shows the physical meaning of the magnetic flux density.

In terms of units, the magnetic flux density has units of (newton/coulomb)/ (meter/second) or newton/(ampere · meter). This unit is called a *tesla* [T] or weber/meter^2 [Wb/m^2]. The weber is equal to newton.meter/ampere. The common SI unit is the tesla [T] while the [Wb/m^2] is used sometimes to emphasize that $\mathbf{B}$ is a flux density. Both units are appropriate and commonly used.

In addition to the magnetic flux density, we define the magnetic field intensity $\mathbf{H}$ using the constitutive relation

$$\mathbf{B} = \mu\mathbf{H} \quad [\text{T}] \tag{8.3}$$

where μ is called the *magnetic permeability* (or, in short, permeability) of materials. In free space, the relation is

$$\mathbf{B} = \mu_0\mathbf{H} \quad [\text{T}] \tag{8.4}$$

where μ_0 is the *permeability of free space*. The magnitude of μ_0 is $4\pi \times 10^{-7}$ and its unit is the henry/meter or [H/m]. This unit will become obvious as we discuss the magnetic field intensity $\mathbf{H}$ and the magnetic flux density $\mathbf{B}$. At this point it is useful to recall that a similar relation ($\mathbf{D} = \varepsilon\mathbf{E}$) connects the electric field intensity $\mathbf{E}$ and the electric flux density $\mathbf{D}$. The relations in **Eqs. (8.3)** and **(8.4)** are fundamental and we will discuss them at some length in the following chapter. At this point, it is sufficient to indicate that both can be shown to be correct by performing whatever experiments are necessary to do so. This approach will allow us to introduce the important laws of Biot and Savart, and Ampere.

While Eq. (8.1) may be viewed as defining the magnetic flux density $\mathbf{B}$, this definition is not convenient for calculation because it involves the vector product of $\mathbf{v}$ and $\mathbf{B}$ rather than $\mathbf{B}$ directly and is defined on individual moving charges. We will, therefore, seek a more convenient mathematical relation, one that will allow calculation of the magnetic flux density directly from the current I. Also, the magnetic flux density is material dependent (through μ). For this reason, we calculate $\mathbf{H}$ rather than $\mathbf{B}$, at least until we had a chance to discuss behavior of materials in the magnetic field, which we will do in the following chapter. For these reasons, most of the discussion in this chapter is in terms of $\mathbf{H}$ rather than $\mathbf{B}$. This does not create much difficulty since $\mathbf{B} = \mu\mathbf{H}$ can be used to calculate $\mathbf{B}$. We will return to using $\mathbf{B}$ in latter parts of this chapter and in the next chapter because the force in the magnetic field depends on $\mathbf{B}$ and force is fundamental in magnetic applications.

There are, in fact, two relations that accomplish the task of calculating the magnetic field intensity: the Biot–Savart and Ampere's laws. We start by describ-

This relation is known as the Lorentz-Coulomb force equation (sometimes only as the Lorentz force equation). Note that a stationary charge ($v = 0$) only experiences an electric force, whereas a moving charge experiences both an electric and a magnetic force.

The relation $q\mathbf{v} \times \mathbf{B}$ may be viewed as the defining relation for the magnetic flux density $\mathbf{B}$ since all other quantities are known or measurable. This is similar to using the relation $\mathbf{F} = q\mathbf{E}$ to define the electric field intensity as $\mathbf{E} = \mathbf{F}/q$ (force per unit charge). Although we will define the magnetic flux density in a different way shortly, the discussion here is useful in that it shows the physical meaning of the magnetic flux density.

In terms of units, the magnetic flux density has units of (newton/coulomb)/ (meter/second) or newton/(ampere · meter). This unit is called a *tesla* [T] or weber/meter^2 [Wb/m^2]. The weber is equal to newton.meter/ampere. The common SI unit is the tesla [T] while the [Wb/m^2] is used sometimes to emphasize that $\mathbf{B}$ is a flux density. Both units are appropriate and commonly used.

In addition to the magnetic flux density, we define the magnetic field intensity $\mathbf{H}$ using the constitutive relation

$$\mathbf{B} = \mu\mathbf{H} \quad [\mathrm{T}] \tag{8.3}$$

where μ is called the *magnetic permeability* (or, in short, permeability) of materials. In free space, the relation is

$$\mathbf{B} = \mu_0\mathbf{H} \quad [\mathrm{T}] \tag{8.4}$$

where μ_0 is the *permeability of free space*. The magnitude of μ_0 is $4\pi \times 10^{-7}$ and its unit is the henry/meter or [H/m]. This unit will become obvious as we discuss the magnetic field intensity $\mathbf{H}$ and the magnetic flux density $\mathbf{B}$. At this point it is useful to recall that a similar relation ($\mathbf{D} = \varepsilon\mathbf{E}$) connects the electric field intensity $\mathbf{E}$ and the electric flux density $\mathbf{D}$. The relations in **Eqs. (8.3)** and **(8.4)** are fundamental and we will discuss them at some length in the following chapter. At this point, it is sufficient to indicate that both can be shown to be correct by performing whatever experiments are necessary to do so. This approach will allow us to introduce the important laws of Biot and Savart, and Ampere.

While Eq. (8.1) may be viewed as defining the magnetic flux density $\mathbf{B}$, this definition is not convenient for calculation because it involves the vector product of $\mathbf{v}$ and $\mathbf{B}$ rather than $\mathbf{B}$ directly and is defined on individual moving charges. We will, therefore, seek a more convenient mathematical relation, one that will allow calculation of the magnetic flux density directly from the current I. Also, the magnetic flux density is material dependent (through μ). For this reason, we calculate $\mathbf{H}$ rather than $\mathbf{B}$, at least until we had a chance to discuss behavior of materials in the magnetic field, which we will do in the following chapter. For these reasons, most of the discussion in this chapter is in terms of $\mathbf{H}$ rather than $\mathbf{B}$. This does not create much difficulty since $\mathbf{B} = \mu\mathbf{H}$ can be used to calculate $\mathbf{B}$. We will return to using $\mathbf{B}$ in latter parts of this chapter and in the next chapter because the force in the magnetic field depends on $\mathbf{B}$ and force is fundamental in magnetic applications.

There are, in fact, two relations that accomplish the task of calculating the magnetic field intensity: the Biot–Savart and Ampere's laws. We start by describ-

The current dI in an infinitesimally thin section of the sheet (cut lengthwise as in **Figure 8.11b**) is

$$dI = \frac{I}{b}\, dx' = 100\, dx' \quad [\text{A}]$$

The magnetic field intensity at a distance r from a long line carrying a current dI (see **Example 8.1**) is

$$d\mathbf{H} = \hat{\Phi}\frac{dI}{2\pi r} \quad \left[\frac{\text{A}}{\text{m}}\right]$$

Now we place this infinitesimal segment at a general location x' on the sheet, as shown in **Figure 8.11c**. The distance r is the distance to the point at which we wish to calculate the magnetic field intensity. Because as x' changes, the angle θ and, therefore, the direction of the field intensity change, it is easier to calculate the vertical and horizontal components separately and then integrate each. From symmetry alone, we know that the vertical component must cancel because symmetric segments have equal and opposite vertical components. The horizontal component is

$$dH_x = \frac{dI}{2\pi r}\,\cos(90 - \theta) = \frac{b\,dI}{2\pi r^2} \quad \left[\frac{\text{A}}{\text{m}}\right]$$

Substituting $r = (x'^2 + b^2)^{1/2}$ and rearranging terms, we get

$$dH_x = \frac{b\,dI}{2\pi\,(x'^2 + b^2)} \quad \left[\frac{\text{A}}{\text{m}}\right]$$

Substituting $dI = 100\, dx'$ and $b = 0.1$ and integrating between $x' = -b/2 = -0.5$ m and $x' = b/2 = 0.5$ m gives

$$H_x = \frac{10}{2\pi} \int_{-0.5}^{0.5} \frac{dx'}{(x'^2 + 0.1^2)} = \frac{10}{2\pi}\left[\frac{1}{0.1}\tan^{-1}\frac{x'}{0.1}\right]_{-0.5}^{0.5}$$

$$= \frac{50 \times 2.7468}{\pi} = 43.72 \quad \left[\frac{\text{A}}{\text{m}}\right]$$

The resulting magnetic field intensity is parallel to the sheet (in the x direction in **Figure 8.11c**).

8.4 AMPERE'S LAW

Ampere's law,[5] also called the Ampere circuital law states that

> "the circulation of H around a closed path C is equal to the current enclosed by the path."

---

[5] Ampere's law is a simplified form of the Biot–Savart law. Named after Marie Andre Ampere (See Footnote 1 in Chapter 7). Ampere derived this law from study of the solenoidal (circular) nature of the magnetic field of straight wires.

That is,

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = I_{enclosed} \quad [\text{A}]$$ (8.16)

The circulation is defined by the line integral around a closed contour of the scalar product $\mathbf{H} \cdot d\mathbf{l}$. Thus, in fact, only the component of $\mathbf{H}$ tangential to the contour of integration is included in the calculation.

Why do we need another law which, as stated, is derived from the Biot–Savart law (even though we show no proof)? The reason is simple and is the same we used earlier to justify Gauss' law as opposed to the use of Coulomb's law; Ampere's law is much easier to apply to some problems. In particular, highly symmetric current configurations are very easy to evaluate using Ampere's law, whereas they may be more complex using the Biot–Savart law. Also, in the next section we will show that Ampere's law is, in fact, one of the postulates of the magnetic field. In a practical sense, Ampere's law is another tool which we can use whenever it makes sense to do so.

When applying Ampere's law, we must remember that under normal circumstances, the unknown quantity is $\mathbf{H}$. Since $\mathbf{H}$ is inside the integral sign, we must find a closed contour, enclosing the current I such that the component of $\mathbf{H}$ tangential to the contour is constant along the contour. Under these conditions, $\mathbf{H} \cdot d\mathbf{l} = H_{tan}\,dl$ and H_{tan} can be taken outside the integral sign. Evaluation of the integral is now possible. Thus, the requirements here are that $\mathbf{H}$ be tangential to a contour and constant along the contour. These conditions are satisfied for highly symmetric current configurations. These include the following:

(1) Current in an infinite (or as an approximation, in a very long) filament.

(2) Current or current density in an infinite (or very long) solid or hollow cylindrical conductor. Normally, the current is uniformly distributed in the conductor, but any symmetrical current distribution is allowed.

(3) Infinite sheet of current (or very large, flat current sheet).

(4) Infinite (or very long), thick conductor with symmetric current or current density distribution.

(5) Multiple conductors in a symmetrical configuration.

(6) Nonsymmetric current distributions which are a superposition of symmetric current distributions.

The key requirement for the application of Ampere's law is symmetry of current or current density distribution. The law applies to other current configurations, but it is not generally possible to find a contour over which the tangential magnetic field intensity is constant and, therefore, evaluate the field. The following examples show various important aspects of application of Ampere's law.

312

For $r > b$: Contour C_3 encloses N positive currents and N negative currents. Thus, the total net current enclosed is zero and the magnetic field intensity and magnetic flux density are zero.

Notes.

(1) The magnetic field is contained entirely within the torus. This property is used in many applications, including high-quality transformers and other coils which require containment of the field. Because of the field containment, the magnetic field generated in a torus does not affect or interfere with other devices.

(2) The effect of having 100 turns is to multiply the current by 100. In other words, from Ampere's law, the total current enclosed in the contour is $100I$. If this current is due to a single turn carrying 100 A, 10 turns carrying 10 A each, or 10,000 turns carrying 0.01 A each, the result is the same. For this reason, we will often use the term *ampere·turns* to indicate the total current in a device.

(3) This problem, although easily solved using Ampere's law, would be almost impossible to solve using the Biot–Savart law; integration along all current paths would be very tedious.

▼ **EXAMPLE 8.10   Application: Magnetic field inside a long solenoid**

A solenoid is a coil, wound on a circular (sometimes rectangular) form. The turns of the coil are most often tightly wound and normally in a single layer, whereas the coil is very long. Short, multilayer solenoids are called coils.

A very long solenoid is wound with a turn density of 1 turn per mm length of the solenoid ($n = 1000$ turns/meter). The current in the solenoid is $I = 1$ A. Calculate the magnetic field intensity and magnetic flux density everywhere in space. The current and dimensions are shown in **Figure 8.15a** (in cross section).

Solution. Although the coil is cylindrical, the cross section shown in **Figure 8.15a** is identical through any longitudinal cut that includes the axis of the coil. The configuration looks as if we had two planar current sheets, with currents in opposite directions. From the right-hand rule, we note that the fields of the two opposing current layers are in opposite directions outside the solenoid but are in the same direction inside the solenoid. Thus, the field outside the solenoid must be zero.
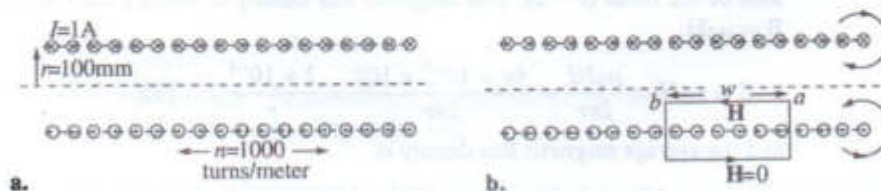

FIGURE 8.15 A long solenoid. (a) Dimensions and properties (axial cross section shown). (b) Contour used for calculation.

A contour is chosen as shown in **Figure 8.15b**. The width of the contour is w and is arbitrary. The total current enclosed by the contour is nwI. Since the field outside the solenoid is zero and inside it must be along the axis (from the right hand rule and symmetry considerations), Ampere's law gives

$$\oint_{C_2} \mathbf{H} \cdot d\mathbf{l} = H \int_a^b dl_2 = Hw = nwI \rightarrow H = nI$$

Thus, the magnetic field intensity inside the solenoid is

$$H = nI = 1000 \qquad \left[\frac{A}{m}\right]$$

The magnetic flux density is found by multiplying the magnetic field intensity by the permeability of free space:

$$B = \mu_0 H = \mu_0 nI = 4\pi \times 10^{-7} \times 1000 = 1.257 \times 10^{-3} \qquad [T]$$

Like the torus in **Example 8.9**, the field of a long solenoid is zero outside the solenoid. Inside the solenoid, the field is constant. Practical solenoids are finite in length and for these, the result obtained here is only an approximation. Note also that an infinite solenoid may be viewed as a torus with infinite radius (see **Exercise 8.2**). Note, again, that the same result would be obtained if, for example, the turn density were 100 turns/m while the current were 10 A. What is important for the solution is the term nI, which is the number of ampere · turns per unit length of the solenoid. For this reason, a solenoid may also be made of a single, bent sheet, carrying a given current.

■**EXERCISE 8.2**

Show that the magnetic field intensity of a circular cross-section torus of average radius a, equals the field of an infinite solenoid as the average radius of the torus tends to infinity.

8.5 MAGNETIC FLUX DENSITY AND MAGNETIC FLUX

We started this chapter with the Lorentz force on a moving charge and showed that this force is proportional to the magnetic flux density **B** (see **Eq. (8.1)**). The relation between the magnetic field intensity **H** and the magnetic flux density **B** for general materials was also given as $\mathbf{B} = \mu\mathbf{H}$ in **Eq. (8.3)**. The permeability μ is material related. We will discuss this in more detail in the following chapter. For now it is sufficient to say that every material has a given and measurable permeability μ and we may assume that permeability is known, even though it may not be a constant value.

[25]   K., Christ, "Kalibrierung von Magnet-Injektoren für Benzin-Direkteinspritzsysteme mittels Körperschall." Dr. Ing. Thesis, Karlsruher Instituts für Technologie (KIT), Karlsruher, Alemania, 2011.

Christ, K., "Kalibrierung von Magnet-Injektoren für Benzin-Direkteinspritzsysteme mittels Körperschall." Dr. Ing. tesis, Karlsruher Instituts für Technologie (KIT), KIT Karlsruhe, Alemania, 2011.

Neben der Kraftstoffeinspritzung wird zur Bildung des Luft-Kraftstoffgemisches im Brennraum eine entsprechende Menge an Frischluft benötigt. Diese wird über den Ansaugstutzen, die Drosselklappe und die Lufteinlassventile bereitgestellt. Die Dosierung dieser Luftmenge ist ebenso entscheidend für eine saubere Verbrennung wie die Kraftstoffeinspritzung. Der Vorgang der Luftzufuhr sowie der Ausstoß des verbrannten Gemischs sind jedoch nicht Gegenstand der vorliegenden Arbeit. Details zu den Grundlagen finden sich in [6, 8]. Aktuelle Forschungsvorhaben, wie die voll-variable Ventilführung oder alternative Ansätze zum Ventilhub, werden in [15, 35] behandelt. Die Verbrennung, die nach der Kraftstoffinjektion ausgeführt wird, wird durch die Zündkerze eingeleitet. Der Prozess der Kraftstoffverbrennung sowie die Abgasnachbehandlung werden im Rahmen dieser Forschungsarbeit nicht abgehandelt. Das Vorhaben konzentriert sich vollständig auf den Vorgang der Kraftstoffinjektion mit Hilfe von Magnetspulen-Injektoren. Der Aufbau, die Ansteuerung und das Arbeitsverhalten solcher Injektoren werden im nächsten Abschnitt behandelt.

2.2 Magnetspulen-Injektor

Das Verständnis des Aufbaus und der Abläufe im Magnet-Injektor während der Kraftstoffeinspritzung bildet die Grundlage für die später angestellten Untersuchungen. Betrachtet werden ausschließlich Magnetspulen-Injektoren [1], wie sie in modernen BDE-Motoren zum Einsatz kommen. Im Vergleich zu Piezo-Injektoren haben Magnetspulen-Injektoren den Vorteil einer einfacheren Ansteuerung und sind ungefähr um den Faktor 2,5 kostengünstiger [14]. Allerdings erzielen Piezo-Injektoren im Allgemeinen eine höhere Genauigkeit in der Kraftstoffdosierung [43]. Wie die Präzision der Magnet-Injektoren durch ein Kalibrierungsverfahren erhöht werden kann, wird in Abschnitt 2.3 besprochen. Zunächst werden aber die in dieser Arbeit betrachteten Magnet-Injektoren vorgestellt.

2.2.1 Aufbau

In Abbildung 2.2 ist ein schematisches Schnittbild eines der betrachteten Magnetspulen-Injektoren dargestellt [52]. Der Injektor besteht im Wesentlichen aus einer Injektornadel, welche von einer Feder auf die Injektordüse gepresst wird. Die Aufgabe der Düse ist es, den Kraftstoff möglichst

Verständnis: comprensión
Ablauf: transcurso, desarrollo
während: mientras, durante
an/stellen: poner en marcha.
betrachten: examinar, considerar, observar

ausschließlich: exclusivamente
Kostengünstig: barato.
Allerdings: por supuesto
erzielen: obtener, realizar, conseguir
Abschnitt: segmento, sección, inciso,
besprochen: discutir, tratar, comentar,

Abbildung: figura, diagrama, gráfico, ilustación
Schnittbild: vista en sección

9

315

Kraftstoffdruck p_{rail}

Spannhülse

Magnetkern

Magnetanker

Feder

Steuerstrom $I(t, T_i)$

Injektornadel

Injektordüse

Kraftstoffmenge q

Handschriftliche Notizen (rechts):
Sachverhalt: hecho, circunstancia
Gegenstand: asunto, tema, materia.
zahlreich: numeroso, copioso.
Im Rahmen: como parte de
vorliegend: presente
Ausführung: realización, ejecución, acabado.
Eisenkern: núcleo de hierro.
ausbilden: formar, desarrollar,
samt: acompañado de, junto con.
empfindlichkeit: sensibilidad
starr: fijo, rígido,
verschweißt: soldado => schweißen: soldar
lagern: reposar, almacenar, alojar
überwiegen: prevalecer, predominar, preponderar
entgegengesetzt: opuesto, contrario

Abbildung 2.2 Aufbau eines der betrachteten Magnet-Injektoren [52]

Handschriftliche Notizen:
abbauen: desarmar, descomponer, debilitar.
zurücksetzen: reducir

fein im Brennraum zu zerstäuben. Dieser Sachverhalt ist Gegenstand zahlreicher anderer Forschungsarbeiten. Alle Injektoren, die im Rahmen der vorliegenden Arbeit untersucht werden, sind mit einer Mehrlochdüse ausgerüstet. Die Eigenschaften und die verschiedenen Ausführungsformen der Injektordüse werden in [7, 14] vorgestellt. Wird nun eine Spannung an die Magnetspule angelegt, bildet sich über den Eisenkern ein Magnetfeld aus, das den Anker samt Nadel nach oben zieht. Bei einigen der betrachteten Injektoren ist zur Reduzierung der Druckwellenempfindlichkeit der Anker nicht starr mit der Nadel verschweißt, sondern über eine kleine Feder gelagert. Überwiegt die Magnetkraft die entgegengesetzte Federkraft, hebt sich die Nadel und die Düse wird freigegeben. Der unter Hochdruck stehende Kraftstoff strömt durch die Düse in den Brennraum und wird dort zerstäubt. Die Kraftstoffeinspritzung dauert solange an, bis der Steuerstrom zurückgesetzt wird und sich das Magnetfeld abbaut. Die Feder drückt die Nadel auf die Düse und blockiert diese schließlich. Nach dem Schließen der Injektordüse kommt es zu Druckpulsierungen, die Kräfte am Magnetanker ausüben. Um damit verbundene unerwünschte Effekte ausschließen zu können, ist eine niedrige Druckwellenempfindlichkeit des Injektors wünschenswert.

Handschriftliche Notizen:
schließlich: finalmente, en definitiva, por último. ausüben: ejercer, desempeñar. damit: con ello
verbunden: asociado, conectado ausschließen: excluir, dejar fuera, descartar.
wünschenswert: deseable,

10

2.2.2 Ansteuerung

[handwritten annotations: Ze nach: dependiendo de ... hinterlegt: depositado. ablesen: electuar lectura, leer.]

Die Ansteuerung des Magnetspulen-Injektors erfolgt direkt über das Motorsteuergerät (ECU - electronic control unit). Je nach Betriebspunkt wird eine bestimmte Einspritzmenge q_{soll} bei einem Kraftstoffdruck p_{rail} angefordert. Aus dem dafür im Motorsteuergerät hinterlegten Kennfeld wird die entsprechende Ansteuerdauer T_i abgelesen. Die Ansteuerdauer wird in Form eines zweiwertigen Steuersignals $s(t, T_i)$ an die Injektorendstufe übergeben. Diese generiert einen injektorspezifischen Stromverlauf $I(t, T_i)$, wie in Abbildung 2.3 dargestellt. Der Injektorstrom durchläuft in

[handwritten: zweiwertig: bivalente. endstufe: etapa de salida. übergeben: entregar, transmitir, dar]
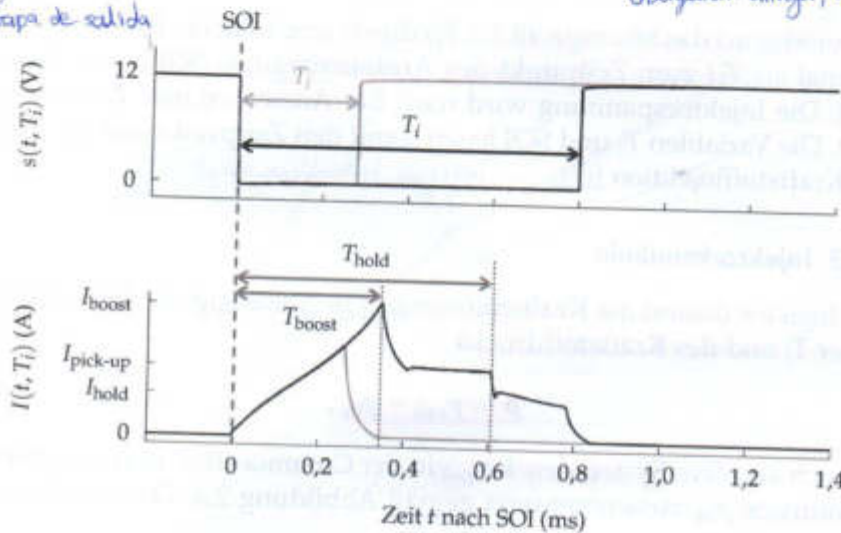

Abbildung 2.3 Steuersignal $s(t, T_i)$ (oben) und daraus generierter Injektorstrom $I(t, T_i)$ (unten)

Abhängigkeit der Ansteuerdauer ein festgelegtes Stromprofil. Das Stromprofil wird durch eine Injektorendstufe generiert und ist in einzelne Phasen unterteilt:

[handwritten: einzeln: separados, singular.]

- **Booster Intervall**: Ein in der Injektorendstufe integrierter Booster-Kondensator liefert beim Ansteuerbeginn SOI kurzzeitig eine Spannung in der Größenordnung um $U = 65$ V. Die Entladung des Kondensators führt zu einem starken Stromanstieg bis zu einer definierten Stromstärke I_{boost}. Durch diesen Stromanstieg wird gewährleistet, dass sich das Magnetfeld zum Anheben der Injektornadel

[handwritten: ntladung: descarga. römstärke: amperaje]

11

2.2.2 Ansteuerung

(handwritten: Je nach: dependiendo de   hinterlegt: depositado.   ablesen: alectuar lectura, leer.)

Die Ansteuerung des Magnetspulen-Injektors erfolgt direkt über das Motorsteuergerät (ECU - electronic control unit). Je nach Betriebspunkt wird eine bestimmte Einspritzmenge q_{soll} bei einem Kraftstoffdruck p_{rail} angefordert. Aus dem dafür im Motorsteuergerät hinterlegten Kennfeld wird die entsprechende Ansteuerdauer T_i abgelesen. Die Ansteuerdauer wird in Form eines zweiwertigen Steuersignals $s(t, T_i)$ an die Injektorendstufe übergeben. Diese generiert einen injektorspezifischen Stromverlauf $I(t, T_i)$, wie in Abbildung 2.3 dargestellt. Der Injektorstrom durchläuft in

(handwritten left: zweiwertig: bivalente   endstufe: etapa de salida)

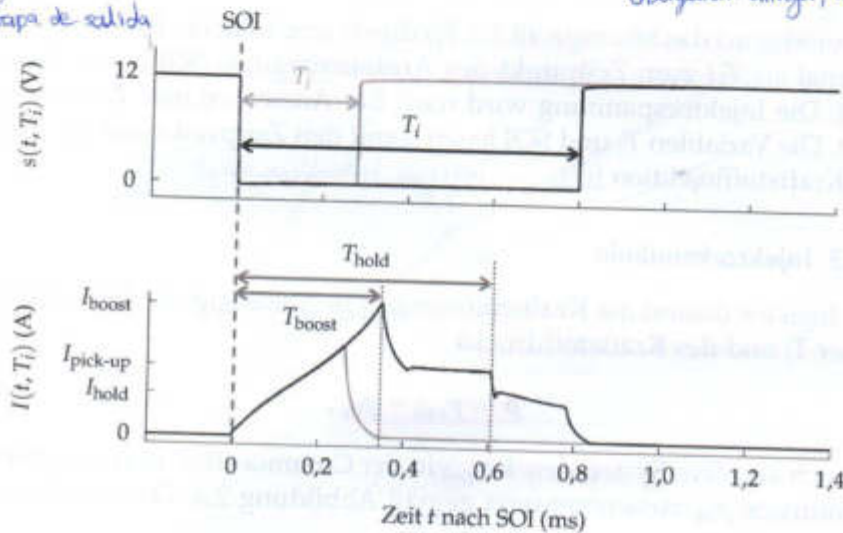*(handwritten right: übergeben: entregar, transmitir, dar)*

Abbildung 2.3 Steuersignal $s(t, T_i)$ (oben) und daraus generierter Injektorstrom $I(t, T_i)$ (unten)

Abhängigkeit der Ansteuerdauer ein festgelegtes Stromprofil. Das Stromprofil wird durch eine Injektorendstufe generiert und ist in einzelne Phasen unterteilt:

(handwritten: einzeln: separado, singular.)

- **Booster Intervall**: Ein in der Injektorendstufe integrierter Booster-Kondensator liefert beim Ansteuerbeginn SOI kurzzeitig eine Spannung in der Größenordnung um $U = 65$ V. Die Entladung des Kondensators führt zu einem starken Stromanstieg bis zu einer definierten Stromstärke I_{boost}. Durch diesen Stromanstieg wird gewährleistet, dass sich das Magnetfeld zum Anheben der Injektornadel

(handwritten: Entladung: descarga.   Stromstärke: amperaje)

11

Verzögerung: retraso, retardo.

schnell aufbaut und es zu keinen unerwünschten Verzögerungen beim Einspritzbeginn kommt.

- **Pick-up Intervall**: Nach dem Booster-Intervall, für $t > T_{boost}$, wird die Injektorspannung auf 12 V reduziert, was der Spannung des Fahrzeugakkus entspricht. Der Injektorstrom fällt daher auf den Wert $I_{pick-up}$.

- **Holding Intervall**: Für $t > T_{hold}$ hat der Anker seine maximale Höhe erreicht und liegt am Magnetkern an. Um Energie zu sparen, wird der Injektorstrom auf I_{hold} gesenkt.

Aktiviert wird das Stromprofil $I(t,T_i)$ durch eine fallende Flanke im Steuersignal $s(t,T_i)$ zum Zeitpunkt des Ansteuerbeginns SOI (start of injection). Die Injektorspannung wird nach der Ansteuerdauer T_i zurückgesetzt. Die Variablen T_i und SOI legen somit den Zeitpunkt und die Dauer der Kraftstoffinjektion fest. festlegen: fijar, determinar,
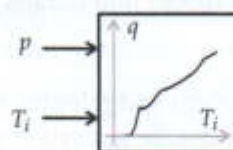
2.2.3 Injektorkennlinie

Der Injektor dosiert die Kraftstoffmenge q in Abhängigkeit der Ansteuerdauer T_i und des Kraftstoffdrucks

$$p = p_{rail} - p_{br},\qquad(2.1)$$

der sich aus dem Systemdruck p_{rail} in der Common-Rail und dem Brennraumdruck p_{br} zusammensetzt gemäß Abbildung 2.4. Das injektorindi-

zusammensetzen (sich): ensamblar, montar.   gemäß: de acuerdo a. Zusammenhang: nexo, relación, conexión. verhalten: comportamiento, conducta,

Abbildung 2.4 Zusammenhang zwischen Eingangs- und Ausgangsgrößen des Magnet-Injektors

viduelle Betriebsverhalten wird durch die Injektorkennlinie wiedergegeben. Außer von den Eingangsgrößen ist das Verhalten des Injektors auch von der Injektortemperatur Θ_{inj} und der Anpresskraft F_n, mit der

12

das Ventil auf dem Zylinderkopf montiert ist, abhängig. Die Zustands-
größen Θ_{inj} und F_n werden im Folgenden, wenn nicht explizit erwähnt,
als konstant angenommen. Die Injektorkennlinie beschreibt die Abhän-
gigkeit der Ist-Menge q_{ist} von der Ansteuerdauer T_i bei einem konstan-
ten Kraftstoffdruck p. Die Kennlinien von drei Injektoren desselben Typs
und gleicher Charge für die Kraftstoffdrücke $p = 4\,MPa$, $p = 12\,MPa$ und
$p = 20\,MPa$ im Ansteuerbereich $T_i = 0\ldots 2\,ms$ sind in Abbildung 2.5 dar-
gestellt. Vor allem im Kleinmengenbereich $q < 5\,mg$ treten erhebliche Ab-
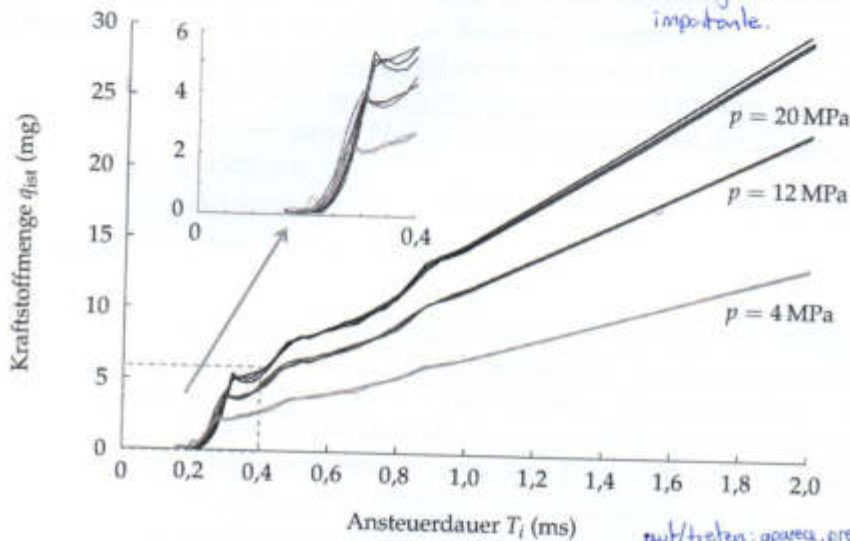
Abbildung 2.5 Gemessene Injektorkennlinien von jeweils drei baugleichen In-
jektoren für verschiedene Kraftstoffdrücke

weichungen im Betriebsverhalten verschiedener Injektoren auf. Die rela-
tiven Abweichungen in der eingespritzten Kraftstoffmenge von Injektor
zu Injektor können dabei über 30% betragen [43].

2.3 Injektorkalibrierung

Die Ursachen der Abweichungen in der Kraftstoffeinspritzung liegen
in den produktionsbedingten Bauteiltoleranzen und der Montage des
Magnetspulen-Injektors. Diese beträchtlichen Exemplar-Streuungen wer-
den durch Verschleißerscheinungen aufgrund der starken Belastung

13

3 Injektionsvorgang

Eine detaillierte Auseinandersetzung mit dem Ablauf des Einspritzvorgangs bildet die Grundlage für die spätere Analyse der Körperschallsignale. Dabei ist ein Verständnis der Vorgänge im Injektor während der Kraftstoffeinspritzung notwendig, um die Körperschallemissionen interpretieren zu können. Der Vorgang der Kraftstoffinjektion wird daher im Folgenden genauer betrachtet. Anhand eines speziellen Messaufbaus, der in Kapitel 4 vorgestellt wird, ist die Aufnahme der Nadelbewegung über der Zeit möglich. Anhand dieser Nadelbewegung lässt sich der Einspritzvorgang detailliert erklären.

3.1 Ablauf

Der von der Ansteuerdauer T_i abhängige Strom $I(t, T_i)$ stellt die Stellgröße des Magnet-Injektors dar. Über die Spule und den Eisenkern wird ein Magnetfeld aufgebaut, welches den Anker anzieht. Daraus resultiert die Nadelbewegung $h(t)$. Der Einspritzvorgang lässt sich anhand der Nadelbewegung in verschiedene Phasen unterteilen. Während der Kraftstoffeinspritzung treten bestimmte Ereignisse auf, die für die spätere Schätzung der tatsächlichen Einspritzdauer bzw. Öffnungsdauer relevant sind. Diese sollen nun anhand von Abbildung 3.1 in den folgenden Teilabschnitten genauer beschrieben werden.

3.1.1 Ansteuerbeginn (SOI)

Wie in Abschnitt 2.2.2 gezeigt, wird zum Zeitpunkt des Ansteuerbeginns t_{SOI} (SOI = start of injection) der Steuerstrom $I(t)$ durch Anlegen einer Injektorspannung $U(t)$ aktiviert. Das Magnetfeld im Injektor beginnt über den Eisenkern eine Magnetkraft F_m aufzubauen.

3.1.2 Einspritzbeginn (BOI)

Die bestromte Zylinderspule bildet über den Eisenkern ein Magnetfeld aus, das den Anker anzieht. Nach der Anzugsdauer T_{an} überwiegt die

entgegen: contrariamente, en oposición
daraufhin: a continuación,
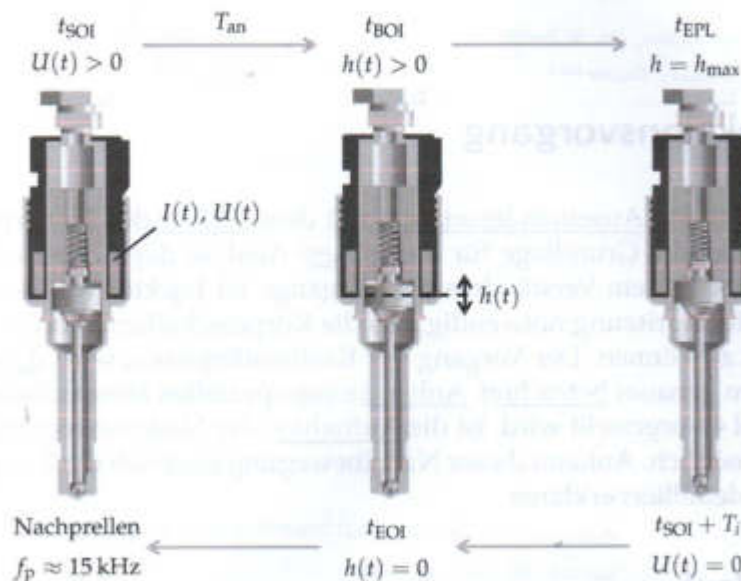
Abbildung 3.1 Veranschaulichung der Nadelbewegung während des Einspritz-
vorgangs

Magnetkraft F_m die <u>entgegen</u> gerichtete Federkraft F_f und der Magnetan-
ker beginnt die Nadel zu heben. Die Injektordüse wird <u>daraufhin</u> frei-
gegeben und der unter Druck stehende Kraftstoff strömt aus. Der Zeit-
punkt t_BOI (BOI = begin of injection) steht für den Einspritzbeginn. Der
Vorgang des Anhebens der Injektornadel ist druckabhängig, da sich die-
se entgegen des ausströmenden Kraftstoffs bewegt, siehe Abbildung 3.2
(oben). Die Druckabhängigkeit ist am Verlauf des Nadelhubs für die
Kraftstoffdrücke $p = 4\,\mathrm{MPa}$, $p = 12\,\mathrm{MPa}$ und $p = 20\,\mathrm{MPa}$ bei einer festen
Ansteuerdauer $T_i = 0{,}5\,\mathrm{ms}$ dargestellt. Im Falle eines geringeren Kraft-
stoffdrucks hebt sich die Injektornadel zeitlich früher und schneller, was
an der Steigung des Hubverlaufs zu sehen ist. Für sehr geringe Ansteuer-
dauern im Kleinmengenbereich ist t_BOI abhängig von der Ansteuerdauer
T_i. Dies ist im unteren Diagramm von Abbildung 3.2 zu sehen.

Anschlag: choque, tope, golpe.

3.1.3 <u>Anschlag</u> am Magnetkern (EPL)

Überwiegt die Magnetkraft die Federkraft ausreichend lange, so trifft der
Anker auf den Magnetkern. Da die Injektornadel nicht starr mit dem An-
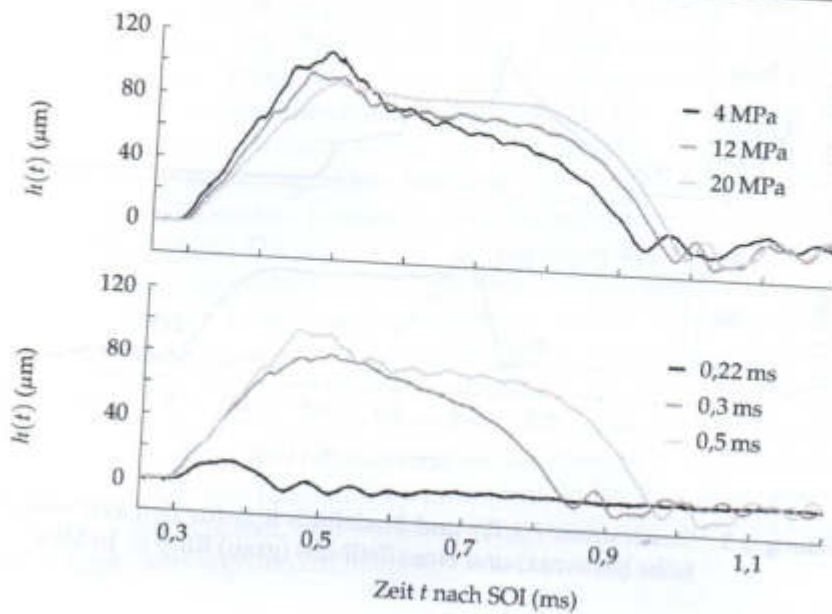
Abbildung 3.2 Nadelhub $h(t)$ für verschiedene Drücke p bei $T_i = 0{,}5\,\text{ms}$ (oben) und für verschiedene Ansteuerdauern T_i bei $p = 4\,\text{MPa}$ (unten)

ker verbunden ist, kommt es zu einem <u>Überschwinger</u>, welcher deutlich im Verlauf von $h(t)$ zu sehen ist. Die Injektornadel bewegt sich so weit nach oben, bis die Federkraft überwiegt und die Nadel wieder zurück auf den Anker presst. Der Zeitpunkt des Ankeranschlags wird im Folgenden stets mit t_{EPL} (EPL = end of pintle lift) bezeichnet. Wie in Abbildung 3.2 zu erkennen, ist t_{EPL} druckabhängig.

3.1.4 Einspritzende (EOI)

Zum Zeitpunkt $t_{SOI} + T_i$ wird die Injektorspannung $U(t)$ zurückgesetzt. Der Injektorstrom und somit auch die Magnetkraft <u>fallen ab.</u> Aufgrund der überwiegenden Federkraft beginnt die Nadel auf die Düse zu fallen. Zum Zeitpunkt t_{EOI} (EOI = end of injection) des <u>Auftreffens</u> der Nadel auf der Düse wird der Einspritzvorgang beendet. Abbildung 3.3 zeigt den Injektorsteuerstrom $I(t, T_i)$ und die resultierende Nadelbewegung $h(t)$ für $T_i = 0{,}28\,\text{ms}$ und $T_i = 0{,}8\,\text{ms}$ bei jeweils $p = 10\,\text{MPa}$. Im Falle von $T_i = 0{,}28\,\text{ms}$ tritt das EPL-Ereignis nicht auf. Deutlich zu erkennen ist eine dominante Schwingung im Verlauf von $h(t)$
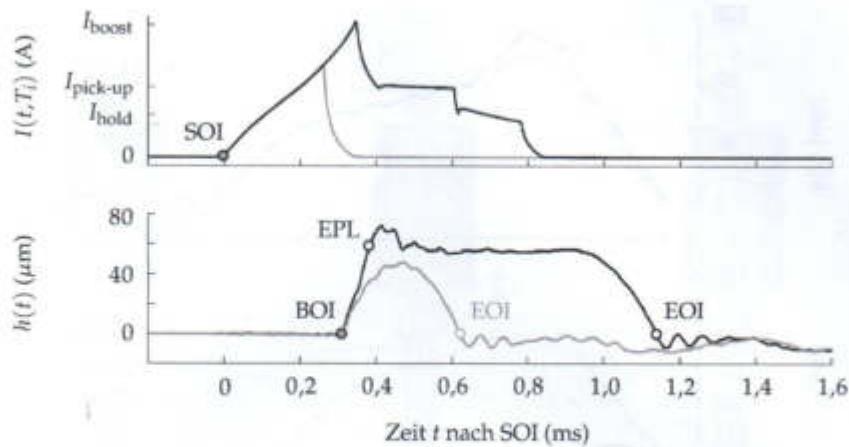
Abbildung 3.3 Steuerstrom $I(t,T_i)$ und Nadelhub $h(t)$ für den Fall eines Voll-
hubs (schwarz) und eines Teilhubs (grau) für $p = 10\,\text{MPa}$

nach dem Einspritzende. Es handelt sich hierbei um das Nachprellen der
Injektornadel auf der Düse. Im Falle des vorrangig in dieser Arbeit be-
trachteten Magnet-Injektors beträgt die Frequenz dieser Schwingung ca.
$f_\text{P} \approx 15\,\text{kHz}$. Die Prellfrequenz f_P spielt in den weiteren Betrachtungen
dieser Arbeit eine sehr entscheidende Rolle.

Die aus den Nadelbewegungen $h(t)$ ermittelten Auftrittszeitpunkte t_BOI,
t_EPL und t_EOI der eben beschriebenen Ereignisse sind in Abhängigkeit
der Ansteuerdauer T_i in Abbildung 3.4 aufgetragen, wobei der Kraftstoff-
druck p konstant ist. Es sind deutliche Zusammenhänge in der (T_i, t)-
Ebene zu erkennen. Während die Zeitpunkte $t_\text{BOI}(T_i)$ und $t_\text{EPL}(T_i)$ ab be-
stimmten Ansteuerdauern konstant sind, ist an t_EOI der typische Verlauf
der Injektorkennlinie aus Abbildung 2.7 zu erkennen. Die Injektorkenn-
linie kann bezüglich der Ansteuerdauer in einzelne Bereiche unterteilt
werden. Diese werden nun näher erläutert.

3.2 Teilhub und Vollhub

Bei einem konstanten Kraftstoffdruck lässt sich der gesamte Ansteuerbe-
reich in drei Teilbereiche unterteilen. Für alle folgenden Betrachtungen
gilt daher stets die Fallunterscheidung:

[26]   (2007) Singh Pundir, A. [online]. Disponible: https://www.alliedmarketresearch.com/gasoline-direct-injection-system-gdi-market

[27]   (2021) Hantek Electronic Co., Ltd [online]. http://www.hantek.com/products/detail/13170

[28]   "Hantek CC-65 current clamp data sheet," Hantek Electronic Co., Ltd, Qingdao, China.

APPLICATOIN NOTES

1. In the case of DC current, the output is positive when the current flows from the upside to the underside of the clamp. The red banana plug is positive.

2. In the case of DC current measurement, the current clamp can not possibly zero properly, due to the hysteresis effect. To eliminate this, open and close the jaws several times and then zero again.

APPLICATION SAFETY

1. The voltage of the conductor measured with the current clamp should be little than 300V DC or 240V rms AC,

2. We are forbidden to measure the conductor whose insulation has break down or wear down, in order to avoid doing harm to us.

SPECIFICATIONS GENERAL

Captured Conductor diameter Size: 9mm maximum

Low Battery Indicator: Red LED lighting

Operating Temperature: 0℃ to 50℃, 70% R.H.

Storage Temperature: -20℃~+70℃, 80% R.H.

Battery type: 9V DC, NEDA 1604, 6F22, 006P

Battery Life: 100 hours typical with carbon-zinc

Weight: 250 gram typical

Dimensions: 195 mm (H) x 70 mm (W) x 33 mm (D)

Output: wire with a straight banana plug or BNC connector

ELECTRICAL (At 23±5℃, 70% R.H. maximum)

Efficient Measure Range:

1mV/10mA: 10mA to 2 A, multimeter range is set to DC or rms AC 200mV

1mV/100mA: 100mA to 20 A, multimeter range is set to DC or rms AC 200mV

20A to 65 A, multimeter range is set to DC or rms AC 2V

Precision:

System Accuracy: The Precision of the current clamp add the Precision of the digital multimeter or scope.

For example, if the conductor carries 100 mA current, the current clamp range will be set to 1mV/10mA and the output is 10mV. Assuming that the Precision is 1.5%, the output is10.15mV maximum and 9.85mV minimum.

If the digital multimeter connected with current clamp is set to 200mV and the Precision is 0.5%, the readings is 10.20mV maximum and 9.8mV minimum. (10.15mV×(1+0.5%)=10.20mV, 9.85mV×(1-0.5%)=9.80mV)

Accuracy:

DC Current: 1m/10mA

±(1.5%±5mA) 10mA ~ 20A

DC Current: 1mV/100mA

±(2%±20mA) 100mA ~ 40A

±(4%±0.3A) 40A ~ 65A

AC Current: 1mV/10mA

±(2%±30mA) 100mA ~ 10A (40Hz ~ 2KHz)

±(4%±30mA) 100mA ~ 10A (2KHz ~ 10KHz)

±(6%±30mA) 100mA ~ 10A (10KHz ~ 20KHz)

±(8%±30mA) 10A ~ 15A (40Hz ~ 20KHz)

AC Current: 1mV/100mA

±(2%±30mA) 100mA ~ 40A (40Hz ~ 1KHz)

±(4%±30mA) 100mA ~ 40A (1KHz ~ 2KHz)

±(6%±30mA) 100mA ~ 40A (3KHz ~ 5KHz)

±(8%±0.3A) 40A ~ 65A (40Hz ~ 20KHz)

Load Resistance: 10kΩ Typical

Temperature coefficient: 0.1×(defined accuracy) per degree C (0℃to18℃, 28℃to50℃)

SAFETY INFORMATION

The instrument complies with class II, overvoltage CAT II of the EN 61010-1, and EN 61010-2-032 standards. The pollution of Current clamp is degree 2 in accordance with IEC 664 indoor use. If the equipment is used in a manner not specified, the protection provided by the equipment may be impaired.

This product complies with the requirements of the following European Community Directives: 89/336/EEC (Electromagnetic Compatibility) and 73/23/EEC (Low voltage) as amended by 93/68/EEC (CE marking).

[29]   SAE International. (2013). [Online]. Disponible: https://doi.org/10.4271/AS40401

[30] SAE International. (2018). [Online]. Disponible: https://doi.org/10.4271/J2713_201808

[31]   (2021) Arduino Official Store. [Online]. Disponible: https://store.arduino.cc/usa/arduino-uno-rev3

[33]   Zhang, X., Palazzolo, A., Kweon, C., Thomas, E. et al., "Direct Fuel Injector Power Drive System Optimization," SAE *Int. J. Engines* 7(3): 2014, doi: 10.4271/2014-01-1442.

conditions. Injection timing is typically retarded as engine load is increased. Pilot injection strategies have been widely used to reduce combustion noise and improve engine idling. Typical pilot injection quantities range from about 1 to 5 mg for advanced diesel engines. Fast pintle movement is critical in multiple injections because multiple injections should be made in each cycle, while the first injection minimally influencing the following injection. Increasing the current ramp rate of the solenoid coil would increase the pulling motion of the pintle, given that in theory the magnetic force is proportional to the square of supply current to the injector. In reality, however, it will reach a point where increasing the magnetic force rate is impeded by the limited current increasing rate, which is determined by inductance. Other factors such as hydraulic hysteresis, magnetization hysteresis and saturation of the magnetic material also impose limitations on pintle response time.

Repeatable fuel injection leads to less shot-to-shot variations of fuel injection quantity and injection rate. Repeatable fuel injection can be achieved when the pintle moves rapidly to a desired position (typically fully open position) to avoid random fuel injection quantity and rate. Higher current may ensure the repeatability of the injection by creating more than enough magnetic force to hold the pintle at its desired position.

However, higher current would create more heat in the solenoid coil, which increases the temperature of the surrounding magnetic material, thus degrading the magnetic material property, i.e. lowered saturation limit and more hysteresis. Injector performance is highly dependent on temperature in the solenoid which was observed during our injector dry test in which the threshold current increased significantly as the injector body temperature was increased. Ideally, an injector power drive system should use the least amount of current, while achieving fast response and repeatable pintle movement.

In the past, research has been performed mostly in two different areas: (1) numerical modeling of fuel injector dynamics and (2) power drive design for direct fuel injectors that typically deal with power strategies to the injector solenoid coil. Numerical simulation codes for diesel injection systems were developed by various authors [1, 3, 4]. All these codes include or partially include coupled multi-physics model of hydraulics, fluid, mechanical and electromagnetic. Digesu's model coupled one-dimensional hydraulics with a two-dimensional (2D) axisymmetric electromagnetic finite element code to provide magnetic force in terms of exciting current and working gap [1]. In his study, needle lift, pressures (pipes, control volume, rail and accumulative volume) and flow rate were matched with experimental data. The authors investigated how pipe design affects the needle lift profile due to the hydraulic wave, and showed a moderate dependence on pipe parameters. Ficarella developed a code to evaluate the instability phenomena in a common rail injection system to predict injection characteristics [2]. It was found that the most challenging part of the modeling lies in the control valve whose

residual motion causes re-opening of the valve after a pilot injection. Coppo's code was also based on a common-rail injector with a combined pipe, fluid, mechanical and electromagnetic model [3, 4]. Experimentally measured exciting current signals were input into the model, which included magnetic saturation effects. Good agreement was observed between calculated and experimental profiles of rail pressure, needle lift, and injection quantity. Hu developed a one-dimensional (1D) model to correlate the pressure profile with the measured one in which no electromagnetic model was used and a measured magnetic force profile was employed in the 1D model [5]. Ando developed a 1D electromagnetic model coupled with simplified rigid body and fluid models [6]. The main emphasis of this model was on the electromagnetic effects including magnetization time lag.

For power drive optimization, Tsai designed a three-stage electrical drive for high-pressure gasoline direct injection (GDI) injectors. In the design, the first two stages used a power Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) to switch between power supplies, while the 3rd stage used pulse-width-modulation (PWM) control. The author provided the current amperages for each stage (12/5/3A) but did not fully explain why those values were chosen. The author observed that the pintle closed faster with the 3rd stage PWM control than with a direct voltage control method [7]. Tsai, in another paper, performed a parametric study to investigate the effect of voltage, current, frequency and fuel pressure on injection quantity [8].

From the literature review, it was found that there is a lack of integral method to model from injector power source to pintle dynamics, as well as driver design details for controlling fuel injectors. The goal of this paper is to find the optimal power strategy to drive fuel injectors, specifically in our case, it is a gasoline direct fuel injector (GDFI). The methodology of fulfilling this goal is as follows: using electromagnetic-mechanical-fluid coupled simulation to reveal the impact of injector current shapes on injection dynamics, as well as to achieve fast response and repeatability of pintle movement with the least power consumption or the minimal driving current. Through fluid and mechanical model, the electromagnetic force requirement is formed; and through electromagnetic circuit analysis, the corresponding current requirement, which contains optimal parameters for multi-stage current shape, is established under different supply pressures. Finally, the simulation results are validated using a photo detector system.

The paper is organized as follows: Firstly, injector power drive system basics, electromagnetic system, mechanical system and fluid system are introduced separately in four different sections. These four sections are the building blocks for the entire system. Secondly, a coupled simulation is run using one of the power strategies (current shapes) to reveal the detailed dynamics of pintle movement, flow rate and pressure forces on the pintle under different supply pressures. The insight of current requirement is gained after this run of simulation.

Thirdly, two more power strategies (current shapes) are applied to illustrate the failure of maintaining the injection repeatability if the current requirement is violated. Fourthly, the optimal power strategy is discussed in detail, which includes the reason of using the multi-stage current shape and its optimal parameter settings. Lastly, the simulation results are validated using a photo detector system.

Injector Power Drive System Basics

Direct fuel injectors are designed for fast response under high pressures, which requires careful selection of their power drives. The power input to the injector coil creates a magnetic force to open the pintle valve. Since magnetic force is proportional to the square of supply current, the pintle will only move when supply current reaches certain levels where magnetic force overcomes opposing forces (including pressure, spring and friction forces). The threshold of the pintle movement depends on the threshold level of the excitation current in the solenoid coil. Therefore, how fast the pintle opens depends on how fast the current level can reach the required current threshold. However, there is always a time constant for current rise due to the inductance in the solenoid coil. The current rising rate needs to be fast to achieve fast opening of the pintle.

The governing equation for a simple resistor-inductor circuit (RL circuit) is

$$V = i \cdot R + L \cdot \frac{di}{dt}$$

(1)

where V is the voltage across the coil; R is the resistance; and L is the inductance. The step voltage input response of the current is obtained by solving (1) to obtain:

$$i(t) = \frac{V}{R} \cdot (1 - e^{-\frac{R}{L}t}) = \frac{V}{R} \cdot (1 - e^{-\frac{1}{\tau}t})$$

(2)

where $\frac{L}{R}$ is the electrical subsystem time constant τ. The inductance was measured to be 1.9 millihenry (mH) at 1 kHz and 3.9 mH at 0.12 kHz at the maximum air gap. In our case,

$$\tau = \frac{L}{R} = \frac{(1.9 \sim 3.9)mH}{1.5\Omega} = 1.26 \sim 2.6ms$$

(3)

This means that it takes 1.26~2.6 ms to reach 63.2% of the steady-state current $\frac{V}{R}$; where R, L are measured values.

Figure 1 illustrates the times to reach 8.5 A for different voltage inputs. The reason for 8.5 A in the analysis will be discussed in the later section where the magnetic force requirement is specified.
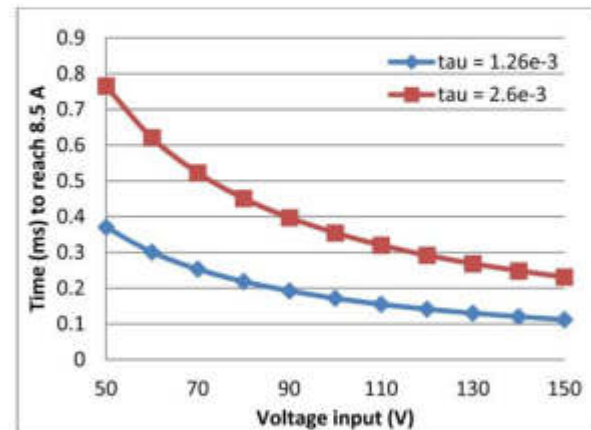
Figure 1. Time to reach 8.5 A with different time constants

A PWM servo-power amplifier was used to provide current to the fuel injectors. Nowadays, PWM amplifiers become the preferred way of driving fuel injectors. Compared with power MOSFETs, PWM power amplifiers are more power efficient and easier to use.

A PWM power amplifier consists of a DC power supply, capacitor (for short pulsation application), and control signal. The DC power supply delivers power to the PWM via the capacitor. Through the PWM frequency and duty cycle, the output current (or voltage) is modulated to match the input control signal shape at higher power level. For instance, in our system, 0~10 volts control signal was fed into PWM and the output from the PWM was 0~15 A current shape that closely followed the input control signal. This approach is extremely helpful because real applications utilize complex control strategies of the fuel injectors. One of the main benefits of a PWM amplifier is that one can easily adjust the control signal amplitude and time duration with multiple stages to regulate the multiple stage current output to the fuel injector. Another benefit of the PWM control was discussed in Tsai's paper [7] which showed that the last stage PWM control could shorten the closing delay time of the fuel injector pintle. Figure 2 illustrates the injector power and control system utilized in the present paper.
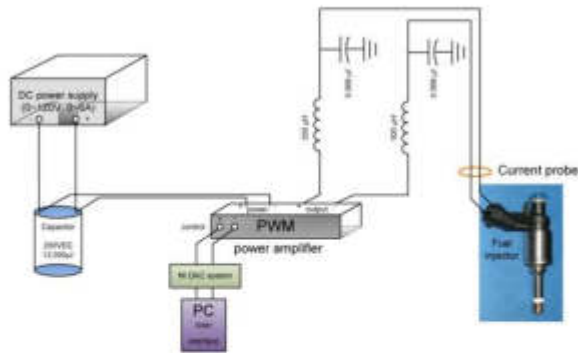
Figure 2. Injector power and control system

Electromagnetic System

Electromagnetic analysis is an essential part of the multi-physics model. The fuel injector used in this study is a BOSCH GDI injector (HDEV 5.2), which is a plunger type solenoid electromagnetic device. A coil is imbedded inside the injector body, with a magnetic conducting material forming a closed magnetic path when the coil is energized. The motion of the pintle is caused by the magnetic field created by the current flowing in the coil.

The electromagnetic model includes an electrical circuit model for the injector coil and a magnetic model for magnetic flux and force calculation.

The electrical circuit for the injector coil is basically a RL circuit. The PWM power amplifier was not modeled in this study due to the circuit's low frequency nature compared to the switching frequency of PWM, which is 22 kHz in our case. For direct comparison, the time constant of the injector RL circuit is about 2 ms, while the switching time is only 0.02 ms. Thus, the model was simplified to use three power sources: one voltage source with a time-controlled switch, together with two pulse current sources, to obtain a two-stage current shape. In the actual system, a two-stage control voltage signal is fed into the PWM power amplifier and then the output is a two-stage current shape. The electrical properties of the fuel injector coil and PWM power amplifier are listed in Table 1.

The electrical circuit was modeled in Maxwell circuit editor, as shown in Figure 3. It included a voltage source with a switch controlling on and off time, a pulse current source for the peak current, and another pulse current source for the holding current. This circuit model was imported into a 2D axisymmetric finite element code Maxwell as external excitation source for the coil. The time constant τ in the RL circuit (see equation 3) determines the maximum current increasing rate. The inductance, which is represented as winding1 in the circuit, changes as the geometry changes, which takes into account of the pintle motion in the magnetic circuit. Figure 4 shows the inductance profile with respect to the change of air gap. The inductance from Maxwell was calculated through the

relationship $L = \dfrac{d\lambda}{di}$ at each time step in the transient analysis. It is shown that the change of the inductance due to pintle motion counts for about 10% of coil's total inductance.

Table 1. Electrical properties of the fuel injector coil and PWM power amplifier used in this study

Resistance		$1.5\,\Omega$
Inductance		1.9 mH at 1 kHz; 3.9 mH at 0.12 kHz
Number of turns		160
PWM switching frequency		22 kHz
PWM power range	Peak current	50 A
	Continuous current	25 A
	Supply voltage	40-190 VDC
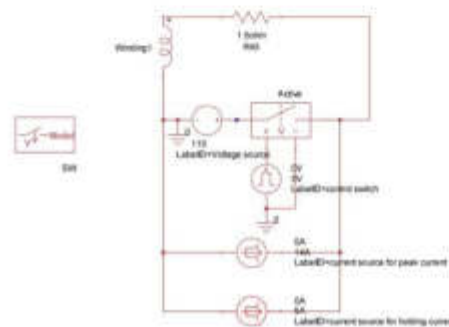
Figure 3. Electrical circuit in Maxwell circuit editor.
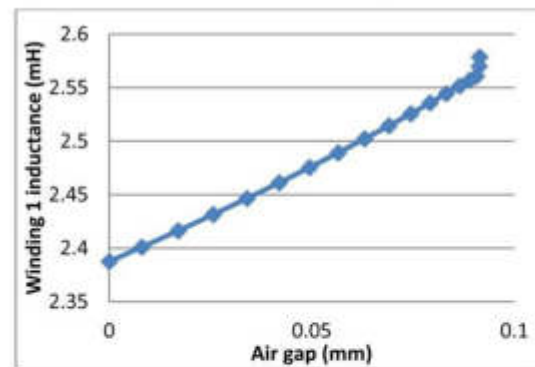
Figure 4. Inductance at different air gaps from Maxwell.

This inductance plays an important role in determining the current rise time. A 1D magnetic analysis is provided below to compare with the inductance calculation from Maxwell.

From Ampere's law $\displaystyle\oint_c \overline{H} \times d\overline{l} = I_{net}$

$$H_1 \cdot \Delta x + H_2 \cdot w + H_3 \cdot w = N \cdot I$$

(4)

332

where, the parameters in Equation (4) are illustrated in Figure 5.
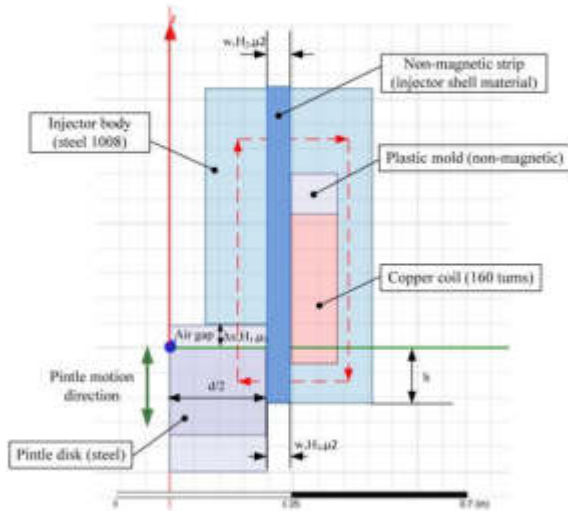
Figure 5. Flux path in the magnetic circuit analysis

Apply Gauss M around pintle disk, $\oint_s \vec{B} \times d\vec{a} = 0$

which yields

$$\mu_1 \cdot H_1 \cdot \frac{\pi}{4} d^2 - \mu_2 \cdot H_3 \cdot \pi d \cdot h = 0$$

(5)

Assume no fringing and leakage,

$$H_2 \cong H_3$$

(6)

Also, the permeability in the air μ_1 and the one in the non-magnetic strip μ_2 are similar: $\mu_1 \cong \mu_2$;

Apply flux linkage $\lambda = N \cdot \int_s \vec{B} \times d\vec{a}$ on the top plane of the pintle disk,

$$\lambda = \left[\frac{N^2 \cdot \mu_1 \cdot \frac{\pi}{4} \cdot d^2 \cdot h}{\frac{d \cdot w}{2} + \Delta x \cdot h} \right] \cdot i$$

(7)

$$L = \frac{N^2 \cdot \mu_1 \cdot \frac{\pi}{4} \cdot d^2 \cdot h}{\frac{d \cdot w}{2} + \Delta x \cdot h}$$

where, inductance

The geometric parameters were obtained by measuring the dimensions with a cut-open injector. Table 2 lists the parameters used in the calculation. The calculated inductance (2.4 ~ 2.6 mH) correlates well with the measured inductance, which ranged from 1.9 mH at 1 kHz to 3.9 mH at 0.12 kHz.

Table 2. Parameters for inductance calculation

Non-magnetic strip width w [mm]	0.8
Pintle disk diameter d [mm]	10
Pintle height h [mm]	4
Air gap Δx [mm]	0 ~ 0.1
Permeability μ_1 [H/m]	1.256e-6
Number of coil turns	160
Calculated inductance using eqn (7) [mH]	2.4 ~ 2.6

Figure 6 shows elapsed times for different current targets at the calculated inductances with the same 110 V voltage source using Equation (2).
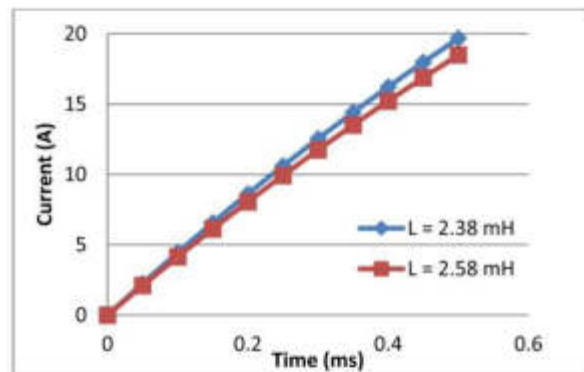
Figure 6. Current level vs. elapsed times at the calculated inductances

Figure 7 compares the calculated current shape from Maxwell and the measured one with the 1st stage on-time of 400 µs and 2nd stage on-time of 600 µs. The calculated current profile, especially the rise time (which is the key to the opening delay time), fits well with the measured current.
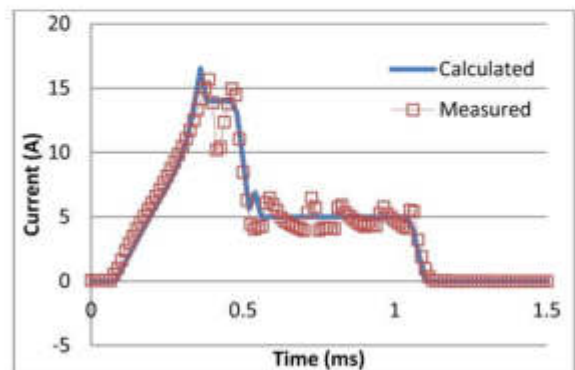
Figure 7. Maxwell calculated and measured current shapes

Zhang et al / SAE Int. J. Engines / Volume 7, Issue 3 (August 2014)

In the electrical circuit simulation, the current limit was set by the PWM power amplifier and the voltage source was turned off when the current reached the current limit. This was done to capture the characteristics of the rising curve and the current limit. Another pulse current source was used to provide the 2nd stage current. In practice, once a desired current shape is determined, it is relatively easy to adjust the current magnitude and duration by changing the DC power supply voltage, control signal amplitude, current limit and command signal's time durations.

In Figure 8, the 2D axisymmetric model in Maxwell shown in Figure 5 was simulated in transient. Magnetic density B vector plot was shown at time t = 0.4 ms at which the 1st stage current reached its steady-state.
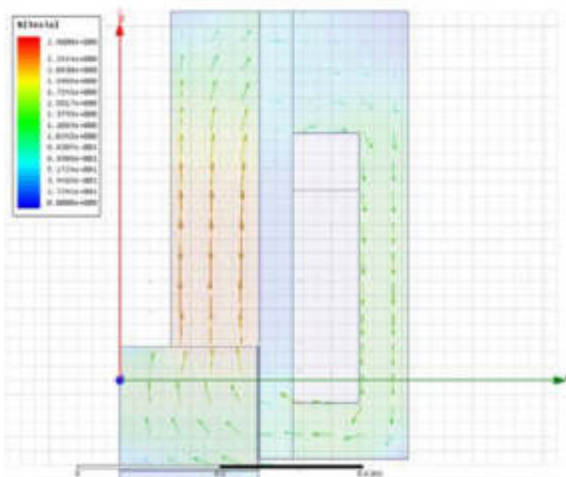
Figure 8. Magnetic flux density vector plot at t = 0.4 ms from 2D Maxwell transient with the two-stage current shown in Figure 7

The magnetic field density was calculated using the ampere's law and the constitutive relationship between flux density B and magnetic field strength H: $B = \mu \cdot H$. Then, Maxwell force law was used to calculate the magnetic force, given the magnetic flux density, permeability and area.

$$F_{mag} = \frac{B^2 A}{2\mu} = \frac{(\mu \cdot H_1)^2 A}{2\mu}$$

(8)

where, B is the flux density at H_1; A is the area on the upper surface of pintle disk; μ is the permeability of air.

Solving Equations (4, 5, 6) for H_1, then substituting H_1 into Equation (8), the magnetic force becomes

$$F_{mag} = \frac{\mu \cdot N^2 \cdot h^2 \cdot A}{2(\frac{d \cdot w}{2} + \Delta x \cdot h)^2} \cdot i^2$$

(9)

The magnetic force was also calculated using the finite element code Maxwell. Figure 9 compares the magnetic forces calculated with Equation (9) and Maxwell at different air gaps and current levels.
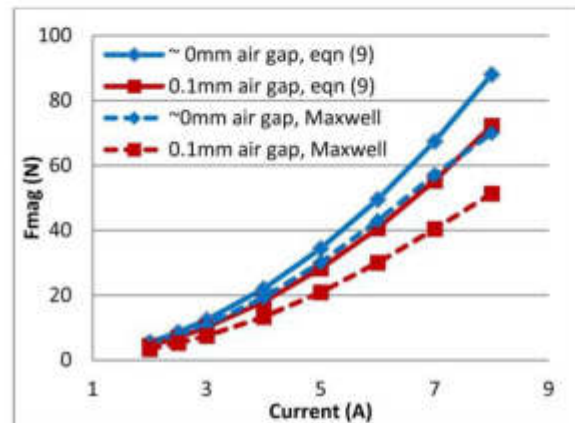
Figure 9. Calculated magnetic forces at different air gaps and current levels with Equation (9) and Maxwell

The 2D axisymmetric finite element code Maxwell considered the non-linearity of the B-H curve of the magnetic material and the fringing loss. Therefore, the magnetic force calculated using Maxwell was less than the theoretical force calculated from Equation (9) as excitation current increased. However, the two force curves are close enough to justify the results from Maxwell.

Figure 9 shows the magnetic forces calculated using finite element code Maxwell with the current shape shown in Figure 7 under three different conditions: with pintle motion and without pintle motion at extreme pintle positions.

Figure 10 shows that the magnetic force with moving pintle (solid) falls in-between the magnetic forces at extreme pintle positions without pintle motion (dotted and dashed). For simplicity, the magnetic force with non-moving pintle at its minimum air gap (~0mm) was used in the following coupled analysis assuming that the magnetic force is only a function of time given a certain current shape.

334
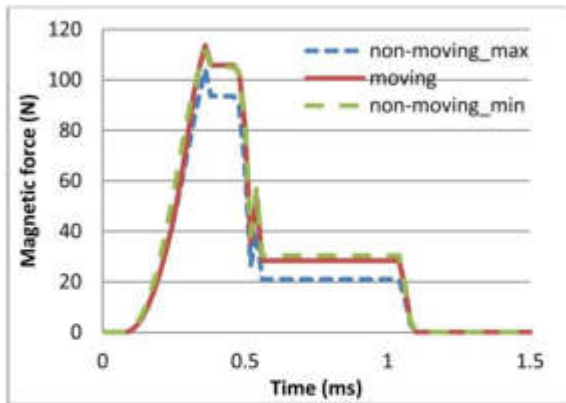
Figure 10. Magnetic force profile with and without pintle movement using Maxwell

Besides current and air gap, another important factor that affects the magnetic force is the magnetic saturation. The magnetic saturation imposes the maximum resultant magnetic force due to the flux saturation in the magnetic circuit. Therefore, the magnetic force cannot increase infinitely with the increasing current. Figure 11 shows the magnetic forces at different currents, pintle positions, and saturation levels. Amp*turns is the multiplication of current (amp) and coil's number of turns. In our case, the total number of turns was 160.

The resultant magnetic forces of less than 1000 amp-turns were similar for different saturation materials. However, the difference of the resultant magnetic forces could be as much as 15 N at the same current (1500 amp*turns) and air gap (0.1mm), only with the change of magnetic material from 1.5 Tesla (T) saturation to 2 T saturation. Therefore, in order to obtain the largest possible magnetic force, it is preferred to use high saturation material in the magnetic path.
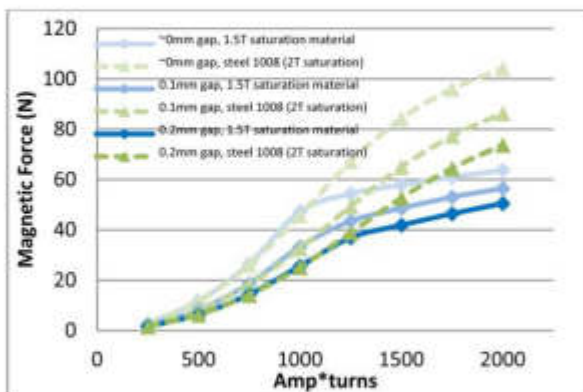
Figure 11. Saturation study for different materials and air gaps

Figure 12 shows the magnetic force versus the current (Amp*turns) using steel 1008 with 0.1 mm air gap. 1250 amp*turns (7.8 A at 160 turns) created about 50 N magnetic force.
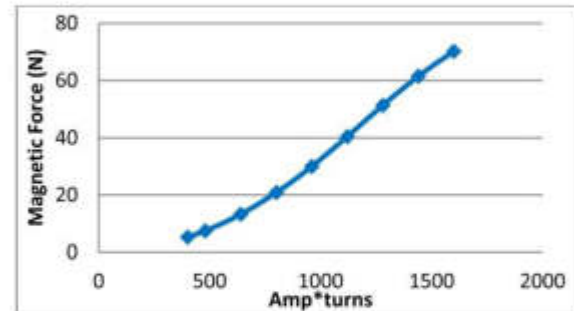
Figure 12. Magnetic force vs. current at 0.1 mm air gap with steel 1008 (2T saturation)

Mechanical System

The mechanical system model predicts the rigid body motion of the pintle. The force illustration is shown in Figure 13. At the beginning, the pintle sits on the valve seat where there are five forces acting on it: pressure, gravity, spring, contact friction and supporting forces. The pressure force is due to the pressure on the pintle. Spring force is due to the return spring's compression. Contact friction force takes into account of the initial friction due to its contact with the valve seat. And the support force from valve seat, represented as spring and damper, is the equal and opposite force of all other forces when pintle rests on the pintle valve. The magnetic force increases as the current in the coil increases. In the event of pintle movement, magnetic force overcomes pressure, gravity, spring and contact friction forces. The pintle motion is limited by the stator, which is represented as another set of spring and damper at the distance of 0.09mm. This distance was measured with a miniature dial indicator in the event of pintle motion under the condition where there was no fluid inside the fuel injector.
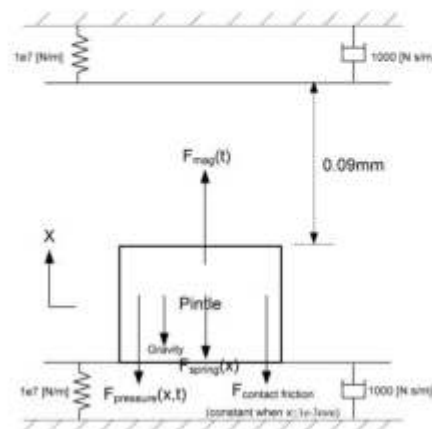
Figure 13. Mechanical system of the pintle with force illustration

335

The non-moving pintle at ~0mm air gap with steel 1008 and power strategy shown in <u>Figure 7</u> was used to create the magnetic force profile shown in <u>Figure 14</u>. The magnetic force $F_{mag}(t)$ can be expressed by segmenting the force profile and curve-fitting at different segments:

$$F_{mag}(t) = \begin{cases} 0, & 0ms < t < 0.08ms \\ 110.6 \cdot e^{-\left(\frac{t-0.3821}{0.1241}\right)^2} + 20.68 \cdot e^{-\left(\frac{t-0.2471}{0.08861}\right)^2}, & 0.08ms < t < 0.37ms \\ 105, & 0.37ms < t < 0.47ms \\ -914.9 \cdot t + 536.9, & 0.47ms < t < 0.55ms \\ 30.35, & 0.55ms < t < 1.05ms \\ -528.5 \cdot t + 579.1, & 1.05ms < t < 1.09ms \\ 0, & t > 1.09ms \end{cases}$$

(10)

where t in the expression is in [ms].
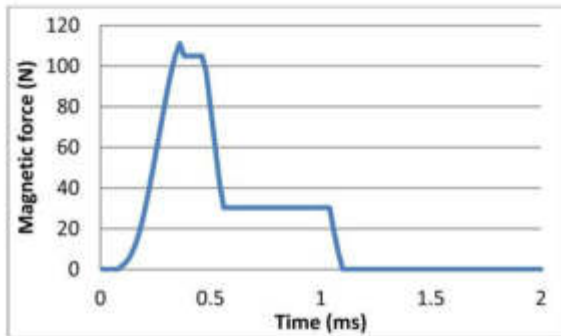
Figure 14. Magnetic force profile

The spring force F_{spring} can be expressed as,

$$F_{spring}(x) = F_0 + K_{spring} \cdot x$$

(11)

where, $F_0 = -6.5\,N$ is the initial compression force; $K_{spring} = -12140\,[N/m]$ is the spring constant, which is obtained by performing the return spring test.

The initial contact friction force is,

$$F_{initial\ contact\ friction}(x) = \begin{cases} -13.8\,N, & x \le 1e-3mm \\ 0, & x > 1e-3mm \end{cases}$$

(12)

The initial spring compression and initial contact friction forces were determined by performing the threshold current test which was conducted to investigate the minimum current to open the pintle valve under non-fluid condition. This single valued friction force may change as the fluid pressure applies. Future high fidelity model should investigate further on this force.

When the pintle moves 0.09 mm distance upwardly, a barrier is enforced by applying a stiff spring and damper to simulate that the pintle hits the stator. The force acting on pintle from the upper wall can be expressed mathematically as follows,

$$F_{upper\ wall}(x) = \begin{cases} -K*(x-0.09mm) - c*\dot{x}, & x > 0.09mm \\ 0, & x < 0.09mm \end{cases}$$

(13)

When the pintle moves back to the zero position (i.e. closed position), another barrier is enforced to simulate the supporting force from the bottom wall,

$$F_{bottom\ wall}(x) = \begin{cases} -K*x - c*\dot{x}, & x < 0 \\ 0, & x > 0 \end{cases}$$

(14)

Where K is the wall stiffness and K = 1e7 [N/m] for the current case; c is the damping coefficient and c = 1000 [N s/m] in this study.

The pressure force $F_{pressure}(x, t)$ on the pintle in vertical direction (i.e. pintle moving direction) was calculated using a CFD program in its transient simulation. This pressure force was also calculated statically at x=0 (i.e. pintle at its closed position) using the relation $F_{pressure}(x = 0) = P \cdot A_x$, where P is the fuel supply pressure (assuming uniform pressure within the fluid region) and A_x is the area in contact with fluid projected in the pintle moving direction. The static pressure force at different inlet pressures is shown in <u>Figure 15</u>. The pressure force, together with the initial spring and contact friction force, becomes the initial opposing forces that the magnetic force needs to overcome during the pintle opening event.
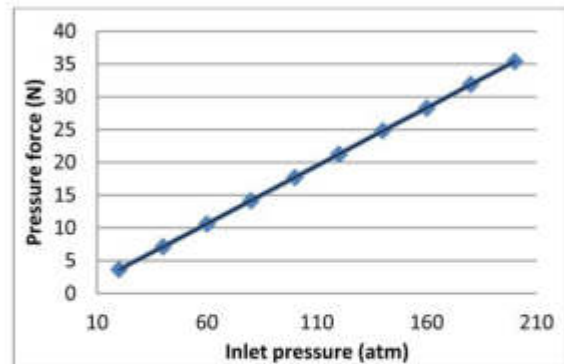
Figure 15. Calculated static pressure forces at pintle closed position at different inlet pressures

The pintle's equation of motion can be expressed as (ignoring damping of the fluid),

$$m \cdot \ddot{x} = F_{mag}(t) + F_{spring}(x) + F_{contact\ friction}(x) + F_{upper\ wall}(x) + F_{bottom\ wall}(x) + F_{pressure}(x,t)$$

(15)

All forces besides pressure force were incorporated into the rigid body dynamics as external forces in the CFD program. The pressure force was calculated at every time step in the transient simulation as the pintle position changed. In return, this pressure force was fed back to the rigid body motion as part of the external forces on the pintle. By doing so, the coupled electromagnetic-mechanical-fluid simulation was predicted.

Fluid System

Unlike other papers which included upstream pipe systems to consider pressure drops, mass transfer loss coefficients and pressure wave effects, this paper does not address the piping systems located away from the pintle. The effect of feeding pipe parameters on needle lift and flow rate can be found in [1], which investigated the impact of wave propagation on the injection event. Due to the scope of this paper, the rail dynamics' effect on pintle movement is not considered. But this can be an improvement of the fidelity of the fluid model in future study.

In this paper, the purpose of studying the fluid system was to investigate the pressure distribution around the nozzle and pintle valve during the pintle opening event. In order to ensure the accuracy of this fluid model, mass flow rates between calculated and measured data were compared at different inlet pressures. A sensitivity study of the nozzle diameter was performed to match the average mass flow rates. This multi-physics model was used to predict the injection spray event at different pressures and with different power strategies.

The CFD program used in this study was ANSYS CFX. The mesh preparation was done using the ICEM CFD mesh generation software. Only one sixth of the fluid region was modeled due to its periodicity. The pintle valve seat facets were simplified with smooth curvatures. A small inner radius was created instead of a singular point to make the mesh to be strictly hexahedral elements.

Mesh information is listed in Table 3. Typically, the quality of the mesh should be adequate for computation as long as the element angle is greater than 9~18 degrees and its determinant 2×2×2 is greater than 0.2. Different meshes were tested to study the mesh size effect on the mass flow rate at two different boundary conditions. Steady-state cavitation flow simulation at 200 atm supply pressure was carried out using different mesh sizes at pintle's fully open position. Figure 16 shows the mass flow rate tends to converge as the element number goes up. The mesh listed in Table 3 was found to be adequate to represent the geometry and generate convergent results.

The geometry of the modeled fluid region is shown in Figure 17.

Table 3. Mesh information for fluid model

Total elements	Total nodes	Lowest angle and the percentage	Lowest determinant 2x2x2 and the percentage
16097	12948	9~13.5 (0.703%)	0.25~0.3 (0.009%)
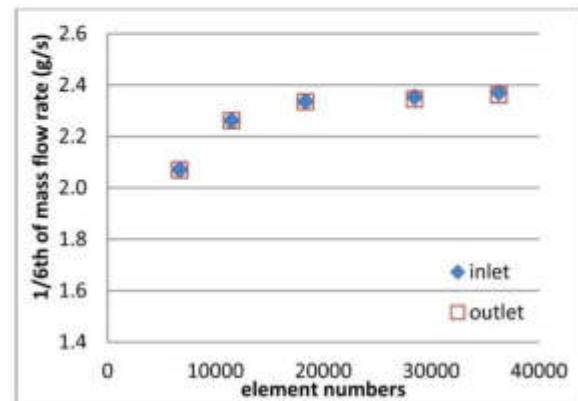
Figure 16. Grid independence study
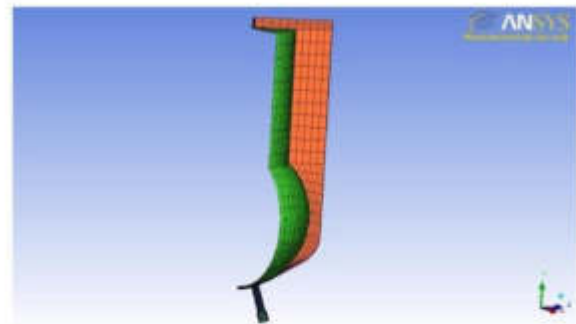
Figure 17. Meshed fluid region from ICEMCFD

Notice that this geometry only represents the initial fluid region. During the transient simulation in ANSYS CFX, the mesh will be deformed and the gap between the ball and valve seat will change. Since the software does not allow a zero or negative element, the complete separation of fluid region is not possible during the transient simulation. Therefore, a small gap (2.54e-3mm) was used to create the valve "closed" condition. The detail of this initial geometry at the ball valve location is shown in Figure 18.

337
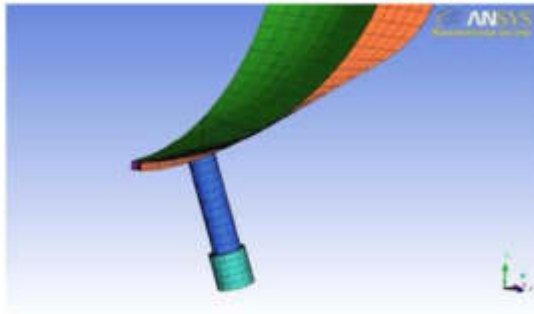
Figure 18. Small gap between ball and valve seat at the "closed" position

Transient cavitation flow simulation was performed in ANSYS CFX where a multiphase homogeneous model was used. In order to correlate with the experiments, Jet A fuel and air at 25°C were used as fluid and vapor phases, respectively. Jet A fuel was selected in the simulation because the injector power drive system optimization was performed for direct fuel injectors for military applications. Note that the saturation pressure of the Jet A fuel is 240 Pa. The fluid temperature was set at 25°C and the entire fluid region was assumed to be isothermal. The turbulence model was K-epsilon.

Mesh deformation strategy was used to incorporate rigid body motion in the transient simulation. The rigid body motion was applied to the pintle ball. The external forces on the ball were: spring force, electromagnetic force, initial contact friction force, and upper and bottom wall contact forces. The details of the force expressions were discussed in the previous section. Only a vertical (in the pintle moving direction) degree of freedom was allowed. Some key physical parameters are listed in Table 4.

Table 4. Key physical parameters for the fluid model

1/6th of the pintle mass	Pintle ball radius	Nozzle diameter	Initial pintle displacement
0.507 g	1.524mm	0.1524mm	2.54e-3mm

Boundary conditions, listed in Table 5, were specified as inlet, outlet and wall. Total and static pressures were used as inlet and outlet boundary conditions, respectively. The rest which includes the ball, valve seat, two periodic sides, and nozzle wall, were specified as no-slip wall boundary. The transient simulation's total run time was 2 ms and the time step was 0.25 μs.

Table 5. Boundary conditions

Inlet	Total pressure: 20 ~ 200 atm
Outlet	Avg. static pressure: 0 atm (reference pressure is 1 atm)
Others	No-slip wall

RESULTS AND DISCUSSIONS

Three different current shapes (power strategies) were used as the source of the pintle dynamics. The corresponding magnetic force profiles obtained using Maxwell was coupled with fluid model in ANSYS CFX transient. Pintle displacement, pressure force on the pintle and mass flow rate profiles during one injection event were obtained. The simulation revealed the details of the pressure forces' change with the pintle movement which led to the discovery of the magnetic force requirement for injection repeatability. This force requirement was further traced back in the Maxwell electromagnetic model to reveal the current shape requirement.

Two-stage current shapes were found to be the optimal power strategy for driving the fuel injector under different supply pressures. Key parameters, such as the 1st stage current, the 1st stage on-time, and the 2nd stage current were obtained according to the magnetic force requirement. The optimization goal was to use the minimal current to drive the fuel injector and achieve repeatability under different supply pressures.

Finally, simulation results that showed successful operation under the optimal power strategy and simulation results that failed to operate the injector consistently were validated experimentally using a photo detector measurement system.

Power Strategy (a)

The current profile in Figure 7 was named as the power strategy (a) which was used in the simulation to understand the dynamics of the pintle and the pressure force during the pintle opening event. The results gave insights into the optimal power strategy to drive the fuel injector under different operating pressures.

The resultant magnetic force from Maxwell was expressed in Equation (10) which used the power strategy (a) current profile.

After running cases for different inlet pressures, three major results were obtained from ANSYS CFX transient simulation: pintle displacement, pressure force, and mass flow rate.
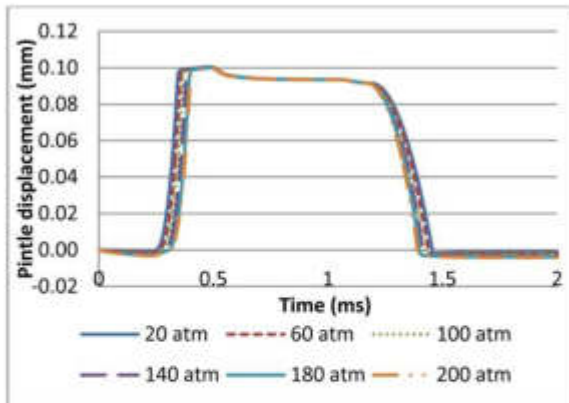
Figure 19. Pintle displacement profiles at different inlet pressures with power strategy (a)
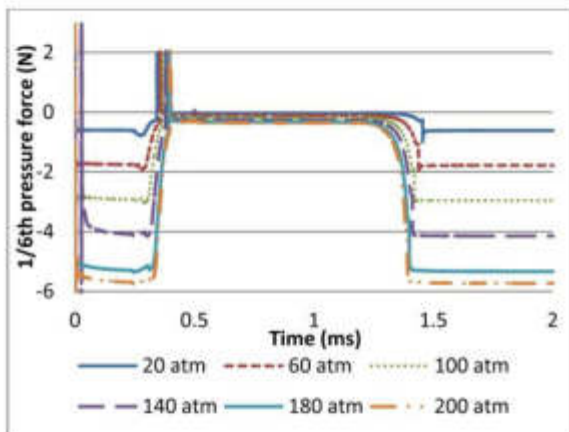
Figure 20. 1/6th of pressure forces on the pintle at different inlet pressures with power strategy (a)
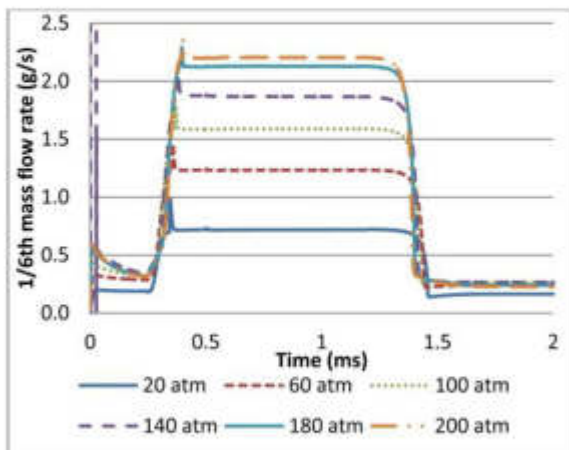
Figure 21. 1/6th of mass flow rates at different inlet pressures with power strategy (a)

As the inlet pressure was increased in Figure 19, the opening delay time became longer due to the increased initial pressure force on the pintle, as shown in Figure 20. This initial pressure force increased proportionally with the inlet pressure. As the current in the coil was built up (after the opening command signal), the magnetic force overcame the pressure force plus other resistant forces (i.e. spring force, contact friction force) and the pintle valve was opened. As the pintle moved, the pressure underneath the pintle ball increased, which caused the drop of net pressure force on the pintle ball. After the pintle reached its maximum position (upper wall of the stator), it hit the wall and was pushed back a little due to the stiffness of the wall. (Depending on the stiffness and damping coefficients, this could cause residual motion. For simplicity, this paper used a large damping coefficient to damp out this residual motion. The details of this phenomenon can be found in [2]. This push-back motion compressed the fluid region underneath the pintle ball, leading to a positive force spike as seen in Figure 20. After the pintle completely opened, it stuck to the upper wall of the stator and the pressure force on the pintle dropped to nearly zero. This was the time when steady flow formed, as can be seen from the "plateau" in Figure 21. In this figure, the mass flow rate was not strictly zero when the pintle was at its "closed" position due to the small gap used for the initial mesh. Regardless of this "leakage" flow at the "closed" position, the mass flow rate profile corresponded well with the pintle displacement profile.

In Figure 19, the pintle opening delay time started from 0.25 ms at 20 atm and increased to 0.31 ms at 200 atm. This increasing opening delay time was caused by the increasing initial pressure forces on the pintle ball as shown in Figure 20, from 3.6 N (total pressure force) at 20 atm to 35.4 N at 200 atm. Since the power strategy was the same, it took more time for the magnetic force to reach higher levels to overcome this increased initial pressure force. Therefore, in order to have minimum opening delay time, the 1st stage current increasing rate needed to be high so that it could reach the same current level with less time. Due to the limitation of the power supply's maximum voltage rating and the relation between the time to

$$t \propto \frac{1}{V}$$

reach certain current level and supply voltage: , the current increasing rate remained the same for all power strategies given the same voltage supply. Therefore, higher voltage rating power supplies are recommended for fast pintle opening.

After the current increasing rate was determined, the next question was how much current was required for the 1st stage to initiate the pintle opening. As discussed earlier in section 3, the magnetic force is needed to overcome the initial pressure, spring and contact friction forces. Table 6 listed the required magnetic forces at different inlet pressures. The relationship between the magnetic force and current was given in Figure 11. Using 0.1mm as the initial air gap and 160 as the number of coil turns in Maxwell, the corresponding threshold current levels were calculated using the force and current relationship

339

1148   *Zhang et al / SAE Int. J. Engines / Volume 7, Issue 3 (August 2014)*

in Figure 11. The resultant threshold current values were listed on the far right column of Table 6. For example, it required 8.4 A to open the pintle valve in the case of 200 atm inlet pressure.

Table 6. Initial opposing forces, required magnetic forces and threshold current levels under different pressures

Inlet pressure (atm)	Initial pressure force (N)	Initial spring force (N)	Initial contact friction force (N)	Combined forces (required magnetic forces) (N)	Threshold current (A)
0	0	6.5	13.8	20.3	4.9
20	3.7	6.5	13.8	24	5.4
40	7.1	6.5	13.8	27.4	5.8
60	10.6	6.5	13.8	30.9	6.2
80	14.2	6.5	13.8	34.5	6.5
100	17.7	6.5	13.8	38	6.8
120	21.2	6.5	13.8	41.5	7.2
140	24.8	6.5	13.8	45.1	7.5
160	28.3	6.5	13.8	48.6	7.7
180	31.8	6.5	13.8	52.1	8.0
200	35.4	6.5	13.8	55.7	8.4

After this threshold current was determined for different inlet pressures, the next question was about the on-time of the 1st stage current. Although the threshold current was straight forward, the on-time of the 1st stage required careful observation from the pressure force profile shown in Figure 20.

As shown in Figure 20, the pressure force was the maximum when the pintle was at its "closed" position. As soon as the magnetic force overcame the initial opposing forces, this pressure force dropped dramatically to nearly zero. Depending on the supply fuel pressure, the pintle opening time varied a small amount. The time also varied only a small amount between when the pintle opened and when it reached the fully open position. This time duration between when pintle opened and when it reached the fully open position was important because only when the pintle reached the fully open position, did the steady flow form and the pressure force drop to nearly zero. And only when the pressure force dropped to nearly zero could we possibly drop the current to the 2nd stage holding current, which was only required to overcome the spring force.

If the 1st stage current dropped sooner than it required to keep the pintle motion going until it reached the fully open position, the pintle might experience early fall back due to the inadequate magnetic force compared to the combined pressure and spring force during this opening process. This created inconsistent injection. This situation was modeled using 0.25 ms on-time of the 1st stage current. We called this power strategy (b), with a current profile shown in Figure 22. The pintle displacement, pressure force and mass flow rate plots under power strategy (b) are shown in Figure 23, 24, 25.

Power Strategy (b)
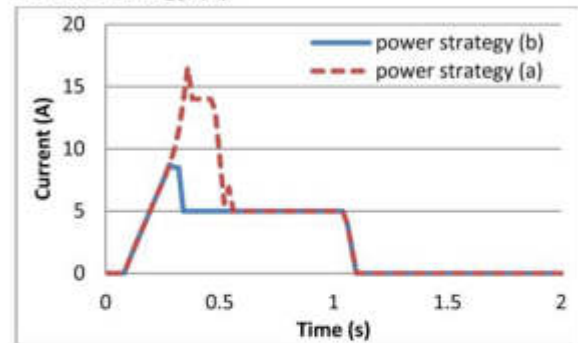
Figure 22. Current profile comparison between power strategy(a) and (b)
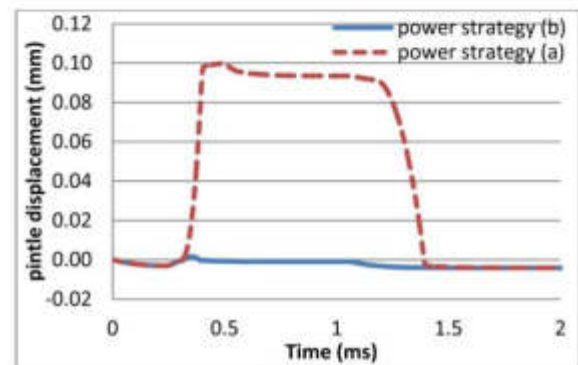
Figure 23. Pintle displacement profiles with power strategy (b) and (a) at 200 atm supply pressure
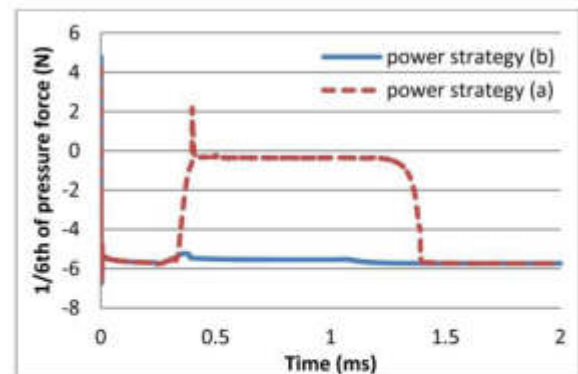
Figure 24. 1/6th of pressure forces on the pintle with power strategy (b) and (a) at 200 atm supply pressure
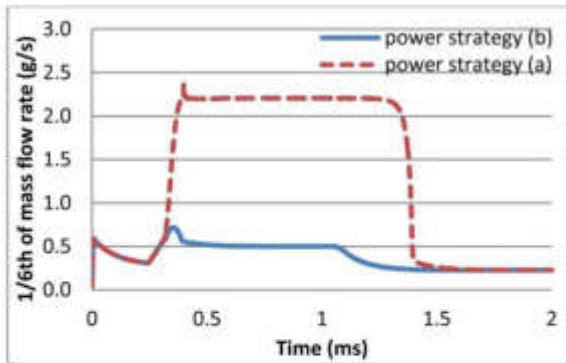
Figure 25. 1/6th of mass flow rates with power strategy (b) and (a) at 200 atm supply pressure

Power strategy (b) failed to completely open the pintle at 200 atm inlet pressure. The pintle was first initiated but it fell back before it reached the upper wall of the stator. Therefore, even though the 1st stage current level might be enough to overcome the threshold opposing forces (i.e. pressure force, spring force and initial contact friction force), it might create early fall (or random position) of the pintle as shown in the 200 atm inlet pressure, if the current dropped too soon (i.e. before the pintle reached the upper wall of the stator). This caused injection inconsistency. The ideal current of the 1st stage should remain at a certain level until the pintle reaches its maximum position to ensure consistency of fuel injection. The time that it took to fully open the pintle was estimated to be at most 0.4 ms based on the case with 200 atm supply pressure in Figure 20.

After the pintle was fully opened (i.e. the 1st stage), the magnetic force at the 2nd stage could be reduced significantly due to the drop of the pressure force at the pintle fully open position. Still, the magnetic force should be large enough to overcome the spring force at the pintle fully open position to hold the valve open. If the 2nd stage current level creates less magnetic force than the compressed spring force, the pintle might experience early fall as well. This is illustrated using current profile shown in Figure 26 which is named as power strategy (c). The pintle displacement, pressure force and mass flow rate plots with power strategy (c) under 200 atm supply pressure are shown in Figure 27, 28, 29.

With power strategy (c), pintle fell back earlier due to the inadequate 2nd stage magnetic force (at 3 A) to hold the pintle at its fully open position. Therefore, even though the 2nd stage current could be dropped, it required a minimum value (>3A, which created >7.5 N magnetic force at the fully open position. The spring force at the fully open position was 7.7 N) so that the magnetic force could still overcome the spring and pressure forces at the pintle fully open position. As discussed earlier, steady flow was formed and the pressure force dropped to nearly zero when the pintle was at its fully open position.

Power Strategy (c)
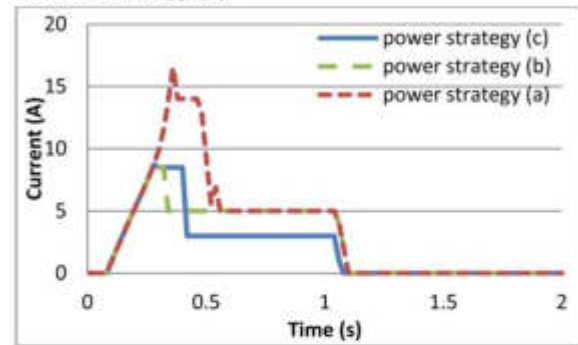
Figure 26. Current profile comparison among power strategy (a), (b) and (c)
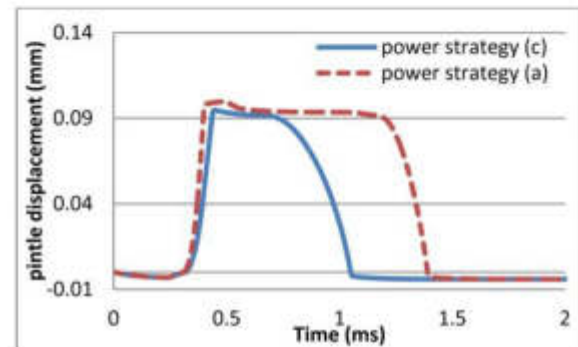
Figure 27. Pintle displacement profiles with power strategy (c) and (a) at 200 atm supply pressure
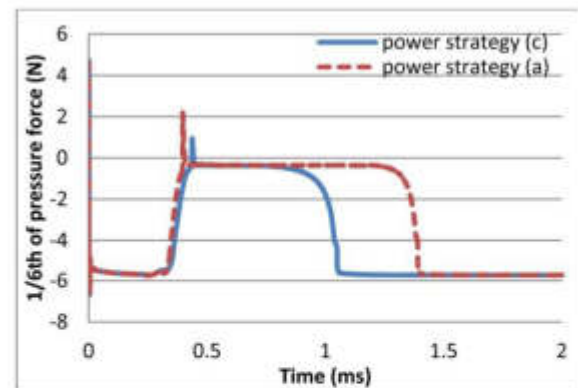
Figure 28. 1/6th of pressure forces on the pintle with power strategy (c) and (a) at 200 atm supply pressure

341
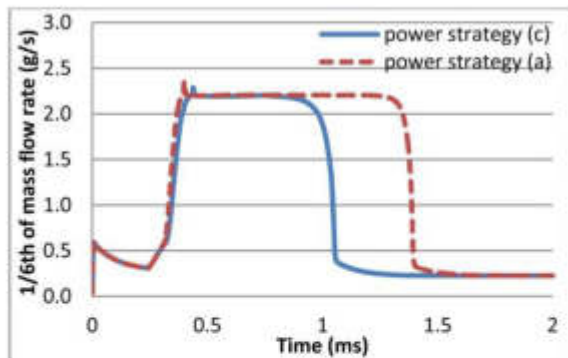
Figure 29. 1/6th of mass flow rates with power strategy (c) and (a) at 200 atm supply pressure

Optimal Power Strategy

After the trials of different power strategies, the magnetic force requirement that led to the current requirement gradually formed. From the previous discussion, there are three key parameters that determine whether or not the pintle could be successfully opened and held open: threshold current, the 1st stage on-time and the 2nd stage current.

These key parameters are illustrated in Figure 30. $I_{threshold}$ at different inlet pressures can be referred to Table 6. For example, $I_{threshold}$ = 8.4 A at 200 atm supply pressure.

$I_{pick-up}$ can be lowered based on the fact that once the pintle motion is initiated, the initial contact friction force drops to zero. $I_{pick-up}$ can be determined using the corresponding current value of the combined force minus the initial contact friction force at different inlet pressures in Table 6. For example, $I_{pick-up}$ = 7.5 A at 200 atm supply pressure.

Ihold is the holding current for almost all pressure cases after pintle reaches the fully open position. Based on the spring force calculation (i.e. F_{spring} =7.7 N at pintle fully open position), the magnetic force at the pintle fully open position requires I_{hold} > 3 A. Refer to Figure 10 for ~0 mm air gap and steel 1008 case.
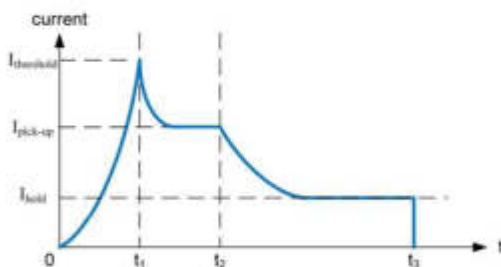
Figure 30. Current requirement illustration during one injection event

The three key current levels shown in Figure 30 are often mentioned as peak, pick-up and holding currents in the user's manual of fuel injectors. The previous discussion explained the reason for having these three key parameters. Threshold current is the current to overcome the initial contact friction, pressure force and spring force. Pick-up current is the current to keep the pintle opening motion going until it reaches fully open position. Holding current is the current to keep the pintle at its fully open position.

In a real application, the pick-up stage lasts a very short amount of time, the threshold and pick-up stages are often combined to form the 1st stage. Holding stage is to be called as the 2nd stage. Figure 31 illustrates how the current shape looks like for implementation.
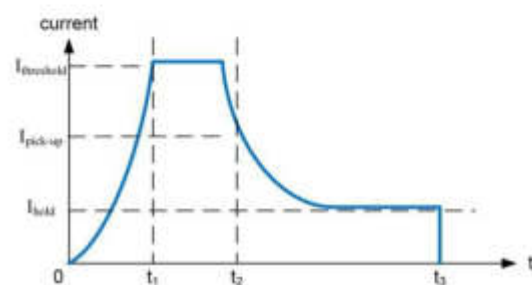
Figure 31. Current profile illustration for implementation

On the time axis, t_1 is the time when the pintle motion is initiated; t_2 is the time when the pintle reaches its fully open position; t_3 is the time when the injection is to be stopped. Among these times, t_2 is the time that ensures the pintle motion continues from when the motion gets initiated to the time when the pintle reaches its fully open position. This time is called the 1st stage on-time. It is around 0.4 ms based on the results from test and simulation. t_3 simply defines the total spray duration time. Despite the fact that the closing delay time is also closely related to pressure and magnetic forces (i.e. higher inlet pressure has the tendency to push the pintle back quicker, thus creating less closing delay time), it was found in the accompanying paper [9] that, the closing delay time is almost constant regardless of t_3, given adequate t_2 value and the spray duration time is proportional to total command signal on-time.

Validation

Validation Under Power Strategy (a)

JP-8 fuel was used in the validation experiment. JP-8 fuel was selected because the injector power drive system was designed to control a heavy fuel direct fuel injector for military applications. However, any other liquid fuel should be a sound substitute in the study. The test and simulation results on average mass flow rate during an opening event (total injection

quantity divided by injection duration time) were compared at different supply pressures in Figure 32. Power strategy (a) was used for all cases. The total injection quantity was measured by a scale after a specified number of injections.
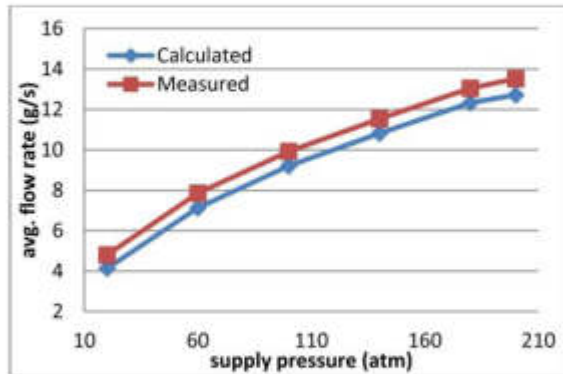
Figure 32. Average mass flow rate comparison during one injection event with power strategy (a) at different supply pressures

Accuracy of pintle displacement profiles was validated with the experimental data taken using a photo detector, which is an indicator of pintle motion. This approach assumes little time delay between the pintle motion and the detected spray. The discrepancy between the photo detector measured data and computer simulated pintle movement data are mainly due to this assumption. Overall, because of the high velocity spray under high supply pressure, the results from measuring the spray should give insights to the pintle behavior, which is the main purpose of this validation work.

The photo detector is an electronic device which emits an infrared beam to the receiver end. The change of the output of the receiver circuitry indicates whether or not there is an obstacle blocking the path. In our case, the fuel spray is the obstacle that causes the change of the photo detector output signal. The details of using the photo detector to measure the injection event can be found in the accompanying paper [9]. Figure 33 shows the control voltage, current and photo detector signal during an injection event with injection rate at 100 Hz. Four (4) volts of the photo detector output indicates no obstacles in the path, while nine (9) volts indicates obstacles are detected. Therefore, the "plateau" duration indicates the total spray duration time.

The spray temporal characteristics (i.e. opening delay and spray duration times) between calculated and measured data are compared in Table 7.

As shown in Table 7, the measured delay times did not follow the monotonous increasing trend of the calculated values as the supply pressure was increased. The discrepancy between the calculated and measured delay times was mainly due to the delay time between the first drop of spray coming out from the nozzle and the time when the photo detector sensed it. As

the supply pressure was increased, the velocity of the first drop of spray increased as well. This shortened the delay time at higher pressures. Therefore, even though the pintle was opened quicker at lower pressures, it took longer time for the photo detector to detect the spray. The end result was about the same delay time (~0.4 ms) from photo detector measurement for all pressures.
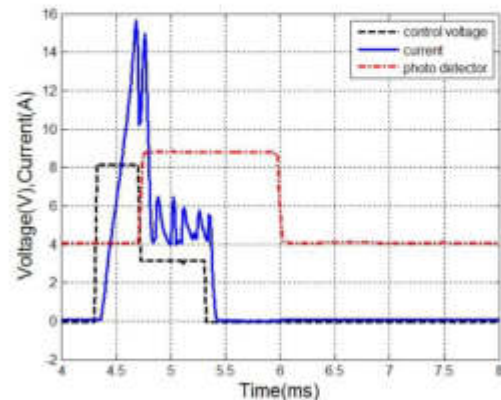
Figure 33. Test with photo detector during one spray event with power strategy (a) at 200 atm

Table 7. Spray temporal characteristics comparison between calculated and measured data

supply pressure (atm)	opening delay time (ms)		spray duration (ms)	
	calculated	measured	calculated	measured
20	0.25	0.43	1.21	1.25
60	0.27	0.41	1.17	1.31
100	0.28	0.39	1.14	1.31
140	0.29	0.39	1.12	1.30
180	0.31	0.39	1.09	1.28
200	0.31	0.40	1.08	1.28

The spray duration time followed the trend of calculated values at pressures higher than 20 atm. The discrepancy between the absolute values of the calculated and measured data was mainly due to the delay time when the pintle was closed and when the last drop of spray passed the photo detector. For the case at 20 atm supply pressure, the spray intensity was lower than its high pressure counterpart, this lowered intensity of the spray might render less stagnation time of the photo detector output.

Overall, there were small errors between the calculated and measured spray temporal characteristics (opening delay and spray duration times) due to the differences of pintle motion and spray detected at the photo detector. However, the close match of average mass flow rates and spray temporal characteristics at different supply pressures validates the multi-physics model during injector opening event.

343

Validation of Predicted Phenomena Under Power Strategy (b) and (c)

The purpose of the multi-physics model developed in this study was to optimize the power strategy to drive fuel injectors. This gave insights into the requirement of threshold current, the 1st stage on-time and the 2nd stage current. The multi-physics model predicted the failure of the power strategy (b) and (c). We then tested the injector using the similar power strategies and confirmed the inconsistency of the photo detector output, which was the indication of the early fall of pintle as we predicted in the simulation.

Figures 34 and 35 compare two power strategies with the same 1st stage current, 2nd stage current, total control signal on-time, but different 1st stage on-times. Figure 34 shows the consistency (same spray duration time between pulses) of the case with 0.3 ms 1st stage on-time, while Figure 35 shows the inconsistency (different spray duration times between pulses) of the case with 0.16 ms 1st stage on-time. The latter case is the demonstration of power strategy (b) in the simulation, which failed to open the pintle to its fully open position. The early fall back of the pintle was the cause of inconsistent spray duration times.

Figures 34 and 36 compare two power strategies with the same 1st stage current, 1st stage on-time, 2nd stage on-time, but different 2nd stage current levels. Figure 34 shows the consistency (same spray duration time between pulses) of the case with 3.5 A 2nd stage current, while Figure 36 shows the inconsistency (different spray duration times between pulses) of the case with 2.5A 2nd stage current. The latter case is the demonstration of power strategy (c) in the simulation, which failed to hold the pintle at its fully open position. The early fall back of the pintle was the cause of inconsistent spray duration times.
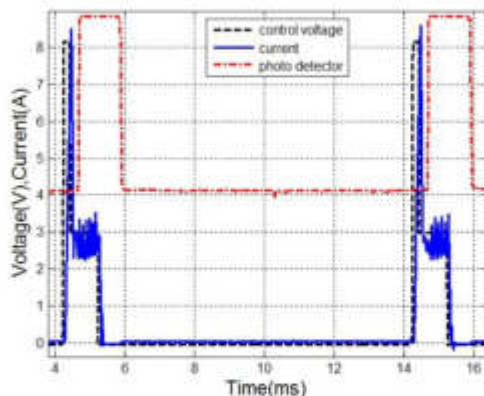
Figure 34. Photo detector output with 1st stage current of 8 A, 1st stage on-time 0.2 ms, 2nd stage current of 3.5 A, 2nd stage on-time 0.8 ms; at 200 atm supply pressure
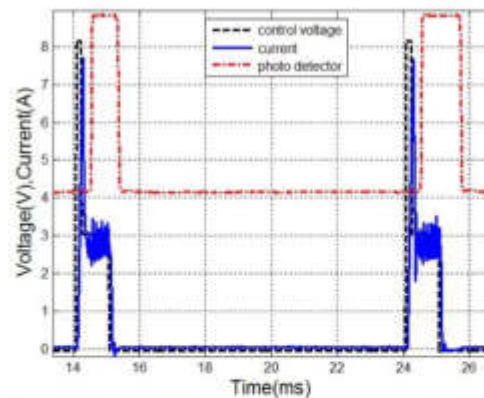
Figure 35. Photo detector output with 1st stage current of 8 A, 1st stage on-time 0.16 ms, 2nd stage current of 3.5 A, 2nd stage on-time 0.84 ms; at 200 atm supply pressure
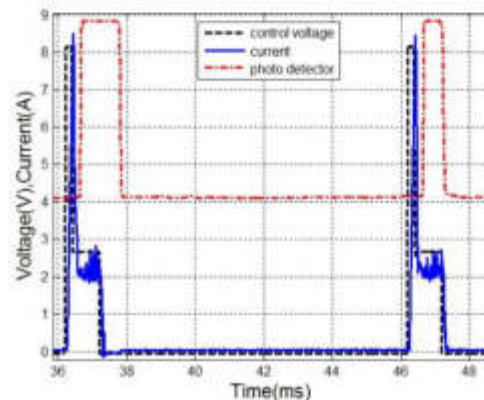
Figure 36. Photo detector output with 1st stage current of 8 A, 1st stage on-time 0.2 ms, 2nd stage current of 2.5 A, 2nd stage on-time 0.8 ms; at 200 atm supply pressure

CONCLUSION

This paper presents a coupled electrical-electromagnetic-mechanical-fluid model of a direct fuel injector and its power drive system. A simplified electrical circuit model was built to create the current profile to the injector coil from a two-stage PWM current controlled approach. The circuit model was then imported into the finite element code Maxwell for an electromagnetic force calculation. Pintle transient movement and the non-linearity of B-H curve were considered in the model. The force calculated using Maxwell was compared with the 1D electromagnetic analysis. Then, the magnetic force profile was obtained with power strategy (a). A detailed mechanical model, which included rigid body motion of pintle, spring force, initial contact force, pressure force and spring-damper system as wall barrier, was formed. One-sixth of the coupled electromagnetic-mechanical-fluid model using the CFD code ANSYS CFX was built and analyzed. Pintle displacement, pressure force and mass flow rate transient responses were obtained. The pressure force gave insights into the power requirement during the pintle opening event.

344

Different power strategies were used in the computer model to illustrate the cause of injection inconsistency. Finally, a photo detector based spray temporal measurement system was used to validate the results from computer simulation.

The optimal power requirement for a two-stage PWM current controlled approach was obtained to improve the injection repeatability at minimum power consumption:

1. The 1st stage current needs to be able to overcome initial opposing forces such as spring, pressure and contact friction force ($I_{threshold} > 8$ A at 200 atm supply pressure). The threshold current can be lowered at lower operating pressures. It is always recommended to use higher voltage rating power supplies for fast pintle opening.

2. The 1st stage on-time needs to be long enough to keep the pintle movement going until it reaches its fully open position ($t_2 \approx 0.4$ ms). Otherwise the pressure force will push the pintle back to create early fall back, which is one cause of injection inconsistency;

3. The 2nd stage current can be lowered significantly due to the drop of pressure force at pintle fully open position, yet it needs to maintain at certain level to be able to overcome the spring force and hold the pintle at its fully open position ($I_{hold} > 3$ A). Otherwise the pintle will experience an early fall as well due to the spring force, which is another cause of injection inconsistency. The 2nd stage current requirement does not increase as the supply pressure increases due to the fact that once the pintle is at its fully open position, the steady flow forms and the pressure force on the pintle drops to nearly zero.

4. The 2nd stage on-time has not been discussed due to the fact that the closing delay time does not vary with on-times and the total injection time is proportional to control signal's total on-time. The injection time can be easily adjusted by changing the 2nd stage on-time.

The fluid model used in this paper only included geometries at the near-nozzle region. The upstream piping system was not considered in the model. Furthermore, due to the inaccessibility of the real-time pintle displacement, pressure force on the pintle and flow rate measurement, only average flow rate data and photo detector data (indicator of pintle displacement) were used to validate the results from computer simulation. More careful matching of these instantaneous profiles is needed for future investigation.

This paper also has given insights into how to design a high-performance fuel injector:

1. High magnetic saturation material is preferred to build fuel injectors to create larger magnetic force to initiate the pintle motion.

2. The return force (spring force in this study) determines the closing delay time. Thus, large return force would create less closing delay time.

3. The fluid model can be used to better understand the physics during the fuel injection. For example, the relationship between nozzle diameter and mass flow rate; the effect of geometry changes on the cavitation zone at various operating pressures. The cavitation can thus be reduced or eliminated by carefully choosing the right geometries and pintle motions.

ACKNOWLEDGEMENTS

REFERENCES

1. Digesu, P. and Laforgia D., "Diesel electro-injector: A numerical simulation code". Journal of Engineering for Gas Turbines and Power- Transactions of the Asme, 117(4): 792-798, 1995. Doi:10.1115/1.2815466.

2. Ficarella, A. and Laforgia D., "Experimental and numerical investigation on cavitating flows in diesel injection systems": Meccanica, 33(4): 407-425, 1998. Doi:10.1023/A:1004329902598.

3. Coppo, M., Dongiovanni C., and Negri C., "Numerical analysis and experimental investigation of a common rail-type diesel injector". Journal of Engineering for Gas Turbines and Power- Transactions of the Asme, 126(4): 874-885, 2004. Doi:10.1115/1.1787502.

4. Coppo, M. and Dongiovanni C., "Experimental validation of a common-rail injector model in the whole operation field". Journal of Engineering for Gas Turbines and Power-Transactions of the Asme, 129(2): 596-608, 2007. Doi:10.1115/1.2432889.

5. Hu, Q., et al., "Modelling of dynamic responses of an automotive fuel rail system, part I: Injector". Journal of Sound and Vibration, 245(5): 801-814, 2001. Doi:10.1006/jsvi.2000.3605.

6. Ando, R., Koizumi M., and Ishikawa T., "Development of a simulation method for dynamic characteristics of fuel injector". Ieee Transactions on Magnetics, 37(5): 3715-3718, 2001. Doi:10.1109/20.952697.

7. Tsai, W.C. and Yu P.C., "Design of the Electrical Drive for the High-Pressure GDI Injector in a 500cc Motorbike Engine". International Journal of Engineering and Industries, volume 2, Number 1, March, 2011. doi:10.4156/ijei.vol2.issue1.9.

8. Tsai, W.C. and Wu Z.H., "Use of Taguchi Method to Optimize the Operating Parameters of a High-Pressure Injector Driving Circuit". Mechanical and Electronics Engineering Iii, Pts 1-5, 130-134: 2795-2799, 2012. Doi:10.4028/www.scientific.net/AMM.130-134.2795.

9. Zhang, X., Palazzolo, A., Kweon, C., Thomas, E. et al., "Direct Fuel Injector Temporal Measurements," SAE Technical Paper 2014-01-1444, 2014.

ERRATUM

Zhang, X., Palazzolo, A., Kweon, C., Thomas, E. et al., "Direct Fuel Injector Power Drive System Optimization," *SAE Int. J. Engines* 7(3):1137-1153, 2014, doi:10.4271/2014-01-1442.

1.   As published, Figure 4 contained errors. The correct figure is

Figure 4. Inductance at different air gaps from Maxwell.

2. Under figure 6, $\mu 1 \approx \mu 2$, the approximation sign ($\approx$) was not published correctly.

3. In previously published Figures 22 and 26, the horizontal axis title incorrectly read 'Time (s)' but should read 'Time (ms)'. The correct figures are

Figure 22. Current profile comparison between power strategy (a) and (b)

Figure 26. Current profile comparison among power strategy (a), (b) and (c)

1154

346

[34]   Tutaj, J., Fijalkowski, B., "A new Fuel-Injection Mechatronic Control Method for Direct-Injection Interal Combustion Engines" Sciendo, doi: 10.2478/ama-2018-0042, 2018.

A NEW FUEL-INJECTION MECHATRONIC CONTROL METHOD
FOR DIRECT-INJECTION INTERNAL COMBUSTION ENGINES

Józef TUTAJ[*], Bogdan FIJAŁKOWSKI[**]

[*]Institute for Automotive Vehicles and Combustion Engines, Department of Mechanical Engineering, Cracow University of Technology,
Al. Jana Pawła II 37, 31-864 Kraków, Poland
[**]Institute of Technology, State Higher Vocational School in Nova Sandec, ul. Staszica 1, 33-300 Nowy Sącz, Poland

pmtutaj@cyf-kr.edu.pl, pmfijalk@cyf-kr.edu.pl

Abstract: In this paper, a novel fuel-injection mechatronic control method and system for direct injection (DI) internal combustion engines (ICE) is proposed. This method and system is based on the energy saving in a capacitance using DC-DC converter, giving a very fast ON state of the fuel injectors' electro-magnetic fluidical valves without an application of the initial load current. A fuel-injection controller for the DI ICEs that provides a very short rising time of an electromagnet-winding current in an initial ON state of the fuel-injector's electromagnetic fluidical valves, which improves a fuel-injection controller reliability and simplify its construction, is presented. Due to a number of advantages of afore -mentioned fuel-injection mechatronic control method and system, it may be utilised for the DI ICEs with fuel injectors dedicated to all types of liquid and/or gas fuels, for example, gasoline, diesel-oil, alkohol, LPG and NPG.

Key words: Direct Injection, Internal Combustion Engine, Fuel Injection, Injection Control System

1. INTRODUCTION

Currently on automotive market, more and more automotive vehicles use internal combustion engines (ICE) with a spark-ignition and direct-fuel injection. The idea for directly injection consists in forming the mixture directly in a combustion chamber. To achieve this, the electromagnetic fuel injector (EFI) is placed in the combustion chamber and by an intake fluidical valve is flow only an air.

A special injection mechano-hydraulical (M-H) pump generates the fuel injection at high pressure (5 to 12 MPa). The task of the fuel injector is dosing and spraying of a fuel into the smallest particles to provide an adequate mixing of the fuel and air in a specific area of the combustion chamber. After activation of an electromagnet of the EFI by an appropriate current pulse, the fuel is injected into the combustion chamber due to the pressure difference inside and outside fuel injector.

The idea of direct injection (DI) of a gasoline is quite old (Bosch 2014, Zhao 2016), but mass production took place only since 1995 with the introduction of ICEs with the gasoline direct injection (GDI) by Mitsubishi.

Electronically controlled fuel injection is currently used more widely, both in ICEs with the spark ignition with different markings of the supply system (GDI, FSI, IDE, HPI, JTS, etc.), as well as – Diesel ICEs (HDI, JTD, CDI, DCI, TDI etc.).

EFIs for ICEs with the direct fuel injection need to convey the current pulse with the proper value of the instantaneous power to the winding of the fuel injection electromagnet in the initial phase of fuel injection's opening. due to the operation to be performed by moving the needle fuel injector with a specific mass in a relatively short ON state (Achleitner et al., 2007; Husted et al., 2014).

The purpose of this article is to present a new method of control and powering of the injector of the direct injection system, which uses the energy stored in the additional capacitor to fast force the winding current, and does not need to be taken into account when controlling the so-called "initial load phase" time (Fig. 2).

The paper presents the results of electric current measurements of the injector during experimental tests.

1.1. A fuel injecttion mechatronic control for direct-injection internal combustion engines

In conventional fuel-injection mechatronic control, the microcontroller generates the control voltage for the solenoid coil of the EFI equals 60 - 90 V that is necessary to activate the EFIs of ICEs with the direct fuel injection. The internal structure of the EFI is shown in Fig. 1.

The control voltage is so high so as to minimize a response time of the EFI's mechatronic-control signal and get a short time of the fuel injection, which should be less than time of the fuel injection for ICEs with the fuel injection into the intake manifold (Fig. 2).

1.2. An initial load phase

The winding of the fuel-injector's electromagnet is powered by a pulse-width-modulated (PWM) voltage of value equals 12 – 14 V, in order to stabilize the current value of about 1 A.

An initial load phase prepares the gasoline fuel injector fuel to open in a short time, through the initial activation of the fuel

276

injector (approx. 1.5 ms before the beginning of opening of the EFI) to generate the initial value of the EFI's electromagnet-winding current.

The EFI's solenoid coil consists of an electrical circuit RL, and therefore a presence of the inductance L (several mH) in depending on the type of the EFI and its resistance R (0.5-2 Ω) causes that at the EFI's solenoid-coil supply voltage of a constant value, the current rises from zero to a steady-state value limited by the resistance R of the EFI's solenoid coil.

The initial load phase allows bringing faster floating of the EFI's needle during opening of the electromagnetic fluidical valve by the primary saving of an electrical energy in the inductance.
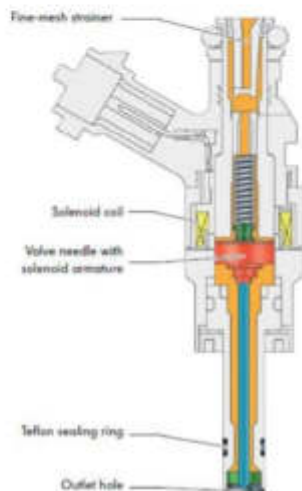
Fig. 1. Electromagnetic fuel injector (Bosch, 2014)
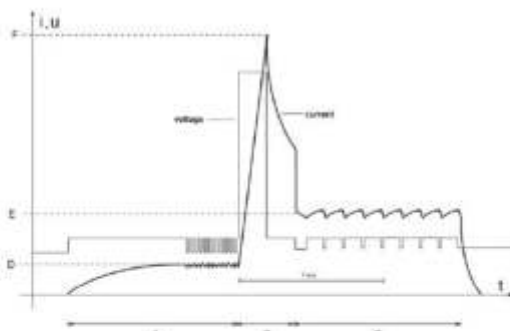
Fig. 2. Waveforms of the fuel-injector's current and voltage: A - initial load phase; B - excitation phase; C - hold-up phase; D - initial load current; F - excitation current; E - hold-up current

1.3. An excitation phase

An excitation phase is for causing a rapid increase of the EFI's needle by temporarily increasing the current flowing through the fuel-injector's solenoid coil. The EFI's electromagnet winding is supplied on a following ways: by means of an electrical valve (electronic switch) that for a short time commutate the voltage

of about 80-90 V, which increases the winding's current to an instantaneous value of about 11-12 A. Afterwards, the EFI's solenoid coil is switched to the voltage of 12-14 V, and causes a decrease in the forcing current impulse. It should be noted that the relatively high current flows only for a very short time of about 0.5 ms without any causing any thermal overload of the electromagnet's winding.

The high instantaneous value of the power-pulse control will cause decreasing of the opening time of the EFI. The width of the voltage pulse (the voltage wave-form of the EFI's electromagnet solenoid coil supply in Fig. 1) must be controlled with high accuracy, because even a small change in the duration of the voltage strongly influences the instantaneous value of the electro-magnet-winding current.

1.4. A holding-up phase

Holding-up phase allows you to continue a power supply (opening of the EFI), limiting the saved electrical energy needed to held the EFI's needle in the open position by a few milliseconds, depending on the required fuel dose. The EFI's solenoid coil is supplied as follows: with the PWM controlled voltage of about 12-14 V; and the current of about 2.5-3 A. The duration of the holding phase (1-5 ms) determines the size of the dose of fuel injected into the combustion chamber of the ICE. Loss of current pulse closes the EFI.

The above-described method for controlling fuel injectors for DI ICEs by many car manufacturers has significant drawbacks. The use of an initial-load-phase's current pulse (the prediction pulse opening the fuel injector) causes as well an additional fuel-injector's thermal load – an additional heating of the windings of the electromagnet coil during the initial-load phase of the EFI's opening, as troublesome fuel-dose control. Because the synchronization pulses control of fuel injectors should take into account the additional lead time of the initial-load phase and convert a position angle of the crankshaft), for the EFI's opening at a predetermined angular position of the crankshaft of the ICE.

A second shortcoming of this control method is a disadvantageous shape of the current pulse in the electromagnet winding of the fuel-injector during an excitation phase (Fig. 2).

The shape of the current pulse in the excitation-phase control like a 'peak' that results in generation of harmonics with higher frequencies and increasing of larger eddy-current losses in a magnetic circuit of the EFI. The exact stabilization of the excitation-phase duration due to the fact that in the case of un-controlled extension of this impulse may be destroyed the EFI (the electromagnet-winding current - resulting from the higher supply voltage of 80-90 V and lower impedance could be achieved in the long run value of tens of amperes).

2. A NEW METHOD AND FUEL MECHATRONIC CONTROL SYSTEM FOR DIRECT-INJECTION INTERNAL COMBUSTION ENGINES

In Fig. 3 is shown a fuel-injection mechatronic control system for DI ICEs that is based on an electrical energy accumulation in a capacitor and fast transfer of the power to the EFI's electromagnet solenoid coil.
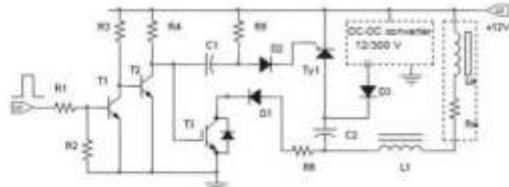
Fig. 3. Schematic diagram of the fuel-injector controller for DI ICEs

The capacitor C2 collects electricity using the diode D3 is charged from the output voltage 12/300 V of the DC-DC converter. With the emergence of the input control pulse is actuated at once the thyristor Ty1 and isolated-gate bipolar transistor (IGBT) T3. The energy stored in the capacitor C2 is in a very short period of time transferred to the additional inductance L1 and the fuel-injector's electromagnet winding. After discharging the capacitor C2, a function to maintain the current of the EFI's electromagnet through the IGBT T3 by means of the two-terminal circuit consisting of the high-voltage pulse diode D1 (400 V, 5 A), and additional resistor R6. A value of the resistor R6

determines a value of the holding current (Fig. 3). Since the energy stored in the capacitor C2 depends only on the output voltage of the DC-DC converter and its capacitance, and therefore the value of this energy can be easily stabilized and the EFI is protected from damage even in the case of a wrong control pulse appearance (e.g. due to interference). After an exchange of energy between the capacitor C2 and the EFI's electromagnet winding Rw, Lw during the holding-up phase, the commutation system consisting of components: L1, R6, D1, T3, holds up the current in the EFI's electromagnet winding. Very fast exchange of energy between the capacitor C2 and the additional inductance L1 and EFI's electromagnet winding Lw, Rw lasts approximately 0.3 ms (instead of approx. 0.5 ms in a conventional fuel-injection control system). It means that there is no need for an initial-load-phase pulse of the EFI, which simplifies the control and reduces heat losses in the EFI's electromagnet winding (no initial-load-phase current of approx. 1 (A) during approximately 1.5 (ms). An additional advantage (in addition to faster opening of the EFI) is preferable to form of a forcing pulse shape, which reduces the losses in the magnetic circuit of the EFI. The schematic diagram of the 12V / 250-300V DC-DC converter is shown in Fig. 4.
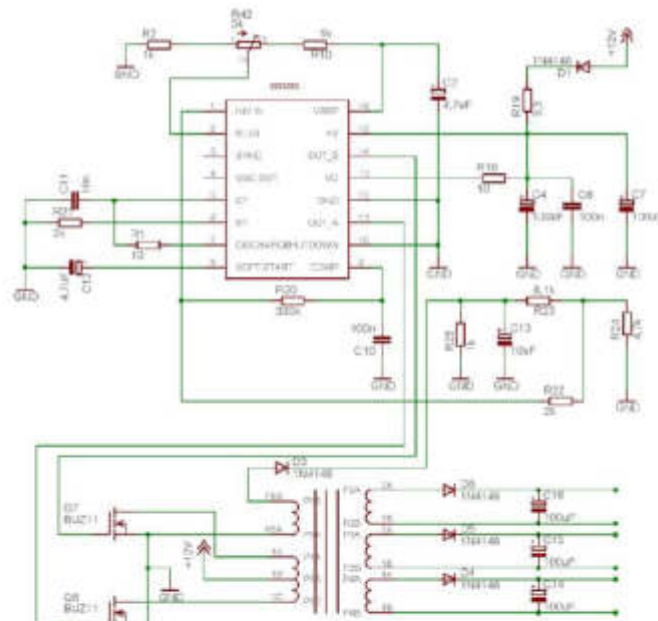
Fig. 4. Circuit diagram of the 12V / 250-300V DC-DC converter

An auxiliary voltage of 250 - 300 V is generated by the DC-DC converter with a feedback loop to stabilize the output voltage using the application specified integrated circuit (ASIC) type SG3525.

2.1. Mathematical model of the pressure dynamics

In this paper, we exercise a mathematical model describing

the pressure dynamics in the common rail (CR) rail that is projected and experimentally validated in di Gaeta et al. (2009, 2011, 2012). This mathematical model depicts the electrical dynamics of the EFI (fluidical valve), ignoring the effects owing to the movement of its plunger (i.e. inductance variations and back-electromotive force), and the actuation circuit utilized to drive the EFI fluidical valve) as well. According with this alternative the pressure, say p(t) bar, in the CR depends on the mechano-hydraulical (M-H) pump rotational speed and EFI current i(A), and

it can be de-composed into two components, namely the mean pressure term, say pa(t), and the residual pressure, p(t), relating the ripple around the mean value. The CR dynamical system is then given by:

$$\frac{di}{dt} = \frac{R}{L}i + \frac{V_b}{L}\left(\frac{a\delta(t)+b}{100}\right), \tag{1a}$$

$$p_\alpha(t) = c(N)i + d(N), \tag{1b}$$

$$p(t) = p_\alpha(t) + \eta(t), \tag{1c}$$

where: δ (%) is the duty cycle expressed in percentage terms of the PWM signal used to actuate the EFI hydraulical valve, a and b are parameters of the EFI actuation circuit, L (H) and R (Ω) are the EFI (hydraulical valve) inductance and the electric resistance of solenoid coil, respectively, whereas V_b (V) is the chemo-electrical (CH-E) storage battery voltage supplying the power circuit and N (rpm) is the rotational speed of the M-H pump that is equal to N + Nc/2 with Nc being the ICE rotational speed.

The terms c(N) and d(N) in (1) depend on the M-H pump rotational speed N and have to be experimentally identified. According to the approach in di Gaeta et al. (2009), we mathematically model them as third-order polynomials of the form:

$$c(N) = \sum_{k=0}^{3} c_k \left(\frac{N}{10^3}\right)^k, \, d(N = \sum_{k=0}^{3} d_k \left(\frac{N}{10^3}\right)^k. \tag{2}$$

2.2. GDI injector's mathematical model

Since for the design of basic automotive control loops, namely air-to-fuel ratio (AFR) control, the control variable to be designed is the mass of fuel to be injected, a relation between this control variable and the time of injection is essential for the integration of the CR in the entire ICE control system (Gupta et al., 2011; Jiangjian et al., 2007; Tang et al., 2009; Yan and Wang, 2011).

In this paper, we advise and validate by means of experimental data a simple but effective static mathematical model for the injected mass that obtains the following mathematical structure:

$$m_{inj} = T_{inj}\, r_1\, \sqrt{\Delta P} + (q_1\Delta P + q_0). \tag{3}$$

where: m_{inj} (mg) is the actual injected fuel mass, r_1, q_0 and q_1 are the mathematical model parameters to be tuned via experiments while ΔP (bar) is the pressure difference between the upstream EFI (rail pressure) and the downstream EFI (ambient pressure during tests).

Once the mathematical model of the injected fuel mass (Eq. 3) is available, the inverse relation describing the injection calibrated time as a function of a desired fuel mass is:

$$T_{inj} = \frac{m_d - q_1\Delta P - q_0}{r_1\sqrt{\Delta P}}, \tag{4}$$

where: m_d (mg) is the demanded fuel mass to be sprayed into ICE cylinder at the injection pressure ΔP determined by the AFR control duty (Corno et al., 2008; Gaeta et al., 2012).

3. THE MEASUREMENT WAVEFORMS OF THE FUEL-INJECTOR'S CURRENT PULSES

In Fig. 5 are shown waveforms of the current pulses for DI ICEs of fuel for two cases of the fuel-dose control.

In waveforms can be observed a very short forcing-current-

pulse rise time to peak value of approx. 11 A, without the use initial-load phase and the ability to obtain short time of opening of the EFI.
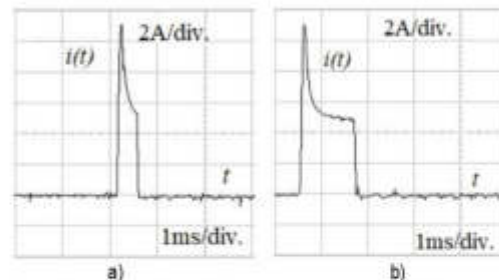
Fig. 5. Waveforms of the fuel-injector's current pulses for DI ICEs of fuel for two cases of the fuel-dose control: a) a small amount of fuel, b) a large amount of fuel

4. CONCLUSION

In the paper the fuel-injection mechatronic controller for DI ICEs that provides a very short rise time of the EFI's electro-magnet solenoid coil current in the excitation phase of opening without the use of initial-load phase, which improves the reliability of the electronic driver and simplifies its construction, is presented.

Thanks to the numerous advantages of the above-described fuel-injection mechatronic control system and method for controlling the fuel injection can be utilized for DI ICEs with the EFIs suitable for different liquid and/or gas fuels, for example, gasoline, diesel-oil, alcohol, methane, liquefied petroleum gas (LPG) and so on.

REFERENCES

1. **Achleitner E., Bcker H., Funaioli A.** (2007), Direct injection systems for otto engines, *SAE Technical Paper*, 2007-01-1416.
2. **Berndorfer A., Breuer S., Piock W., Bacho P.** (2013), Diffusion Combustion Phenomena in GDi Engines caused by Injection Process, *SAE Technical Paper*, 2013-01-0261.
3. **Bosch R.** (2014), *Automotive Electrics and Automotive Electronics*, Springer Vieweg, 5th Edition, Springer, Germany,
4. **Boudy F., Seers P.** (2009), Impact of physical properties of bio-diesel on the injection process in a common-rail direct injection system, *Energy Conversion and Management*, 50(12), 2905-2912.
5. **Chatlatanagulchai W., Yaovaja K., Rhienprayoon S., Wannatong K.** (2010), Gain-scheduling integrator-augmented sliding-mode control of common-rail pressure in dieseldual-fuel engine, *SAE Technical Paper*, 2010-01-1573.
6. **Chen H., Gong X., Liu Q. F., Hu Y. F.** (2014), Triple-step method to design nonlinear controller for rail pressure of gasoline direct injection engines, *IET Control Theory & Applications*, 8(11), 948-959.
7. **Chladny R.R., Koch C.R.**, (2008), Flatness-based tracking of an electromechanical variable valve timing actuator with disturbance observer feedforward compensation, *IEEE Transactions on Control Systems Technology*, 16(4), 652 – 663.
8. **Como M., Savaresi S.M., Scattolini R., Comignaghi E., Sofia M., Palma A., Eduardo Sepe E.** (2008), Modelling, parameter identification and dynamics analysis of a common rail injection system for gasoline engines, *Proceedings of the 17th IFAC World Congress*, 8481-8486, Seoul, Korea.

9. Gaeta A., Fiengo G., Palladino A., Giglio V. (2009), A control oriented model of a common-rail system for gasoline direct injection engine, *Proceedings of the 28th Chinese Control Conference*, 6614-6619, Shanghai, China.

10. Gaeta A., Montanaro U. , Fiengo G. , Palladino A & Giglio V. (2012), A model-based gain scheduling approach for controlling the common-rail system for GDI engines, *International Journal of Control*, 85(4), 419-436.

11. Gaeta, A., Montanaro, U., Giglio, V. (2011), Model-based Control of the AFR for GDI Engines via Advanced Co-simulation: An Approach to Reduce the Development Cycle of Engine Control Systems, *Journal of Dynamic Systems, Measurement, and Control*, 133, 061006 (1-17).

12. Giorgetti N., Ripaccioli G., Bemporad A., Kolmanovsky I., and Hrovat D. (2006), Hybrid model predictive control of direct injection stratified charge engines, *IEEE/ASME Transactions on Mechatronics*, 11(5), 499-506.

13. Gupta, V.K., Zhang Z., Sun Z. (2011), Modeling and control of a novel pressure regulation mechanism for common rail fuel injection systems, *Applied Mathematical Modelling*, 35(7), 3473-3483.

14. Hoffmann G., Befrui B., Berndorfer A., Piock W. and Varble D. (2014), Varble Fuel System Pressure Increase for Enhanced Performance of GDi Multi Hole Injection Systems, *Delphi Automotive Published 04/01/2014 Copyright © 2014 SAE International J. Engines.*

15. Husted H., Spegar T., Spakowski J. (2014), The Effects of GDi Fuel Pressure on Fuel Economy, *SAE Technical Paper*, 2014-01-1438.

16. Jiangjian A., Gao Xiyan B., Yao Chunde C. (2006), An Experimental Study on Fuel Injection System and Emission of a Small GDI Engine, *Proceedings of the 2nd IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications*, 1-6.

17. Lino P., Maione B., Rizzo A. (2007), Nonlinear modelling and control of a common rail injection system for diesel engines, *Applied Mathematical Modelling*, 31(9), 1770-1784.

18. Montanaro U., Gaeta A., Giglio V. (2011), An MRAC approach for tracking and ripple attenuation of the common rail pressure for GDI engines, *Proceedings of the 18th IFAC World Congress*, 4173-4180, Milano, Italy.

19. Ra Y., Loeper, P., Andrie, M., Krieger, R., David E. Foster D. , Reitz R. and Durrett R. (2012), Gasoline DICI Engine Operation in the LTC Regime Using Triple-Pulse Injection, *SAE Int. J. Engines*, 5(3), 1109-1132.

20. Sellnau M., Sinnamon J., Hoyer K., Husted H. (2012), Full-Time Gasoline Direct-Injection Compression Ignition (GDCI) for High Efficiency and Low NOx and PM, *SAE Int. J. Engines*, 5(2), 300-314.

21. Sellnau M., Sinnamon J., Hoyer K., Kim, J., Cavotta M. and Harry Husted H. (2013), Part-Load Operation of Gasoline Direct-Injection Compression Ignition (GDCI) Engine, *SAE Technical Paper*, 2013-01-0272.

22. Sun Z-Y., Zhao J., ShiLeonid Y. and Ma G. (2015), Numerical investigation on transient flow and cavitation characteristic within nozzle during the oil drainage process for a high-pressure common-rail DI diesel engine, *Energy Conversion and Management*, 98, 507-517.

23. Tang H.J., Weng L., Dong Z.Y., Yan R. (2009), Adaptive and learning control for SI engine model with uncertainties, *IEEE/ASME Transactions on Mechatronics*, 14, 93-104.

24. Wen-Chang T., Peng-Cheng Y., (2011), Design of the Electrical Drive for the High-Pressure GDI Injector in a 500cc Motorbike Engine, *International Journal of Engineering and Industries*, 2(1), 70-83.

25. Yan F., Wang J. (2011), Common rail injection system iterative learning control based parameter calibration for accurate fuel injection quantity control, *International Journal of Automotive Technology*, 12(2), 149-157.

26. Zhao H. (2016) Advanced Direct Injection Combustion Engine Technologies and Development, *Elsevier*, Amsterdam.

351

[35]   Glasmachers H., Melbert, J., Koch, A., "Sensorless Movement Control of Solenoid Fuel Injectors," SAE Technical Paper Series, *New SI Engine and Component Design 2006* No. 2006-01-0407, 2006.

SAE TECHNICAL
PAPER SERIES

Sensorless Movement Control of Solenoid Fuel Injectors

H. Glasmachers and J. Melbert
Ruhr-University of Bochum, Germany

A. Koch
Siemens VDO Automotive AG, Germany

SAE *International*™

The Engineering Meetings Board has approved this paper for publication. It has successfully completed SAE's peer review process under the supervision of the session organizer. This process requires a minimum of three (3) reviews by industry experts.

For permission and licensing requests contact:

SAE Permissions
400 Commonwealth Drive
Warrendale, PA 15096-0001-USA
Email: permissions@sae.org
Tel: 724-772-4028
Fax: 724-776-3036

SAE.

Global Mobility Database®

All SAE papers, standards, and selected
books are abstracted and indexed in the
Global Mobility Database.

For multiple print copies contact:

SAE Customer Service
Tel: 877-606-7323 (inside USA and Canada)
Tel: 724-776-4970 (outside USA)
Fax: 724-776-0790
Email: CustomerService@sae.org

Copyright © 2006 SAE International

Persons wishing to submit papers to be considered for presentation or publication by SAE should send the manuscript or a 300 word abstract to Secretary, Engineering Meetings Board, SAE.

Printed in USA

2006-01-0407

Sensorless Movement Control of Solenoid Fuel Injectors

H. Glasmachers and J. Melbert
Ruhr-University of Bochum, Germany

A. Koch
Siemens VDO Automotive AG, Germany

ABSTRACT

Precise fuel metering and high linear flow ranges (LFR) are the key issues for advanced injection systems for diesel and gasoline engines. They ensure lower emission, lower noise and higher fuel efficiency. Fast-switching solenoid injectors show high valve needle impact velocities and thus bouncing at injector closing. This causes multiple parasitic injections which reduce injection precision and deteriorates emission and efficiency. Mechanical and hydraulic anti-bounce concepts reduce the injection dynamic range while electronic open-loop controlled concepts are instable. For this reason a new sensorless electronic closed-loop anti-bounce solution able to reduce bouncing effectively and providing robust soft landing under all operating conditions was developed.

INTRODUCTION

Current and future emission legislation limits together with the demand for low fuel consumption require a more effective combustion in modern automotive engines. Improved fuel injection plays a major role in this respect. Advanced fuel injection profiles for direct gasoline injection and diesel engines are subdivided into several phases: a pre-injection, a main injection and a post-injection phase.

There are two competing systems on the market: solenoid injectors and piezo-driven injectors. Compared to traditional solenoid injectors recently introduced piezo-driven injectors combine advantages in both dynamics and linearity. Nevertheless, further improvements in the switching behavior of solenoid injectors by means of electronic control strategies still show significant potential. Increased dynamics with shorter opening and closing times are bought at the price of higher injector valve needle velocities leading to higher impact impulses and increased bouncing effects after closing.

A solenoid fuel injector is usually driven by a standard current profile as shown in Figure 1. A high current pulse with short rise time accelerates the valve needle for fast

opening. A low hold current is sufficient to keep the valve in this condition. After current drop the needle is driven back by a mechanical spring towards its starting position. There is a substantial delay from current end until total closing which causes non-linearity between current pulse width and injection volume and determines the minimum level. Even for reduced spring forces significant bouncing after valve impact is observed. Further parasitic injections occur and the effective closing phase is much longer than the opening phase.
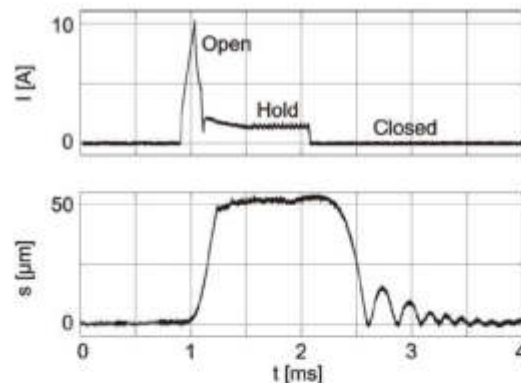
Fig. 1: Injector current profile and valve needle lift

Bouncing results from the mechanical construction of the injector. On principle, such injectors consist of a spring forcing the needle into the closed position and a solenoid together with the needle acting as an armature for opening. This spring-mass system is mainly damped by fuel. In case of injector closing there is no electric drive for needle movement control.

Figure 1 presents an elongation measurement at 50 bar fuel pressure. It shows that bouncing only occurs at injector closing. Bouncing can reach a maximum level of up to one third of total injector lift. It depends on injector

354

closing velocity. Bouncing is mainly a function of needle mass, spring stiffness and fuel pressure and also depends on mechanical tolerances of the injector and temperature of fuel and mechanics [4]. Thus, bouncing is a non-constant effect. Methods for calculation of bouncing influence rather than avoiding the disturbances have been published [1, 2].

Bouncing can be reduced by introducing hydraulic flow resistance into the fuel support. This leads to a limitation of upper injection volume per time and affects the final application. A controlled anti-force from the breaking current in the coil after lift off can compensate excessive spring force and is able to almost completely eliminate bouncing. However, the system is sensitive to parameter variation. This sensitivity does not allow open-loop control for the breaking current parameters.

This paper describes a new method for closed-loop control without any additional sensor. Solely analysis of solenoid current and voltage in the electronic control unit is required for closed-loop parameter calculation. In a similar way bounce suppression performance can be evaluated.

TEST BENCH
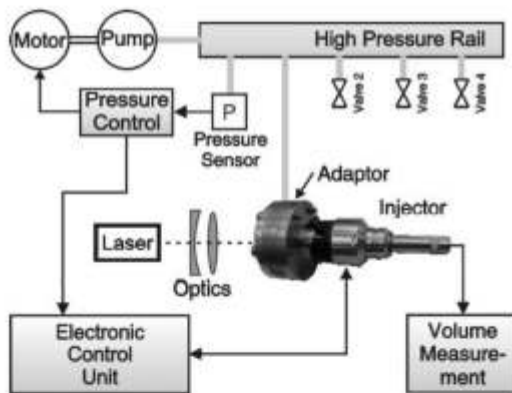
Fig. 2:   Test bench setup

In order to observe and control injector needle movement a test bench has been implemented. This test bench is capable of providing variable fuel pressures at the injector and an opportunity to observe needle movement during injection. Figure 2 illustrates the test bench setup. It contains components from the target engine (injector, fuel rail, fuel pump, pressure sensor and pressure valve). An electronic control unit keeps the rail pressure at constant values between 0 and 120 bar. This test bench provides realistic pressure fluctuations for optimal investigation in injector movement control.

A laser interferometer combined with an adapted waveguide and optics is used for the online needle

observation. Flexible current profiles can be generated from a programmable electronic power stage. The injector drive circuitry is shown in figure 3. It also allows to derivate control signals from the injector coil. It works with supply voltages of up to 80 V and can handle current levels of up to ±20 A. The output stage configuration is a full bridge topology for four-quadrant operation concerning coil voltage and current.
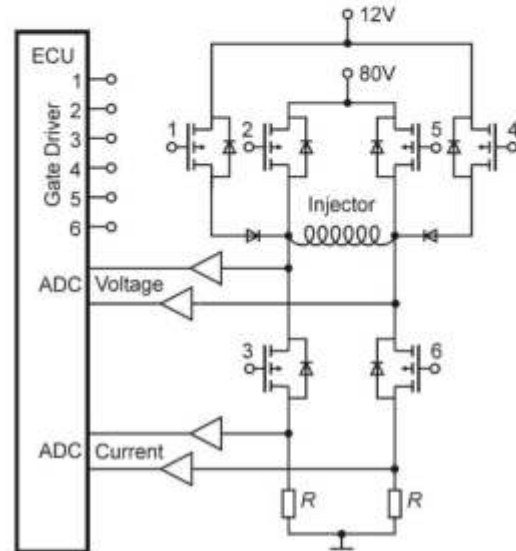
Fig. 3:   Injector drive circuitry

BOUNCE REDUCTION APPROACHES

There are various approaches to reduce injector bouncing. The most evident mechanical approach limits the flow diameter for fuel support. This reduces transient hydraulic forces on the needle during closing and decreases needle impact energy. Consequently, the maximum flow rate is also affected. This concept allows only a compromise between bouncing magnitude and fuel dynamics and cannot completely eliminate bouncing over the full operating range. Furthermore, such flow resistance elements must have tight mechanical tolerances and lead to higher production cost.

ELECTRONIC OPEN-LOOP CONTROLLED APPROACH

Electronic methods decelerate the injector needle after lift off by adding auxiliary energy to the injector during injector closing [5]. In this investigation a breaking current is used. This causes a magnetic force in the direction opposite to the spring force. Delay time after hold current switch off, peak value and duration are the breaking current parameters.

355
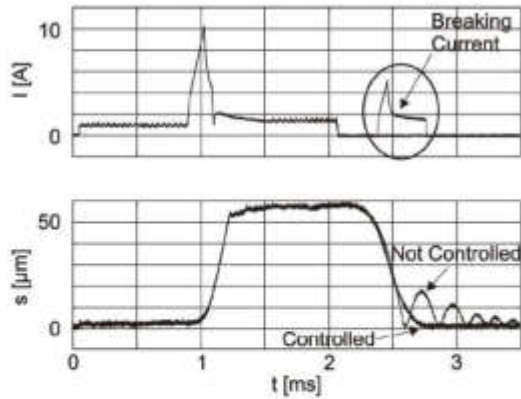
Fig. 4: *Bouncing reduced by breaking current*

Figure 4 shows injector needle movement at 50 bar fuel pressure for optimum parameters. Without any breaking current needle movement shows significant ripple. By applying an appropriate current waveform bouncing can be almost completely eliminated. Although the point of first touch down is slightly delayed, the effective opening time respectively the injected volume is reduced as all parasitic injections are fully eliminated.

A zoomed image of the relevant impact phase is presented in Figure 5. The illustration starts at the end of hold phase. Needle velocity and needle elongation are shown. Needle impact velocity without control is about 0.2 m/s and it is close to zero in the controlled mode. Open-loop controlled bounce reduction under stable conditions can reach an impact velocity reduction of more that 95%. It is possible to almost eliminate bouncing for the full pressure range from 40 bar to 120 bar.
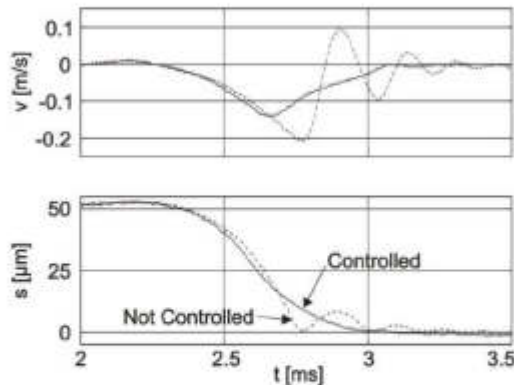
Fig. 5: *Zoom of bouncing reduced by breaking current*

SENSITIVITY

Breaking current parameters for proper impact are influenced by variable system parameters such as spring stiffness, fuel pressure, mechanical tolerances of the injector and temperature of fuel and mechanics. Fuel pressure variation is the predominant effect. In order to investigate this effect, movement variation as a function of pressure variation is analyzed. For each pressure there is an optimized set of parameters. Pressure variation at fixed parameters leads to suboptimal needle movement behavior. Reduction of pressure causes excessive breaking energy which is applied too early. Increase in pressure results in insufficient breaking energy which is applied too late. To achieve optimum bounce reduction at one given pressure, breaking current must be applied at one specific point in time with the right amount of energy.

Figure 6 illustrates bouncing for fixed breaking current parameters optimized for 50 bar under pressure variation of $\pm$ 5 bar.
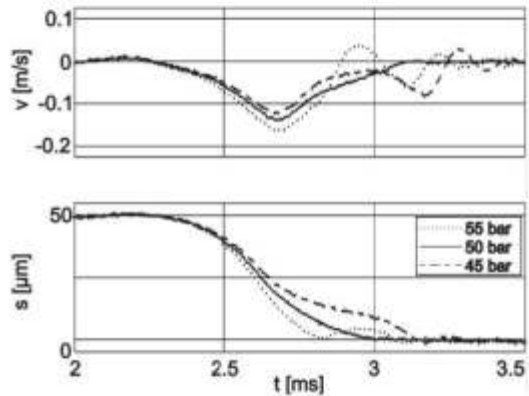
Fig. 6: *Sensitivity of breaking current parameters vs. pressure variations*

The solid line represents the correct pressure-to-breaking-current ratio. If there is breaking energy excess at lower pressure the needle velocity is reduced to almost zero before needle landing. This leads to unwanted additional closing delay and bouncing at the actual impact (dash-dotted line). Increased pressure results in a lack of breaking energy. Now the breaking current cannot decelerate the needle sufficiently for soft landing. This results in high impact velocities and bouncing (dotted line).

356
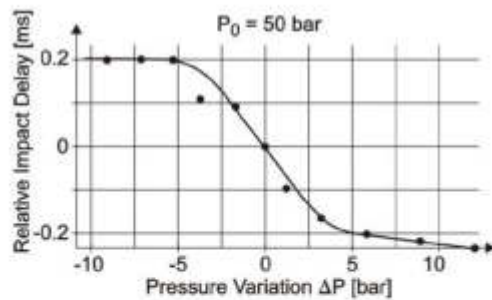
Fig. 7: *Additional impact delay versus pressure variation*

Figures 7 and 8 illustrate the influence of pressure deviation on impact delay and impact velocity. For negative pressure deviations there is a further delay in addition to the expected delay of needle soft landing. Needle impact velocity rises nearly parabolically in the ±4 bar range. It remains nearly constant at lower pressure and rises proportionally with higher pressure. Bouncing caused by too high impact velocity and additional needle closing delay at low pressure makes this point of operation unacceptable. High pressure variations result in nearly the same impact behavior as no bounce reduction.
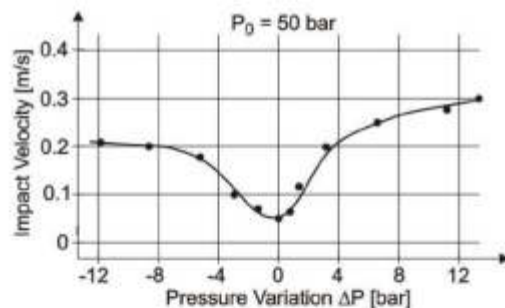
Fig. 8: *Impact velocity versus pressure variation*

From the fuel metering point of view bouncing is even worse than additional injector closing delay caused by bounce reduction, because the area below the elongation curves without bouncing is always lower than that of curves with bouncing. This area is nearly proportional to the amount of injected fuel. For this reason only the optimal amount of breaking energy leads to a suitable needle movement behavior. System sensitivity towards pressure variations of a few bar exclude open-loop controlled approaches for stable bounce reduction. Sensorless closed loop solutions without consideration of fuel pressure are already described [5]. Fuel pressure uncertainty at the injector due to rail pressure waves widely reduces precision of bounce prediction approaches. This limits the engine control unit to determine the precise amount of injected fuel.

CLOSED-LOOP CONTROLLED APPROACH

As fuel pressure variation represents the major disturbing effect a closed-loop controlled approach requires precise pressure information at each injector. The rail pressure control loop applies one pressure sensor which does not fulfill the requirements. Pressure pump effects and local variations in the rail caused by the different injections avoid the use of this displaced sensor. However, needle velocity information could also be used in order to determine optimum breaking current parameters. An implementation of a velocity or position sensor inside a fuel injector in series production is quite impossible due to enormous cost. One possible solution is to extract information from electrical solenoid signals. For this sensorless closed-loop approach injector current and voltage and their derivations can be used.

NEEDLE VELOCITY MEASUREMENT

There are several electromagnetic effects during needle movement which can be used to determine needle velocity. After the end of the hold current phase when the valve needle leaves the coil there is still residual magnetism left in the coil. This energy is reduced by eddy currents over a time period of some 100 microseconds. This energy is a function of the previous hold current level.

The influence of needle velocity on coil current and voltage can either be measured as a variation of induced voltage with open coil ends or as a variation of eddy current with short circuit coil ends. Both measurements are enabled by analog signal conditioning before analog-to-digital conversion in the test bench control electronics.

For a comparison, the efficiency of the two procedures can be described by their sensitivity. Sensitivity is the relative variation of the measured current or voltage versus pressure. Monitoring of the current of the shorted coil offers 4 times higher sensitivity compared to the induced open connection voltage.

Results presented in this paper are reached by di/dt-measurement for shorted coil. Closed-loop soft landing versus pressure fluctuations requires adaption of breaking current starting time and breaking current peak level. Proper breaking current parameters are taken from a lookup-table in a standard automotive microcontroller ST10 from ST-Microelectronics.

Analysis of the coil signals has shown that increased needle velocity increases the signal sensitivity. This relationship calls for a late observation point of time, as needle velocity rises with time during the injector closing. However, there must still be sufficient time for reaction in terms of breaking current fed into the coil before impact.

Valve needle closing velocity is not only a proper parameter for controlling pressure variations. It is also affected by aging, temperature variations and derivations of mechanical valve parameters. Thus closed-loop

controlled systems based on velocity information are capable of controlling needle movement versus any other disturbance. Needle velocity measurement is the basic information for closed-loop control.

CLOSED-LOOP CONTROLLED RESULTS

For implementation of closed-loop control in the test bench the conventional electronics for valve driving is extended by a signal conditioning stage which determines the short circuit coil current gradient after the hold phase. The breaking current parameters are derived from this observation.

Figure 9 shows results of closed-loop controlled needle movement under pressure variations. The diagram compares the needle movement for one uncontrolled actuation and for multiple controlled events. Mean pressure is 50 bar with a variation of ±10 bar.
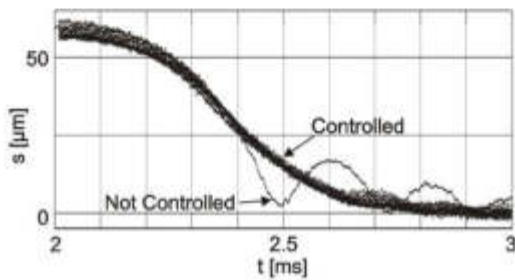
Fig. 9:   Closed-loop control results

All controlled events do not show bouncing as a result of different breaking current profiles. The system ensures sufficient robustness against pressure variations of about ±10 bar from one injection to the next. This is sufficient to control the instantaneous fluctuations at the individual valve. The rail pressure sensor is still used for rough pre-information and adaption of the current monitoring period.

Figure 10 presents needle elongation over time at different rail pressures. The system bounces at the full pressure range from 40 bar to 120 bar (dotted line) and bouncing can always be eliminated (solid line) by an optimized controlled breaking current.
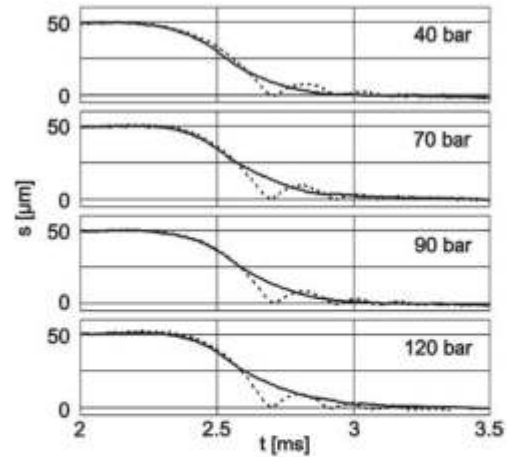
Fig. 10:   Bouncing at different pressures

BOUNCE DETECTION

Ensuring soft needle landing under all operating conditions must be fulfilled by the control system. Variations in environmental conditions such as temperature, system parameter drift, production deviations and aging require further long-term correction – particularly concerning the relatively high sensitivity towards parameter variations. Closed-loop long-term parameter correction is also useful to avoid extensive calibrations in production caused by production deviations. Parameter correction function can generally distress mechanical tolerances. If injector needle bouncing is observed the effect of the anti-bounce intervention can be analyzed and used for long-term parameter optimization. For needle bounce detection the same physical effects as for velocity measurement can be used. Instead of using the remaining magnetization of hold current now the residual magnetization of breaking current is observed. Extending the free-wheeling phase of breaking current permits analysis of bounce reduction performance. Soft impact control is not affected by this extension.

358
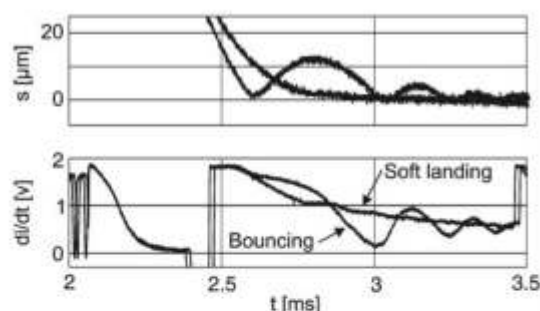
Fig. 11: *Bounce analysis*

Figure 11 illustrates bounce detection via di/dt-measurement after breaking current intervention. The two top curves show the needle movement during impact phase, one for a well controlled event and one for needle bouncing. The di/dt-curves below reflect the needle movement behavior. Bouncing can easily be detected and bounce intensity can be extracted. If breaking current parameters have been miscalculated due to system drift parameter correction is now possible. Bounce detection is a closed-loop system able to ensure system stability towards long-term parameter variations as well as to spontaneous limited pressure variations. By means of di/dt-observation an effective tool was found which allows fast reactions versus limited pressure variations for soft needle impact and adaptive long term compensation of system parameter drift.

CONCLUSION

Injector needle bouncing can be almost eliminated by applying an additional breaking current to a standard current profile after the hold phase. Parameters for breaking current can be derived from injector needle closing velocity. This information is taken from coil voltage or current measurement without the use of any further sensor. System sensitivity requires closed-loop control for safe and stable operation. Even compared to ideal needle movement the effective valve opening time of closed-loop control has not been extended. Closed-loop needle movement control not only eliminates bouncing, it also relaxes mechanical tolerance requirements. Closed-loop controlled long-term parameter correction is possible by the analysis of needle impact. This ensures high reliability and stable operation under all conditions. Thus, by little extra effort in signal conditioning on the ECU and by savings in injector design minimum system cost can be achieved.

ACKNOWLEDGMENTS

REFERENCES

[1] Becker, U.: *Verfahren zur Schaetzung des Nadelhubs eines Magnetventils*, Patent No. DE 19834405 (German)

[2] Ando, R., Koizumi, M., Ishikawa, T.: *Development of a Simulation Method for Dynamic Characteristics of Fuel Injector.* IEEE TRANSACTIONS ON MAGNETICS, VOL. 37, NO. 5, SEPTEMBER 2001 3715

[3] Robert Bosch GmbH: *Ottomotormanagement*, ISBN 3-8348-0037-6

[4] Passarini, L. C., Nakajima, P. R.: *Development of a High-Speed Solenoid Valve: an Investigation of the importance of the Armature Mass on the Dynamic Response.* ABCM, Vol. XXV, No. 4.

[5] Wright, D. O., Czimmek, P. R.: *Apparatus and method for electronically reducing the impact of an armature in a fuel injector.* US Patent No. 6,128,175

CONTACT

[1] Dipl.-Ing. Holger Glasmachers, Phd-Student
Holger.Glasmachers@lems.rub.de
Ruhr-Universitaet Bochum
Lehrstuhl fuer elektr. Mess- und Schaltungstechnik
IC-FW 03/453
Phone: +49 234 322 6983
Fax: +49 234 321 14168
Universitaetsstr. 150
Germany - 44780 Bochum

[2] Prof. Dr.-Ing. Joachim Melbert
Joachim.Melbert@lems.rub.de
Ruhr-Universitaet Bochum
Lehrstuhl fuer elektr. Mess- und Schaltungstechnik
IC-FW 03/461
Phone: +49 234 322 3137
Fax: +49 234 321 14168
Universitaetsstr. 150
Germany - 44780 Bochum

[3] Dipl.-Ing. Achim Koch
achimkoch@siemens.com
Siemens VDO Automotive AG
Phone: +49 941 790 3036, +49 172 8937 167
Fax: +49 941 790 13 3036
Siemensstrasse 12
Germany - 93055 Regensburg

http://www.lems.ruhr-uni-bochum.de

[36]   Tagliatela Scafati, F., Pirozzi, F., Cannavacciuolo, S., Alloca, L. et al., "Real Time Control of GDI Fuel Injection During Ballistic Operation Mode," SAE Technical Paper 2015-24-2428, doi: 10.4217/2015-24-2428, 2015.

SAE INTERNATIONAL

Real Time Control of GDI Fuel Injection during Ballistic Operation Mode	2015-24-2428
	Published 09/06/2015

Ferdinando Taglialatela Scafati, Francesco Pirozzi, and Salvatore Cannavacciuolo
APG-PT&S STMicroelectronics

Luigi Allocca and Alessandro Montanaro
Istituto Motori CNR

Abstract

Gasoline direct injection (GDI) combustion with un-throttled lean stratified operation allows to reduce engine toxic emissions and achieve significant benefits in terms of fuel consumption. However, use of gasoline stratified charges can lead to several problems, such as a high cycle-to-cycle variability and increased particle emissions. Use of multiple injection strategies allows to mitigate these problems, but it requires the injection of small fuel amounts forcing the traditional solenoid injectors to work in their "ballistic" region, where the correlation between coil energizing time and injected fuel amount becomes highly not linear.

In the present work a closed-loop control system able to manage the delivery of small quantities of fuel has been introduced. The control system is based on a particular feature found on the coil voltage command signal during the de-energizing phase. On the basis of this feature, the injector needle closing time and then, in turn, the actual amount of fuel injected can be calculated.

Results showed that the proposed control system, through a proper management of ballistic injections, has the potential to increase the minimum fuel injection capabilities of GDI solenoid injectors.

Introduction

Future emission regulations require the development of gasoline combustion engines with improved efficiency in order to obtain a strong reduction of the toxic emissions coupled to the reduction of fuel consumption and hence carbon dioxide emissions. The greatest fuel consumption benefit is achieved by means of systems such as the gasoline direct injection (GDI) combustion with un-throttled lean stratified operation. In this mode, the fuel is injected later in the compression stroke [1] allowing stable combustion of ultra-lean mixtures. However, use of gasoline stratified charges can lead to several problems. In particular, due to the oxygen excess in the combustion stroke, the NO_x emission levels are generally higher than in the port-fuel injection (PFI) engine or in homogeneous charge

direct injection. Moreover, short time for mixture preparation and spray wall impingements are responsible for a high cycle-to-cycle variability and high particle emissions. On the other hand, the reduction of the particulate at the exhaust of gasoline direct engines represents a crucial aspect also considering the introduction of EU6 emission legislation that strongly pushes toward a reduction of the particulate emitted by the engines.

A potential effective way to mitigate the problems of GDI stratified operation and reduce the wall impingement is the use of multiple fuel injections, splitting up the total fuel injection into several smaller (and shorter in duration) shots. A first effect of this approach is the reduction of the jet penetration into the combustion chamber thus reducing the wall wetting and decreasing the particulate formation. Moreover, it has been demonstrated [2] that the use of multiple injection strategies allows to retard the angle of 50% mass fraction burned (MBF50), placing it close to the thermodynamic optimum. This allows to have lower combustion peak temperatures with a reduction of NO_x emissions.

However, this is not easy to obtain by the traditional GDI solenoid injectors. Indeed, the management of small injections forces GDI solenoid injectors to work in their so called "ballistic" mode. The ballistic behavior appears at small injection pulse-width when the pulse is cut-off before the valve fully lifts up. During ballistic operation, the correlation between the electrical command and the injected amount of fuel is not linear, the valve motion is unstable and the fuel delivering cannot be controlled with optimum precision.

The present work introduces a closed-loop control system able to manage the delivery of small fuel amounts and to increase the minimum fuel injection capabilities of GDI solenoid injectors in order to extend their use to the multiple fuel injection strategies. Based on the features derived from the injector current signal, the proposed control architecture calculates the needle closing timing providing a real time information about the actual amount of fuel

injected. This latter can then be compared to absolute target values and, as a result, the length of the electrical command can be consequently adjusted.

Experimental Equipment

A Bosch GDI multi-hole solenoid injector has been used for the experimental tests. The injector was driven by means of a smart power device (STMicroelectronics- *L9781*) equipped with a dedicated Graphical User Interface, which allowed to program the injector command current levels and to select the desired coil energizing times. The current profile used to drive the injector included a *pick current* level of 12 A and two *hold current* levels, the first one at 5 A and the second one at 2.5 A.

Specific test points on the Evaluation Board of L9781 allowed to acquire, during electrical actuation, three different types of coil voltage signals:

- *High-side voltage signal*, representing the *source* voltage of the High Side *MOSFET B* of L9781.
- *Low-side voltage signal*, representing the *drain* voltage of the Low-Side *MOSFET* of L9781.
- *Differential voltage signal*, obtained by the difference between High-side and Low-side voltage signals. It represented the actual coil voltage command signal.

The gasoline mass flow rate has been measured by means of an AVL Fuel Injection Gauge Rate meter working on the Bosch tube principle [3, 4]. The fuel was injected into a small chamber kept at the constant pressure of 5 bar and connected to a pipeline of constant internal section, 6 m long. The AVL GM12D pressure transducer, located downstream the nozzle, collected the pressure variations produced by the incoming fuel. The generated pressure variation, Δp, was proportional to the injected quantity, q, through the relationship:

$$q = \frac{\Delta p \cdot A_{tube}}{a \cdot \rho} ,$$

(1)

where ρ was the fuel density, A_{tube} the inner section of the pipe and a the speed of the sound in the fuel at ambient temperature. The time resolution of the acquisition chain was 0.01 ms, consistent with the 130 kHz natural frequency cut of the transducer (0.0076 ms). The instantaneous value of q gave the injected quantity in the interval of 0.01 ms, while the total amount per shot was the integral over the entire pulse duration. The injection rate measurements have been averaged over 100 shots to absorb the cycle-to-cycle variations while the total amount has been compared with that one collected at the discharge pipe and weighted by a precision balance (OHAUS 410GX1 - range 410 g, precision 1.0 mg).

Results

Figure 1 shows the characteristic curves of the injection rate obtained for two different injectors belonging to the same family. The graphs represent a sample of the typical dispersion. For both the injectors, the correlation between the pulse command duration and the injected fuel amount appears linear and repetitive for pulse durations greater than 400 ms and for injected fuel greater than 2.8 mg. For command pulse widths lower than 400 ms, the correlation becomes highly non-linear and, for some pulse durations, may also invert the trend and decrease the fuel quantity at the increasing of the electrical command. In the nonlinear region, for very short injection command pulses, the injector needle starts the closure before having reached the widest lift position: this effect is known as *ballistic mode*. It has been reported [5] that the causes of the non-linear behavior mainly rely on the inertia of the injector spring-mass system, the reduction of the electro-magnetic forces exerted by the coil, friction variations and more. All these have an unpredictable effect on the dynamic of the needle lift due to manufacturing tolerances and ageing effects, too. This is the reason for which, up to now, the nonlinear region has not been used in the traditional injection strategies for commercial GDI engines.

In order to achieve the desired injection target also during ballistic operation, and to extend the use of solenoid injectors to short injections, real time information about the actual fuel amount delivered at ballistic are needed. On the basis of these information, the injector energizing time can then be adjusted in real time by means of a closed loop algorithm to obtain the requested injected quantity.
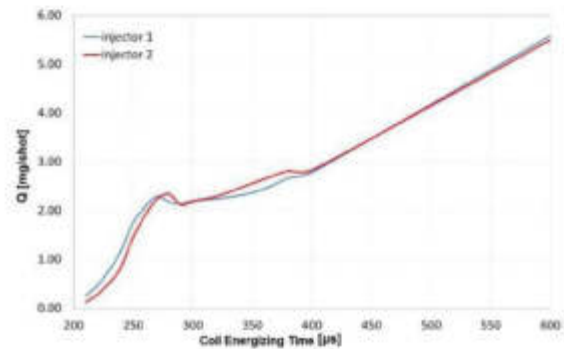
Figure 1. Characteristic curves of the injection rate obtained for two injectors belonging to the same family.

The graphs in figure 2 and figure 3 illustrate a comparison between the injector *low-side* voltage signal and the corresponding injected mass flow rate obtained for a coil energizing time of 300 μs. The analysis of the two curves depicts that the voltage signal shows an inflection at the same time when the injected mass flow rate is annulled. This feature was confirmed also for other coil energizing times.
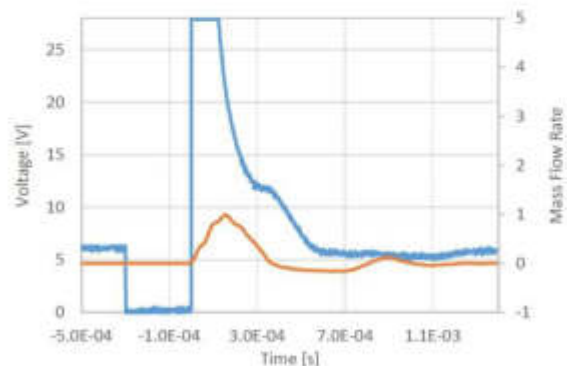
Figure 2. Comparison between injector low-side voltage signal and the corresponding injected mass flow rate, for a coil energizing time of 300 μs.

361
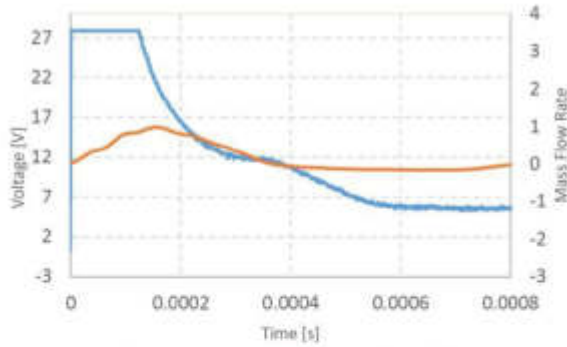
Figure 3. Comparison between injector low-side voltage signal and the corresponding injected mass flow rate for a coil energizing time of 300 μs

A similar correlation was also found comparing the mass flow rate curve with the *differential* voltage command signal. Also this latter signal showed an inflection when the fuel mass flow reached its zero value (see figure 4).
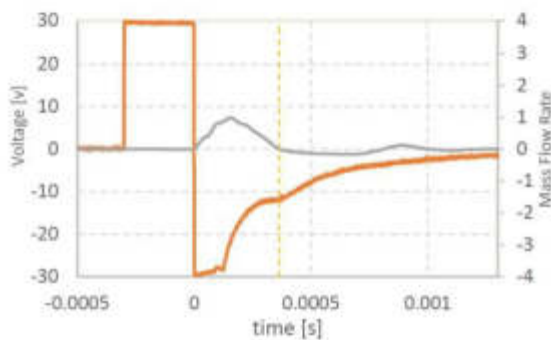
Figure 4. Comparison between differential voltage command signal and the corresponding injected mass flow rate, for a coil energizing time of 300 μs

For both low-side and differential voltage signals, the inflection occurred during the *switch-off phase*, when the injector coil was de-energized and a self-induction voltage was created. In this phase, the voltage signal typically includes, other than a contribution due to decaying eddy currents, a contribution due to the movement of the needle in the de-energized coil, which has an amplitude depending on the needle speed. As this speed reaches its maximum value directly before the needle closing, an inflection on the voltage signal appears at this time. On the basis of this feature, it comes out that a proper processing of the injector voltage signal allows to have real-time information about the needle closing time and then, in turn, on the actual amount of injected fuel. This can then be used to real-time adjust fuel quantity to the desired target value during the ballistic operation mode.

Algorithm for Real-Time Compensation of Ballistic Injections

Figure 5 shows a possible control architecture for real time detection and compensation of the fuel quantity injected during ballistic operation.
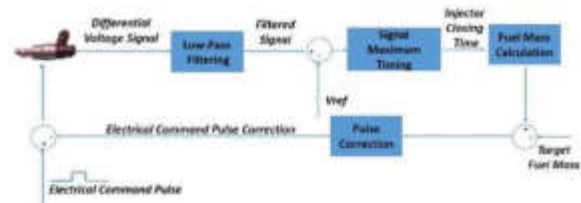
Figure 5. Closed-loop control algorithm for real time compensation of fuel quantity injected during ballistic operation.

A first part of this control algorithm was implemented and tested in this work.

As indicated in the figure, the *differential* voltage signal in the de-energized coil was firstly filtered by means of a low-pass filter to reduce the superimposed noise. Then, the filtered signal was compared with a reference voltage signal (*Vref*).

The reference signal *Vref* represented the voltage induced in the de-energized coil using an electrical command that did not produce any movement of the injector needle (e.g. due to a too short command pulse). So, this signal exclusively was related to the self-induced voltage due to the decaying eddy currents and it did not include the above mentioned component due to the needle motion. Then, the difference between the voltage signal and the reference signal allowed to separate the contribution due to the needle movement and, as a consequence, the needle closing time could be determined with a higher level of accuracy. In a real application, a reference voltage signal can be either generated and implemented during the calibration phase of the engine or periodically generated by the ECU in order to take also into account the ageing effects of the injector.

The signal obtained by the difference between the *differential* coil voltage signal and the reference voltage signal was then processed to calculate the time of its maximum value (or minimum, depending on the sign of the difference). This time corresponded to the injector needle closing time.

Being able to measure the time of needle closing means being able to measure the injected fuel mass. In fact, it can be demonstrated that for solenoid injectors, when there is no variation in opening delay, fuel mass amount is directly correlated with needle closing time. For the injectors used in the experimental tests of this work, the relation between needle closing time and the corresponding amount of fuel injected is depicted in figure 6.

The graph clearly shows that a linear correlation (R-Square = 0.98) exists between injection duration (i.e. the time elapsing from the starting of the electrical command to the complete needle closing) and the corresponding mass of fuel injected. This correlation does not depend on the particular injector (of the same family) used and it is also valid at ballistic.

362
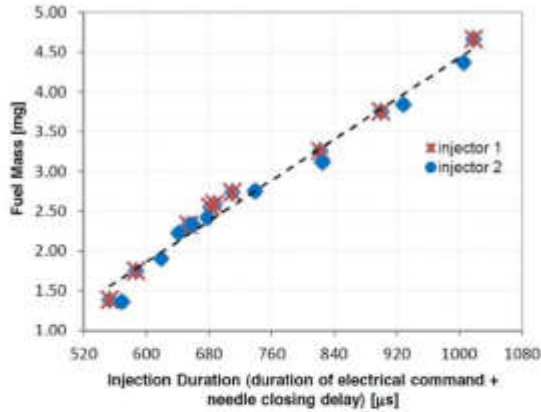
Figure 6. Correlation between injection time duration and delivered fuel mass for the injectors used for experimental tests.

On the basis of this correlation, the algorithm, once that injection duration has been detected from the voltage signal, calculates the actual amount fuel amount injected. This fuel mass value can be subsequently compared with a target value previously defined and, as a result of this comparison, a correction value for the coil energizing time can be defined. However, this latter part of the algorithm (related to the error compensation) has not be performed at this stage and will be the object of a future development.

The whole algorithm proposed in this paper has been implemented on a STM32 microcontroller (produced by STMicroelectronics) using its evaluation board *STM32F3 DISCOVERY*. In the following sections, details on implementation of some parts of the algorithm are provided.

Differential Voltage Signal Filtering

In order to reduce superimposed noise on the difference voltage signal, a proper digital filter has been designed. To this aim, a spectral analysis of the signal was performed by means of a *Fast Fourier Transform* (FFT) (see figure 7).
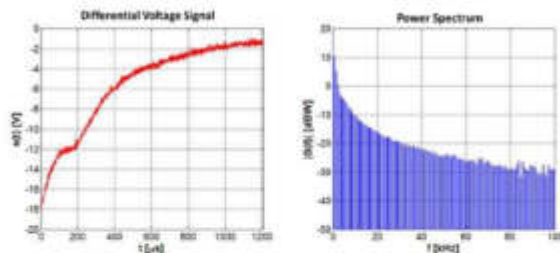
Figure 7. FFT of differential voltage signal for a coil command time of 300 μs

Figure 8 shows the behavior of the integral over the frequency of Power Spectral Density (PSD) of the *differential* voltage signal normalized to the maximum power. The figure refers to a coil energizing time of 300 μs, but a similar behavior has been also found for other electrical command pulse durations.

On the basis of results of the spectral analysis, a low-pass moving average filter with 40 coefficients and a cut-off frequency of 12 kHz has been chosen. As can be deduced from the analysis of the graph in figure 8, this latter was the frequency at which power spectrum of differential voltage signal reached the 95% of the maximum power.
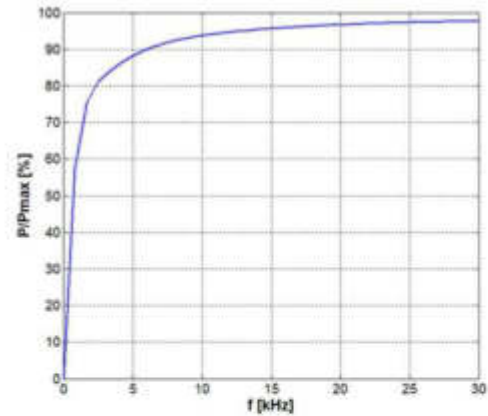
Figure 8. Integral of PSD of differential voltage signal normalized to the maximum power. Electrical command pulse = 300 μs

Calculation of Injector Needle Closing Time

Figure 9 shows an example of signals obtained by a difference between the filtered *differential* voltage signal and the reference signal (*Vref*). As mentioned above, the injector needle closing time has been calculated by determining the time at which the maximum of this difference signal occurred.
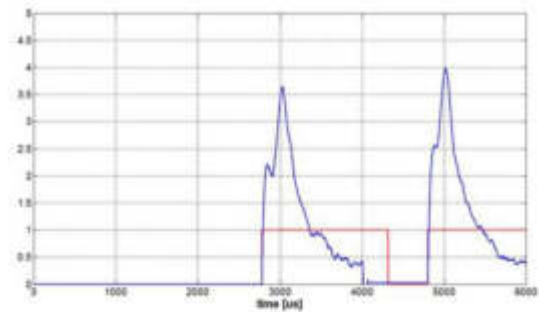
Figure 9. Signals obtained by the difference between the filtered differential voltage signal and a reference voltage (Vref). The corresponding electrical command pulses, translated to the beginning of the corresponding difference signals, are also shown in the figure (red line).

The implementation on the *STM32* microcontroller of the algorithm for the calculation of injector needle closing time required the use of the following peripherals of the microcontroller:

- a **comparator**, to enable the functioning of the Analog to Digital Converter (ADC) when a voltage threshold level was achieved. Use of the comparator allowed to restrict the acquisition of the coil voltage signals to the *switch-off phase*, i.e. when the injector coil was de-energized.
- an **Analog to Digital Converter** (ADC), devoted to the acquisition of the coil voltage signals.
- a **Timer**, used to count the time elapsing between the end of the electrical command pulse and the needle closing time.

Results about the performances of the implemented algorithm are illustrated in table 1. In particular, the table shows, for several coil energizing times, the results of a comparison between the time at

363

which mass flow rate was annulled (t_{flow}) and corresponding estimated needle closing time (t_{needle}) calculated by means of the proposed algorithm.

Table 1 Comparison between the time at which mass flow rate is annulled and the corresponding estimated needle closing time

T_{CMD} (μs)	$|tflow - tneedle|$ (μs)
230	15
235	10
240	2
250	12
260	6
270	7
280	2
290	4
300	12
320	10
340	7
380	6
400	5

As reported in the table, the maximum error between the estimated needle closing time and the time at which mass flow rate is annulled was 15 μs; the average error was 7.5 μs, with a standard deviation of 4 μs. These results confirm that the proposed algorithm has the capability to predict the actual fuel amount delivered during ballistic operation allowing a real time correction to the desired injection target. Implementation of the correction part of the algorithm and its testing on a GDI engine will be objects of a future development.

Conclusions

Multiple injection strategies for GDI engines require the injection of small fuel amounts, forcing the traditional solenoid injectors to work in their "ballistic" region, where the correlation between coil energizing time and injected fuel amount becomes highly not linear.

In the present work, a closed-loop control system able to manage the delivery of small quantities of fuel has been introduced. The control system, through a proper management of the ballistic injections, allows to increase the minimum fuel injection capabilities of GDI solenoid injectors

The control system is based on a particular feature found on the coil voltage command signal during the de-energizing phase. In particular, it has been found that, in correspondence of the injector needle closing time, an inflection on the coil voltage signal appears.

On the basis of this feature, the proposed algorithm through a proper processing of the injector voltage signal, allows to have real-time information about the needle closing time and then, in turn, on the actual amount of fuel injected.

The whole algorithm has been implemented on a STM32 microcontroller using some dedicated peripherals.

Details about implementation of some parts of the algorithm have been provided in the present work. In particular, details about coil voltage signal filtering and injector needle closing time detection have been described.

The implemented algorithm was able to detect the time at which fuel mass flow rate was annulled with an average error of 7.5 μs and a standard deviation of 4.5 μs.

References

1. Schwarz, C., Schünemann, E., Durst, B., Fischer, J. et al., "Potentials of the Spray-Guided BMW DI Combustion System," SAE Technical Paper 2006-01-1265, 2006, doi:10.4271/2006-01-1265.

2. King, J., Schmidt, L., Stokes, J., Seabrook, J., Nor, F., Sahadan, S. "Multiple injection and boosting benefits for improved fuel consumption on a Spray Guided Direct Injection gasoline engine", Proceedings of the FISITA 2012 World Automotive Congress.

3. Bosch, W., "The Fuel Rate Indicator: A New Measuring Instrument For Display of the Characteristics of Individual Injection," SAE Technical Paper 660749, 1966, doi:10.4271/660749.

4. Wallace, I. "Injection Rate Gauge: Pass Off Information and User Instructions" - Fuel & Engine Management Systems, Graz - December 2002.

5. Parotto, M., Sgatti, S., and Sensi, F., "Advanced GDI Injector Control with Extended Dynamic Range," SAE Technical Paper 2013-01-0258, 2013, doi:10.4271/2013-01-0258.

Abbreviations

GDI - gasoline direct injection

PSD - power spectral density

MBF50 - 50% of mass fraction burned

ADC - analog to digital converter

[37] Cvetkovic, D., Cosic, I., Subic, A., "Improved Performance of Electromagnetic Fuel Injector Solenoid Actuator Using a Modelling Approach," *International Journal of Applied Electromagnetism and Mechanic*, Vol. 27, pg. 251-273, IOS Press, Australia, 2008.

Improved performance of the electromagnetic fuel injector solenoid actuator using a modelling approach

Dean Cvetkovic[a,*], Irena Cosic[a] and Aleksandar Subic[b]

[a]*School of Electrical and Computer Engineering, RMIT University, GPO Box 2476V, Melbourne, VIC 3001, Australia*

[b]*School of Aerospace, Mechanical and Manufacturing Technology, RMIT University, GPO Box 2476V, Melbourne, VIC 3001, Australia*

Abstract. The aim of this investigation was to develop a specific modelling approach capable of reducing the size of the fuel injector solenoid device while improving its response time and attraction force. Several developed modelling and simulation procedures have focused on various aspects of solenoid component modifications in order to develop an evolution process of miniaturising a fuel injector solenoid, using the latest finite element method (FEM) tool software. The specific factors that influenced the optimum operation of the fuel injector solenoid were the geometrical shape of individual solenoid components, material properties, air-gap constraints, boundary conditions, current source conditions, mass constants, and damping coefficients of the plunger. The attraction force distribution in the main air-gap was directly influenced by the taper angle of the 2D and 3D plunger pole faces, plunger length and the permanent magnetism. The precise definition of the electro-mechanical motion of plunger was of enormous importance in reducing the fuel injector solenoid response time, closely related to stroke and mass of the plunger, spring characteristics, motion and rebound of the plunger. Using the developed approach, the initial size of the fuel injector has been reduced by 35%, the attraction force increased by 26% and the response time reduced by 76%. However, by frequently repeating the design trials and conducting a thorough experimental investigation, the final minimum response time was achieved by the virtual rebound delay model. The reduction in response time from the 'optimal' experimental to virtual model was by 35%. The simplicity and effectiveness of the developed methods, allowed for quick and accurate evaluation.

Keywords: Modelling, miniaturisation, response time, electro-mechanical, fuel injector, solenoid, FEM, experiment

1. Introduction

Today's small, high-speed engine demands a lightweight and compact fuel-injectors. In recent years, the automotive industry focus has been on compact fuel injector solenoids. If the assembly of the existing fuel injector could be made simpler and smaller in size, significant reduction of costs involved in the production of fuel injectors would be achieved, with lower power consumption, fast acting operation, lower maintenance and lower noise and vibration levels [1,2].

Traditional tools for simulating circuits are clumsy at best in predicting the behaviour of fuel injector solenoid components. The use of ordinary circuit equations in order to predict the behaviour of fast acting fuel injector solenoids are often inaccurate and time consuming. However, sophisticated computational

*Corresponding author. Tel.: +61 3 9925 9641; Fax: +61 3 9925 2007; E-mail: dean.cvetkovic@rmit.edu.au.

tools that combine circuit theory for analysing currents, voltages, equations of motion to compute mechanical interactions, and Finite Element Analysis (FEA) to compute magnetic fields and compensate for non-linearities, have proved to be extremely useful in the design of electromagnetic solenoid actuators for improved fuel injector performance [3–6].

In this research investigation, modelling and simulation of fast acting fuel injector solenoids was undertaken using Ansoft's Maxwell 2D/3D and Electro-mechanical System Simulator (EMSS) software tool in order to simultaneously reduce its size and improve its operational performance [4–6]. The general procedure for designing the fast acting fuel injector solenoid began with roughing out a first-pass model for FEA [7,8]. This included picking preliminary geometry for the fuel injector solenoid, such as the coil, and associated mechanical parts, particularly any that moved. In the particular case of this research study, a computational model of the basic physical fuel injector solenoid was constructed. This model was in the form of a 2D and 3D cross section. In order to get the FEA process rolling, it was essential to make several assumptions, such as the range of drive currents, materials and their attributes, and boundary conditions, which dictated how the magnetic fields behave at solenoid component interfaces and at the edges [9,10]. The design of fuel injector solenoids frequently involved calculating attraction force and response time. The analyses to provide a single simulation of both the electromechanical component and circuitry of fuel injector solenoid system was typically done using EMSS in order to answer the questions of whether the plunger closes quickly enough, how much energy the device needs, and what waveshape typifies the transient current running through the coil after closure.

2. Theoretical background of solenoid actuators

Solenoid actuators are electro-mechanical devices that convert electrical energy into mechanical energy associated with linear motion [11]. Solenoids consist of magnetic components and other material characterised by a non-linear behaviour. As solenoid component dimensions get smaller, non-linear effects have a bigger impact on operational performance. Generally, solenoid devices are used in a wide range of mechanical applications, such as fuel-injection systems, starting motors and ABS systems [12]. In this research investigation, a solenoid was applied in a fast acting fuel injection system. The purpose of a solenoid, specially designed electromagnet was to control how long the injector is open and therefore how much fuel goes into the system.

There is more focus now on simple devices such as solenoid actuators to make the fuel injectors more efficient and smaller, which would result in smaller fuel injectors, and eventually in smaller engines. However, by miniaturising the solenoid actuator it is of greatest interest to improve the operational performance. The operational performance of the fuel injector solenoid depends on two major factors, such as the response time and attraction force of the plunger [13,14]. The fast actuation of fuel injector solenoid is influenced by the response time or the elapsed time from the initial switch on, until the plunger is in the fully attracted position. The energising time for a solenoid to complete a given stroke, measured from the beginning of the initial pulse to the seated or energised position, is influenced by the attraction force of the plunger. The stronger the attraction force is, so would the response time or the energised time be minimised. From this relation the fuel consumption can be minimised and the overall engine operation can be improved.

Qualitative mathematical relations describe the behavior of solenoid actuator system in terms of the ordinal relationships, such as more or less, bigger or smaller, between quantities [15]. In order to compute the maximum attraction force or shortest response time, analysing the qualitative relations

Table 1
The list of qualitative mathematical reasoning statements

Reasoning statements	Description
1. Coil → Turns + MMF	The increase in number of coil turns would result in an increase of the magnetomotive force (MMF).
2. Stroke Length → Air-Gap − MMF + Force	If the air-gap between the plunger and plugnut is minimised, the MMF will increase, resulting in a higher attraction force.
3. Input Voltage + Current + MMF + Force	If the input voltage and current are both increased; the MMF is increased, resulting in higher attraction force.
4. Coil → Type → Resistance − Current + MMF + Force	Depending on the coil type, the reduction of resistance, results in current, MMF and attraction force.
5. Coil → Type → Resistance + Current − Response Time − Force	If the coil resistance is increased (heat increase), the current flow would be reduced, resulting in a reduction of attraction force.
6. Plunger − Response Time − Force	Reducing the weight of the plunger would require smaller attraction force and therefore reducing the response time.
7. Input Voltage → Current − MMF − Response Time	By reducing the input voltage, it results in the reduction of the current flow and MMF. As a result, the response time will be reduced.
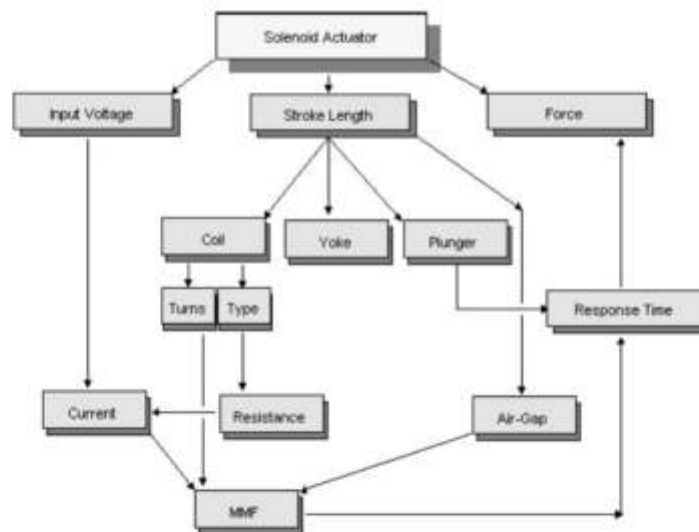
Fig. 1. Relationships and its interdependencies between solenoid model parameters.

could determine how changes in one parameter can result in changes in other parameters (Fig. 1 and Table 1).

The first four reasoning statements focus on the virtual manipulation of solenoid input parameters, such as coil turns, air-gap, input voltage, current and coil type. The output of each of these statements is
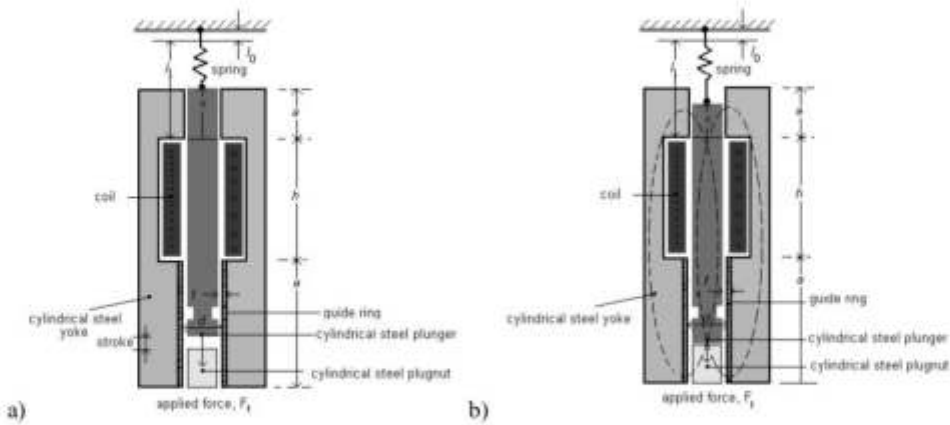
Fig. 2. Solenoid actuator's coil in de-energised (a) and energised (b) positions. The solenoid consists of a cylindrical plunger of mass M, which moves vertically in a free-space guide rings of thickness t and mean diameter d. The plunger is supported by a spring of stiffness K. Its unstrached length is l_0. A mechanical load force F_t is applied to the plunger via the mechanical system connected to it. The coil has N turns and resistance R. Considering the electrical system of the solenoid, its terminal voltage is v_t, and its current is i. The effects of magnetic leakage and reluctance of the steel are negligible. The reluctance of the magnetic circuit is that of the two guide rings in series, with the flux directed radially through them, represented by the dotted flux lines Φ.

mainly appropriate for MMF and attraction force results. By applying specific input values (more or less) to a solenoid system, the output will either be increased or decreased. The last three statements focus much more on the practical solenoid principles and response time output behavior. These relationships are central to modelling and simulation of fuel injector solenoids.

The operation of a solenoid actuator is based on two main principles, energised and de-energised operation [16]. When the switch is open, as shown in Fig. 2(a), current does not flow through the coil. In this case, the plunger is forced upwards by the spring. There is a gap (equal to the maximum stroke length) between the plunger and the stationary plugnut. When the switch is closed, as shown in Fig. 2(b), current flows through the coil. A magnetic field is induced through the center of the coil, through the stationary plugnut, yoke, into the plunger, and across the air-gap between the plunger and plugnut to complete the magnetic circuit. An attraction force is produced in the air-gap, which pulls the plunger against the return spring until the gap is closed. Once the switch is opened, the magnetic field collapses and the force holding the plunger against the plugnut no longer exists, permitting the return spring to return the plunger to its rest position.

The magnetic circuit of the solenoid actuator is made up of several components, including plunger, plugnut, coil and yoke. The analogy of the magnetic circuit to the resistance of the electrical circuit is termed reluctance. Therefore, reluctance is a measure of the opposition offered by a magnetic circuit to the setting up of the flux. Ohm's law for an electrical circuit is analogous to that for a magnetic circuit [17]:

$$Reluctance = MMF/Flux \tag{1}$$
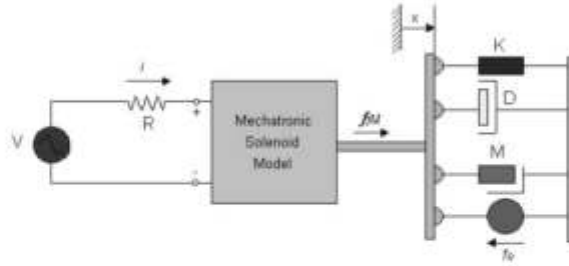
Fig. 3. Model of EMSS (Electro Mechanical System Simulator).

Since $t \ll d$, the flux density in the guide rings is nearly constant with the respect to radial distance. The reluctances of the upper and lower gaps are:

$$R_1 = \frac{t}{\mu_0 \cdot \pi \cdot dx} \quad , \quad R_2 = \frac{t}{\mu_0 \cdot \pi \cdot da} \tag{2}$$

The field is assumed to be concentrated in the area between the upper end of the plunger and the lower end of the upper guide ring [17]. The total reluctance is

$$R_T = R_1 + R_2 = \frac{t}{\mu_0 \cdot \pi \cdot d} \left(\frac{1}{a} + \frac{1}{x} \right) = \frac{t}{\mu_0 \cdot \pi \cdot da} \frac{a+x}{x} \tag{3}$$

The permeance is

$$P = \frac{1}{R} = \frac{\mu_0 \cdot \pi \cdot da}{t} \frac{x}{a+x} \tag{4}$$

Hence, the inductance is

$$L = N^2 P = \frac{\mu_0 \cdot \pi \cdot da \cdot N^2}{t} \frac{x}{a+x} = L' \frac{x}{a+x}, \text{ where } L' = \frac{\mu_0 \cdot \pi \cdot da \cdot N^2}{t} \tag{5}$$

The attraction force between the plunger and the stationary plugnut depends on the efficiency and characteristics of the magnetic circuit. It is important, therefore, that the effects of internal and external changes affecting these characteristics are understood and described in detail.

The mechatronic system of the solenoid actuator consists of three main subsystems: mechanical, electrical, and magnetic subsystem (solenoid model), as shown in Fig. 3.

The coupling term of the magnetic force f_{fld}, produced by the magnetic field when the coil is energised, and induced electro motive force (emf), e, can be expressed in terms of air gap distance, x, and current of the coil, i [18]. The magnetic force acting downward in this case, on the plunger, in the negative direction of x is

$$f_{fld} = \frac{-\partial W'_{fld}(i,x)}{\partial x} = -\frac{1}{2}i^2 \frac{dL}{dx} = -\frac{1}{2}L' \frac{ai^2}{(a+x)^2}, \text{ where } W'_{fld}(i,x) = \int_0^i \lambda(i,x) \cdot di \tag{6}$$

369

where W' is co-energy obtained by integrating the flux linkage against current, and flux linkage variable dependent, λ_l [17].

The establishment of the voltage across the coil windings is an important factor in the formulation of the electric circuit equations. The electrical subsystem is represented by a voltage source and resistance R. The source could be represented by a current source and a parallel conductance G. The lumped parameter network model is given by

$$V = IR + \frac{d\lambda_l}{dt} = IR + \sum_{i=1}^{N} \frac{d\Phi i}{dt} \tag{7}$$

where V is source voltage, I is solenoid current, R is coil resistance, N is number of turns, Φ_i is flux linkage of coil turn i, and λ_i is the total flux linkage of the winding.

$$\Phi_i = \iint_{S_i} B \cdot ds \tag{8}$$

where S_i is area enclosed by coil turn i, and B is flux density in the area, S_i.

Let the flux linkage λ_i be $\lambda = L(x)i$, then the external equation becomes

$$V = iR + \frac{L(x)di}{dt} + \frac{idL(x)}{dx} \cdot \frac{dx}{dt} \tag{9}$$

The counter emf induced in the coil is

$$e = \frac{d}{dt}(Li) = L\frac{di}{dt} + i\frac{dL}{dx}\frac{dx}{dt}, \quad e = L'\frac{x}{a+x}\frac{di}{dt} + L'\frac{ai}{(a+x)^2}\frac{dx}{dt} \tag{10}$$

the term Ldi/dt is the self-inductance voltage term, and $idL(x)/dx \cdot dx/dt$ is the velocity voltage term, which is used for energy transfer between the electrical and the magnetic subsystems. Hence, by combining the equations stated previously, the voltage equation for the electrical system is

$$v_t = iR + e = iR + L'\frac{x}{a+x}\frac{di}{dt} + L'\frac{ai}{(a+x)^2}\frac{dx}{dt} \tag{11}$$

The mechanical subsystem is represented by the spring K, damping D, and mass M, which relate force to displacement, velocity, and acceleration as follows:

$$f_K = K(l_1 - x), \quad f_M = M \cdot \frac{d^2x}{dt^2} \text{ and } f_D = D \cdot \frac{dx}{dt} \tag{12}$$

where x is the value of l_1 with the spring unscratched and the applied mechanical force being $f_0 = 0$. The force-balance equation for the mechanical portion of the system is

$$f_{fld} = f_K + f_D + f_M + f_0, \text{ hence } f_{fld} = K(l_1 - x) + D \cdot \frac{dx}{dt} + M \cdot \frac{d^2x}{dt^2} + f_0 \tag{13}$$

By substituting the magnetic force in the differential equation of motion for the mechanical system, the following expression is retained:

$$-f_t - \frac{1}{2}L'\frac{ai^2}{(a+x)^2} = M\frac{d^2x}{dt^2} + D\frac{dx}{dt} + K(l_1 - x) + f_0 \tag{14}$$

The Eqs (11) and (14) are valid only so long as the upper end of the plunger is well within the upper guide ring, (e.g. $0.1a < x < 0.9a$) [17]. This is the standard working range of the solenoid.

In this paper, a cylindrical solenoid was described. The objective of these mathematical description of solenoid actuator was to derive the dynamic equations of motion of this type of electromechanical system in order to understand and apply the critical parameters used for EMSS solenoid model. The investigation of the solenoid response time was built on the knowledge of dynamic equations of motion. Therefore, the mathematical formulation was important to determine specific parameters, such as spring, damping and mass, which relate to magnetic force.

3. Computational modelling of fuel injector solenoids

Computational modelling and simulation of fast acting fuel injector solenoids involved the design of an actual physical model system and an analysis for the purpose of either understanding the behavior of the system or evaluating various scenarios (within the limits imposed by a criterion or set of criteria). It was necessary to identify the relationships between the design parameters of the solenoid device and their effects on solenoid's performance. This was achieved by deriving a set of mathematical equations relating the structural, functional and performance parameters of the solenoid design. There were several factors which affected the switching time of injector solenoid, such as: effect of weight, attraction force and plunger cross sectional area, applied voltage, stroke and magnetic path length, number of coil turns, and other factors [19].

The main objective of this research was to improve the operational performance of fuel injector solenoid. The operational performance depends on three major factors. The first factor focused on reducing the size (miniaturisation of fuel injector solenoid in relation to maximising the attraction force) of GM fuel injector solenoid in relation to maximising the attraction force. This objective was achieved by conducting several modelling and simulation procedures in order to develop an evolution process of miniaturising a GM fuel injector solenoid. The second factor focused on reducing the response time of GM fuel injector by developing a number of modelling and simulation procedures. The third factor focused on the experimental method and investigation which included the following procedures: to determine the pulse and delay width of all the possible train pulses generated from user interface; and to estimate the output response time of fuel injector at various delay and duty-cycles by varying the supply and driver voltage.

The development of an evolution process of miniaturising a GM fuel injector solenoid consisted of four procedures or four specific objectives. The four procedures, consisted of modelling and simulation of several fuel injector miniaturised solenoid models, characterised with a particular: plunger pole shape (moving component) and plugnut (stationary component) (2D model); plunger length (2D model); yoke (2D model) and introducing the permanent magnet in a 3D modelling environment; adopted the permanent magnet principles to eventually develop a miniaturised fuel injector solenoid model in a 2D environment and determine the maximum attraction force possible.

Considering the second factor of the general objective, the operational performance of GM fuel injector solenoid was improved by reducing the response time for a range of fuel injector models. The first part of this objective was achieved by developing three modelling and simulation procedures in order to reduce the response time. These procedures were undertaken by developing electro-mechanical models that included features such as stroke and mass of the plunger; spring characteristics, motion and rebound of the plunger.

371
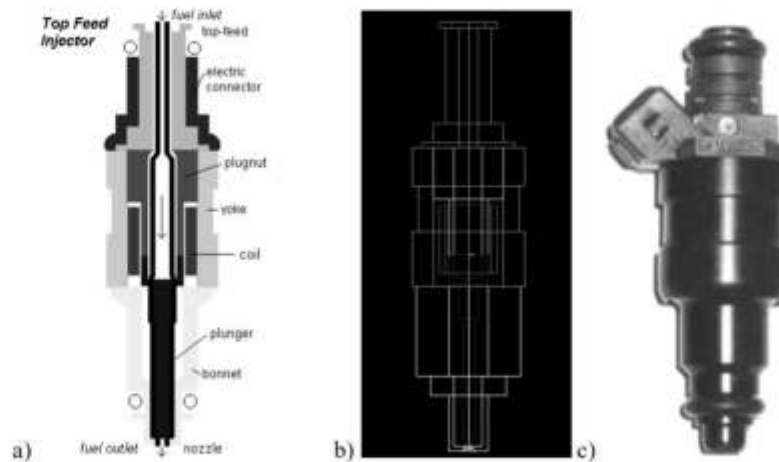
Fig. 4. 2D Cylindrical cross-sectional (a), (b) 3D and (c) physical fuel injector solenoid model.

The original GM fuel injector solenoid actuator has been accurately measured in order to determine the essential component dimensions, such as length and width of plunger, coil, yoke and etc. From that point, the collected data was used to model a virtual 2D and 3D fuel injector solenoid model. The dimensions of the virtual 2D and 3D model were identical to a physical fuel injector, as shown in Fig. 4. Also, the input parameters, such as the weight of the plunger and the number of coil turns had to be determined. The purpose of modelling 2D and 3D fuel injector solenoid actuators was to build a base model that will be passed through a miniaturisation process.

3.1. Miniaturisation procedures: Modelling and simulation of various plunger pole shapes

The main objective of plunger pole shape procedure was to examine the influence of various plunger pole shapes in relation of the attraction force characteristics. In order to improve the thrust or attraction force of solenoid actuators, there were several features that need to be evaluated. From the theoretical background, the reluctance needs to be decreased and the flux density on the plunger increased. However, from the modelling and simulation point of view, it was essential to simulate the 2D Maxwell axisymmetric plunger pole shape solenoid model in order to determine the virtual attraction force [6,14].

Solenoids can provide different force versus displacement profiles by varying the angle or the shapes of the plunger pole or other surfaces between the two members, such as plugnut [20,21]. In general, a flat surface provides high forces at small gaps, while a sharper angle can provide higher forces at large gaps, as shown in Fig. 5.

3.2. Miniaturisation procedures: Modelling and simulation of various plunger-length models

The development of plunger length solenoid models was similar in principle to the plunger pole shape solenoid models. In order to determine the force-stroke characteristics for various plunger lengths, a particular method was used to determine the highest attraction force at a specific stroke for a set plunger
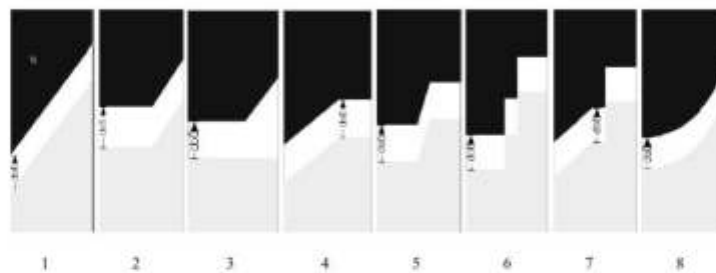
Fig. 5. Different plunger pole shapes of the solenoid actuator. In order to examine a more uniform force/air gap relation, eight pole and plunger contact areas were geometrically modelled using the standard 2D Maxwell axisymmetric solenoid. The geometry has systematically been reduced in relation to GM fuel injector solenoid as required. The affects of internal and external changes in the magnetic circuit, affected force characteristics.
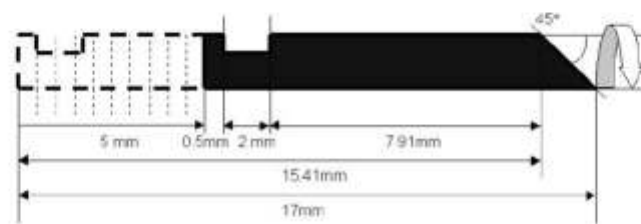
Fig. 6. Plunger length miniaturisation method for solenoid actuator model. From the theoretical background (Section 2), it can be seen that the strength of the magnetic field is different at each position inside the coil [1–3]. The most effective actuation or the strongest attraction force acting on the plunger would most probably occur in the center of the coil. The modelling and simulation of the solenoid actuators was therefore conducted in order to determine the strongest attraction force possible.

length, as shown in Fig. 6. As in the case of the pole shape model, the following objectives were targeted: miniaturisation of the plunger, and determining the highest attraction force possible in the solenoid actuator.

3.3. Miniaturisation procedures: Modelling and simulation of 3D solenoid actuator model

The 3D solenoid actuator model was developed using a similar principle as the one used to develop the plunger pole shape and plunger length models. In the modelling and simulation of the 3D solenoid actuator, three models were used to miniaturise the size certain aspects of solenoid components, such as yoke. The three models included the three main features with the following objectives: to introduce as few components as possible in the solenoid actuator model; to miniaturise the yoke component and to assign a permanent magnet to the solenoid actuator. The initial 3D model was developed to create simple geometry and introduce as few components as possible in the solenoid actuator system (Fig. 7).

Considering that there was a gap created when the top and bottom sleeve was excluded from the solenoid actuator system, it was possible to insert a permanent magnet (Neo35) in that space to replace the top sleeve. The purpose of inserting permanent magnet into the solenoid actuator was to determine if the attraction force would improve. The size of the solenoid actuator was reduced when the top and
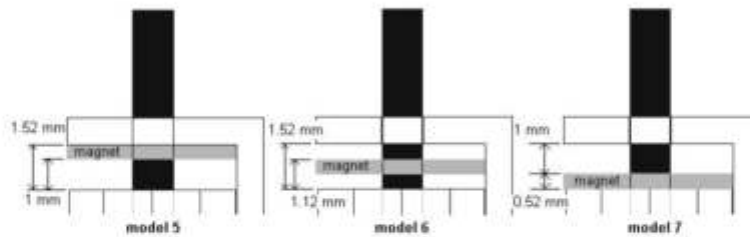
Fig. 7. 3D permanent magnet solenoid models. This model did not have top and bottom sleeves (plugnut). Also the coil and plunger were shifted in (-z) direction by 1 mm (by the thickness of the bottom sleeve). The next model that was investigated had the miniaturised yoke length by 8.4%. Due to miniaturisation of yoke length by 8.4%, the attraction force of this model was decreased to 1.65 N.
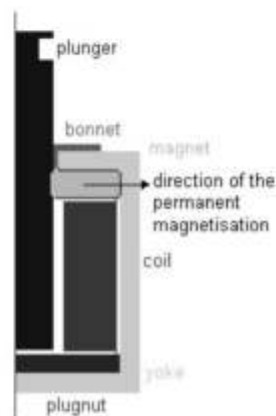
Fig. 8. Direction of the permanent magnet inside the miniaturised 2D solenoid model. The permanent magnet, positioned between the coil, yoke and plunger, was identical to 3D permanent magnet solenoid model. The type of material was assigned as Neo35. The overall stroke of the solenoid actuator was 3 mm, with the air gap of 0.165 mm. In order to define the vector properties such as magnetisation and polarisation, the vector's angle of magnetisation from the global x-axis was set to 0° (pointing to the right hand side).

bottom sleeves were excluded and the attraction force was increased to 2.15 N at 0 mm stroke. The permanent magnet solenoids (also known as magnetic latching or self-holding solenoids) utilised the attraction force advantage of a high performance permanent magnet [22].

When the coil is energised, the plunger moves toward the pole piece. Once energised, and plunger movement has occurred, the plunger will remain in the energised (holding) position [16]. The coils magnetic field and the permanent magnetic field are additive. To release the solenoid from the 'hold' position, the coil field has to cancel the permanent magnet field or reverse its direction, and the current has to flow in the opposite direction. Figure 8 shows miniaturised permanent magnet component that was displaced at various positions between the yoke and coil.

The magnet was shifted in $(-z)$ direction by 0.52 mm, (or by the thickness of the magnet). Considering the miniaturisation of the solenoid was established, it was necessary to determine the maximum attraction

force developed by the three models. The computational results indicated the highest attraction force of 2.49 N at stationary plunger position. Therefore, the location of the magnet component definitely affected the attraction force factor of the solenoid actuator.

3.4. Miniaturisation procedures: Modelling and simulation of 2D solenoid actuator with permanent magnet

The miniaturised 3D solenoid actuator model was constructed with the permanent magnet in order to gain the maximum attraction force possible. The further miniaturisation of permanent magnets solenoid actuators was conducted in the 2D Maxwell environment. The attraction force was investigated in relation to 2D solenoid model size (Fig. 8). The plunger type used in this model was of the flat pole shape, as shown in Fig. 5. When developing a miniaturised permanent magnet model, the feature of magnet orientation was of great importance.

Four different direction vectors have been investigated, including $0°$, $90°$, $180°$ and $270°$. This section has two major objectives: to determine the highest attraction force that can be developed by the four vector directions; and to use this particular direction to model four additional models with different permanent magnet properties assigned to the geometrical magnet object.

3.5. Reduction of response time: Standard EMSS fuel injector solenoid

The behavioral modelling of a linear electromachanical fuel injector solenoid actuation system consists of subsystems and components belonging to different physical domains, such as magnetic, electrical and mechanical subsystems. The electrical subsystem or the driver circuit consisted of DC voltage pulse 12 V, coil resistor element 14.5 Ω, number of coil turns 315 and a coil current of 827 mA. The mechanical subsystem consisted of loads on mechanical device, such as externally applied force -0.01813 N, gravitational force acting on a mass and the mass of the plunger. The force source value was determined from the plunger mass 1.85 g times the acceleration due to gravity 9.8 kg-m/sec^2 or -0.01813 N. The Standard EMSS Model was investigated on the plunger weight basis. Mechanical subsystem was set at four different weights, 1.55, 1.70, 1.85, and 2.00 g. The magnetic subsystem (fuel injector solenoid model) was connected to both of these subsystems via four ports. The goals of the system simulation were to calculate the minimum response time for the solenoid to close after the coil was energised.

3.6. Reduction of response time: Spring characteristic EMSS fuel injector solenoid

The main objective of the spring model was to creating a more realistic virtual model. Ideal spring parameters exhibit negligible mass and mechanical losses. A linear ideal spring was included to produce a force directly proportional to the relative displacement of the two ends of the spring, as shown in Fig. 9. Since the force at one end of the spring must be equal and opposite to the force at the other end, the force was transmitted through the spring in a manner similar to the way current passeed through an inductor.

The principle of the switch operation was relatively simple. Energising the solenoid coil generated a magnetic force, which pulled the plunger in a closed position. The long spring's relaxed length equaled the length of the plunger stroke of 0.30 mm. The short spring (at rest and out of contact when the plunger was in the fully open position) prevented the plunger from "hammering" down against the plugnut (plunger stop). An electrical switch in series with the short spring models caused this effect by closing when the plunger reached the short spring position. When the plunger contacted the short spring, it was near the minimum gap position, and the switch closed to activate the short spring force. Switching
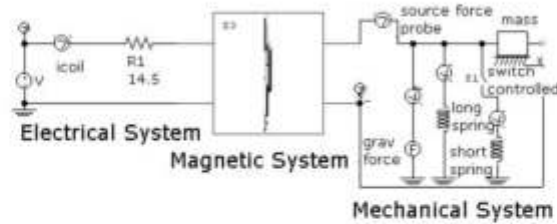
Fig. 9. Spring model of the solenoid fuel injector. The electromachanical fuel injector solenoid system consisted of magnetic, electrical and mechanical subsystems. The mechanical model consisted of an externally applied force, gravitational force acting on a mass, and two springs with different relaxed lengths. The long spring's relaxed length equals the length of the plunger stroke of 0.3 mm.

occured automatically as a result of setting the positive and negative control terminals of the switch to monitor plunger position. De-energising the coil allowed the long spring to push the plunger back open. In this part of spring EMSS fuel injector solenoid modelling, the plunger position moved from 0.30 mm to 0 mm. The plunger contacted the short spring at a gap of 0.08 mm. From the initial plunger position of 0.30 mm to a position of 0.08 mm, the short spring was out of the circuit. At that position only long spring operated. As a load, the long spring with the setting force of -1.71 N and a spring constant of 7.80 N/m was used. The short spring used load of 6.53 N, and spring constant of 81.60 N/m. The position threshold was at 0.08 mm.

3.7. Reduction of response time: Plunger's rebound decay model

The plunger's rebound decay model was developed due to constant plunger opening and closing movement in the nozzle of the injector. There were several factors that influenced the motion of the plunger, such as the effects of liquid-induced forces, spring preload forces, and the most important rebound forces. The rebound force represented both the elastic and inelastic collision forces between the plunger and plugnut. These particular factors were included in order to increase the realism of the developed virtual approach for reducing the response time of the fuel injector solenoids. It was also very important to stabilise the plunger motion at its closing time. Considering that the plunger was an integral part of the injection nozzle, its smooth and rapid operation played an important part in the overall design of the fuel injector [23,24]. The smooth and rapid operation 'bounce effects' of plunger was investigated using Maxwell Spice tool as shown in Fig. 10.

This circuit operated on the principle that when the stop material compresses, there is a spring-like restoring force proportional to the amount of compression. Also, there was a linear damping coefficient representing energy lost in compressing the material. The starting position or the stroke was 0.30 mm. The circuit sensed when the mass traveled beyond 0 mm, then developed a signal proportional to the amount of negative displacement. The Voltage Controlled Current Source (VCCS) $G1$ in Fig. 10 used position as a control voltage, with a gain of -1. The current represented mass position was forced through a nonlinear resistor to produce the control voltage. The VCCS $G2$ injected the spring force into the mass. Its gain of 900 represented the stop's spring constant K, N/m. The voltage-controlled switch simulated rebound damping. Without this damping, the position would show multiple bounces. After the system reached steady state, the source and spring forces balanced.

The investigation of such models has been used as a validation of general procedures and techniques developed. In the case of EMSS modelling, the manipulation of physical parameters of the fuel injector
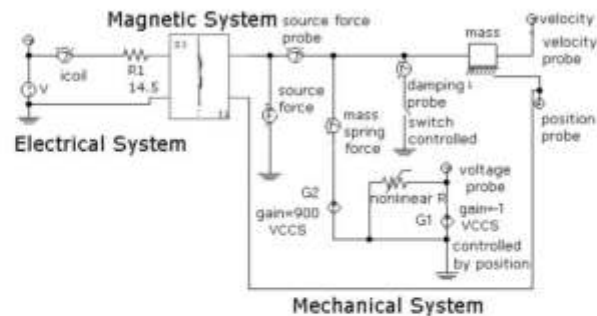
Fig. 10. EMSS rebound delay circuit model. In most of the mechanical system applications, the mechanical stops that are included in mass and inertia models were represented by very high damping values (damping coefficients). These were switched on when the mass position exceeded one of the position limits. Most of the time, they behaved as ideal stops. However, a real stop could cause rebound or have different cushioning behavior, called 'bounce effects'.

solenoid (such as mass and the starting position of the moving plunger, minimum and maximum positions, and damping coefficients) has been included in order to improve the fast acting fuel injector solenoid.

3.8. Reduction of response time: Experiment

At the experimental stage of the project, an investigation needed to be conducted utilising hardware and software control to monitor and test a GM fuel injector performance at various parameters. Combining the National Instruments® (NI) LabView® software package and data acquisition (DAQ) technique, a customised control system has been designed and developed. Data acquisition was implemented with NI 'low-cost' digital I/O 24 channel, 5 V/TTL DAQ card (National Instruments Corporation, USA), connected to CB-50LP (50-pin) I/O connector block via the PSH27-50F-D1 cable. From the connector block, digital pulse of 5 V/TTL was applied to an Electronic Driver Circuit (EDC), where the 5 V/TTL signal was initially amplified to 12 VDC to drive the fuel injector. The control and generation of the ON/OFF pulse length, used to drive the fuel injector actuation, was implemented via LabView software. A graphical user interface (GUI) in LabView environment was designed with the delay and duty cycle operations. The delay function determined the length of the OFF pulse in msec, or the time the fuel injector should remain in switched OFF position. Whereas, the duty cycle function determined the length of the complete pulse period, or ON/OFF pulse length in msec. The manipulation of both of these functions, the duty cycle or the pulse width modulation (PWM), resulted in the generation of the various waveform patterns in driving the fuel injector. The manipulation of the delay and duty cycle operations was implemented in from the digital pulse generation arithmetic in order to generate a particular waveform.

In the field of electronically controlled fuel injection systems, it is imperative that electromagnetic solenoid be capable of high-speed operation and have consistently reproducible stroke characteristics [23]. The necessity of high speed operation requires little explanation when one considers that an engine operating at 2000 rpm must have fuel injection into each cylinder of a multicylinder engine at 10 msec intervals while the entire injection pulse occurs over only a 1 msec period. The solenoid fuel injectors that operate at lower speeds or longer response period, result in enormous quantities of fuel being delivered to each cylinder at an inappropriate time advance, which can affect the overall performance

of the engine. An electronic control driver unit was developed to translate 5 V/TTL pulses, generated from the LabVIEW GUI, into 12 VDC pulses to drive the injector solenoid. The pulse generated from the software driver was initially, sent to electronic circuit, which consisted of dual monolithic power MAX628 MOSFET driver chip, acting as a translator from TTL input to high voltage/current output. The MOSFET's high speed minimises power losses in switching power drivers. The fast rise and fall times of the driver chip is typically 20 nsec with 1000 pF load. The output signal from MAX628 was connected to the ground of the power MOS transistor, which was intended for use in fast switching of the fuel injectors. The power MOS transistor was connected to one side of the injector solenoid coil and a silicon rectifier diode was connected in parallel with the solenoid coil in order to provide satisfactory coil suppression [23]. When designing an electronic circuit driver for high-speed pulse trains, it is important to consider the type of coil suppression used. The main reason for the use of such diode across the coil is to de-energise the circuit, by suppressing the instantaneous surge voltage that develops when the circuit is opened [16]. In other words, diode is used to protect the transistor against the backlash. A diode across the coil may provide satisfactory coil suppression, but it causes a slower collapse of the magnetic field, resulting in increasing the delay time interval. The EDC was powered from the Topward (Dual-tracking), DC power source. The monitoring and measurement of the output signal from the fuel injector was observed on the Tektronix, 2432A, 250 ms/s, digital oscilloscope. The output signal displayed on the oscilloscope is used to monitor and observe the behaviour of the fuel injector response time. The captured measurements were then acquired into Matlab® software tool (Mathworks, USA) via the PC/GPIB+ card for accurate measure of the output signal amplitude and pulse width. Matlab script was written to load the exported data to be plotted and tabulated.

4. Results

4.1. Miniaturisation of virtual fuel injector solenoid

The first section of this paper asseses modelling and simulation of various 2D and 3D models, which have undertaken the process of miniaturisation. Throughout the miniaturisation process, each step has been evaluated separately, so that every feature of fuel injector solenoid was miniaturised as much as possible until the main condition was satisfied. The second section assesses the response time of the original (GM) fuel injector solenoid. The virtual and experimental results were compared to determine the minimum response time operation possible.

The initial process of miniaturisation of the fuel injector solenoids has been undertaken by modelling the essential surface area features of the plunger pole shape design. Since the pole face area and flux density were related to the attraction force, the MMF was also related to magnetic flux density. The attraction force was proportional to the square of the magnetic flux density in the air gap. Therefore, it could be determined that the attraction force was dependent on the square of MMF, on the effective cross-section of the plunger, and on the inverse square of the air gap. The initial computation involved the analysis of the flux line plots of plunger pole shapes. These computational plots were very important for establishing the shortest magnetic flux path possible across the air gap. 8-plunger pole shape solenoid models were analysed. From the resultant plots of linear flux paths across the air gap for shapes 4, 5, and 7, it was possible to predict that these shapes will have the highest attraction force at 0 mm stroke. The strength of the flux density was very important across the air gap, in order to achieve the highest attraction force, and these three plots have shown that characteristic. The next part was to generate the parametric solution, sweeping the plunger displacement from 0 to 5 mm. Considering that standard 2D
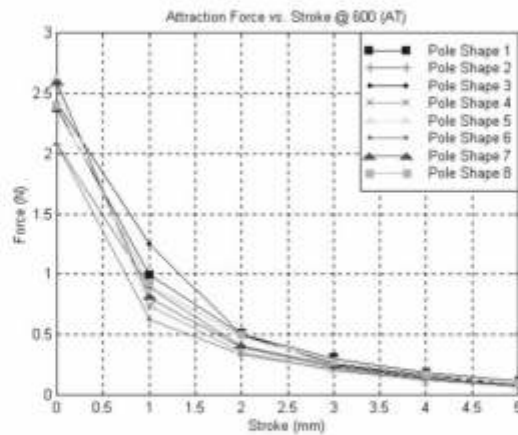
Fig. 11. Parametric force vs. stroke computations at 600 and 480 AT. From all these plots, it can be concluded that at the maximum MMF, and stationary plunger position, shapes (4), (5), and (7) show the strongest attraction force. As for middle range plunger displacements, shape (3), is the most effective. Considering the surface area of all the pole shapes, shape (1) was the highest, and as a result, it was most suitable for longer plunger displacements. However, shape (3), has the lowest surface area, and it's best operation in terms of attraction force was at middle range plunger displacement.

solenoid model, analysed in this case, operated at 552 AT, the closest force results could be found by correlating 600 AT plot, as shown in Fig. 11.

The design of plunger length solenoid was the next step in the miniaturisation process. The method was established to determine the highest attraction force at a particular stroke for a particular plunger length. Considering that the maximum magnetic field could be found in the center of the coil, the plunger would exhibit the maximum attraction force as a result. The air-gap is positioned at the middle of the coil, where the maximum force can occur. There were (11) different solenoid models plotted at 600 AT for force versus stroke relationship. At the stationary plunger position, models (5), (6), and (7) developed the maximum force, just over 2.77 N. Plunger lengths in this case were 19, 19.5 and 20 mm. As plunger displacement was incremented to 1 mm, models (8) and (9) developed the maximum force of 2 N. At 2 mm stroke, the maximum force was developed by models (10) and (11), 21.5 and 22 mm plunger length. In comparison with the plunger pole shape approach, the attraction force was increased. This increase in force was due to an increase in plunger length. From this approach, it has been determined that the attraction force was not necessarily increased as the plunger length was increased. The plunger length needed to be ideal to achieve the maximum attraction force. Since the optimum solution to plunger length has been developed, the miniaturisation has not been satisfied. Therefore, it was necessary to come up with the new miniaturisation 3D solenoid model approach to achieve maximum attraction force.

The 3D solenoid actuator has been developed as a complete new design. Some of the design differences are the rectangular shape of the plunger, compared to cylindrical shape used in 2D solenoid models. By incorporating three important approaches, seven models have been developed. For the modelling and simulation of the miniaturised 3D solenoid, it is important to introduce as few components as possible, miniaturise the yoke component, and assign a permanent magnet to the solenoid actuator. Models (5), (6), and (7) had a permanent magnet of smaller thickness, and the magnet component was gradually

displaced at three different positions between the yoke and top of the coil. Model (6) had the highest force at 0 mm stroke, 2.50 N. This proved that inclusion of permanent magnets in solenoids increased the attraction force. Overall, the fuel injector solenoid has been miniaturised by now, and the force level has reached the approximate constant level in comparison to 2D solenoid models. Now, the permanent magnet concept was adopted back into 2D modelling and simulation of the final miniaturised solenoid device.

The fuel injector solenoid was miniaturised even further in this way, and finally this miniaturised model was capable of developing the highest attraction force possible. The plunger type used in this model was of the flat pole shape, with the permanent magnet stationed in between the coil, yoke, and plunger, similar to the 3D permanent magnet solenoid actuator. Since the MMF of the system was 134 AT and the overall stroke of the solenoid actuator was 3 mm, with the air gap of 0.17 mm, it was obvious that every parameter has been minimised to a certain extent. The dynamic magnetic flux distribution of the simulated permanent magnet solenoid was the best indication of suitability and quality of solenoid design. The strongest attraction force was for $0°$ direction. At $0°$, 200 AT, the maximum force was -3.04 N. Considering that $0°$ was the preferred permanent magnet direction, the following investigation involved four additional magnet properties, such as NdFe30, NdFe35, SmCo24, and SmCo28. It has been determined that model with Neo35 permanent magnet achieved the highest attraction force, but with the highest energy consumption. From Fig. 12 it has been found that Neo35 permanent magnet has the most suitable and highest attraction force, compared to other properties.

In this type of approach, the miniaturisation process was flexible, accurate, and the CPU time could be significantly reduced. What was also most important about this particular miniaturisation approach was that it could be adopted for other solenoid actuator applications. Since the miniaturisation method has been established, the dimensions of essential solenoid components, (such as coil, yoke, and plunger lengths and widths) have been compared and illustrated in Fig. 13(a). From the initial plunger pole shape to permanent magnet model, the yoke length was miniaturised from 24 mm to 10mm; coil length from 12 mm to 6 mm; plunger width from 3.5 mm to 1.5 mm; and plunger length from 32 mm to 14 mm. In terms of attraction force, the strongest overall force was computed for the plunger length model from 0.2 to 5 mm stroke, as shown in Fig. 13(b). However, the 2D permanent magnet model exhibited the strongest force at the initial stroke displacement from 0 to 0.2 mm. These results clearly show that reducing the size of solenoid components does in fact reduce the attraction force. However, with the case of adding the permanent magnet, this attraction force could be improved for the shorter stroke displacement which is sufficient enough for the actual fuel injection operation.

4.2. Reduction of virtual fuel injector solenoid response time

The initial step in finding the fast response solution involved the effects of the plunger weight of solenoid actuator. The weights ranged from 1.55 to 2.00 g. At final stroke displacement, 0.30 mm, the smaller weights of the moving plunger, 1.55 g, gave the longest response time. However, for shorter strokes, from 0.20 to 0 mm, the response time was shorter. In case of a plunger of 2.00 g, at 0.30 mm stroke, the response time was shortest. But at shorter strokes, from 0.20 to 0 mm, the response time was longest. The weight of the moving plunger had a significant effect on the response time of the solenoid.

In the spring characteristic model, the plunger position moves from 0.30 mm to 0 mm. The plunger contacts the short spring at a gap of 0.08 mm. From the initial plunger position of 0.30 mm to a position of 0.08 mm, the short spring was out of the circuit. At that position only long spring operated. At the initial position, the response time was 1.19379 msec and at 0.25 mm, the response was 0.717 msec. In
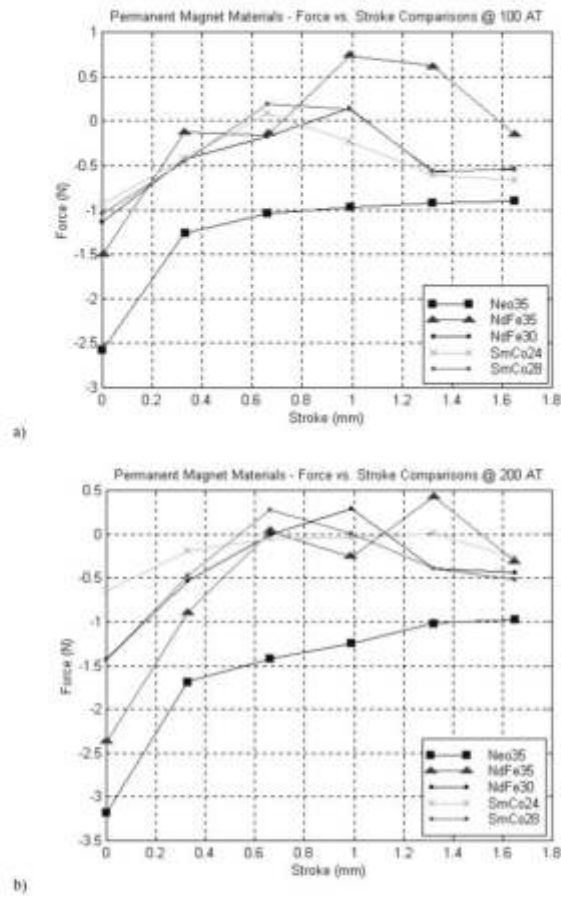
Fig. 12. Force vs. stroke characteristic for various permanent magnet properties. (a) At the stationary plunger position, 100 AT, Neo35 property developed the force of −2.6 N, and at longer strokes, it reached saturated level, around −1 N until its operation is completed. The other properties exhibited irregularities in attraction force over the complete plunger displacement. (b) At 200 AT, the maximum attraction force was 3.2 N. The dynamic characteristics (magnetic flux density) of the permanent magnet solenoids from 0 to 3 mm stroke were analysed to achieving the maximum attraction force.

comparison to standard EMSS models developed, the response time was minimised by 0.05 msec, as shown in Figs 14 and 15. Figure 14 illustrates the response time minimisation process by comparing the three principle models developed in this research paper.

In the rebound delay model, a 1.67 N step force was applied to a plunger mass of 1.85 g. Considering the operation principle of the rebound delay circuit, there was a spring-like restoring force proportional to the amount of compression. Due to this operation, there was a linear damping coefficient representing energy lost in compressing the material. Similar to the spring model, the starting position or the stroke

a)

b)
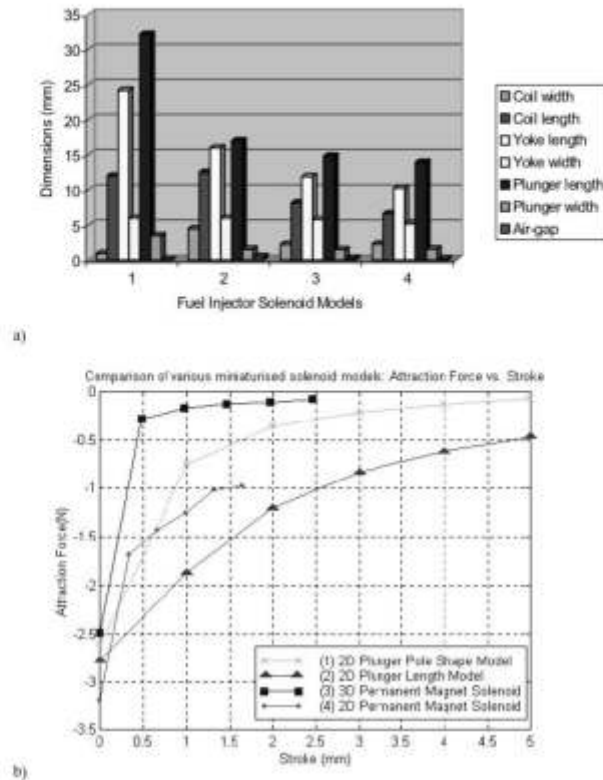
Fig. 13. Miniaturisation process (a) and force vs. stroke comparison (b) for various 2D and 3D solenoid models. They include miniaturised dimensions of the coil, yoke and plunger length. As a result of this miniaturisation, the computation of the respective attraction force for these models has also been improved. The development of the 2D permanent magnet solenoid led to a new miniaturised fuel injector solenoid with the maximum attraction force.

was 0.30 mm. This type of control operated on the feedback principle, and in terms of response time, the results seemed to be effective. Initially, the response time was 1.12052 msec, and at 0.25 mm, the response time was 0.404 msec, as shown in Figs 14 and 15. Overall, these results proved the effectiveness of the 'bounce effect' model in terms of rebound delay, which resulted in shorter response time.

The rebound EMSS model was more complex than the previous standard and spring models. It gave the minimum response time. Comparison of the response times achieved using the three models is displayed in Fig. 15.

4.3. Experimental results: Reduction of fuel injector solenoid response time

The response time of the injector solenoid was greatly influenced by the supply voltage. Measurements of the response times were conducted by sweeping the supply voltages (5–20 Vdc). The dual power

382
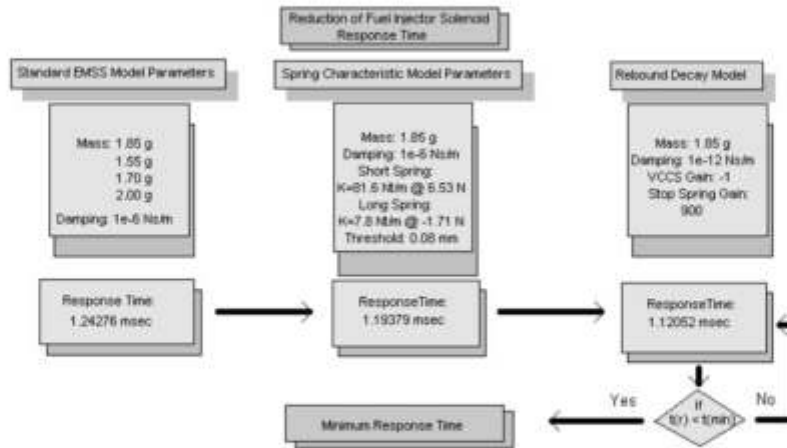
Fig. 14. Process of reducing the response time of the fuel injector solenoid.
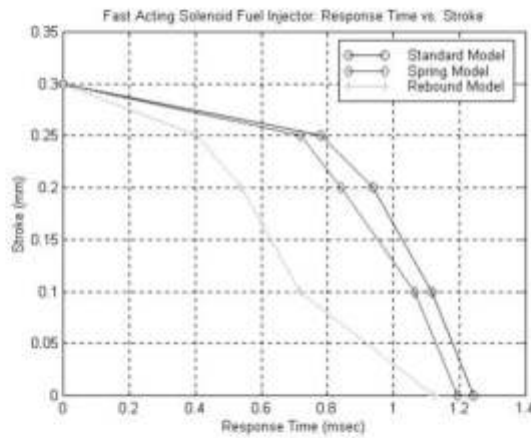
Fig. 15. Comparison of the response times achieved using the three virtual models.

supply was initially set at 5 Vdc, and the response time readings were investigated in detail for the three generated pulse levels, (delay-duty cycle configuration: 4-4, 0-3, 8-0). For example, at 12 Vdc, the delay (0) and duty-cycle (3) or (0-3), the duty cycle (%) was 13.97%. The reduction in response time was examined by finding the most optimum duty-cycle (%), driver and supply voltage used to drive the solenoid coil.

It can be seen from Fig. 16, that the lower the voltage, the shorter the response times will be. Also,
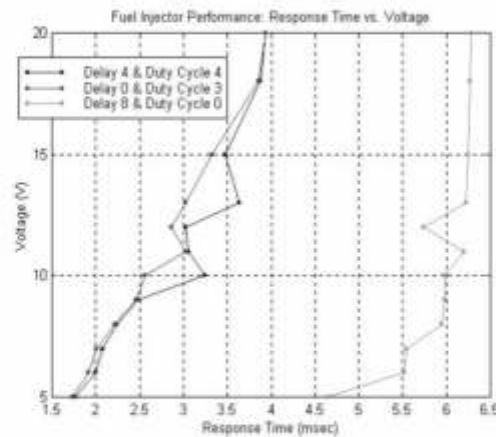
Fig. 16. Measured experimental response times at several supply and driver voltages for three pulse trains.

the higher the supply voltage, the longer the return speed of the plunger is required. When altering the supply voltage, several conditions needed to be considered. For example, a drop in the supply line voltage can prevent the solenoid from closing by reducing its force until it cannot overcome the load. The ambient temperature also needed to be checked. If the ambient temperature was too high, the coil would lose its ability to dissipate heat, resulting in the resistance increase, current flow decrease, and the solenoid will not close, resulting in coil over-heating. It is believed that for each degree Celsius above or below 20°C, the resistance of the coil's copper wire changes by 0.393%. The coil over-heating could also be caused by the high voltage supply which can generate an excessive holding current.

In our EDC design, a silicon rectifier diode was connected in parallel with the fuel injector's coil in order to provide satisfactory coil suppression, or to de-energise the injector coil. This type of coil suppression causes slower collapse of the magnetic field, increase in the delay time interval, and eventually, longer response time. Typically, the higher the supply voltage, the longer it will take for the diode to de-energise the fuel injector coil. It takes longer to collapse the magnetic field surrounding the fuel injector coil. However, the lower the supply voltage, the faster the de-energising of the coil can take place, resulting in faster response time. From Fig. 16, pulse trains (4-4) and (0-3) were identical, except for the differences in response between the 10 and 15 Vdc. The fastest recorded response time was 1.75 msec at 5 V. The slowest response time was approximately 4 msec, observed at 20 V. At lower driver voltage, and lower supply voltage, the response time was fastest around 2.1 msec. The slowest response time was observed at higher driver and supply voltages (around 3.9 msec).

5. Discussion

The proposed methodology was designed using a 'stepwise' approach towards modelling the solenoid, with each step serving to improve the accuracy and refine the model. As a first step, a qualitative mathematical relation model was designed describing the behaviour of solenoid actuator system in terms of the ordinal relationships followed by the mathematical algebraic equations representing the

basic functionality of magnetic, electric and mechanical circuits and deriving the dynamic equations of motion which could be utilised in the modelling design of electromechanical subsystem. The reasoning statements, based on the virtual manipulation of solenoid input parameters could be further improved using fuzzy logic techniques in any future research work. The refinement of structural, functional and performance parameters of the solenoid design was the second and third step of the approach. These parameters are part of the subsystems and components belonging to different physical domains, such as the electromagnetic and electromechanical. The electromagnetic subsystem approaches focused on the miniaturisation and improvement in attraction force of the solenoid actuator. The electromechanical subsystem was mainly applied in reducing the response time of actuation.

The initial process of miniaturisation of the fuel injector solenoids has been undertaken by modelling the essential surface area features of the plunger pole shape design. Our initial modelling and simulations revealed that solenoids can provide different force versus displacement profiles by varying the angle or the shapes of the plunger pole or other surfaces between the two members, such as plugnut. In general, a flat surface provides high forces at small gaps, while a sharper angle can provide higher forces at large gaps, as shown in Fig. 3. The design of plunger length solenoid was the next step in the miniaturisation process. The method was established to determine the highest attraction force at a particular stroke for a particular plunger length. The results showed that the increase in attraction force was due to an increase in plunger length. From this approach, it has been determined that the attraction force was not necessarily increased as the plunger length was increased. In order to improve our miniaturisation design, a 3D solenoid actuator was developed. The computational results proved that inclusion of permanent magnets in the solenoid increased the attraction force despite the optimisation of yoke geometry. At this stage the permanent magnet (Neo35) concept was adopted back into 2D 'final' miniaturised solenoid model which generated the maximum attraction force when compared to other permanent magnet properties.

Reduction of fuel injector response time has been achieved through virtual modelling and simulation. The virtual fuel injector solenoid modelling and simulation consisted of three EMSS models. The three EMSS models included the standard EMSS fuel injector solenoid model, spring characteristic EMSS model, and plunger rebound decay EMSS model (bounce effect). The reduction of response time was certainly achieved by the three EMSS models. Finally, the minimum virtual response time was 1.12 msec, developed by EMSS Rebound Delay Model. Throughout this research, a new approach for miniaturisation, force improvement and respose time reduction was flexible, accurate and allowed for simple adaptability to other solenoid actuator applications.

The experiment was conducted using the actual GM fuel injector, previously modelled and simulated using EMSS software. LabView generated duty cycle statements. The voltage generated was transformed in the electronic driver unit, EDC, to actuate the fuel injector solenoid at a range of specific regimes. The reduction of response time in the GM fuel injector solenoid was achieved by focusing on particular features such as the manipulation between the MOSFET driver and supply voltages, and delay and duty cycle levels. Initially, the minimum response time was measured at the constant supply voltage of 12 Vdc at the pulse train (3-7). The resultant response time was 2.98 msec. Since the minimum response time was required, MOSFET drive voltage was reduced to 5 V and the supply voltage remained constant 12 V. Generated pulses from the LabView user interface also had to be adjusted to pulse train (4-4). Due to these adjustments, the response time was slightly reduced to 2.95 msec. Next, the delay-duty cycle level was kept constant, and only the supply level was reduced to 6 V. As a result, the response time was reduced to 2.1 msec. During the experiment, it was obvious that the manipulation between the pulse rate and supply voltage had a lot to do with the reduction of response time. Therefore, once again the pulse rate has been left untouched, the supply was set to 5 V, and the driver was 10 V. The result of this

Table 2
Comparison of response times obtained from virtual models and experimental investigation

	Supply – Driver Voltage (Vdc)	Delay – Duty Cycle (index)	Response time (msec)
Experimental	12–10	3–7	2.98
	12–5	4–4	2.95
	6–5	4–4	2.10
	5–10	4–4	1.76
	5–10	0–3	1.73
Virtual	Standard model	–	1.24
	Spring model	–	1.19
	Rebound decay model	–	1.12

investigation was surprising. The response time was measured to be 1.76 msec. The process reached the end when the supply and driver voltages were set as the last condition and delay-duty cycle was set to (0-3). This condition reduced the response time to the minimum possible level of 1.73 msec (Table 2).

The experiment proved to have a significant effect on the reduction of the response time. However, the reduction of response time was certainly achieved by the three EMSS models, with Rebound delay model being 1.12 msec. Throughout this research, a number of modelling and simulation procedures have been repeatedly conducted in order to achieve a fast acting operation of the fuel injector solenoid. The simulation based on mathematical solenoid model served to support these experimental results. The modelling and simulation procedures have been successful in reducing the response time that somehow seemed difficult to be achieved by the experimental method, under given conditions. However, by repeating the design and experimental trials, the final minimum response time was achieved by the virtual Rebound delay model.

In order to further improve the existing methodology, the additional subsystem and domains would greatly enhance the performance of fuel injector solenoid actuator, such as hydraulic, thermal and digital. Also, further improvements in the temperature rise of the coil and magnetic hysteresis which can have a significant influence on the dynamic performance of solenoid actuators could be investigated in the future in the form of optimization techniques and possibly applied in future experiments. Due to a complexity and nature of multiple-factor dependency on the overall performance of fuel injector solenoids, it was decided that a stepwise approach was best suited.

6. Conclusion

Modelling and simulation of fast acting fuel injector solenoids was the main focus of this research which was later compared with the experimental investigation. As the main object of this research investigation was a petrol fuel injector solenoid, the main objective of this research was to improve the operational performance of fuel injector solenoid. The operational performance consisted of two major factors that were improved. Using the developed approach, the initial size of the fuel injector has been reduced by 35%, the attraction force increased by 26% and the response time reduced by 76%. The modelling and simulation procedures have been successful in reducing the response time. However, by frequently repeating the design trials and conducting a thorough experimental investigation, the final minimum response time was achieved by the virtual rebound delay model. The reduction in response time from the 'optimal' experimental to virtual model was by 35%. The simplicity and effectiveness of the developed methods, allowed for quick and accurate evaluation. The correlation of all the modelling simulation and experimental results obtained in this research investigation have greatly improved the operational

performance of the fuel injector solenoid. The utilisation of the operational performance processes, made up of individual procedures in a stepwise approach has been the essential tool in achieving the given general and specific objectives. Future virtual modelling research could be enhanced to reflect in the actual experimental operational performance of fuel injector solenoids. Likewise, experimental methods could be developed in the miniaturisation process of fuel injector solenoids with the close relationship to virtual modelling and simulation.

References

[1] Y. Xu and B. Jones, A Simple Means of Predicting the Dynamic Response of Electromagnetic Actuators, *Mechatronics* 7(7) (1997), 589–598.

[2] T. Kajima and Y. Kawamura, Development of a High-Speed Solenoid Valve: Investigation of Solenoids, *IEEE Transactions on Industrial Electronics* 42(1) (February 1995), 1–7.

[3] Y.B. Cheng and J.K. Sykulski, *CAD and Optimisation of Electromechanical Actuators*, United Kingdom: University of Southampton Press, 1991.

[4] D. Cvetkovic and A. Subic, Modelling and Simulation of a Miniaturised Actuating Solenoid Device, *IASTED International Conference 2000, Applied Simulation and Modelling*, Banff, Alberta, Canada, July 2000, 70–76.

[5] D. Cvetkovic and A. Subic, Miniaturisation of Electro-Mechanical Actuators for Fuel Injection Systems, *NAFEMS World Congress 2001, The Evolution of Product Simulation*, Lake Como, Italy, April 2001, 415–426.

[6] D. Cvetkovic, *Modelling and Simulation of Fast Acting Fuel Injector Solenoids*, RMIT University, Master of Engineering Thesis, Australia, 2002.

[7] G. Meunier, D. Shen and J.-L. Coulomb, Modelisation of 2D and Axisymmetric Magnetodynamic Domain by the Finite Elements Method, *IEEE Transactions on Magnetics* 24(1) (January 1988), 166–169.

[8] H. Yamashita, V. Cingoski, M. Mikami and K. Kaneda, Visual Computing Concept in Finite Element Analysis, *IEEE Transactions on Magnetics* 33(2) (March 1997), 967–973.

[9] P. Alotto, P. Girdinio and M. Nervi, Mesh Adaptation in Finite Element Analysis of 2D Steady State Time Harmonic Eddy Current Problems, *IEEE Transactions on Magnetics* 32(3) (May 1996), 1361–1364.

[10] P. Hammond and J.K. Sykulski, *Engineering Electromagnetism – Physical Processes and Computation*, UK: Oxford University Press, 1994.

[11] K. Takeuchi, M. Shimizu and K. Okazaki, Fast Actuator Modeling by Finite Element Method, *IEEE Transactions on Magnetics* 30(6) (November 1994), 4284–4286.

[12] A.W. Cox, H. Garmestani, D. Markiewicz and I.R. Dixon, Power Series Stress Analysis of Solenoid Magnets, *IEEE Transactions on Magnetics* 32(4) (July 1996), 3012–3015.

[13] Y. Mitsutake, Y. Ishihara and K. Hirata, Dynamic Response Analysis of a Linear Solenoid Actuator, *IEEE Transactions on Magnetics* 33(2) (March 1997), 1634–1637.

[14] T. Weiland, 2D and 3D Calculation of Forces, *IEEE Transactions on Magnetics* 30(5) (September 1994), 3467–3470.

[15] R. Rong and D.A. Lowther, Adapting Design Using Dimensional Models of Electromagnetic Devices, *IEEE Transactions on Magnetics* 32(3) (May 1996), 1437–1440.

[16] T. Kajima, Dynamic Model of the Plunger Type Solenoids at Deenergizing State, *IEEE Transactions on Magnetics* 31(3) (May 1995), 2315–2323.

[17] D.A. Bell, *Fundamentals of Electric Circuits*, Magnetism, A Prentice-Hall Company, Reston, Virginia, USA, 1978, 221–226.

[18] T. Kabashima, A. Kawahara and T. Goto, Force Calculation Using Magnetizing Currents, *IEEE Transactions on Magnetics* 24(1) (January 1988), 451–454.

[19] Y.B. Cheng, J.K. Sykulski and R.L. Stoll, Force Optimisation in DC Actuators, *COMPEL-The International Journal for Computation and Mathematics in Electrical and Electronic Engineering* 13(1) (1994), 163–166.

[20] S-B. Yoon, J. Hur, Y.-D. Chun and D.-S. Hyun, Shape Optimization of Solenoid Actuator Using the Finite Element Method and Numerical Optimization Technique, *IEEE Transactions on Magnetics* 33(5) (September 1997), 4140–4142.

[21] D. Dyck, D.A. Lowther, Z. Malik, R. Spence and J. Nelder, Response Surface Models of Electromagnetic Devices and their Application to Design, *IEEE Transactions on Magnetics* 35(3) (May 1999), 1821–1824.

[22] P. Campbell, *Permanent Materials and Their Applications*, UK, Cambridge University Press, 1994.

[23] K.-Y. Yuan and S.-C. Chen, New Algorithm for Coupled Solutions of Electric, Magnetic, and Mechanical Systems in Dynamic Simulation of Solenoid Actuators, *IEEE Transactions of Magnetics* 26(3) (May 1990), 1189–1197.

[24] M.F. Rahman, N.C. Cheung and K.W. Lim, Position Estimation in Solenoid Actuators, *IEEE Transactions on Industry Applications* 32(3) (May/June 1996), 552–559.

[39]   Ito, A., Kawano, M., Fujita, S., "ECU Structure Strategy to Detect Lift Timing of GDI Solenoid Injectors with High Precision," SAE Technical Paper 2017-01-1628, doi: 10:4217/2017-01-1628, 2017.

SAE INTERNATIONAL.

ECU Structure Strategy to Detect Lift Timing of GDI Solenoid Injectors with High Precision	2017-01-1628
	Published 03/28/2017

Atsushi Ito, Masahiro Kawano, and Shohei Fujita
DENSO Corporation

Abstract

In gasoline direct injection (GDI) systems, various injection types are needed to reduce emissions and improve fuel consumption. This requires high-precision injection in the region in which the amount of injection is small. Achieving injection of a small amount of fuel using GDI solenoid injectors requires the use of the half-lift region. In this region, however, the variation in the injection amount tends to increase due to the variation in the lift behavior of the injectors, posing the problem of how to achieve high-precision injection. To reduce the variation, we analyzed the lift timing out of the injector current and voltage signal with the ECU in an attempt to adjust the amount of injection. To detect the injector needle-opening timing, we focused on the characteristic point of the injector current signal appearing when the injector needle opens, and to detect the injector needle-closing timing, we focused on the characteristic point of the injector voltage signal appearing when the injector needle closes. Using the ECU, we analyzed the current and voltage signal, developing an algorithm allowing the characteristic points of both signals to be detected. Detection requires a wide dynamic range and high precision. The application specific integrated circuit (ASIC) can acquire and process the current and voltage signal with high precision and at high sampling speed. Accordingly, we adopted a functional allocation based on the use of ASIC. As a result, the ECU has potential to detect the timing of the open-close needle with high precision without increasing the processing load of the micro-controller significantly.

Introduction

Future emission regulations require the development of gasoline combustion engines that will achieve reduction in emissions and improvement in fuel consumption. For this reason, expectations are placed on gasoline direct injection (GDI) systems to provide compatibility between reduction in emissions and improvement in fuel consumption. To be compatible with both the requirements, various injection types such as multiple injections, injection with a small amount of fuel, and short interval are needed. In the region in which fuel is injected in a small amount, high-precision injection is required.

To achieve injection with a small amount of fuel using GDI solenoid injectors, it is necessary for GDI solenoid injectors to work in the half-lift region before the valve fully lifts up [1]. In the half-lift region, however, the variation in the amount of injection tends to increase due to the variation in the lift behavior of the injectors, posing the problem of how to achieve high-precision injection.

The variation in the lift behavior of the injectors correlates with the timing of the open-close injector needle. If the ECU can detect the timing of the open-close needle, the amount of injection can be adjusted. Focusing on the characteristic point of the injector current and voltage signal appearing in the timing of the open-close needle, we tried to develop a method of detecting the needle-opening and needle-closing timing. To detect the injector needle-opening timing, we focused on the characteristic point of the injector current signal appearing when the injector needle opens, and to detect the injector needle-closing timing, we focused on the characteristic point of the injector voltage signal appearing when the injector needle closes.

Micro-controllers and ASICs are electronic parts capable of detecting the characteristic point. The criteria in deciding whether the role of detecting the characteristic point is allocated to micro-controller or ASIC are the processing load of the micro-controller and the ASIC chip size. When micro-controller is used to achieve the detection, the processing load matters because sampling the current signal (or voltage signal) at high speed and analyzing the characteristic point continuously is required. With the control specification of ECU becoming increasingly complex recently, the processing load poses a serious problem. On the other hand, when ASIC is used to achieve the detection, it becomes possible to acquire the current signal (or voltage signal) with high precision and at high sampling speed for the purpose of analysis. As a result of adding circuits, however, the chip size increases.

With the performance and the processing load taken into consideration, it is important to decide the functional allocation between micro-controller and ASIC at an early stage. The reason for this is that after designing ASIC, the functional allocation is difficult to change. Even when a change is allowed, a substantial delay in the schedule will be needed. Therefore, the functional allocation should be decided before starting on ASIC design.

This paper reports the following:

- The ECU structure needed to detect the timing of the open-close injector needle
- The strategy for the functional allocation between micro-controller and ASIC
- The optimization method of the functional allocation based on the simulation of processing load

ECU Structure

As shown in Figure 1, the system consists of ECU and injector. The ECU is equipped with injector driving function and driving signal analyzing function, with the latter detecting the timing of the open-close injector needle.
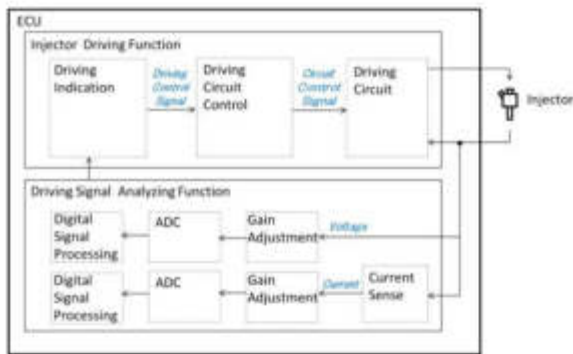
Figure 1. System

The injector needle-closing timing is detected by analyzing the characteristic point in the voltage signal occurring in the low-side driving voltage in the injector. A typical voltage signal detected is shown in Figure 2. The driving signal analyzing function for detecting the needle-closing timing consists of gain adjusting section, ADC (analog to digital converter) section, and digital signal processing section. With the voltage signal to be detected having a wide dynamic range between 0 and 80 V, the gain is adjusted in the gain adjusting section. The voltage after the gain adjustment is put to ADC in the ADC section at resolution of 10 bits and at sampling cycle of 1MHz. To detect the needle-closing timing with high precision, it is necessary to detect a minute change in the voltage of less than 50mV per 1μsec. The gain of the gain adjusting section was set so that this change in the voltage might be detected with the above-mentioned resolution of the ADC section. Digitized signals are analyzed in the digital signal processing section. The signal analyzing algorithm developed by us independently is implemented in the digital signal processing section, which allows us to analytically determine the detection timing highly correlated with the needle-closing timing.
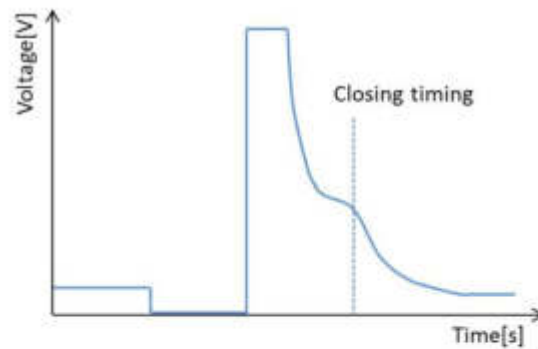
Figure 2. Injector low-side voltage signal

The injector needle-opening timing is detected by analyzing the characteristic point in the current signal occurring in the low-side driving current in the injector. A typical current signal detected is shown in Figure 3. The driving signal analyzing function for detecting the needle-opening timing consists of current sense section, gain adjusting section, ADC section, and digital signal processing section. With the current signal to be detected requiring a wide dynamic range between 0 and 20 A, the current is detected in the current detection section and the gain adjusting section by means of differential amplification. The current after the gain adjustment is put to ADC in the ADC section at resolution of 10 bits and at sampling cycle of 1MHz. To detect the needle-opening timing with high precision, it is necessary to detect a minute change in the current at less than 20 mA per 1 μsec. The gain of the gain adjusting section was set so that this change in the current might be detected with the above-mentioned resolution of the ADC section. Digitized signals are analyzed in the digital signal processing section. The signal analyzing algorithm developed by us independently is implemented in the digital signal processing section, which allows us to analytically determine the detection timing highly correlated with the needle-opening timing.

Figure 3. Injector current signal

The feasibility of detecting the needle-opening timing and the needle-closing timing by means of this ECU structure was demonstrated on an actual unit. The environment of the feasibility study on an actual unit is shown in Figure 4. The digital signal processing section was allocated to the test chip on the evaluation board, and the results of detecting the needle-opening and the needle-closing timing are stored in the RAM in the test chip. The results stored in the RAM in the test chip can be checked via the communications line on the evaluation board by means of a personal

389

computer. The actual timing of the needle-opening and the needle-closing is determined by measuring the injection rate. Injectors manufactured by DENSO CORPORATIION were used in the feasibility study. We choose Injector A and B which has variation of closing timing. We choose Injector C which has variation of opening timing. We studied that, by comparing the needle-opening and the needle-closing timing measured by means of the injection rate with the needle-opening and the needle-closing timing detected by the evaluation board, the needle-opening and the needle-closing timing can be detected with this ECU structure.
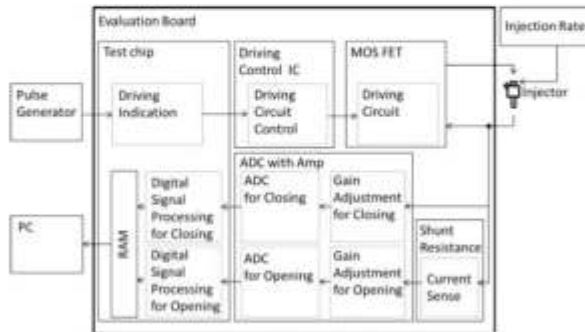
Figure 4. Experimental Environment Structure

The results of the detection of needle-closing timing are shown in Figure 5. The horizontal axis denotes the actual needle-closing timing measured by means of injection rate, with the vertical axis denoting the needle-closing timing detected by analyzing the voltage signal on the evaluation board. There is correlation between the actual closing timing and the closing timing detected on the evaluation board.
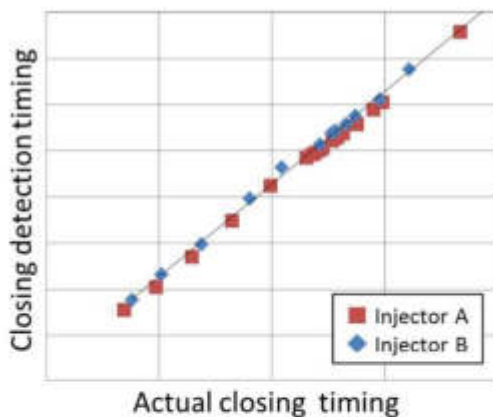
Figure 5. Correlation between closing detection timing and actual closing timing

The results of the detection of needle-opening timing are shown in Figure 6. The horizontal axis denotes the actual needle-opening timing measured by means of injection rate, with the vertical axis denoting the needle-opening timing detected by analyzing the current signal on the evaluation board. There is correlation between the actual opening timing and the opening timing detected on the evaluation board.
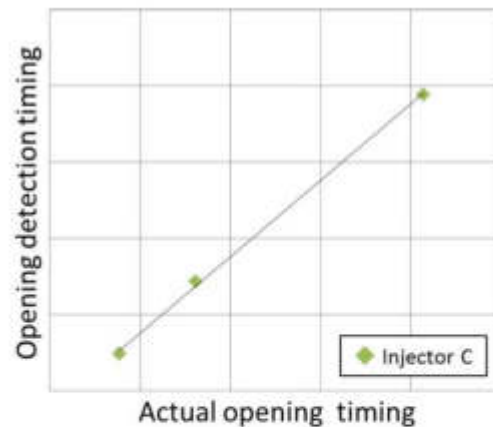
Figure 6. Correlation between opening detection timing and actual opening timing

Errors of 10µsec or less existed between the actual timing and the timing detected on the evaluation board. The above results suggest that it is highly possible to detect the needle-opening timing and the needle-closing timing with this ECU structure.

ECU Functional Allocation Strategy

The functions necessary to detect the needle-opening and needle-closing timing are provided by current sense section, gain adjusting section, ADC section, and digital signal processing section as shown in ECU structure. To achieve each functions with high precision by means of the ECU, it is necessary to allocate them between the micro-controller and ASIC.

Current Sense Section and Gain Adjusting Section

In driving signal analysis for needle-opening, current sense section and gain adjusting section are required a wide dynamic range between 0 and 20 A and to achieve the precision to detect a minute change in the current of less than 20 mA. The analog input of a general-purpose micro-controller is connected directly to ADC via the switch function and has no such functions as those described above. For this reason, it is necessary to form an amplifying circuit consisting of a general-purpose operational amplifier and resistors outside micro-controller and connect it to the analog input. With the acquisition of resistors with identical characteristics being difficult, however, it is difficult to achieve required precision. On the other hand, a similar amplifying circuit is built when the functions are allocated to ASIC; with resistors formed on the same chip, resistors with identical characteristics can be obtained and high precision can be achieved. Accordingly, we considered that it is the optimum measure to allocate current sense section and gain adjusting section to ASIC. In driving signal analysis for needle-closing, gain adjusting section is required a wide dynamic range between 0 and 80 V and to achieve the precision to detect a minute change in the voltage less than 50 mV. We considered that, as in the case of needle-opening, the allocation of the functions to ASIC is the optimum measure.

390

ADC Section

In the driving signal analysis for both needle-opening and needle-closing, ADC section is required to perform ADC at the resolution of 10 bits and at the sampling cycle of 1MHz. When the functions are allocated to micro-controller, ADC contained in micro-controller is used. On the assumption that this converter is used exclusively with a single input, a sufficient resolution and a sufficient sampling cycle can be obtained. However, exclusive use is difficult because a general-purpose micro-controller is structured to convert as many as several tens of analog inputs sequentially with several ADCs. When functions are allocated to ASIC, on the other hand, it is optimized ADC with a sufficient resolution and a sufficient sampling cycle and it is possible to be occupied with driving signal analysis. For these reasons, we concluded that allocating ADC section to ASIC is the optimum means.

Digital Signal Processing Section

The criteria in deciding whether digital signal processing section is allocated to micro-controller or ASIC are the processing load on the micro-controller and ASIC chip size. Allocating digital signal processing section to ASIC does not cause the processing load to be applied to micro-controller, but this risks ASIC chip size increasing. On the other hand, micro-controller excels in its calculating capability in digital signal processing section; however, an increase in the calculation to execute causes the concern over an increase in the processing load. It is necessary to work out a plan for functional allocation on the basis of an understanding of the features of micro-controller and ASIC. Allocating digital signal processing section to micro-controller (Case A, shown in Figure 7) is the most effective in restricting the chip size because the digital processing section does not accompany ASIC. Conversely, this requires the output from ADC section to be communicated from ASIC to micro-controller, resulting in the communication of a vast amount of data. This increases the processing load on micro-controller, causing a risk of exceeding the limit of the processing performance of micro-controller. In contrast to this, allocating digital signal processing section to ASIC (Case B, shown in Figure 8) results in a small amount of data to be handled since the data communicated from ASIC to micro-controller is limited to needle-opening and needle-closing timing information. This in turn allows the processing load on micro-controller to be limited, easily keeping the processing load within the limits of micro-controller's processing performance.
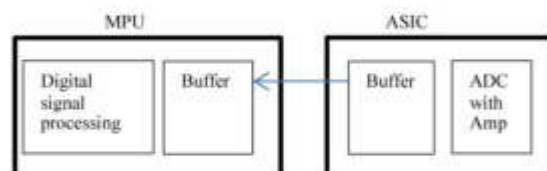
Figure 7. Functional Allocation (Case A)
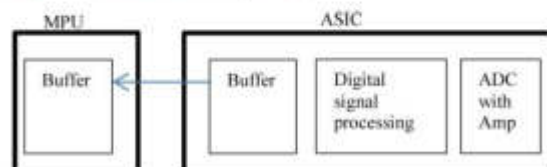
Figure 8. Functional Allocation (Case B)

Simulation-based Optimization Method for the Functional Allocation

Since the processing performance of micro-controller is limited by CPU clock frequency, RAM, and the like, it is necessary to select the functional allocation as a result of quantitatively determining that the processing load on digital signal processing section is within the processing performance of micro-controller to be adopted. For digital signal processing section, we tried to verify the feasibility of functional allocation by quantifying the processing load through simulation.

Method of Verifying the Processing Load

ELS method is well known to verify the processing load quantitatively by the partition of chipset ASIC/MCU conventionally. For verification based on ELS method, however, we must code the software and implement it on virtual ECU. That means not the verification of the optimum functional allocation but the verification of the software. In other words, if we are not able to achieve requirements of the processing load, the only alternative is to reduce the processing load by means of software, and in the worst case, the chipset must be redesign, which causes substantial rework. In this paper, we adopted the processing performance prediction method as the method of verifying the functional allocation discussed in the previous section. This method is the technology of DENSO CORPORATION that allows the processing load of the software on virtual ECU to be predicted with high accuracy, without coding the software. The key components are twofold. One of them is the software modeling and another is ECU chipset modeling. In the software modeling, more specifically, the characteristics of the existing software are profiled in detail, with the key design parameters affecting the processing performance alone extracted to equip the modeling with flexibility. On the other hand, ECU chipset modeling is capable of expressing versatile new ECU chipset models by combining common models and setting key design parameters affecting the processing performance. The combination of these two key components has allowed the processing load to be verified these case of the chipset partition [2]. With a model based on the original software and another simulating digital signal processing section provided for the study this time, we calculated by means of simulation the processing load for the case in which digital signal processing section is integrated into the routine of the original software. We verified the feasibility of the two functional allocations, Cases A and B, by comparing the results of calculation of the processing load obtained and the limit value for the processing performance of micro-controller to be adopted. As described above, the difference in the data quantity between these two functional allocation plans is noteworthy. In the simulation this time, accordingly, we calculated the processing load in each functional allocation plan by performing calculation with the specific parameters characterizing the communication load being changed (Table 1).

Table 1. Setting Parameters of simulation

Clock frequency	320MHz
Pipeline stage	Frontend:3 /Backend:3
Fetch hit ratio of cache	0.95
Interconnect bus division ratio [LMem/ Flash/ GMem/ Peripheral]	1/ 4/ 2/ 4
Latency [LMem/ Flash/ GMem/ IO]	3/ 5/ 6/ 8

Results of the Calculation of the Processing Load and Consideration

Figure 9 shows the rates of the increase in the processing load for Cases A and B with reference to the values prior to the change assumed to be 1. In Case A, it turned out that functional allocation to micro-controller is difficult because the result of the calculation of the processing load exceeds the processing performance of micro-controller. As estimated with this case, a vast amount of data communication from ASIC to micro-controller is considered to be the major reason. In Case B, it turned out that the functional allocation to micro-controller is possible because the result of calculation of the processing load is below the processing performance of micro-controller. In this case, the reduction of the communication load is considered to be the major contributing factor, since the data sent from ASIC to micro-controller is limited to the timing information as estimated.

Based on the above discussion, we concluded that the allocation of digital signal processing section to ASIC is the optimum measure.
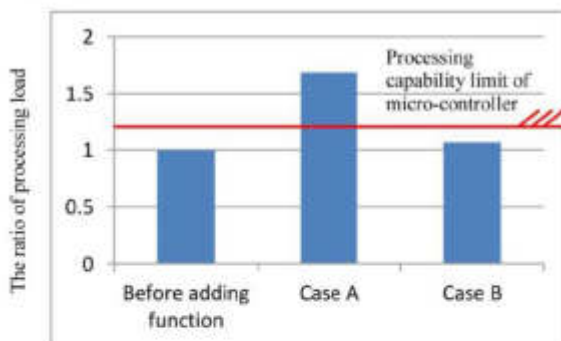
Figure 9. Comparison of Processing load between before adding function, Case A and Case B

Conclusions

Achieving injection of a small amount of fuel using GDI solenoid injectors requires a high-precision injection in the half-lift region. Achieving high precision required the detection of the injector needle lift timing correlated with the amount of injection. The detection was attained by acquiring the characteristic point of the injector current appearing when the injector needle opens and that of the injector voltage appearing when the injector needle closes.

The detection of the injector needle lift timing was attained by the system consisting of current detecting section, gain adjusting section, ADC section (with resolution of 10 bits and sampling cycle of 1 MHz), and digital signal processing section. In the digital signal processing section, signal analyzing algorithm is implemented.

There is correlation between the actual needle lift timing and the needle lift timing detected on the evaluation board. We will verify them with enough sample data.

With the current sense section, the gain adjusting section, and the ADC section needing a wide dynamic range and high precision, the functional allocation not to micro-controller but to ASIC is required.

Simulation was used to determine the functional allocation for the digital signal processing section with the aim of taking the processing load on the micro-controller into consideration. Allocating the digital signal processing section to ASIC, in which the processing load is less affected, is the optimum measure.

The ECU functional allocation in this paper has potential to detect the injector needle lift timing with high precision.

References

1. Scafati, F., Pirozzi, F., and Cannavacciuolo, S., "Real Time Control of GDI Fuel Injection during Ballistic Operation Mode," SAE Technical Paper 2015-24-2428, 2015, doi:10.4271/2015-24-2428.

2. Takashi, A., Hideki, S., and Soichiro, A., "MULTI-CORE PROCESSORS ARCHITECTURE DESIGN METHOD FOR NEXT GENERATION AUTOMOTIVE CONTROLLER," FISITA 2016 World Automotive Congress, F2016-VESA-002.

Abbreviations

GDI - gasoline direct injection

ASIC - application specific integrated circuit

ADC - analog to digital converter

[39]   Wieclawski, K., Maczak, J., Szczurowski, K., "Electric Current Waveform of the Injector as A Source of Diagnostic Information," *Journal of Sensors*, Vol. 20, 4151, doi: 10.3390/s20154151, MDPI, Polonia, 2020.

sensors

MDPI

Electric Current Waveform of the Injector as a Source of Diagnostic Information

Krzysztof Więcławski *[iD], Jędrzej Mączak[iD] and Krzysztof Szczurowski

Faculty of Automotive and Construction Machinery Engineering, Warsaw University of Technology, 02-524 Warsaw, Poland; jedrzej.maczak@pw.edu.pl (J.M.); krzysztof.szczurowski@pw.edu.pl (K.S.)
* Correspondence: krzysztof.wieclawski@pw.edu.pl; Tel.: +48-(22)-2348117

check for updates

Abstract: The article discusses the method of evaluation of the fuel injector operation based on the observation of the electric current parameters, which were measured with a current transducer using the Hall effect, during the dosing process. This method relies on comparison of the electric current-related values of the examined injector with the model characteristics, which are representing the properly functioning injector. A model of the fuel injector in the form of the electric current waveform that describes the changes in the electric current and voltage during its work is presented in this article. Complex equations describing the fuel injector model under discussion account for the characteristics of the current variations, with no damage-induced modifications. Due to these, the modeled electric current/voltage waveform mirrors the real conditions. The use of a mathematical model describing the voltage–current phenomena occurring during the injector operation allows determining the actual beginning and duration of the injection. The model can also be used to develop new injector diagnostic methods that can be implemented in the engine controller (ECU).

Keywords: injector; diagnostics; inductance

1. Introduction

This article discusses the method of evaluation of the fuel injector operation based on the observation of the electric current parameters during the dosing process. This method relies on comparison of the electric current-related values of the examined injector with the model characteristic, representing the properly functioning injector. Figure 1 shows a cross-section of an electromagnetic fuel injector.
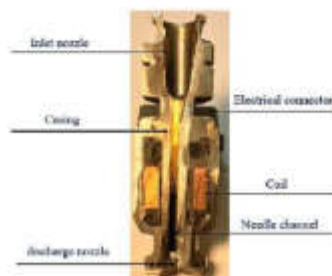
Figure 1. Cross-section of an electromagnetic fuel injector.

A model of the fuel injector in the form of the electric current waveform describing changes in electric current and voltage during its work is presented in a further section of this article.

393

Complex equations describing the fuel injector model under discussion account for the characteristics of the current variations, with no damage-induced modifications. The developed model represents a properly functioning injector at the predetermined control parameters.

Object modeling is widely used in diagnostics and control. Modeled are both objects and processes, where the model plays a crucial role. The fuel injector is such an assembly, playing a vital part in the supply system of a combustion engine. Modeling the control process requires developing the mechanical model, describing the needle displacements, determining the hydraulic model describing variations in the fuel flow, and developing the model of the changes in electrical values (Hung et al.) [1]. In this work, the current flow through the injector coil was described based on Kirchhoff's law, and the basis of the mathematical model was Newton's second law. The model built in this way based on a real object was verified in the ANSYS Maxwell and Simplorer environment. In [2], apart from determining the influence of control parameters on the resulting magnetic force, the flow of magnetic flux through the injector core was modeled. In [3,4], the simulation results were described, thanks to which the influence of the size and position of individual ferromagnetic elements on the shape of the magnetic flux, including the air gap and input voltage, was determined. Czarnigowski et al. [5] stressed that the electric current waveform can be used to compare the functioning of different injectors because it is a precise verification tool. Yao et al. [6] performed an analysis of the dynamics of the injector needle oscillation, using the mathematical model of the injector implemented in the Amesim software [7]. Factors influencing needle movement and injection quality were determined. Thanks to this, a method of optimizing the injector structure was proposed. Yan et al. [8], describing the injector dosing process, divided it into six phases, analyzing each of them and presenting theoretical corresponding calculation models. The agreement between the forecasted model and the experiment was confirmed based on current characteristics and electromagnetic parameters. (In this article, the injection phases have been divided into five ranges describing changes in current and eight ranges representing changes in electric voltage). Electromagnetic field simulations based on the injector model carried out by Wanatabe et al. [9] allowed a description of the phenomena occurring during dosing and estimation of actuator response—the needle. The magnetic field geometry was also analyzed, showing the direction of changes in the injector design in order to increase the dosing efficiency. Kay et al., in article [10], describe the influence of external factors on the method of atomizing fuel. The analysis of injection operation was carried out by dividing the injection into three main phases. Based on the fact that the quality of the injection is significantly dependent on the electric current, Xue [11] presents the optimization of the dosing process by modeling the shape of the electric current flowing through the injector coil. This resulted in the best response time of the actuator. The model includes magnetic hysteresis and the magnetostriction of ferromagnetic materials. Hung et al. [12], using differential equations, modeled the operation of the fuel injector. Thanks to the electric and mechanical model, they determined the impact of the return spring rigidity, needle mass, and number of turns of the electric coil on the characteristics of the injector, current size, electromagnetic force, and dynamics of the needle. It was proven in [13] that the use of thermodynamic models and their analysis can be used to improve the efficiency of fuel combustion and the use of energy contained in it. Kusakabe et al. [14] presented a simulation model that takes into account the magnetic resistance of the injector core, nonlinearities in its dynamics, and magnetic force. Based on the simulations, it was found that the observed nonlinearity results from the residual magnetic force. The electric current parameters were shaped, which resulted in a reduction of unfavorable oscillations of the needle.

The approaches outlined above allows for a complementary optimization of the injector's work. In the literature, different approaches to the modeling can be found.

Modeled are the flow through the injector nozzle and the needle displacements, for the relationship between the fuel flux and the injection pressure and temperature to be determined [15–17]. After the characteristics have been determined, they are verified for different control parameters. Mathematical models are used to evaluate the system's efficiency and to detect various kinds of damage. Injectors and fuel systems are modeled in special programming environments, such as Matlab/Simulink, ANSYS, or

394

Amesim [18], due to which the system's properties and its functioning during operation [6,19,20] are determined. Modeling enables assessing the gasoline injection and determining its most important determinants [21], as well as evaluating the influence of the design changes on the control result, i.e., the obtained engine power and quality of the combustion process. Dutka et al. [22] can detect failures in the intake system on the basis of mathematical modeling with state equations using the Kalman filter. The behavior of the needle shutting off the fuel flow is analyzed by means of computational models, which facilitates later management of the injector work [23]. Modeling that leads to the detection of damages within the system improves the reliability and safety of its functioning [24,25], by providing the tools enabling a fast diagnostics. The object-based models allow for a detailed detection of injector damages, such as blockage of the outlet channel. The process of injector control is modeled [26], and the influence of control on the solid particle emissions by the combustion engine's exhaust system is determined [27]. The effect of the fuel composition on the efficiency of the supply system and combustion process is investigated. By means of modeling the combustion processes, the possibilities of combining different fuels are tested, which results in improving the efficiency of utilizing the energy of the fuel with a simultaneous optimization of the exhaust gas composition [28,29]. Characteristics of the differential current of the current-related waveforms are analyzed, which enables the fault detection in injectors [30,31]. Modeling is useful as far as diagnostics is concerned for performing analyses, as well as determining the performance of a working element during changes of the work parameters and control values, due to which both the diagnostics and the process of the injector control can be advanced.

The mathematical model of an electromagnetic valve is presented in this paper, discussing changes in the values of the electric current and voltage in the coil, which are characteristic for both the parameter values and time phases.

This model is based on the exponential function, describing an increase in the current and voltage decay in the DC serial circuit in the circuit controlling the injector. The current and voltage equations have been successively formed, relating to the adequate sections of the work. Subsequently, the obtained results have been compared to the actual results obtained in the course of the laboratory experiment at the test bench mimicking the injector's working conditions.

2. Laboratory Test Stand

Fuel injectors were tested on a stand dedicated for testing car injectors (Figure 2). This stand was expanded with a specially constructed computer-based measuring and control unit that managed the dosing and enabled the change of control parameters, as well as their observation and recording. This unit uses a programmable controller that controls a specially designed and manufactured for this project control and measuring module (Figure 3), which can replace the engine controller in the implementation of the injector control function and allows the observation of voltage and electric current. This allows automatization of the recording and obtaining a repeatability of measurements.

Figure 2. Experimental stand.

Figure 3. Control and measuring module.

The MP-S module allows controlling the injector time parameters with microsecond accuracy. Recorded voltage and current waveforms as well as control signals are saved on the controller disk for later analysis. The system ensures a high reproducibility of results, since the beginning and duration times of the control signals are strictly controlled. In conjunction with process automation, this allows the creation of a large statistical data resource.

3. Fuel Injection Phases

The injector coil circuit can be considered as a serial circuit of the connected electrical resistance R and inductance L (with the electromotive force ε_0). The circuit $R - L$ is described by Kirchhoff's Law:

$$RI + L\frac{dI}{dt} = \varepsilon_0. \tag{1}$$

Solving the following differential equation results in obtaining the equation describing the current flow (2) and voltage decay (3) in the $R - L$ circuit (injector):

$$I(t) = \frac{\varepsilon_0}{R}\left(1 - exp\left(-\frac{R}{L}t\right)\right) \tag{2}$$

$$|U_L|(t) = \varepsilon_0\left(exp\left(-\frac{R}{L}t\right)\right) \tag{3}$$

where exp—exponential function, I—electric current [A], U_L—electric voltage on inductance [V], L—inductance [H], R—resistance [Ω], ε_0—electromotive force [V], and t—time [s].

Inductance of the injector core depends on the geometry of the coil core and on the magnetic permeability:

$$L = \frac{\mu_0 \ \mu_R N^2 S}{l} \tag{4}$$

where N—number of turns, S—surface area of the coil cross-section $\left[\text{m}^2\right]$, μ—magnetic permeability of the material, μ_0—magnetic permeability of the vacuum $4\pi \ 10^{-7}$ [H/m], μ_R—relative magnetic permeability of the material, and l—length of the coil core [m].

Figure 4 shows an electric current-related waveform of a dosing injector that was obtained in the course of the experiment. The current–voltage waveforms of all types of injectors are of similar character, form, and ranges. The differences are only in the scale of quantities shown in the electric current, voltage, and injection pressure waveforms.
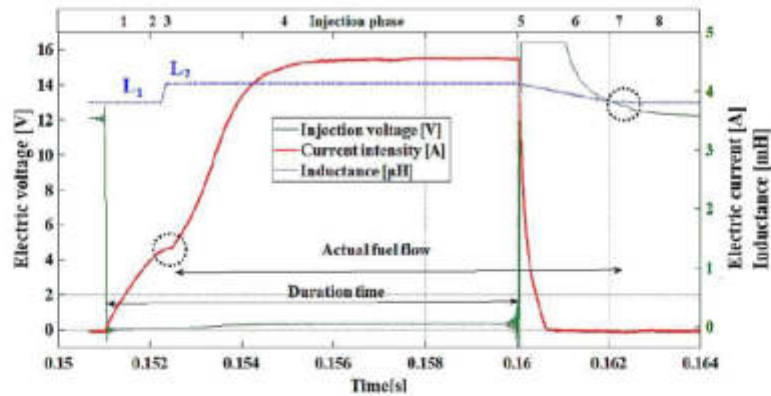
Figure 4. Recorded time-related waveform of injector current–voltage.

In Figure 4, the dotted circles denote points of opening and closing the injector nozzle by the moving needle, which were determined on the basis of changes in the inductance (L_1, L_2) and the related change of time constants τ_1 and τ_2.

$$\tau = \frac{L}{R} \tag{5}$$

A change of the time constant and inductance determines ranges corresponding to the successive differential equations describing the current–voltage changes during the fuel injection. Observing the change of the derivative of the current-related waveform regarding the increase in electric current (Figure 4) enables determination of the actual needle-lifting point (from 0.1524 s to 1.526 s) with a microsecond accuracy. However, analysis of the voltage decay allows for an observation that the change in the derivative of the waveform enables ascertainment of the real point of the needle settling in the nozzle, which finishes off the flow (around 0.1622 s). As can be seen in Figure 4, the real points of opening (injection phase no. 3: $dI/dt = 0$) and closing the flow (injection phase no. 7: $dU/dt = 0$), are significantly delayed relative to the signal controlling the injector.

A detailed analysis of waveforms allows determining the specific phases of the injector work and the assignment of differential equations corresponding to the successive phases of the work cycle. These phases are marked in the table in Figure 4 in its top section. On the axis on the left side of this figure, there is a range of electric current shown, on the vertical axis on the right, there is a range of voltage in the coil. The horizontal axis denotes time in seconds.

Individual phases of the injector work may be identified through change in the core inductance (ΔL):

- phase 1—nozzle closed, core inductance equals L_1,
- phase 2—opening of the nozzle, change in core inductance from L_1 to L_2,
- phase 3—nozzle open, core inductance equals L_2,
- phase 4—closing of the nozzle: core inductance L_2 changes to L_1,
- phase 5—nozzle closed core inductance equals L_1.

4. Analysis

4.1. Analysis of Changes in Voltage Waveforms

Below, the successive ranges of the time-related current–voltage waveforms are discussed, together with their mathematical and physical description, and conclusions related to the possibilities of the

397

diagnostics of the fuel injector are presented. The specific slope and position of the curve representing the model were obtained by inserting the following coefficients in the model equations:

- location coefficient (f_p),
- directional coefficient (f_k), and
- influence of pressure (f_{press}).

The purpose of the pressure coefficient (f_{press}) [32] is to obtain the modeled electric current value in the range according to the actual value. Depending on the modeled waveform segment, this coefficient takes different values. For example, for ranges 1 and 2 (Figure 4), f_{press} coefficient will be less than unity, because the maximum current in that range is less than the classic Kirchhoff equation. This is the range of the characteristic bending of the electric current waveform, defining the needle rise (Figure 4). This action is opposed by fuel pressure; therefore, the current in this range is higher the higher the fuel pressure, which is associated with a higher value of the coefficient.

The values of directional and location coefficients depend not only on electric current parameters, but also on the position in the space between the ordinate and abscissa axis of the modeled waveform section. These coefficients are a function depending on the coil resistance, inductance, electromotive force, electric current strength, and forces resisting the needle lifting. The successive waveform sections must begin at the points where the preceding sections end, so that the total waveform is continuous. A given set of coefficients is current for specific values of injector control parameters. In practice, the method of determining the coefficients is as follows: we measure the current waveform; then, knowing its characteristic points (given in Section 4), one should choose the coefficients experimentally, reflecting the shape and values of the waveforms in subsequent points.

The described model was based on the Kirchhoff equations. The Kirchhoff equation in a classic form describes the increase in current and loss of electrical voltage in the $R-L$ circuit, which does not perform work. The current description of the injector (needle and coil of the injector) performing the work must be expanded at least by a pressure factor, thanks to which the received current at the needle lifting point will be in line with the real one (e.g., Sections 1 and 2, Figure 4).

- Phase 0: $U_{L,0}(t)$

  Before the current impulse is started (t_{inj}), the voltage in the injector coil equals the supply voltage:

  $$U_{L,0} = \varepsilon_0 = 12\,\text{V} \tag{6}$$

- Phases 1, 2, 3, 4: $U_{1,2,3,4}(t)$

  After initiation of the current impulse, the voltage in the coil drops from the supply voltage value ($\varepsilon_0 = 12\,\text{V}$) to zero:

  $$U_{L,1,2,3,4}(t_{inj}) = 0 \tag{7}$$

  Such a state is maintained throughout the whole injection duration. This allows for the determination of the moment when the controller starts the process of fuel injection.

- Phase 5: $U_{L,5}(t)$

  After completion of the injection, as a result of releasing the energy compensated in the coil, there occurs an inductive voltage spike, exceeding the value of the source voltage ε_0. The inductive voltage spike results from the rapid decay of the current (Section 5, R increases to ∞; I decreases to 0):

  $$U_{L,5}(t) = \frac{dI}{dt}\,L. \tag{8}$$

  The high value of the current derivative (rapid decay to zero) multiplied by the coil inductance causes a generation of the voltage that may be a few times greater than the source voltage.

398

- Phase 6: $U_{L,6}(t)$

  After the voltage spike caused by the current decay, the voltage decays exponentially, going to zero, in accordance with the equation:

  $$U_{L,6}(t) = \left(f_{press}\, \varepsilon_0 \left(exp\left(-\frac{R}{L_2(e^{-t})}\, t\right)\right)\right) + f_p; \quad \left[L_2 \text{ changes to } L_1; \; U_L \text{ changes to } U_{L_6}\right]. \qquad (9)$$

  The injector needle pushed by the force of the spring (F_S) and by the force resulting from the fuel pressure $\left(F_p\right)$, through the whole time of phase 6, shifts toward the nozzle, allowing the fuel flow. The motion of the needle being lifted due to the magnetic force is very quick. The needle returning from the coil core takes longer because the magnetic force (F_m) is much greater than the force of the spring and of the force resulting from the fuel pressure:

  $$\left(F_S + F_p\right) \ll F_m. \qquad (10)$$

  As a result of the gradual lowering, the needle rests against the nozzle, which stops the flow of fuel.

- Phase 7: $U_{L,7}(t)$

  At this point, the change in the core inductance finishes in accordance with the relationship: $L_2\left(e^{-t}\right)$ tends to L_1. This is a transient state on the waveform of the voltage decay in the coil. The needle stoppage in the nozzle seat can be observed as a temporary (approximately 0.0002 s) flattening of the waveform of the decaying voltage:

  $$U_{L,7}(t) = \left(f_{press}\, \varepsilon_0 \left(exp\left(-\frac{R}{L_1}\, t\right)\right)\right) + f_p; \quad \left[U_L \text{ changes to } U_{L_7}\right]. \qquad (11)$$

  Observation of this range in real time, through control of the derivative of the voltage waveform dU/dt, enables determination of the time of delay in closing the nozzle with the needle. The value of the voltage at which the needle settles in the nozzle determines the efficiency of the injector spring and determines the correctness of the mechanical system's operation: the needle and the guide sleeve.

- Phase 8: $U_{L,8}(t)$

  After bringing the needle to a halt and stopping the fuel flow, the voltage decay takes place in accordance with inductance L_1, going to zero, as shown in the Equation (11). Phase 8 is terminated at a point where the voltage in the coil equals zero: $U_L = U_{L0}$. Figure 5 shows a modeled voltage waveform of an injector in accordance with the presented equations (6–11). The waveform shown in a dotted green line is an actual waveform. The continuous line represents the modeled voltage waveform. In the table, in the top part of the figure, the successive ranges assigned to the current equations have been marked.

In the measured waveform, the inductive spike part has been replaced with the horizontal line due to the upper limit of the measuring transducer used during the experiment. The exponential function models an increase in the electric current because the Euler's number with the negative exponent, subtracted from one, is an asymptotic going to the set value ($I(t)$) with the delay dependent on the time constant. An equation with the negative, adequately big number in the exponent means a fast decrease of the quantity to zero (U_L).

399
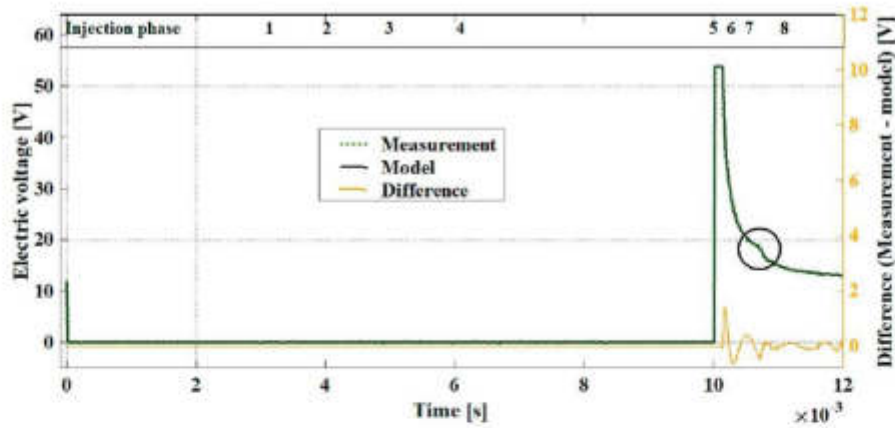
Figure 5. Measured and modeled voltage waveforms of the injector with the phases of differential equations marked.

4.2. Analysis of Changes in Electric Current Waveforms

Below, the changes in the current waveforms are discussed considering individual injection phases.

- Phase 1: $I_1(t)$

    At the point $t = 0.151$ s (Figure 4), a current impulse starts determined by a given injection time. At this point, the value of current I = 0. The current increases exponentially, activating the magnetic flux due to which the magnetic force F_m is generated. At phase 1, the magnetic force F_m is smaller than the forces counteracting the lifting of the needle F_0:

$$F_m < F_0. \tag{12}$$

The injector inductance is L_1, and the time constant is τ_1. The nozzle is closed by the needle; thus, the flow does not occur. The duration of phase 1 is about 1 ms, and the equation describing the current is as follows:

$$I_1(t) = \left(f_{press} \frac{\varepsilon_0}{R} \left(1 - exp\left(-\frac{R}{L_1} t \right) \right) \right); \quad [L = L_1]. \tag{13}$$

The way the current grows over time depends on the time constant τ. For example, the time constant for the coil in the $R-L$ circuit amounts to (5). The value of the time constant can be read from the current plot for the value $t = \tau$:

$$I(t) = \frac{\varepsilon_0}{R} \left(1 - exp\left(-\frac{t}{\tau} \right) \right) = \frac{\varepsilon_0}{R} \left(1 - exp(-1) \right) = 0.632 \frac{\varepsilon_0}{R}. \tag{14}$$

A smaller value of the time constant means a faster increase in the electric current intensity $I(t)$. The injector time constant (τ), determined on the basis of the resistance and coil inductance is an indicator that can be used while monitoring the electric efficiency of the coil. In the case of the injector, the time constant needs to be adopted as a specific value reached by the current after the voltage impulse has been turned on. It will determine the angle of the line denoting the increasing current intensity against the horizontal time axis (θ). The quantity $I(\tau)$ from Equation (14) cannot be applied, because it is preceded by the transient state (needle lifting) and the time constant has an altered value. The best solution is adopting a defined characteristic of the inclination of the curve denoting the increasing current intensity (angle (θ)), being a consequence of an injector's

qualities. Any change in the time constant ensues change in the increase in the current intensity at the range $I_1(t)$ and $I_4(t)$. This parameter can be monitored by observing the derivative of the current intensity at this range: dI/dt, or changing the time range from driving the injection pulse to the point of raising the needle (waveform bends $=> dI/dt \to 0$). Observation of the moment of occurrence of the whole transient state allows for evaluation of the magnetic force generated by the coil of the injector, i.e., the efficiency of its operation. Shifts in this range in time may indicate disturbances in the movement of the needle.

- Phase 2: $I_2(t)$

At the point $t = 0.1524$ s (Figure 4), there occurs a flattening of the line denoting current intensity—the beginning of the transient state. This is a point at which the magnetic flux has already generated the magnetic force F_m of the magnitude, which counterbalanced the forces in opposition to the lifting of the needle. Equation (13) is valid here. The duration of this range is approximately 50 μs. The whole transient state lasts about $100 \div 200$ μs (Figure 1—a black indicator, $t = 0.1525$ s). The beginning of the transient state on the current waveform is the current intensity at the point of the needle lifting I_{op}. Changes in the value of the current intensity at this point indicate electrical changes in the injector coil. Shorted circuits in the coil cause a reduction in the generated magnetic flux, which in turn leads to an increase in the current. A greater current is needed for the generation of the magnetic flux, implying that the magnetic force is enough to overcome the resistance of the needle movement. Another cause for the change in the current I_{op} is the changes in the fuel pressure before the injector (injection pressure). Differentiation between the discussed causes is possible based on the signal from the fuel pressure sensor.

- Phase 3: $I_3(t)$

This phase takes place in the middle of the transient state. At $t = 0.1525$ ms, the magnetic force F_m overcame the force F_0. At this point, the process of the injector needle lifting starts. Sliding the needle into the core is synonymous with the change in the inductance of the injector coil (ΔL, Figure 4). The equation defining the current in this phase is as follows:

$$I_3(t) = \left(f_{press}\, \varepsilon_0 \left(exp\left(-\frac{R}{L_2(1-e^{-t})}t \right) \right) \right) + f_p\;; \quad \left[L_1 \text{ changes to } L_2; \frac{dI}{dt} = 0 \right] \tag{15}$$

where L_2 is the injector inductance after lifting the needle [H], $(L_2 > L_1)$. The derivative of the waveform tends to zero, hence the denotation $I(t)$ according to Equation (15). Due to the short time of this range, the decrease in the current is slight, but it can be observed as the flattening or bending of the line in the plot (Figure 4—black indicator (circle drawn with a dotted line)). In this phase, the fuel flow starts. In the transient state, there occurs a change in the injector core inductance, resulting from the summation of the masses of the core and the needle's ferromagnetic material.

- Phase 4: $I_4(t)$

The last phase of the transient state in the current-related waveform. The waveform of the current transitions from the decrease through the state in which the derivative of the waveform equals to zero to a rapid growth (Figure 4), which is in accordance with the equation:

$$I_4(t) = \left(f_{press}\, \frac{\varepsilon_0}{R} \left(1 - exp\left(-\frac{R}{L_2}t \right) \right) \right) - f_p\;; \quad \left[L = L_2; I \text{ tends to } \frac{\varepsilon}{R} \right]. \tag{16}$$

The final phase of the transient state overlaps with the beginning of the homogenous increase in current to the steady state (maximal for a given injection time). Two sections are combined here: the end of the transient state and an increase in the current intensity to its maximum. The maximal value at the preset duration of injection depends on the duration of this impulse. The value of the current intensity tends to quotient $\frac{\varepsilon_0}{R}$ asymptotically in its exponential waveform. The needle is lifted, the injector coil core has an inductance equal to L_2, and the fuel flow is continued. Resistance (R) in the Euler's number exponent is far greater than inductance (L); e.g., the injector

used in the experiment ratio R to L amounts to: 729.06 ($R/L = 14.8/0.0203 = 729.06$. Therefore, the expression: $1 - exp\left(-\frac{R}{L_2}t\right)$ has this effect in which the value of current tends to the defined value very fast.

- Phase 5: $I_5(t)$

After the injection time is over, the electric current intensity rapidly decays to zero. Disconnecting the circuit means an increase in resistance (R) to infinity, and the electromotive force increases its value (inductance peak) significantly above the level of the source voltage, which is a key factor in the speed of the current decay, and it takes place in accordance with the equation:

$$I_5(t) = \left(f_{press}\,\varepsilon_0\left(exp\left(-\frac{f_k R}{L_2}t\right)\right)\right) + f_{p}, \quad [R \text{ tends to } \infty \Rightarrow I \text{ tends to } 0;\ \varepsilon \uparrow;\ L = L_2;]. \quad (17)$$

This equation concludes the description of changes in the electric current at $I(t) = 0$.

Figure 6 illustrates the measured and modeled electric current waveform of the injector in accordance with presented Equations (13)–(17). The biggest differences result from the oscillations of the recorded measurement. At the time point $t = 0.00102$ s, as a result of a sudden decay of current after switching off the control impulse, the electric current drops out of range. At this point, the difference between the model and measurement is the largest (0.07 A). However, this point has no diagnostic significance, so it can be omitted.
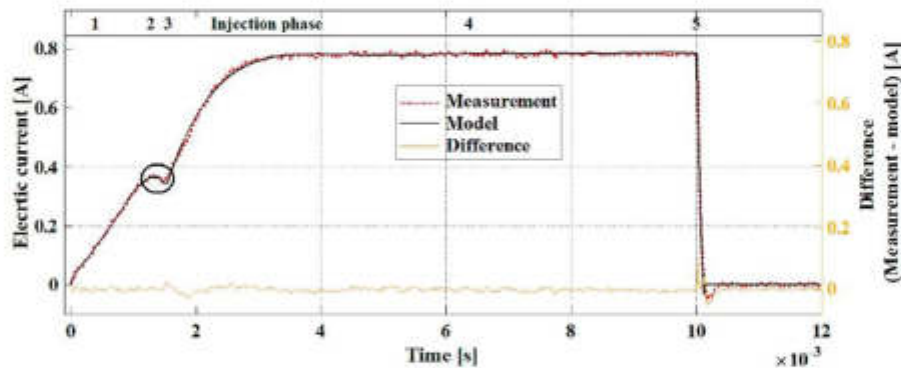
Figure 6. Measured and modeled electric current waveforms of the injector with the phases of differential equations marked.

The greatest differences are obtained for the first transient state ($\Delta = (0.019\ A + |-0.035\ A|) = 0.054$ A, Figure 7). This is a satisfactory result, and the obtained model characteristics can be a reference for determining the correctness of the injector operation. Below is the calculated correlation estimator for both waveforms (r_{pm}) (18). The correlation was calculated for the transient range, because it is the range that is most important for early injector diagnostics.

The linear correlation coefficient estimator is defined in the following equation (Pearson's correlation coefficient):

$$r_{pm} = \frac{\sum_{i=1}^{n}(p_i - \bar{p})(m_i - \bar{m})}{\sqrt{\sum_{i=1}^{n}(p_i - \bar{p})^2}\,\sqrt{\sum_{i=1}^{n}(m_i - \bar{m})^2}}\ ;\ r_{pm} \in [-1, 1]. \quad (18)$$

As a result of the calculations, the correlation coefficient $r_{pm} = 0.9935$ was obtained. The result indicates a strong fit of the model to the measurement. The maximum deviation for the value of 0.42 A (maximum transient in the current waveform) is 0.0027 A. The mathematical model of the injector can

be developed with any accuracy depending on the adopted initial values (coefficients) in the equations describing it.
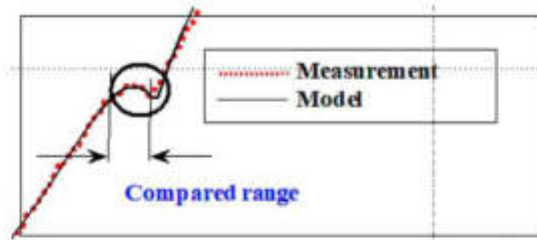
Figure 7. The transient section of the model and measurement for which the correlation was calculated (fragment of the graph of Figure 6).

Figure 8 shows the comparison of the modeled and measured time-related current–voltage waveforms of the dosing injector, which were obtained due to the presented differential equations describing the successive ranges. The waveform that was created based on the differential equations is assigned to successive ranges of the measured waveform of the current intensity and voltage as the function of time.
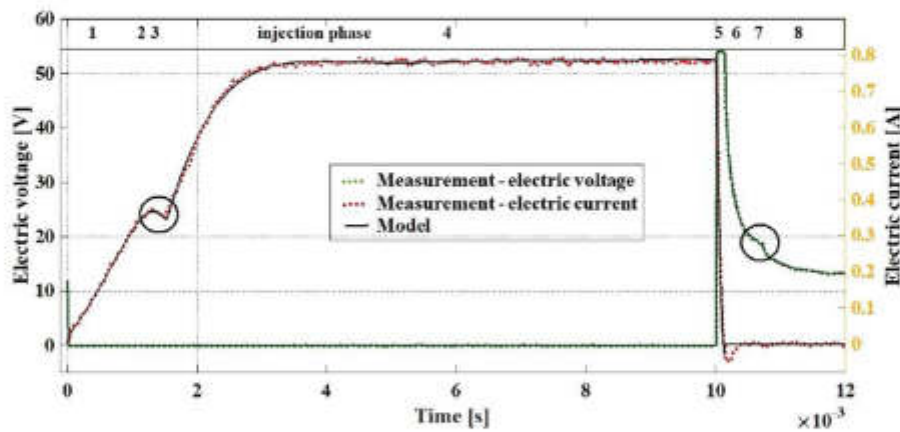
Figure 8. Measured and modeled electric current and voltage waveforms.

5. Model—Supported Analysis of Injector Operation

The developed mathematical model describes the current–voltage phenomena in successive phases of the injector work. The way of describing individual sections that make up the electric current waveform of the dosing injector and the need to glue them into the continuous form results from its complicated shape and the fact of the existence of transient states. The increase in electric current and loss of voltage in the $R-L$ circuit is described by the Kirchhoff equation. This is not true in the case where the electric circuit performing the work is considered and the derivative of the current wave changes the sign. The change of sign of the derivative of the electric current waveform can be observed at the points where the needle changes its position. This results from the work being performed by the magnetic force generated by the injector coil. When the injector needle is lifted (the beginning of the fuel injection), the magnetic permeability of the injector's core increases (μ_R), the magnetic resistance (R_m) decreases, inductance increases (L_1 changes to L_2), and the value of the coil time constant (τ)

decreases. The conversions take place over a short time, which in consequence causes a characteristic "bend" at the needle lifting point, and this is defined by the change of sign of the derivative of the current waveform (Figure 1). Precise reflection of the function representing the current waveform, whose character varies over successive phases, by means of the differential equation, requires the usage of the coefficients inserted in the developed equations (f_{press}, f_k, f_p). Due to these, the modeled electric current/voltage waveform mirrors the real conditions.

A correctly identified model of the fuel injector in the form of an electric current waveform representing a properly working injector, after being implemented in the engine controller, allows the determination of characteristic quantities that can be treated as reference values for comparison with the actual values recorded during the injector engine operation. These quantities for the discussed injector are:

- electric current value at the point of injector core inductance increase (Figure 8; $I(t) = 0.35$ A) and the time phase of this point (Figure 8; $t = 0.0014$ s),
- electric current value in the steady state (Figure 8; $I(t) = 0.78$ A),
- value of coil voltage at the inductance reduction point (Figure 8; ($U_L(t) = 18$ V) and the time phase of this point (Figure 8; $t = 0.0016$ s).

The above points on the injector current–voltage characteristics are important for controlling the injector operation, as they describe the phases of the actual fuel flow. An automatic search of these values through the engine controller (ECU) consists of observing the derivative of the current waveform in the current rise phase. The inductance increase point (spike stroke—beginning of fuel flow) determines the negative value of the derivative of this waveform. Such detection is possible after implementing elements that differentiate the observed parameter in the controller. Determining the fuel flow end point (lowering the needle to the seat and thus reducing the injector core inductance) is done to observe the derivative of the voltage waveform in the coil. After determining these points, the engine controller will have information about the characteristic values to determine the actual start and duration of the injection. These quantities can be compared with the model values obtained for current operating parameters, constituting valuable diagnostic information about the condition of the injector.

By analyzing the behavior of the injector model and comparing the modeling results with measurements made on a real object, we could identify the phenomena occurring in the work of the injector and thus the phases of its work. The correct operation of the injector is described by a mathematical model that is defined by its coefficients. Changes in the values of the model coefficients, as well as a change in the value and time phases of the other parameters, indicate the occurrence of changes caused by damage. A set of changes or individual differences can be assigned to specific damages and determined automatically thanks to simple comparisons performed by the appropriate algorithm implemented in the engine controller.

The model mapping of electrical phenomena occurring in the injector is very accurate; therefore, diagnostic methods based on a comparison of assumed and actual current waveforms describing the correct and incorrect functioning of the injector should also respond quickly to changes in technical condition. The possibility of detecting changes resulting from damage to the injector or fuel system based on such observations exists because the equations describing the dosing pattern of the injector contain such parameters as the core inductance, coil resistance, electromotive force, and current. Hence, we conclude that changes in current waveforms must result from the modification of any of the parameters listed. Thus, you can detect the change in the resistance of the connector or electric coil, the change in supply voltage, and the short circuit in the electrical circuit. The magnetic force that lifts the spire results from the flowing current, so this force can also be determined by observing the current waveform. The change in the form of reducing the fuel pressure in the system will be visible by reducing the value of the current at the needle's raising point, because the magnetic flux resulting from the flowing electric current at a lower needle resistance will carry out its raising action earlier.

404

Damage in the form of a blocked needle results in a steady increase in the electric current, without changing the derivative of the waveform, exactly as in the theoretical $R-L$ circuit. The electrical signal, which is both the power and control signal of the injector coil, contains information about the electrical and mechanical status of the injector; therefore, observation and analysis of the waveforms allows detecting both electrical and mechanical damage to the injectors.

The method of fuel injector diagnostics carried out online by the controller based on the occurrence of differences between the model and measurement could be extremely useful for a quick diagnostics of faults. Up until now, despite the extensive on-board diagnostics (OBD II), which verifies the engine control and power supply system, some damages, especially in their the initial phase, are not precisely determined by the on-board diagnostics [33,34].

6. Conclusions

In this paper, a mathematical model of the injector was presented based on the analysis of the electric current and voltage waveforms in the injector circuit. The presented model is based on selecting characteristic points of the waveform, and it was verified on the test stand with the tools used in the diagnostics of the fuel injectors. The characteristics of a properly working injector were nearly identical to those of the model. Therefore, it was proved that the electric current waveform of the injector could be a valuable source of diagnostic information about the phases of injector operation and its technical state.

The article does not present a verified diagnostic method. Experiments confirming it were carried out, but they still require scrupulous development. The authors' intention was to present a model, a tool with which new diagnostic methods can be created. These tools result directly from the differential equations described and from the explanations of physical phenomena occurring during injector dosing.

The presented considerations result from the analysis of differential equations describing the $R-L$ circuit and principles binding the electric quantities. The structure of this article is intended to describe the tools that can be used in diagnostics and to explain the physical phenomena based on the equations presented. Individual sections of the course contain diagnostically useful information. The taxonomy described in this article allows the implementation of the presented tools in the engine controller. Inherent in the operation of fuel injectors are delays in its operation, resulting from physical and mechanical properties. The timing of the electric impulse controlling the injector needle is not identical to the resulting timing of fuel flow, which must be considered in the control. The basic information resulting from the observation of the current waveforms (Figure 4) is the possibility of determining the phase of the actual fuel flow. The engine controller (ECU) can monitor the flow phases by looking for changes in the derivative of the current waveform. Changes in the flow phases not resulting from the control indicate damage to the fuel system. By observing the flow phases, you can also determine the correctness of its functioning. This information can be used to increase the accuracy and uniformity of fuel supply.

The discussed model can be applied in an injector control system by developing the operation algorithm based on modifications within the subsequent cycles of the dosing process, the duration of the injection, and the phase of the injection, in the case when differences between the actual and model characteristics are detected [35]. Thanks to the accurate characteristics of the electric current variations, any possible differences between the model characteristics and the actual measurements for the determined control parameters will prove the existence of unwelcome modifications in the fuel system or in the injector itself. The differences, depending on their type, may be mapped onto the specific failures within the system. The algorithm is easy enough to be implemented in the engine's ECU and will enable identification of the fuel system malfunctions in real time. Such implementation requires that the design of the control module should expanded by elements defining the derivatives of an electric current waveform with appropriate frequency that is adequate to the speed of the electric current changes taking place in the injector coil. The occurrence of discrepancies between the

405

measurement and the model may be determined automatically, supporting the on-board diagnostics (OBD II). Then, the controller could come into action, decreasing the danger resulting from the malfunctioning injector. Thus, the function of the injector control unit will relate to the diagnostic function, due to which they complement each other, constituting a mutual-support system. Thanks to this, the quality of the control system and engine power supply will be increased. This affects the safety of the vehicle use and ensures engine operation in accordance with the requirements of ecology, extending the operation of systems dependent on the work of the injector and the engine itself.

References

1. Hung, N.B.; Lim, O.T. A simulation and experimental study on the operating characteristics of a solenoid gas injector. *Adv. Mech. Eng.* **2019**, *11*. [CrossRef]
2. Leach, F.; Davy, M.H.; Henry, M.P.; Tombs, M.; Zhou, F. A New Method for Measuring Fuel Flow in an Individual Injection in Real Time. *SAE Int. J. Engines* **2018**, *11*, 687–696. [CrossRef]
3. Czarnigowski, J. Experiments on the Effect of Pressure and Voltage Supply on Pulse Injector Opening Time. *SAE Tech. Pap.* **2014**, *1*, 9. [CrossRef]
4. Mitukiewicz, G.; Burdzik, R.; Leyko, J. Relationship between LPG fuel and gasoline injection duration for gasoline direct injection engines. *Fuel* **2015**, *153*, 526–534. [CrossRef]
5. Czarnigowski, J.; Jakliński, P.; Zyska, T. An empirical model of current in the pulse gas injector's circuit. *Prz. Elektrotechniczny* **2014**, *90*, 195–198.
6. Yao, X.; Zhang, Z.; Kong, X.; Yin, C. Dynamic Response Analysis and Structure Optimization of GDI Injector based on Mathematical Model. *Int. J. Reliab. Qual. Saf. Eng.* **2017**, *25*, 1850008. [CrossRef]
7. Chai, B.; Gao, W. Simulation on fuel injection system for EUP based on AMESim. In Proceedings of the 2010 Second International Conference on Computer Modeling and Simulation, Sanya, China, 22–24 January 2010; Volume 3, pp. 456–458.
8. Yan, F.; Zou, H.; Xiao, Q. Simulated Analysis on Dynamic Process of Electronic-controlled Injector. Available online: http://en.cnki.com.cn/Article_en/CJFDTotal-WHGY200412022.htm (accessed on 26 July 2020).
9. Watanabe, H.; Ichise, S.; Nagaoka, T.; Tsuchiya, T. Development of Compact and High Performance Fuel Injector using Electromagnetic Field Simulation. *Soc. Automot. Eng. Jpn.* **2005**, *32*, 0019.
10. Kay, P.J.; Bowen, P.; Gold, M.; Sapsford, S.M. Studies of gasoline direct—injesction sprays at elevated ambient gas temperatures and pressures. *At. Sprays* **2012**, *22*, 305–331. [CrossRef]
11. Xue, G.; Zhang, P.; He, Z.; Li, D.; Yang, Z.; Zhao, Z. Displacement model and driving voltage optimization for a giant magnetostrictive actuator used on a high-pressure common-rail injector. *Mater. Des.* **2016**, *95*, 501–509. [CrossRef]
12. Hung, N.B.; Lim, O.; Yoon, S. Effects of Structural Parameters on Operating Characteristics of a Solenoid Injector. *Energy Procedia* **2017**, *105*, 1771–1775. [CrossRef]
13. Puškár, M.; Bigoš, P.; Kelemen, M.; Tonhajzer, R.; Šima, M. Measuring method for feedback provision during development of fuel map in hexadecimal format for high-speed racing engines. *Measurement* **2014**, *50*, 203–212. [CrossRef]
14. Kusakabe, R.; Abe, M.; Ehara, H.; Ishikawa, T.; Mayuzumi, T.; Miyake, T. Injection Quantity Range Enhancement by Using Current Waveform Control Technique for DI Gasoline Injector. *SAE Int. J. Engines* **2014**, *7*, 560–567. [CrossRef]
15. Cammalleri, M.; Pipitone, E.; Beccari, S.; Genchi, G. A mathematical model for the prediction of the injected mass diagram of a S.I. engine gas injector. *J. Mech. Sci. Technol.* **2013**, *27*, 3253–3265. [CrossRef]

16. Mohapatra, C.; Jacobsohn, G.; Baldwin, E.; Schmidt, D. Modeling Sealing in Transient Injector Simulations. In *Fluids Engineering Division Summer Meeting, Waikoloa, HI, USA, 30 July–3 August 2017*; Paper No: FEDSM2017-69309, V01AT04A009; ASME: New York, NY, USA, 2017; p. 10.
17. Yasukawa, Y.; Ishii, E.; Yoshimura, K.; Ogura, I.K. Fuel Spray Analysis Near Nozzle Outlet of Fuel Injector During Valve Movement. In *Tagung Diesel- und Benzindirekteinspritzung 2016*; Springer Vieweg: Wiesbaden, Germany, 2017; pp. 345–362.
18. Shuai, G.; Baiyu, X.; Yunfen, H.; Hong, C. Injection quantity control for GDI engines. In Proceedings of the 25th Chinese Control and Decision Conference (CCDC); Guiyang, China, 25–27 May 2013; pp. 3593–3597.
19. Marčič, S.; Marčič, M.; Praunseis, Z. Mathematical Model for the Injector of a Common Rail Fuel-Injection System. *Engineering* **2015**, 7, 307–321. [CrossRef]
20. Liu, Q.F.; Wang, C.Y.; Hu, Y.F.; Chen, H. Flatness-based Feedforward and Feedback Control for Fuel Rail System of Gasoline Direct Injection Engine. *IFAC Pap.* **2016**, 49, 775–780. [CrossRef]
21. Payri, R.; Bracho, G.; Gimeno, J.; Bautista, A. Rate of injection modelling for gasoline direct injectors. *Energy Convers. Manag.* **2018**, 166, 424–432. [CrossRef]
22. Dutka, A.; Javaherian, H.; Grimble, M.J. Model-based engine fault detection and isolation. In Proceedings of the American Control Conference, St. Louis, MO, USA, 10–12 June 2009; pp. 4593–4600.
23. Kekez, M.; Radziszewski, L.; Sapietova, A. Application of Artificial Intelligence Methods to Modeling of Injector Needle Movement in Diesel Engine. *Procedia Eng.* **2017**, 177, 303–306. [CrossRef]
24. Isermann, R. Model-Based Fault Detection and Diagnosis—Status and Applications. *Annu. Rev. Control* **2005**, 39, 71–85. [CrossRef]
25. Byungho, L.; Guezennec, Y.; Rizzoni, G. Model-Based Fault Diagnosis of Spark-Ignition Direct-Injection Engine Using Nonlinear Estimations. *Trans. J. Engines* **2005**, 114, 190–200.
26. Lee, G.S.; Sung, H.J.; Kim, H.C. Multiphysics analysis of a linear control solenoid valve. *J. Fluids Eng.* **2013**, 135, 011104. [CrossRef]
27. Wang, C.; Xu, H.; Herreros, J.M.; Wang, J.; Cracknell, R. Impact of fuel and injection system on particle emissions from a GDI engine. *Appl. Energy* **2014**, 132, 178–191. [CrossRef]
28. Flekiewicz, M.; Kubica, G. The influence of selected gaseous fuels on the combustion process in the SI engine. *Transp. Probl.* **2017**, 12, 135–146. [CrossRef]
29. Karagiorgis, S.; Glover, K.; Collings, N. Control Challenges in Automotive Engine Management. *Eur. J. Control* **2007**, 13, 92–104. [CrossRef]
30. Sebok, M.; Jurcik, J.; Gutten, M.; Korniak, D.; Roj, J.; Zukowski, P. Diagnostics and measurement of the gasoline engines injection system. *Prz. Elektrotechniczny* **2015**, 1, 77–80. [CrossRef]
31. Farooqi, Q.R.; Snyder, B.; Anwar, S. Real Time Monitoring of Diesel Engine Injector Waveforms for Accurate Fuel Metering and Control. *J. Control Sci. Eng.* **2013**, 2013. [CrossRef]
32. Więclawski, K.; Mączak, J.; Szczurowski, K. Electric Current as a source of information about control parameters of indirect injection fuel injector. *Maint. Reliab.* **2020**, 22, 449–454.
33. Komorska, I.; Wołczyński, Z.; Borczuk, A. Model-based analysis of sensor faults in SI engine. *Combust. Engines* **2017**, 169, 146–151.
34. Dąbrowski, Z.; Madej, H. Masking mechanical damages in the modern control systems of combustion engines. *J. KONES* **2006**, 13, 53–60.
35. Isermann, R. *Mechatronic Systems: Fundamentals*; Springer: Berlin, Germany, 2007; p. 624.

[40]   Barkhordarian, V., "Power MOSFET Basics," International Rectifier, Application note, El Segundo, California, S.F.

International
I⊗R Rectifier

Power MOSFET Basics
By Vrej Barkhordarian, International Rectifier, El Segundo, Ca.

www.irf.com

Power MOSFET Basics
Vrej Barkhordarian, International Rectifier, El Segundo, Ca.

Discrete power MOSFETs employ semiconductor processing techniques that are similar to those of today's VLSI circuits, although the device geometry, voltage and current levels are significantly different from the design used in VLSI devices. The metal oxide semiconductor field effect transistor (MOSFET) is based on the original field-effect transistor introduced in the 70s. Figure 1 shows the device schematic, transfer characteristics and device symbol for a MOSFET. The invention of the power MOSFET was partly driven by the limitations of bipolar power junction transistors (BJTs) which, until recently, was the device of choice in power electronics applications.

Although it is not possible to define absolutely the operating boundaries of a power device, we will loosely refer to the power device as any device that can switch at least 1A. The bipolar power transistor is a current controlled device. A large base drive current as high as one-fifth of the collector current is required to keep the device in the ON state.

Also, higher reverse base drive currents are required to obtain fast turn-off. Despite the very advanced state of manufacturability and lower costs of BJTs, these limitations have made the base drive circuit design more complicated and hence more expensive than the power MOSFET.
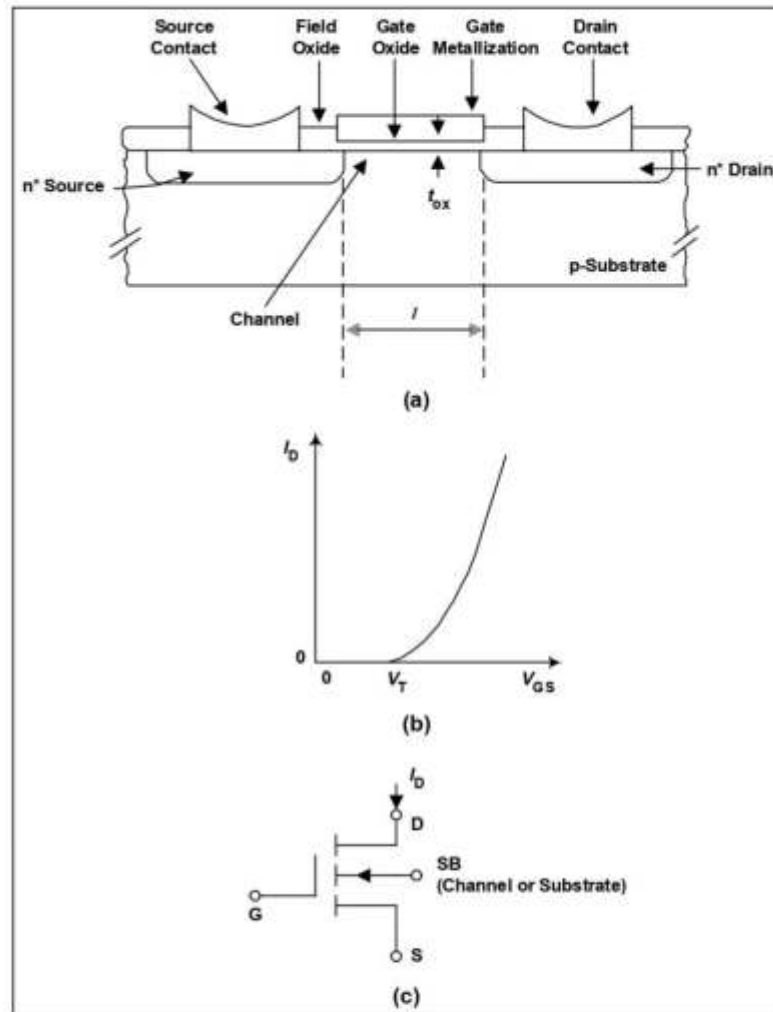
Figure 1. Power MOSFET (a) Schematic, (b) Transfer Characteristics, (c) Device Symbol.

Another BJT limitation is that both electrons and holes contribute to conduction. Presence of holes with their higher carrier lifetime causes the switching speed to be several orders of magnitude slower than for a power MOSFET of similar size and voltage rating. Also, BJTs suffer from thermal runaway. Their forward voltage drop decreases with increasing temperature causing diversion of current to a single device when several devices are paralleled. Power MOSFETs, on the other hand, are majority carrier devices with no minority carrier injection. They are superior to the BJTs in high frequency applications where switching power losses are important. Plus, they can withstand simultaneous application of high current and voltage without undergoing destructive failure due to second breakdown. Power MOSFETs can also be paralleled easily because the forward voltage drop increases with increasing temperature, ensuring an even distribution of current among all components.
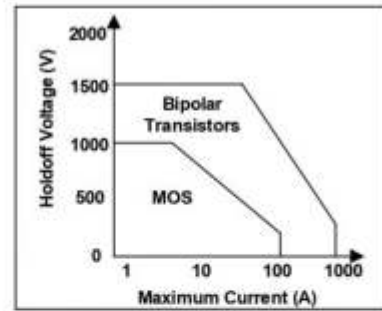
Figure 2. Current-Voltage Limitations of MOSFETs and BJTs.

However, at high breakdown voltages (>200V) the on-state voltage drop of the power MOSFET becomes higher than that of a similar size bipolar device with similar voltage rating. This makes it more attractive to use the bipolar power transistor at the expense of worse high frequency performance. Figure 2 shows the present current-voltage limitations of power MOSFETs and BJTs. Over time, new materials, structures and processing techniques are expected to raise these limits.
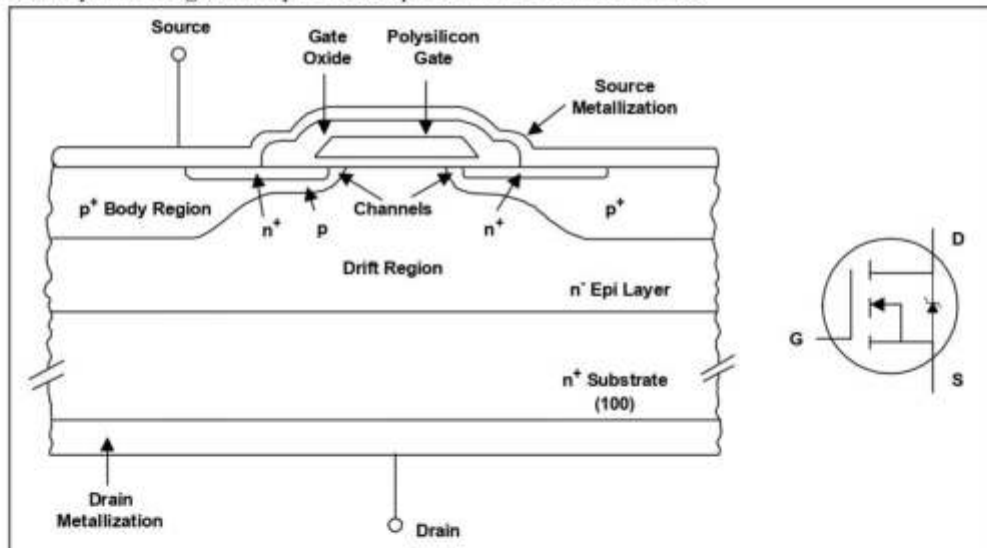
Figure 3. Schematic Diagram for an n-Channel Power MOSFET and the Device.

410

Figure 3 shows schematic diagram and Figure 4 shows the physical origin of the parasitic components in an n-channel power MOSFET. The parasitic JFET appearing between the two body implants restricts current flow when the depletion widths of the two adjacent body diodes extend into the drift region with increasing drain voltage. The parasitic BJT can make the device susceptible to unwanted device turn-on and premature breakdown. The base resistance RB must be minimized through careful design of the doping and distance under the source region. There are several parasitic capacitances associated with the power MOSFET as shown in Figure 3.

C_{GS} is the capacitance due to the overlap of the source and the channel regions by the polysilicon gate and is independent of applied voltage. C_{GD} consists of two parts, the first is the capacitance associated with the overlap of the polysilicon gate and the silicon underneath in the JFET region. The second part is the capacitance associated with the depletion region immediately under the gate. C_{GD} is a nonlinear function of voltage. Finally, C_{DS}, the capacitance associated with the body-drift diode, varies inversely with the square root of the drain-source bias. There are currently two designs of power MOSFETs, usually referred to as the planar and the trench designs. The planar design has already been introduced in the schematic of Figure 3. Two variations of the trench power MOSFET are shown Figure 5. The trench technology has the advantage of higher cell density but is more difficult to manufacture than the planar device.
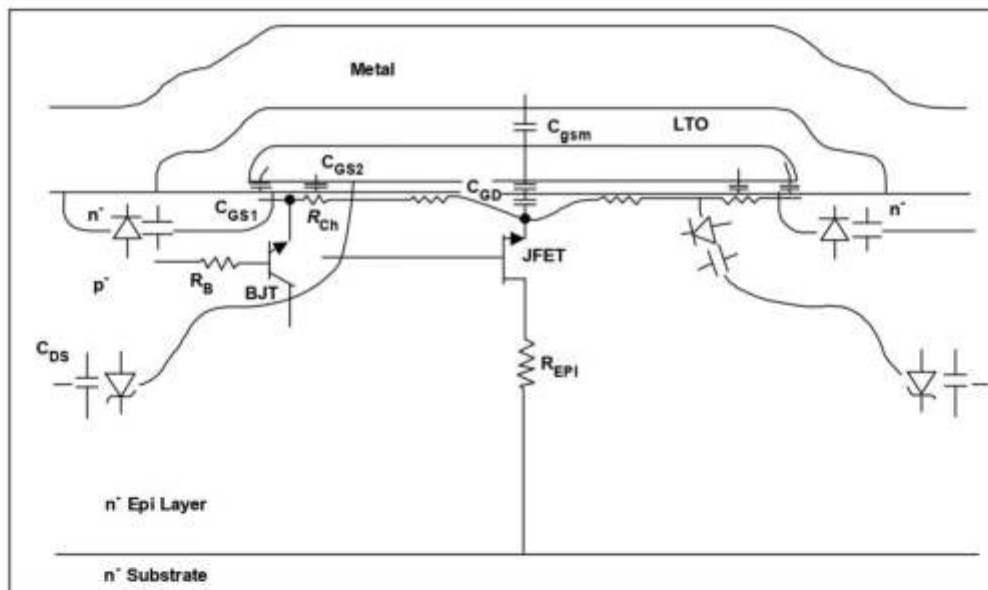
Figure 4. Power MOSFET Parasitic Components.

BREAKDOWN VOLTAGE

Breakdown voltage, BV_{DSS}, is the voltage at which the reverse-biased body-drift diode breaks down and significant current starts to flow between the source and drain by the avalanche multiplication process, while the gate and source are shorted together. Current-voltage characteristics of a power MOSFET are shown in Figure 6. BVDSS is normally measured at 250μA drain current. For drain voltages below BV_{DSS} and with no bias on the gate, no channel is formed under the gate at the surface and the drain voltage is entirely supported by the reverse-biased body-drift p-n junction. Two related phenomena can occur in poorly designed and processed devices: punch-through and reach-through. Punch-through is observed when the depletion region on the source side of the body-drift p-n junction reaches the source region at drain voltages below the rated avalanche voltage of the device. This provides a current path between source and drain and causes a soft breakdown characteristics as shown in Figure 7. The leakage current flowing between source and drain is denoted by I_{DSS}. There are tradeoffs to be made between $R_{DS(on)}$ that requires shorter channel lengths and punch-through avoidance that requires longer channel lengths.

The reach-through phenomenon occurs when the depletion region on the drift side of the body-drift p-n junction reaches the epilayer-substrate interface before avalanching takes place in the epi. Once the depletion edge enters the high carrier concentration substrate, a further increase in drain voltage will cause the electric field to quickly reach the critical value of 2×10^5 V/cm where avalanching begins.
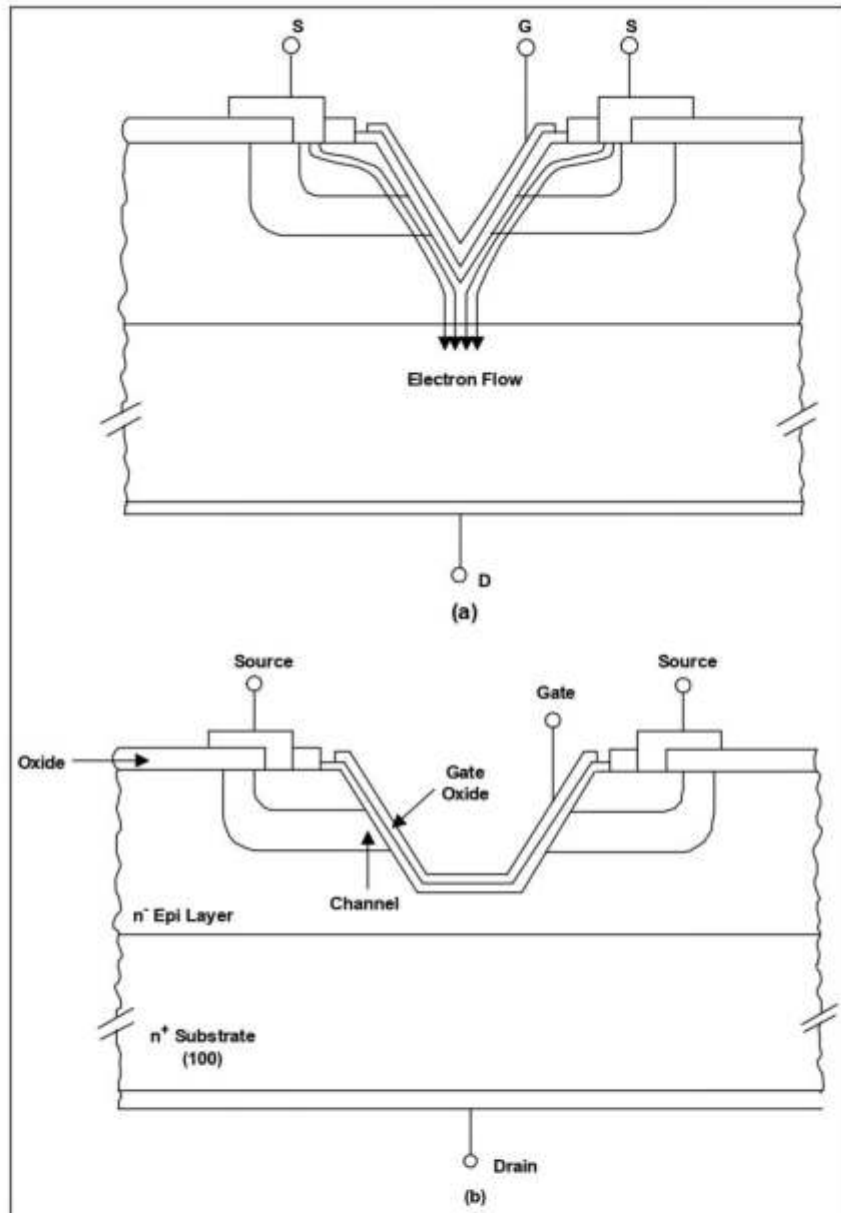
Figure 5. Trench MOSFET (a) Current Crowding in V-Groove Trench MOSFET, (b) Truncated V-Groove MOSFET

ON-RESISTANCE

The on-state resistance of a power MOSFET is made up of several components as shown in Figure 8:

$$R_{DS(on)} = R_{source} + R_{ch} + R_A + R_J + R_D + R_{sub} + R_{wcml} \qquad (1)$$

where:

R_{source} = Source diffusion resistance
R_{ch} = Channel resistance
R_A = Accumulation resistance
R_J = "JFET" component-resistance of the region between the two body regions
R_D = Drift region resistance
R_{sub} = Substrate resistance

Wafers with substrate resistivities of up to 20mΩ-cm are used for high voltage devices and less than 5mΩ-cm for low voltage devices.

R_{wcml} = Sum of Bond Wire resistance, the Contact resistance between the source and drain Metallization and the silicon, metallization and Leadframe contributions. These are normally negligible in high voltage devices but can become significant in low voltage devices.

Figure 9 shows the relative importance of each of the components to $R_{DS(on)}$ over the voltage spectrum. As can be seen, at high voltages the $R_{DS(on)}$ is dominated by epi resistance and JFET component. This component is higher in high voltage devices due to the higher resistivity or lower background carrier concentration in the epi. At lower voltages, the $R_{DS(on)}$ is dominated by the channel resistance and the contributions from the metal to semiconductor contact, metallization, bond wires and leadframe. The substrate contribution becomes more significant for lower breakdown voltage devices.

Figure 6. Current-Voltage Characteristics of Power MOSFET

TRANSCONDUCTANCE

Transconductance, gfs, is a measure of the sensitivity of drain current to changes in gate-source bias. This parameter is normally quoted for a V_{gs} that gives a drain current equal to about one half of the maximum current rating value and for a VDS that ensures operation in the constant current region. Transconductance is influenced by gate width, which increases in proportion to the active area as cell density increases. Cell density has increased over the years from around half a million per square inch in 1980 to around eight million for planar MOSFETs and around 12 million for the trench technology. The limiting factor for even higher cell densities is the photolithography process control and resolution that allows contacts to be made to the source metallization in the center of the cells.

Channel length also affects transconductance. Reduced channel length is beneficial to both gfs and on-resistance, with punch-through as a tradeoff. The lower limit of this length is set by the ability to control the double-diffusion process and is around 1-2mm today. Finally the lower the gate oxide thickness the higher gfs.

THRESHOLD VOLTAGE

Threshold voltage, V_{th}, is defined as the minimum gate electrode bias required to strongly invert the surface under the poly and form a conducting channel between the source and the drain regions. V_{th} is usually measured at a drain-source current of 250μA. Common values are 2-4V for high voltage devices with thicker gate oxides, and 1-2V for lower voltage, logic-compatible devices with

Figure 7. Power MOSFET Breakdown Characteristics

thinner gate oxides. With power MOSFETs finding increasing use in portable electronics and wireless communications where battery power is at a premium, the trend is toward lower values of RDS(on) and Vth.

DIODE FORWARD VOLTAGE

The diode forward voltage, VF, is the guaranteed maximum forward drop of the body-drain diode at a specified value of source current. Figure 10 shows a typical I-V characteristics for this diode at two temperatures. P-channel devices have a higher VF due to the higher contact resistance between metal and p-silicon compared with n-type silicon. Maximum values of 1.6V for high voltage devices (>100V) and 1.0V for low voltage devices (<100V) are common.

Figure 8. Origin of Internal Resistance in a Power MOSFET.

POWER DISSIPATION

The maximum allowable power dissipation that will raise the die temperature to the maximum allowable when the case temperature is held at 25°C is important. It is give by Pd where:

$$P_d = \frac{T_{jmax} - 25}{R_{thJC}} \qquad (2)$$

T_{jmax} = Maximum allowable temperature of the p-n junction in the device (normally 150°C or 175°C) R_{thJC} = Junction-to-case thermal impedance of the device.

DYNAMIC CHARACTERISTICS

When the MOSFET is used as a switch, its basic function is to control the drain current by the gate voltage. Figure 11(a) shows the transfer characteristics and Figure 11(b) is an equivalent circuit model often used for the analysis of MOSFET switching performance.
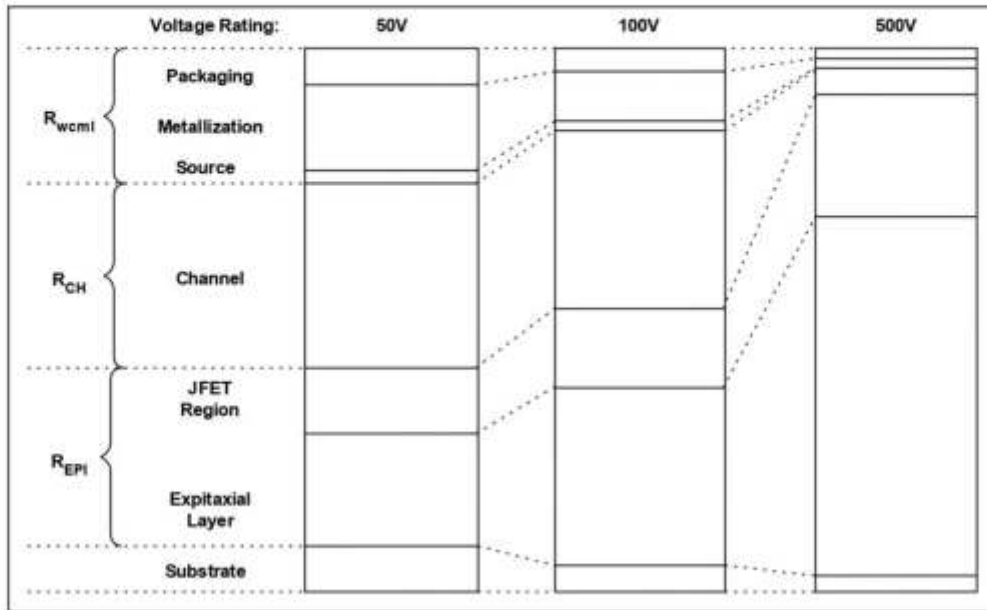
Figure 9. Relative Contributions to $R_{DS(on)}$ With Different Voltage Ratings.

The switching performance of a device is determined by the time required to establish voltage changes across capacitances. R_G is the distributed resistance of the gate and is approximately inversely proportional to active area. L_S and L_D are source and drain lead inductances and are around a few tens of nH. Typical values of input (C_{iss}), output (C_{oss}) and reverse transfer (C_{rss}) capacitances given in the data sheets are used by circuit designers as a starting point in determining circuit component values. The data sheet capacitances are defined in terms of the equivalent circuit capacitances as:

$C_{iss} = C_{GS} + C_{GD}$, C_{DS} shorted

$C_{rss} = C_{GD}$

$C_{oss} = C_{DS} + C_{GD}$

Gate-to-drain capacitance, C_{GD}, is a nonlinear function of voltage and is the most important parameter because it provides a feedback loop between the output and the input of the circuit. C_{GD} is also called the Miller capacitance because it causes the total dynamic input capacitance to become greater than the sum of the static capacitances.

Figure 12 shows a typical switching time test circuit. Also shown are the components of the rise and fall times with reference to the V_{GS} and V_{DS} waveforms.

Turn-on delay, $t_{d(on)}$, is the time taken to charge the input capacitance of the device before drain current conduction can start. Similarly, turn-off delay, $t_{d(off)}$, is the time taken to discharge the capacitance after the after is switched off.
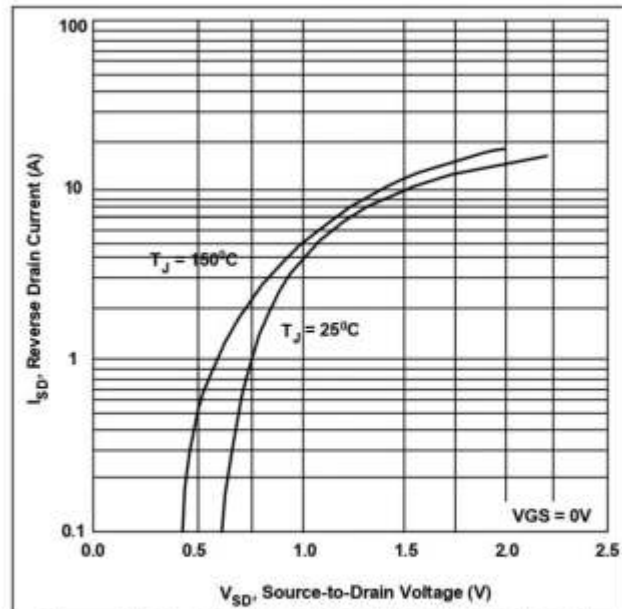
Figure 10. Typical Source-Drain (Body) Diode Forward Voltage Characteristics.
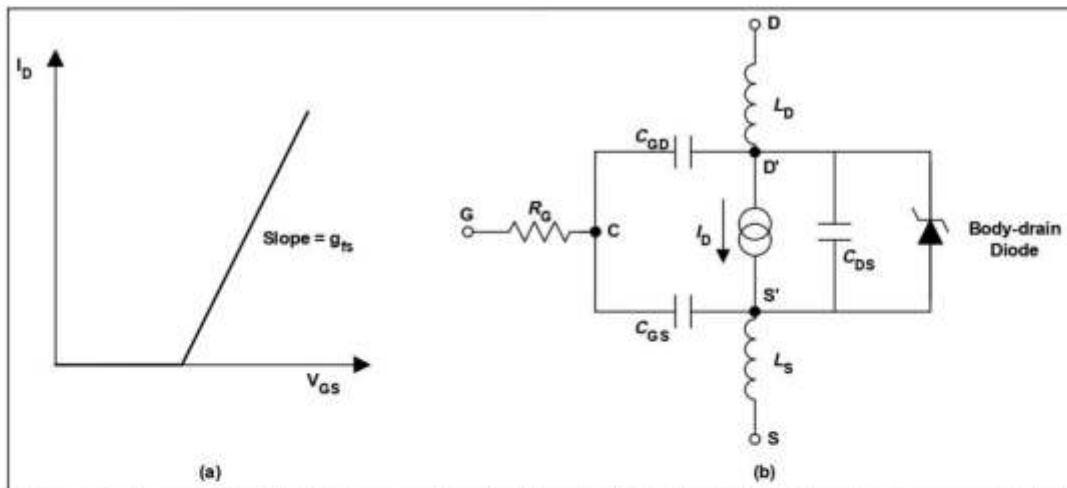
Figure 11. Power MOSFET (a) Transfer characteristics, (b) Equivalent Circuit Showing Components That Have Greatest Effect on Switching

416

GATE CHARGE

Although input capacitance values are useful, they do not provide accurate results when comparing the switching performances of two devices from different manufacturers. Effects of device size and transconductance make such comparisons more difficult. A more useful parameter from the circuit design point of view is the gate charge rather than capacitance. Most manufacturers include both parameters on their data sheets. Figure 13 shows a typical gate charge waveform and the test circuit. When the gate is connected to the supply voltage, V_{GS} starts to increase until it reaches V_{th}, at which point the drain current starts to flow and the C_{GS} starts to charge. During the period t_1 to t_2, C_{GS} continues to charge, the gate voltage continues to rise and drain current rises proportionally. At time t_2, C_{GS} is completely charged and the drain current reaches the predetermined current I_D and stays constant while the drain voltage starts to fall. With reference to the equivalent circuit model of the MOSFET shown in Figure 13, it can be seen that with C_{GS} fully charged at t_2, V_{GS} becomes constant and the drive current starts to charge the Miller capacitance, C_{DG}. This continues until time t_3.
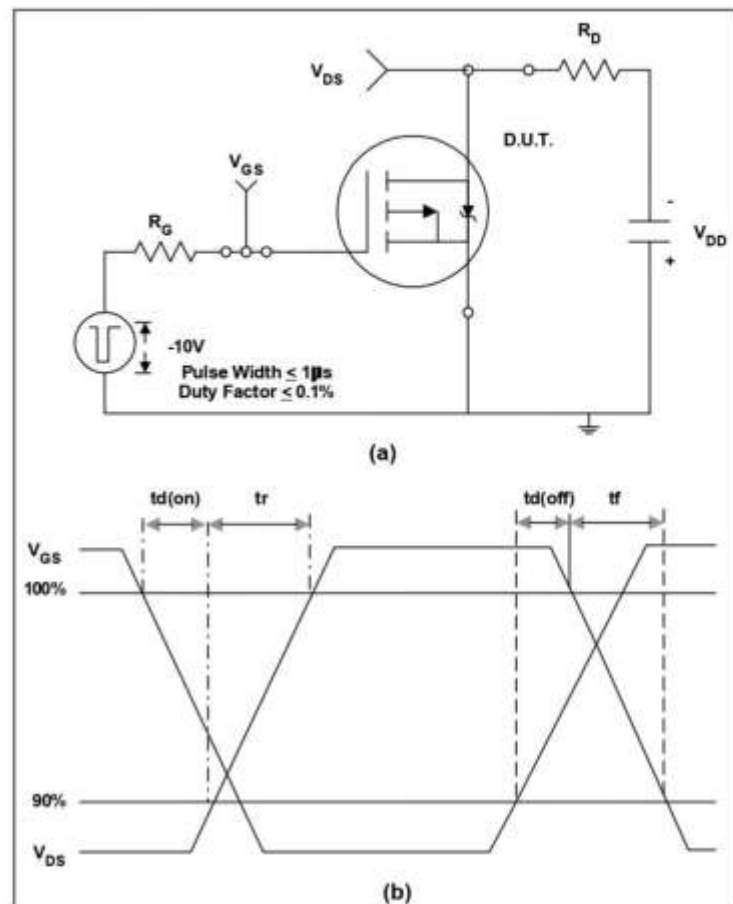
Figure 12. Switching Time Test (a) Circuit, (b) VGS and VDS Waveforms

417

Charge time for the Miller capacitance is larger than that for the gate to source capacitance C_{GS} due to the rapidly changing drain voltage between t_2 and t_3 (current $= C$ dv/dt). Once both of the capacitances C_{GS} and C_{GD} are fully charged, gate voltage (V_{GS}) starts increasing again until it reaches the supply voltage at time t_4. The gate charge ($Q_{GS} + Q_{GD}$) corresponding to time t_3 is the bare minimum charge required to switch the device on. Good circuit design practice dictates the use of a higher gate voltage than the bare minimum required for switching and therefore the gate charge used in the calculations is Q_G corresponding to t_4.

The advantage of using gate charge is that the designer can easily calculate the amount of current required from the drive circuit to switch the device on in a desired length of time because $Q = CV$ and $I = C$ dv/dt, the $Q = $ Time x current. For example, a device with a gate charge of 20nC can be turned on in 20μsec if 1ma is supplied to the gate or it can turn on in 20nsec if the gate current is increased to 1A. These simple calculations would not have been possible with input capacitance values.

dv/dt CAPABILITY

Peak diode recovery is defined as the maximum rate of rise of drain-source voltage allowed, i.e., dv/dt capability. If this rate is exceeded then the voltage across the gate-source terminals may become higher than the threshold voltage of the device, forcing the device into current conduction mode, and under certain conditions a

Figure 13. Gate Charge Test (a) Circuit, (b) Resulting Gate and Drain Waveforms.

catastrophic failure may occur. There are two possible mechanisms by which a dv/dt induced turn-on may take place. Figure 14 shows the equivalent circuit model of a power MOSFET, including the parasitic BJT. The first mechanism of dv/dt induced turn-on becomes active through the feedback action of the gate-drain capacitance, C_{GD}. When a voltage ramp appears across the drain and source terminal of the device a current I_1 flows through the gate resistance, R_G, by means of the gate-drain capacitance, C_{GD}. R_G is the total gate resistance in the circuit and the voltage drop across it is given by:

$$V_{GS} = I_1 R_G = R_G C_{GD} \frac{dv}{dt} \qquad (3)$$

When the gate voltage V_{GS} exceeds the threshold voltage of the device V_{th}, the device is forced into conduction. The dv/dt capability for this mechanism is thus set by:

418

$$\frac{dv}{dt} = \frac{V_{th}}{R_G C_{GD}} \qquad (4)$$

It is clear that low V_{th} devices are more prone to dv/dt turn-on. The negative temperature coefficient of V_{th} is of special importance in applications where high temperature environments are present. Also gate circuit impedance has to be chooses carefully to avoid this effect.

The second mechanism for the dv/dt turn-on in MOSFETs is through the parasitic BJT as shown in Figure 15. The capacitance associated with the depletion region of the body diode extending into the drift region is denoted as C_{DB} and appears between

Figure 14. Equivalent Circuit of Power MOSFET Showing Two Possible Mechanisms for dv/dt Induced Turn-on.

the base of the BJT and the drain of the MOSFET. This capacitance gives rise to a current I_2 to flow through the base resistance R_B when a voltage ramp appears across the drain-source terminals. With analogy to the first mechanism, the dv/dt capability of this mechanism is:

$$\frac{dv}{dt} = \frac{V_{BE}}{R_B C_{DB}} \qquad (5)$$

If the voltage that develops across R_B is greater than about 0.7V, then the base-emitter junction is forward-biased and the parasitic BJT is turned on. Under the conditions of high (dv/dt) and large values of R_B, the breakdown voltage of the MOSFET will be limited to that of the open-base breakdown voltage of the BJT. If the applied drain voltage is greater than the open-base breakdown voltage, then the MOSFET will enter avalanche and may be destroyed if the current is not limited externally.

Increasing (dv/dt) capability therefore requires reducing the base resistance R_B by increasing the body region doping and reducing the distance current I_2 has to flow laterally before it is collected by the source metallization. As in the first mode, the BJT related dv/dt capability becomes worse at higher temperatures because R_B increases and V_{BE} decreases with increasing temperature.

Figure 15. Physical Origin of the Parasitic BJT Components That May Cause dv/dt Induced Turn-on

419

[41] Havanur, S., "Power MOSFET Basics: understanding the Turn/On Process," Vishay Siliconix, Application note AN850, 2015.

VISHAY. VISHAY SILICONIX

www.vishay.com

| Power MOSFETs | Application Note AN850 |

Power MOSFET Basics: Understanding the Turn-On Process

by Sanjay Havanur

The question of how to turn on a MOSFET might sound trivial, since ease of switching is a major advantage of field-effect transistors. Unlike bipolar junction transistors, these are majority carrier devices. One does not have to worry about current gain, tailoring the base current to match the extremes of h_{fe} and variable collector currents, or providing negative drives. Since MOSFETs are voltage driven, many users assume that they will turn on when a voltage, equal to or greater than the threshold, is applied to the gate. However, the question of how to turn on a MOSFET or, at a more basic level, what is the minimum voltage that should be applied to the gate, needs reappraisal with more and more converters being controlled digitally. While digital control offers the next level of flexibility and functionality, the DSPs, FPGAs, and other programmable devices with which it is implemented are designed to operate with low supply voltages. It is necessary to boost the final PWM signal to the level required by the MOSFET gate. This is where things begin to go wrong, because of the misconceptions about what really turns on a MOSFET. Many digital designers look at the gate threshold voltage and jump to the conclusion that, just as with their digital logic, the MOSFET will change state as soon as the threshold is crossed.

TABLE 1 - GATE THRESHOLD SPECIFICATION FOR SIR826ADP (T_J = 25 °C, unless otherwise noted)						
PARAMETER	SYMBOL	TEST CONDITION	MIN.	TYP.	MAX.	UNIT
Static						
Gate-Source Breakdown Voltage	V_{DS}	V_{GS} = 0 V, I_D = 250 µA	80	-	-	V
V_{DS} Temperature Coefficient	$\Delta V_{DS}/T_J$	I_D = 250 µA	-	47	-	mV/°C
$V_{GS(th)}$ Temperature Coefficient	$\Delta V_{GS(th)}/T_J$		-	-5.7	-	
Gate-Source Threshold Voltage	$V_{GS(th)}$	V_{DS} = V_{GS}, I_D = 250 µA	1.2	-	2.8	V

First, the threshold voltage $V_{GS(th)}$ is not intended for system designers. It is the gate voltage at which the drain current crosses the threshold of 250 µA. It is also measured under conditions that do not occur in real-world applications. In some cases a fixed V_{DS} of 5 V or higher may be used as the test condition, but is usually measured with gate and drain shorted together as stated. This does not require searching for fine print, it is clearly stated in the datasheet. Table 1 shows the $V_{GS(th)}$ specification and test conditions for the SiR826ADP. In many applications there are concerns about the so called "induced" gate voltage, such as in the low-side MOSFET of a synchronous buck. Again, taking the gate voltage above the threshold does not automatically drive the device into a shoot-through-induced failure. $V_{GS(th)}$ is a MOSFET designer's parameter and defines the point where the device is at the threshold of turning on. It is an indication of the beginning, nowhere near the end. Gate voltage should be held below the threshold in the off state to minimize the leakage. But during turn-on, system designers can, and should, ignore it altogether.

Another curve given on the datasheet refers to the MOSFET turning on with increasing gate voltage - the transfer characteristics. This is illustrated for the same SiR826ADP device in figure 1. However, the transfer characteristics are more a measure of current variation with respect to temperature and applied gate voltage. The V_{DS} is maintained at a constant but high value, sometimes as much as 15 V, and not always disclosed in the datasheet. In the curve shown below, at a current of 20 A it is not enough to apply 3.2 V to the gate. The combination would maintain a V_{DS} of 10 V typical and a continuous dissipation of 200 W. The transfer curve is quite important when the MOSFET is operated in the linear mode, but has little relevance for switching operations.

APPLICATION NOTE

Power MOSFET Basics: Understanding the Turn-On Process
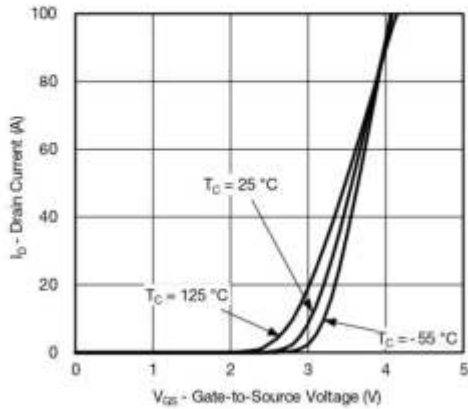
Fig. 1 - Transfer Characteristics
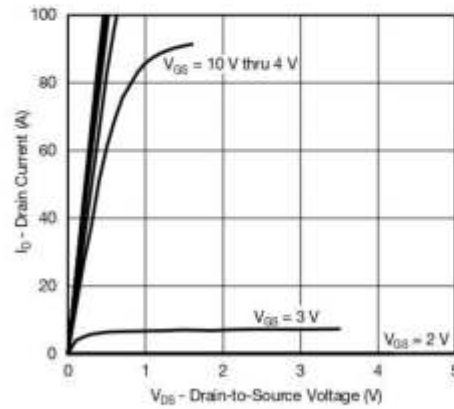
Fig. 2 - Output Characteristics

The curve that has data with the MOSFET fully on is called the output characteristics, as shown in figure 2. Here, the MOSFET forward drop is measured as a function of current for different values of V_{GS}. Designers may refer to this curve to ensure that the gate voltage is sufficient. For each gate voltage where $R_{DS(on)}$ is guaranteed, there is a range where the V_{DS} drop maintains strict linearity with current, beginning from zero. For lower values of gate voltage, as the current is increased the curve loses linearity, goes through a knee, and flattens out. Figure 3 shows the detailed output characteristics for gate voltages between 2.5 V to 3.6 V. MOSFET users usually think of this as the linear mode. However, device designers refer to the gray area as the current saturation region - for the given gate voltage, the current that can be delivered has reached its saturation limit. Any increase in applied V_{DS} will be sustained with only a slight increase in the current, whereas even a slight change in current can lead to a relatively large increase in V_{DS}. For higher gate voltages, when the MOSFET has been fully turned on, any operating point will be located in the area shaded in green to the left, marked as the resistive (or ohmic) region. Note that all curves are typical, with no minimum or maximum limits, and derived at 25 °C. At lower temperatures the gate voltage required to keep the device in the resistive region will be higher, increasing at the rate of 0.3 %/°C.
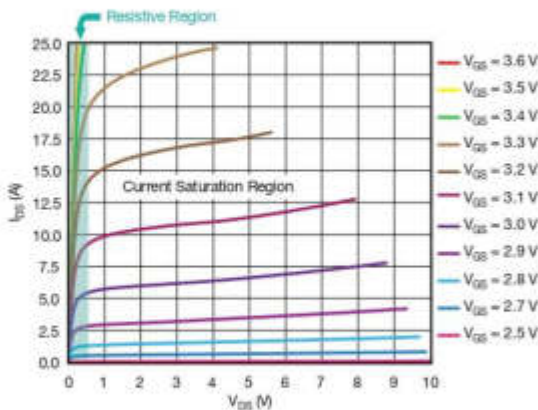
Fig. 3 - Expanded Output Characteristics
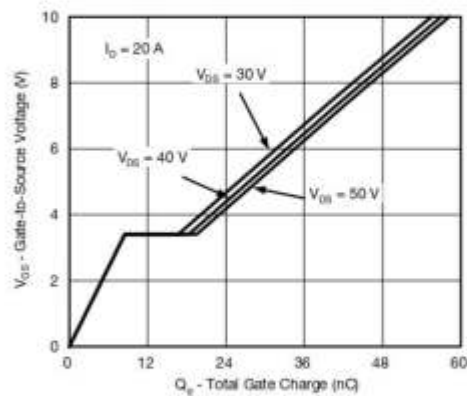
Fig. 4 - Gate Charge Characteristics

APPLICATION NOTE

421

Power MOSFET Basics: Understanding the Turn-On Process

When confronted with the output characteristics, designers invariably demand to know the $R_{DS(on)}$ at their particular operating conditions. Typically it will be at a combination of V_{GS} and I_{DS}, where the curve has strayed from the straight and narrow line into the gray area. For example, in the case above it could be a gate voltage $V_{GS} = 3.1$ V and a startup current of 10 A. They understand $R_{DS(on)}$ will be higher than specified, but can the MOSFET manufacturer provide an approximate indication? Since both V_{DS} and I_{DS} are available on the curve, there is a temptation, often succumbed to, to divide the two and arrive at the "effective" $R_{DS(on)}$. Unfortunately there is no $R_{DS(on)}$ to calculate here. It does not exist under the given conditions. Any segment of the load line that represents a resistance must pass through the origin in a linear fashion. One can of course model the load line in its entirety as a non-linear resistance. If nothing else it will ensure that any understanding of real-world behavior is maintained at the origin (0, 0).

The real clue to turning on the MOSFET is provided by the gate charge curve shown in figure 4. While the curve is routinely offered for every MOSFET, its implications are not always understood by designers. In addition, recent developments in MOSFET technology, like trench and shielded gates and charge-compensating superjunction structures, demand a fresh appraisal of this information. To start with, the term "gate charge" itself is somewhat misleading. The linearized and segmented curve does not look like the charging voltage of any capacitor, no matter how non-linear its value. In reality the gate charge curve represents a superposition of two capacitors which are not in parallel, have different values, and carry different voltages. In the literature, the effective capacitance as seen from gate terminal is defined as

$C_{iss} = C_{gs} + C_{gd}$.

While this is a convenient entity to measure and specify in the datasheet, it is worth noting that C_{iss} is not a physical capacitance. It would be a misconception to imagine that the MOSFET is turned on by simply applying a voltage to "the gate capacitance C_{iss}." As shown in figure 5, prior to turn-on the gate source capacitance C_{gs} is uncharged, but the gate drain capacitance C_{gd} has a negative voltage / charge which needs to be removed. Both capacitors are non-linear, whose values can vary widely with respect to applied voltage. The switching characteristics, therefore, are dependent more on their stored charges rather than the capacitance value at any given voltage.
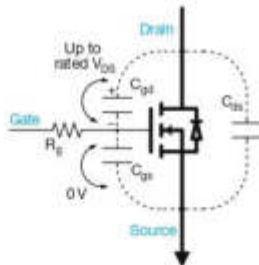
Fig. 5 - Gate Capacitances with Initial Voltages
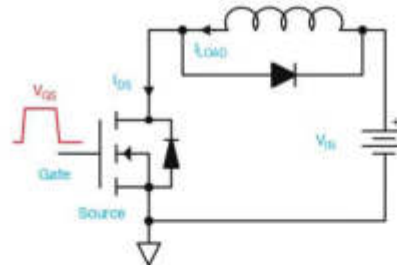
Fig. 6 - Simplified Inductive Turn-On Circuit

Since the two component capacitances that make up C_{iss} are physically different and are charged to different voltages, the turn-on process also has two stages. The exact sequence is different for inductive and resistive loads; however, in most practical applications the load is heavily inductive and can be described using the circuit model shown in fig 6. With reference to the timing diagram in figure 7,

T0 - T1: C_{gs} is charged from zero to $V_{GS(th)}$. There is no change in V_{DS} or I_{DS}.

T1 - T2: Current begins to rise in the device as the gate voltage rises from $V_{GS(th)}$ to the plateau voltage V_{gp}. I_{DS} rises from 0 A to the full load current, but there is no change in V_{DS}. The charge associated is the integral of C_{gs} from 0 V to V_{gp} and specified in the datasheets as Q_{gs}.

T2 - T3: The flat region between T2 and T3 is also known as the Miller plateau. Prior to turn-on, C_{gd} is charged to supply voltage V_{IN} and holds it till I_{DS} has peaked to I_{LOAD} at T2. Between T2 and T3, the negative charge of $(V_{IN} - V_{gp})$ is converted to the positive charge corresponding to the plateau voltage V_{gp}. This is also seen as the fall of the drain voltage from V_{IN} to near zero. The charge associated is approximately the integral of C_{gd} from zero to V_{IN} and specified in the datasheets as Q_{gd}.

T3 - T4: As the gate voltage rises from V_{gp} to V_{GS}, there is very little change in V_{DS} or I_{DS}. However, the effective $R_{DS(on)}$ reduces marginally with rising gate voltage. At some voltage above V_{gp}, the manufacturers feel confident enough to guarantee an upper limit on the effective $R_{DS(on)}$.

APPLICATION NOTE

Power MOSFET Basics: Understanding the Turn-On Process

When the load is inductive, the rise of the current in the MOSFET channel must be completed before the voltage begins to fall. At the beginning of the plateau, the device is off with simultaneous high current and voltage across drain and source. Between T2 and T3, a charge of Q_{gd} is delivered to the gate, and at the end of it the MOSFET characteristic has changed from constant current to constant resistance mode. During this entire transition there is no significant change in the gate voltage V_{gp}, which is why it is not meaningful to associate the turning on of a MOSFET with any specific gate voltage.

Similar analysis can be made for turn-off, where the same two charges must be removed from the gate in reverse order. While the sum of Q_{gs} and Q_{gd} guarantees that the MOSFET will be fully on, it does not guarantee how fast. The rate of change in voltage or current is determined by the rate at which the gate charge components are applied or removed, which is nothing but the gate drive current. While fast rise and fall times are necessary to reduce switching losses, they also introduce system-level problems of high peak voltages, ringing, and EMI, especially during inductive turn-off.
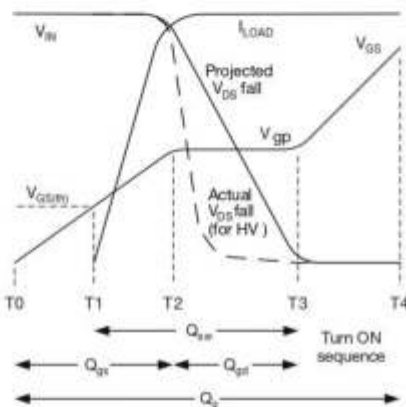
Fig. 7 - Gate Charge Components and Timings
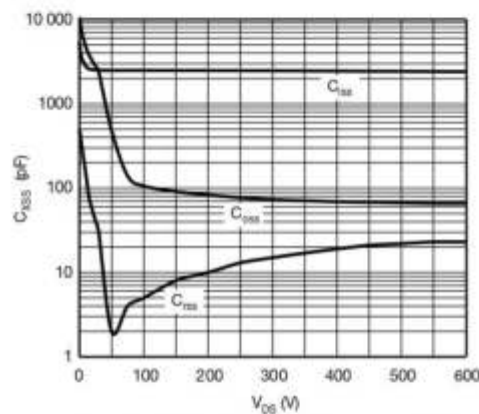
Fig. 8 - Capacitance Variations w.r.t. Applied V_{DS}

The simplified linear voltage fall shown in figure 7 assumes a constant value of C_{gd}, which is rarely true for real MOSFETs. In particular, the C_{gd} of a high-voltage superjunction MOSFET shows extreme non-linear behavior, as illustrated in figure 8 for the SiHF35N60E. There is more than a 200:1 variation in C_{rss} value within the first 100 V. As a result, the actual fall time of voltage against the gate charge curve may look more like the red dashed line in figure 7. The rise and fall times, and their corresponding dV/dt values, depend more on the C_{rss} values at higher voltages rather than the integral of the entire curve specified as Q_{gd}. While comparing devices across different design platforms, users need to be aware that a MOSFET with half the Q_{gd} will not necessarily switch twice as fast or have half the switching losses. Depending on the shape of the C_{gd} curve and its value at higher voltages, it is quite possible to have a device that has low Q_{gd} in the datasheet but shows no increase in switching speed.

CONCLUSIONS

In the real world, turning on a MOSFET is not an event but a process. Designers have to stop thinking of applying a $V_{GS(th)}$, or some other voltage, as an input at the gate which will toggle the output from high to low $R_{DS(on)}$. Questions like what is the $R_{DS(on)}$ at a gate voltage below some value or the other, are not valid because it is not the gate voltage per se that turns on a MOSFET. It is the two charges Q_{gs} and Q_{gd}, injected into the device through the gate pin, that do the job. The gate voltage will rise above $V_{GS(th)}$ and V_{gp} in the process, but that is secondary. The speed with which a modern power MOSFET turns on or off is also not a simple function of Q_{gs} or Q_{gd}. A detailed study of both the gate charge curve and capacitance characteristics is necessary to compare switching speeds, especially for superjunction MOSFETs.

423

[42]   "IRFZ44N data sheet," International Rectifier, California, USA.

PD - 94787B

International
IΩR Rectifier

IRFZ44NPbF

HEXFET® Power MOSFET

- Advanced Process Technology
- Ultra Low On-Resistance
- Dynamic dv/dt Rating
- 175°C Operating Temperature
- Fast Switching
- Fully Avalanche Rated
- Lead-Free

$V_{DSS} = 55V$

$R_{DS(on)} = 17.5m\Omega$

$I_D = 49A$
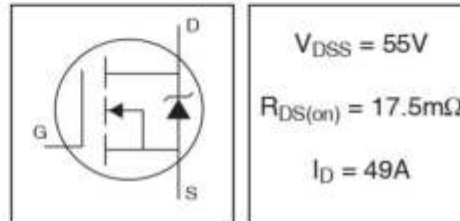
Description

Advanced HEXFET® Power MOSFETs from International Rectifier utilize advanced processing techniques to achieve extremely low on-resistance per silicon area. This benefit, combined with the fast switching speed and ruggedized device design that HEXFET power MOSFETs are well known for, provides the designer with an extremely efficient and reliable device for use in a wide variety of applications.

The TO-220 package is universally preferred for all commercial-industrial applications at power dissipation levels to approximately 50 watts. The low thermal resistance and low package cost of the TO-220 contribute to its wide acceptance throughout the industry.

TO-220AB

Absolute Maximum Ratings

	Parameter	Max.	Units
I_D @ T_C = 25°C	Continuous Drain Current, V_{GS} @ 10V	49	A
I_D @ T_C = 100°C	Continuous Drain Current, V_{GS} @ 10V	35	
I_{DM}	Pulsed Drain Current ①	160	
P_D @T_C = 25°C	Power Dissipation	94	W
	Linear Derating Factor	0.63	W/°C
V_{GS}	Gate-to-Source Voltage	± 20	V
I_{AR}	Avalanche Current①	25	A
E_{AR}	Repetitive Avalanche Energy①	9.4	mJ
dv/dt	Peak Diode Recovery dv/dt ②	5.0	V/ns
T_J	Operating Junction and	-55 to + 175	°C
T_{STG}	Storage Temperature Range		
	Soldering Temperature, for 10 seconds	300 (1.6mm from case)	
	Mounting torque, 6-32 or M3 srew	10 lbf•in (1.1N•m)	

Thermal Resistance

	Parameter	Typ.	Max.	Units
$R_{\theta JC}$	Junction-to-Case	—	1.5	°C/W
$R_{\theta CS}$	Case-to-Sink, Flat, Greased Surface	0.50	—	
$R_{\theta JA}$	Junction-to-Ambient	—	62	

www.irf.com

1

09/21/10

IRFZ44NPbF

International
IGR Rectifier

Electrical Characteristics @ $T_J = 25°C$ (unless otherwise specified)

	Parameter	Min.	Typ.	Max.	Units	Conditions
$V_{(BR)DSS}$	Drain-to-Source Breakdown Voltage	55	—	—	V	$V_{GS} = 0V, I_D = 250\mu A$
$\Delta V_{(BR)DSS}/\Delta T_J$	Breakdown Voltage Temp. Coefficient	—	0.058	—	V/°C	Reference to 25°C, $I_D = 1mA$
$R_{DS(on)}$	Static Drain-to-Source On-Resistance	—	—	17.5	$m\Omega$	$V_{GS} = 10V, I_D = 25A$ ④
$V_{GS(th)}$	Gate Threshold Voltage	2.0	—	4.0	V	$V_{DS} = V_{GS}, I_D = 250\mu A$
g_{fs}	Forward Transconductance	19	—	—	S	$V_{DS} = 25V, I_D = 25A$④
I_{DSS}	Drain-to-Source Leakage Current	—	—	25	μA	$V_{DS} = 55V, V_{GS} = 0V$
		—	—	250		$V_{DS} = 44V, V_{GS} = 0V, T_J = 150°C$
I_{GSS}	Gate-to-Source Forward Leakage	—	—	100	nA	$V_{GS} = 20V$
	Gate-to-Source Reverse Leakage	—	—	-100		$V_{GS} = -20V$
Q_g	Total Gate Charge	—	—	63		$I_D = 25A$
Q_{gs}	Gate-to-Source Charge	—	—	14	nC	$V_{DS} = 44V$
Q_{gd}	Gate-to-Drain ("Miller") Charge	—	—	23		$V_{GS} = 10V$, See Fig. 6 and 13
$t_{d(on)}$	Turn-On Delay Time	—	12	—		$V_{DD} = 28V$
t_r	Rise Time	—	60	—	ns	$I_D = 25A$
$t_{d(off)}$	Turn-Off Delay Time	—	44	—		$R_G = 12\Omega$
t_f	Fall Time	—	45	—		$V_{GS} = 10V$, See Fig. 10 ④
L_D	Internal Drain Inductance	—	4.5	—	nH	Between lead, 6mm (0.25in.) from package and center of die contact
L_S	Internal Source Inductance	—	7.5	—		
C_{iss}	Input Capacitance	—	1470	—		$V_{GS} = 0V$
C_{oss}	Output Capacitance	—	360	—	pF	$V_{DS} = 25V$
C_{rss}	Reverse Transfer Capacitance	—	88	—		$f = 1.0MHz$, See Fig. 5
E_{AS}	Single Pulse Avalanche Energy②	—	530③	150⑥	mJ	$I_{AS} = 25A, L = 0.47mH$

Source-Drain Ratings and Characteristics

	Parameter	Min.	Typ.	Max.	Units	Conditions
I_S	Continuous Source Current (Body Diode)	—	—	49	A	MOSFET symbol showing the integral reverse p-n junction diode.
I_{SM}	Pulsed Source Current (Body Diode)①	—	—	160		
V_{SD}	Diode Forward Voltage	—	—	1.3	V	$T_J = 25°C, I_S = 25A, V_{GS} = 0V$ ④
t_{rr}	Reverse Recovery Time	—	63	95	ns	$T_J = 25°C, I_F = 25A$
Q_{rr}	Reverse Recovery Charge	—	170	260	nC	$di/dt = 100A/\mu s$ ④
t_{on}	Forward Turn-On Time	Intrinsic turn-on time is negligible (turn-on is dominated by L_S+L_D)				

Notes:
① Repetitive rating; pulse width limited by max. junction temperature. (See fig. 11)

② Starting $T_J = 25°C, L = 0.48mH$
$R_G = 25\Omega, I_{AS} = 25A$. (See Figure 12)

③ $I_{SD} \leq 25A, di/dt \leq 230A/\mu s, V_{DD} \leq V_{(BR)DSS}$. $T_J \leq 175°C$

④ Pulse width ≤ 400μs; duty cycle ≤ 2%.

⑤ This is a typical value at device destruction and represents operation outside rated limits.

⑥ This is a calculated value limited to $T_J = 175°C$.
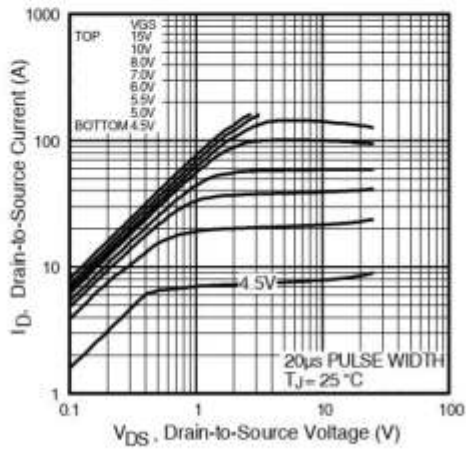
Fig 1. Typical Output Characteristics
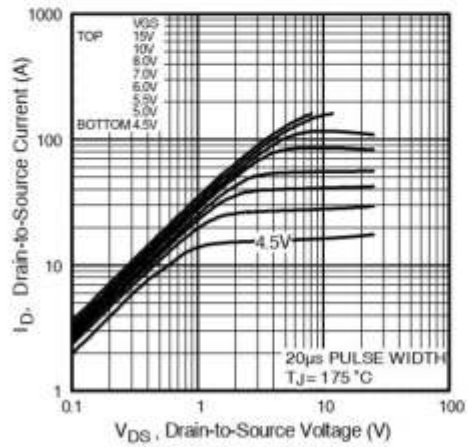
Fig 2. Typical Output Characteristics
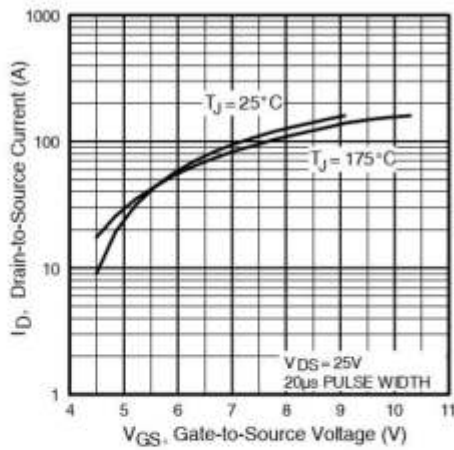
Fig 3. Typical Transfer Characteristics
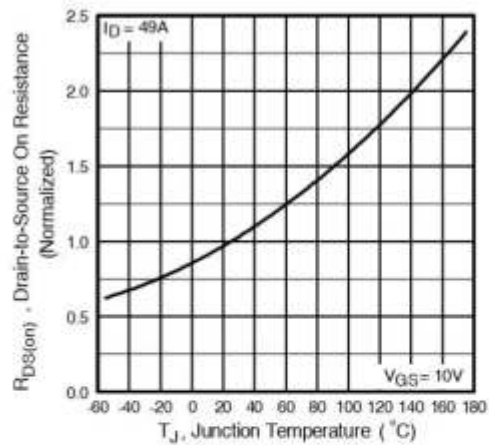
Fig 4. Normalized On-Resistance
Vs. Temperature

www.irf.com

3

426

IRFZ44NPbF
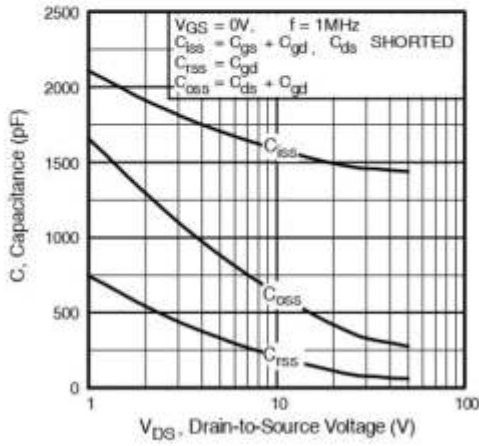
Fig 5. Typical Capacitance Vs. Drain-to-Source Voltage
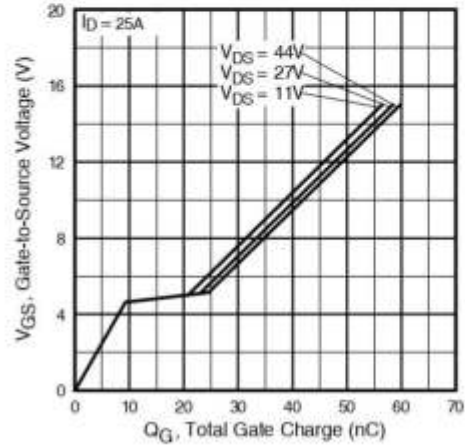
Fig 6. Typical Gate Charge Vs. Gate-to-Source Voltage
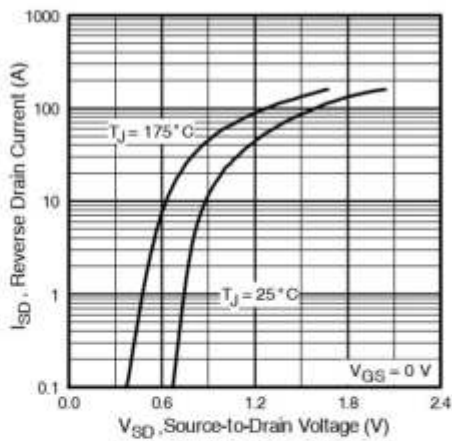
Fig 7. Typical Source-Drain Diode Forward Voltage
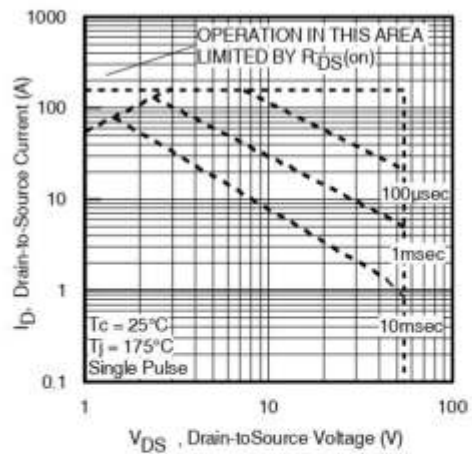
Fig 8. Maximum Safe Operating Area

4

www.irf.com
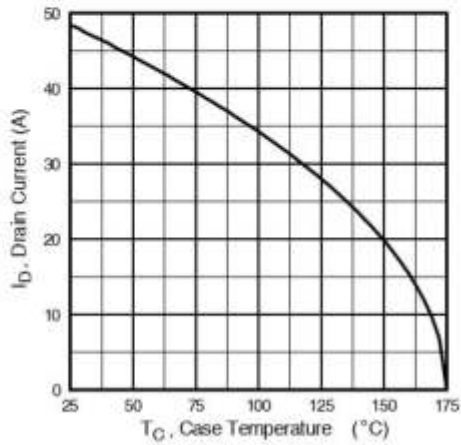
Fig 9. Maximum Drain Current Vs. Case Temperature
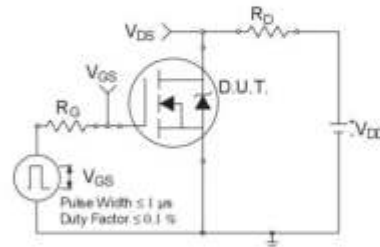
Fig 10a. Switching Time Test Circuit
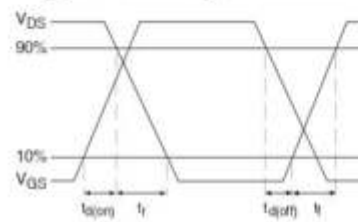
Fig 10b. Switching Time Waveforms

Notes:
1. Duty factor D = t_1 / t_2
2. Peak $T_J = P_{DM} \times Z_{thJC} + T_C$
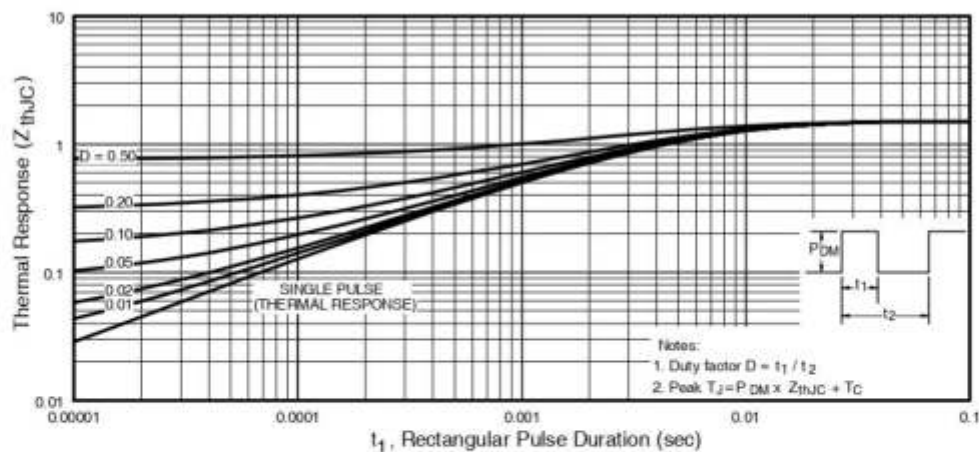
Fig 11. Maximum Effective Transient Thermal Impedance, Junction-to-Case

IRFZ44NPbF

International
IQR Rectifier
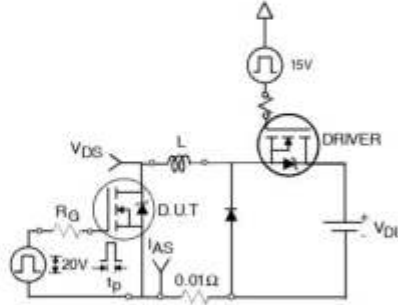
Fig 12a. Unclamped Inductive Test Circuit
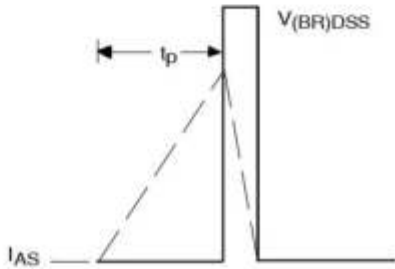
Fig 12b. Unclamped Inductive Waveforms

Fig 12c. Maximum Avalanche Energy
Vs. Drain Current

Fig 13a. Basic Gate Charge Waveform
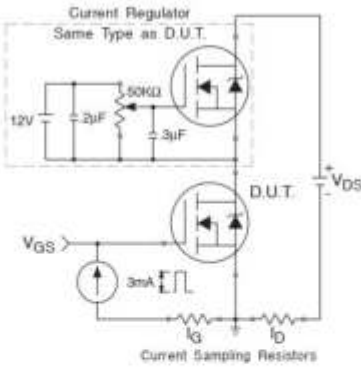
Fig 13b. Gate Charge Test Circuit

6

www.irf.com

Peak Diode Recovery dv/dt Test Circuit

D.U.T*

Circuit Layout Considerations
• Low Stray Inductance
• Ground Plane
• Low Leakage Inductance
  Current Transformer

• dv/dt controlled by R_G
• I_{SD} controlled by Duty Factor "D"
• D.U.T. - Device Under Test

R_G

V_{GS}

V_{DD}
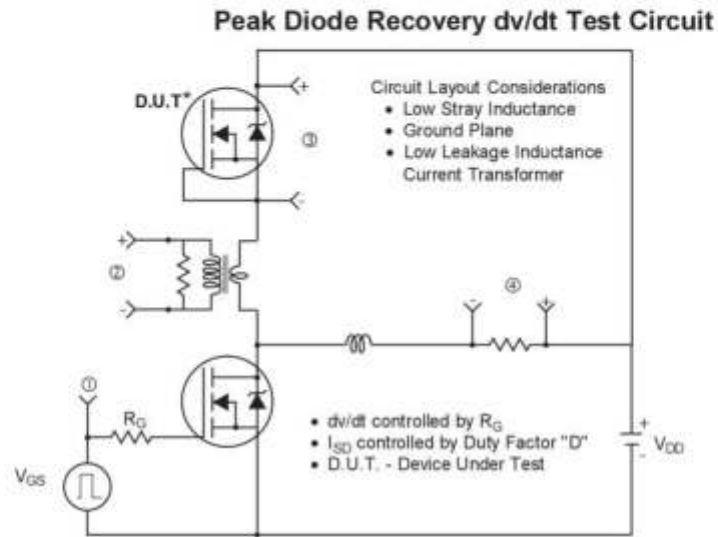
* Reverse Polarity of D.U.T for P-Channel

① Driver Gate Drive

$D = \dfrac{P.W.}{Period}$

P.W.   Period

$[V_{GS}=10V]$ ***

② D.U.T. I_{SD} Waveform

Reverse Recovery Current

Body Diode Forward Current
di/dt

③ D.U.T. V_{DS} Waveform

Diode Recovery dv/dt

$[V_{DD}]$

Re-Applied Voltage

Body Diode Forward Drop

④ Inductor Current

Ripple ≤ 5%

$[I_{SD}]$
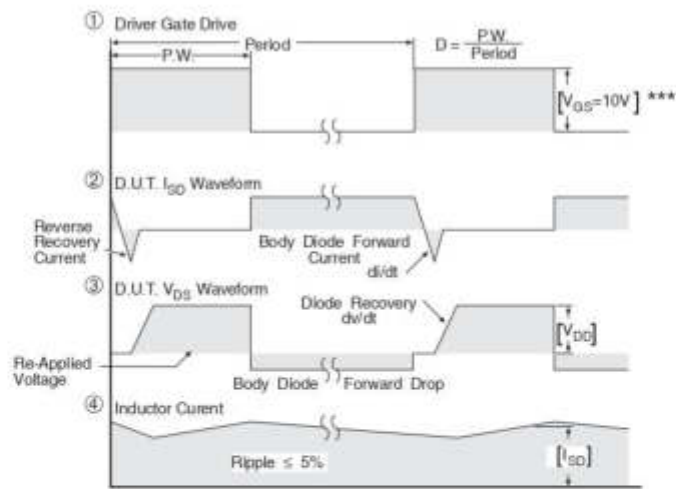
*** V_{GS} = 5.0V for Logic Level and 3V Drive Devices

Fig 14. For N-channel HEXFET® power MOSFETs

430

IRFZ44NPbF

International
IGR Rectifier

TO-220AB Package Outline (Dimensions are shown in millimeters (inches))
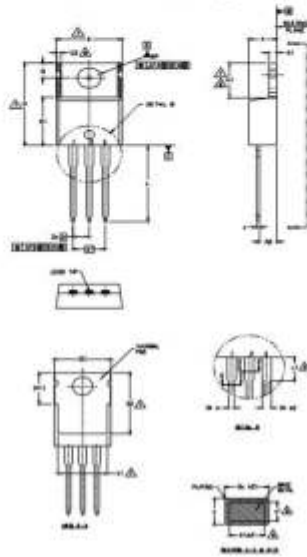
NOTES:
1.- DIMENSIONING AND TOLERANCING AS PER ASME Y14.5 M- 1994.
2.- DIMENSIONS ARE SHOWN IN INCHES [MILLIMETERS].
3.- LEAD DIMENSION AND FINISH UNCONTROLLED IN L1.
4.- DIMENSION D, D1 & E DO NOT INCLUDE MOLD FLASH. MOLD FLASH SHALL NOT EXCEED .005" (0.127) PER SIDE. THESE DIMENSIONS ARE MEASURED AT THE OUTERMOST EXTREMES OF THE PLASTIC BODY.
5.- DIMENSION b1, b3 & c1 APPLY TO BASE METAL ONLY.
6.- CONTROLLING DIMENSION : INCHES.
7.- THERMAL PAD CONTOUR OPTIONAL WITHIN DIMENSIONS C1+1.02 & E1
8.- DIMENSION E2 X H1 DEFINE A ZONE WHERE STAMPING AND SINGULATOR IRREGULARITIES ARE ALLOWED.
9.- OUTLINE CONFORMS TO JEDEC TO-220, EXCEPT A2 (max.) AND D2 (min.) WHERE DIMENSIONS ARE DERIVED FROM THE ACTUAL PACKAGE OUTLINE.

SYMBOL	MILLIMETERS		INCHES		NOTES
	MIN.	MAX.	MIN.	MAX.	
A	3.56	4.83	.140	.190	
A1	0.51	1.40	.020	.055	
A2	2.03	2.92	.080	.115	
b	0.58	1.01	.015	.040	
b1	0.38	0.97	.015	.038	5
b2	1.14	1.78	.045	.070	
b3	1.14	1.73	.045	.068	5
c	0.36	0.61	.014	.024	
c1	0.36	0.56	.014	.022	5
D	14.22	16.51	.560	.650	4
D1	8.38	9.02	.330	.355	
D2	11.68	12.88	.460	.507	7
E	9.65	10.67	.380	.420	4,7
E1	6.86	8.89	.270	.350	7
E2	-	0.76	-	.030	8
e	2.54 BSC		.100 BSC		
e1	5.08 BSC		.200 BSC		
H1	5.84	6.86	.230	.270	7,8
L	12.70	14.73	.500	.580	
L1	-	6.35	-	.250	3
øP	3.54	4.08	.139	.161	
Q	2.54	3.42	.100	.135	

LEAD ASSIGNMENTS

HEXFET
1- GATE
2- DRAIN
3- SOURCE

IGBTs, CoPACK
1- GATE
2- COLLECTOR
3- EMITTER

DIODES
1- ANODE/OPEN
2- CATHODE
3- ANODE

TO-220AB Part Marking Information

EXAMPLE: THIS IS AN IRF1010
LOT CODE 1789
ASSEMBLED ON WW 19, 2000
IN THE ASSEMBLY LINE "C"

Note "P" in assembly line position indicates "Lead - Free"

INTERNATIONAL RECTIFIER LOGO
PART NUMBER
DATE CODE YEAR 0 = 2000 WEEK 19 LINE C
ASSEMBLY LOT CODE
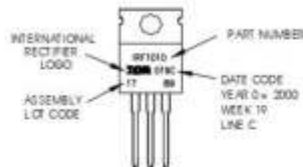
Notes:
1. For an Automotive Qualified version of this part please see http://www.irf.com/product-info/auto/
2. For the most current drawing please refer to IR website at http://www.irf.com/package/

Data and specifications subject to change without notice.
This product has been designed and qualified for the Industrial market.
Qualification Standards can be found on IR's Web site.

International
IGR Rectifier

IR WORLD HEADQUARTERS: 233 Kansas St., El Segundo, California 90245, USA Tel: (310) 252-7105
TAC Fax: (310) 252-7903
Visit us at www.irf.com for sales contact information. 09/2010
www.irf.com

8

431

[43]   "IR2110 data sheet," International Rectifier, California, USA.

International
IƠR Rectifier

IR2110(S)PbF/IR2113(S)PbF

HIGH AND LOW SIDE DRIVER

Features

- Floating channel designed for bootstrap operation
  Fully operational to +500V or +600V
  Tolerant to negative transient voltage
  dV/dt immune
- Gate drive supply range from 10 to 20V
- Undervoltage lockout for both channels
- 3.3V logic compatible
  Separate logic supply range from 3.3V to 20V
  Logic and power ground ±5V offset
- CMOS Schmitt-triggered inputs with pull-down
- Cycle by cycle edge-triggered shutdown logic
- Matched propagation delay for both channels
- Outputs in phase with inputs

Product Summary

V_{OFFSET} (IR2110)	500V max.	
(IR2113)	600V max.	
I_O+/-	2A / 2A	
V_{OUT}	10 - 20V	
$t_{on/off}$ (typ.)	120 & 94 ns	
Delay Matching (IR2110)	10 ns max.	
(IR2113)	20ns max.	

Packages

14-Lead PDIP
IR2110/IR2113

16-Lead SOIC
IR2110S/IR2113S

Description

The IR2110/IR2113 are high voltage, high speed power MOSFET and
IGBT drivers with independent high and low side referenced output chan-
nels. Proprietary HVIC and latch immune CMOS technologies enable
ruggedized monolithic construction. Logic inputs are compatible with
standard CMOS or LSTTL output, down to 3.3V logic. The output
drivers feature a high pulse current buffer stage designed for minimum
driver cross-conduction. Propagation delays are matched to simplify use in high frequency applications. The
floating channel can be used to drive an N-channel power MOSFET or IGBT in the high side configuration which
operates up to 500 or 600 volts.

Typical Connection

up to 500V or 600V

V_{DD} V_{DD} V_B
HIN   HIN V_S TO LOAD
SD   SD
LIN   LIN V_{CC}
V_{SS} V_{SS} COM
V_{CC} HO   LO

(Refer to Lead Assignments for correct pin configuration). This/These diagram(s) show electrical
connections only. Please refer to our Application Notes and DesignTips for proper circuit board layout.

www.infineon.com/gatedriver   1
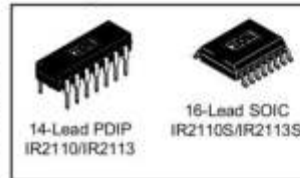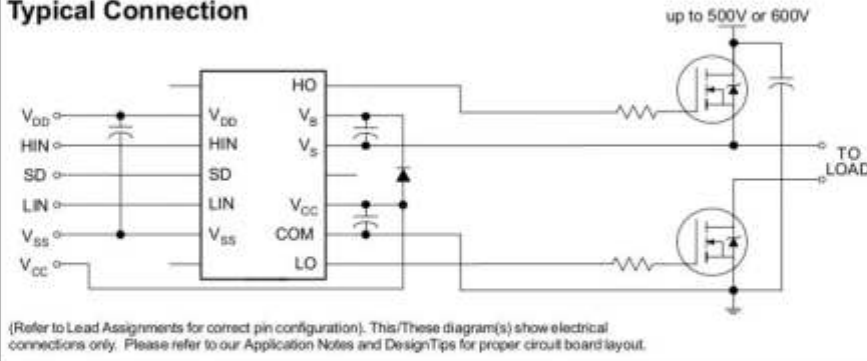
Absolute Maximum Ratings

Absolute maximum ratings indicate sustained limits beyond which damage to the device may occur. All voltage parameters are absolute voltages referenced to COM. The thermal resistance and power dissipation ratings are measured under board mounted and still air conditions. Additional information is shown in Figures 28 through 35.

Symbol	Definition		Min.	Max.	Units
V_B	High side floating supply voltage (IR2110)		-0.3	525	V
	(IR2113)		-0.3	625	
V_S	High side floating supply offset voltage		V_B - 25	V_B + 0.3	
V_{HO}	High side floating output voltage		V_S - 0.3	V_B + 0.3	
V_{CC}	Low side fixed supply voltage		-0.3	25	
V_{LO}	Low side output voltage		-0.3	V_{CC} + 0.3	
V_{DD}	Logic supply voltage		-0.3	V_{SS} + 25	
V_{SS}	Logic supply offset voltage		V_{CC} - 25	V_{CC} + 0.3	
V_{IN}	Logic input voltage (HIN, LIN & SD)		V_{SS} - 0.3	V_{DD} + 0.3	
dV_s/dt	Allowable offset supply voltage transient (figure 2)		—	50	V/ns
P_D	Package power dissipation @ T_A ≤ +25°C	(14 lead DIP)	—	1.6	W
		(16 lead SOIC)	—	1.25	
R_{THJA}	Thermal resistance, junction to ambient	(14 lead DIP)	—	75	°C/W
		(16 lead SOIC)	—	100	
T_J	Junction temperature		—	150	°C
T_S	Storage temperature		-55	150	
T_L	Lead temperature (soldering, 10 seconds)		—	300	

Recommended Operating Conditions

The input/output logic timing diagram is shown in figure 1. For proper operation the device should be used within the recommended conditions. The V_S and V_{SS} offset ratings are tested with all supplies biased at 15V differential. Typical ratings at other bias conditions are shown in figures 36 and 37.

Symbol	Definition		Min.	Max.	Units
V_B	High side floating supply absolute voltage		V_S + 10	V_S + 20	V
V_S	High side floating supply offset voltage	(IR2110)	Note 1	500	
		(IR2113)	Note 1	600	
V_{HO}	High side floating output voltage		V_S	V_B	
V_{CC}	Low side fixed supply voltage		10	20	
V_{LO}	Low side output voltage		0	V_{CC}	
V_{DD}	Logic supply voltage		V_{SS} + 3	V_{SS} + 20	
V_{SS}	Logic supply offset voltage		-5 (Note 2)	5	
V_{IN}	Logic input voltage (HIN, LIN & SD)		V_{SS}	V_{DD}	
T_A	Ambient temperature		-40	125	°C

Note 1: Logic operational for V_S of -4 to +500V. Logic state held for V_S of -4V to -V_{BS}. (Please refer to the Design Tip DT97-3 for more details).
Note 2: When V_{DD} < 5V, the minimum V_{SS} offset is limited to -V_{DD}.

www.infineon.com/gatedriver

Dynamic Electrical Characteristics

V_{BIAS} (V_{CC}, V_{BS}, V_{DD}) = 15V, C_L = 1000 pF, T_A = 25°C and V_{SS} = COM unless otherwise specified. The dynamic electrical characteristics are measured using the test circuit shown in Figure 3.

Symbol	Definition		Figure	Min.	Typ.	Max.	Units	Test Conditions
t_{on}	Turn-on propagation delay		7	—	120	150		V_S = 0V
t_{off}	Turn-off propagation delay		8	—	94	125		V_S = 500V/600V
t_{sd}	Shutdown propagation delay		9	—	110	140	ns	V_S = 500V/600V
t_r	Turn-on rise time		10	—	25	35		
t_f	Turn-off fall time		11	—	17	25		
MT	Delay matching, HS & LS turn-on/off	(IR2110)	—	—	—	10		
		(IR2113)	—	—	—	20		

Static Electrical Characteristics

V_{BIAS} (V_{CC}, V_{BS}, V_{DD}) = 15V, T_A = 25°C and V_{SS} = COM unless otherwise specified. The V_{IN}, V_{TH} and I_{IN} parameters are referenced to V_{SS} and are applicable to all three logic input leads: HIN, LIN and SD. The V_O and I_O parameters are referenced to COM and are applicable to the respective output leads: HO or LO.

Symbol	Definition	Figure	Min.	Typ.	Max.	Units	Test Conditions
V_{IH}	Logic "1" input voltage	12	9.5	—	—		
V_{IL}	Logic "0" input voltage	13	—	—	6.0		
V_{OH}	High level output voltage, V_{BIAS} - V_O	14	—	—	1.2	V	I_O = 0A
V_{OL}	Low level output voltage, V_O	15	—	—	0.1		I_O = 0A
I_{LK}	Offset supply leakage current	16	—	—	50		V_B=V_S = 500V/600V
I_{QBS}	Quiescent V_{BS} supply current	17	—	125	230		V_{IN} = 0V or V_{DD}
I_{QCC}	Quiescent V_{CC} supply current	18	—	180	340	μA	V_{IN} = 0V or V_{DD}
I_{QDD}	Quiescent V_{DD} supply current	19	—	15	30		V_{IN} = 0V or V_{DD}
I_{IN+}	Logic "1" input bias current	20	—	20	40		V_{IN} = V_{DD}
I_{IN-}	Logic "0" input bias current	21	—	—	1.0		V_{IN} = 0V
V_{BSUV+}	V_{BS} supply undervoltage positive going threshold	22	7.5	8.6	9.7		
V_{BSUV-}	V_{BS} supply undervoltage negative going threshold	23	7.0	8.2	9.4		
V_{CCUV+}	V_{CC} supply undervoltage positive going threshold	24	7.4	8.5	9.6	V	
V_{CCUV-}	V_{CC} supply undervoltage negative going threshold	25	7.0	8.2	9.4		
I_{O+}	Output high short circuit pulsed current	26	2.0	2.5	—		V_O = 0V, V_{IN} = V_{DD} PW ≤ 10 μs
I_{O-}	Output low short circuit pulsed current	27	2.0	2.5	—	A	V_O = 15V, V_{IN} = 0V PW ≤ 10 μs

Functional Block Diagram
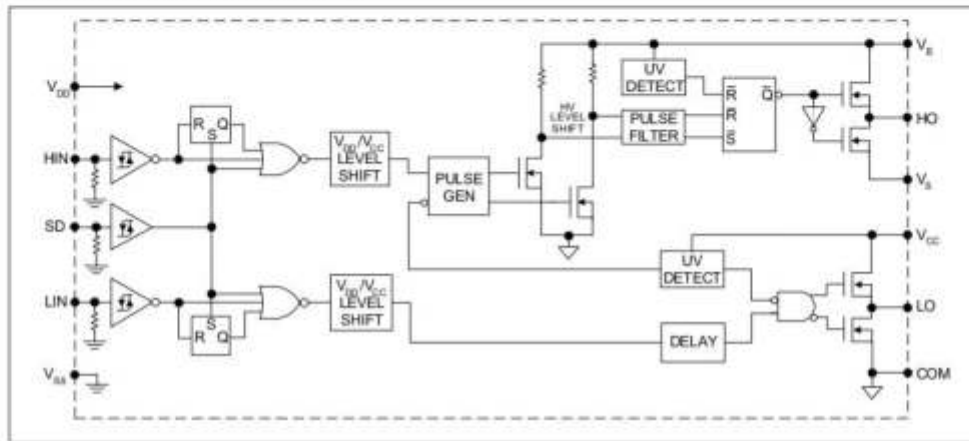
Lead Definitions

Symbol	Description
V_{DD}	Logic supply
HIN	Logic input for high side gate driver output (HO), in phase
SD	Logic input for shutdown
LIN	Logic input for low side gate driver output (LO), in phase
V_{SS}	Logic ground
V_B	High side floating supply
HO	High side gate drive output
V_S	High side floating supply return
V_{CC}	Low side supply
LO	Low side gate drive output
COM	Low side return

[44]   Iwamoto, Y., Noma, K., Nakayama, O., Yamauchi, T., Ando, H., "Development of Gasoline Direct Injection Engine," SAE Technical Paper 970541, DOI: https://doi.org/10.4271/970541
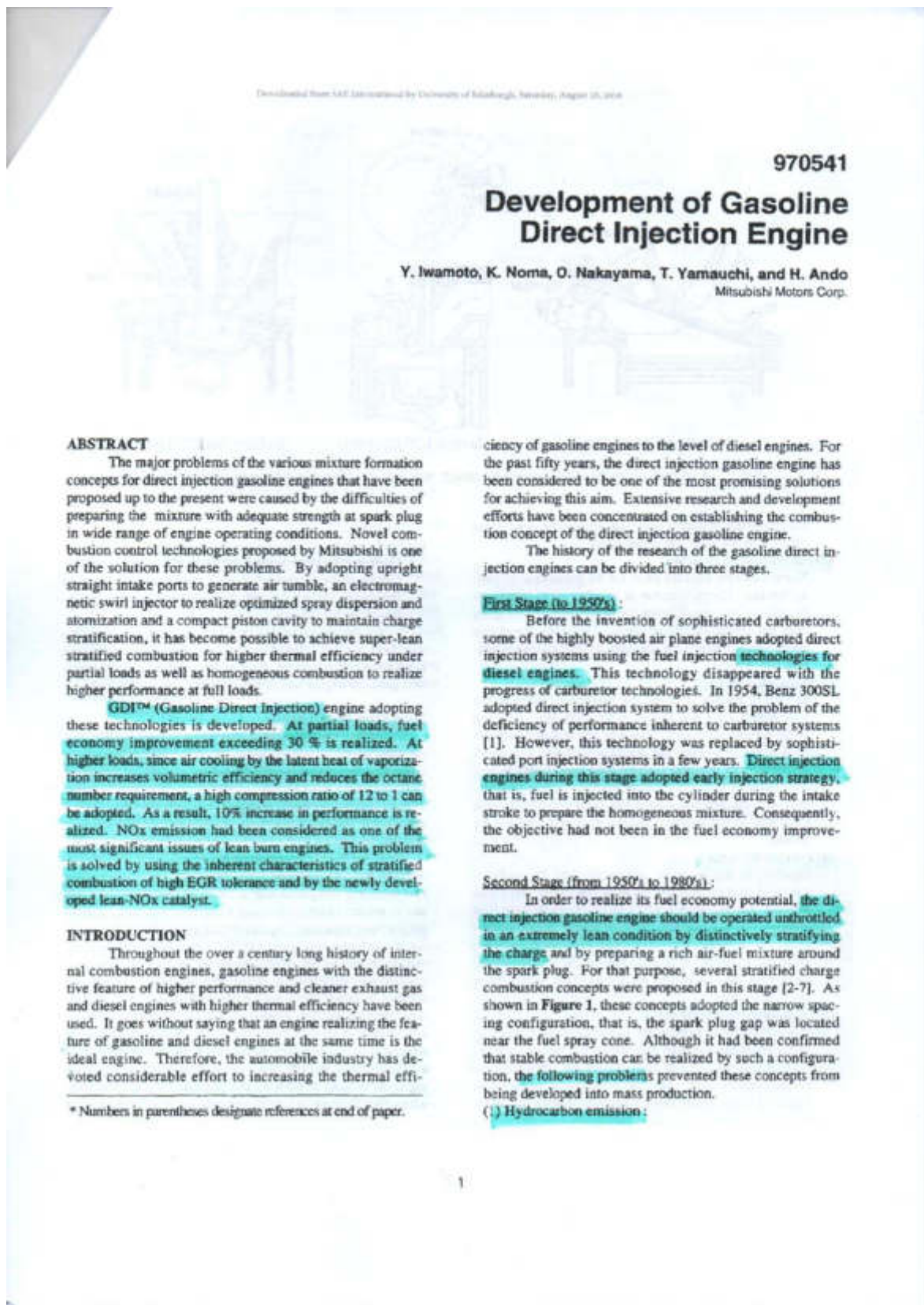
970541

Development of Gasoline Direct Injection Engine

Y. Iwamoto, K. Noma, O. Nakayama, T. Yamauchi, and H. Ando
Mitsubishi Motors Corp.

ABSTRACT

The major problems of the various mixture formation concepts for direct injection gasoline engines that have been proposed up to the present were caused by the difficulties of preparing the mixture with adequate strength at spark plug in wide range of engine operating conditions. Novel combustion control technologies proposed by Mitsubishi is one of the solution for these problems. By adopting upright straight intake ports to generate air tumble, an electromagnetic swirl injector to realize optimized spray dispersion and atomization and a compact piston cavity to maintain charge stratification, it has become possible to achieve super-lean stratified combustion for higher thermal efficiency under partial loads as well as homogeneous combustion to realize higher performance at full loads.

GDI™ (Gasoline Direct Injection) engine adopting these technologies is developed. At partial loads, fuel economy improvement exceeding 30 % is realized. At higher loads, since air cooling by the latent heat of vaporization increases volumetric efficiency and reduces the octane number requirement, a high compression ratio of 12 to 1 can be adopted. As a result, 10% increase in performance is realized. NOx emission had been considered as one of the most significant issues of lean burn engines. This problem is solved by using the inherent characteristics of stratified combustion of high EGR tolerance and by the newly developed lean-NOx catalyst.

INTRODUCTION

Throughout the over a century long history of internal combustion engines, gasoline engines with the distinctive feature of higher performance and cleaner exhaust gas and diesel engines with higher thermal efficiency have been used. It goes without saying that an engine realizing the feature of gasoline and diesel engines at the same time is the ideal engine. Therefore, the automobile industry has devoted considerable effort to increasing the thermal effi-

* Numbers in parentheses designate references at end of paper.

ciency of gasoline engines to the level of diesel engines. For the past fifty years, the direct injection gasoline engine has been considered to be one of the most promising solutions for achieving this aim. Extensive research and development efforts have been concentrated on establishing the combustion concept of the direct injection gasoline engine.

The history of the research of the gasoline direct injection engines can be divided into three stages.

First Stage (to 1950's) :

Before the invention of sophisticated carburetors, some of the highly boosted air plane engines adopted direct injection systems using the fuel injection technologies for diesel engines. This technology disappeared with the progress of carburetor technologies. In 1954, Benz 300SL adopted direct injection system to solve the problem of the deficiency of performance inherent to carburetor systems [1]. However, this technology was replaced by sophisticated port injection systems in a few years. Direct injection engines during this stage adopted early injection strategy, that is, fuel is injected into the cylinder during the intake stroke to prepare the homogeneous mixture. Consequently, the objective had not been in the fuel economy improvement.

Second Stage (from 1950's to 1980's) :

In order to realize its fuel economy potential, the direct injection gasoline engine should be operated unthrottled in an extremely lean condition by distinctively stratifying the charge and by preparing a rich air-fuel mixture around the spark plug. For that purpose, several stratified charge combustion concepts were proposed in this stage [2-7]. As shown in Figure 1, these concepts adopted the narrow spacing configuration, that is, the spark plug gap was located near the fuel spray cone. Although it had been confirmed that stable combustion can be realized by such a configuration, the following problems prevented these concepts from being developed into mass production.

(1) Hydrocarbon emission :

1

(a) Benz 300SL [1] (1954)   (b) Texaco TCP [2] (1951)   (c) Ford PROCO [3] (1968)
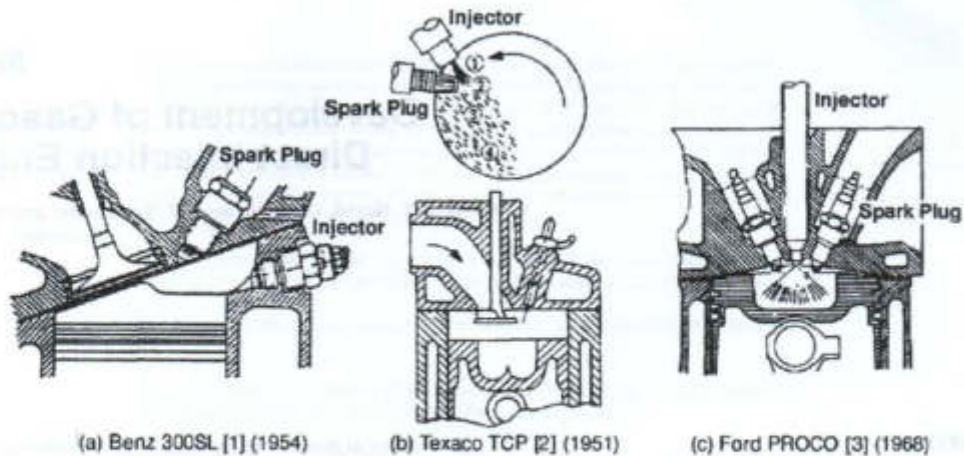
Fig.1 Example of combustion system for direct injection gasoline engines in the earlier stages

Large amount of hydrocarbon was emitted because it was difficult to complete the combustion.

(2) Limited operation zone :

These concepts adopted swirl for the promotion of fuel-air mixing. The momentum of swirling air is in proportional to engine speed, while the momentum of fuel spray does not depend on engine speed. Therefore, the engine speed range in which the adequate air-fuel mixing is realized was limited.

(3) Spark plug fouling :

In case of narrow spacing, the liquid fuel spray or the over-rich mixture was located at the spark plug, resulting in the formation of soot. Soot was accumulated between plug gap and caused the ignition fouling. Although higher energy ignition systems have sometimes shown the effects of reduced ignition fouling, problem associated with the deterioration of the durability of the spark plug due to the higher ignition energy could not be solved.

(4) Poor performance :

Injection timing variation range realized by the mechanical fuel injection equipment employed in this stage was limited, and the switching from late to early injection strategy was not possible. Consequently, these engines should be operated with the stratified charge even in the highest load range. In order to prevent the soot emission, air-excess ratio should be maintained high, resulting in poor performance.

(5) Dilution of lubricating oil :

It was difficult to prevent the liquid gasoline droplets to impinge on the cylinder liner or on the piston surface. Gasoline on the cylinder liner diluted the lubricating oil on the cylinder liner. Gasoline on the piston surface was captured in the piston crevice and also diluted the lubri-cating oil.

(6) Soot emission and deposit accumulation :

Over-rich mixture around the spark plug caused the formation and emission of soot. Liquid fuel film layer on the piston surface caused the accumulation of combustion chamber deposit.

By the comprehensive studies performed by many researchers, the difficulties to solve these problems had become a common understanding. Research activities on direct injection gasoline engine had fallen off at the end of this stage.

Third Stage (from 1990's) :

At this stage, fuel economy improvement has become the most important subject for automobile industry, because it is the key factor for the energy saving and the reduction of CO_2, one of the most harmful green house effect gas. In order to meet this requirement, energetic research activities started to establish the direct injection gasoline engine technologies that can be applied to the practical engines in the real world.

It has been a common target of research activities to develop a direct injection gasoline engine realizing greater fuel economy compared with a diesel engine at partial loads and to realize better performance than the conventional MPI (Multi Point Injection) engines at high loads. In order to realize its fuel economy potential, the direct injection gasoline engine should be operated unthrottled in an extremely lean condition by distinctively stratifying the charge. In order to achieve its higher performance potential at high loads, the direct injection gasoline engine should be operated under stoichiometric or slightly rich conditions. When the charge is stratified, soot is generated in the rich zone. Sufficient excess air should be provided around the combustion zone containing soot, in order to burn-up the generated soot. Therefore, when the average mixture strength is stoichiometric or

2

437
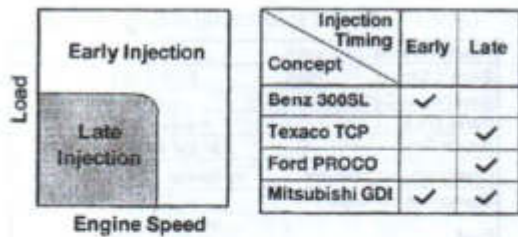
Fig.2 Operation modes of direct injection engines

slightly rich, that is, when the equivalence ratio is larger than unity, the mixture should be homogeneous so as to suppress the soot formation. Accordingly, as shown in **Figure 2**, the direct injection gasoline engines at this stage should realize the following goals :

(1) stable and distinctive stratification by late injection mode at partial load,
(2) completely homogeneous charge by early injection mode at higher load,
(3) rapid and smooth switching of late and early injection modes.

Comprehensive research has been carried out in the laboratories of Mitsubishi Motors Corporation. It was found that these goals could be realized by a novel mixture preparation concept. Mitsubishi Motors Corporation adopted this concept for its GDI™ (Gasoline Direct Injection) engine, and started the mass production in August of 1996.

In this paper, the basic concept, measures for realizing the concept and the characteristics of Mitsubishi GDI engine will be described.

BASIC STRATEGIES OF MITSUBISHI GDI [8, 9]

The novel mixture formation concept is illustrated in **Figure 3**. In place of narrow spacing layout, it adopts wide spacing layout. Fuel spray is not directed toward the spark plug. It is directed to the piston surface and after impinging on the spherical piston cavity, it is reflected toward the spark plug. By adopting this layout, the interval between the end of injection and the spark ignition long enough for promoting the fuel vaporization and the mixing with the surrounding air can be realized. Consequently, problems of the previous stage caused by the liquid fuel or the over-rich mixture around the spark plug can be solved. The principal factor controlling the mixing is the fuel spray or gaseous mixture reflection on the cavity wall which is subject to the fuel spray momentum. Unlike the methods controlling the mixing by swirl, it is hardly affected by the engine speed. This guarantees the adequate mixing in wide speed range.

Fundamental technologies developed to realize the adequate mixing control by wide spacing layout are illustrated in **Figure 4**. The following new technologies are employed :

(1) upright straight intake ports generating an intense reverse tumble, that is, a tumble with a rotational direction oppo-

Fig.3 Wide spacing concept adopted by Mitsubishi GDI

(a) Reverse Tumble   (c) Swirling Spray

Fig.4 Base technologies for Mitsubishi GDI

3

438

[45]   Fu-Quan, Z., Ming-Chia, L., "A Review of Mixture Preparation and Combustion Control Strategies for Spark-Ignited Direct-Injection Gasoline Engines," SAE Technical Paper 970627, DOI: https://doi.org/10.4271/970627

970627

A Review of Mixture Preparation and Combustion Control Strategies for Spark-Ignited Direct-Injection Gasoline Engines

Fu-Quan Zhao and Ming-Chia Lai
Wayne State University

David L. Harrington
General Motors Research and Development Center

ABSTRACT

The current extensive revisitation of the application of gasoline direct-injection to automotive, four-stroke, spark-ignition engines has been prompted by the availability of technological capabilities that did not exist in the late 1970s, and that can now be utilized in the engine development process. The availability of new engine hardware that permits an enhanced level of computer control and dynamic optimization has alleviated many of the system limitations that were encountered in the time period from 1976 to 1984, when the capabilities of direct-injection, stratified-charge, spark-ignition engines were thoroughly researched. This paper incorporates a critical review of the current worldwide research and development activities in the gasoline direct-injection field, and provides insight into new areas of technology that are being applied to the development of both production and prototype engines.

The advantages and disadvantages of the emerging technologies that are being utilized to develop gasoline direct-injection combustion systems, such as electronic common-rail injection, variable swirl and tumble control, computational fluid dynamics, and laser diagnostics of fuel sprays and combustion are discussed in detail within the context of the total industry effort to enhance the attainable brake specific fuel consumption and emissions. Significant emphasis is placed upon extracting the key findings of each investigation and discussing the relative potential of the technique for meeting the major requirements of the gasoline direct-injection engine. Areas of general consensus that are evident among the hundreds of researchers who are currently investigating this field are noted, as are apparent conflicts that must be resolved in the near future by obtaining additional data. Also noted and discussed are a number of critical research areas for which it is obvious that insufficient effort and resources are being devoted. One example of such an area is the development of a proven laboratory test for quantifying the deposit formation tendencies of direct-injection gasoline injectors.

Based upon the review of all available technical publications, the trends and directions of the current effort on gasoline direct-injection engine are noted, and the remaining difficult developmental areas are discussed. For planning purposes, the totality of the literature is used to formulate likely engine scenarios regarding the development

of specific engine configurations, and to outline technological path options and trends that may be expected in the displacement of the port fuel injection engine by the gasoline direct-injection engine.

INTRODUCTION

The injection of gasoline directly into the cylinder of a four-stroke, spark-ignition engine is a concept that offers a number of unique potential advantages over even the most sophisticated current production engines that inject fuel into the intake port. A comparison of the basic gasoline direct-injection (GDI) and port-fuel-injected (PFI) combustion systems is illustrated schematically in Fig. 1. It may be seen that the major difference between the PFI engine and the GDI engine is the additional mixture preparation stage comprised of a port film of liquid fuel. The thermodynamic potential for reducing the specific fuel consumption, coupled with the advantages of quicker starting, enhanced transient response, and more precise control of the mixture air/fuel ratio, have resulted in a renewed worldwide interest in the GDI engine as an automotive powerplant. The primary developments since the extensive examination of these engines during the 1976-1984 time period include the replacement of pressure-activated poppet injectors by the electronic fuel injector, the incorporation of the common-rail injection system, and the ready availability of extensive new capabilities for computer-controlled optimization of the in-cylinder flow field, cam-phasing, throttle positioning and exhaust-gas-recirculation (EGR).

There are two current strategic objectives in the continued long-range development of the four-stroke, gasoline, spark-ignition engine for automotive applications; the first is associated with fuel economy and the second deals with emissions. The fuel economy objective is the attainment of a step-change improvement in the level of brake specific fuel consumption (BSFC) that is provided by current spark-ignition engines. The emission objective is to significantly reduce the levels of unburned-hydrocarbon (UBHC) and NOx emissions to production-feasible mean levels for ultra-low-emission-vehicle (ULEV) regulatory compliance. There can be varying degrees of design complexity associated with direct-injection gasoline engines. Perhaps the simplest embodiment of a GDI system is an engine that

1

- excessive rate of formation of injector deposits and/or ignition fouling
- excessive light-load UBHC emissions
- excessive high-load NOx emissions
- high local NOx production under part-load, stratified-charge operation
- three-way catalysts cannot be utilized to full advantage
- increased fuel system wear due to the combination of high fuel pressure and low fuel lubricity
- increased rates of cylinder bore wear
- soot formation for high-load operation
- increased electrical power and voltage requirements of the injectors and drivers.

If ULEV emissions regulations and corporate-averaged-fuel-economy (CAFE) requirements can be achieved using PFI engines without the requirement of complex and expensive new hardware, the market penetration rate for GDI engines will be reduced, as the GDI engine will require relatively sophisticated fuel injection hardware and control systems. Two key technical limitations for GDI engines have been excessive UBHC emissions and the fact that three-way catalysts could not be used. Operating the engine under lean conditions reduces the engine-out NOx emissions, but this generally cannot achieve the 90% reduction level that can be attained using a three-way catalyst. Much work is underway worldwide to develop lean-NOx catalysts, but at this time the attainable conversion efficiency is still much less than that of the three-way catalyst. The excessive UBHC emissions at light load represent a significant research problem to be solved.

In spite of very real concerns and difficulties, the GDI engine offers an expanded new horizon for future applications as compared to the well-developed PFI engine. New technologies and computer-control strategies are currently being invoked by a number of automotive companies to re-examine the extent to which the potential benefits of the GDI engine can be realized in a production engine (Kume et al., 1996; Harada et al., 1997; Anderson et al., 1996; Takagi, 1996; Shimotani et al., 1996; Karl et al., 1997; Lake et al., 1996; Buchheim and Quissek, 1996; Douaud, 1996; Fansler et al., 1995; Hauser et al., 1995; Seiffert, 1996). Mitsubishi has very recently introduced a production GDI engine into the Japanese market, as has Toyota, and Mitsubishi plans to introduce a version of their GDI engine into the European market in 1997 (Ando, 1996b).

The information in this paper will provide the reader with a comprehensive review of the mixture dynamics and combustion control strategies that may be utilized in four-stroke, spark-ignition, direct-injection, gasoline engines. The current state of knowledge, as exhibited in more than one hundred recent key publications, many as yet untranslated, is discussed in detail, and the critical research and development needs for the near future are identified.

FUEL SYSTEMS FOR GASOLINE DIRECT-INJECTION

Unthrottled operation with the load controlled by the fuel quantity has been shown to be a very efficient operation mode for the internal combustion engine, as the pumping loss is significantly reduced and the volumetric efficiency is increased. This combustion approach is very successful in the diesel engine, as ignition occurs spontaneously at points within the combustion chamber where the mixture is well prepared for autoignition. The fixed location of the ignition source in the spark-ignition (SI) engine, however, makes it quite difficult to operate in the unthrottled mode for other than full load. This imposes a critical additional requirement on the mixture formation process of this type of engine in that the mixture cloud that results from fuel vaporization and mixing must be controlled both spatially and temporally in order to obtain stable combustion. Preparing the required mixture distribution inside the combustion chamber for a wide range of engine operating conditions is quite difficult, as the fuel/air mixing process is influenced by many time-dependent variables. The development of a successful combustion system depends upon the optimized design of the fuel injection system and the proper matching of the system components to control the in-cylinder flow field and burn rate.

The fuel injection systems of early DISC engines were derived from the basic diesel injection system (Baranescu, 1983; Duggal et al., 1984; Enright et al., 1988; Iida, 1992; Wood, 1978). For example, the Texaco TCCS engine (Alperstein et al., 1974) utilized a diesel-type injector that produced a spray with relatively poor atomization and fuel/air mixing quality, and with high penetration rates relative to sprays from current pressure-swirl atomizers. The Ford PROCO engine (Scussel et al., 1978) used an outwardly opening pintle atomizer with vibration to enhance the fuel atomization. However, the poppet opening pressure was on the order of 2 MPa, which is quite low. Gasoline injection using single-hole or multihole, narrow-angle sprays from diesel injection systems nearly always result in substantial UBHC emissions due to the compactness of the spray and the high penetration velocity. In order to avoid this problem, combustion systems using direct wall impingement and fuel film formation such as the MAN-FM (Urlaub and Chmela, 1974) were developed. For early DISC engine experiments using multi-hole injectors, the singularity of the ignition source proved to be a definite problem, as compared to a diesel combustion system where multiple ignition sites occur simultaneously. Another severe problem for the diesel-based DISC injection system was the lack of variability of injection characteristics between part load and full load. For full-power DISC engine operation, injection of maximum fuel quantities early in the intake stroke resulted in significant wall impingement of the fuel.

In recent years, significant progress has been made in the development of advanced, computer-controlled fuel injection systems, which has had much to do with the expansion of research and development activities related to GDI engines. In this section, the key components that are utilized in the fuel system, and their important roles in the mixture preparation and combustion processes, are described in detail.

General Fuel System Requirements . The fuel injection system in a GDI engine is a key component that must be matched with the specific in-cylinder air flow field to provide the desired mixture cloud over the entire operating range of the engine. For all operating conditions a well-atomized fuel spray must be produced, and, for the efficient combustion of a stratified mixture, a stable and compact spray geometry is necessary (Kume et al., 1996). A GDI fuel system needs to provide for at least two, and possibly three, distinct operating modes. For unthrottled, part-load operation, the injection system should provide the capability for late
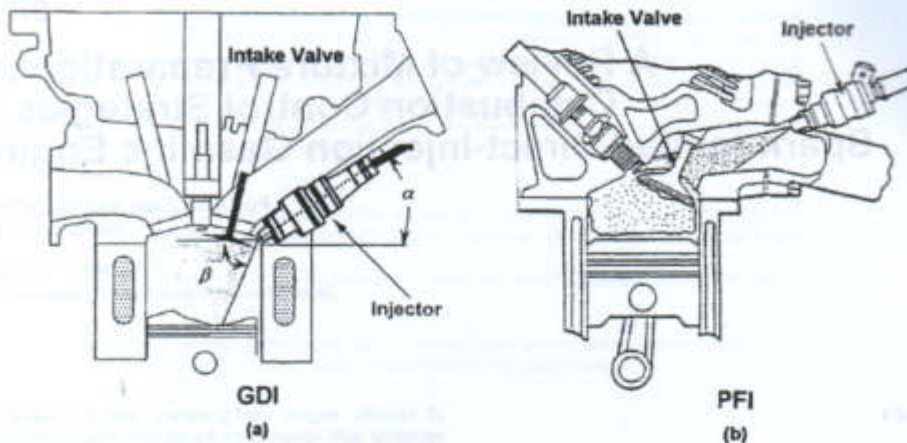
3

Fig. 1 Comparison of GDI and PFI mixture preparation systems; (a) GDI engine
(Zhao et al., 1997a); (b) PFI engine (Ohyama et al., 1992)

operates in the early-injection, homogeneous, stoichio-metric mode. This engine must utilize throttling for load control, thus the full fuel-economy potential that is associated with the elimination of throttling will not be realized. Engines that operate exclusively in this mode do, however, experience the thermodynamic advantages of intake charge cooling, as well as the advantages of more rapid starting and fuel cutoff on deceleration. The next level of complexity in GDI engine design is to utilize leaner homogeneous mixtures with reduced throttling for some degree of load control, which permits the realization of additional potentials. Making use of the full potential of the GDI concept requires a design that generates a stable, stratified mixture using an overall-lean air/fuel ratio for operating at part load, and which can transition smoothly to heavy or full-load operation by injecting increased volumes of fuel at earlier times in the cycle. The most complex GDI engine would operate in all three modes at various times, with transitions from one mode to another. This would require not only a complex control system, but an array of fast and reliable sensors and a sophisticated set of algorithms in order to maintain acceptable levels of drive-ability during mode transitions.

The PFI engine has evolved since 1983, with many incremental improvements having been incorporated on an annual basis. Enhancements such as sequentially-timed injection, computer algorithms for transient metering, four-valves per cylinder, multiple roller camshafts, variable cam phasing, turbocharging, and even supercharging have become features on PFI engines. In spite of these signifi-cant developments, the current high-tech PFI engine still requires, and will continue to require, throttling for basic load control, and still has, and will continue to have, an operating film of liquid fuel in the intake port. These two basic PFI operating requirements represent major impedi-ments to achieving significant breakthroughs in PFI fuel economy or emissions. Continuous incremental improve-ments in the older PFI technology will be made, but it is unlikely that the long-range fuel economy and emission objectives can be simultaneously achieved.

Even though throttling is a well-established and reliable mechanism of load control in the PFI engine, the thermo-dynamic loss associated with throttling is very substantial. Any system that utilizes this method to adjust load levels will experience the inherent loss, and will have significant efficiency limitations at low levels of engine load. The GDI engine, in theory, does not have these two significant limitations, nor the performance boundaries that are associated with them. The theoretical advantages of the GDI engine over the current PFI engine are summarized as follows :

- improved BSFC (up to 30% improvement)
  ◊ less pumping loss (unthrottled, stratified mode)
  ◊ less heat losses (unthrottled, stratified mode)
  ◊ higher compression ratio (charge cooling with injection during induction)
  ◊ lower octane requirement (charge cooling with injection during induction)
  ◊ increased volumetric efficiency (charge cooling with injection during induction)
  ◊ fuel cutoff during a deceleration
- improved transient response
  ◊ less acceleration-enrichment required
- more rapid starting with less cold enrichment required
  ◊ starting on first or second cranking cycle
- selective emission advantages
  ◊ reduced cold-start UBHC emissions
  ◊ reduced CO_2 emissions
  ◊ more precise air/fuel ratio control.

In spite of the important potential advantages noted above, the practical GDI engine does have certain inherent problems that are similar to those of the direct-injection stratified-charge (DISC) engines. The commercial success of GDI as an automotive production powerplant has been delayed by the following areas of concern :

- difficulty in controlling the stratified-charge combustion over the required operating range
- complexity of control and injection technologies required for stepless load changes

2

injection at high rates during the compression stroke into a cylinder pressure of up to 2 MPa, which requires a relatively high fuel injection pressure. The injection pressure is also very important for obtaining both effective spray atomization and the required level of spray penetration. A higher fuel injection pressure is effective in reducing the mean droplet diameter of the spray approximately as the inverse square root of the pressure differential, whereas the use of a lower pressure generally reduces the pump parasitic load and injector noise, and increases the reliability of the fuel pump system. The use of a very high fuel injection pressure, such as 20 MPa, will enhance the atomization but will most likely generate an overpenetrating spray, resulting in fuel wall wetting. The fuel pressures that have been selected for most of the current prototype GDI engines range from 4 to 7 MPa, which are quite low when compared with diesel injection systems of 50 to 140 MPa, but are relatively high in comparison with typical PFI injection pressures of 270 to 450 kPa. A constant fuel-line pressure is utilized in a common-rail system for most of the current GDI applications; however, a strategy using a variable fuel injection pressure does offer an alternative method of obtaining the required flow range while reducing the dynamic-range requirements of the injector itself (Pontoppidan et al., 1996) and for meeting differing fuel spray requirements corresponding to a range of engine loads (Matsushita et al., 1996).

Fuel injection systems for full-feature GDI engines must have the capability of providing both late injection for stratified-charge combustion at part load, as well as early injection for homogeneous-charge combustion at full load. At part load, a well-atomized compact spray or mixture plume is desirable to achieve rapid mixture formation and controlled stratification. At full load, a well-dispersed fuel spray or mixture plume is desirable to ensure a homogeneous charge even for the largest fuel quantities. This is generally achieved by early injection at low cylinder pressure, similar to the mode of open-valve fuel injection in the PFI engine. These fuel system requirements are more comprehensive than those of either the diesel or the PFI injection systems. Based upon recent experimental investigations, it has been concluded that common-rail injection systems with electromagnetically-activated injectors can meet these requirements (Pischinger and Walzer, 1996; Buchheim and Quissek, 1996). In principle, these systems are comparable to those utilized in current PFI and advanced diesel engines, with the GDI application having an intermediate fuel pressure level, high-frequency noise level, and associated cost. A subdivided or split injection during the intake and compression strokes may also be a viable option. In the Toyota D-4 GDI combustion system (Matsushita et al., 1996) a two-stage injection strategy is utilized to improve the transition between part-load and full-load operation.

During the engine cold crank and start, the high pressure fuel pump generally cannot deliver fuel at sufficient pressure due to the short time available and the low engine cranking speed. A bypass valve is used in the Mitsubishi GDI engine to allow the fuel to bypass the high-pressure regulator under such a condition. As a result, the electric feed pump supplies the fuel directly to the fuel rail at a fuel pressure of 335 kPa. After the engine speed increases and the high pressure pump fully primes the system, the bypass valve is closed and the high pressure regulator begins to regulate

the fuel pressure to 5.0 MPa.

Gasoline has a lower lubricity, viscosity, and a higher volatility than diesel fuel, generally resulting in more concerns regarding system friction and wear, greater potential leakage and the need for enhanced cooling in fuel pumps and injectors. It should be noted, however, that hydrodynamic lubrication may be used at high fuel pressures to compensate for the low viscosity (Iwamoto et al., 1997). A very wide range of fuel quality is available in the field, thus it is important that a robust GDI fuel system be developed and proven for production engines.

Fuel Injector Requirements - The fuel injector may be said to be the most critical element in the GDI fuel system, and should have the following attributes. Many of the required characteristics of the GDI injector also correspond to those of the port fuel injector, which are (Zhao et al., 1995) :

- accurate fuel metering
- desirable spray pattern for the application
- minimal spray skew
- good spray symmetry over the operating range
- minimal drippage and fuel leakage, particularly for cold operation
- small pulse-to-pulse variation in fuel quantity and spray characteristics
- small sac volume
- good low-end linearity between the dynamic flow and the fuel pulse width
- minimal variation in the above parameters from unit to unit.

n certain critical areas, the requirements of the GDI injector significantly exceed those of the port fuel injector. These requirements are :

- significantly improved level of atomization
- expanded dynamic range
- enhanced resistance to deposit formation
- ability to operate at higher injector body and tip temperatures
- stable operation at elevated fuel and ambient pressures
- avoidance of needle bounce that creates unwanted secondary injections.

The GDI injector should be designed to deliver a precisely metered fuel quantity with a symmetric and highly repeatable spray geometry, and must provide a highly atomized fuel spray having a Sauter mean diameter (SMD) of generally less than 25 μm, and with a droplet diameter corresponding to the 90% volume point (DV90) not exceeding 45 μm. The DV90 statistic is a quantitative measure of the largest droplets in the spray. Smaller values than these are even more beneficial, provided sufficient spray penetration is maintained for good air utilization. The SMD is also denoted by the symbol D32. The fuel pressure required is at least 4 MPa for a single-fluid injector, with 5 to 7 MPa being more desirable if the late-injection, stratified mode is to be invoked. Even if successful atomization could be achieved with fuel pressures less than 4 MPa, significant metering errors could result from the variation of metering pressure differential with cylinder pressure. In general, the smaller the injector sac volume, the fewer large peripheral drops that will be generated when the injector opens. The sac volume within the injector tip is basically a volume of fuel that is not at the fuel-line pressure, therefore it retards

4

the acceleration of the main portion of the injected fuel and degrades both the fuel atomization and the resulting combustion. Needle bounce is to be avoided, as a secondary injection generally results in uncontrolled atomization consisting of larger droplets of lower velocity. It also reduces the fuel metering accuracy and contributes to increases in the UBHC and particulate emissions.

The ability to deliver the required fuel with a short fuel pulse, which corresponds to a higher rate of injection, is much more important for the GDI engine than for the PFI engine, particularly for light-load stratified-charge operation. Therefore, much more significance is attached to the low-pulse-width region of the GDI injector, effectively increasing the importance of the injector dynamic range requirement. The optimal design of the injector to resist coking is also one of the important requirements of the GDI injector, as is discussed in the section on injector deposits. Often over-looked are the voltage and power requirements of the injector solenoids and drivers. A number of prototype GDI injectors have power requirements that would be unaccept-able for a production application. It is also worth noting that it is advantageous to injector packaging to have the body as small as possible. This provides more flexibility in optimizing the injector location and in sizing and locating the ports and valves.

In spite of decades of continuous development on diesel multi-hole injectors, it has been shown that these nozzle-type injectors are not readily adaptable to GDI applications. A multi-hole nozzle used in a GDI engine application generally results in an unstable flame kernel when ignited by a single fixed spark gap. The rich mixture zones are close to the lean mixture zones, thus the flame front does not propagate uniformly through the combustion chamber. With the multi-hole nozzle, the number of nozzle holes is an important factor in determining engine combustion perfor-mance. The hole distribution that is effective in ensuring good spray dispersion and reliable flame propagation between the sprays generally provides the best engine performance. The effect of the cone angle of individual spray plumes on engine performance was studied by Fujieda et al. (1995). A multi-hole GDI injector similar to that used in diesel engines was designed and tested. It was found that reducing the individual nozzle flow area and increasing the number of holes can extend engine lean limit.

Currently, the most widely utilized GDI injector is the needle-type, high-pressure, swirl-spray unit, which delivers a conical spray. This type of injector can be regarded as a multi-hole nozzle with an infinite number of holes, and a uniform distribution of the fuel over the cone circumference may be obtained. As a consequence, wall wetting at full load can be minimized for an appropriate injector position and an optimized spray cone angle (Fraidl et al., 1996). The needle-type, swirl-spray injector is designed to apply a strong rotational momentum to the fuel in the injector nozzle that is in addition to the axial momentum. In a number of nozzle designs, liquid flows through a series of tangential holes or slots into a swirl chamber. The liquid emerges from the discharge orifice as an annular sheet that spreads radially outward to form an initially hollow-cone spray. The initial spray cone angle ranges from a minimum of 25 to almost 180°, depending on the requirements of the application, with a delivered SMD ranging from 15 to 25 μm. In the swirl-type injector, the pressure energy is effectively transformed into rotational momentum, which

enhances atomization, but limits spray penetration. The swirl nozzle generally produces a spray having a narrower distribution of drop sizes (DV90-DV10) than is obtained with the standard hole-type nozzle, with the best atomization occurring at high delivery pressures and wide spray angles. An additional advantage is that the injector designer can customize the spray penetration by altering the swirl ratio with only small changes in the atomization level, thus providing the necessary variability of spray configuration in order to meet different stratification requirements. However, the surface roughness of the orifice wall tends to generate streams or fingers of fuel in the fuel sheet exiting the nozzle, resulting in the formation of a locally rich air/fuel mixture. In order to minimize such mixture inhomogeneity, precise control of the nozzle tip quality is required (Tomoda et al., 1997).

In regards to the relative advantages of inwardly-opening versus outwardly-opening needles, the inwardly-opening needle generally provides better pulse-to-pulse repeatability of the spray cone geometry, especially when a flow guide bushing is present at the needle tip, although the outwardly-opening geometry has enhanced leakage resistance (Schapertons et al., 1991; Pontoppidan et al., 1997).

Engineers at Mitsubishi successfully applied the pressure-swirl injector to meet the requirements of their GDI engine, with the swirl generated by a swirler tip located upstream of the injector hole. The effect of the swirl level on the droplet size and spray structure was studied extensively by Kume et al. (1996) and the results are shown in Fig. 2. It was found that if the swirl can be intensified, a comparatively lower fuel pressure will be sufficient for achieving an acceptable level of atomization. Enhancement of the swirl also promotes air entrainment, and a vortex ring initially generated near the injector tip grows to a large scale toroidal vortex during the last portions of the injection event. Droplet velocity measurements using a phase Doppler anemometry (PDA) showed that the axial velocity component decreases with distance from the injector tip, whereas the swirl component remains fairly constant. The decrease in the axial velocity is caused by drag on the fuel droplets that are moving relative to the ambient air. In contrast, the swirl-component drag is not significant because the ambient air rotates with the fuel droplets. In the case of reduced ambient pressure as would occur with early injection, the fuel spray has a wide hollow-cone structure. In the case of higher ambient pressure, the higher drag force changes the spray into a narrow solid-cone shape having a reduced tip penetration.

Fuel Spray Characteristics – As is the case for diesel combustion, the fuel spray characteristics are of significant importance to GDI combustion systems. Parameters such as the spray cone angle, mean drop size, spray penetration, and fuel delivery rate are known to be critical, and the optimum matching of these parameters to the air flow field and spark location usually constitute the essence of the GDI engine development project. The primary fuel spray characteristics of a port fuel injector generally have much less influence on the subsequent combustion event, mainly due to the integrating effects of the residence time on the closed valve, and due to the secondary atomization that occurs as the induction air flows through the valve opening. For direct injection in both GDI and diesel engines, however, the mixture preparation time is significantly less

Medición de pulsos de inyección Chevrolet Cruze.

Medición pulsos de inyección Mazda CX-3

Solenoide modelo

Construcción de circuito

Circuito y doble solenoide terminados