



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

FACULTAD DE ESTUDIOS
SUPERIORES CUAUTITLAN

" ANALISIS DEL FUNCIONAMIENTO
FISICO DEL TRANSISTOR MOS Y SUS
LIMITACIONES EN SU APLICACION A LA
MICROELECTRONICA "

298073

T E S I S

QUE PARA OBTENER EL TITULO DE
INGENIERO MECANICO ELECTRICISTA

P R E S E N T A N :

CLAUDIA ALEJANDRA LOPEZ RODRIGUEZ

LUIS MARTIN RESENDIZ MENDOZA

ASESOR:

M. EN A. I. PEDRO GUZMAN TINAJERO



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLAN
 UNIDAD DE LA ADMINISTRACION ESCOLAR
 DEPARTAMENTO DE EXAMENES PROFESIONALES

ASUNTO: VOTOS APROBATORIOS

FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLAN



DEPARTAMENTO DE EXAMENES PROFESIONALES

DR. JUAN ANTONIO MONTARAZ CRESPO
 DIRECTOR DE LA FES CUAUTITLAN
 PRESENTE

ATN: Q. Ma. del Carmen García Mijares
 Jefe del Departamento de Exámenes
 Profesionales de la FES Cuautitlán

Con base en el art. 28 del Reglamento General de Exámenes, nos permitimos comunicar a usted que revisamos la TESIS:

"Análisis del funcionamiento físico del transistor
 MOS y sus limitaciones en su aplicación a la
 microelectrónica".

que presenta la pasante: Claudia Alejandra López Rodríguez
 con número de cuenta: 9110054-6 para obtener el título de:
 Ingeniera Mecánica Electricista

Considerando que dicho trabajo reúne los requisitos necesarios para ser discutido en el EXAMEN PROFESIONAL correspondiente, otorgamos nuestro VOTO APROBATORIO.

ATENTAMENTE

"POR MI RAZA HABLARA EL ESPIRITU"

Cuautitlán Izcalli, Méx. a 22 de Agosto de 2001

PRESIDENTE

Ing. Ramón Osorio Galicia

VOCAL

Ing. Margarita López López

SECRETARIO

M. AT. Pedro Guzmán Tinajero

PRIMER SUPLENTE

Ing. Vicente Magaña González

SEGUNDO SUPLENTE

Ing. Albino Arteaga Escamilla



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLAN
UNIDAD DE LA ADMINISTRACION ESCOLAR
DEPARTAMENTO DE EXAMENES PROFESIONALES

ASUNTO: VOTOS APROBATORIOS
FACULTAD DE ESTUDIOS
SUPERIORES CUAUTITLAN



DEPARTAMENTO DE
EXAMENES PROFESIONALES

DR. JUAN ANTONIO MONTARAZ CRESPO
DIRECTOR DE LA FES CUAUTITLAN
PRESENTE

ATN: Q. Ma. del Carmen García Mijares
Jefe del Departamento de Exámenes
Profesionales de la FES Cuautitlán

Con base en el art. 28 del Reglamento General de Exámenes, nos permitimos comunicar a usted que revisamos la TESIS:

"Análisis del funcionamiento físico del
transistor MOS y sus limitaciones en su
aplicación a la microelectrónica".

que presenta el pasante: Luis Martín Reséndiz Mendoza
con número de cuenta: 9201536-8 para obtener el título de :
Ingeniero Mecánico Electricista

Considerando que dicho trabajo reúne los requisitos necesarios para ser discutido en el EXAMEN PROFESIONAL correspondiente, otorgamos nuestro VOTO APROBATORIO.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"

Cuautitlán Izcalli, Méx. a 22 de Agosto de 2001

PRESIDENTE	Inq. Ramón Osorio Galicia	
VOCAL	Inq. Margarita López López	
SECRETARIO	M. AF. Pedro Guzmán Tinajero	
PRIMER SUPLENTE	Inq. Vicente Magaña González	
SEGUNDO SUPLENTE	Inq. Albino Arteaga Escamilla	

Queremos agradecer especialmente a nuestro asesor de Tesis, M. en A. I. Pedro Guzmán Tinajero, por todo su apoyo, ayuda y sobre todo por la confianza que ha depositado en nosotros.

Así, como a los profesores integrantes del jurado

Ingeniero Ramón Osorio Galicia

Ingeniero Margarita López López

Ingeniero Vicente Magaña González

Ingeniero Albino Arteaga Escamilla

por su apoyo y confianza.

A la universidad:

Por que ser universitarios es uno de los más grandes honores que hemos tenido en mi vida.

CLAUDIA;

QUIERO DAR LAS GRACIAS:

A mi familia:

A mis padres: Ma. De Lourdes Rodríguez B. y Juan Cruz López C.

A mis hermanas: Vanessa y Mónica López Rodríguez.

A mi tía: Cecilia T. López Contreras.

A mis primos: Valeria, Jose Luis, Alma y Emanuel Solís López

y a mis abuelas: Remedios Becerra y Victorina Contreras

por todo el amor que me han dado, por su apoyo incondicional, por su lealtad, por su ayuda en todo momento, por los regaños necesarios, por soportar mi carácter, por enseñarme a ser una mejor persona, por sus consejos, por los buenos y malos momentos, por consolarme, por apoyarme para seguir siempre hacia adelante no solo en mis estudios profesionales, sino en toda mi vida, mil gracias, son ustedes mis personas favoritas.

A Carlos A. Rivera Guevara:

por que alcanzar está meta no hubiera sido lo mismo sin tu ayuda incondicional, sin la confianza que siempre me has tenido y sin todo el apoyo que me has brindado tanto en las buenas como en las malas situaciones.

A mi gran amigo y colega Luis:

No solo por darme su amistad durante todos estos años, sino por soportarme el paso, que no es nada fácil.

A mis amigos Felipe Acuatla, Mario A. Malvaez, Fernando Valdivieso, Marco Antonio López H. y a Noé S. Badillo:

por compartir conmigo todos estos años de escuela y sobre todo por ser mis amigos.

A todos los profesores que durante todos mis estudios me dieron parte de ellos mediante todas sus enseñanzas.

A la Dra. Magali Estrada y al Dr. Antonio Cerdeira, que con sus clases enriquecieron grandemente este trabajo de tesis.

LUIS MARTÍN;

A mi Madre Gloria Mendoza:

Porque siempre me ha apoyado, su ejemplo de tenacidad me ha dado siempre la fuerza para luchar por las cosas que valen la pena.

A mi Padre Luis Reséndiz:

Por su cariño y espero que te des cuenta que todo los sacrificios tienen una recompensa y de que todo es posible con voluntad.

A mis hermanas Maribel y Lorena:

Por haber compartido conmigo todo lo bueno y lo malo que nos ha tocado vivir y porque siempre están conmigo para celebrar mis triunfos y reconfortar los fracasos.

A Claudia López:

Por su amistad, por estar cuando he necesitado de una amiga tanto en los buenos y malos momentos, porque este solo es un poso más en nuestro futuro.

A mi amigo, Felipe Acuautla.

Por ser siempre un gran amigo, por compartir alegres momentos, además de contar siempre con su amistad; y por toda su ayuda.

A mis amigos, Fernando O. Valdivieso, Noe S. Badillo, Mario A. Malvaez, Alicia del Angel, Claudia Guzman, Fernando Bello, Sandra Morales, Magdalena Balderay, Rocío Gonzalez, Carlos Rivera, Marco A. López.

Por enseñarme que hay cosas que no se aprenden en los libros y porque sin su apoyo y su amistad no habría llegado a esta instancia.

A Rocío Garcia, Rocío Pontifes.

Por su ejemplo de tratar de ser siempre mejor en todos los aspectos.

A Mamá Chona y Tía Carmela.

Por su ayuda en los momentos difíciles.

A mis padrinos, Elena Alfredo; Primos, Beatriz, Juan, Alejandro,
Fernando, Antonio, Rocio.

Por que siempre me han alentado a seguir adelante.

INDICE

INTRODUCCIÓN	I
CAPITULO 1: FÍSICA DE SEMICONDUCTORES	
1. Semiconductores.	1
1.1 Clasificación de los materiales desde el punto de vista eléctrico.	1
1.2 Estructura electrónica de los materiales sólidos.	3
1.3 Conductores semiconductores y aislantes.	11
1.4 Semiconductores intrínsecos.	14
1.4.1 Portadores de carga: concepto de hueco.	16
1.4.2 Interpretación del esquema de bandas de energía.	17
1.4.3 Fenómenos de conducción.	21
1.4.4 Fenómenos de excitación de portadores.	23
1.5 Semiconductores extrínsecos.	25
1.5.1 Semiconductor de tipo n.	26
1.5.2 Semiconductor de tipo p.	27
1.6 Ley de acción de masas.	31
1.7 Procesos de transporte de carga en semiconductores.	35
1.8 Cálculo de la concentración de portadores.	36
1.9 Efecto de la temperatura en la concentración de portadores.	38
1.9.1 Distribución de energía de los portadores.	38
1.9.2 Cálculo de la concentración de portadores.	48

1.9.3 Concentración de portadores intrínsecos.	51
1.10 Determinación del nivel de Fermi.	52
1.11 Procesos de conducción en semiconductores.	57
1.12 Procesos de difusión.	63
1.13 Constancia del nivel de Fermi.	64
1.14 Tiempo de vida de los portadores.	65
CAPITULO 2: ESTRUCTURA METAL-ÓXIDO-SEMICONDUCTOR.	
2 Estructura metal-aislante-semiconductor.	69
2.1 La estructura MOS ideal.	69
2.2 Potencial de superficie.	78
2.3 Capacitancia de la estructura MOS ideal.	84
2.3.1 Cálculo de la carga acumulada.	84
2.3.2 Capacitancia de la estructura MOS ideal. Efecto de la tensión aplicada.	89
2.4 Desviaciones del comportamiento ideal.	95
CAPITULO 3: TRANSISTOR MOS	
3 Introducción.	101
3.1 El transistor de metal-óxido-semiconductor.	103
3.2 Modos de funcionamiento de los transistores MOS.	104
3.2.1 El MOS de deserción.	105

3.2.2	El MOS de acrecentamiento.	106
3.3	Descripción del funcionamiento del MOS.	109
3.4	Cálculo de las características intensidad voltaje del MOS.	113
3.5	Modelo circuital del MOS.	118
3.6	Circuito equivalente del MOS para señales pequeñas.	120
3.7	El MOS como inversor.	122
3.8	Circuitos integrados MOS.	126
3.8.1	El inversor MOS integrado.	126
CAPITULO 4: LIMITACIONES DEL TRANSISTOR MOS		
4	Transistor MOS: sus limitaciones.	129
4.1	Velocidad de respuesta.	129
4.2	Resistencia en serie.	131
4.3	Corriente subumbral.	132
4.4	Escalado ("Scaling").	136
4.4.1	Escalado con campo constante.	137
4.4.2	Escalado a voltaje constante.	139
4.4.3	Escalado a voltaje cuasi-constante.	139
4.4.4	Escalado generalizado.	139
4.5	Efecto de las geometrías reducidas.	140
4.5.1	Velocidad límite.	141
4.5.2	Canal corto.	145

4.5.3 Canal estrecho.	148
4.5.4 Portadores calientes.	149
4.5.5 Ruptura.	150
CONCLUSIONES	153
ABREVIATURAS	160
BIBLIOGRAFÍA	163

INTRODUCCIÓN

No cabe duda que, desde la invención del transistor en 1947, la Electrónica ha sido la ciencia que más ha influido en el desarrollo tecnológico e industrial de nuestro siglo. Las modernas tecnologías de las comunicaciones, de las computadoras o del espacio no podrían haber alcanzado el desarrollo actual si no fuera por los avances logrados en el campo de la microelectrónica entre ellos, la integración de varios millones de componentes en una sola pastilla ("chip"), como las utilizadas en las modernas memorias semiconductoras.

La estructura multicapas que se forma en un semiconductor del tipo Metal-Óxido-Semiconductor (MOS) es una de las más importantes en la técnica de los semiconductores. Este tipo de estructura es utilizada para el estudio de las propiedades superficiales de los semiconductores y además tiene una gran importancia práctica en dispositivos como los capacitores MOS, los transistores MOS y los circuitos integrados MOS, base de la microelectrónica moderna.

Desde la aparición del transistor MOS las áreas de aplicación del mismo se han ido ampliando y especializando. Actualmente las principales áreas de aplicación son:

1. Transistores de muy alta impedancia de entrada.
2. Transistores de potencia.
3. Circuitos integrados digitales de alta escala de integración: lógicos y de memorias.
4. Circuitos integrados analógicos digitales

El área de mayor impacto científico, tecnológico y económico ha sido la de los circuitos integrados de muy alta escala de integración, donde en la actualidad se alcanzan niveles de integración de 3 a 10 millones de transistores en un "chip" para los circuitos lógicos, y decenas de millones de transistores en las memorias. Esto ha sido posible por las reducidas dimensiones que pueden tener los transistores MOS, la fiabilidad de la tecnología de fabricación, los reducidos consumos de energía y las velocidades relativamente altas de operación, mayores de 300 MHz.

Dentro del área de los circuitos integrados uno de los principales problemas que se presentan al ir aumentando la densidad de integración, es la reducción de las dimensiones geométricas del transistor MOS, lo que conlleva a la aparición de nuevos factores que influyen en el funcionamiento del dispositivo, por lo que varían sus parámetros de trabajo, es decir, bajo esta condición el dispositivo presenta limitaciones.

Por lo cual, el **objetivo** de este trabajo de tesis es hacer un análisis detallado del comportamiento físico del Transistor MOS, el cual nos permitirá entender su funcionamiento no como un elemento circuital, sino como un dispositivo físico, y a su vez poder estudiar las diferentes limitaciones que se presentan cuando se aborda el problema del escalamiento de las dimensiones de este transistor.

El conocimiento de las limitaciones nos permitirá proponer algunas soluciones que ayuden a resolver el problema anteriormente planteado, así como dar las bases para estudios posteriores de posgrado en esta área de investigación.

Tomando en cuenta lo descrito anteriormente este trabajo se divide en los siguientes cuatro capítulos.

En el capítulo 1 se hace un resumen de los conceptos básicos de la física de los semiconductores necesarios para el análisis de cualquier dispositivo semiconductor.

El estudio de la estructura Metal-óxido-Semiconductor (MOS) también es de suma importancia en este análisis, ya que esta estructura es la base de todos los dispositivos de efecto de campo, como se verá en el capítulo 2.

Los capítulos 3 y 4 son la parte principal de este trabajo; en el capítulo 3 se hace un análisis del funcionamiento físico del transistor MOS, sus principales características, parámetros y su funcionamiento como elemento en el circuito inversor, base de la microelectrónica. Finalmente en el capítulo 4 se desarrolla el estudio de las diferentes limitaciones que afectan el funcionamiento de este dispositivo al ser disminuido en su geometría, para poder así dar algunas soluciones a este problema.

CAPITULO 1

FÍSICA DE SEMICONDUCTORES

1 SEMICONDUCTORES.

Los semiconductores ocupan un lugar prominente en el conjunto de los materiales. Esto se debe al alto grado de desarrollo que se ha alcanzado en el conocimiento de sus propiedades básicas así como también en el de sus aplicaciones. Podemos decir que hoy en día los semiconductores son piezas básicas en toda la tecnología electrónica, la cual en los últimos años ha mostrado un crecimiento espectacular, abarcando el campo de los procesadores, las comunicaciones, la robótica etc.

1.1 CLASIFICACIÓN DE LOS MATERIALES DESDE EL PUNTO DE VISTA ELÉCTRICO.

Desde el punto de vista eléctrico, los materiales suelen dividirse en tres categorías atendiendo a su conductividad: conductores, semiconductores y aislantes. En la Fig. 1.1 se han ordenado algunos de los materiales típicos según el valor de su conductividad¹.

¹ Los valores indicados en la Fig. 1.1 han de tomarse como aproximados, ya que la conductividad es una magnitud sujeta a la influencia de muchos factores, tales como el estado de agregación del material, su estructura cristalina, temperatura, etc.

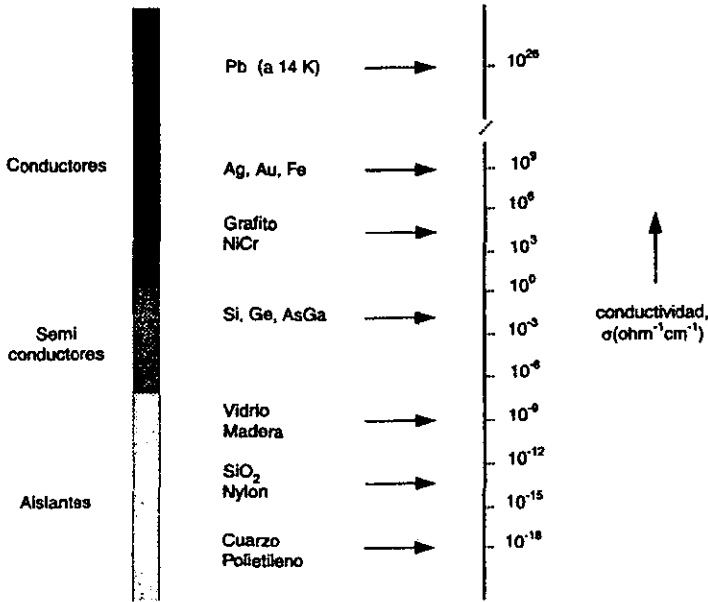


Fig. 1.1. Representación de los valores de la conductividad en algunos materiales típicos.

El primer dato a destacar en la Fig. 1.1 es que la conductividad es una de las magnitudes físicas que admite mayor espectro de variación. Dejando aparte el caso de los materiales superconductores, la conductividad puede variar en más de 25 órdenes de magnitud al pasar de los materiales aislantes como el vidrio y los plásticos a los materiales conductores, tales como el cobre y la plata. Esto da lugar a que los diferentes tipos de materiales puedan presentar fenómenos eléctricos muy diversos. En los materiales conductores, los fenómenos predominantes al aplicar un campo eléctrico son los de conducción, debido al movimiento de electrones libres en el interior del material arrastrados por el campo eléctrico aplicado.

Los semiconductores forman un grupo de materiales que presentan un comportamiento intermedio entre los conductores y los aislantes. Los semiconductores en estado puro y a temperaturas bajas presentan una conductividad relativamente baja por lo que sus propiedades se asemejan a las de los aislantes. Sin embargo, la conductividad de los materiales es una función creciente con la temperatura de forma que a la temperatura ambiente la mayoría de los semiconductores presentan una conductividad apreciable, aunque siempre menor que la de los metales. Incluso a una temperatura dada, es posible variar a voluntad la conductividad de estos materiales si se les añade una cantidad controlada de impurezas de determinados elementos químicos. Es precisamente esta característica la que ha permitido desarrollar una gran variedad de componentes y dispositivos electrónicos basados en los materiales semiconductores.

1.2 ESTRUCTURA ELECTRÓNICA DE LOS MATERIALES SÓLIDOS.

Nos podemos preguntar, a que obedecen las diferencias de comportamiento electrónico entre unos materiales y otros. Como la mayor parte de las propiedades de estado sólido, estas diferencias están originadas por la diferente composición química y estructura electrónica de enlace de los átomos que forman el material. De esta manera es necesario revisar los aspectos más esenciales que determinan la estructura de enlace de los materiales.

Considerando primero la estructura electrónica de **átomos aislados**, esto es sin interacción entre ellos. La mecánica cuántica nos dice que los electrones de los átomos se mueven alrededor del núcleo con una cierta energía que sólo puede tomar unos valores bien definidos (orbitales o niveles atómicos). El cálculo de la energía asociada a

los niveles atómicos es generalmente complejo, y sólo es posible llevarlo a cabo de forma exacta para el átomo de hidrógeno, formado por un protón y un electrón. En este caso, considerando el núcleo en reposo, la energía del electrón ocupando un nivel n viene dada por la expresión:

$$E = -\frac{q^2 m_0}{8\epsilon_0 n^2 h^2} \quad (1.1)$$

donde q es la carga del electrón, m_0 su masa, ϵ_0 la permitividad del vacío y h la constante de Planck² Introduciendo los valores numéricos de estas magnitudes resulta para la energía del nivel n :

$$E = -\frac{13.6}{n} \text{ eV} \quad (1.2)$$

En el nivel mas bajo ($n=1$) la energía vale -13.6 eV. Este nivel se denomina *estado fundamental* del átomo de hidrógeno. El signo negativo indica que se trata de energía de enlace, esto es la energía necesaria para sacar el electrón desde el estado fundamental hasta una posición fuera de la influencia del núcleo (a distancias infinitas). Esta energía, que para el hidrógeno es de 13.6 eV, se denomina también *energía de ionización*. El estado fundamental del electrón puede ocupar también otros estados de mayor energía (con $n>1$), determinados *estados excitados*, cuando recibe energía suficiente mediante algún proceso de excitación(térmica, luminosa, etc.).

En la Fig. 1.2a se da un esquema de los niveles energéticos del electrón para el átomo de hidrógeno, mostrando así mismo la curva de energía potencial, la cual sigue

² $h = 6.626617 \times 10^{-34}$ joules -s

siendo una ley del tipo: $E = -\frac{q}{4\pi\epsilon_0 r}$, siendo r la distancia del electrón al centro del átomo. Esta curva determina la región espacial en la cual el electrón puede moverse bajo la influencia del núcleo, ya que fuera de los límites de la curva la energía total del electrón sería menor que la energía potencial y por tanto no hay estados posibles para el electrón. Los átomos con menor número de electrones tienen también una estructura de niveles energéticos similar a la del hidrógeno, aunque están distribuidos y ordenados de manera más compleja, ya que el valor de la energía no depende solamente del valor de n sino también de otros números cuánticos. En cualquier caso, en el estado fundamental (es decir en el de la energía más baja), los electrones del átomo se distribuyen entre los niveles más bajos de energía siguiendo el *principio de exclusión de Pauli*, el cual establece que en cada estado cuántico (definido por un conjunto de números cuánticos) no puede haber más de un electrón. Debido a la multiplicidad del spin, este principio implica que en cada nivel de energía sólo puede haber hasta un máximo de dos electrones.

Cuando se trata de **moléculas formadas por dos o más átomos**, solamente los electrones de las capas más externas de cada átomo participan en el enlace interaccionando con el resto de los átomos mientras que el resto de los electrones siguen unidos a sus núcleos respectivos. En un enlace típicamente covalente, tal como el que se presenta en muchas moléculas diatómicas homopolares (H_2 , Cl_2 , etc.), la interacción de los electrones de enlace con el potencial eléctrico de los dos núcleos atómicos da lugar a un desdoblamiento de los niveles de energía originales de estos electrones.

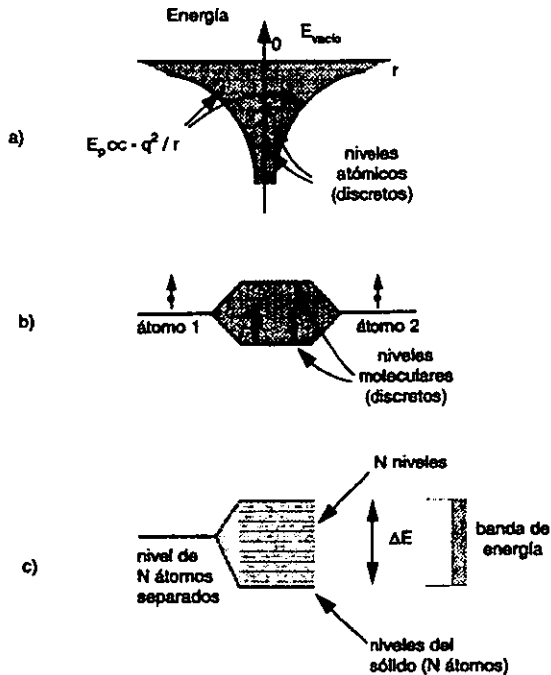


Fig. 1.2 Estructura de los niveles energéticos de electrones: a) en un átomo aislado, b) en una molécula formada por dos átomos, c) para un conjunto de N átomos.

El resultado son, dos niveles nuevos separados por una cierta energía, ΔE , uno de ellos con energía más baja que en nivel original (estado fundamental), y el otro con energía más elevada (estado excitado). Los electrones de enlace se sitúan en los nuevos niveles ocupando primero los de energía más baja, la ocupación se hace siguiendo también el principio de exclusión de Pauli, de forma que si por ejemplo cada uno de los átomos de la molécula aportan un electrón al enlace, los dos electrones se sitúan en el nivel inferior, dejando el nivel superior vacío (susceptible de ser ocupado si la molécula

se encuentra en un estado excitado) (Fig. 1.2b). El estado fundamental corresponde en este caso al de los electrones de enlace moviéndose alrededor de los núcleos de cada átomo, en los llamados *orbitales moleculares*, con una energía de movimiento bien definida. El resto de los electrones pertenecientes a las capas más internas de átomos no participan en el enlace y por tanto se mantiene en los orbitales atómicos originales, con una energía que es prácticamente la misma que la que tenían con los átomos separados.

En el caso de moléculas **aisladas con un cierto número de átomos** unidos entre sí mediante enlace covalente es mucho más complejo, aunque se pueden extrapolar algunas de las conclusiones del modelo descrito anteriormente para moléculas diatómicas. Así, se puede demostrar que en las moléculas simples formadas por N átomos iguales los electrones de las capas más internas se mantiene en sus niveles originales mientras que los electrones de enlace se sitúan en nuevos niveles originados por el desdoblamiento de los últimos niveles atómicos. Como resultado de este desdoblamiento aparecen N niveles nuevos (o sub-niveles) donde se sitúan los electrones de enlace, ocupando primero los de energía más baja. Hay que tener en cuenta que con el principio de exclusión de Pauli, cada uno de estos sub-niveles puede albergar hasta dos electrones debido al spin. Un aspecto interesante a destacar en este caso es que la **diferencia de energía, ΔE , entre el sub-nivel más bajo y el más alto es prácticamente independiente del número total de átomos que forma la molécula.**

Generalmente esta diferencia de energía tiene un valor de unos pocos electrón-volts³ (eV), (Fig. 1.2c).

El fenómeno del desdoblamiento de los niveles atómicos es completamente general y se presenta también en los **sólidos con enlace covalente**. Este es el caso e la mayoría de los semiconductores y también de los metales, en los cuales los número de átomos participantes en el enlace es muy elevado (alrededor de 10^{23} átomos cm^{-3}). En estos materiales, la interacción de los electrones de enlace con el conjunto de los N átomos del sólido da lugar al desdoblamiento de los niveles atómicos originales más elevados en un total de N nuevos sub-niveles. La diferencia de energía entre el sub-nivel máximo y el mínimo sigue siendo de unos pocos eV, por lo que los sub-niveles individuales se encuentran muy próximos entre sí, es decir, separados por una energía extremadamente pequeña, ya que ahora el valor de N es muy elevado. La continuación de energía de los subniveles permite hablar en este caso de una **banda de energía** con una anchura total de unos pocos eV constituida por N subniveles, y con capacidad de alojar hasta 2N electrones provenientes de los electrones de valencia de cada uno de los átomos. Si por ejemplo, cada átomo aporta un solo electrón, la banda de energía queda ocupada hasta la mitad. En cambio, si aportan dos electrones por átomo la banda quedaría completamente ocupada. Esta banda de energía ocupada por los electrones de valencia se denomina **banda de valencia**. Es importante mencionar que los orbitales atómicos originales de los electrones de enlace forman un solapamiento de unos con otros nuevos orbitales que se extienden espacialmente por todo el sólido.

³ 1 eV = 1.60218×10^{-19} J

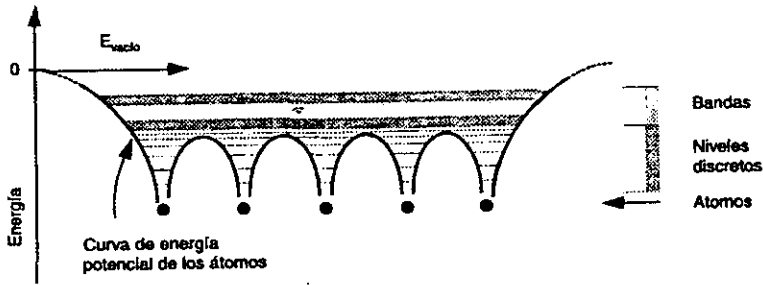


Fig. 1.3 Estructura de los niveles energéticos y de las bandas de energía correspondientes a una red monodimensional de átomos.

Los niveles excitados de los átomos, incluso aunque no estén ocupados por electrones también están sujetos a un desdoblamiento, dando lugar a la formación de una banda de energía cuando se trata de materiales sólidos con un número elevado de átomos. Así, en un sólido con enlace covalente los niveles excitados se convierten en una banda continua de niveles energéticos que se sitúa por encima de la banda de valencia, separada por una zona o **"gap" (banda de energía prohibida)**, de energía donde no existen niveles. Se trata pues de una banda prohibida, en la que no existen estados energéticos posibles para los electrones de enlace.

En la Fig. 1.3⁴ se da un esquema de los niveles energéticos de los electrones de un sólido ideal formado por una red de átomos monodimensional. En la Fig. 1.3 se incluye la curva de energía potencial debida al conjunto de átomos que limita también la extensión espacial del movimiento de los electrones.

⁴ Pag. 232 Ref. 19

En los átomos ocupando posiciones extremas, esta curva se extiende hasta un valor de energía igual a cero, es decir, igual que en el caso de átomos aislados. A menudo a este nivel de energía se le conoce como **nivel de vacío** ($E_{\text{vacío}}$ en la Fig. 1.3). Según se observa, entre cada átomo la curva de energía potencial alcanza un valor máximo por debajo del nivel de vacío, debido a la interacción de los átomos entre sí.

Hay que notar además que en los sólidos, al igual que en las moléculas aisladas, los electrones de las capas más internas se sitúan en niveles discretos limitados espacialmente por la curva de energía potencial de cada átomo, mientras que los de las capas más externas que participan en el enlace se distribuyen entre los niveles que forman la banda de valencia. Por encima de esta banda se encuentra la banda correspondiente a los estados excitados, también denominada **banda de conducción**. Nótese que tanto la banda de valencia como la de conducción no están limitadas espacialmente y por tanto se extienden por toda la cadena de átomos. Esto significa que en un sólido con enlace covalente los electrones de enlace tiene cierta capacidad de movimiento a través de todo el cristal intercambiando su posición unos con otros. A temperaturas próximas al cero absoluto todos los electrones de enlace ocupan los niveles más bajos de la banda de valencia. Sin embargo a temperaturas superiores, algunos electrones pueden ser excitados a otros niveles más elevados dentro de esta banda siempre que existan niveles o estados vacantes. Incluso pueden pasar a la banda superior de energía si los electrones adquieren energía suficiente para saltar el "gap".

Son los electrones pertenecientes a estas bandas de energía (bandas de valencia y de conducción) los que confieren las propiedades eléctricas características de los materiales. La energía de estas bandas, su separación, el número de electrones en cada

banda, etc., está determinado en parte por los factores intrínsecos del material tales como el tipo de enlace, distancia entre los átomos, etc. Y también factores extrínsecos al material, como la temperatura, contenido de impurazas, etc.

1.3 CONDUCTORES, SEMICONDUCTORES Y AISLANTES.

En la Fig. 1.4 se presenta un esquema de la estructura de bandas de energía típica de los materiales conductores, semiconductores y aislantes a temperaturas próximas a 0 K. En este esquema, el eje horizontal representa una de las coordenadas espaciales dentro de la red de átomos, mientras que el eje vertical indica la energía total de los electrones dentro de cada banda. En este tipo de diagramas normalmente se prescinde por comodidad del nivel de vacío, ya que a menudo lo que interesa conocer es la diferencia de energía del electrón en relación con los bordes superior e inferior de la banda de valencia y de conducción, respectivamente. Según se ha mencionado anteriormente las dos bandas representadas son las últimas que pueden estar ocupadas por los electrones de enlace del material y reciben el nombre de banda de valencia, la energía inferior, y banda de conducción la de energía más elevada. Ambas bandas están separadas por una zona de energía prohibida de valor E_g denominada banda prohibida.

En las proximidades del cero absoluto los electrones de la banda de valencia ocupan los niveles más bajos de energía, formando parte del enlace entre los átomos (zona rayada en la Fig. 1.4). En el caso de que la banda no se encuentre totalmente ocupada, los niveles más altos de energía dentro de la banda permanecerán vacíos. En cambio, a temperaturas más elevadas, una fracción apreciable de los electrones puede ser excitada a niveles con energía superior siempre que estos niveles (o posiciones de

enlace) se encuentren vacantes. Esta circunstancia, es decir, la existencia de niveles vacantes en una banda de energía es la que permite que los electrones se puedan mover dentro del cristal bajo la acción de un campo eléctrico. Esta condición viene impuesta por las leyes de la mecánica cuántica, ya que el movimiento de los electrones implica una ganancia en energía cinética y por tanto en su energía total. Este aumento de energía sólo se produce si existen niveles de energía vacantes en la banda de valencia, de forma que los electrones en su movimiento puedan pasar a ellos. El movimiento de los electrones en el interior del cristal es bastante complejo, ya que no sólo actúan las fuerzas del campo eléctrico aplicado sino también las debidas a la interacción de los electrones con los átomos del cristal.

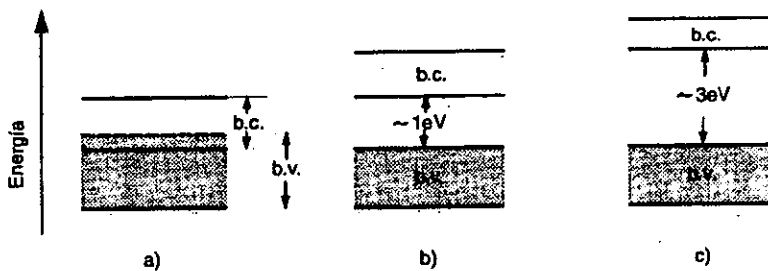


Fig. 1.4 Estructura de bandas en materiales de tipo: a) conductor, b) semiconductor, c) aislante, a temperaturas próximas de cero absoluto.

En el caso opuesto, si la banda de energía se encuentra totalmente ocupada la aplicación de un campo eléctrico no implica necesariamente un desplazamiento neto de electrones en la dirección del campo, aún cuando los electrones puedan tener una cierta movilidad dentro de la banda. La ausencia de niveles vacantes da lugar a que el

desplazamiento de un electrón en la dirección del campo este compensado siempre por el de otro electrón en sentido opuesto. De todo esto se deduce que la **presencia de electrones en una banda de valencia o de conducción, en la cual existen niveles o estados vacantes a los cuales el electrón se pueda trasladar, es una de las condiciones que se exige para que los electrones de esa banda puedan participar en los procesos de conducción al aplicar un campo eléctrico.**

En los semiconductores con enlace típicamente covalente, cada átomo aporta también un número determinado de electrones para formar el enlace con los átomos vecinos. A temperaturas próximas a las de cero absoluto todos los electrones de valencia participan en el enlace de unos átomos con otros y la banda de valencia se halla completamente llena, es decir, sin estados vacantes, mientras que la conducción está completamente vacía, por lo que en estas condiciones no puede haber conducción. Sin embargo, la energía de enlace de los electrones es relativamente pequeña de forma que a temperaturas ordinarias (300K) una fracción apreciable de electrones puede romper el enlace y pasar a la banda de conducción donde existe un gran número de estados vacantes. Estas vacantes, junto con las generadas en las bandas de valencia, hacen que los electrones puedan participar en los procesos de conducción cuando se aplica un campo eléctrico. La energía necesaria para romper el enlace corresponde con la energía de la banda prohibida, con un valor alrededor de 1 eV, para la mayoría de los semiconductores (Fig. 1.4b). En un semiconductor típico como el silicio el número de electrones que pueden pasar a la banda de conducción a la temperatura ambiente es del orden de 10^{10} electrones/cm³. Por ello su conductividad aunque apreciable, será mucho más baja que la de los metales. Podemos decir que los materiales semiconductores a

temperaturas bajas tiene un comportamiento típico de los materiales aislantes, mientras que a temperaturas medias o altas su comportamiento se acerca más al de los metales, al poseer un cierto número de electrones disponibles para la conducción.

1.4 SEMICONDUCTORES INTRÍNSECOS.

Los llamados *semiconductores intrínsecos*, son aquellos semiconductores en estado puro y perfectamente cristalizados. El Silicio es uno de los elementos semiconductores típicos empleados en la fabricación de la mayor parte de los dispositivos electrónicos de estado sólido. En estado cristalino los átomos de este material ocupan posiciones tetraédricas en una red similar a la del diamante, compartiendo cuatro electrones con sus átomos vecinos en un enlace de tipo covalente según se indica en la Fig. 1.5a⁵.

En la Fig. 1.5b se da una representación bidimensional de la estructura y del enlace químico de un semiconductor típico, de valencia cuatro, como el silicio (Si) o el germanio (Ge), en los cuales cada átomo comparte dos electrones con otro átomo vecino. Otros compuestos, también con características semiconductoras están formados por la combinación de elementos del grupo III y V de la tabla periódica, como arseniuro de galio (GaAs), fosfuro de galio (GaP), etc., o bien por la combinación de elementos del grupo II y del VI, tales como el sulfuro de zinc (ZnS) o el telurio de cadmio (CdTe). En todos los casos tanto la estructura cristalina del compuesto como su estructura de enlace es muy similar a la del silicio o germanio.

⁵ La teoría de la red cristalina y los tipos de cristales se puede consultar en la Ref. 18 capítulo 1.

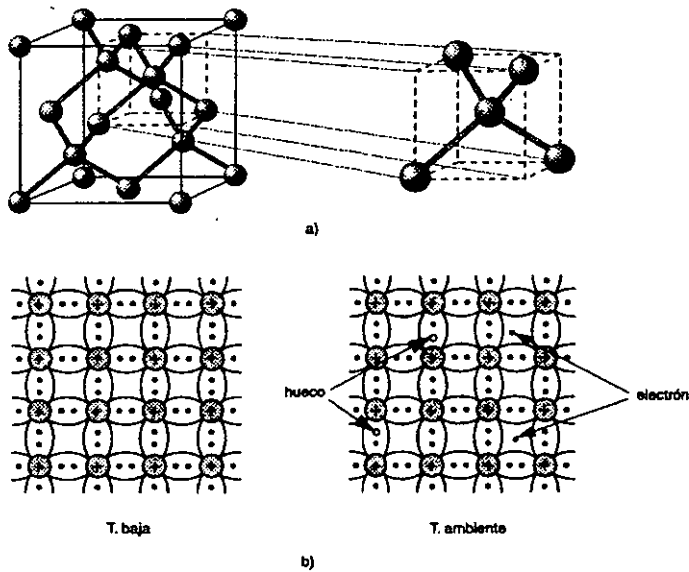


Fig. 1.5 a) Estructura de enlace del silicio, b) representación bidimensional de la estructura de enlace del silicio. Fig. izquierda: a bajas temperaturas mostrando todos los enlaces saturados. Fig. derecha: a temperatura ambiente, con algunos enlaces sin saturar dando lugar a huecos y electrones libres .

A temperaturas bajas, próximas al cero absoluto, todos los enlaces de los átomos se encuentran saturados con los electrones correspondientes, por lo que la banda de valencia se encuentra totalmente ocupada. Si embargo, la energía necesaria para romper el enlace es relativamente pequeña del orden de 1.1 eV, para el silicio y 0.7 eV para el germanio. Los electrones pueden recibir esta energía mediante excitación térmica, por ejemplo. Los procesos de excitación térmica ocurren cuando se eleva la temperatura del material. Algunos electrones entonces ganan energía a partir de las vibraciones de los átomos, en cantidad suficiente para romper el enlace, pasando desde

la banda de valencia hasta la de conducción, donde existen numerosos estados o niveles de energía vacantes (Fig. 1.6b).

1.4.1 PORTADORES DE CARGA: CONCEPTO DE HUECO.

La excitación de un electrón a la banda de conducción implica la ruptura de un enlace en algún punto del cristal, donde a su vez se origina un **estado vacante** que además presenta una deficiencia de carga negativa (equivalente a una carga positiva de magnitud igual a la de la carga del electrón). Esta deficiencia de carga asociada a la vacante posee una cierta movilidad en el interior del cristal. De manera gráfica, la movilidad de los huecos se explica si se tiene en cuenta que los electrones que se encuentran en enlaces próximos a la vacante pueden saltar a esa posición vacante dejando tras sí una nueva vacante o enlace sin saturar. Este proceso da lugar a un desplazamiento de la vacante en sentido opuesto al del electrón que efectúa el salto. Mediante un intercambio repetido de la vacante con los electrones de enlaces próximos se origina un movimiento de la vacante de un punto a otro del cristal con un consumo de energía muy pequeño. Es más, la deficiencia de carga negativa asociada a un nivel vacante de la banda de valencia mantiene su entidad una vez que ha sido formada. Debido a ello, los niveles vacantes de la banda de valencia tienen un comportamiento muy similar al de los electrones de la banda de conducción. En realidad, el desplazamiento de estas vacantes electrónicas se hace mediante un proceso más complejo de naturaleza cuántica. Sin embargo, el modelo anterior puede ser suficiente para entender el comportamiento de los niveles vacantes de energía en la banda de valencia.

Estas características de los niveles o estados vacantes denominados también **HUECOS**, permite considerarlos como partículas inmersas en un mar de electrones de enlace dentro de la banda de valencia, es decir, en proporción mucho menor que la de los electrones en esta banda. A la temperatura ambiente solamente un electrón de cada 10^{12} de la banda de valencia, en el caso del silicio, rompe su enlace por excitación térmica para pasar a la banda de conducción, dejando tras sí el correspondiente nivel vacante. Además, debido a que los huecos son capaces de moverse en el interior del cristal también se les puede asociar una energía cinética de movimiento (en realidad se trata de la energía de los electrones que se desplazan en sentido opuesto ocupando las posiciones del hueco). A partir de estos hechos se desprende la importancia del hueco como entidad, con un comportamiento similar al de los electrones que se mueven en la banda de conducción. Tal es así que, desde el punto de vista cuantitativo, el hueco es también considerado como una partícula que se mueve con una energía dentro de la banda de valencia y que posee una carga igual a la del electrón pero de signo opuesto, es decir positiva. Debido a ello los huecos pueden ser arrastrados con una cierta movilidad por un campo eléctrico externo. Incluso es posible también asociar a los huecos con una cierta masa denominada **masa efectiva**. Estas características hacen que tanto los huecos como los electrones de un semiconductor intrínseco sean denominados indistintamente **portadores de carga o portadores intrínsecos**.

1.4.2 INTERPRETACIÓN DEL ESQUEMA DE BANDAS DE ENERGÍA.

En el esquema de bandas de energía, las escala vertical representa la **energía total**, E , de los niveles electrónicos en la banda de valencia o de conducción. En este

esquema el valor de E_g corresponde a la energía mínima necesaria para romper un enlace en el cristal, lo que a su vez implica que tanto el electrón como el hueco generado en el proceso quedan en reposo en los niveles de energía correspondientes al fondo de la banda de conducción, E_c , y al tope de la banda de valencia E_v , respectivamente. En este proceso de ruptura de un enlace, cualquier exceso de energía absorbida sobre el valor de E_g obedece bien sea a que el electrón liberado procede de un nivel inferior a E_v en la banda de valencia, o incluso a que el nivel de destino en la banda de conducción posee una energía mayor que E_c . En uno u otro caso se imparte al electrón o al hueco creado una energía adicional que se traduce en energía de movimiento a través del cristal. Así por ejemplo, si E representa la energía final del electrón una vez que pasa a la banda de conducción, la diferencia $E-E_c$ se interpreta como la energía cinética asociada al movimiento del electrón en la banda de conducción. Del mismo modo, si el hueco generado en la banda de valencia se encuentra en un nivel de energía E (por supuesto diferente a la del electrón) la diferencia E_v-E representa la energía cinética del hueco. En uno y otro caso, puesto que E representa la energía total de la partícula, esto es la suma de las energías cinética y potencial, $E=E_{cin}+E_{pot}$, los valores de E_c y E_v representan a su vez la energía potencial (asociada al enlace) medida desde un cierto nivel de referencia para los electrones en la banda de conducción y los huecos en la banda de valencia, respectivamente (Fig. 1.6).

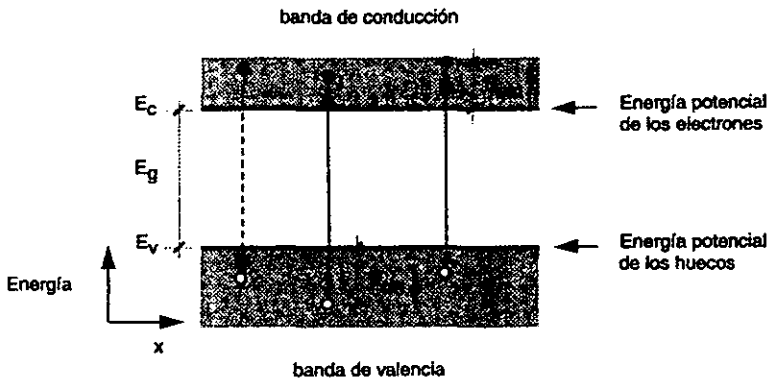


Fig. 1.6 Diagrama de bandas de energía de un semiconductor mostrando las componentes de energía cinética y energía potencial de los electrones y huecos .

Así pues, en el diagrama de bandas de energía, la energía cinética de los electrones de la banda de conducción viene dada por el valor de E medido desde E_c . Por el contrario para los huecos la energía cinética viene dada por el valor de E medido desde E_v hacia abajo. Dado que E_c y E_v corresponden al valor de la energía potencial, se puede tomar para esos niveles un origen arbitrario. Por esta razón, es muy frecuente omitir el nivel de referencia en un esquema de bandas, aunque a veces en los casos que se hace necesario se toma como referencia el nivel de vacío o infinito, mencionado más arriba. Este nivel se halla generalmente situado varios eV por encima del fondo de la banda de conducción.

Las leyes de mecánica cuántica predicen para un electrón, que se mueve en la banda de conducción sometido al potencial periódico de los átomos, una relación entre su energía cinética, $E_{cin}=E-E_c$, y el momento cuántico, p , similar a la que existe para un

electrón libre ($E_{cin} = \frac{p^2}{2m_0}$, con m_0 = masa del electrón libre). Para un electrón con energía E dentro de la banda de conducción tendremos:

$$E = E_c + \frac{p^2}{2m_e^*} \quad (1.3a)$$

siendo p el llamado momento cristalino, esto es el equivalente cuántico del vector momento y m_e^* la masa efectiva del electrón. Debido a que el electrón durante su movimiento, esta interaccionando constantemente con los átomos de la red cristalina, la masa efectiva no es una constante en el sentido estricto, sino que depende de factores tales como la estructura cristalina, la disposición de los átomos en una dirección determinada, etc. Por esta razón m_e^* esta influenciada incluso por la dirección del movimiento del electrón. Algo similar sucede para los huecos, para los cuales se puede considerar que poseen una cierta masa efectiva, m_h^* , determinada por una ecuación análoga a la anterior, es decir:

$$E = E_v + \frac{p^2}{2m_h^*} \quad (1.3b)$$

En el caso del Silicio, por ejemplo, la masa efectiva de los electrones moviéndose en la dirección $[100]^6$ tiene un valor dado por: $m_e^* = 0.19 m_0$. Generalmente, la masa efectiva de los huecos suele ser algo mayor que la de los electrones. El que las masas efectivas de electrones y huecos sean diferentes no debe sorprender, si se considera que el movimiento de los electrones en la banda de conducción es relativamente libre (aún

⁶ Teoría de las redes cristalinas y tipos de cristales, Cap. 1 Ref. 19

cuando interaccionan con los átomos de la red), en cambio el de los huecos implica un intercambio de enlaces entre los átomos de la red.

1.4.3 FENÓMENOS DE CONDUCCIÓN.

Tanto los electrones en la banda de conducción como los huecos en la banda de valencia pueden participar directamente en los procesos de conducción. Aplicando sobre una barra semiconductor uniforme de longitud L una diferencia de potencial $V=V_2-V_1$, con $V_2>V_1$. Este potencial eléctrico se superpone al potencial al que están sometidos los electrones en el interior del cristal y hace que su energía potencial varíe en la cantidad $-qV_2$ para los electrones que se encuentran en un extremos del semiconductor y en $-qV_1$ para los que se encuentran en el extremo contrario. Si para mayor simplicidad se hace $V_1=0$, podemos decir que la energía potencial de los electrones disminuye uniformemente desde uno de los lados hasta el otro, hasta alcanzar el valor de $-qV_2=-qV$. En un esquema de bandas de energía esto implica que las líneas que representan los valores E_v y E_c deben representarse inclinadas según se presenta en la Fig. 1.7a, con una caída total igual a $q\Delta V$. Además, si el semiconductor es uniforme, el campo eléctrico en su interior asociada al potencial aplicado debe ser constante, con un valor E dado por $E=\Delta V/L$. Esto quiere decir que la pendiente de las bandas de energía representadas en la Fig. 1.7a coinciden en valor absoluto con el valor del campo eléctrico. La pendiente de la banda de energía en cada punto del semiconductor es proporcional al valor del campo eléctrico en ese punto, cambiando de signo.

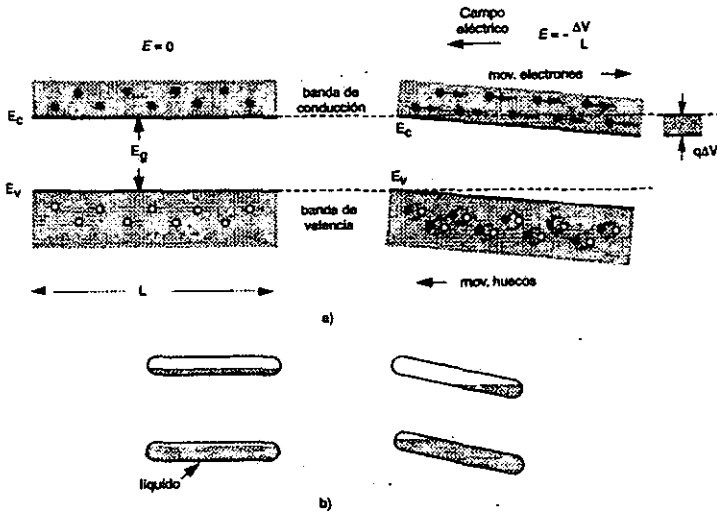


Fig. 1.7 a) Esquemas de bandas de energía de un semiconductor antes y después de aplicar un campo eléctrico. b) Analogía mecánica del movimiento de los electrones en la banda de conducción (tubo superior) y de los huecos en la banda de valencia (tubo inferior).

El movimiento de los electrones se verifica en dirección opuesta al campo, desplazándose siempre hacia los puntos de energía potencial más baja. Así pues, la aplicación del campo eléctrico hace que los electrones de la banda de conducción se muevan dentro de la banda bajando hacia los puntos de menor potencial. Igualmente, cuando se trata de la banda de valencia también puede existir movimiento de electrones siempre que exista un hueco o estado vacante en sus proximidades. Cuando esto ocurre el electrón viaja del mismo modo hacia los puntos de energía potencial más baja, intercambiando su posición con el hueco correspondiente, lo cual se traduce en definitiva en un desplazamiento del hueco en sentido contrario al del electrón. De esto se concluye

que los huecos de la banda de valencia se desplazan en la dirección del campo eléctrico o, lo que es lo mismo, hacia valores de energía potencial más elevada.

1.4.4 FENÓMENOS DE EXCITACIÓN DE PORTADORES.

En un semiconductor que se encuentra en equilibrio térmico a una temperatura dada, existe un proceso continuo de excitación de electrones desde la banda de valencia a la de conducción. En este proceso se rompe un enlace y se crea el hueco correspondiente en la banda de valencia. Los procesos de excitación están a su vez compensados por procesos de recombinación que actúan en sentido opuesto en los cuales un electrón de la banda de conducción se desexcita y pasa a ocupar un lugar vacante de la banda de valencia, con lo que desaparece un hueco. De todo esto se desprende que **en un semiconductor intrínseco, en equilibrio térmico, la concentración de electrones presentes en la banda de conducción, n , debe ser igual a la de huecos en la banda de valencia, p , es decir: $n=p$.** Es más, el valor de n y p debe ser constante con el tiempo si la temperatura del material es constante, a pesar de que continuamente existen procesos de generación y de recombinación de pares electrón-hueco.

En los procesos de excitación térmica los electrones de enlace ganan energía de la red como consecuencia de las vibraciones de los átomos. Estas vibraciones dan lugar a la ruptura de un cierto número de enlaces produciéndose electrones "libres" en la banda de conducción y el correspondiente número de huecos en la banda de valencia. Si llamamos n_i a la concentración de electrones (o huecos) en la banda de conducción (valencia), el valor de n_i debe ser más elevado cuando mayor sea la temperatura del

crystal, a que en este caso la energía de las vibraciones de la red es mayor. Así pues, para un semiconductor intrínseco que se encuentra en equilibrio térmico a una cierta temperatura T podemos escribir:

$$n = p = n_i(T) \quad (1.4)$$

El valor de n_i también depende, obviamente, del valor de la energía de la banda prohibida, E_g , ya que cuanto menor sea E_g mayor es el número de electrones que tienen energía suficiente de excitación para pasar desde la banda de valencia a la de conducción a una temperatura dada.

Un hecho a destacar en los semiconductores es que los electrones excitados a la banda de conducción y los huecos generados en la banda de valencia no se encuentran en una posición estática sino que están en continuo movimiento en el interior del cristal. La energía de movimiento procede también de la energía térmica del cristal, es decir, la impartida por las vibraciones de los átomos de la red del semiconductor.

En algunos semiconductores como el arseniuro de galio (GaAs), para que se efectúe la transición de una banda a otra sólo se requiere que la energía transferida al electrón sea igual o mayor que la energía de la banda prohibida. Por el contrario, los semiconductores como el silicio o el germanio requieren por consideraciones de conservación del momento no sólo el aporte necesario en energía sino también la transferencia de una cierta cantidad de movimiento al electrón.

1.5 SEMICONDUCTORES EXTRÍNSECOS.

En un semiconductor intrínseco, la concentración de portadores intrínsecos – electrones o huecos- está determinada por la temperatura del semiconductor $n_i = n_i(T)$, siendo $n_i(T)$ una función creciente con la temperatura y de naturaleza cuasi-exponencial. Así pues, la conductividad de un semiconductor puro es una magnitud muy sensible a los cambios de temperatura. Desde el punto de vista práctico, la variación de la conductividad con la temperatura constituye un serio inconveniente para la utilización de los semiconductores en estado puro en la fabricación de dispositivos electrónicos, ya que en este caso lo que se pretende es que los dispositivos tengan un comportamiento lo más estable posible con la temperatura. Aparte de ello, interesa también disponer de semiconductores en los que la conducción esté determinada por un solo tipo de portadores, bien sea electrones o huecos.

Existe un procedimiento para obtener un valor relativamente constante del número de portadores de un semiconductor a la temperatura ambiente. Este procedimiento consiste en la introducción de átomos de diferente valencia, en una proporción adecuada, dentro de la red del material semiconductor. En los semiconductores +4, como germanio o el silicio, los átomos añadidos suelen ser de valencia +3, boro (B) o aluminio (Al), por ejemplo, o de valencia +5, fósforo (P), arsénico (As), antimonio (An), etc. La conductividad pasa entonces a estar dominada por la concentración y la naturaleza de los átomos añadidos, también dominados impurezas, y el semiconductor se denomina **extrínseco o dopado**.

1.5.1 SEMICONDUCTOR DE TIPO N.

En la Fig. 1.8a se muestra la estructura de enlace de un semiconductor como el silicio de valencia +4, dopado con una cierta proporción de átomos de antimonio. La valencia de este elemento es +5. Cuando la red del semiconductor esta libre de defectos y formada por un cristal único o monocristal, los átomos de antimonio ocupan posiciones en sustitución de silicio cediendo también cuatro electrones para compartir con sus átomos vecinos de silicio en un enlace covalente. Cada átomo de antimonio tiene además un quinto electrón que no participa en el enlace y que queda por tanto muy débilmente ligado al propio átomo. Por esta razón a temperaturas próximas a la del ambiente este electrón recibe suficiente energía térmica para romper su enlace con el antimonio quedando en completa libertad para moverse a través del cristal. El átomo de impureza queda entonces ionizado con una carga positiva. Desde un punto de vista energético esta situación se representa mediante el paso del electrón desde un cierto nivel de energía (E_d), correspondiente a la energía de enlace con el átomo de antimonio, a la banda de conducción con E_c situado en el interior del "gap" a unas centésimas de eV de energía por debajo de E_c .

A este tipo de impurezas, que como el antimonio son capaces de ceder electrones a la banda de conducción, se les denomina **donadoras**. Los electrones así generados, denominados "extrínsecos", una vez en la banda de conducción son completamente indistinguibles de los electrones "intrínsecos", y por tanto participan en pie de igualdad en los procesos de conducción.

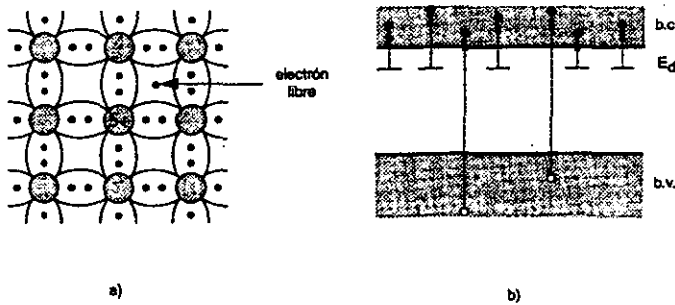


Fig. 1.8 a) Esquema de la estructura de enlace químico del silicio dopado con átomos pentavalentes. b) Esquema correspondiente de la estructura de bandas de energía de los electrones.

Se puede observar en la Fig. 1.8b que la concentración total de electrones, n , en el semiconductor extrínseco puede ser completamente diferente a la de huecos, p , a diferencia de los semiconductores intrínsecos en los que siempre se cumple que $n=p$ (ec. 1.4). La introducción de impurezas de tipo donador en cantidad suficiente hace que a la temperatura ambiente sean $n > p$, y el semiconductor se dice entonces que es de **tipo n** .

1.5.2 SEMICONDUCTOR DE TIPO P.

Si en lugar de impurezas de valencia +5, se añaden átomos de valencia +3, como el boro, ocurre una situación en ciertos aspectos similar a la anterior (Fig. 1.9a). Los átomos ocupan también en este caso, por consideraciones del tamaño, posiciones sustitucionales, y comparten los tres electrones de la capa más externa con el resto de los átomos vecinos de Silicio. La deficiencia de un electrón para completar el enlace hace que electrones vecinos que también participan en el enlace puedan ocupar esa vacante

si son activados mediante energía. La energía necesaria para la activación suele ser muy pequeña de forma que a la temperatura ambiente prácticamente todos los átomos de impurezas están ionizados, es decir han recibido un electrón "extra" para completar el enlace. Mediante este proceso se origina igual número de vacantes de electrones de enlace que átomos de impureza añadidos. El comportamiento eléctrico de estas vacantes es, para todos los efectos, igual al de los huecos generados térmicamente en la banda de valencia (Fig. 1.9b). Así pues, mediante la adición de impurezas de valencia inferior, impurezas *aceptadoras* o *aceptoras*, aparece en la banda de valencia a la temperatura ambiente un número de huecos aproximadamente igual al de átomos de impurezas debido a la excitación de electrones desde la banda de valencia a los niveles de energía correspondientes a los átomos de impureza. El átomo de impureza queda entonces ionizado, esto es, con carga negativa. La energía necesaria para este proceso, representada por E_a , corresponde a la diferencia de energía entre el borde superior de la banda de valencia y el nivel de energía del átomo ionizado (E_a). En este caso se tiene que $p > n$ y el semiconductor se denomina de tipo p.

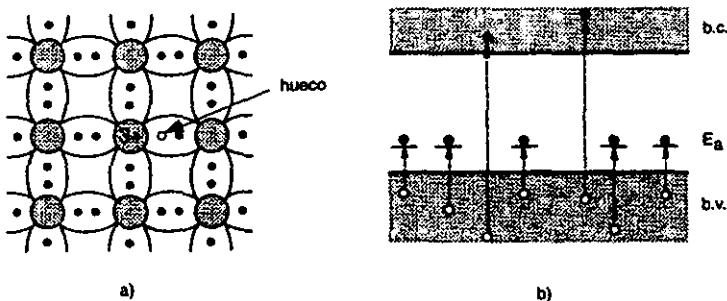


Fig. 1.9 a) Esquema de la estructura de enlace químico del silicio dopado con átomos trivalentes. b) Esquema correspondiente de la estructura de bandas de energía de los electrones.

Generalmente, el valor de la concentración de impurezas (donadoras o aceptoras) y su energía de excitación ($E_c - E_d$) ó ($E_a - E_v$), es tal que, a la temperatura ambiente, la concentración de portadores extrínsecos generados, n o p , es mucho mayor que la concentración de portadores intrínsecos, n_i . En el silicio, por ejemplo, que tiene alrededor de 10^{22} átomos cm^{-3} , la concentración de portadores intrínsecos es del orden de 10^{10} cm^{-3} . Así pues, si se pretende que la concentración de portadores extrínsecos sea 10^4 veces mayor, es suficiente añadir una concentración de átomos de impureza en la proporción de 1 por cada 10^8 átomos de Silicio. La energía de excitación debe ser muy pequeña (alrededor de una centésimas de eV) con objeto de que a la temperatura ambiente todas las impurezas estén ionizadas. Este es el caso de las impurezas normalmente añadidas al Silicio, como el boro, fósforo, Arsénico etc. Por supuesto la pureza del semiconductor de partida debe ser tal que la proporción de impurezas no deseadas sea muy inferior a las del dopaje, es decir, de unas partes por billón, con objeto de que las propiedades de conducción estén determinadas exclusivamente por el dopaje. Estos requerimientos de pureza, así como de la cristalinidad del material son verdaderamente importantes para determinar con precisión las propiedades eléctricas de los semiconductores. Es evidente además que mediante la adición de impurezas de forma controlada se puede obtener un valor prefijado de antemano de la concentración de portadores, y por tanto de la conductividad del material en un rango más amplio de temperaturas.

En la Fig. 1.10 se ha representado de manera cualitativa la variación de la concentración total del número de portadores en función de la temperatura para el caso

del silicio dopado con átomos de arsénico, en una concentración de 10^{15} átomos cm^{-3} (silicio tipo-N). A temperaturas muy bajas la proporción de átomos de arsénico ionizados es muy pequeña por lo que la concentración de portadores (electrones) es también muy pequeña (tramo *a* de la curva). A temperaturas muy bajas, inferiores a los 100K, comienza la ionización de los átomos de impureza, por lo que el valor de n aumenta rápidamente con la temperatura. Si se aumenta todavía más la temperatura llega un momento en que prácticamente todos los átomos de arsénico quedan ionizados, con lo que n alcanza un valor estable. Esto ocurre desde una temperatura ligeramente superior a 100K hasta unos 600K aproximadamente (tramo *b*). Por encima de esta temperatura comienza a dominar la ionización de los átomos de silicio aumentando de forma considerable el número de portadores intrínsecos n_i . El valor de n aumenta de nuevo iniciándose el comportamiento intrínseco (tramo *c*). Así pues, un semiconductor puede comportarse como intrínseco incluso cuando está dopado para ello basta aumentar la temperatura suficiente.

El rango de operación útil de un semiconductor es el comprendido en el tramo *b* de la Fig. 1.10, también llamado **rango de saturación**, ya que en esta región la concentración de portadores tiene un valor constante dentro de un margen amplio de temperaturas que incluye la temperatura ambiente. Además el valor de la concentración y el signo de los portadores es controlable a voluntad dentro de este rango. Basta por ello la adición de impurezas,ceptoras o donadores, en la proporción debida. Es éste, quizás, uno de los aspectos más característicos de los semiconductores, ya que mediante el proceso de dopaje es posible prefijar la conductividad del material en

un valor determinado. La técnicas utilizadas para la introducción de estas impurezas en el interior de un semiconductor es bastante compleja.

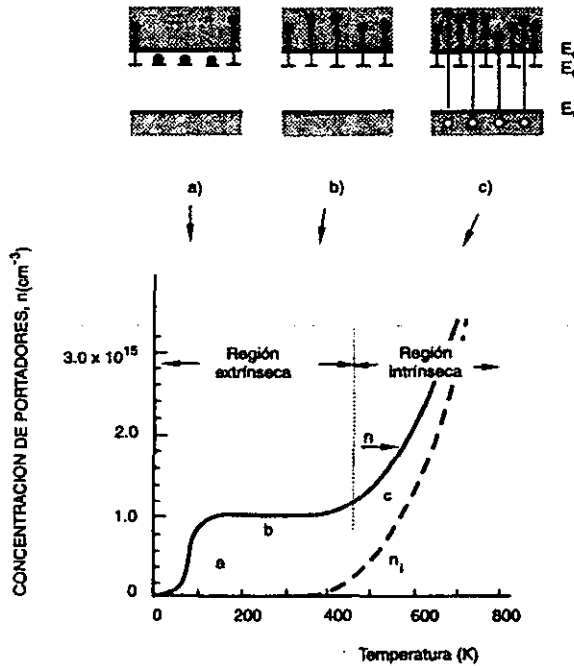


Fig. 1.10 Variación con la temperatura de la concentración total de portadores para el silicio tipo n, dopado con $N_d=1 \times 10^{15}$ cm⁻³ (línea continua). Se muestra también la variación de portadores, n_i , para el caso en el que el semiconductor fuera intrínseco (línea a trazos). En la parte superior se indican los procesos de excitación correspondientes a cada rango de temperatura.

1.6 LEY DE ACCIÓN DE MASAS.

Los electrones en la banda de conducción y los huecos en la banda de valencia generados térmicamente adquieren una cierta energía cinética que les permite moverse en el interior del cristal. Sin embargo, el tiempo de vida de estos portadores no es infinito

ya que una fracción importante de ello esta sometida constantemente a procesos de recombinación, mediante los cuales un electrón puede pasar a ocupar el nivel correspondiente a un hueco, desapareciendo tanto el electrón como el hueco y liberando al mismo tiempo una determinada energía. Estos procesos ocurren por ejemplo cuando un electrón pasa por las inmediaciones de un hueco, aunque existen también otros mecanismos posibles de recombinación. Si el semiconductor se halla en equilibrio térmico los procesos de recombinación están compensados a su vez por los de generación de nuevos pares electrón-hueco debido a la excitación térmica. Suponiendo un semiconductor intrínseco a una cierta temperatura, T. Si se denomina $R(T)$ y $G(T)$ al número de portadores que se recombinan y se generan por unidad de tiempo, respectivamente, es evidente que en equilibrio térmico tendremos (ver Fig. 1.11)

$$R(T) = G(T) \quad (1.5)$$

Las funciones $R(T)$ y $G(T)$ presentan una variación creciente con la temperatura, ya que a medida que aumenta T mayor es el número de electrones que se excita a la banda de conducción y lógicamente mayor es también el número de electrones que pierden energía y pasan a la banda de valencia. En cualquier caso, la ecuación (1.5) implica que en un semiconductor en equilibrio térmico la concentración de electrones en la banda de conducción y de huecos en la de valencia se mantiene constante con el tiempo.

Cuando un semiconductor se encuentra dopado, los procesos de generación y recombinación alteran profundamente la concentración de portadores (electrones y huecos) en el interior del semiconductor respecto al caso intrínseco. Así, cuando se

añade una concentración N_d de impurezas donadoras en suficiente cantidad tendremos a la temperatura ambiente $n \approx N_d$, es decir, el valor de n se incrementa hasta igualar aproximadamente el valor de N_d . El aumento en la concentración de electrones en la banda de conducción origina a su vez una disminución en el número de huecos ya que existe una mayor probabilidad de procesos de recombinación. Algo parecido ocurre al añadir impurezas aceptoras a un semiconductor con una concentración N_a . En este caso el valor de p aumenta hasta el valor de N_a es decir $p \approx N_a$, mientras que el de n disminuye. Considerando el caso de un semiconductor tipo n con una concentración N_d de impurezas donadoras y queremos calcular la concentración de huecos en el semiconductor. Podemos suponer en este caso (aunque el razonamiento también es válido si se trata de un semiconductor tipo P) que la velocidad de recombinación, R , es proporcional al número de electrones presentes en la banda de conducción, n , y al de huecos en la banda de valencia, p , esto es:

$$R = r \cdot n \cdot p \quad (1.6)$$

Siendo r una constante de proporcionalidad. Teniendo en cuenta la relación (1.5), y que para cualquier temperatura R y G han de ser constantes, la ecuación anterior conduce a:

$$n \cdot p = cte(T) \quad (1.7)$$

Es decir, el producto no ha de ser una constante dependiente de la temperatura, exclusivamente, para un semiconductor determinado.

En realidad la ecuación es válida no sólo en el caso extrínseco sino también en el caso intrínseco. Es más, para un material semiconductor dado, bien sea puro o dopado,

el valor de la ecuación (1.7) ha de ser el mismo en un caso u otro. Para calcular el valor de esta constante podemos extender la ecuación (1.7) al caso intrínseco en el que se cumple $n = p = n_i$ (ecuación 1.4). Se obtiene entonces: $n_i^2(T) = cte(T)$. Llevando el valor de la constante a la ecuación (1.7) resulta la conocida **ley de acción de masas**:

$$n \cdot p = n_i^2(T) \quad (1.8)$$

La ecuación es totalmente general y por tanto válida tanto para los semiconductores intrínsecos como para los extrínsecos. A temperatura ambiente (300K) el valor de n_i resulta ser $1.45 \times 10^{10} \text{ cm}^{-3}$ para el Silicio y $1.79 \times 10^6 \text{ cm}^{-3}$ para el Arseniuro de Galio. En el caso del Silicio puro si se añade por ejemplo una concentración de impurezas donadoras, $N_d = 10^{14} \text{ cm}^{-3}$, supuesto que todas están ionizadas a la temperatura ambiente, se tendrá para la concentración de electrones: $n = 10^{14} \text{ cm}^{-3}$ y para la de huecos: $p = n_i^2/n = (1.45 \times 10^{10})^2 / 10^{14} = 2.1 \times 10^6 \text{ cm}^{-3}$. Así pues, a medida que aumenta la concentración de electrones la de huecos se reduce en la misma proporción. Desde el punto de vista físico este proceso se explica si se tiene en cuenta que una pequeña fracción de los electrones procedentes de los átomos donadores contribuye a saturar una parte importante de los niveles o estados vacantes, disminuyendo por tanto la concentración de huecos.

Los datos anteriores muestran como es posible variar la concentración de electrones y huecos en un semiconductor mediante la adición de una proporción relativamente pequeña de átomos de impureza. Efectivamente, en el caso presente la concentración de impurezas añadidas (10^{14} átomos de impureza cm^{-3}) representa una fracción despreciable frente a la concentración de átomos de silicio del material (5×10^{22}

átomos de silicio cm^{-3}). Aún así, la concentración de electrones aumenta sobre la de huecos en varios ordenes de magnitud.

La ecuación (1.8) indica además que en un semiconductor extrínseco de tipo n por ejemplo, con una concentración elevada de electrones en la banda de conducción, el número de huecos en la banda de valencia será muy pequeño. Se dice entonces que los electrones son los *portadores mayoritarios* y los huecos son *portadores minoritarios*. En semiconductores de tipo p, en condiciones análogas, tendremos la situación opuesta.

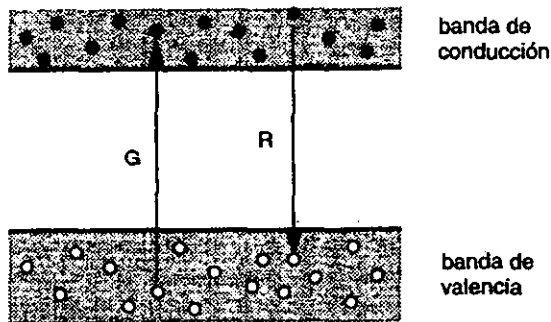


Fig. 1.11 Esquema de los procesos de generación G , y de recombinación R , de electrones y huecos en un semiconductor.

1.7 PROCESOS DE TRANSPORTE DE CARGA EN SEMICONDUCTORES.

Quizás lo más destacable de este capítulo es la descripción del comportamiento del semiconductor en términos de la "teoría de bandas", la cual permite explicar, al menos a nivel cualitativo, las características eléctricas de estos materiales. Sin embargo, los aspectos más cuantitativos de este problema, es decir, el número de cargas presentes en las bandas de valencia y de conducción y cómo varía su concentración con

la temperatura u otros factores externos, aún no han sido abordados. Es más, interesa conocer también cómo contribuyen los portadores de carga a los procesos de conducción o lo que, de una manera más general, se llama *procesos de transporte*.

1.8 CÁLCULO DE LA CONCENTRACION DE PORTADORES A LA TEMPERATURA AMBIENTE.

Antes de entrar en un cálculo detallado de la concentración de portadores y su variación con la temperatura es posible hacer un cálculo sencillo de su valor a la temperatura ambiente haciendo uso de la ley de acción de masas. Se tendrá el caso más general de un semiconductor extrínseco dopado a la vez con una concentración N_d de impurezas donadoras y N_a de impurezas aceptoras. Según se vio en el apartado 1.5, las impurezas donadoras cuando ceden un electrón quedan cargadas positivamente, mientras que las impurezas aceptoras cuando reciben un electrón quedan cargadas negativamente. A temperaturas muy bajas, inferiores a 100 K, solamente una fracción de esas impurezas se encuentran ionizadas. Si se le llama N_d^+ y N_a^- a la concentración de impurezas donadoras y aceptoras, respectivamente, que están ionizadas a una temperatura dada, la condición de **neutralidad de carga** exige que para esa temperatura se cumpla la relación:

$$N_d^+ + p = N_a^- + n \quad (1.9)$$

donde n y p es la concentración de electrones y huecos a esa temperatura. La ecuación anterior expresa que la concentración total de carga positiva ha de ser igual que la negativa, ya que en conjunto el semiconductor es neutro.

El cálculo de N_d^+ y N_a^- para cualquier temperatura es relativamente complejo. Sin embargo, cuando se trata de la temperatura ambiente normalmente todas las impurezas están ionizadas por lo que se tendrá entonces: $N_d^+ = N_d$ y $N_a^- = N_a$. La expresión anterior se reduce a:

$$N_d + p = N_a + n(T = T_{amb}) \quad (1.10)$$

aunque en esta ecuación los valores de p y n no tienen por qué coincidir con los de la ecuación anterior. Suponiendo n , la ecuación (1.10) junto con la (1.8) forma un sistema de dos ecuaciones con dos incógnitas que permite obtener separadamente los valores de n y p a la temperatura ambiente.

Para ilustrar mejor este cálculo se supondrá un semiconductor extrínseco tipo n en el cual $N_a = 0$. Imponiendo la condición de que N_d sea suficientemente grande para que $n \gg p$, de las ecuaciones (1.8) y (1.10) resulta:

$$\begin{aligned} n &\approx N_d \\ p &\approx \frac{n_i^2}{N_d} \quad (\text{tipo n}) \end{aligned} \quad (1.11)$$

Análogamente para un semiconductor tipo p, con $N_d = 0$, la condición $p \gg n$ lleva a:

$$\begin{aligned} p &\approx N_a \\ n &\approx \frac{n_i^2}{N_a} \quad (\text{tipo p}) \end{aligned} \quad (1.12)$$

En semiconductores compensados, es decir, dopados a la vez con impurezas de tipo donador y aceptor, ocurre una situación intermedia. Así, si $N_a = N_d$, el semiconductor se comporta como intrínseco con $n = p$, mientras que si $N_d > N_a$, con la condición adicional de que el valor $N_d - N_a$ sea mucho mayor que n_i , entonces n viene dado por la

diferencia $N_d - N_a$. Un resumen del cálculo de n y p a partir de la ley de acción de masas y del principio de neutralidad de carga para diferentes tipos de semiconductores viene dado en la tabla 1.1.

1.9 EFECTO DE LA TEMPERATURA EN LA CONCENTRACIÓN DE PORTADORES.

1.9.1 DISTRIBUCIÓN EN ENERGÍA DE LOS PORTADORES .

Los electrones y huecos en un semiconductor tienen un comportamiento estadístico, de forma que su movimiento dentro de las bandas de conducción y de valencia se realiza con un cierto intercambio de energía entre ellos. Como resultado de este intercambio de energía, el conjunto de electrones en cada una de las bandas se distribuye entre los distintos niveles de la energía. Sin embargo, esta distribución no es uniforme en las diferentes regiones de energía dentro de la banda, ya que hay determinadas zonas o intervalos de energía donde la concentración de electrones puede ser mayor que en otras regiones de energía. La razón es debida a la diferencia en la cantidad de niveles disponibles para el electrón en cada intervalo de energía.

Incluso, en un intervalo determinado, la probabilidad de ocupación en un intervalo puede ser mayor que en otro.

Tabla 1.1

CÁLCULO DE LA CONCENTRACIÓN DE PORTADORES EN DIFERENTES TIPOS DE SEMICONDUCTORES A LA TEMPERATURA AMBIENTE⁷

Semiconductor	Concentración de portadores	
tipo n	$n \approx Nd$	$p \approx n_i^2 / Nd$
tipo p	$p \approx Na$	$n \approx n_i^2 / Na$
compensado (*)	$Nd = Na$	$n = p = ni$
	$Nd > Na(*)$	$n \approx Na - Nd$ $p \approx n_i^2 / (Nd - Na)$
	$Na > Nd(*)$	$p \approx Na - Nd$ $n \approx n_i^2 / (Na - Nd)$

Supongamos un intervalo pequeño de energía, ΔE , centrado alrededor del valor E de energía dentro de la banda de conducción. Si $\Delta n(E)$ representa la concentración total de electrones cuya energía está comprendida en ese intervalo, la densidad de electrones por unidad de energía vendrá dado por el cociente $\Delta n(E)/\Delta E$. En el límite de intervalos infinitamente pequeños (pero conteniendo un gran número de niveles) este cociente tiende a un valor que se denominará $n(E)$. Así pues, el producto $n(E)dE$ representa la fracción de electrones cuya energía está comprendida entre los valores E y

⁷ Pag. 31 Ref. 17.

$E+dE$. La función $n(E)$ se denomina **función de distribución** de los electrones, ya que nos dice cómo se distribuyen los electrones en los diferentes intervalos infinitesimales de energía dentro de las bandas del semiconductor.

Para calcular la densidad de electrones $n(E)$ en un semiconductor basta conocer la densidad de niveles disponibles en ese intervalo de energía y contar cuántos de ellos están ocupados por electrones. Así pues, podemos escribir $n(E)$ simplemente como:

$$n(e) = Z(E) \cdot f(E) \quad (1.13)$$

donde $Z(E)$ es la *densidad de estados posibles* en la banda (de conducción o de valencia), es decir la concentración de estados o niveles por unidad de energía que existe para un valor E de la energía y que son susceptibles de ser ocupados por los electrones, y $f(E)$ representa la **probabilidad de ocupación** de esos estados, o dicho en otras palabras, la fracción de estados o niveles que se encuentran ocupados.

En semiconductores, al igual que en los metales, la probabilidad de ocupación de niveles por electrones está regida por la **estadística de Fermi-Dirac**⁸. Esta estadística se aplica a partículas que cumplen con el principio de exclusión de Pauli, como es el caso de los electrones. En la estadística de Fermi-Dirac, la función $f(E)$ está dada por:

$$f(E) = \frac{1}{1 + \exp\left[\frac{(E - E_f)}{kT}\right]} \quad (1.14)$$

siendo E_f un parámetro denominado **nivel de Fermi** [4], k la constante de Boltzmann y T la temperatura absoluta del sistema en consideración. La forma de la función $f(E)$, para distintas temperaturas incluida la del cero absoluto, está representada en la Fig. 1.12

⁸ Ver Ref. 19 pag. 168

para un sistema de electrones. En este sistema el nivel de Fermi se ha situado arbitrariamente en 2.0 e V. Conviene señalar que el nivel de Fermi en un semiconductor es una magnitud que depende mucho de las características del semiconductor así como de otros factores externos, tales como la temperatura.

Observando que, de acuerdo con (1.14) la función de probabilidad, $f(E)$, varía entre cero (estado vacante) y la unidad (estado ocupado). El nivel de Fermi, E_f , representa la energía para la cual la probabilidad de encontrar el electrón en ese nivel de energía vale 1/2. Para niveles con una energía inferior a la del nivel de Fermi la probabilidad es siempre superior a 1/2, mientras que para una energía superior, la probabilidad es menor que 1/2, decreciendo con la energía. La función de probabilidad es también dependiente de la temperatura. Así, a 0 K la probabilidad tiene un valor constante e igual a la unidad hasta la energía del nivel de Fermi. A partir de este valor la probabilidad disminuye abruptamente a cero. Esto quiere decir que a esta temperatura todos los niveles por debajo (o por encima) del nivel de Fermi se hallan ocupados (o vacíos) de electrones. A temperaturas superiores a 0 K la función de distribución tiene una variación más suave, aunque la variación más pronunciada aparece alrededor del nivel de Fermi en un intervalo de energía del orden de kT .

En analogía con la ecuación (1.13) se puede escribir para la *función de distribución de los huecos* en la banda de valencia, $p(E)$:

$$p(E) = Z(E)[1 - f(E)] \quad (1.15)$$

ya que un hueco representa un estado de energía vacante y por tanto la probabilidad de encontrar un hueco a la energía E vendrá dada por la función $[1 - f(E)]$.

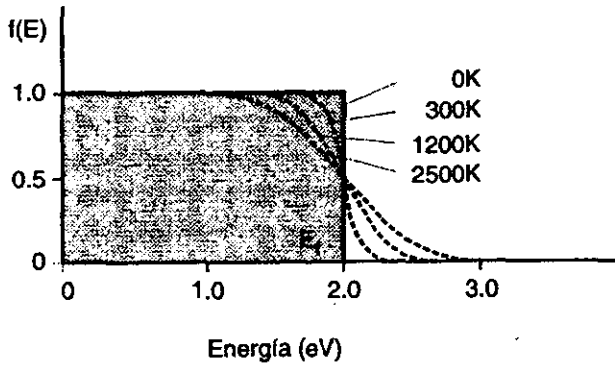


Fig. 1.12 Función de probabilidad de Fermi-Dirac, $f(E)$, a diferentes temperaturas, para un sistema electrones cuyo nivel de Fermi E_f está situado a 2.0 eV.

En analogía con la ecuación (1.13) se puede escribir para la *función de distribución de los huecos* en la banda de valencia, $p(E)$:

$$p(E) = Z(E)[1 - f(E)] \quad (1.15)$$

ya que un hueco representa un estado de energía vacante y por tanto la probabilidad de encontrar un hueco a la energía E vendrá dada por la función $[1 - f(E)]$.

El primer factor en las ecuaciones (1.13) y (1.15) representa, según se ha dicho, la densidad de estados o niveles posibles para el electrón. Esta función depende también de la función creciente proporcional a $(E - E_c)^{1/2}$ tal como se presenta en la Fig. (1.13a). Esto quiere decir que en el fondo de la banda de conducción no hay niveles susceptibles de ser ocupados por los electrones, mientras que a medida que subimos dentro de la banda la densidad de niveles aumenta de forma cuadrática con la energía. En los semiconductores intrínsecos y en los aislantes, $Z(E)$ puede ser una función bastante

compleja y difícil de calcular. Generalmente se admite, por extrapolación del caso de los metales, que en las zonas próximas al tope de la banda de valencia, E_v , y al fondo de la banda de conducción, E_c , $Z(E)$ varía según las ecuaciones:

$$Z(E) = \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_c)^{1/2} \quad (1.16)$$

$$Z(E) = \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_v - E)^{1/2} \quad (1.17)$$

para las bandas de conducción y valencia, respectivamente (en las expresiones anteriores h es la constante de Planck y m_e^* y m_h^* son las masas efectivas de los electrones y huecos, respectivamente, definidas en la sección 1.4.2. En la región intermedia de energía, zona de energía prohibida, si el semiconductor es intrínseco la función $Z(E)$ es cero en toda la región (fig. 1.14). En cambio si el semiconductor es extrínseco, la función $Z(E)$ puede adoptar valores diferentes de cero en las proximidades de los niveles aceptores o donadores.

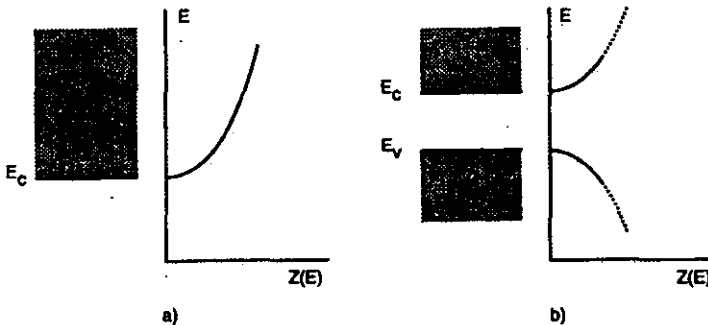


Fig. 1.13 Densidad de estados posibles, $Z(E)$, para los electrones en la banda de conducción de un metal (a) y en las proximidades de las bandas de valencia y conducción de un semiconductor intrínseco (b)

La Fig. 1.14 muestra de forma gráfica el cálculo de la función de distribución para los electrones en el caso de un metal (Fig. 1.14a), y para los electrones y huecos en el caso de un semiconductor (Fig. 1.14b), utilizando las ecuaciones (1.13) y (1.15). En este último caso el nivel de Fermi se halla situado generalmente entre las bandas de valencia y de conducción, mientras que en el caso de un metal el nivel de Fermi está situado dentro de la banda de conducción. Esta diferencia en la posición del nivel de Fermi es debida en cierto modo a que la banda de conducción del metal está mucho más poblada de electrones que la de un semiconductor.

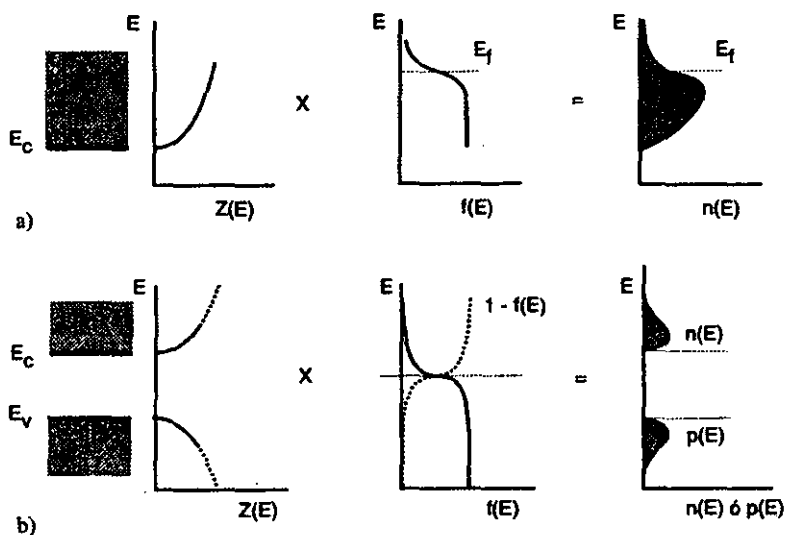


Fig. 1.14 Esquema del cálculo de la función de distribución de electrones, $n(E)$, en un metal (a) y de los electrones y huecos en un semiconductor (b).

Conocida la función $Z(E)$ es posible calcular la concentración de electrones en la banda de conducción y de huecos en la banda de valencia mediante integración de las ecuaciones (1.13) y (1.15):

$$n = \int_{E_c}^{E_{max}} Z(E) f(E) dE \quad (1.18)$$

$$p = \int_{E_{min}}^{E_v} Z(E) [1 - f(E)] dE \quad (1.19)$$

siendo E_{max} la energía del borde superior de la banda de conducción, y E_{min} la energía del borde inferior de la banda de valencia. En el supuesto de ser conocida la energía del nivel de Fermi se puede calcular la concentración total de los electrones o de los huecos a partir de las expresiones anteriores. En realidad el cálculo sólo puede llevarse a cabo bajo ciertas aproximaciones, ya que $Z(E)$ sólo se conoce en las proximidades del fondo de la banda de conducción y en el borde superior de la banda de valencia.

En la Fig. 1.15 se da un esquema de las funciones $Z(E)$ y $f(E)$, junto con la variación posible para $n(E)$ y $p(E)$, en diferentes tipos de semiconductores. En la columna de la izquierda, Fig. 1.15a, se muestran las curvas correspondientes a la densidad de estados posibles, función $Z(E)$. Observando que esta función es prácticamente la misma en todos los casos, excepto en la región de energías dentro de la banda prohibida, donde la función se hace cero si el semiconductor es intrínseco, o bien adopta valores finitos en ciertos niveles para tener en cuenta la existencia de estados donadores y aceptores (según sea el semiconductor n ó p) capaces de albergar electrones.

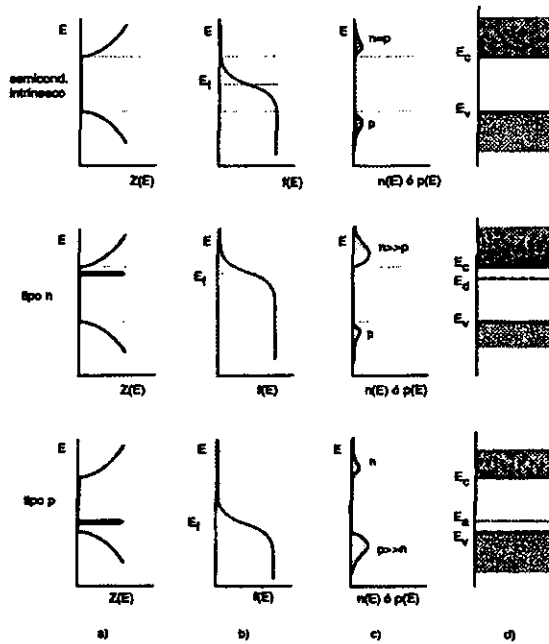


Fig. 1.15 Curvas representativas de: a) la densidad de estados posibles, b) la probabilidad de ocupación y c) la distribución en energía de los portadores en diferentes tipos de semiconductores.

La Fig. 1.15b representa, asimismo, la función de probabilidad $f(E)$ que es también similar en todos los casos excepto en lo que respecta al nivel de Fermi, cuya posición varía según el tipo de semiconductor. En semiconductores intrínsecos, E_f se encuentra aproximadamente en el centro de la banda prohibida. En los extrínsecos tipo n ó tipo p, este nivel se halla situado en las proximidades del fondo de la banda de conducción o del tope de la de valencia, respectivamente, aunque su posición exacta depende también de la temperatura. Finalmente, en la Fig. 1.15c se ha representado las funciones $n(E)$ y $p(E)$. Según se aprecia las funciones de distribución $n(E)$ y $p(E)$ toman valores finitos, hasta

un cierto máximo, por encima del borde de la banda de conducción y por debajo del tope de la banda de valencia (se supone en todos los casos que el semiconductor está a la temperatura ambiente), disminuyendo ambas funciones después del máximo a medida que se aleja de estos límites. Para un semiconductor intrínseco, $n(E)$ y $p(E)$ tienen la misma forma, con valores relativamente pequeños ya que el número de portadores generados térmicamente es muy bajo. En cambio las funciones $n(E)$ y $p(E)$ toman valores mucho más elevados cuando se trata de semiconductores extrínsecos. Una representación cualitativa de cómo se distribuyen los portadores para cada tipo de semiconductor se da en la Fig. 1.15d.

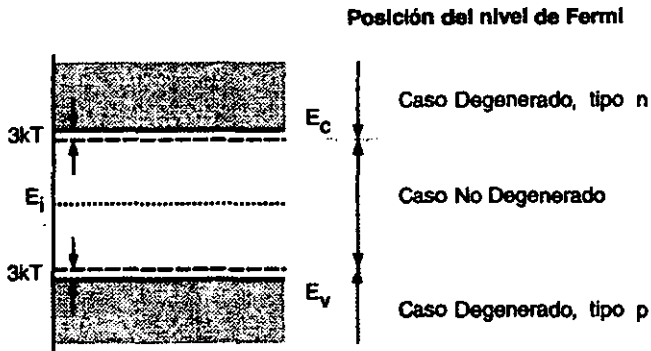


Fig. 1.16 Esquema de los diferentes casos que se puede presentar en un semiconductor según sea la posición del nivel de Fermi en relación con las bandas de energías. Se indica también la posición aproximada del nivel de Fermi intrínseco E_i .

1.9.2 CÁLCULO DE LA CONCENTRACIÓN DE PORTADORES.

Mediante la integración de las expresiones (1.18) y (1.19) es posible obtener una expresión general para la concentración total de portadores a cualquier temperatura (área de las curvas $n(E)$ y $p(E)$ en la Fig. 1.15c). El cálculo es relativamente complejo, aunque se puede llegar a una ecuación simple bajo ciertas aproximaciones. Así, por ejemplo, para energías superiores o inferiores al nivel de Fermi en $3kT$, la función de Fermi-Dirac se puede aproximar a una función exponencial. En estas condiciones es posible hacer un cálculo aproximado de las integrales (1.18) y (1.19) para obtener n y p , aunque obviamente el cálculo está limitado al caso en que el nivel de Fermi del semiconductor se halle en el interior del "gap" y alejado de los bordes de la banda de valencia y de conducción en una distancia en energía superior a $3kT$. Esto ocurre con frecuencia en numerosos dispositivos electrónicos, y se dice entonces que el semiconductor es *no degenerado*. La condición de no-degeneración implica necesariamente que la concentración de electrones y huecos en las bandas de conducción y valencia no sea muy elevada y se presenta siempre que el dopaje del semiconductor no sea excesivamente alto y que la temperatura no sobrepase un cierto límite (fig. 1.16). Dentro de esta aproximación, las integrales (1.18) y (1.19) dan como resultado:

$$n = N_c \exp\left[-\frac{E_c - E_f}{kT}\right] \quad (1.20)$$

para la concentración total de electrones en la banda de conducción, y

$$p = N_v \exp\left[-\frac{E_f - E_v}{kT}\right] \quad (1.21)$$

para los huecos en la banda de valencia. Los parámetros N_c y N_v se denominan **densidad efectiva de estados** en la banda de conducción y de valencia, respectivamente. Su valor viene dado por las expresiones:

$$N_c = 2 \left(\frac{2\pi m_e^*}{h^2} \right)^{3/2} \quad (1.21)$$

$$N_v = 2 \left(\frac{2\pi m_h^*}{h^2} \right)^{3/2} \quad (1.22)$$

Tanto N_c como N_v se pueden interpretar como el límite máximo de estados posibles en las bandas de conducción y de valencia, respectivamente, de forma que si todos estuvieran ocupados su número representaría la concentración máxima posible de electrones y huecos, en cada caso. En realidad, si el semiconductor se considera no degenerado esta condición no se puede alcanzar, ya que implicaría que el nivel de Fermi estaría situado en los bordes de la banda de conducción o de valencia, según el caso. A la temperatura ambiente N_c y N_v tienen un valor muy próximo y es del orden de 10^{19} cm^{-3} . En cualquier caso, los valores de n y p se refieren siempre a la concentración total de portadores, tanto si se trata de semiconductores intrínsecos como extrínsecos.

La utilización de las ecuaciones (1.20) y (1.21) requiere fundamentalmente el conocimiento de la energía del nivel de Fermi. Su valor está determinado por el contenido de impurezas del semiconductor y otros parámetros del material. Las expresiones (1.20) y (1.21) se pueden utilizar también en sentido inverso, es decir, para determinar la

posición del nivel de Fermi una vez que se conocen los valores de n ó p a través de la ley de acción de masas y el principio de neutralidad de carga (ecuaciones (1.8) y (1.9), respectivamente).

Con objeto de obtener expresiones alternativas de n y p es conveniente introducir el **nivel de Fermi intrínseco**, E_i , que es el nivel de Fermi que tendría el semiconductor si fuera intrínseco, es decir, con una concentración de portadores igual a n_i , el cual se halla situado aproximadamente a la mitad de la banda prohibida, según se esquematiza en la Fig. 1.15. Conocido el valor de E_i , la concentración de portadores intrínsecos, n_i , se puede calcular a partir de las expresiones anteriores, haciendo en (1.20) $E_f = E_i$, obteniéndose:

$$n_i = N_c \exp\left[-\frac{E_c - E_i}{kT}\right] \quad (1.23)$$

Al mismo tiempo, la expresión (1.20) también se puede poner de la forma:

$$n = N_c \exp\left[-\frac{E_c - E_i}{kT}\right] \exp\left[\frac{E_f - E_i}{kT}\right] \quad (1.24)$$

con lo cual, utilizando la relación (1.23), resulta:

$$n = n_i \exp\left[-\frac{E_f - E_i}{kT}\right] \quad (1.25)$$

El mismo procedimiento, usado para determinar la concentración de huecos lleva a:

$$p = n_i \exp\left[-\frac{E_i - E_f}{kT}\right] \quad (1.26)$$

Las dos expresiones anteriores son de mucha utilidad, ya que el nivel de Fermi intrínseco se emplea muy a menudo como nivel de referencia en un esquema de bandas

de energía. Esto ocurre sobre todo cuando se tienen semiconductores de la misma naturaleza pero con diferente dopaje. En estos casos, E_i es común a todos ellos, ya que su valor no depende del dopaje del semiconductor.

1.9.3 CONCENTRACIÓN DE PORTADORES INTRÍNSECOS.

Tratándose de semiconductores no degenerados, la concentración de portadores intrínsecos, n_i , se puede obtener fundamentalmente a partir de las ecuaciones (1.20) y (1.21), junto con la ley de acción de masas (ecuación (1.8)):

$$n_i^2 = np = N_c N_v \exp\left[-\frac{E_c - E_v}{kT}\right] \quad (1.27)$$

o bien, de forma más simple

$$n_i = (N_c N_v)^{1/2} \exp\left[-\frac{E_g}{2kT}\right] \quad (1.28)$$

siendo $E_g = E_c - E_v$ la energía de la banda prohibida. La ecuación (1.28) muestra que, independientemente del tipo de semiconductor, la concentración de portadores intrínsecos es una función exponencial decreciente de la anchura de la banda prohibida. La Fig. 1.17 muestra la variación de la concentración de portadores en un semiconductor intrínseco con el inverso de la temperatura en una escala semilogarítmica, para tres semiconductores típicos: silicio, germanio y arseniuro de galio. Notando que, de acuerdo con la expresión (1.28), en este tipo de representación la variación de n se convierte en una línea recta de pendiente negativa proporcional a E_g . Para el Si a temperatura ambiente resulta $n_i \approx 1.45 \times 10^{10} \text{ cm}^{-3}$. Esta concentración de portadores es muy baja cuando se le compara con la de los metales (del orden de 10^{22} ó 10^{23} cm^{-3}), por lo que el

silicio puro a temperatura ambiente puede ser considerado como un aislante. Algo similar ocurre con el arseniuro de galio, e incluso también con el germanio.

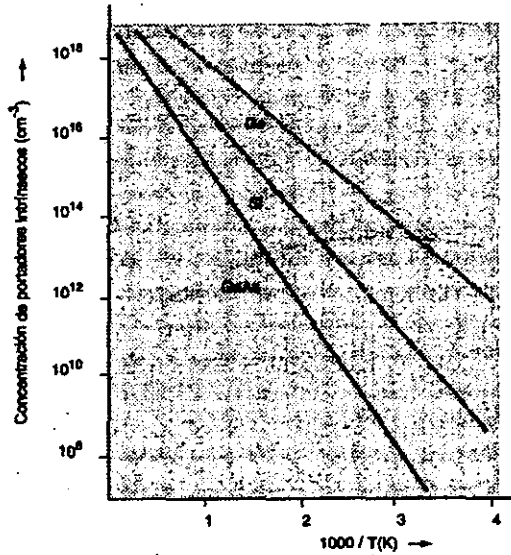


Fig. 1.17 Dependencia de la concentración de portadores intrínsecos con el inverso de la temperatura absoluta, para el germanio, silicio y arseniuro de galio.

1.10 DETERMINACIÓN DEL NIVEL DE FERMI EN UN SEMICONDUCTOR.

La determinación de la concentración de portadores en un semiconductor requiere el conocimiento de la posición de E_f . El valor de E_f puede obtenerse en cada caso concreto imponiendo la condición de neutralidad de carga, aunque sólo es posible obtener E_f de forma explícita dentro de ciertas aproximaciones.

Considerando el caso simple de un semiconductor tipo n: no compensado, es decir $N_a=0$. Desde un punto de vista práctico es conveniente distinguir tres rangos de temperaturas:

1. A temperaturas suficientemente bajas la concentración de portadores está controlada por el proceso de ionización de las impurezas, de forma que sólo una cierta fracción N_d^+ de los niveles donadores se encuentra ionizado. En este *rango*, denominado *extrínseco*, el cálculo de N_d^+ es complejo y por tanto es difícil conocer la posición del nivel de Fermi. Sin embargo, se puede demostrar que en este caso el nivel de Fermi se encuentra situado aproximadamente a media distancia entre el fondo de la banda de conducción y la energía de los niveles donadores, aunque el valor exacto depende entre otros factores de la concentración de impurezas, la temperatura, etc. En cualquier caso, la concentración de electrones en la banda de conducción se puede aproximar por la expresión:

$$n \approx \left[\frac{N_d N_c}{2} \right]^{1/2} \exp \left[-\frac{E_d}{kT} \right] \quad (1.29)$$

siendo E_d la energía de los niveles donadores medida desde el fondo de la banda de conducción. De acuerdo con esta expresión a medida que aumenta la temperatura crece la concentración de portadores de una manera casi exponencial como consecuencia de la excitación de electrones desde los niveles donadores a la banda de conducción.

Una vez calculado n a través de la ecuación anterior, la concentración de huecos minoritarios en la banda de valencia se obtiene mediante la ecuación (1.8).

Una ecuación similar a la ecuación (1.29) se puede desarrollar también para el caso análogo de un semiconductor tipo p. La ecuación (1.29) es comparable a la ecuación (1.28), válida para semiconductores intrínsecos. De hecho, los procesos de excitación desde los niveles de impureza son completamente análogos a los de excitación desde la banda de valencia a la de conducción, la única diferencia es que la energía necesaria para la excitación es mucho menor en el primer caso, ya que $E_d \ll E_g$. Debido a ello, es evidente que a temperaturas bajas la concentración de portadores en un semiconductor está determinada por la ecuación (1.29). En esta región de temperaturas una representación del logaritmo de n en función del inverso de T debe dar una recta de pendiente negativa y proporcional a E_d .

2. A temperaturas intermedias existe el llamado *rango de saturación* en el cual todas las impurezas se encuentran ionizadas siendo n esencialmente constante con la temperatura e igual a N_d (ecuación (1.11)). Esta característica hace que este rango de temperaturas, que puede extenderse entre 100 y 500°K, sea de gran utilidad cuando se desea controlar las propiedades de conducción de un semiconductor. La condición de $n = N_d$ permite además calcular de forma sencilla la posición del nivel de Fermi a partir de la ecuación (1.20). Se obtiene entonces:

$$E_c - E_f = kT \ln \left[\frac{N_c}{N_d} \right] \quad (1.30)$$

Igualmente, para un semiconductor tipo p. con $p = N_a$, se obtiene a partir de la ecuación (1.21):

$$E_f - E_v = kT \ln \left[\frac{N_v}{N_a} \right] \quad (1.31)$$

De estas expresiones se deduce que cuanto mayor sea la concentración de impurezas donadoras (o aceptoras) más próximo se halla el nivel de Fermi a la banda de conducción (o de valencia).

3. Finalmente, a medida que se eleva la temperatura por encima del rango de saturación, llega un momento en que la concentración de portadores intrínsecos se hace importante, de forma que su concentración puede alcanzar un valor sensiblemente superior al de los portadores puramente extrínsecos. En este rango de temperaturas altas, denominado *rango intrínseco*, la concentración de portadores, determinada por la ecuación (1.28), aumenta muy rápidamente con la temperatura. Del mismo modo, la posición del nivel de Fermi corresponde a la de un semiconductor intrínseco, E_i , y por tanto puede calcularse imponiendo en las ecuaciones (1.20) y (1.21) la condición de que en un semiconductor intrínseco el número de electrones en la banda de conducción ha de ser igual al de huecos en la banda de valencia, esto es $n = p$. Tendremos entonces a partir de estas ecuaciones:

$$E_f = E_i = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln \frac{N_v}{N_c} \quad (1.32)$$

Dado que N_c y N_v son del mismo orden de magnitud, la expresión anterior pone de manifiesto que el nivel de Fermi intrínseco se halla prácticamente en el

medio de la banda de energía prohibida mostrando una dependencia muy débil con la temperatura.

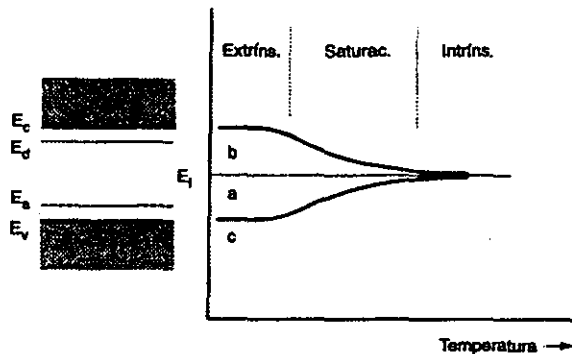


Fig. 1.18 Posición del nivel de Fermi en función de la temperatura para diferentes tipos de semiconductores: a) intrínseco, b) tipo n, c) tipo p.

En la Fig. 1.18 se presenta la variación del nivel de Fermi en los tres rangos de temperaturas indicados para distintos tipos de semiconductores. Como se puede apreciar, en el rango extrínseco el nivel de Fermi está situado a medio camino bien sea entre los niveles donadores y el fondo de la banda de conducción cuando el semiconductor es tipo n (curva b), o bien entre los niveles aceptores y el tope de la banda de valencia si es tipo p (curva c). A medida que aumenta la temperatura la variación del nivel de Fermi se rige por las ecuaciones (1.30) y (1.31), rango de saturación, las cuales predicen un descenso (o aumento) del nivel de Fermi si el semiconductor es de tipo n (o de tipo p). Finalmente cuando se alcanza el rango intrínseco a altas temperaturas el nivel de Fermi se aproxima al nivel intrínseco, E_i , es decir el nivel de Fermi que tendría el semiconductor si fuera intrínseco (curva a).

1.11 PROCESOS DE CONDUCCION EN SEMICONDUCTORES .

El transporte de carga en un semiconductor presenta más variedades que en un metal. Esto se debe a dos razones básicas:

1. Los semiconductores generalmente poseen dos tipos de portadores, los cuales se mueven en sentidos opuestos cuando se aplica un campo eléctrico, y
2. En un semiconductor puede haber fuertes variaciones locales en la concentración de portadores, las cuales a su vez dan lugar a desplazamientos locales de carga a través de un mecanismo de difusión.

Considerando un semiconductor en equilibrio térmico, a una temperatura T , sin aplicar un campo eléctrico. Los electrones y huecos poseen en estas condiciones una cierta energía cinética como consecuencia de su excitación térmica a los niveles de la banda de conducción o de valencia. La energía cinética de los electrones (o huecos) está determinada por la diferencia entre su energía total y la del borde de la banda de conducción (o valencia, cambiada de signo). Debido a ésta energía cinética, los electrones y huecos se hallan en continuo movimiento a través de la red, interaccionando con los átomos. En los procesos de interacción con los átomos, los portadores pueden perder o incluso ganar cierta cantidad de energía, aunque en conjunto la función de distribución de los electrones y huecos, y por tanto su energía cinética media se mantiene constante. Para tener una idea del orden de magnitud de la velocidad media de los electrones, $\langle v \rangle$, se puede hacer un cálculo sencillo suponiendo que toda la energía térmica transferida por la red, $1/2 kT$ por cada grado de libertad, se convierte en energía

cinética de traslación. Para un electrón moviéndose con tres grados de libertad tendremos por tanto:

$$\frac{1}{2} m_e \langle v^2 \rangle = \frac{3}{2} kT \quad (1.33)$$

donde m_e es la masa efectiva para la conductividad, con un valor diferente a la masa efectiva que aparece en el cálculo de la densidad efectiva de estados (m_e^*). A la temperatura ambiente ($\cong 300^\circ\text{K}$) la ecuación anterior arroja un valor de $\langle v \rangle \cong 10^5 \text{ ms}^{-1}$ para los electrones moviéndose en la banda de conducción del silicio o del arseniuro de galio. Aunque la velocidad media puede parecer elevada, hay que tener en cuenta que los electrones cambian constantemente su dirección debido a las colisiones con los átomos de la red y con las impurezas del dopaje. Resulta así un movimiento completamente aleatorio que no supone un desplazamiento neto de carga (Fig. 1.19a). La distancia media entre colisiones se caracteriza por el **recorrido libre medio**, y el tiempo entre colisiones se denomina **tiempo libre medio**⁹, *tc*. Este último parámetro resulta del orden de 10^{-12} s .

Cuando se aplica un campo eléctrico E en una dirección determinada tenemos una situación diferente ya que el electrón está sujeto a una aceleración debida a una fuerza de magnitud qE en dirección opuesta al campo. Sin embargo, una parte de la energía ganada del campo eléctrico es cedida a la red cristalina debido a las colisiones de los electrones con los átomos que están en reposo. Como consecuencia de ello, los electrones experimentan en su conjunto un arrastre en dirección opuesta al campo, ya que a la velocidad $\langle v \rangle$, debida al movimiento térmico, se superpone ahora la velocidad

de arrastre, v_e (Fig. 1.19b). Se puede de nuevo hacer un cálculo sencillo de la velocidad media de arrastre de los electrones, $\langle v_e \rangle$, igualando el impulso mecánico del campo eléctrico entre dos colisiones, dado por $-q E t_c$, a la variación media de la cantidad de movimiento, $m_e \langle v_e \rangle$. Resulta así:

$$\langle v_e \rangle = -\frac{q t_c}{m_e} E \quad (1.34)$$

El factor de proporcionalidad entre $\langle v_e \rangle$ y E es una nueva magnitud, característica del movimiento considerado para los electrones, denominada **movilidad**, μ_e . Para el mecanismo descrito de colisión de los electrones con los átomos de la red, esta magnitud está dada por:

$$\mu_e = \frac{q t_c}{m_e} \quad (1.35)$$

De acuerdo con esta expresión, la ecuación (1.34) resulta:

$$\langle v_e \rangle = -\mu_e E \quad (1.36)$$

con una ecuación análoga (con signo positivo) para los huecos. Uno de los factores que más afecta el valor de la movilidad es la temperatura del semiconductor, ya que la temperatura produce un aumento de la amplitud de vibración de los átomos de la red, lo cual a su vez hace que el tiempo libre medio entre colisiones sea cada vez más pequeño. Esto da lugar a una disminución de la movilidad con la temperatura. Sin embargo, en semiconductores muy dopados la interacción de los electrones con las impurezas puede dominar sobre otros mecanismos de colisión produciendo una variación de t_c con la temperatura más compleja.

⁹ Cap. 5 Ref. 18.

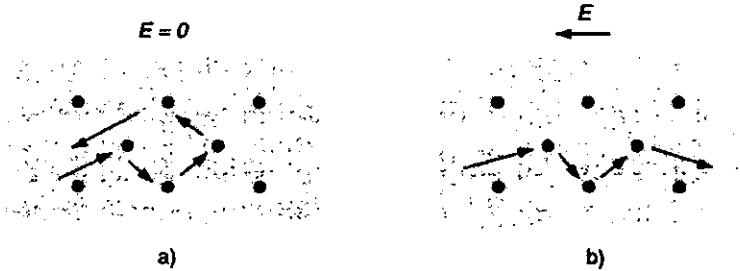


Fig. 1.19 Esquema del movimiento de un electrón en la banda de conducción a través de la red de átomos del semiconductor, a) en ausencia de campo eléctrico, b) con un campo eléctrico aplicado E .

Al estudiar el fenómeno de conducción, la presencia de un campo eléctrico constante en el interior del semiconductor normalmente es debida a la aplicación de una diferencia de potencial, ΔV , en sus extremos a través de sendos contactos metálicos (Fig. 1.20a). Esto implica que los electrones en el lado positivo tienen una energía potencial más baja, de valor igual a $-q\Delta V$, que en el lado negativo. En un diagrama de energías esta situación se representa por una inclinación uniforme de las bandas de energía del semiconductor en esa misma cantidad, tal como se indica en la Fig. 1.19b. Por contra, las bandas de energía de los contactos metálicos quedan inalteradas, ya que se supone que el campo eléctrico en el interior del metal es nulo y por tanto no existe caída de potencial. Dejando aparte el transporte de carga desde el metal al semiconductor, se puede visualizar en un esquema de energías el movimiento de arrastre de los electrones entre dos colisiones dentro del semiconductor mediante un desplazamiento horizontal dentro de la banda de conducción, implicando con ello una

ganancia en energía cinética a costa de una pérdida en energía potencial. En el proceso de colisión hay una pérdida de energía cinética (que se transfiere a la red en forma de calor dando lugar al efecto Joule), cayendo el electrón (en energía) prácticamente al fondo de la banda de conducción (flecha vertical). Después de esta pérdida de energía el electrón se acelera de nuevo ganando energía del campo hasta que sufre una nueva colisión. Para los huecos ocurre una situación similar aunque viajando en sentido opuesto y aproximándose tras cada colisión al tope de la banda de valencia.

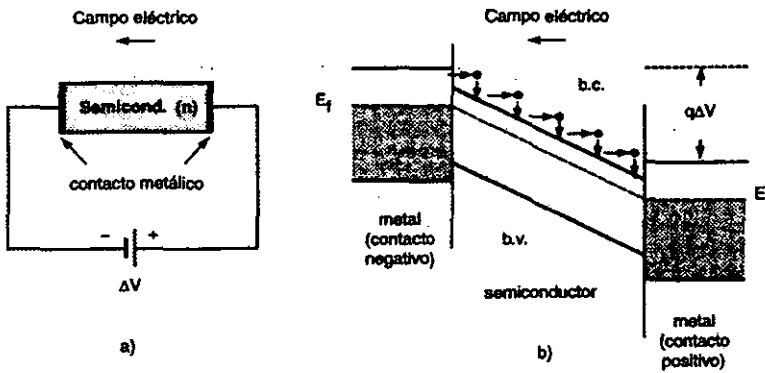


Fig. 1.20 Esquema del movimiento de portadores en un semiconductor: a) la aplicación de una diferencia de potencial en sus extremos, ΔV , da lugar a un desnivel de las bandas de energía en la cantidad $q\Delta V$. b) Movimiento de los electrones en la banda de conducción por la acción de la caída de potencial en el semiconductor.

Es importante mencionar que la velocidad de arrastre de los electrones puede ser muy diferente a la de los huecos. El desplazamiento de los huecos implica cambios en los enlaces interatómicos mientras que los electrones se mueven con relativa libertad dentro de la nube electrónica a través de los átomos. Por ello es de esperar, como ocurre en el caso del silicio, que la movilidad de los huecos sea mucho menor que la de los

electrones. Los valores experimentales de la movilidad obtenidos para los semiconductores típicos, Si, Ge y GaAs vienen recogidos, junto con otros datos característicos, en la tabla 1.2.

Para un conjunto de electrones desplazándose en el interior de un semiconductor con una velocidad media de arrastre determinada, $\langle v_e \rangle$, la densidad de corriente J_e , es decir la corriente eléctrica I_e por unidad de área viene dada por:

$$J_e = -qn\langle v_e \rangle = qn\mu_e E \quad (1.37)$$

Análogamente para los huecos, si su concentración es p y la movilidad μ_h , tendremos:

$$J_h = qp\mu_h E \quad (1.38)$$

La densidad de corriente total. $J = J_e + J_h$ debida a ambas contribuciones será por tanto:

$$J = J_e + J_h = (qn\mu_e + qp\mu_h)E \quad (1.39)$$

o bien:

$$J = \sigma E \quad (1.40)$$

siendo σ la **conductividad** del semiconductor definida por el factor entre paréntesis de la ecuación (1.39):

$$\sigma = qn\mu_e + qp\mu_h \quad (1.41)$$

La conductividad es también un parámetro que depende no sólo de las características del semiconductor, sino también de la temperatura. La concentración de portadores aumenta exponencialmente dentro de determinados rangos de temperatura. A su vez la movilidad puede disminuir con la temperatura, aunque no lo suficiente como para compensar el aumento debido a la concentración de portadores. Esto hace, en conjunto, que la conductividad de un semiconductor en el rango intrínseco aparezca

como una función creciente con la temperatura, justo al contrario de lo que sucede con los metales en los que, debido a que en ellos la concentración de portadores es constante, la conductividad disminuye con la temperatura.

Tabla 1.2

PARÁMETROS CARACTERÍSTICOS DE ALGUNOS SEMICONDUCTORES¹⁰

	Si	Ge	GaAs
Átomos/cm ³	5.0x10 ²²	4.4x10 ²²	4.4x10 ²²
Densidad (gr/cm ³)	2.3	5.3	5.3
Permitividad efectiva	11.9	16.0	13.1
Densidad efectiva. De estados (eV)			
N _c	2.8x10 ¹⁹	1.0x10 ¹⁹	4.7x10 ¹⁷
N _v	1.0x10 ¹⁹	6.1x10 ¹⁹	7.0x10 ¹⁸
Energía de la banda prohibida	1.12	0.66	1.43
Conc. Portadores intrínscos, T _{amb} (cm ⁻³)	1.5x10 ¹⁰	2.4x10 ¹³	1.8x10 ⁶
Masa efectiva			
m _e /m ₀	0.26	0.12	0.09
m _v /m ₀	0.38	0.23	—
Movilidad (cm ² V/s)			
μ _e	1.5x10 ³	3.9x10 ³	8.5x10 ³
μ _h	0.6x10 ³	1.9x10 ³	0.4x10 ³

1.12 PROCESOS DE DIFUSIÓN.

Los procesos de difusión son muy frecuentes en los semiconductores y ocurren cuando, por cualquier circunstancia, se produce una variación de la distribución portadores de un punto a otro en el interior del material. Así, por ejemplo, si se tiene el caso hipotético de un semiconductor extrínseco preparado con una distribución no homogénea de impurezas. Inicialmente la concentración de portadores, para cualquier

intervalo de energía, será mayor en aquellas zonas donde el contenido de impurezas sea también mayor, ya que a temperatura ambiente todas las impurezas estarán ionizadas.

Así, la variación local de la concentración de portadores de un punto a otro origina un proceso de difusión, mediante el cual los portadores de carga se mueven desde las zonas donde la concentración es mayor a aquellas donde es menor, tendiendo a igualar la concentración en todos los puntos del cristal. El fenómeno de la difusión tiene su origen en el movimiento aleatorio de los electrones dentro del semiconductor durante el cual están continuamente intercambiando su energía. Quizás, el símil más parecido es el de la difusión de un fluido a través de una membrana que separa dos recintos de diferentes concentraciones, o el caso de la difusión del calor de un punto a otro debida ala diferencia de temperaturas. De hecho, las ecuaciones que rigen el fenómeno de la difusión son muy similares en todos estos casos. Así, por ejemplo, si la concentración de huecos en un semiconductor es una función de la posición x en el interior del semiconductor, es decir $p=p(x)$, existe entonces una corriente de huecos cuyo valor por unidad de superficie, J_h , es proporcional a la derivada de la función $p(x)$.

1.13 CONSTANCIA DEL NIVEL DE FERMI.

Es importante mencionar que el proceso de difusión no implica cambios en la energía al del sistema siempre que el semiconductor se encuentre aislado y en equilibrio térmico. En consecuencia, la energía del nivel de Fermi, E_f se debe mantener constante en el interior del semiconductor. Por tanto, el desplazamiento de las bandas de energía da lugar a un corrimiento paralelo del nivel hasta que queda horizontal

¹⁰ Pag. 49 Ref. 17.

Este resultado es aplicable no sólo al caso de un semiconductor único sino a cualquier semiconductor formado por la unión de materiales de diferentes tipos, conductores o no, siempre que el sistema esté aislado y en equilibrio termodinámico. De hecho se puede establecer como un principio fundamental que **en condiciones de equilibrio el nivel de Fermi en todo el interior de un material aislado (o de un grupo de materiales en contacto), es decir, sin campo eléctrico aplicado, tiene un valor constante.** Este principio es importante ya que permite hacer el análisis del comportamiento del transistor MOS.

Evidentemente, si el sistema de electrones se encuentra en equilibrio térmico, niveles con la misma energía deben tener también la misma probabilidad de ocupación en cualquier punto del cristal. O a la inversa, niveles que tengan la misma probabilidad (como es el caso del nivel de Fermi, con probabilidad $1/2$) han de tener la misma energía, independientemente de la posición en el interior del semiconductor. Se puede, pues, concluir que **un sistema en el que la distribución de electrones se encuentra inicialmente fuera del equilibrio evoluciona hasta que el nivel de Fermi es constante en el interior del semiconductor.**

De acuerdo con estos postulados, se puede decir que la curvatura en las bandas de energía tiende a mantener constante en cada punto del cristal la probabilidad de ocupación de cada nivel de energía permitido.

1.14 TIEMPO DE VIDA DE LOS PORTADORES.

Si el semiconductor se encuentra en una situación de equilibrio térmico, en el cual la tasa de recombinación está siempre compensada por la de generación

mantiéndose constantes las concentraciones n y p . Sin embargo, no siempre se tienen situaciones de equilibrio ya que existen muchos procesos de generación de portadores que tienden a aumentar los valores de n y p sobre los de equilibrio térmico. Estos procesos de generación, son por ejemplo, los debidos a la excitación de portadores cuando el material se ilumina con radiación electromagnética de energía suficiente.

En estos casos por cada fotón absorbido se rompe un enlace y se crea un par electrón-hueco. También se puede crear un exceso de portadores sobre el valor de equilibrio cuando se ponen en contacto dos semiconductores de signo opuesto. A todos estos procesos que originan un exceso de portadores (Δn , Δp) sobre los valores de equilibrio (n_0 , P_0) se les denomina **procesos de inyección**. Los procesos de inyección son muy comunes en el funcionamiento de los dispositivos semiconductores. Frecuentemente, en los casos denominados de *baja inyección*, el exceso de portadores solamente modifica la concentración de portadores minoritarios mientras que la de los mayoritarios no se altera apreciablemente.

Para entender mejor el comportamiento de los portadores en los procesos de inyección consideremos el caso de un semiconductor tipo n formado por una placa de espesor pequeño, con objeto de que la radiación pueda penetrar en todo su interior sin sufrir atenuación. Cuando el semiconductor se ilumina durante un corto espacio de tiempo con radiación electromagnética de energía suficiente se origina en este tiempo un exceso de portadores, dn y dp , en las bandas conducción y valencia, respectivamente (Fig. 1.20a). Supondremos además que se trata de un proceso de baja inyección en el cual la concentración de portadores mayoritarios prácticamente no se altera mientras que

la de minoritarios aumenta sensiblemente ($\Delta p \gg P_0$). En el instante $t = 0$, cuando cesa la irradiación, el sistema tratará de volver al equilibrio compensando el exceso de huecos mediante procesos de recombinación. En estas circunstancias, podemos establecer, en primera aproximación, que la disminución con el tiempo de la concentración de huecos en cualquier instante, $p(t)$, es proporcional al exceso de huecos, Δp , en ese instante, según una ecuación del tipo:

$$\left. \frac{\partial p(t)}{\partial t} \right]_{recomb} = -\frac{\Delta p}{\tau_h} \quad (1.42)$$

con una ecuación similar para semiconductores tipo p. La constante de proporcionalidad que aparece en el denominador del segundo miembro, " τ_h " recibe el nombre de **tiempo de vida de portadores minoritarios**, huecos en este caso. Se puede demostrar en un análisis más riguroso que este tiempo de vida coincide con el tiempo medio que sobreviven los portadores después del proceso de inyección hasta que se alcanza el equilibrio por recombinación (es obvio que no todos los portadores duran el mismo tiempo y que unos desaparecen antes que otros).

Los tiempos de vida de los portadores minoritarios, μ_e y μ_h dependen mucho de las características del semiconductor y en particular de la concentración de átomos de determinadas impurezas, N_t , que se añaden intencionadamente para actuar como **centros o trampas de recombinación**. Cuanto mayor sea N_t menor es el valor de μ_e y μ_h . Sin embargo, la presencia de trampas de recombinación permite que una partícula sea capturada en esa posición permaneciendo en ella durante un cierto tiempo hasta que llega otra de signo opuesto y se efectúa la recombinación. Los tiempos de vida media,

que son del orden de 10^{-6} o incluso menores cuando se añaden trampas, juegan un papel fundamental en el tiempo respuesta de los dispositivos semiconductores.

CAPITULO 2

ESTRUCTURA METAL-ÓXIDO-SEMICONDUCTOR

2 ESTRUCTURAS METAL-AISLANTE-SEMICONDUCTOR.

Las estructuras metal-aislante-semiconductor, y en particular aquellas en las que el semiconductor es el silicio y el aislante su óxido, SiO_2 (denominadas **estructuras MOS**) constituyen la base de diversos dispositivos de integración a muy alta escala, entre ellos las memorias semiconductoras. El caso más general sería aquel en el que el material aislante no es el óxido de silicio sino cualquier otro material constante. A los dispositivos correspondientes se les denomina **MIS**. Uno de estos materiales es por ejemplo el nitruro de silicio, Si_3N_4 , que también presenta unas excelentes cualidades como aislante. Sin embargo, debido a su importancia será tratada en este capítulo la estructura MOS debido a que los transistores de efecto de campo basados en esta estructura (**MOSFET**) constituyen hoy en día uno de los dispositivos más importantes dentro de la microelectrónica.

2.1 LA ESTRUCTURA MOS IDEAL.

En la Fig. 2.1a¹ se presenta una estructura metal-óxido-semiconductor (MOS), formada por una sucesión de varias capas: la primera de ellas es un metal (electrodo superior), depositada sobre otra capa aislante de óxido de silicio, que a su vez recubre

¹ Ver Fig. 7.1, Ref. 18.

la superficie del semiconductor (el sustrato de silicio). La estructura se completa con una última capa metálica depositada sobre la cara inferior del sustrato de silicio formando un contacto óhmico con el silicio (electrodo inferior). En esta estructura, la capa de óxido se obtiene normalmente por oxidación directa del silicio a temperatura elevada. El espesor de la capa aislante suele variar entre unas decenas de nanómetros en los dispositivos actuales. Los electrodos de contacto, superior e inferior, sirven de contacto de salida al exterior del dispositivo cuando se aplica sobre ellos una tensión externa, generalmente con el electrodo inferior conectado a tierra (potencial de referencia).

El diagrama de bandas de la estructura MOS ideal en equilibrio (esto es, sin ningún voltaje externo aplicado) se muestra en la Fig. 2.1b para el caso de un semiconductor tipo p. Tratándose de un sistema formado por tres materiales diferentes es preciso elegir un nivel de referencia común para las bandas de energía, el cual será el nivel de vacío.

La estructura de la Fig. 2.1b se denomina ideal porque se refiere al caso en que las funciones de trabajo del metal, $q\phi_m$, y del semiconductor, $q\phi_s$, son iguales, lo que implica la igualdad de los niveles de Fermi respectivos, es decir $(E_f)_m = (E_f)_s$. Puesto que las funciones del trabajo del metal y del semiconductores son iguales, no existe transferencia neta de carga entre el metal y el semiconductor cuando se ponen en contacto a través del aislante. En el caso de la estructura MOS ideal se supone también que el óxido es un aislante perfecto por lo que no puede fluir carga a través de él, incluso cuando se aplica una diferencial de potencial en los electrodos de contacto. El

caso ideal implica asimismo que no hay cargas atrapadas en el óxido ni siquiera en la superficie óxido semiconductor.

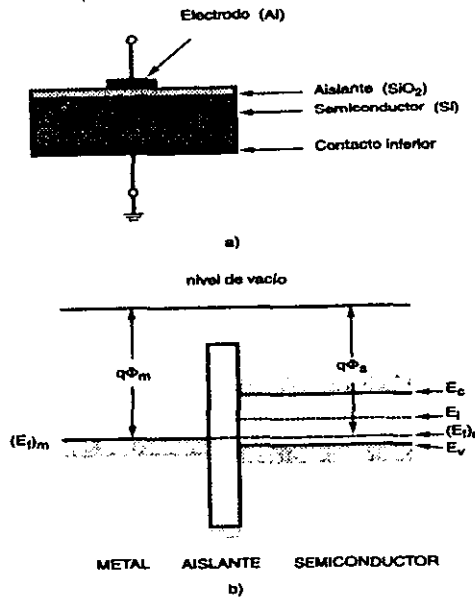


Fig. 2.1 a) Esquema de una estructura metal-óxido-semiconductor (MOS).
 b) Esquema de bandas de energía de la estructura MOS con un semiconductor de tipo p, en equilibrio, suponiendo que $\phi_m = \phi_s$.

Supongamos que se aplica un potencial V_G al electrodo metálico (denominado puerta o compuerta) situado sobre el aislante, con el contacto inferior conectado a tierra. Debido a que no puede existir flujo de carga a través del aislante, los niveles de Fermi del metal y del semiconductor quedan desnivelados en una cantidad igual a qV_G . Si se tratara de dos contactos metálicos aplicados directamente sobre el aislante el diagrama de bandas después de aplicar la tensión se tendría una caída de tensión constante al

pasar de un electrodo a otro. Sin embargo, la presencia del semiconductor, con un número limitado de portadores, hace que exista también una caída de tensión en el propio semiconductor. Debido a ello, las bandas de energía del semiconductor sufren una cierta curvatura sobre todo en la región de la superficie próxima del aislante. En cualquier caso, el nivel de Fermi en el semiconductor se mantiene constante en todo su interior, ya que los portadores se mantienen en equilibrio térmico en el interior del material.

En función del valor de la tensión externa aplicada se puede distinguir diferentes casos:

ii) Caso de acumulación:

Supongamos primeramente que la tensión aplicada el electrodo de compuerta es negativa, esto es, $V_G < 0$. Esto hace que el nivel de Fermi de metal $(E_f)_m$, suba en relación al del semiconductor $(E_f)_s$, en la cantidad qV_G . En este caso, los huecos del semiconductor (portadores mayoritarios) son atraídos hacia la superficie próxima hacia el aislante produciéndose una acumulación de carga positiva en la interfase. La acumulación de huecos en la superficie del semiconductor implica que en esta región la banda de valencia ha de estar más cerca del nivel de Fermi (cuya posición es constante) que en el interior. El resultado final es una pequeña curvatura hacia arriba de las bandas de energía del semiconductor en las proximidades de la superficie en contacto con el

aislante, tal como se muestra en la Fig. 2.2a². En esta situación se dice que la superficie del semiconductor esta en **acumulación de carga**.

Puesto que se debe mantener el balance de carga en la estructura, la carga positiva acumulada en la superficie del semiconductor, Q_s , debe estar compensada por otra carga igual y de sentido contrario (electrones) en la superficie del contacto metálico, Q_m . Debido a que en ambos casos se trata de carga libre (huecos y electrones) la acumulación de carga se produce dentro de una franja muy estrecha, a ambos lados del aislante, tal como se indica en la Fig. 2.2b.

iii) Caso de agotamiento:

Cuando se aplica un voltaje V_G positivo en el electrodo metálico ocurre lo contrario del caso anterior, es decir, el nivel de Fermi del metal $(E_f)_m$ baja en relación al del semiconductor, $(E_f)_s$. Sin embargo, cualitativamente es muy diferente a la interior, ya que en la interfase del semiconductor con el aislante se produce un desplazamiento de los huecos hacia el lado opuesto de la interfase del semiconductor dejando en las proximidades del aislante una región de carga espacial, formada por la carga negativa no compensada debida a los átomos de impurezas aceptoras que están ionizados. Como consecuencia de la disminución de la concentración de huecos esta región de carga espacial y la banda de valencia del semiconductor se encuentra en la superficie más alejada del nivel de Fermi que en el interior, por lo que las bandas del semiconductor se doblan hacia abajo en la forma indicada en la Fig. 2.2c. La región de

² Ver Fig. 1, cap. 5, Ref. 11

carga espacial es muy similar a la que se produce en la unión p-n, en la cual la carga negativa de los iones se distribuyen de modo prácticamente uniforme sobre una cierta distancia, w , a la interfase. La carga total por unidad de superficie (transversal a la estructura), Q_{agot} , contenida en la región de carga espacial vendrá dada por:

$$Q_{\text{agot}} = -qN_a w \quad (2.1)$$

donde N_a es la concentración de impurezasceptoras en el semiconductor. La carga negativa del semiconductor a su vez está compensada por otra carga igual y de sentido contrario acumulada en la superficie del metal Q_m (Fig. 2.1d). En esta situación y debido a la ausencia de carga libre en la región de carga espacial, se dice que la interfase del semiconductor con el aislante está en **agotamiento de carga**.

iv) Caso intrínseco:

Conforme el voltaje V_G se hace más positivo las bandas del semiconductor se doblan cada vez más hacia abajo, debido al aumento de la carga negativa sin compensar que dejan los huecos. Al mismo tiempo, y por efecto de la ley de acción de masas, la disminución de huecos da lugar a un aumento de la concentración de electrones (portadores minoritarios) en la superficie. Al principio (es decir, en la situación inicial de agotamiento) la proporción de electrones en la interfase es muy pequeña en comparación a la de los iones negativos, pero puede alcanzar una concentración apreciable cuando V_G aumenta por encima de un cierto valor. De hecho existe un potencial crítico, V_c , para el cual la concentración de electrones en la superficie, n_s , se hace igual a la concentración de portadores intrínsecos del semiconductor, n_i . Esto

ocurre cuando la curvatura de las bandas de energía es tal que el nivel de Fermi intrínseco en la superficie del semiconductor $E_{i, \text{sup}}$, coincide con el nivel de Fermi, E_f , en el interior del semiconductor y n_s vendrá dado por:

$$n_s = n_i \exp \left[\frac{(E_f - E_{i, \text{sup}})}{kT} \right] \quad (2.2)$$

por tanto, la condición de $E_f = E_{i, \text{sup}}$ implica $n_s = n_i$, según se ha comentado. En esta situación, la superficie del semiconductor es de **tipo intrínseco**.

v) Caso de inversión:

Si se sigue aumentando V_G , las bandas del semiconductor se curvan aún más hacia abajo y el nivel intrínseco en la superficie, $E_{i, \text{sup}}$, queda por debajo de E_f , tal como se indica en la Fig. 2.2e. Según se deduce de la ecuación anterior, la concentración de electrones en la superficie es entonces mayor que n_i . En cambio, la concentración de huecos en la superficie, p_s , será menor que n_i , ya que el valor de p_s varía de forma análoga a (2.2.) según la ecuación:

$$p_s = n_i \exp \left[\frac{(E_f - E_{i, \text{sup}})}{kT} \right] \quad (2.3)$$

Puesto que en este caso el número de electrones o portadores minoritarios en la superficie es mayor que el número de huecos o portadores mayoritarios en el interior se dice que la superficie del semiconductor está en **inversión de carga**.

vi) Caso de inversión fuerte:

Cuando la superficie del semiconductor está en situación de inversión, el aumento de la concentración de electrones en la superficie con el voltaje aplicado da lugar a que exista un cierto voltaje de umbral para el cual la concentración de electrones alcanza un valor igual a la de huecos mayoritarios en el interior del semiconductor. Se dice entonces que la superficie se halla en condiciones de *inversión fuerte de carga*, con una concentración de carga negativa en la superficie muy elevada.

Nótese que la carga de electrones, Q_{im} , debida a la inversión se acumula en una región muy estrecha (espesor entre diez y cien Å) en la superficie del semiconductor, tal como se esquematiza en la Fig. 2..2f. Ello es debido a que, en contraste a la carga fija de la zona de agotamiento, se trata en este caso de *carga libre* (electrones), la cual se encuentra sometida a la fuerza de atracción de carga de signo opuesto situada al otro lado del óxido, Q_m . Es más, en condiciones de fuerte inversión, un aumento en la tensión positiva aplicada al electrodo metálico produce un gran aumento de la concentración de electrones en la superficie, mientras que la carga espacial no aumenta apreciable. Esto se explica por el carácter exponencial de las funciones (2.2.) y (2.3), así como por el apantallamiento³ producido por los electrones en la capa de inversión, los cuales, una vez que alcanzan una concentración elevada, impiden una mayor penetración del campo eléctrico hacia el interior del semiconductor. ***Es entonces en estas condiciones de fuerte inversión, cuando la capa de agotamiento alcanza su anchura máxima, W_m .***

³ Ver Ref. 20.

Un aspecto verdaderamente notable de la condición de inversión fuerte es la presencia de carga libre, electrones, en una pequeña franja de la superficie en proporción igual o superior a la de huecos mayoritarios en el interior del semiconductor. Esta franja de la superficie, que adquiere así propiedades conductoras en cierto modo similares a las del interior del semiconductor, está limitada por dos regiones de mucho menor conductividad esto es, la capa aislante de óxido a un lado y la región de agotamiento del semiconductor en el lado opuesto. Sin embargo, los portadores de esta franja también denominada *canal* son electrones minoritarios en contraste al interior en el que los portadores son los huecos mayoritarios. ***El efecto más importante de la formación del canal en la superficie es la posibilidad de controlar a voluntad la concentración de electrones y por tanto la conductividad del canal a través del potencial aplicada en el electrodo de compuerta.*** Esta capacidad de control del proceso de conducción en la superficie con portadores de signo opuesto constituye la base del funcionamiento de los transistores de efecto de campo (MOSFET).

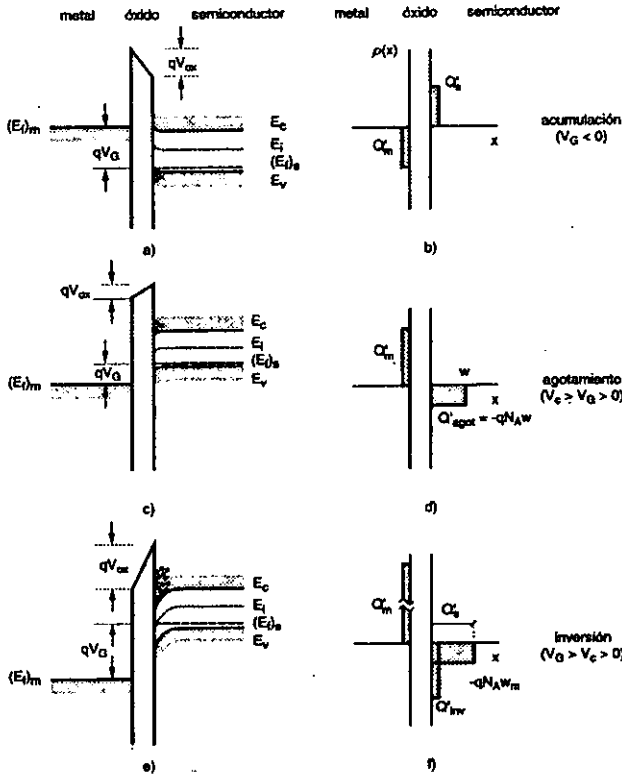


Fig. 2.2 Esquema de las bandas de energía y de las cargas acumuladas a ambos lados de la capa de óxido en una estructura MOS ideal, con semiconductor tipo p, en los casos de acumulación ($V_G < 0$), agotamiento ($V_G > V_G > 0$) e inversión ($V_G > V_G > 0$).

2.2.POTENCIAL DE SUPERFICIE.

Para describir el comportamiento del semiconductor se introduce un nuevo parámetro, ϕ_s , denominado **potencial de superficie**, el cual se define para un semiconductor tipo p a través de la ecuación:

$$q\phi_s = E_{i,vol} - E_{i,sup} \tag{2.4}$$

Según indica esta expresión el producto $q\phi_s$, representa el desplazamiento del nivel intrínseco de la superficie, $E_{i,sup}$, respecto de la posición del nivel intrínseco en el interior del semiconductor que se expresara ahora por $E_{i,vol}$.

Para el interior del semiconductor se introduce también el parámetro ϕ_i , mediante la ecuación:

$$q\phi_i = E_{i,vol} - E_f \quad (2.5)$$

El parámetro ϕ_i , conocido como **potencial de volumen**, tiene un valor fijo en todo el semiconductor y esta determinado únicamente por las características del material (concentración de impurezas). Dado que el nivel intrínseco está por debajo (encima) del nivel de Fermi en un semiconductor tipo n (tipo p), el valor de ϕ_i será negativo o positivo según sea el signo de los portadores. Se puede calcular el valor de ϕ_i , mediante las ecuaciones (1.25) y (1.26) (Capítulo 1) para la concentración de portadores en el interior de un semiconductor. Para un semiconductor de tipo p, por ejemplo, con $p = N_a$ se tendrá por ejemplo:

$$p = n_i \exp\left(\frac{q\phi_i}{kT}\right) = N_a \quad (2.6)$$

por tanto

$$q\phi_i = kT \ln\left(\frac{N_a}{n_i}\right) \quad (2.7)$$

O bien, si el semiconductor es de tipo n, con una concentración de impurezas N_d , tendremos:

$$n = n_i \left[\exp\left(\frac{-q\phi_i}{kT}\right) \right] = N_d \quad (2.8)$$

Por lo que en este caso:

$$q\phi_i = -kT \ln\left(\frac{N_d}{n_i}\right) \quad (2.9)$$

La Fig. 2.3 muestra los valores de ϕ_s y de ϕ_i para un semiconductor tipo p, en el cual existe una cierta curvatura de las bandas de energía en la superficie. Según se aprecia en esta figura, el valor de $q\phi_s$ es igual al doblamiento de las bandas de energía en la superficie del semiconductor cuando, por ejemplo, se aplica una tensión externa en las terminales de la estructura MOS.

En función de los parámetros ϕ_s y ϕ_i las ecuaciones (.2.2) y (2.3) toman la forma:

$$n_s = n_i \exp\left[\frac{q(\phi_s - \phi_i)}{kT}\right] \quad (2.10)$$

$$p_s = n_i \exp\left[\frac{q(\phi_i - \phi_s)}{kT}\right] \quad (2.11)$$

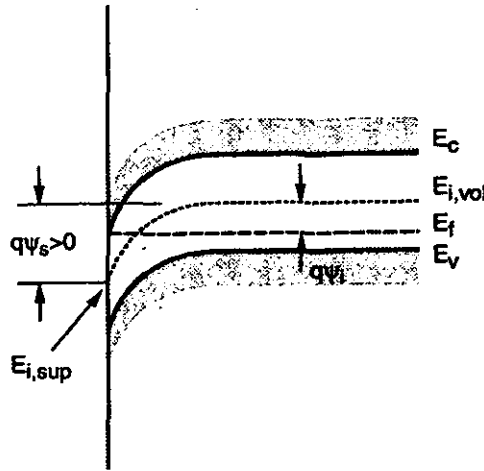


Fig. 2.3 Diagrama de la curvatura de las bandas de energía en una estructura MOS con semiconductor tipo p, indicando los valores Ψ_s y Ψ_b definidos de acuerdo con las ecuaciones (2.4) y (2.5). La situación que muestra la Fig. corresponde a la de inversión, con $\Psi_s > \Psi_b$.

De acuerdo con estas dos ecuaciones, cuando $\phi_s = \phi_i$, la concentración de electrones en la superficie es igual a la de los huecos y al mismo tiempo igual a la de portadores intrínsecos, n_i . Para $\phi_s < \phi_i$, se produce una situación de agotamiento de carga, en la cual la concentración de huecos en la superficie es menor que en el interior aunque se cumple $p_s > n_s$. Para $\phi_s > \phi_i$, las condición es de inversión, esto es, la concentración de electrones en la superficie supera a la de huecos $n_s > p_s$. Al aumentar ϕ_s , (debido al aumento de la tensión en el electrodo metálico) la concentración de electrones en la superficie puede alcanzar el valor de la concentración de huecos en el

interior del material, N_a . Esta condición corresponde a la de fuerte inversión la cual se presentará cuando φ_s cumpla la condición que se presenta en la ecuación (2.10):

$$n_i \exp\left[\frac{q(\varphi_s - \varphi_i)}{kT}\right] = N_a \quad (2.12)$$

o bien:

$$q(\varphi_s - \varphi_i) = kT \ln\left(\frac{N_a}{n_i}\right) \quad (2.13)$$

Comparando esta ecuación (2.7) resulta para la condición de inversión fuerte:

$$\varphi_s = 2\varphi_i \quad (2.14)$$

Es decir, el potencial de superficie ha de ser igual al doble del potencial φ_i . Al potencial necesario que es preciso aplicar al electrodo de compuerta para conseguir esta condición de inversión fuerte se le denomina **voltaje de umbral**, V_T . Evidentemente, si el potencial aplicado en el electrodo es todavía más positivo que el voltaje de umbral la concentración de electrones en la superficie aumentará por encima del valor de N_a .

Tabla 2.1

**COMPORTAMIENTO DE UNA ESTRUCTURA MOS IDEAL
DE SILICIO TIPO P ($\phi_s > 0$), AL VARIAR
LA TENSIÓN APLICADA AL ELECTRODO METALICO⁴**

Caso	Tensión	Situación	Concentración de portadores	Potencial de superficie
i)	$V_G < 0$	acumulación	$p_s > p_{vol}$	$\phi_s < 0$
ii)	$V_c > V_G > 0$	agotamiento	$n_s < p_s < p_{vol}$	$\phi_s > 0, \phi_s < \phi_i$
iii)	$V_c = V_G > 0$	intrínseco	$n_s = n_i = p_s$	$\phi_s > 0, \phi_s = \phi_i$
iv)	$V_G = V_c > 0$	inversion	$n_s > n_i > p_s$	$\phi_s > 0, \phi_s > \phi_i$
v)	$V_T = V_G > 0$	inversión fuerte	$n_s = N_a$	$\phi_s > 0, \phi_s = 2\phi_i$

En una estructura MOS con semiconductores tipo n, la aplicación de un voltaje al electrodo de compuerta conduce a los mismos casos que han sido estudiados anteriormente (acumulación, agotamiento, inversión, etc.). Sin embargo, cuando el semiconductor es de tipo n es preciso aplicar tensiones opuestas a las aplicadas en semiconductores tipo p. Así, por ejemplo, si la tensión aplicada la electrodo metálico es positiva se produce la acumulación de electrones en la interfase óxido-semiconductor. En esta situación las bandas del semiconductor se doblan hacia abajo en la región de la

interfase con el aislante. Así mismo, al aplicar tensiones negativas crecientes se suceden entonces las diferentes condiciones en la superficie del semiconductor: agotamiento, caso intrínseco, inversión e inversión fuerte, con un doblamiento progresivo de las bandas hacia arriba. En la tabla 2.1 se ha hecho un resumen de las diferentes situaciones que se pueden presentar en una estructura MOS con un semiconductor tipo p al variar el voltaje aplicado al electrodo metálico de compuerta, mostrando en cada caso el valor que adquiere el potencial de superficie.

2.3 CAPACITANCIA DE LAS ESTRUCTURA MOS IDEAL.

La estructuras MOS se puede considerar como un condensador capaz de almacenar carga en ambos lados del aislante, esto es, en el contacto de compuerta, Q_m , y en el propio semiconductor, Q_s .

2.3.1 CÁLCULO DE LA CARGA ACUMULADA.

En una estructura MOS ideal formada con un semiconductor tipo p, la aplicación de una tensión de compuerta negativa da lugar a la formación de una capa de carga superficial en el lado del metal (electrones) y de carga positiva en el lado del semiconductor (huecos). Si el semiconductor está suficientemente dopado, éste constituye una fuente inagotable de carga positiva por lo que la carga de los huecos acumulados en la interfase con el aislante será proporcional al voltaje aplicado entre los

⁴ Tabla 7.1, Ref. 18.

contactos. Este caso corresponde al de acumulación y la estructura se comporta como un condensador ideal.

Si el voltaje aplicado al electrodo de la compuerta es positivo la situación es diferente. En el caso de agotamiento, la carga espacial negativa en el lado del semiconductor procede fundamentalmente de los átomos aceptores que están ionizados, con una concentración igual a N_a . Al tratarse de carga fija su distribución se extiende en el interior del semiconductor hasta un espesor w dado por la ecuación (2.1). Se puede hacer en este caso el cálculo de la carga total por unidad de superficie, Q_{agot} , a través de la ecuación de Poisson⁵ utilizando el potencial de superficie φ_s , obteniéndose:

$$w = \left(\frac{2\varepsilon\varphi_s}{qN_a} \right)^{1/2} \quad (2.15)$$

siendo ε la constante dieléctrica del semiconductor. Mediante la ecuación (2.1) se tendrá para Q_{agot} :

$$Q_{agot} = -(2q\varepsilon N_a \varphi_s)^{1/2} \quad (2.16)$$

Estas dos expresiones muestran tanto el espesor de la región de carga espacial como la densidad de carga fija debida a los átomos aceptores ionizados aumenta con la raíz cuadrada del potencial de superficie.

Para valores positivos de compuerta suficientemente elevados, el semiconductor alcanza la situación de inversión fuerte. Esto ocurre cuando el potencial de superficie toma el valor de $\varphi_s = 2\varphi_i$. En estas condiciones, la carga negativa acumulada en la

⁵ Desarrollo de esta ecuación : Cap. 3, Ref. 18. y ecuación [4.5], Cap.4, Ref. 18.

región de superficie del semiconductor, Q_s , tienen dos componentes, esto es, la carga fija de agotamiento, Q_{agot} , y la carga de inversión propiamente, Q_{inv} , debida a electrones es decir:

$$Q_s = Q_{agot} + Q_{inv} \quad (2.17)$$

Al inicio de la inversión fuerte, la concentración de portadores minoritarios en la interfase misma del semiconductor con el óxido es muy elevada. Sin embargo, la concentración de electrones disminuye de forma muy abrupta al separarse de la interfase, esto es debido a la exponencial de la ecuación (2.10), por lo que la carga total de inversión es en estas condiciones muy pequeña frente a la carga de agotamiento. De esta manera se puede escribir la densidad de carga Q_s acumulada en la superficie del conductor en condiciones de inversión fuerte:

$$Q_s = Q_{agot}(inv) \quad (2.18)$$

Esta carga de agotamiento esta distribuida a lo largo de un cierto espesor, al cual alcanza un valor próximo a la anchura máxima, w_m , cuando la inversión es fuerte. En función de ϕ_s , el valor de w_m se puede obtener a partir de la ecuación (2.15):

$$w_m = \left(\frac{4\epsilon\phi_s}{qN_a} \right)^{1/2} \quad (2.19)$$

con ϕ_s , dado por la ecuación (2.7). Análogamente, para la carga de agotamiento en situación de inversión fuerte:

$$Q_{agot}(inv) = -(aq\epsilon N_a \phi_s)^{1/2} \quad (2.20)$$

A partir de estos datos es posible incluso determinar el potencial de umbral, esto es, el potencial de compuerta necesario para llevar el semiconductor a la condición de inversión fuerte. Teniendo en cuenta que la tensión V_G aplicada se reparte entre la caída de tensión en el óxido, V_{ox} , y la caída en la región de carga espacial del semiconductor, ϕ_s (ver la Fig. 2.4) se puede escribir

$$V_G = V_{ox} - \phi_s \quad (2.21)$$

A su vez V_{ox} , es proporcional a la densidad de carga (positiva), Q_m , en el contacto de compuerta a través de la ecuación

$$V_{ox} = \frac{Q_m}{C_{ox}} \quad (2.22)$$

siendo C_{ox} la capacidad por unidad de superficie asociada a la capa de óxido aislante. Por otro lado, Q_m es igual y de signo contrario a la densidad de carga acumulada en el semiconductor en la interfase con el aislante ($Q_m = -Q_s$). Según hemos visto, tanto en el caso de agotamiento como al inicio de la inversión fuerte la contribución más importante a la carga negativa del semiconductor es la carga fija en la región de agotamiento, de densidad Q_{agot} por lo que tendremos:

$$Q_m = -Q_{agot} \quad (2.23)$$

Utilizando la ecuación (2.16) para Q_{agot} resulta para V_G

$$V_G = \frac{(2q\epsilon N_a \phi_s)^{1/2}}{C_{ox}} + \phi_s \quad (2.24)$$

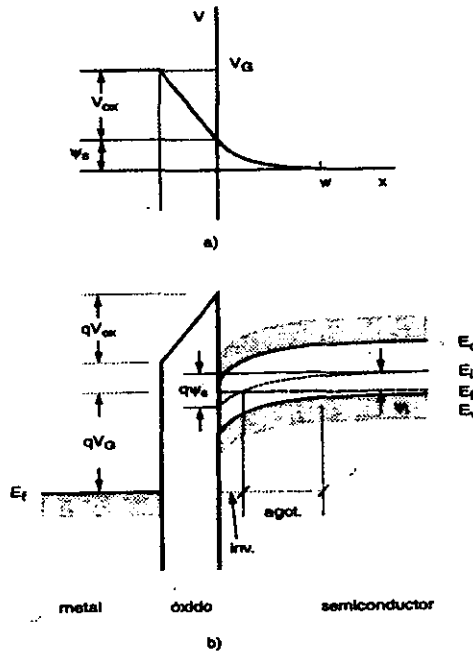


Fig. 2.4 Variación del potencial eléctrico en una estructura MOS tal y como la indicada en la parte inferior, cuando se aplica una tensión positiva V_G en el contacto de compuerta. El caso mostrado presenta una situación de inversión en un semiconductor tipo p.

Se llega así a un resultado importante, ya que esta expresión permite calcular el potencial de superficie, ϕ_s , en función del voltaje de compuerta V_G , tanto en la situación de agotamiento como al inicio de la inversión fuerte. El voltaje de umbral V_T , se obtiene haciendo $\phi_s = 2\phi_i$, en la ecuación anterior por lo que tendremos finalmente:

$$V_T = \frac{(4q\epsilon N_o \phi_i)^{1/2}}{C_{ox}} + 2\phi_i \tag{2.25}$$

con ϕ , dado por la ecuación (2.7). este resultado indica que el potencial de umbral varía inversamente con la capacidad de la estructura MOS. Así, si se pretende obtener un potencial de umbral bajo el espesor de la capa de óxido ha de ser lo más pequeño posible con objeto de obtener una capacidad elevada. A su vez, V_T está también determinado por las características del semiconductor, y para un material dado por su concentración de impurezas.

2.3.2 CAPACIDAD DE LA ESTRUCTURA MOS IDEAL: EFECTO DE LA TENSIÓN APLICADA.

De los resultados anteriores se desprende que en una estructura MOS polarizada en condiciones de agotamiento o de inversión no existe una relación lineal entre el voltaje aplicado, V_G , y la carga acumulada a uno y otro lado de la estructura debido a que una parte del voltaje, ϕ_s , cae en el propio semiconductor. Es más, la carga acumulada en cada caso depende estrechamente de los parámetros de la estructura. Por este motivo, la medida de la variación de la capacidad en función del voltaje (curvas V-C) de una estructura MOS permite obtener una gran información sobre la estructura interna de bandas y la distribución de carga tanto en el semiconductor como en la interfase óxido-semiconductor. Esta información se obtiene al comparar las curvas V-C experimentales con las que presenta una estructura MOS ideal. La medida de la capacidad⁶ se realiza superponiendo a la señal alterna utilizada por el puente o medidor

⁶ Ver "Técnicas experimentales para la obtención de las curvas C-V, Ref. 21

de capacidades una tensión continua V_G aplicada a los extremos de la estructura MOS. La señal alterna suele ser muy pequeña comparada con la tensión en continua, generalmente de unas pocas centésimas de volt y de frecuencia variable (hasta un mega Hz).

Supongamos una estructura MOS ideal formada por un semiconductor de tipo p a la que se le aplica un voltaje V_G negativo. En este caso la estructura estará en acumulación con una distribución de carga análoga a la de la Fig. (2.2b), es decir, con un exceso de portadores mayoritarios (huecos) en la interfase óxido-semiconductor. Un pequeño voltaje alterno superpuesto al continuo causará pequeñas variaciones de carga en la superficie metálica y en la del semiconductor por lo que la capacidad medida $C_{acum} = \frac{dq}{dV}$, será prácticamente la que corresponde a la del óxido aislante con espesor d . Esto ocurre incluso a las frecuencias más altas de la señal aplicada en la medida (alrededor de 1 MHz), ya que los portadores mayoritarios tienen una velocidad de respuesta muy rápida frente a las variaciones del campo eléctrico aplicado, alcanzando el equilibrio en un tiempo inferior a unos 10^{-10} s. Dado que el espesor d es normalmente mucho menor que las dimensiones laterales de la estructura MOS podemos utilizar para la capacidad la expresión correspondiente a un condensador plano-paralelo, es decir

$$C_{acum} \approx C_{ox} = \frac{\epsilon_{ox} S_d}{d} \quad (2.26)$$

donde ϵ_{ox} es la constante dieléctrica del óxido y S el área correspondiente al electrodo metálico de la compuerta. Este resultado es válido siempre que el voltaje de polarización sea negativo ya que la estructura MOS se encuentra en situación de acumulación. En

estas circunstancias podemos se puede decir que la capacidad del sistema es constante con la frecuencia e independiente del voltaje aplicado. Además su valor coincide con el que corresponde a la capacidad del óxido aislante. Esta región de voltajes de compuertas negativos corresponde al tramo AB de la Fig. 2.5 en la cual se ha representado la variación relativa de la capacidad (esto es del cociente C/C_{ox}) de una estructura MOS tipo p en función de la tensión aplicada en la compuerta.

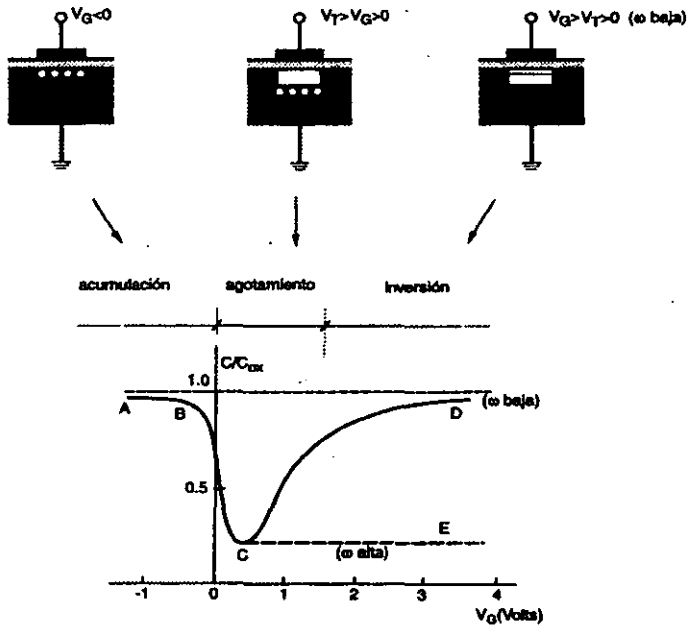


Fig. 2.5 Característica C-V normalizada para una estructura MOS tipo p. En la parte superior se muestra la distribución de la carga libre que contribuye a la capacidad en los diferentes rangos del voltaje aplicado al electrodo de compuerta.

Para valores de V_G positivos, es decir, en el rango de agotamiento (Fig. 2.2d), la carga fija del semiconductor, de signo negativo, Q_{agot} , está distribuida en una región de anchura w dada por la ecuación (2.1). En este caso, de nuevo son los portadores mayoritarios los que ensanchan o reducen la región de carga espacial alrededor de su anchura de equilibrio w , según varía la señal alterna superpuesta al valor de V_G . Esto quiere decir que para señales alternas la carga variable que entra en juego se halla a una distancia w de la capa de óxido. Así pues, el aislante equivalente que separa de las fluctuaciones de carga está compuesto ahora por dos capas, una de espesor d y constante dieléctrica ϵ_{ox} , que corresponde al óxido, y la otra de espesor w y constante dieléctrica ϵ correspondiente al semiconductor. Desde un punto de vista electrostático, la capacidad de estas dos capas es similar a la de dos condensadores en serie, uno de capacidad C_{ox} dada por (2.26) y otro de capacidad C_s , debida al semiconductor, cuyo valor es: $C_s = \epsilon S/w$. Por tanto, la capacidad de la estructura MOS será:

$$C_{agot} = \frac{C_{ox}C_s}{C_{ox} + C_s} = \frac{C_{ox}}{1 + \frac{C_{ox}}{C_s}} = \frac{C_{ox}}{1 + \left(\frac{\epsilon_{ox}}{\epsilon_d}\right) \frac{d}{w}} \quad (2.27)$$

Observando que puesto que la anchura de la capa de carga espacial w crece al hacerse V_G más positivo, la capacidad de la estructura MOS decrecerá conforme V_G se hace más positivo. Según se ve en la Fig. 2.5, la curva de variación de la capacidad con el voltaje presenta entonces una cierta disminución a partir del valor máximo dado por la ecuación (2.26) (tramo BC de la curva).

Cuando el voltaje V_G es suficientemente positivo el semiconductor se encuentra en condiciones de inversión fuerte y se crea una capa de electrones (portadores minoritarios) justo en la interfase con el óxido (Fig. 2.2f). Esta carga de electrones afecta profundamente la medida de la capacidad. Efectivamente, la capacidad del sistema depende ahora de la frecuencia de la señal alterna de medida, ya que los procesos de generación y recombinación necesarios para proveer o eliminar portadores minoritarios en respuesta a la señal alterna son relativamente lentos comparados con el tiempo de respuesta de los portadores mayoritarios. En el caso de que la frecuencia de medida sea muy baja ($\omega \rightarrow 0$), los portadores minoritarios tienen tiempo suficiente para ser generados o aniquilados por efecto de la señal alterna. Por tanto, la capacidad de la estructura MOS debe ser la misma que la que se obtiene en la región de acumulación aunque ahora la carga situada a ambos lados del óxido es de diferente naturaleza: carga positiva en el lado del metal, y carga negativa, fundamentalmente debida a electrones procedentes de la capa de inversión, en el lado del semiconductor. El valor de la capacidad tenderá a C_{ox} de nuevo, como se indica en la Fig. 2.5, en la curva de baja frecuencia (tramo CD). Se puede ver que la capacidad vuelve a crecer después de pasar por un valor mínimo correspondiente a la región de agotamiento.

Cuando la frecuencia de la señal alterna de medida de la capacidad es alta (en la región de MHz) las fluctuaciones de carga no afectan a la capa de inversión, ya que el tiempo necesario para formar o modificar la capa de inversión mediante procesos de recombinación puede ser del orden de centésimas de segundo, es decir, mucho mayor que el periodo de la señal de medida. Esto implica que la concentración de portadores

minoritarios en la capa de inversión permanece fija en el valor correspondiente al voltaje continuo V_G . Las fluctuaciones de carga se reducen en este caso únicamente a los portadores mayoritarios y aparecen alrededor de la terminación de la capa de agotamiento en el interior del semiconductor. De aquí que la capacidad, C_{inv} , obtenida tenga un valor análogo al deducido para el caso de agotamiento, ecuación (2.27):

$$C_{inv} = \frac{C_{ox}}{1 + \left(\frac{\epsilon_{ox}}{\epsilon_d}\right) w_m} \quad (2.28)$$

Donde w_m es la anchura máxima de la región de agotamiento dada por la ecuación (2.19). En esta expresión, el espesor de la capa tiene un valor máximo w_m , una vez establecida la inversión fuerte, el espesor de la capa de agotamiento permanece prácticamente constante. El valor de la capacidad a altas frecuencias toma en este caso de inversión un valor bajo tal como se muestra en la curva C-V en la Fig. 2.5 (tramo CE, a trazos).

De lo anterior se desprende que la determinación experimental de la variación de la capacidad con el voltaje aplicado en la compuerta constituye una herramienta fundamental para estudiar las características eléctricas de las estructuras MOS. Así por ejemplo, la posición del mínimo en la curva de baja frecuencia coincide con el voltaje de umbral V_T , de la estructura. Es preciso notar que, cuando se trate de un semiconductor tipo n la curva C-V es similar a la anterior pero con una forma simétrica a la curva de la Fig. 2.5 respecto al eje de las ordenadas. Esto es debido a que los voltajes necesarios

para polarizar la estructura en cada uno de los casos estudiados tienen signo opuesto a las que se utilizan en una estructura MOS formada por un semiconductor tipo P.

2.4 DESVIACIONES DEL COMPORTAMIENTO IDEAL.

De todas las estructuras MOS, la formada por un metal (generalmente aluminio), el aislante SiO_2 y el silicio como semiconductor es la más ampliamente estudiada por su importancia tecnológica. En esta parte del capítulo se verán las diferencias de comportamiento entre esta estructura y la estructura MOS ideal estudiada anteriormente. Estas diferencias son debidas fundamentalmente a los siguientes factores: i) diferencia entre las funciones de trabajo del metal y el semiconductor, ii) existencia de cargas eléctricas en el óxido, iii) la aparición de las denominadas "trampas" o defectos localizados en la interfase SiO_2 -Si.

i) Diferencia entre las funciones de trabajo

En la estructura MOS ideal se supone que las funciones de trabajo del metal y el semiconductor son iguales; sin embargo esto no ocurre siempre en la práctica. Así por ejemplo, el diagrama de bandas de energía de la Fig. 2.6 para el sistema Al- SiO_2 -Si(p) en el estado inicial, es decir, antes de alcanzarse el equilibrio termodinámico, se desprende de la función de trabajo del metal, $q\phi_m$, es menor que la del silicio tipo p, $q\phi_s$. Cuando se establece el equilibrio los niveles de Fermi se sitúan a la misma energía por lo que las bandas se doblan como se indica en la Fig. 2.6b, quedando la superficie metálica cargada positivamente y la del semiconductor negativamente. Es preciso hacer notar que para alcanzar el equilibrio tiene que existir en el sistema un medio de

intercambio de carga de un lado a otro del aislante. Este medio puede ser un conductor metálico conectado entre el metal y el semiconductor que facilite el traspaso de electrones desde el metal hacia el semiconductor o viceversa, hasta que se igualen los niveles de Fermi. El exceso de carga, positiva en el metal y negativa en el semiconductor en el caso que estamos considerando, se localiza justo en las interfaces metal-aislante y aislante-semiconductor, provocando a su vez una caída de potencial en el aislante, igual a $\phi_s - \phi_m$.

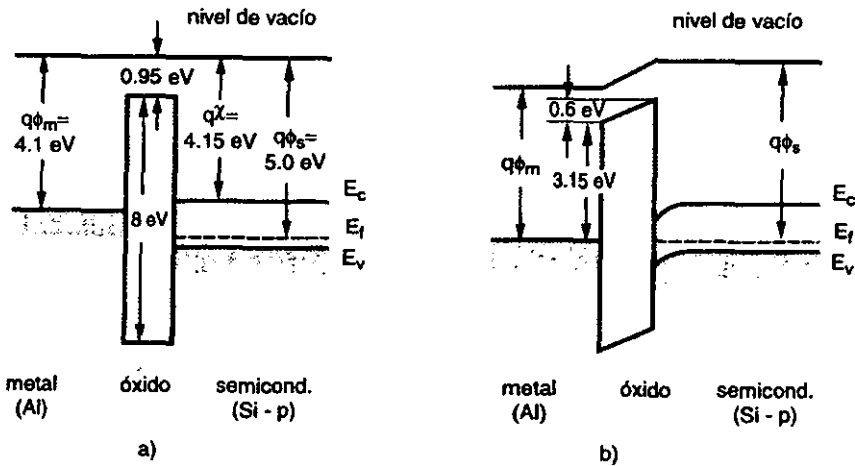


Fig. 2.6 Esquema de bandas de energía de una estructura MOS de aluminio, con silicio tipo p $N_a = 2 \times 10^{15} \text{ cm}^{-3}$; a) antes y b) después de alcanzarse el equilibrio termodinámico.

Para conseguir que las bandas de energía en el metal y el semiconductor estén planas de nuevo habría que aplicar una diferencia de potencial que compense exactamente la curvatura debida al efecto de la diferencia entre las funciones de trabajo. La carga almacenada en las interfases sería entonces cero y los campos eléctricos en el

óxido y en el semiconductor también se harían cero. Por este motivo, al voltaje aplicado para alcanzar esta situación se le denomina **voltaje de banda plana**, V_{FB} , y viene dado por

$$V_{FB} = \phi_m - \phi_s \quad (2.31)$$

Notando que cuando la estructura MOS se encuentra en la condición de banda plana no está en equilibrio. El efecto de las diferencias de funciones de trabajo del metal y el semiconductor sobre las curvas C-V ideal es una traslación a lo largo del eje de voltajes de magnitud V_{FB} como se indica en la Fig. 2.7. Ello es debido a que el sistema, una vez que se encuentre en la situación de banda plana, con un voltaje aplicado igual a V_{FB} , tiene un comportamiento igual a la estructura MOS ideal (Fig. 2.7).

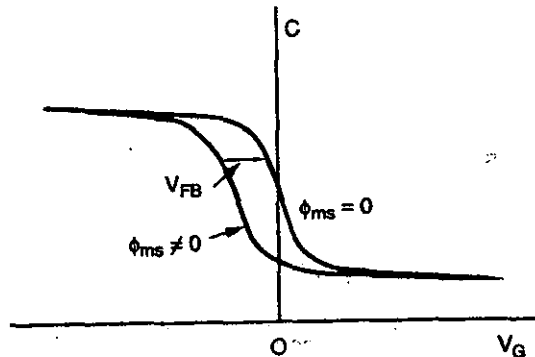


Fig. 2.7 Comparación de las curvas C-V típicas de alta frecuencia de una estructura MOS de silicio tipo p cuando $\phi_m \neq \phi_s$ ($\phi_{ms} \neq 0$) y $\phi_m = \phi_s$ ($\phi_{ms} = 0$)

ii) Cargas eléctricas en el óxido

Una causa importante de estabilidad en la curva C-V de una estructura MOS es debida a los iones móviles que pueden estar presentes como impurezas en el óxido. En particular, una de las impurezas que origina mayores problemas es el sodio (Na), ya que en forma de ion tiene una gran movilidad. Efectivamente los iones de sodio pueden trasladarse fácilmente a través del óxido por efecto de un campo eléctrico, incluso aunque el voltaje aplicado sea relativamente bajo. Este fenómeno da lugar a un desplazamiento de la curva C-V en el eje de voltajes que puede alcanzar en estas circunstancias varias decenas de volts.

La causa del desplazamiento de la curva C-V se debe a que la carga del ion induce una carga igual y opuesta repartida entre el electrodo metálico y el semiconductor. Esta ultima carga hace que el potencial V_{FB} varíe en una cantidad ΔV_{FB} , motivando el desplazamiento de la curva C-V. De hecho, para distinguir si el desplazamiento de la curva C-V es debido a impurezas tales como el sodio, frecuentemente se aplica un voltaje a la estructura y se calienta a temperaturas de unos cien o doscientos grados centígrados con objeto de dar más movilidad a los iones. Diversos métodos han sido desarrollados para neutralizar el efecto adverso de los iones Na^+ . Uno de los más populares consiste en difundir átomos de fósforo (P) en la superficie del óxido antes de la metalización, formando una fina capa de un vidrio denominado fosfosilicato, en el cual quedan atrapados los iones de sodio próximos a la superficie. El hecho de permanecer los iones de sodio cercanos al metal y alejados del semiconductor hace que no se induzca apenas carga en el semiconductor y por tanto no afecten a la curva C-V.

Aún en los casos de eliminación de las cargas móviles en el óxido la curva C-V puede mostrar una translación en el eje de voltaje respecto al caso de las curvas ideales el efecto se cree que es debido a cargas positivas fijas localizadas a unos 300Å de la interfase Si-SiO₂. Esta carga es independiente del espesor del óxido y del tipo de semiconductor sin embargo puede depender de las condiciones de oxidación utilizadas para crear la capa de óxido así como de la orientación de las superficie del silicio alcanzando la densidad de carga un valor de alrededor de 10¹⁰ veces la carga del electrón por cm² para la cara (100) y de 5x10¹⁰ para la cara (111). Parece Ser que el origen de la carga se debe a que, cuando se termina el proceso de oxidación del silicio para formar el SiO₂, algunos átomos de silicio de la interfase con el silicio quedan ionizados y sin reaccionar con el oxígeno. Si a la carga por unidad de superficie se le denomina Q_f su contribución, ΔV_{FB} , a la traslación de las curvas C-V viene dada por la expresión:

$$\Delta V_{FB} = \frac{Q_f}{C_{ox}} \quad (2.32)$$

siendo C_{ox} la capacidad por unidad de superficie asociada a la capa de óxido.

iii) Estados superficiales

Los **estados superficiales o trampas interfaciales** son estados electrónicos de energía distribuidos en la banda prohibida del semiconductor. El origen de las trampas superficiales se atribuye generalmente a los enlaces químicos no saturados que evidentemente tiene que existir en la superficie del semiconductor. Al igual que ocurría con la carga fija, la superficie (100) del silicio presenta una densidad de estado

superficiales alrededor de la tercera parte de la que exhibe la superficie (111). Una forma muy efectiva de reducir el número de estados superficiales de la estructura MOS consiste en calentar la estructura a unos 400°C en una atmósfera de hidrógeno. El hidrógeno se acopla a los enlaces de silicio no saturados dejándolos eléctricamente inactivos. Obviamente la carga por unidad de superficie acumulada en las trampas, Q_{it} causa, al igual que se vio para la carga fija del óxido, una traslación de la curva C-V en una cantidad ΔV_{FB} dada por Q_{it}'/C_{ox}' . El valor de Q_{it} para la orientación (100) del silicio puede ser tan bajo como 10^{10} veces la carga del electrón por cm^2 lo cual corresponde a alrededor de una carga atrapada por cada 10^5 átomos de la superficie del silicio.

CAPITULO 3

TRANSISTOR MOS

3 INTRODUCCIÓN.

Fue en 1930 cuando el investigador americano Julius Edgar Lilienfeld descubre que se puede gobernar la conducción en un cristal aplicándole un campo eléctrico perpendicular y registra varias patentes relativas a lo que hoy se considera como el primer transistor de efecto de campo. Ya para finales de los años 30 un físico de los laboratorios Bell, William B. Shockley se interesa por estudiar los dispositivos de estado sólido, y en particular por una teoría propuesta por Walter Schottky relativa al efecto de rectificación de la corriente alterna observando en una unión metal-semiconductor.

Shockley vislumbra de entrada la posibilidad de amplificar la señal a través de una zona de carga espacial o de empobrecimiento, que se extendería por el semiconductor, por debajo del metal. Sus primeros ensayos para realizar un dispositivo de este tipo, con cobre y óxido de cobre, en 1939, fueron infructuosos. Tras la segunda guerra mundial reanuda sus estudios acompañado por John Bardeen y Walter Brattain utilizando, germanio y crea en 1947 el primer transistor de puntas.

En 1948 Schockley y Gérald L. Pearson , en Bell Laboratories, observan la acción de un campo en una unión PN de silicio. Después en 1952, Schockley pública la teoría del transistor de efecto de campo.

El silicio comienza a ganar terreno al germanio; su gama de temperatura es mayor y su tratamiento más fácil. En 1960. Dawon Kahng y John Atalla. También de Bell Laboratories, proponen una estructura de silicio en la que un electrodo de control aislado provoca la creación de un canal conductor entre uniones PN.

Mas o menos al mismo tiempo, el investigador de RCA Dr. Paul K. VJeimer desarrolla la idea de elaborar funciones lógicas con la ayuda de MOS de película delgada de sulfuro y de seleniuro de cadmio. El primer circuito integrado MOS parece obra de Hofstein y Heiman, que trabajaban para los laboratorios de investigación electrónica RCA.

Se trata de una red de ocho pares de transistores de canal N, interconectados para formar una doble puerta de cuatro entradas.

Vendrá seguido por muchos otros que, tras la fase de ensayo del MOS, dieron lugar a toda una industria especializada orientada, desde 1970, hacia la lógica y las memorias, y más tarde hacia la electrónica lineal, la opto electrónica e incluso las hiperfrecuencias.

Sin embargo es en la industria de los circuitos integrados donde es importante el empleo de este transistor. Esto se debe a las siguientes consideraciones:

1. El MOS tiene una estructura más simple que el transistor bipolar. Por ello ocupa una superficie menor de superficie de la oblea de silicio.
2. Consume menos potencia que el bipolar. El consumo del transistor MOS puede medirse en microwatt, o incluso en nanowatt, mientras que en el transistor bipolar, se mide en miliwatt.

3. De donde resulta que presenta mayores densidades de Integración, como por ejemplo VLSI.
4. Menor superficie, estructura más simple y mayor densidad de integración contribuyen a hacer los circuitos integrados de transistores MOS más económicos que los bipolares.
5. Además, el MOS posee propiedades que no presenta el transistor bipolar, que le permite funcionar en régimen dinámico.
6. Finalmente, forma parte de los circuitos de transferencia de carga que no tienen equivalente en componentes discretos.

Estas razones fundamentales explican la penetración de los circuitos integrados MOS en el dominio de la informática (lógica y memorias), de las calculadoras de bolsillo, de la microinformática (con los microprocesadores), de las telecomunicaciones, de la instrumentación, de los automatismos, del reloj electrónico, etc.

3.1 EL TRANSISTOR METAL OXIDO SEMICONDUCTOR.

Algunos acrónimos comunes para el transistor MOS son:

- 1) MOST: Metal-Oxide-Semiconductor transistor (Transistor semiconductor Metal-Oxido).
- 2) MOSFET: MOS Field-Effect Transistor (Transistor MOS de Efecto de Campo).
- 3) IGFET-Insulated-Gate Field-Effect Transistor (Transistor Efecto de campo compuesto aislado).

El último nombre fue originado para distinguir el dispositivo del transistor de Unión Efecto de Campo, en el cual la compuerta está separada del resto de la estructura por una unión p-n. Alrededor de los tres acrónimos, IGFET es el más general dado que no especifica el material usado para la compuerta o el aislador. Sin embargo, MOST y MOSFET han venido a significar lo mismo que el IGFET y no implica que el metal, o el dióxido de silicio sean necesariamente usados para la compuerta y aislador¹.

3.2 MODOS DE FUNCIONAMIENTO DE LOS TRANSISTORES MOS.

El MOST tiene dos modos de funcionamiento:

1. De deserción.
2. De acrecentamiento.

En el modo de deserción, existe ya un canal conductor con polarización nula de la compuerta, al polarizar en sentido inverso a la compuerta, se rechazan los portadores presentes en el canal y reduce su conductancia; además, la aplicación de una polarización directa a la compuerta da origen a un acrecentamiento de la concentración de portadores mayoritarios en el canal y al aumento consiguiente de la conductancia de éste. Los dispositivos que exhiben una conductancia de canal apreciable con polarización nula de compuerta se denominan MOST de deserción; los que no presentan conductividad apreciable en las mismas condiciones, se llaman MOST de acrecentamiento. Además, el MOST puede ser de canal n o de canal p.

¹ Pág. 431 Ref. 11

3.2.1 EL TRANSISTOR MOS DE DESERCIÓN.

En el MOST de deserción las regiones de fuente, drenaje y canal son todas del mismo tipo de conductividad, de modo que la corriente de drenaje resulta importante con polarización nula en la compuerta. La Fig. 3.1 ilustra la sección transversal de un MOST de deserción y canal n.

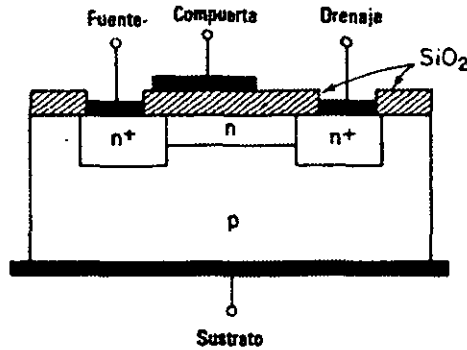


Figura 3.1 Sección transversal de un transistor MOS de deserción de canal n.

Puesto que el canal se forma por difusión de un sustrato de tipo p. Éste forma junturas p-n con las regiones de fuente y de drenaje, altamente impurificadas. Estas junturas equivalen, entonces, a un par de diodos conectados en oposición y en paralelo con el canal. Cuando la compuerta se hace positiva con respecto al sustrato, se atraen hacia el canal electrones adicionales, provenientes de la fuente y del drenaje, y aumenta la conductancia del canal; es decir, predomina la modalidad de acrecentamiento cuando la compuerta es positiva. En cambio si la tensión de compuerta es negativa, se repelen los electrones del canal, con lo que disminuye la conductancia de este y forma una

región de deserción de portadores en la vecindad de la interfaz oxido - sustrato. De esta manera se construye el canal. Si la compuerta se hace lo suficientemente negativa, la zona de deserción se profundiza y se une con la formada en la unión p-n del otro lado del canal. En esta condición, **la conducción entre drenaje y fuente cesa; la tensión de compuerta requerida para alcanzar esta situación es la tensión de estrangulación V_p . El fabricante especifica por lo general V_p como la tensión para la cual I_D se reduce a un valor determinado.**

3.2.2 EL TRANSISTOR MOS DE ACRECENTAMIENTO.

En la Fig. 3.2a se muestra la estructura básica de un transistor MOS de sustrato de silicio tipo p, desarrollada según la tecnología planar². Básicamente consiste en una estructura MOS en la cual el electrodo metálico superior, G, depositado sobre la capa aislante actúa como terminal de compuerta del transistor. Existen además dos regiones pequeñas dopadas fuertemente con impurezas donadoras, es decir de tipo n⁺, situadas a cada lado de la compuerta. Sobre cada una de estas regiones tipo n⁺ se deposita asimismo un electrodo metálico, formando el contacto de fuente S, y el de drenador D, del transistor. Finalmente, al igual que en una estructura MOS simple, sobre la superficie inferior del dispositivo se deposita una capa metálica que se mantiene conectada a tierra. En la Fig. 3.2b se muestra el símbolo del transistor MOS de canal n (como el de la Fig. 3.2a) en el cual la flecha indica el sentido convencional de la corriente (de

² Ver Ref. 22

drenador a fuente) en el modo normal de operación del transistor. En el MOST de canal p, la corriente tiene sentido opuesto y la flecha del dibujo lleva la dirección invertida.

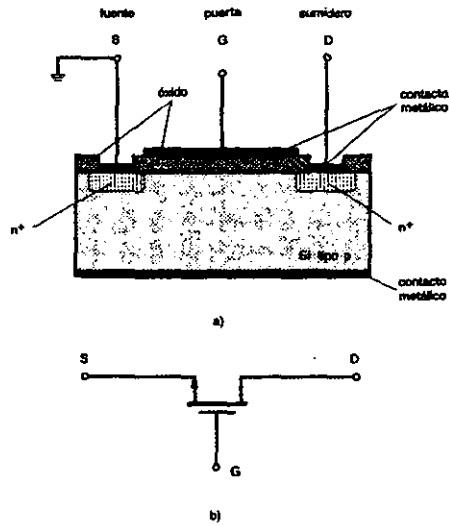


Fig. 3.2 a) Estructura típica de un MOST de canal n. b) Símbolo del MOST de canal n utilizado en circuitos.

Debido a que el canal en la Fig. 3.3 es formado por electrones, el dispositivo es llamado un transistor MOS canal n o transistor NMOS. Si el substrato fuera fabricado de material tipo n y la fuente y el drenador de material tipo p, tendríamos lo que es llamado un canal p o transistor PMOS.

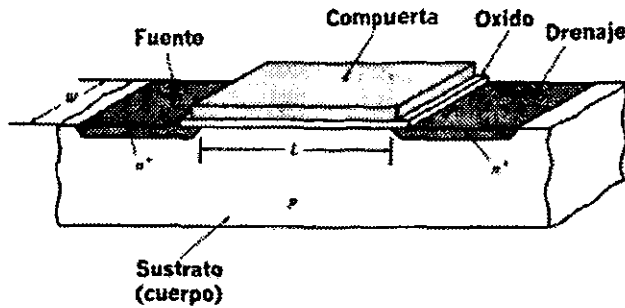


Figura 3.3 Estructura simplificada de un transistor MOS de canal n.

De esta manera, entre el drenaje y la fuente hay dos diodos conectados en oposición y la corriente de drenaje es esencialmente cero cuando la compuerta no está polarizada, independientemente de la polaridad de la tensión aplicada entre los diodos. Cuando la compuerta del dispositivo de canal n se hace ligeramente positiva con respecto al sustrato, los huecos de éste son repelidos por la compuerta y se forma, debajo de ésta, una capa de deserción (de huecos) entre la fuente y el drenaje. Si la compuerta se hace suficientemente positiva, comienzan a ser atraídos electrones libres hacia la compuerta y se forma un canal de tipo n, o capa de inversión en la interfaz óxido - sustrato. Puesto que este canal se extiende entre la fuente y el drenaje, la acción diódica entre éstos electrodos desaparece en la superficie del sustrato y puede haber ahora corriente de drenaje. **La tensión de compuerta necesaria para formar el canal recibe el nombre de tensión de umbral V_T .** Un incremento de la tensión positiva de compuerta por arriba del valor de umbral induce más electrones libres en el canal y acrecienta la corriente de drenaje.

3.3 DESCRIPCIÓN DEL FUNCIONAMIENTO DEL TRANSISTOR MOS.

Consideraremos primero el caso de que el voltaje aplicado a la compuerta es cero, es decir, $V_G = 0$. Las dos regiones de tipo n^+ de fuente y drenador forman con el resto del semiconductor de tipo p uniones p- n^+ conectadas en oposición, por lo que prácticamente no existe paso de corriente entre los electrodos de fuente y drenaje, cualquiera que sea el signo de la tensión aplicada entre ellos.

Supongamos ahora que aplicamos un voltaje positivo suficientemente alto a la compuerta para tener la condición de inversión fuerte en la interfase del semiconductor con el óxido. Esto quiere decir que $V_G \geq V_T$, siendo V_T el voltaje umbral de la estructura MOS. Los portadores minoritarios, electrones en este caso, dan lugar a la formación de un canal conductor de tipo n en la superficie del semiconductor entre la fuente y el drenador con una conductancia mayor cuanto mayor sea el voltaje aplicado en la compuerta. Este canal está limitado en su parte superior por la capa aislante de SiO_2 y en la parte inferior por la región de carga espacial que se forma en el semiconductor bordeando la compuerta y también alrededor de las regiones n^+ de los contactos de fuente y drenaje. En estas circunstancias si se aplica un voltaje positivo de valor pequeño al drenador (véase la Fig. 3.4a³), los electrones fluyen desde la fuente al drenador a lo largo del canal que actúa ahora como si fuera una resistencia de bajo valor. Se obtiene así, en esta región a bajos voltajes de drenador, una variación lineal entre la corriente I_D y el voltaje aplicado V_D .

³ Pag. 245 Ref. 18

Al aumentar V_D , manteniendo $V_D < V_G$, tanto la región de carga espacial que rodea al drenador como aquella que existe a lo largo del canal se ensancha, ya que el contacto inferior del dispositivo está a tierra. Debido a la progresiva caída de tensión desde la fuente al drenador, el ensanchamiento es mayor cuanto más próximos nos hallemos del drenador. Por otra parte, esta caída de tensión a lo largo del canal semiconductor hace que la diferencia de potencial efectiva entre la compuerta y el sustrato sea cada vez más pequeña a medida que se avanza hacia el drenador. Todo ello da lugar a una disminución del número de electrones que están presentes en la capa de inversión próxima al drenador, lo que equivale a su vez a una reducción, también progresiva, de la anchura del canal. Evidentemente, esta reducción es más acusada en la zona del drenador. El efecto global es una disminución de la pendiente en la curva de variación de I_D en función de V_D .

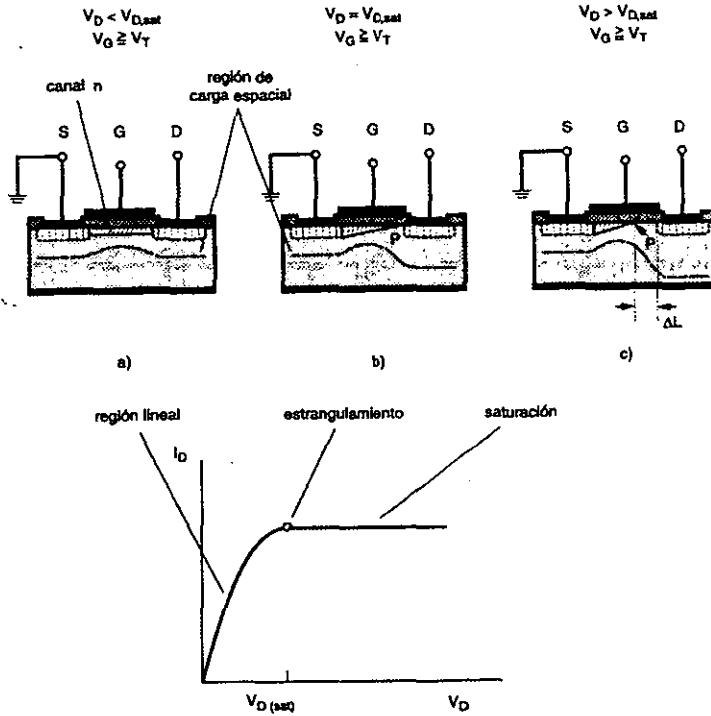


Fig. 3.4 Esquema de la variación de la anchura del canal al aplicar voltajes de drenador crecientes en un MOST de canal n. En la parte inferior se muestra la correspondiente curva característica $I_D - V_D$.

Quando se alcanza un voltaje, tal que la anchura del canal se reduce a cero en el drenador (Fig. 3.4b), se dice que ha ocurrido el estrangulamiento del canal (punto P en Fig. 3.4b). Esto ocurrirá para un voltaje denominado voltaje de saturación, $V_{D,sat}$, el cual ha de cumplir la relación $V_{D,sat} = V_G - V_T$. Para voltajes mayores, la región del canal estrangulada ΔL , aumentará de longitud en la dirección de

la fuente (Fig. 3.4c) y la corriente se mantendrá esencialmente constante, ya que el voltaje en el nuevo punto P de estrangulamiento se mantiene prácticamente igual a V_{Dsat} .

En la Fig. 3.5 se presentan las curvas experimentales I_D - V_D obtenidas al variar la tensión de compuerta V_D , en un transistor MOS de canal n. Obsérvese en esta figura que, para un voltaje V_D dado, al hacer más positiva la tensión V_G aplicada a la compuerta la corriente I_D se hace más elevada, ya que el canal es cada vez más ancho y por tanto el número de portadores más elevado. Hay que notar también que para conseguir que el canal sea conductor es preciso que V_G tenga un valor por lo menos igual al voltaje de umbral (alrededor de un volt) para la formación en la superficie de una capa con fuerte inversión. Evidentemente, en los transistores MOS de silicio de substrato tipo n el canal conductor es de tipo p y el potencial de compuerta ha de ser negativo; asimismo, la tensión V_D debe ser negativa en este caso. Por lo demás el comportamiento es exactamente igual al transistor MOS de canal n discutido anteriormente.

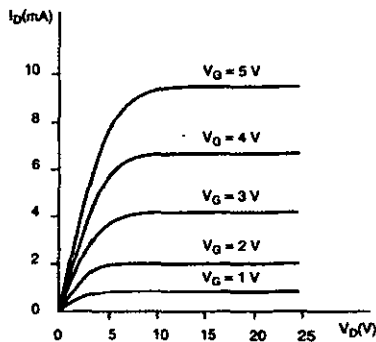


Fig. 3.5 Curvas I_D - V_D típicas de un MOST de canal n.

3.4 CÁLCULO DE LAS CARACTERISTICAS INTENSIDAD - VOLTAJE DEL TRANSISTOR MOS.

Con objeto de obtener una relación sencilla entre la corriente I_D y el voltaje V_D para un transistor MOST de canal n consideraremos el caso ideal, es decir, cuando no hay diferencia entre la función de trabajo del metal de compuerta y la del semiconductor, las cargas eléctricas en el óxido son nulas y no existen estados superficiales. Supondremos además que inicialmente aplicamos en la puerta una tensión V_G superior a la de umbral, esto es, $V_G > V_T$. Esto quiere decir que, incluso sin tensión aplicada en el drenador, el canal se encuentra en condiciones de inversión fuerte.

Consideraremos el caso en el que el transistor se polariza en la región cuasi-lineal de las curvas I_D - V_D . Esto implica que la tensión aplicada al drenador, V_D , (con la fuente a potencial cero) sea menor que la tensión de saturación, V_{Dsat} , es decir $V_D < V_{Dsat}$. Podemos suponer también que la anchura del canal es variable, disminuyendo en la dirección del drenador (en realidad lo que varía es la concentración de carga presente en cada punto del canal). Consecuentemente, la resistividad del canal debe aumentar al pasar de la fuente al drenador. Sabiendo que la corriente en cualquier punto del canal, I_D debe ser independiente de la posición, entonces:

$$dV = I_D dR \quad (3.1)$$

Siendo dR la resistencia de un elemento dx e una posición x a lo largo del canal, con una resistividad ρ variable. El valor de dR vendrá dado por:

$$dR = \rho(x) \frac{dx}{S} = \frac{1}{q\mu_n(x) z y(x)} dx \quad (3.2)$$

Donde $S = zy(x)$ representa el área de la sección transversal del canal de profundidad z y altura $y(x)$ variable (Fig. 3.6). El producto $qn(x)y(x)$ del denominador corresponde a la densidad de carga de inversión contenida en el canal en la posición x , $Q'_{inv}(x)$. Así pues podremos escribir:

$$I_D dx = z\mu_e Q'_{inv}(x) dV \tag{3.3}$$

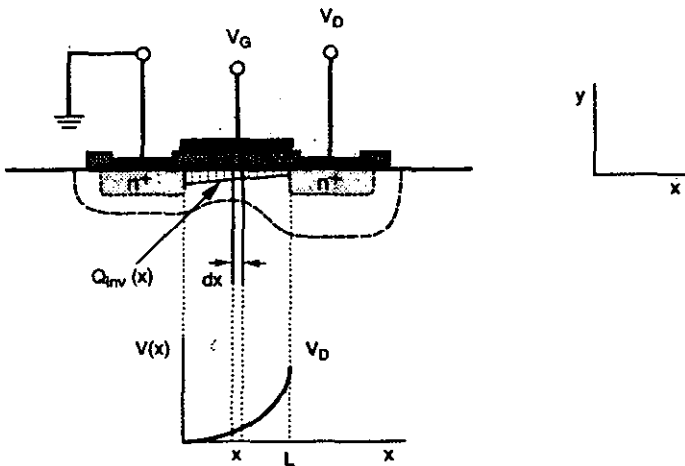


Fig. 3.6 Esquema de la variación de la carga de inversión , $Q'_{inv}(x)$, y de la tensión, $V(x)$, a lo largo del canal en un MOST, tipo n, polarizado en a región lineal ($V_D < V_{Dsat}$)

La carga por unidad de área en la zona de inversión del semiconductor, $Q'_{inv}(x)$, se puede expresar como:

$$Q'_{inv}(x) = Q'_s(x) - Q'_{agot}(x) \tag{3.4}$$

Donde $Q'_s(x)$ es la carga total en el punto de coordenada x de la interfase del semiconductor con el aislante y $Q'_{inv}(x)$ es la carga comprendida en la zona de agotamiento (ambas por unidad de superficie). Ahora bien en la situación de inversión Q'_s está directamente relacionada con la capacidad específica de la estructura MOS, C'_{ox} , a través de la ecuación: $Q'_s = -C'_{ox}V_{ox}$. En esta ecuación V_{ox} es la parte del potencial de puerta que cae en el óxido.

$$Q'_{inv}(x) = C'_{ox}[V_G - \Psi_s(x)] - Q'_{agot} \quad (3.5)$$

Siendo V_G el voltaje aplicado a la puerta y $\Psi_s(x)$ el potencial de superficie correspondiente al punto de coordenada x . Por otra parte, en una estructura MOS con semiconductor tipo p, dopado con una concentración de impurezas N_a , por lo que la expresión anterior resulta:

$$Q'_{inv}(x) = -C'_{ox}[V_G - \psi_s(x)] + [2q\epsilon N_a \psi_s(x)]^{1/2} \quad (3.6)$$

Si denominamos $V(x)$ a la caída de voltaje entre la fuente del MOST y un punto de abscisa x a lo largo del canal, podemos escribir para el potencial de superficie del semiconductor cuando está en inversión fuerte: $\Psi_s(x) \approx 2\psi_i + V(x)$, con ψ_i . Esta expresión de $\Psi_s(x)$ sustituida en (3.6) nos da para $Q'_{inv}(x)$:

$$Q'_{inv}(x) = -C'_{ox}[V_G - 2\psi_i - V(x)] + \{2q\epsilon N_a [2\psi_i + V(x)]\}^{1/2} \quad (3.7)$$

En la Fig. 3.6 se da un esquema cualitativo de la variación de la carga de inversión y del potencial a lo largo del canal en un transistor MOST de canal n polarizado en la región lineal.

Sustituyendo ahora en la ecuación (3.3) el valor de $Q'_{inv}(x)$ dado por la ecuación (3.7) e integrando ambos miembros de la ecuación anterior entre los límites $x=0$ y $x=L$, para $V=0$ y $V=V_D$, respectivamente, resulta finalmente:

$$I_D = \frac{z}{L} \mu_e \left\{ C'_{ox} \left(V_G - 2\psi_i - \frac{V_D}{2} \right) V_D - \frac{2}{3} (2q\epsilon N_a)^{1/2} \left[(2\psi_i + V_D)^{3/2} - (2\psi_i)^{3/2} \right] \right\} \quad (3.8)$$

En el análisis anterior se ha considerado la movilidad constante con un valor igual al valor medio de la movilidad de los electrones dentro del canal. Debido a que la anchura del canal es muy pequeña, los electrones de la capa de inversión interaccionan fuertemente con la superficie del óxido, disminuyendo su velocidad. Por esta razón la movilidad, μ_e , en el canal es más baja que en el interior del semiconductor, y además puede incluso depender de la anchura del canal.

En la Fig. 3.7⁴ se ha representado en trazo continuo las curvas características normalizadas, intensidad - voltaje de un MOST. Para pequeños valores de V_D la dependencia es de tipo lineal, mientras que para voltajes más elevados la pendiente disminuye hasta el valor cero, correspondiente al punto de abscisa $V_D = V_{Dsat}$. A pesar de las aproximaciones empleadas en el modelo, también en este caso se obtiene un buen acuerdo entre las curvas teóricas y las obtenidas experimentalmente (Fig. 3.5).

En la región lineal, para voltajes V_D mucho menores que $(V_G - V_T)$ la ecuación (3.8) puede aproximarse por:

$$I_D \approx \frac{z}{L} \mu_e C'_{ox} (V_G - V_T) V_D \quad (3.9)$$

⁴Pag. 441 Ref. 11

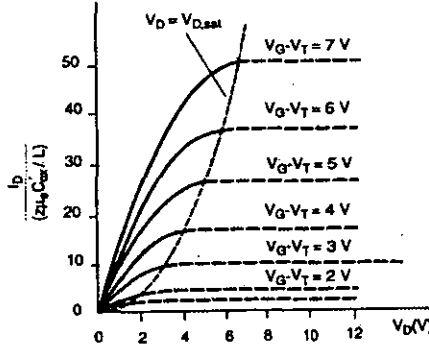


Fig. 3.7 Curvas características $I_D - V_D$ (normalizadas) para un MOST de canal n.

En esta región la conductancia del canal g_D , vendrá dada en forma aproximada por:

$$g_D = \left. \frac{\partial I_D}{\partial V_D} \right|_{V_G} \approx \frac{z}{L} \mu_n C_{ox} (V_G - V_T) \quad (3.10)$$

Y la transconductancia g_m :

$$g_m = \left. \frac{\partial I_D}{\partial V_D} \right|_{V_D} \approx \frac{z}{L} \mu_n C_{ox} V_D \quad (3.11)$$

Para $V_D = V_{D,sat}$, ocurre el estrangulamiento del canal en el drenador. Por encima de este voltaje, la corriente toma un valor constante, $I_{D,sat}$ (líneas de trazo discontinuo en la Fig. 3.7). En el valor de la corriente de saturación, $I_{D,sat}$, puede hallarse obteniendo primero el valor de $V_{D,sat}$ a partir de la ecuación (3.11), por:

$$I_{D,sat} \approx \frac{z\mu_e \epsilon_{ox}}{2dL} (V_G - V_T)^2 \quad (3.12)$$

Siendo una función que depende únicamente de la tensión aplicada en la puerta, V_G .

$$g_m = \frac{\partial I_D}{\partial V_D} \Big|_{V_D} \approx \frac{z\epsilon_{ox}}{dL} (V_G - V_T) \quad (3.13)$$

Las curvas de corriente de drenaje en función de la tensión de drenaje del MOST tienen las siguientes características; I_D varia linealmente con V_{DS} cuando ésta tensión es pequeña, y se satura, después, al aumentar la magnitud de V_{DS} . El canal es prácticamente independiente de V_{DS} cuando esta tensión es pequeña, pero se estrangula en el extremo próximo al drenaje cuando este se hace muy negativo. Puesto que V_T es la tensión entre compuerta y extremo del canal vecino al drenaje necesaria para mantener el canal, se produce una estrangulación cuando $|V_{GS}|$ disminuye hasta $|V_T|$ al aumentar $|V_{DS}|$. Es decir, hay estrangulación cuando,

$$|V_{DS}| \geq |V_{GS}| - |V_T|$$

Cuando V_{DS} (en un MOST de canal p) excede el valor de estrangulación la caída de tensión a lo largo del canal permanece prácticamente constante en el valor de estrangulación y el aumento de V_{DS} es absorbido por la juntura inversamente polarizada drenaje - sustrato.

3.5 MODELO CIRCUITAL DEL TRANSISTOR MOS.

El modelo de la Fig. 3.8 es valido para una amplia gama de frecuencias.

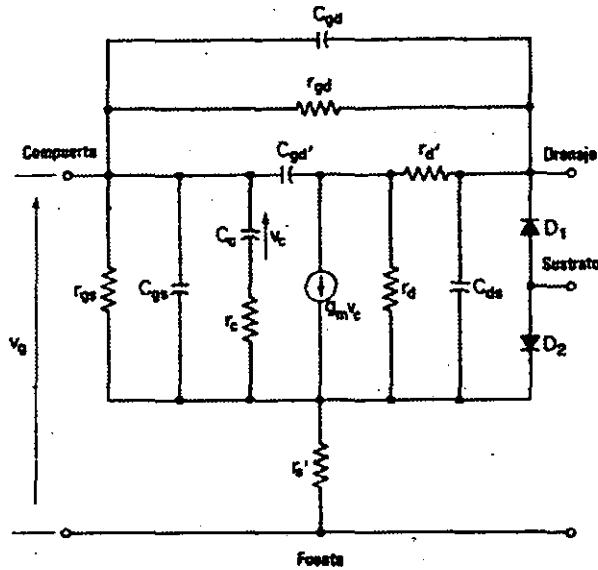


Figura 3.8 Circuito equivalente de fuente común del transistor MOS de canal n. En el dispositivo de canal canal p, se invierten los diodos D_1 y D_2

Los elementos de este circuito equivalente tienen los siguientes significados:

r_{gs} y r_{gd} . Estas resistencias representan las fugas que se producen a través y a lo largo de la capa de óxido desde la compuerta a la fuente y desde la compuerta al drenaje, respectivamente. Son por lo común muy altas y la resistencia de su combinación en paralelo es del orden de $10^{15}\Omega$ o más.

r_c y C_c . Estos elementos representan una aproximación concentrada de la resistencia del canal activo y de la capacitancia de la compuerta metálica. La tensión v_c a través de C_c realiza el control de cargas en el MOST. A baja frecuencia, v_c es igual a la tensión de entrada v_g . La resistencia del canal r_c es recíproca de la pendiente de las curvas $I_D - V_{DS}$ en la región triódica de las características.

g_m y r_d . Estos elementos son la transconductancia de compuerta a drenaje (admitancia de transferencia directa) y la resistencia dinámica de drenaje, respectivamente.

r_d r_s . Estas resistencias representan la porción de la resistencia del canal que no es modulada por la compuerta.

C_{gd} , C_{gs} y C_{ds} . Estos elementos representan las capacitancias respecto de la caja y entre terminales. Además, C_{ds} incluye las capacitancias relacionadas con los diodos D_1 y D_2 .

C_{gd} . Este elemento es la capacitancia intrínseca de compuerta a drenaje.

D_1 y D_2 . Estos diodos son los formatos por las junturas de drenaje y de fuente con el sustrato.

Sustrato. El sustrato forma una juntura pn con el canal y actúa como una segunda compuerta. Por esta razón debe ponerse a tierra para c.a. y a un potencial de c.c. que reduzca al mínimo la diferencia de potencial con respecto a la compuerta.

3.6 CIRCUITO EQUIVALENTE DEL TRANSISTOR MOS PARA SEÑALES PEQUEÑAS.

En los circuitos amplificadores de señales alternas basados en un MOST, el transistor se polariza en la región de saturación utilizando la configuración de fuente común. La señal alterna, v_g , que se pretende amplificar se introduce, superpuesta a la tensión de polarización de compuerta, V_G . En estos circuitos, las variaciones producidas en la tensión de puerta originan a su vez una variación en la corriente de drenador I_d , en la resistencia R_L introducida en el circuito de salida.

La Fig. 3.9a muestra el circuito equivalente de un MOST operando a baja frecuencia. La compuerta G se encuentra separada de los electrodos de la fuente y el

drenaje debido a la resistencia prácticamente infinita de la capa de óxido. En este sentido, el MOST es el transistor que representa mejores características en lo que se refiere a la resistencia de entrada, ya que al ser casi infinita, el dispositivo no consume corriente de la fuente de alimentación que proporciona la señal v_g . Nótese en el circuito equivalente que la señal de salida se obtiene a través de un generador de corriente de valor $g_m v_g$. Hay que señalar que, en lo que se refiere al circuito de salida, el MOST, al igual que otros transistores, se comporta como una fuente de corriente constante. Este hecho es debido a la horizontalidad de las curvas I_D - V_D en la región de saturación.

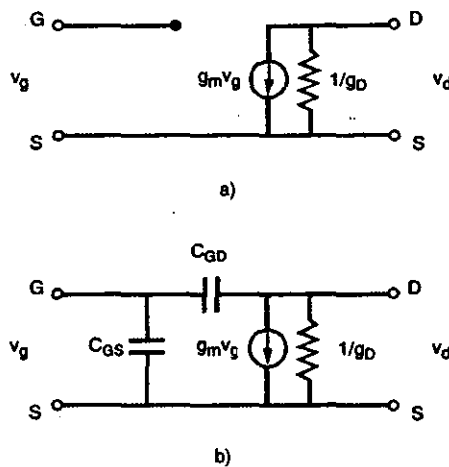


Fig. 3.9 a) Circuito amplificado simple de señales pequeñas para un MOST, de canal n, polarizado en la región de saturación. b) Circuito equivalente del transistor para baja frecuencia.

Cuando se trabaja con señales de alta frecuencia es preciso tener en cuenta las capacitancias parásitas entre el electrodo de compuerta y los de fuente y sumidero (C_{GS} Y C_{GD} en la Fig. 3.9b). Estas capacitancias parásitas pueden tener bastante importancia en el transistor MOST, ya que alteran profundamente el tiempo de respuesta del transistor.

3.7 EL TRANSISTOR MOS COMO INVERSOR.

La Fig. 3.10a muestra un MOST de deserción y canal n conectado como inversor. Cuando v_A es cero, la tensión de compuerta V_G reduce la corriente de drenaje prácticamente a cero, de modo que $v_0 = V_D$. Con v_A positiva, la corriente de drenaje que circula por R_L hace que $v_0 \ll V_D$. El hecho de que no se requiere corriente de compuerta para la conmutación a SI, implica que la carga impuesta a la fuente que suministra v_A es mucho menor, pues R_2 puede ser muy grande y es posible que un inversor de este tipo controle a una gran cantidad de otros circuitos lógicos en paralelo por medio de su salida v_0 . Es decir, los circuitos lógicos MOST tienen una gran capacidad de comando sobre otros circuitos similares, gracias a las menores exigencias de corriente de carga⁵.

⁵ Pag. 42 Ref. 15

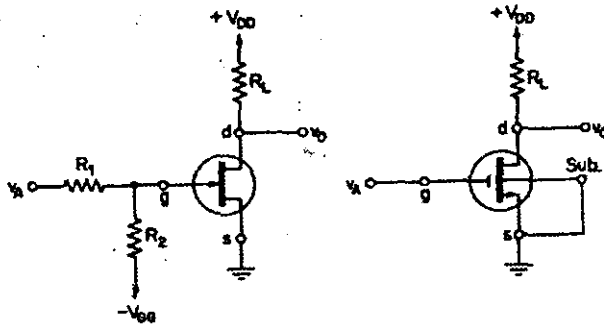


Figura 3.10 Circuitos inversores con transistores MOS. (a) Uso de un transistor MOS de deserción. (b) Uso de un transistor MOS de acrecentamiento de canal n.

La superioridad del inversor con MOST de acrecentamiento se debe a la ausencia de polarización de compuerta en el modo NO lo que permite el acoplamiento directo de distintas etapas sin necesidad de redes para desplazar los niveles de tensión. Esta ventaja es tan importante que casi nunca se usa el MOST de deserción en los circuitos digitales.

En la Fig. 3.11(a) los dos MOST son del mismo tipo general, aunque ellos deben ser elegidos de modo que Q_1 tenga una corriente de saturación de drenaje I_{DS} mucho mayor, para la tensión de compuerta dada (por ejemplo para $V_{GS} = V_D$); en la Fig. 3.11(b), en cambio, se utilizan dispositivos complementarios.

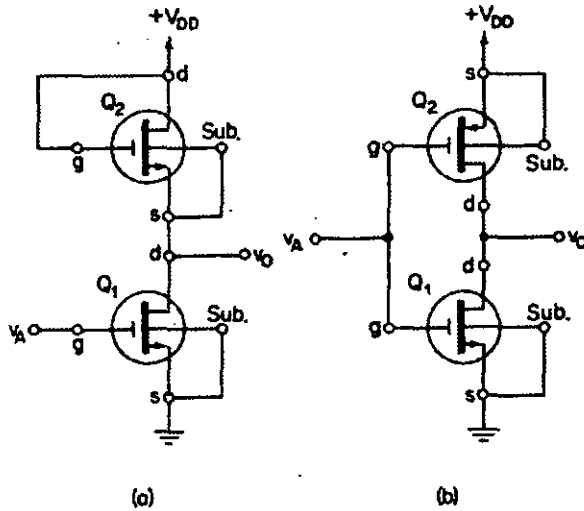


Figura 3.11 Circuitos inversores con transistor MOS de acrecentamiento y resistor MOS como resistor de drenaje. (a) Los dos transistores MOS son del mismo tipo general (b) Los dos transistores MOS son de simetría complementaria.

El transistor de carga Q_2 de la Fig. 3.11(a) tiene su compuerta al mismo potencial que el drenaje y estará en el estado de conducción para toda la tensión de salida v_o tal que

$$V_D - v_o > V_T$$

Donde V_T es la tensión de umbral de Q_2 , no conduciría si

$$V_D - v_o < V_T$$

Si la tensión de entrada v_A se eleva aproximadamente $+V_D$, Q_1 pasa al estado de conducción, con lo que disminuye v_o y aumenta la conducción de Q_2 , la que es menor que la I_{DS} de Q_1 de modo que Q_1 funciona en la región trío y v_o es muy pequeña en

comparación con V_D . Cuando v_A se reduce a cero Q_1 pasa a NO y v_0 se eleva, aproximándose a $(V_D - V_T)$ como limite, pues Q_2 actúa como un resistor MOS.

En el circuito de la Fig. 3.11(b), el resistor MOS de carga Q_1 , es un MOST de acrecentamiento y canal p, Q_2 es un MOST de acrecentamiento y canal n.

Cuando v_A tiene un valor positivo grande, por ejemplo, $+V_D$, Q_1 se encuentra en un estado de alta conductividad, pero Q_2 esta en la condición de corte pues su tensión de compuerta a fuente es cero. Así, la tensión de salida v_0 es prácticamente nula cuando $v_A = +V_D$. En cambio si v_A se reduce a cero, Q_1 pasa a la condición de corte, pero Q_2 ofrece una gran conductancia de drenaje a fuente, puesto que esta sometido a una elevada tensión de polarización directa de compuerta a fuente. De esta manera, si no fuera por la corriente de fuga de la unidad Q_1 , v_0 es prácticamente igual a $+V_D$ cuando v_A es cero. Una notable propiedad de este circuito es que, en cualquier estado de entrada, los transistores solo conducen corriente de fuga (en ausencia de carga), de modo que la potencia disipada es mínima. Esta condición de potencia reducida no se altera cuando el circuito se emplea para controlar otras etapas MOST, pues estas tampoco exigen potencia apreciable. No obstante, fluyen corrientes importantes durante las transiciones, debido a las capacitancias circuitales asociadas; estas corrientes pueden dar lugar a una disipación de potencia importante, especialmente cuando la frecuencia de repetición de la conmutación es grande. Debe observarse también que poniendo a tierra la fuente de Q_2 en lugar de la de Q_1 , se obtiene un circuito que operara con valores negativos de V_D , v_A y v_0 .

3.8 CIRCUITOS INTEGRADOS MOS.

Al examinar el trazado físico, la estructura y los diagramas esquemáticos de los circuitos integrados MOS, se ve la falta total de componentes convencionales. En otros términos, los circuitos integrados MOS solo contienen dispositivos MOS, con total ausencia de resistores, capacitores o diodos individuales.

3.8.1 EL INVERSOR MOS INTEGRADO.

La Fig. 3.12 muestra el diagrama esquemático, un dibujo en escala de la estructura física, y una vista en corte, fuera de escala, de un inversor MOS integrado práctico, tipo p.

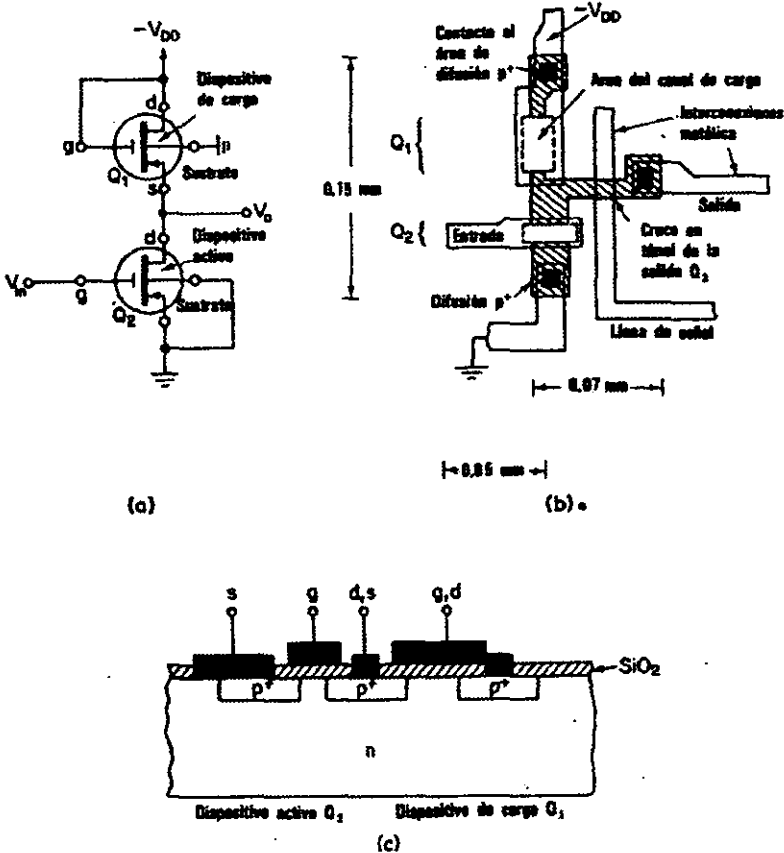


Figura 3.12 Inversor MOS práctico. (a) Diagrama de circuito.(b) Disposición física. (c) Esbozo de la sección transversal.

En la distribución física, Fig. 3.12(b), el dispositivo de carga Q_1 está en la parte superior y el dispositivo activo Q_2 en la parte inferior. La difusión p^+ de abajo sirve como fuente del dispositivo activo, mientras que la difusión p^+ superior es el drenaje del dispositivo de carga y la difusión central actúa a la vez como fuente del dispositivo de carga y como drenaje del dispositivo activo. Esta utilización de un área difundida en

común para más de una función economiza superficie. Otro aspecto que contribuye a la economía de área en los circuitos integrados MOS (transistores y resistores) en comparación con los TBJ y los resistores de difusión, hacen que un circuito MOS requiera un área total igual a aproximadamente la quinta parte de la que necesita para un circuito TBJ comparable.

En la distribución de la Fig. 3.12(b), la compuerta y el drenaje del dispositivo de carga están unidos por mentalización; el metal de la compuerta de cada dispositivo cubre en parte las zonas p^+ , a fin de evitar efectos de borde. La extensión lateral de la difusión central es de particular significado; se trata de un túnel (por debajo de la capa de SiO_2 y la interconexión metálica) que se extiende hasta la interconexión de salida. Estas interconexiones por difusión (cruces aislados) representan otra ventaja a favor de los circuitos integrados MOS. El efecto resistivo de estos cruces por difusión (con resistencias laminares del orden de 100 ohmios por cuadrado) es despreciable, porque se hallan en serie con resistores de carga de alto valor e impedancias de entrada muy elevadas. Es importante observar que son diferentes las geometrías de los dos canales de los dispositivos MOS de la Fig. 3.12⁶.

⁶ Cap 7 Ref. 9

CAPITULO 4

LIMITACIONES DEL TRANSISTOR MOS

4 TRANSISTOR MOS: SUS LIMITACIONES.

El análisis teórico del transistor MOS realizado hasta el momento ha considerado el caso unidimensional y del llamado canal largo, sin embargo el desarrollo de la microelectrónica va mejorando constantemente la tecnología por un lado y por otro disminuyendo las dimensiones físicas de los dispositivos para aumentar la densidad de componentes en un circuito integrado. El resultado es de reducciones considerables en las dimensiones, que conllevan a la aparición de nuevos factores que influyen en el funcionamiento del TMOS y varían sus parámetros de trabajo. Los puntos que se tocarán en este tema tienen relación directa con esto e intentaran dar una idea de algunos elementos que afectan o mejoran el funcionamiento de estos transistores.

4.1 VELOCIDAD DE RESPUESTA.

Cuando se analizan los límites de la velocidad de respuesta de un TMOS. hay tres elementos que deben considerarse:

1. El tiempo de tránsito de los portadores a lo largo del canal.
2. La carga de las capacitancias inherentes al dispositivo.
3. La carga de las capacitancias parásitas.

Con el fin de evaluar el tiempo de tránsito es suficiente utilizar la condición de saturación con las expresiones de primera aproximación¹.

Conociendo que

$$E_y = -\frac{dV}{dy} \quad (4.1)$$

y a partir de los resultados obtenidos para las expresiones de la densidad de corriente j y de $Q_n(y)$, se obtiene que:

$$dV(y) = \frac{(V_g - V_T)^2 dy}{2L(V_g - V_T - V(y))} \quad (4.2)$$

Si se considera que cuando $y = L$, $V = V_{\text{Drain}}$, resulta que el campo será igual a

$$E_y(y) = -\frac{V_g - V_T}{2L} \frac{1}{\sqrt{1 - \frac{y}{L}}} \quad (4.3)$$

Como la velocidad de tránsito para los electrones en función del campo eléctrico es igual a

$$v(y) = -\mu_n E(y) \quad (4.4)$$

y como depende de la coordenada, se puede calcular el tiempo de tránsito integrando a lo largo del canal, o sea:

$$t_v = \int_0^L \frac{dy}{\mu E_y(y)} \quad (4.5)$$

Sustituyendo en la Ecuación 4.5 e integrando se obtiene

¹ Cap. 8 Ref. 11

$$t_r = \frac{4}{3} \frac{L^2}{\mu(V_G - V_T)} \tag{4.6}$$

En la Fig. 3.1 se muestran las capacitancias fundamentales de un transistor MOS, donde las capacitancias de solapamiento C_{GS} y C_{GD} son parásitas.

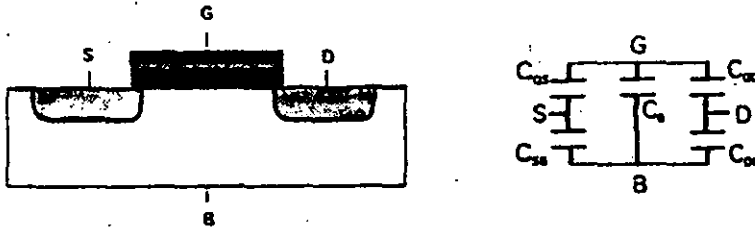


Fig. 4.1 Capacitancias en un transistor MOS, intrínsecas y parásitas.

4.2 RESISTENCIAS EN SERIE.

En un MOST pueden haber resistencias en series en los contactos metal-semiconductor a las zonas de surtidor y drenaje y también por la resistencia propia de estas regiones hasta el canal.

Para la evaluación de este efecto utilizaremos la expresión de primera aproximación para la corriente, o sea

$$I_D = \mu C_o \frac{Z}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \tag{4.7}$$

donde los voltajes V_{GS} y V_{DS} son con respecto a los potenciales en los bordes del canal, por eso si se considera el electrodo externo S puesto a tierra, entonces resultará:

$$V_{DS} = V_{DD} - I_D (R_S + R_D) \quad \text{y} \quad V_{GS} = V_G - I_D R_S \tag{4.8}$$

Sustituyendo la Ecuación 4.8 en 4.7, despejando la corriente y despreciando el termino cuadrático de la corriente, como se ve en la Fig.3.2, se obtiene:

$$I_D = \frac{\mu C_0 Z}{L \left(1 + \mu C_0 \frac{Z}{L} (R_S + R_D) (V_G - V_T) - \mu C_0 \frac{Z}{L} R_D V_{DD} \right)} \left[(V_G - V_T) V_{DD} - \frac{V_{DD}^2}{2} \right] \quad (4.9)$$

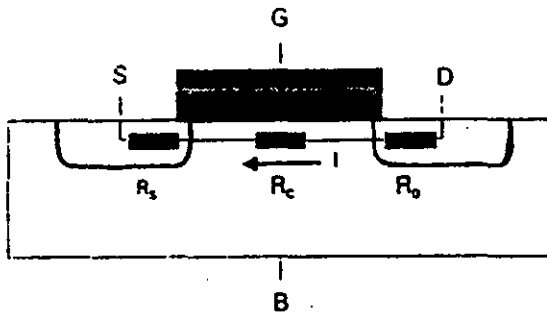


Fig. 4.2 Resistencias en un transistor MOS. R_S - resistencia de contacto y de la región de fuente; R_D - resistencia del contacto y de la región del drenador.

4.3 CORRIENTE SUBUMBRAL .

Como se mostró en párrafos anteriores en el caso en que el potencial superficial varia entre ϕ_f y $2\phi_f$, hay una corriente de drenaje que no se puede calcular con las expresiones que se han deducido, las cuales son válidas solamente a partir de la condición de inversión fuerte. Se realizaran algunos cálculos estimados que permitan estimar la magnitud de esta corriente, que por estar por debajo del voltaje umbral, se conoce como corriente subumbral. La misma corresponde a la región de inversión débil, por eso se considera:

1. La superficie cambió de conductividad, pero la concentración de portadores es pequeña, menor que en el sustrato.
2. Debido a la poca concentración de portadores la caída de potencial a lo largo del canal no puede ser muy alta, por lo cual el campo eléctrico longitudinal es pequeño con respecto al transversal.
3. Para pequeños voltajes de V_D (del orden de 0.1 V) se cierra casi totalmente el canal junto al drenaje por lo cual se puede considerar que la concentración de portadores en la superficie va de n_s junto al S, hasta cerca de 0 junto al D.

Como resultado de estas consideraciones se puede estimar que la corriente que predomina es la de difusión del S al D, y no la de corrimiento, por lo cual se calcula la corriente según:

$$I_{df} = -AD_n \frac{dn}{dy} = AqD_n \frac{dn(y)}{dy} \quad (4.10)$$

Con las consideraciones realizadas $dn \approx n_s$ y $dy = L$, si x es la profundidad del canal de subumbral, entonces

$$I_{df} \approx Zx_i qD_n \frac{n_s}{L} \quad (4.11)$$

Se estima esta profundidad, como se muestra en la Fig. 4.3, partiendo de la concentración de portadores (electrones) en esa región según:

$$n(x) = n_i e^{u-u_i} \quad (4.12)$$

Se sabe que $\Delta n = N_s - n_i$ y $\Delta x = x_i$, por lo cual el análisis de la región x_i muestra que

$$\frac{dn}{ndx} \approx -\frac{N_a - n_i}{N_a} \frac{1}{x_i} = \frac{q}{kT} (E_s) \quad (4.13)$$

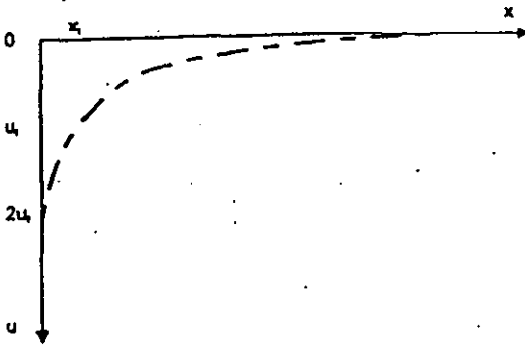


Fig. 4.3 Geometría de la región de subumbral; $u_i < u < 2u_i$

por lo tanto

$$E_s \approx \frac{1}{x_i} \frac{kT}{q} \quad (4.14)$$

O sea, que en este caso prácticamente todo el campo eléctrico cae con un potencial kT/q en la profundidad x_i . Aproximando el valor del campo eléctrico, para el caso en que $u_s \geq u_i$ se tiene que

$$E_s \approx \sqrt{\frac{2qN_a}{k_s \epsilon_0} \phi_s} \quad (4.15)$$

y

$$x_i \approx \frac{kT}{q} \sqrt{\frac{k_s \epsilon_0}{2qN_a \phi_s}} \quad (4.16)$$

Sustituyendo las ecuación 4.10 y 4.16 en la 4.11, se obtiene:

$$I \approx \frac{Z}{L} q D_n \frac{kT}{q} \sqrt{\frac{k_s \epsilon_0}{2q N_a \phi_s}} n_i e^{-\alpha} e^{\eta} \quad (4.17)$$

El análisis de esta expresión indica que esta corriente no depende de V_D . Esto debe mantenerse mientras no se cierre todo el canal, o sea la condición de "punch-through". La dependencia fundamental es la del potencial superficial ϕ_s , lo que es lo mismo que depender del voltaje de compuerta V_G . Esta dependencia es exponencial, como se ve en la Fig. 4.4a, por lo que se adoptó como criterio de análisis de esta región la característica $\ln I_D - V_G$, Fig.3.4b, evaluada por el parámetro de la pendiente de esta curva, conocido como el barrido subumbral S , y que corresponde al voltaje que se requiere para que la corriente varíe una década, o sea:

$$S = \frac{\ln 10}{\frac{d(\ln I_D)}{dV_G}} \quad (4.18)$$

La unidad utilizada para S es de mV /década. El caso ideal es cuando la transición por esta región es muy rápida, o sea, mientras menor es S , más próxima al caso ideal².

² Cap 3 Ref. 15

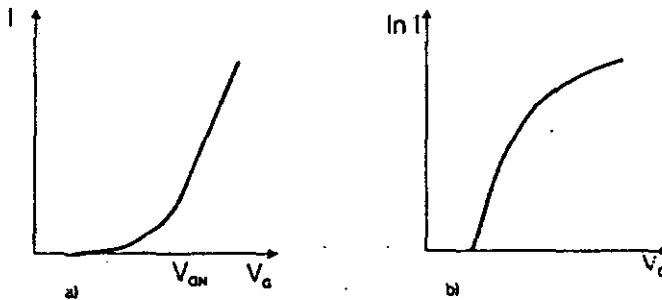


Fig. 4.4 Región de la corriente subumbral: a) en escala normal, donde V_{ON} indica el punto de unión con la región lineal; b) en escala logarítmica, donde se define la S . Para $V_{DS} < 100$ mV

4.4 ESCALADO ("SCALING").

Para poder aumentar la densidad de transistores por unidad de área de silicio se requiere reducir las dimensiones geométricas del MOST básico, lo cual también debe aumentar la velocidad de operación del dispositivo. A menor longitud del canal debe reducirse el tiempo de tránsito de los portadores y reducirse la capacitancia parásita, todo lo cual coadyuva al aumento de la frecuencia de operación.

Esta reducción de las dimensiones se conoce como escalado ("scaling") del transistor. También se incluyen las otras magnitudes como concentraciones y voltajes, todas en forma coherente.

El escalado tiene dos tipos de limitaciones: limitaciones físicas, definidas por la tecnología de fabricación y la menor longitud de canal en la mascarilla que es posible definir y la limitación de funcionamiento (performance), definida por la degradación de los parámetros fundamentales del MOST. Este proceso de escalado se realiza de manera continua desde que comenzó la carrera de desarrollo de los circuitos integrados VLSI en 1970.

Para el escalado se han propuesto, y ejecutado, diferentes formas de abordar el problema como son: con campo constante; con voltaje constante; con voltaje cuasi-constante y el generalizado.

4.4.1 ESCALADO CON CAMPO CONSTANTE.

Se propone reducir las dimensiones horizontales, la profundidad de juntura, el espesor del óxido de la compuerta y los voltajes por un factor de escalado $\lambda > 1$, e incrementar la concentración de impurezas en forma tal que el campo eléctrico en el interior del dispositivo no varíe. Para que la profundidad de la ZCE no varíe se requiere que el V se divida por λ y que la concentración en el sustrato N_B aumente en λ .

Esta propuesta se consideró como de un escalado ideal, evitándose todos los efectos negativos que conllevan el aumento de los campos internos como: aumento del campo crítico; disminución de la movilidad; la inyección de electrones calientes en el óxido de compuerta, etc.

En el caso de las capacitancias propias del MOST el comportamiento es el siguiente: la capacitancia por unidad de área aumenta en λ , el área se reduce en λ^2 , por lo cual la capacitancia real se reduce en λ .

En el caso del voltaje umbral el control es más difícil, ya que V_{FB} no dependen del escalado, requiriéndose un ajuste por implantación.

Para la corriente se puede considerar que la carga de portadores por unidad de área no varía, pero como el área disminuye en λ^2 , entonces la carga efectiva también se

reduce. Al reducir la capacitancia en λ resulta que la corriente se ve reducida finalmente en λ .

La corriente subumbral no es directamente escalable como la corriente normal, lo cual trae por consecuencia que se incremente su efecto relativo.

En lo referente a la potencia disipada se puede concluir que si tanto la corriente como el voltaje se reducen en λ entonces la potencia se reducirá en λ^2 . La potencia por unidad de área se mantiene constante.

Como resultado de mantener el campo constante el tiempo de tránsito del portador por el canal se reduce en λ , con lo cual el importante **factor de mérito potencia-demora** se ve reducido en λ^3 .

Otros factores que tienen que considerarse en estos casos son:

1. La reducción de las líneas de interconexión en anchura y espesor, tanto de metal como de polisilicio, con lo cual aumenta la resistencia por unidad de longitud. Esto conlleva a un aumento de la densidad de corriente en estas líneas en λ . Tiene que considerarse que el límite práctico para el aluminio es de una densidad de corriente de 1 mA/mm^2 .
2. El aumento de las resistencias de contacto, por la reducción de las áreas de contacto.
3. El aumento de las resistencias por cuadrado de las zonas difundidas al reducirse la profundidad de las juntas.

Una de las principales desventajas que tiene éste escalado es la reducción de los voltajes, por considerarse el voltaje de alimentación como una norma por los

constructores de equipamiento, estando el mas generalizado en la técnica digital en 5 v. Se trata de adoptar una nueva norma en 3.3 V para los nuevos componentes de muy alta integración .

4.4.2 ESCALADO A VOLTAJE CONSTANTE .

Para solucionar el problema antes señalado se hace otra propuesta de escalado, en la cual los máximos voltajes externos se mantienen constante, por ejemplo a 5 V. Otras magnitudes como Z, L y N_n se escalan como en el caso anterior. Para evitar grandes campos transversales el espesor del óxido se reduce en un factor menor de λ , igual a β .

4.4.3 ESCALADO A VOLTAJE CASI-CONSTANTE.

Para tratar de evitar los problemas con el incremento del campo eléctrico intenso, se propone una variante en la cual los voltajes externos máximos se reduzcan en el factor β .

4.4.4 ESCALADO GENERALIZADO.

Tratando de conjugar las ventajas y desventajas de cada una de las propuestas anteriores se realiza la siguiente que se conoce como **generalizada**, en la misma, para lograr que la zona de carga espacial se reduzca lo mismo que las dimensiones Z, L y x_0 , se propone que la concentración del substrato varíe de otra forma. La misma se muestra

en la Tabla 3.1, donde se resumen las principales características de cada una de las propuestas anteriores.

En el escalado se trata de tener como un criterio que el comportamiento de canal largo en la región de inversión débil se mantenga, con lo cual se evitan los efectos severos del acortamiento del canal. Para esto se usa un criterio empírico, donde la longitud mínima de canal se relaciona con las dimensiones de las profundidades de juntura, de las ZCE en el S y D y con el espesor de óxido de compuerta. La expresión que refleja estos compromisos es igual a:

$$L_{\min} = 8.8[x_p x_o (x_{ds} + x_{dd})]^{1/3} \text{ [mm]}.$$

Tabla 4.1 Reglas de escalado³

λ - factor de escalado, $1 < \lambda < 10$; $1 < \beta < \lambda$.

Magnitud	Campo constante	Voltaje constante	Voltaje cuasiconstante	Generalizado
Z, L	$1/\lambda$	$1/\lambda$	$1/\lambda$	$1/\lambda$
x_o	$1/\lambda$	$1/\beta$	$1/\lambda$	$1/\lambda$
N_B	λ	λ	λ	λ^2
V, V_T	$1/\lambda$	1	$1/\beta$	$1/\beta$

4.5 EFECTOS DE LAS GEOMETRÍAS REDUCIDAS.

El análisis del transistor MOS realizado hasta el momento considera que el canal se puede considerar "largo", o sea, que la variación de parámetros es fundamentalmente

unidimensional en la dirección perpendicular a la superficie. Cuando se van reduciendo las dimensiones del canal, tanto en la topología como en la profundidad aparecen los efectos bidimensionales. Esta es la condición que se conoce como de "canal corto".

4.5.1 VELOCIDAD LÍMITE.

En un semiconductor en equilibrio la energía cinética del electrón es definida por su temperatura según

$$\frac{1}{2} m_n^* v_t^2 = \frac{3}{2} kT \quad (4.20)$$

donde $m_n^* = 0.6m_0$ es la masa efectiva de conducción y k es la constante de Boltzmann.

En la presencia de un campo eléctrico el electrón experimenta una aceleración que aumenta su energía y añade un componente a su velocidad de desplazamiento v_d . Es conocido que

$$v_d = -\frac{q\tau_{ch}E}{m_n^*} = -\mu_n E \quad (4.21)$$

donde τ_{ch} es el tiempo entre choques;

μ_n es la movilidad efectiva del electrón.

Según la Ecuación 4.21 esta dependencia es lineal, pero ocurre que a medida que el campo se va incrementando se produce un efecto de saturación de la velocidad, alcanzándose para un llamado **campo crítico** la velocidad conocida como **velocidad límite o máxima**. Estos fenómenos son de carácter general para el semiconductor. Para

³ Ref. 2

el transistor MOS nos interesa el comportamiento de los portadores en el canal del transistor. Como se conoce, en esta zona hay dos campos, una transversal que responde por la acumulación de los portadores en la superficie y otro longitudinal que produce la corriente a lo largo del canal.

El estudio experimental del MOST muestra que la velocidad de los portadores en el canal tiene el mismo efecto de saturación, como se ve en la Fig. 4.5.

se ha propuesto una expresión del tipo

$$v_d = \frac{|\mu E_y|}{\left[1 + \left(\frac{E_y}{E_c}\right)^\alpha\right]^{\frac{1}{\alpha}}} \quad (4.22)$$

para ajustarla a la curva de variación de la velocidad, donde μ , E_c y α se determinan empíricamente, y dependen de tipo de material, del portador y de las dimensiones físicas de la zona.

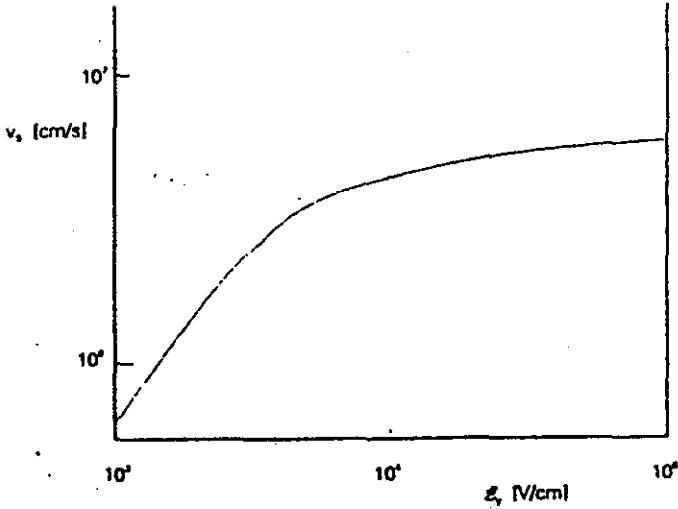


Fig. 4.5 Velocidad de los electrones en el canal del TMOS en dependencia del campo eléctrico longitudinal, mostrando la saturación.

La corriente será

$$I_D dy = -Z\mu Q_n dV_D \quad (4.23)$$

y recordando que $E_y = -dV/dy$, se sustituye en la ecuación 4.23, utilizando la ecuación

3.22 para el caso de $\alpha = 1$, obteniéndose

$$I_D = \mu Z Q_n E_y = Z Q_n \frac{\mu E_y}{1 + \frac{E_y}{E_c}} \quad (4.24)$$

agrupando queda

$$I_D = Z\mu E \left[Q_n - \frac{I_D}{Z\mu E_c} \right] \quad (4.25)$$

e integrando a lo largo del canal, o sea, variando y de 0 a L , lo que corresponde a que V varíe de 0 ($V_s = 0$) hasta V_D , resultará

$$I_D = -\frac{Z}{L} \frac{\mu}{\left[1 + |V_D/LE_c|\right]} \int_0^{V_D} Q_n dV \quad (4.26)$$

Utilizando la expresión de primera aproximación para la corriente, se obtiene que

$$I_D = \frac{Z}{L} C_0 \frac{\mu}{1 + \left|\frac{V_D}{LE_c}\right|} \left[(V_G - V_T) V_D - \frac{V_D^2}{2} \right] \quad (4.27)$$

Si la movilidad superficial es μ_s , entonces se puede considerar que el campo critico sería $E_c = v_{max}/\mu_s$, por lo tanto se puede escribir

$$1 + \left|\frac{V_D}{LE_c}\right| \Rightarrow 1 + \left|\frac{\mu_s V_D}{Lv_{max}}\right|$$

En el caso de saturación se puede calcular a partir de la condición para la cual la conductividad se hace igual a cero, o sea:

$$\frac{dI_D}{dV} = -\frac{\mu Z Q_n}{L(1 + |V_D/LE_c|)} + \frac{Z\mu \int_0^{V_D} Q_n dV}{L^2 E_c [1 + |V_D/LE_c|]^2} = 0 \quad (4.28)$$

por lo tanto

$$I_{Dsat} = \mu Z E_c Q_n (V_{Dsat}) \quad (4.29)$$

Considerando la magnitud de la carga en este caso, a partir de las Ecuación 4.27 y 4.29, se puede despejar la expresión para el V_{Dsat} , resultando:

$$V_{Dsat} = |LE_c| \left[\sqrt{1 + \frac{2(V_G - V_T)}{|LE_c|}} - 1 \right] \quad (4.30)$$

Cuando no se considera la velocidad de saturación, se obtiene que

$$V_{Dsat} = V_G - V_T \quad (4.31)$$

Como se ve la dependencia del V_G en un caso es como raíz cuadrada y en el otro lineal. El no considerar el campo crítico es equivalente a considerar que $E_c \rightarrow \infty$ en cuyo caso las Ecuación 4.30 y 4.31 se igualan.

De esta forma la corriente será proporcional a $V_G - V_T$ cuando se alcanza la velocidad límite.

Con respecto al campo transversal E_x , se demuestra que el incremento del campo conlleva a un decremento de la movilidad.

4.5.2 CANAL CORTO.

En la Fig. 4.6a se muestra el caso que hemos estudiado de canal largo, donde se puede considerar que el cálculo de la zona de carga espacial tiene carácter unidimensional. Consideramos que para los efectos prácticos la ZCE va directamente desde la superficie, o sea, se desprecia la profundidad del canal.

Como se ve en la Fig. 4.6b, al acortarse el canal la ZCE bajo el canal llega a tener una marcada geometría trapezoidal, por lo cual el cálculo de la carga contenida en el mismo varía con la geometría del canal. En la Fig. 4.7a se muestran los elementos geométricos que se utilizan para el cálculo de la ZCE.

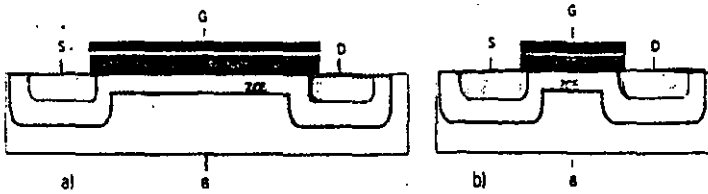


Fig. 4.6 Geometría trapezoidal del canal bajo la compuerta: a) canal largo; b) canal corto.

El calculo se basa en determinar la nueva área del trapecioide L, L', x_{gd} , en lugar del área Lx_{gd} . En la Fig. 4.7b se amplio la zona del extremo del canal, donde se ve cual es el área del triangulo ABC que debe calcularse para restarla.

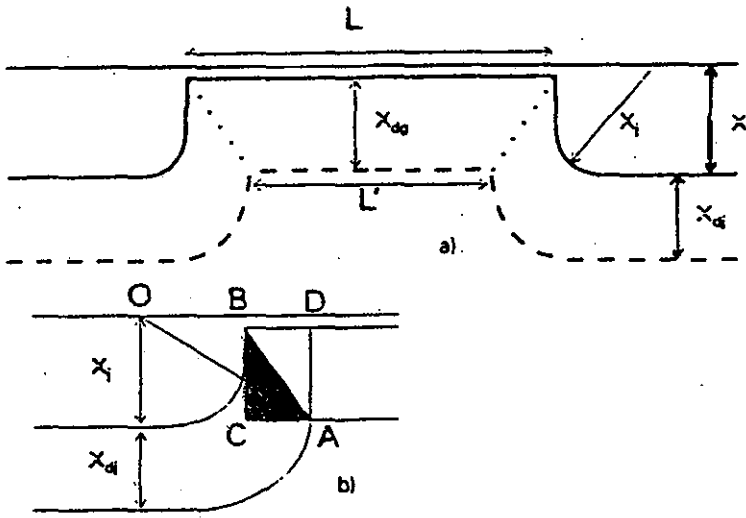


Fig. 4.7 ZCE bajo un canal corto a) dimensiones generales; b) geometría de las zonas junto al S y D.

Se conoce que $BC = x_{qd}$ y como $AO = x_j + x_{dj}$ y $DB = AC$, y considerando que $x_{dj} = x_{dq}$ resulta que

$$(x_j + x_{dq})^2 = x_{dq}^2 + (AC + x_j)^2$$

despejando AC resulta

$$AC = x_j \left[\sqrt{1 + \frac{2x_{dq}}{x_j}} - 1 \right] \quad (4.32)$$

y el área rayado a ambos lados del canal es igual a

$$x_{dq} \cdot x_j \left[\sqrt{1 + \frac{2x_{dq}}{x_j}} - 1 \right] \quad (4.33)$$

La carga total resultante Q'_B será igual a:

$$Q'_B = qN_a Z L x_{dq} - qN_a Z \left[x_{dq} x_j \left(\sqrt{1 + \frac{2x_{dq}}{x_j}} - 1 \right) \right]$$

$$Q'_B = qN_a Z L x_{dq} \left[1 - \frac{x_j}{L} \left(\sqrt{1 + \frac{2x_{dq}}{x_j}} - 1 \right) \right] \quad (4.34)$$

Así se ve reducida la concentración de carga en el substrato bajo el canal en una magnitud que será mayor a medida que se reduzca la L, o que la x_j y x_{dq} aumenten, lo cual ocurre al disminuir la concentración de impurezas en el substrato.

Como el voltaje umbral es igual a:

$$V_T = V_{FB} + 2\phi_f + \frac{Q_B}{C_o}$$

el variar Q_B resultará en una variación de V_T según la expresión

$$\frac{\Delta V_T}{V_T} = \frac{V_T - V_T'}{V_T} = 1 - \frac{Q_B'}{Q_B} = \frac{x_j}{L} \left(\sqrt{1 + \frac{2x_d}{x_j}} - 1 \right) \quad (4.35)$$

O sea, el efecto del canal corto conlleva a una reducción del V_T lo que puede llegar a ser muy significativo. Para reducir este efecto se requiere:

1. Hacer las junturas mas superficiales, o sea menores x_j .
2. Que L sea lo mayor posible.
3. Aumentar la concentración del substrato para reducir x_d .

Otro efecto negativo es el aumento del barrido subumbral S , lo cual no es deseable y menos para transistores pequeños⁴.

4.5.3 CANAL ESTRECHO.

En el caso de que la anchura del canal Z disminuya, también se puede pasar de un análisis unidireccional a uno bidireccional. En la Fig. 3.8 se muestra que hay una ZCE que aparece sombreada, que va mas allá de los bordes del canal, lo cual aumenta la carga efectiva en el canal.

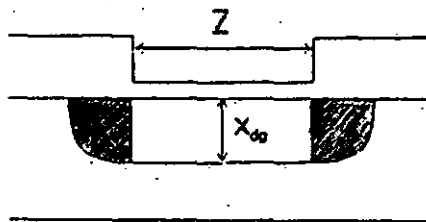


Fig. 4.8 Anchura del canal de un TMOS mostrando la ZCE que se prolonga mas allá de los bordes.

⁴ Cap 5 Ref. 9

Si la carga en la región rectangular bajo el canal es igual a:

$$Q = qN_a ZLx_{ds} \quad (4.36)$$

al considerar las áreas sombreadas esta carga debe aumentar a

$$Q' = Q + \left(2 \frac{\pi x_{ds}^2}{4} \right) qN_a L \quad (4.37)$$

por lo tanto se puede escribir que la carga incrementada será

$$Q' = qN_a ZLx_{ds} \left(1 + \frac{\pi x_{ds}}{2Z} \right) \quad (4.38)$$

Al hacer la misma consideración sobre el efecto de este incremento de carga sobre el V_T se puede calcular que el mismo se incrementa de acuerdo con

$$\frac{\Delta V_T}{V_T} = \frac{\pi x_{ds}}{2Z} \quad (4.39)$$

que es lo contrario que ocurre para canal corto, y depende solamente de la relación x_{ds}/Z .

4.5.4 PORTADORES CALIENTES.

La Ecuación 4.20 muestra la energía cinética media de un electrón para una temperatura dada. Por el efecto de la aceleración del electrón en un campo eléctrico, el mismo aumenta su energía. Si la energía adicional llega a ser igual o superior a la energía en equilibrio, se dice que se tiene un electrón caliente, con una temperatura equivalente superior a la del medio.

La zona del MOST donde los campos son mas intensos es junto al drenador, donde incluso el campo tiene una característica bidimensional para cuyo cálculo se requiere del uso de computadoras. Es en esta zona donde es mayor la probabilidad de que los portadores alcancen la condición de calientes.

Existen varios efectos que pueden provocar los electrones (huecos) calientes, uno de los mas importantes es la inyección de portadores calientes en el óxido de la compuerta, produciendo dos fenómenos, el aumento de la corriente de compuerta que puede variar del orden de fA al de pA, y que parte de estas cargas se quedan atrapadas en el óxido produciendo el conocido efecto de correr el V_T . Como se ve es un efecto negativo.

4.5.5 RUPTURA.

En un transistor MOS la juntura del drenador tiene aplicado un voltaje igual a la suma de V_D y del módulo de V_o . Como toda juntura hay un valor de este potencial aplicado a partir del cual se produce un aumento brusco de la corriente, el cual se conoce como voltaje de ruptura y da por resultado lo que se muestra en la característica de la Fig. 4.9. Otro mecanismo que lleva a la ruptura es el proceso de avalancha que pueden crear los portadores con altas energías en la región del drenador. Al generarse un par electrón-hueco, el electrón lo recoge el drenador y el hueco se va al electrodo del sustrato creando una corriente de sustrato. Esta corriente crea una caída de potencial en el semiconductor polarizando de hecho el sustrato a V_B . Fig. 4.10. El efecto sobre la juntura de la fuente es el de polarizarla en directa, con la consecuente inyección de

portadores en el substrato, generando nuevos pares y produciendo un efecto de avalancha que da lugar a la ruptura. Para canales mas cortos, los campos son mas intensos, con lo cual se incrementa la probabilidad de avalancha, disminuyendo el voltaje de ruptura. Si la energía disipada no sobrepasa la permisible este proceso es reversible.

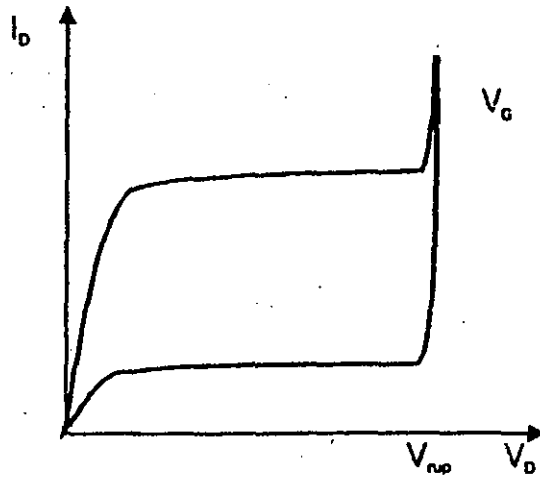


Fig. 4.9 Zona de ruptura en la característica I - V del MOST.

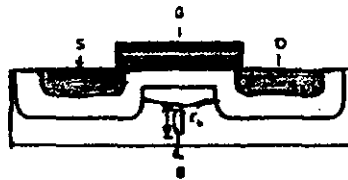


Fig. 4.10 Polarización de la unión de la fuente por efecto de la corriente de substrato.

Hay otro tipo de voltaje de ruptura en el MOST y es el voltaje para el cual el óxido "se poncha", o sea se produce una descarga a través del mismo. Este fenómeno depende en su totalidad de la calidad del óxido de la compuerta, que tiene un valor crítico cuando es bueno de 7×10^6 V/cm. Para evitar este tipo de ruptura que no es reversible, se deben aplicar potenciales de trabajo por debajo de los críticos y se debe tener un cuidado especial en las descargas de alto voltaje que se producen por la carga estática acumulada.

CONCLUSIONES

Los semiconductores forman un grupo de materiales que presentan un comportamiento intermedio entre los conductores y los aislantes; la conductividad de estos materiales es una función creciente de la temperatura de tal forma que a la temperatura ambiente la mayoría de los semiconductores presentan una conductividad apreciable, aunque menor que la de un metal. Incluso a una temperatura dada, es posible variar a voluntad la conductividad de estos materiales si se les añade una cantidad controlada de impurezas de determinados elementos químicos, a este proceso se le llama *dopaje*. De esta manera se pueden clasificar a los semiconductores con respecto a su dopaje, teniendo: semiconductores intrínsecos, que son aquellos semiconductores en estado puro y perfectamente cristalizados; y los semiconductores extrínsecos *tipo n*, que son aquellos que están dopados con átomos *donores* y los semiconductores *tipo p*, que son aquellos que están dopados con átomos *aceptores*.

El estudio de las propiedades de estos materiales se basa en los principios de la física cuántica, la cual hace un análisis estadísticos de la estructura atómica, dando como resultado, el que se pueda usar una estructura de bandas de energía formada por la banda de valencia, banda de conducción y en medio de estas una banda de estados prohibidos o "gap"; lo que sirve para analizar, de hecho, cualquier dispositivo semiconductor, incluyendo el transistor MOS.

Un concepto muy importante en el estudio de los semiconductores es el concepto de hueco, el cual se puede explicar a partir de que la excitación de un electrón a la banda de conducción implica la ruptura de un enlace en algún punto del cristal, donde

a su vez se origina un estado vacante que además presenta una deficiencia de carga negativa (equivalente a una carga positiva de magnitud igual a la de la carga del electrón), y es precisamente a éste estado vacante al cual se le llama hueco.

Todos estos conceptos nos permiten estudiar los fenómenos que se dan en los semiconductores tales como la difusión y la conducción de carga, el cálculo de ésta; al igual que los parámetros característicos del semiconductor como son permitividad relativa, masa efectiva, movilidad, etc.

El estudio de la física de semiconductores nos ayuda a analizar las estructuras metal-aislante-semiconductor, y en particular aquellas en las que el semiconductor es el silicio y el aislante su óxido, SiO_2 (denominadas **estructuras MOS**) la cual esta formada por una sucesión de varias capas: la primera de ellas es un metal (electrodo superior, o compuerta), la segunda es una capa aislante de óxido de silicio, y la tercera es una capa de semiconductor (el sustrato de silicio) está estructura se completa con una capa metálica depositada sobre la cara inferior del sustrato de silicio formando un contacto óhmico con el silicio (electrodo inferior). Los electrodos de contacto, superior e inferior, sirven de contacto de salida al exterior del dispositivo cuando se aplica sobre ellos una tensión externa.

Una estructura metal-óxido-semiconductor (MOS) funciona según sea el valor de la tensión externa aplicada, de está manera se pueden distinguir cinco casos diferentes: 1) *Caso de acumulación*, 2) *Caso de agotamiento*, 3) *Caso intrínseco*, 4) *Caso de inversión* y 5) *Caso de inversión fuerte*.

Uno de los aspectos más importantes bajo la condición de inversión fuerte es la presencia de carga libre, electrones, en una pequeña franja de la superficie en proporción igual o superior a la de huecos mayoritarios en el interior del semiconductor. Esta franja de la superficie, que adquiere así propiedades conductoras en cierto modo similares a las del interior del semiconductor, a la cual se le llama *canal* y está limitada por dos regiones de mucho menor conductividad esto es, la capa aislante de óxido a un lado y la región de agotamiento del semiconductor en el lado opuesto. *El efecto más importante de la formación del canal en la superficie es la posibilidad de controlar a voluntad la concentración de electrones y por tanto la conductividad del canal a través del potencial aplicada del electrodo de compuerta.* Al potencial que es necesario aplicar al electrodo de compuerta para conseguir esta condición de inversión fuerte y que se tenga la formación del canal se le denomina *voltaje de umbral*.

La estructuras MOS también se puede considerar como un condensador capaz de almacenar carga en ambos lados del aislante, esto es, en el contacto de compuerta y en el semiconductor. Considerando que en una estructura MOS polarizada en condiciones de agotamiento o de inversión no existe una relación lineal entre el voltaje aplicado, V_G , y la carga acumulada a uno y otro lado de la estructura, la carga acumulada en cada caso depende estrechamente de los parámetros de la estructura. Por este motivo, la medida de la variación de la capacidad en función del voltaje (curvas C-V) de una estructura MOS permite obtener una gran información sobre la estructura interna de bandas y la distribución de carga tanto en el semiconductor como en la

interfase óxido-semiconductor. Esta información se obtiene al comparar las curvas C-V experimentales con las que presenta una estructura MOS ideal.

Finalmente diremos que existen diferencias de comportamiento entre la estructura MOS ideal y real. Estas diferencias son debidas fundamentalmente a los siguientes factores: *i) diferencia entre las funciones de trabajo del metal y el semiconductor, ii) existencia de cargas eléctricas en el óxido, iii) la aparición de las denominadas "trampas" o defectos localizados en la interfase $\text{SiO}_2\text{-Si}$.*

La base de un transistor MOS, es la estructura MOS (mencionada anteriormente) en donde el electrodo metálico superior G, depositado sobre la capa aislante, actúa como la terminal de compuerta del transistor. Existen además dos regiones dopadas fuertemente con impurezas ya sean donoras o aceptoras de tipo contrario del sustrato, situadas a cada lado de la compuerta.

El transistor MOS presenta dos modos de funcionamiento:

1. De deserción.
2. De acrecentamiento.

En este trabajo se estudió el modo de funcionamiento de acrecentamiento, que es el más utilizado en los circuitos integrados como elemento básico; mientras que el transistor MOS de deserción se utiliza como elemento de carga debido a que a polarización cero en la compuerta, presenta altas corrientes de fuente a drenador.

Las curvas I-V en función de la tensión de drenaje del MOST tienen las siguientes características: I_D varia linealmente con V_{DS} cuando ésta tensión es pequeña, y se satura después, al aumentar la magnitud de V_{DS} . El canal es prácticamente independiente de

V_{DS} cuando esta tensión es pequeña, pero se estrangula en el extremo próximo al drenaje cuando este se hace muy negativo.

Se necesita aplicar un voltaje positivo suficientemente alto a la compuerta para tener la condición de inversión fuerte en la interfase del semiconductor con el óxido. Esto quiere decir que el voltaje de compuerta debe ser mayor al voltaje de umbral de la estructura MOS.

Cuando se alcanza un voltaje, tal que el ancho del canal en el extremo del drenador se hace cero, se dice que ha ocurrido el estrangulamiento del canal; esto ocurrirá para un voltaje denominado voltaje de saturación, V_{Dsat} .

Finalmente el conocer el funcionamiento del Transistor MOS nos permite adentrarnos en el problema de las limitaciones que se presentan con el aumento de la densidad de transistores por unidad de área, es decir, la reducción de las dimensiones geométricas del TMOS básico (escalado), lo que trae como ventaja el aumento de la velocidad de operación del dispositivo y la reducción de las capacitancias parásitas (mayor frecuencia de operación), entre otras cosas.

Se han propuesto, diferentes formas de abordar el problema del escalado: con campo constante; con voltaje constante; con voltaje cuasi-constante y el generalizado.

Con campo constante, se propone reducir las dimensiones horizontales (geométricas), la profundidad de juntura, el espesor del óxido de la compuerta y los voltajes por un factor de escalado $\lambda > 1$, e incrementar la concentración de impurezas en forma tal que el campo eléctrico en el interior del dispositivo no varíe. Para que la

profundidad de la ZCE no varíe se requiere que el V se divida por λ y que la concentración en el sustrato N_B aumente en λ .

Esta propuesta se consideró como de un escalado ideal, evitándose todos los efectos negativos que conllevan el aumento de los campos internos como: aumento del campo crítico; disminución de la movilidad; la inyección de electrones calientes en el óxido de compuerta, etc.

Otros factores que se tienen que considerar en estos casos son:

La reducción de las líneas de interconexión en anchura y espesor, tanto de metal como de polisilicio, con lo cual aumenta la resistencia por unidad de longitud. Esto provoca un aumento de la densidad de corriente en estas líneas.

Para solucionar el problema antes mencionado se hace otra propuesta de escalado, en la cual los máximos voltajes externos se mantienen constantes, por ejemplo a 5 V. Otras magnitudes como Z , L y N_a se escalan como en el caso anterior. Para evitar grandes campos transversales el espesor del óxido se reduce en un factor menor de λ , igual a β .

Para tratar de evitar los problemas con el incremento de la intensidad del campo eléctrico, se propone una variante en la cual los voltajes externos máximos se reduzcan en el factor β ; o sea, el efecto del canal corto conlleva a una reducción del V_T lo que puede llegar a ser muy significativo. Para reducir este efecto se requiere:

1. Hacer las juntas más superficiales, o sea menores x_j .
2. Que L sea lo mayor posible.
3. Aumentar la concentración del sustrato para reducir x_{d1} .

De lo anterior podemos concluir que sin duda el estudio del funcionamiento físico del transistor MOS es fundamental para entender los problemas que se presentan en el escalamiento del dispositivo, y aun más importante, es el hecho de que este conocimiento nos permite dar soluciones a los diferentes problemas que presentan debido a las limitaciones del dispositivo.

ASREVIATURAS

q : carga del electrón, $q=1.602 \times 10^{-19}$ C.

T : temperatura.

h : constante de Planck.

k : constante de Boltzmann, $k=1.38066 \times 10^{-23}$ J/K.

ϵ_0 : permitividad del vacío.

ϵ : constante dieléctrica del semiconductor.

ϵ_{ox} : constante dieléctrica del óxido.

K : grados Kelvin.

p : momentum cuántico.

m_0 : masa del electrón.

m_e^* : masa efectiva del electrón.

m_h^* : masa efectiva del hueco.

n : concentración de electrones.

p : concentración de huecos.

n_i : concentración de portadores intrínseca.

N_d : concentración de impurezas donadoras.

N_a : concentración de impurezas aceptoras.

N_d^+ : concentración de impurezas donadoras ionizadas.

N_a^- : concentración de impurezas aceptoras ionizadas.

N_t : concentración de centros o trampas de recombinación.

E_d : nivel de energía de impurezas donadoras.

E_a : nivel de energía de impurezas aceptoras.

E_i : nivel de Fermi intrínseco.

E_f : nivel de Fermi.

E_g : energía de la banda prohibida "gap".

$R(T)$: número de portadores que se recombinan por unidad de tiempo, en función de la temperatura.

$G(T)$: número de portadores que se generan por unidad de tiempo, en función de la temperatura.

$n(E)$: función de distribución de electrones.

$p(E)$: función de distribución de huecos.

$f(E)$: función de probabilidad.

$\langle v \rangle$: velocidad media de los electrones.

t_c : tiempo libre medio.

v_e : velocidad de arrastre.

$\langle v_e \rangle$: velocidad media de arrastre de los electrones.

μ_e : movilidad de los electrones.

μ_h : movilidad de huecos.

σ : conductividad.

$\tau_{h,e}$: tiempo de vida de los portadores minoritarios (huecos, electrones).

ϕ_m : función de trabajo del metal.

ϕ_s : función de trabajo del silicio.

V_G : voltaje de compuerta.

V_T : voltaje de umbral.

V_{FT} : voltaje de banda plana.

V_p : voltaje o tensión de extrangulación.

V_D : voltaje de drenador a fuente.

w : región de carga espacial, de deserción o de agotamiento.

ϕ_s : potencial de superficie.

ϕ_i : potencial de volumen.

E : campo eléctrico.

BIBLIOGRAFÍA

1. Mohamed I. Elmasry. "Digital MOS integrated circuits II : with applications to processors and memory design", IEEE Press, 1992.
2. Botchek, Charles M. "VLSI : basic MOS engineering" Pacific Technical Group, 1984.
3. Schroder, Dieter K. "Advanced MOS devices" Addison-Wesley, 1987.
4. Uyemura, John P. "Fundamentals of MOS digital integrated circuits", Addison-Wesley, 1988.
5. Mohamed I. Elmasry. "Digital MOS integrated circuits", New York : IEEE Press, 1981.
6. Wang, Niansu "Digital MOS integrated circuits : design for applications ", Prentice Hall, 1989.
7. Davis, J. R. "Instabilities in MOS devices", Gordon and breach science publishers, 1981.
8. Integrated circuits—Catalogs "MOS : integrated circuits", Santa Clara, California : National Semiconductor, 1974.
9. Tsividis, Yannis. "Operation and modeling of the MOS transistor "New York : McGraw-Hill, 1987.
10. A.S Grove, "Physics and Technology of Semiconductor Devices", Wiley, 1967
11. S.M. Sze "Physics of Semiconductor Devices", Wiley, 1981.
12. E. H. Nicollian, J. R. Brews, "MOS (Metal Oxide Semiconductor). Physics and Technology". Wiley, 1982.
13. G. Milnes "Semiconductor Devices and Integrated Electronics", VanNostrand, 1980.

14. R.S. Muller, T.I. Kaminiis "Devices Electronics for Integrated Circuits", 2ed., Wiley, 1986.
15. A. Cerdeira, M. Estrada "Notas para el curso de Microelectrónica", CINVESTAV-IPN , Depto. Ingeniería Eléctrica, 1996
16. Eco, Humberto "Como se hace una tesis: técnicas y procedimientos de investigación, estudio y escritura", Gedisa, 1995.
17. Ibañez Brambila, Berenice "Manual para la elaboración de tesis", Trillas, 1999.
18. José Maria Albella, José Manuel Martínez-Duart "Fundamentos de electrónica física y microelectrónica". Addison-Wesley. Universidad Autónoma de Madrid (1996).
19. J.P. Mckelvey. "Solid State and Semiconductors Physics". Harper & Row (1994).
20. John R. Reitz, Frederick J., Milford Robert, W. Christy, "Fundamentos de la Teoría Electromagnética ", SITESA:Addison-Wesley Iberoamerica 1986.
21. Dr. M. Estrada, "Notas del curso de Tecnología de semiconductores II", CINVESTAV-IPN, Depto. Ingeniería Eléctrica, 1996.
22. M. Estrada, A. Cerdeira, R Martínez, #Circuitos integrados, fundamentos tecnológicos", Pueblo y Educación, 1987.