
Materials y tecnologías de fabricación de circuitos integrados

PID_00206006

Marc Bara Iniesta

Índice

Introducción.....	5
Objetivos.....	8
1. Semiconductores.....	9
1.1. Bandas de energía y conducción eléctrica	9
1.2. Dopaje de materiales semiconductores	15
1.3. La unión PN	17
2. Procesos de fabricación de circuitos integrados de silicio.....	21
2.1. Creación de la oblea	22
2.2. Oxidación	23
2.3. Fotolitografía y máscara	25
2.4. Grabado. Eliminación de capas no deseadas	26
2.5. Impurificación con materiales dopantes (difusión e implantación)	27
2.5.1. Difusión	28
2.5.2. Implantación iónica	28
2.6. Formación de capas delgadas (deposición y epitaxia)	28
2.6.1. Deposición	29
2.6.2. Epitaxia	29
2.6.3. Capas metálicas	30
3. Introducción a la estructura física de circuitos integrados bipolares y CMOS.....	32
3.1. Tecnología bipolar	33
3.2. Tecnología CMOS	34
Resumen.....	38
Glosario.....	39
Bibliografía.....	40

Introducción

La fabricación de circuitos microelectrónicos está basada en la creación de complejas estructuras desarrolladas sobre materiales semiconductores. Estas estructuras se empaquetan en los circuitos integrados (CI). Un circuito integrado contiene multitud de dispositivos electrónicos miniaturizados –principalmente transistores y las interconexiones entre ellos–, todo dispuesto en una única pieza de material semiconductor.

En este módulo veremos qué procedimientos y materiales se necesitan para poder fabricar un CI, que se creará a partir de un sustrato (o varios sustratos) de material semiconductor. Hay que destacar que el hecho de conocer los materiales y las tecnologías de fabricación es muy importante en nuestro caso, para poder tener las ideas claras respecto a la creación y configuración de los CI. Esto nos permitirá avanzar sobre conceptos más abstractos, teniendo asentadas las ideas y procedimientos sobre los cuales se sustentan.

Cuando hablamos de CI interesa disponer de una serie de métricas o términos para poder evaluar su capacidad y complejidad, y de esta forma tener una primera indicación de la naturaleza de los CI. A continuación se introducen los términos más comunes en este sentido:

a) Tamaño del chip. Es simplemente una indicación de la superficie de un CI, por lo tanto, no se trata de una métrica suficiente como para hablar sobre las capacidades de aquel CI. El tamaño varía mucho según la tecnología y el estilo de diseño utilizados.

b) Número de transistores. Es una métrica más precisa para describir la capacidad del CI, aunque el número de transistores que son necesarios para implementar una misma función varía según la familia lógica.

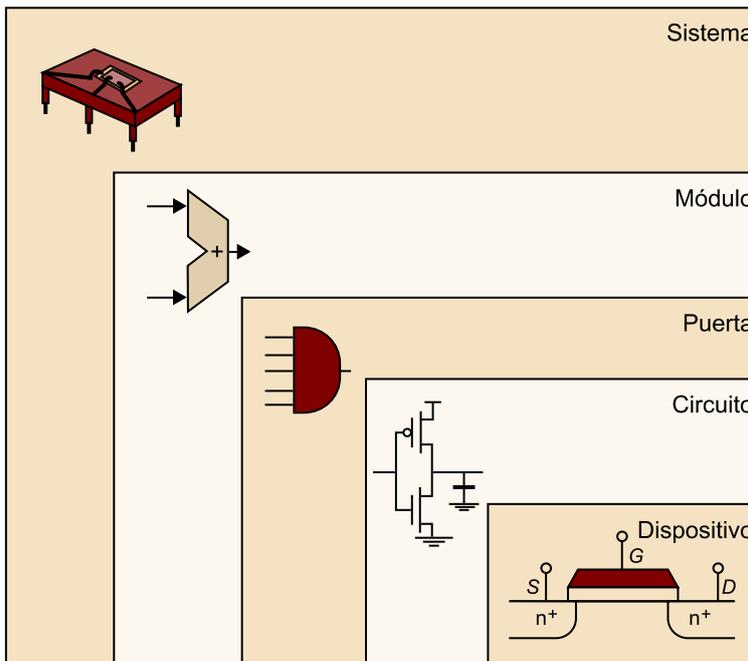
c) Número de puertas. Este indicador intenta valorar la complejidad independientemente de la tecnología de fabricación, dando el orden de magnitud en función del número de puertas lógicas:

- SSI (*small scale integration*): 10-100 puertas. Esta era la escala de integración alcanzada en la década de 1960.
- MSI (*medium scale integration*): 100-1.000 puertas. Años 1970.
- LSI (*large scale integration*): 1.000-10.000 puertas. Años 1980.
- VLSI (*very large scale integration*): 10.000-100.000 puertas). Años 1990.

- ULSI (*ultra large scale integration*): más de 100.000 puertas. Aproximadamente a partir del año 2000.

Estas clasificaciones son solo orientativas para describir la capacidad de un CI. Hay que destacar que la complejidad de un circuito no es un mérito en sí mismo. Todo lo contrario; los diseñadores tienen como reto encontrar la solución más simple, eficiente y elegante que satisfaga los requisitos dados.

Figura 1. Niveles de abstracción de diseño en circuitos digitales



En el proceso de diseño de un circuito integrado encontramos varios niveles de abstracción, como se puede observar en la figura 1. En el nivel más bajo encontramos los dispositivos físicos que realmente hacen posible que el CI ejecute su función. En los niveles de abstracción más elevados lo que se hace conceptualmente es agrupar varios dispositivos según su función, para considerarlos un solo bloque y trabajar con ellos, sin “bajar” al nivel de detalle inferior. En cada nivel, pues, nos encontramos con bloques, funciones y capacidades cada vez más complejas y elaboradas. Estos niveles, de mayor a menor detalle, como se ve en la figura 1, se corresponden con la modelización en el nivel de dispositivos físicos, circuitos, puertas lógicas, módulos y sistemas, respectivamente.

En este módulo explicamos qué materiales y tecnologías se utilizan para fabricar estos dispositivos físicos. Nos situamos, pues, en el nivel de abstracción más detallado. Estos conceptos nos ayudarán a comprender posteriormente de dónde provienen los bloques de niveles superiores y cuál es la base de su funcionalidad.

En todo caso, cabe destacar que un circuito integrado de dimensiones muy pequeñas es capaz de contener multitud de dispositivos miniaturizados y llevar a cabo funciones muy complejas.

Se pueden hacer varios tipos de clasificaciones de CI según el criterio que se utilice. Uno de ellos, que ya se ha comentado anteriormente, es el número de dispositivos que puede contener según sus dimensiones. Otra de estas clasificaciones está basada en el tipo de transistores que incluye. Así, se puede tener un **CI bipolar** si se fabrica sobre la base de transistores bipolares, o se puede tener un **CI MOS** si los transistores que se utilizan son de efecto de campo o más conocidos como MOS (*metal oxide semiconductor*). Al final de este módulo veremos qué estructura física tienen estos tipos de CI y aclararemos estos conceptos nuevos sobre tecnología bipolar y MOS.

Objetivos

Los objetivos que alcanzaréis con el estudio de este módulo son los siguientes:

- 1.** Entender las características generales y el comportamiento básico de los materiales semiconductores y valorar su importancia.
- 2.** Conocer las tecnologías y los materiales empleados en la fabricación de circuitos integrados.
- 3.** Entender el proceso básico de fabricación de los circuitos monolíticos de silicio.

1. Semiconductores

Dentro de un circuito integrado (CI¹) encontramos tres tipos de materiales clasificados según su conductividad eléctrica: aislantes, semiconductores y conductores.

⁽¹⁾Abreviamos *circuito integrado* con la sigla CI.

Los aislantes y los conductores tienen siempre el mismo comportamiento. Por un lado, los aislantes tienen una resistencia idealmente infinita, que hace que por ellos no circule corriente eléctrica. Por otro lado, los conductores tienen una resistencia idealmente nula, de manera que la corriente eléctrica no encuentra oposición a su paso y puede circular libremente. Hay que destacar que, en un caso real, la resistencia de los aislantes toma un valor muy alto, pero no infinito, y por su parte, los conductores tienen una resistencia muy baja, pero no estrictamente igual a cero. En algunos casos hay que tener en cuenta estos efectos no ideales, pero en este apartado simplemente nos interesa destacar estos dos tipos de comportamiento para centrarnos en explicar a continuación el caso intermedio: los semiconductores.

Los semiconductores son unos materiales que nos permiten regular la resistencia que oponen al paso de la corriente eléctrica: en determinadas circunstancias permiten su paso, y en otras situaciones, no lo permiten. La electrónica y la microelectrónica no serían lo que son sin la capacidad de modificar las propiedades conductoras de los semiconductores. Es por este motivo por lo que empezamos con el estudio de estos materiales, su estructura, sus propiedades y las técnicas de manipulación usadas precisamente para poder variar su comportamiento eléctrico.

1.1. Bandas de energía y conducción eléctrica

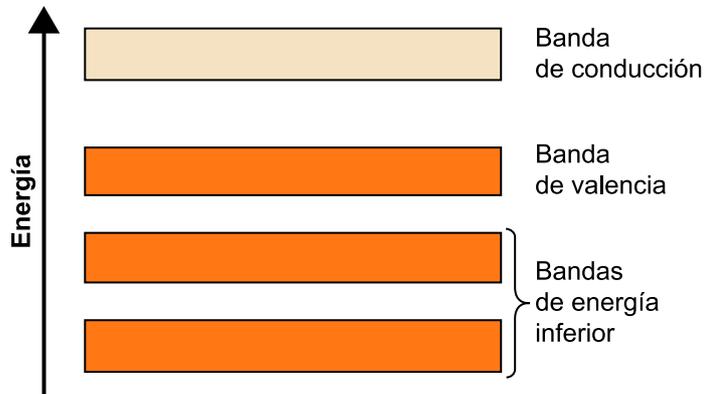
Para comprender el funcionamiento de los semiconductores es necesario revisar brevemente algunos conceptos de física, en concreto referidos a la estructura atómica.

Cuando se unen varios átomos para formar un sólido, los electrones se van disponiendo en niveles de energía nuevos, de manera que cuantos más átomos se unan, más niveles de energía habrá. Cuando el número de átomos que se han unido es muy grande, ya no tiene sentido hablar de niveles discretos de energía (debido a su gran número), sino que los electrones se disponen de forma agrupada en unos ciertos intervalos de energía. Estos intervalos de energía donde se disponen los electrones se denominan **bandas**.

En la figura 2 podéis ver una distribución esquemática de las bandas donde se disponen los electrones cuando tenemos un sólido. Observad que las bandas por donde se distribuyen los electrones no son continuas, sino que hay una

cierta separación (en términos de energía) entre las diversas bandas. Nótese que no tienen por qué estar uniformemente distribuidas, y que los estados posibles de energía se encuentran **solo** en estas bandas; los espacios entre ellas son estados no permitidos.

Figura 2. Distribución de bandas de energía en la estructura atómica de un sólido



En esta figura, además, se hace mención de dos bandas muy importantes:

- La **banda de valencia**, cuando sobre el material no se aplica ninguna energía externa, es la banda de energía más elevada en la que se disponen los electrones. En esta banda, los electrones están ligados al átomo.
- La **banda de conducción** es una banda de energía superior a la de valencia. En la banda de conducción los electrones se pueden mover libremente, y no están ligados a ningún átomo.

Para que un electrón de un átomo abandone la banda de valencia y pase a la banda de conducción, hay que aplicarle una cierta energía, correspondiente a la separación que hay entre estas dos bandas.

Las bandas con energía inferior a la de valencia no afectan a la conducción y, por lo tanto, no las tendremos en cuenta en adelante.

Como se puede ver en la figura 2, el nivel de energía de la banda de conducción está por encima del de la banda de valencia. La banda de energía entre la banda de valencia y la de conducción es la banda de energía prohibida para los electrones, y se denomina **espacio de energía**⁽²⁾.

⁽²⁾En inglés, *energy gap* o *band gap*.

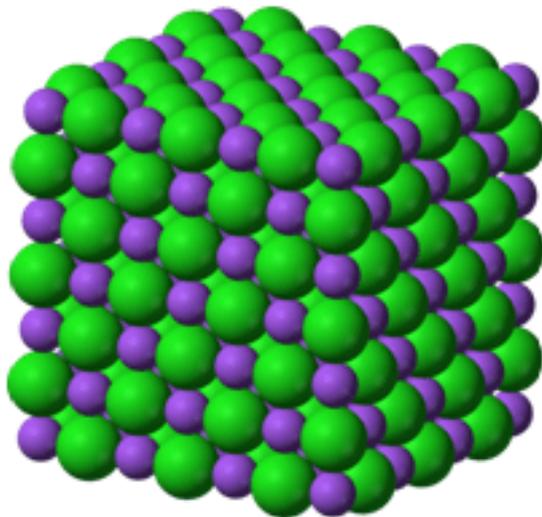
Para que un electrón pase de una banda a otra, tendrá que absorber o emitir energía suficiente, correspondiente a la diferencia de energía entre bandas. Un electrón que llegue a tener suficiente energía para alcanzar la banda de conducción se denomina electrón libre. Serán este tipo de electrones los que contribuirán a la corriente eléctrica.

Hay que darse cuenta también de que este modelo de energías proviene de la propia estructura de los átomos que componen el sólido, y en concreto hay que destacar el concepto de estructura cristalina. La **estructura cristalina** es una forma de ordenar y empaquetar los átomos, moléculas o iones en algunos materiales sólidos. Muchos sólidos se empaquetan de manera ordenada y con patrones de repetición que se extienden en las tres dimensiones del espacio. De hecho, según la distribución espacial de estos átomos, moléculas o iones, los materiales sólidos pueden ser clasificados como:

- **Cristalinos:** Compuestos por átomos, moléculas o iones organizados de una forma periódica en tres dimensiones. Las posiciones ocupadas siguen una ordenación que se repite para grandes distancias atómicas (de largo alcance).
- **Amorfas:** Compuestos por átomos, moléculas o iones que no presentan una ordenación de largo alcance, aunque sí pueden presentar una ordenación de corto alcance.

En el caso de los semiconductores, basados en materiales como por ejemplo el silicio (Si) o el germanio (Ge), nos encontramos ante materiales cristalinos.

Figura 3. Retícula de una estructura cristalina

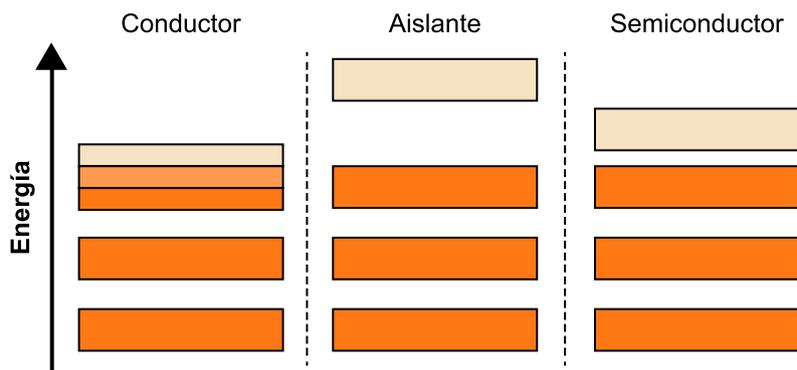


Font: Wikipedia, licencia Creative Commons.

En la figura 3 podemos ver una representación de estructura cristalina, en la que cada átomo ocupa una posición en la retícula, y representamos el tipo de átomo por un color.

Visto este concepto de bandas de energía en una estructura cristalina, ya somos capaces de explicar en estos términos qué son los materiales aislantes, los conductores y los semiconductores. En la figura 4 observamos la distribución de las bandas de energía en todos ellos.

Figura 4. Esquema de la distribución de las bandas de energía en materiales conductores, aislantes y semiconductores



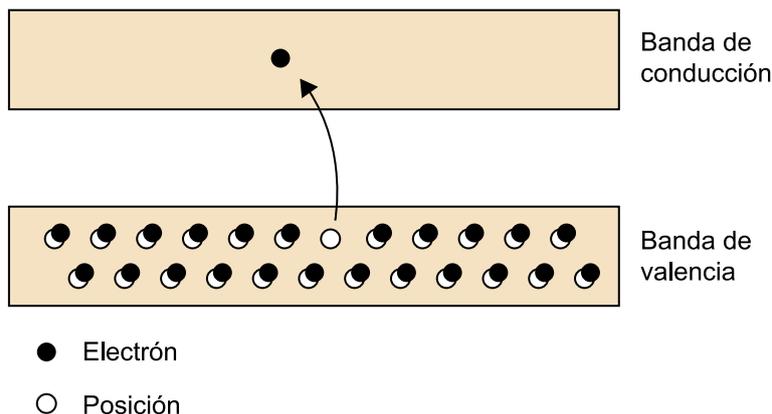
Fijaos en la distancia en términos de energía que hay entre la banda de valencia (la última banda de color más intenso) y la banda de conducción (la banda superior, de color más pálido). En los materiales conductores, la banda de valencia y la de conducción se superponen. En los materiales aislantes, las dos bandas están muy distanciadas (en términos de energía) entre sí. En los materiales semiconductores, están relativamente cercanas.

En los materiales conductores, las dos bandas de energía se llegan a superponer, de modo que siempre hay un número relativamente alto de electrones disponibles (libres) para transportar corriente eléctrica en caso de estar sometidos a un campo eléctrico externo. En los materiales aislantes, las dos bandas de energía están separadas por una banda de energía prohibida muy grande, de modo que en condiciones normales (refiriéndonos básicamente a la temperatura y al valor del campo eléctrico aplicado) no hay prácticamente electrones en la banda de conducción, y por lo tanto, no circulará corriente eléctrica. En cambio, en los semiconductores, la banda de energía prohibida es relativamente pequeña, de forma que en condiciones normales tienen una cierta conductividad, aunque muy menor que la de los conductores.

En este punto vemos, pues, un concepto muy importante: este modelo de bandas de energía nos ha permitido explicar en qué materiales se produce conducción eléctrica y en cuáles no.

En la figura 5 podéis observar cómo cada electrón (círculo oscuro) ocupa una posición (círculo vacío) en la banda de valencia. También podéis ver cómo un electrón al que se ha aplicado la energía necesaria ha dejado la banda de valencia para pasar a la de conducción, donde puede moverse libremente. En esta figura, solo hay dibujadas las bandas de valencia y de conducción porque, como hemos comentado antes, las bandas inferiores no afectan a la conducción.

Figura 5. Electrón que alcanza la banda de conducción



Al pasar a la banda de conducción, el electrón ha dejado una posición “vacía”. El átomo que ha perdido este electrón se queda con un electrón menos que su número de protones, de manera que tiene una carga eléctrica global igual a la de un electrón, pero con signo positivo.

El átomo que ha quedado con carga positiva ejerce una fuerza de atracción sobre los electrones de otros átomos que tiene cerca dentro de la estructura cristalina. Estos electrones tenderán a ocupar la posición vacía que había quedado en este átomo. En muchos casos, esta posición la ocupará algún electrón de un átomo cercano, pero a la vez dejará una posición vacía en el átomo que ocupaba hasta entonces. Esta posición vacía también tenderá a ser ocupada por otro electrón, y así sucesivamente.

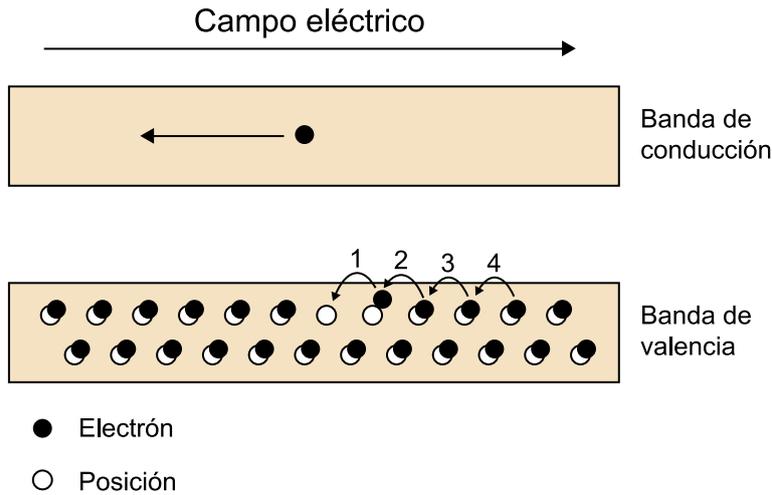
En condiciones normales, tanto el movimiento de electrones libres en la banda de conducción como el movimiento de electrones en la banda de valencia para ocupar posiciones vacías sucesivas se producen en todas direcciones de manera aleatoria, y no se forma una corriente eléctrica. Ahora bien, si el material semiconductor está atravesado por un campo eléctrico (como el que genera una fuente de tensión conectada a los extremos del material), los electrones se moverán en esta dirección concreta, pero en sentido contrario, porque son cargas negativas.

En la figura 6 podéis ver un esquema unidimensional del movimiento de los electrones en un material semiconductor en presencia de un campo eléctrico.

Observad que se producen dos tipos de movimientos:

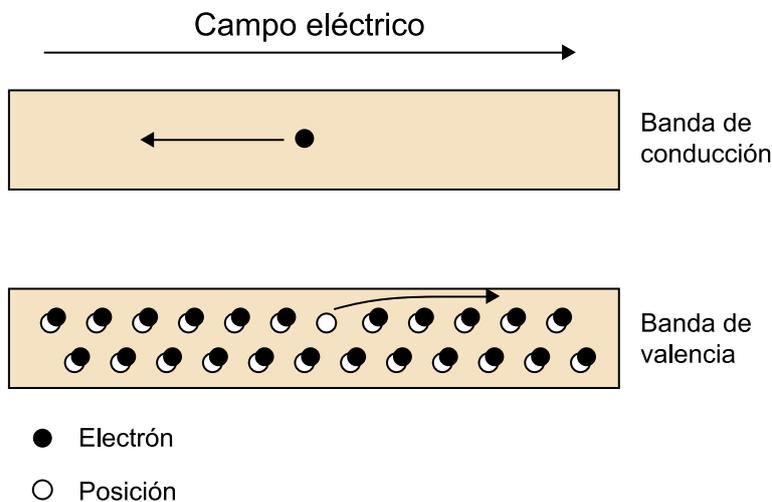
- Los electrones libres de la banda de conducción se mueven libremente en dirección contraria al campo eléctrico.
- Los electrones de la banda de valencia van ocupando sucesivamente las posiciones que van quedando vacías. Este movimiento también se produce en sentido contrario al campo eléctrico.

Figura 6. Movimiento de los electrones en presencia de un campo eléctrico



Fijaos que este segundo tipo de movimiento es equivalente a considerar que en la banda de valencia, en lugar de tener unos electrones sucesivos que se mueven en sentido contrario al campo eléctrico, tenemos una partícula (la posición vacía) que se mueve en el mismo sentido que el campo eléctrico. En la figura 7 podéis ver este movimiento aparente de la posición vacía en la banda de valencia, al que denominaremos **hueco**.

Figura 7. Movimiento aparente de la posición libre en la banda de valencia (hueco)



Los huecos son las posiciones vacías que han dejado los electrones en la banda de valencia. Aunque físicamente un hueco solo es la posición donde antes había un electrón, eléctricamente se la considera una partícula que tiene la misma carga que un electrón, pero con signo positivo.

Hay dos efectos que ya hemos visto hasta ahora que sostienen esta consideración del hueco como partícula eléctrica positiva:

- Cuando un electrón abandona la banda de valencia, deja en su lugar un hueco. Como hemos visto, este hueco atrae a los electrones de los átomos cercanos, como haría una partícula eléctrica positiva.
- Cuando tenemos un campo eléctrico en un material semiconductor, en la banda de valencia se produce un movimiento sucesivo de electrones en sentido contrario al campo eléctrico. Este movimiento es equivalente a un movimiento aparente de un hueco en el mismo sentido que el campo eléctrico, como haría una partícula eléctrica positiva.

Tanto el movimiento de electrones libres en la banda de conducción como el movimiento de huecos en banda de valencia en presencia de un campo eléctrico conforman una corriente eléctrica en el mismo sentido que el campo.

En los materiales semiconductores, la corriente eléctrica se produce por el movimiento de electrones libres o por el movimiento de huecos. Denominamos **portadores** a las partículas (electrones libres o huecos) que en un material semiconductor producen la corriente eléctrica.

Hasta aquí habéis visto qué son los materiales semiconductores y cómo conducen la corriente eléctrica. También hemos destacado la importancia de los portadores, tanto de los electrones libres como de los huecos; su importancia proviene del hecho de que son los causantes de esta corriente eléctrica. En el siguiente subapartado veremos cómo podemos modificar sus características eléctricas para controlar la conducción mediante el dopaje.

1.2. Dopaje de materiales semiconductores

En el subapartado anterior hemos visto cómo se produce la corriente eléctrica en un semiconductor a partir del movimiento de los electrones libres o de los huecos. De todas maneras, hay que tener en cuenta que, para llegar a producir esta corriente eléctrica, hemos tenido que aplicar una cierta energía a un electrón para que deje la banda de valencia y pase a la de conducción y, obtener así un hueco. Si no aplicamos energía a los electrones de valencia no se podrá producir corriente eléctrica puesto que tendremos un material que tendrá todas las posiciones de la banda de valencia ocupadas por electrones (es decir, no tendremos ningún hueco), y la banda de conducción vacía (sin electrones libres). Además, para obtener una corriente eléctrica suficiente, nos interesa tener un número considerable de electrones libres o de huecos en movimiento.

Para mejorar la conducción del material podemos sustituir uno de los átomos de la red cristalina del material semiconductor por un átomo de un material diferente que tenga un electrón de menos. Esto dejará un hueco en el material, que se podrá mover en presencia de un campo eléctrico. También podemos sustituir este átomo por otro que tenga un electrón de más, que se podrá mo-

ver libremente en la banda de conducción si aplicamos un campo eléctrico. Si en lugar de sustituir un solo átomo, lo hacemos con varios, favoreceremos la conducción eléctrica (de electrones libres o de huecos, según el tipo de material que añadamos) en el semiconductor. Esta es la base del dopaje, que veremos en este apartado.

Una de las ventajas de trabajar con semiconductores es que podemos modificar su comportamiento eléctrico añadiendo al semiconductor una pequeña cantidad de átomos diferentes, con el objetivo de favorecer la presencia de portadores, ya sea electrones libres o huecos.

Los átomos que se añaden al semiconductor para aumentar la presencia de portadores se denominan **impurezas**.

En resumen, denominamos dopaje al proceso de añadir a un material semiconductor impurezas que contienen electrones libres o huecos. Un material al cual se ha inyectado impurezas se dice que está **dopado**.

Al dopar un material semiconductor, el número de electrones libres o de huecos que contiene aumenta. Al tener más densidad de portadores, la conductividad de este material aumenta. Es decir, favorecemos el paso de corriente a través del semiconductor.

Hay dos tipos de dopaje, dependiendo del tipo de impurezas que añadimos a un material semiconductor:

- En el **dopaje de tipo P** añadimos al semiconductor impurezas con huecos.
- En el **dopaje de tipo N** añadimos al semiconductor impurezas con electrones libres.

Según qué tipo de dopaje se haya aplicado a un material semiconductor, también se habla, respectivamente, de **semiconductores de tipo P** y **semiconductores de tipo N**.

En los semiconductores, las impurezas se encuentran en una concentración mucho menor que la de los átomos propios del semiconductor. Se puede decir que para dopar un material se añade un número de impurezas muy pequeño. Por ejemplo, se dice que un material tiene un **dopaje alto** o **fuerte** cuando hay un átomo de impureza por cada 10.000 átomos de semiconductor. En los dopajes normales, el número de átomos de impurezas es todavía menor.

1.3. La unión PN

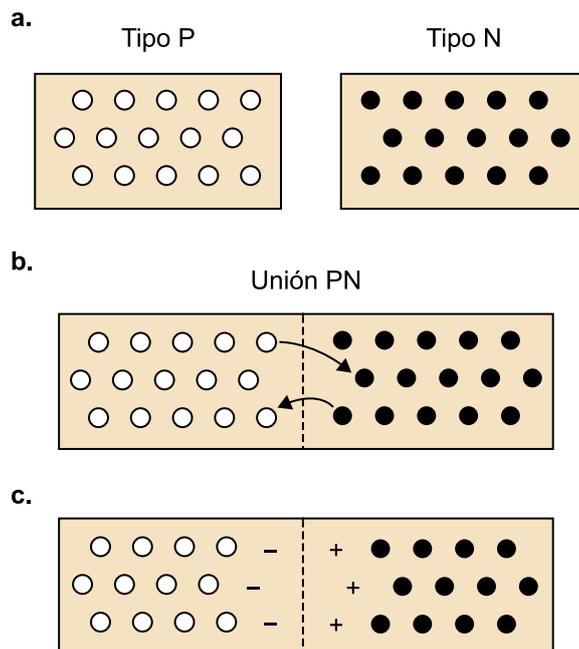
Una **unión PN** es el resultado de unir dos materiales semiconductores, uno con un dopaje de tipo P y el otro con un dopaje del tipo N.

En la figura 8 podemos ver qué pasa cuando unimos un semiconductor de tipo P con uno de tipo N para formar una unión PN. En esta figura hemos representado esquemáticamente las partículas portadoras, es decir, los huecos (con un círculo vacío) y los electrones libres (con un círculo lleno) en la red cristalina.

Recordad

Los electrones libres se mueven en la banda de conducción, mientras que los huecos se mueven en la banda de valencia.

Figura 8. Formación de una unión PN



En la figura 8a podéis observar la distribución de los portadores en los semiconductores P y N antes de unirlos. Como hemos visto, en el semiconductor P los portadores mayoritarios son los huecos (círculos vacíos), mientras que en el semiconductor N los portadores mayoritarios son los electrones libres (círculos llenos).

Ved también

Véase el dopaje de los materiales semiconductores en el subapartado 1.2 de este módulo didáctico.

Cuando unimos los dos semiconductores, se produce un proceso conocido como de difusión entre sus respectivos portadores. Es decir, como se puede ver en la figura 8b, en las zonas cercanas a la unión de los dos materiales, los electrones libres del semiconductor N se mueven hacia el semiconductor P. A la vez, los huecos de la región P se mueven hacia la región N. Esta entrada de portadores a las dos zonas espaciales es el proceso denominado **difusión**, es decir, se “difunden” hacia la otra zona.

Además de esto, y dada la naturaleza de los portadores (son electrones y huecos), en toda la zona de difusión se produce un proceso denominado recombinación. La recombinación es el proceso por el cual un electrón libre (que se encuentra en la banda de conducción) pasa a la banda de valencia y se une con un hueco. Este electrón, pues, queda ligado al átomo en la banda de valencia.

Fijaos en que cuando un electrón libre se recombina con un hueco, obtenemos un electrón en banda de valencia que está ligado al átomo y no conduce corriente eléctrica. Es decir, se pierden dos portadores eléctricos (el hueco y el electrón libre, que se han recombinado).

El semiconductor P, que al principio era eléctricamente neutro, ha perdido huecos (cargas positivas) en la zona cercana a la unión. De este modo, mientras se va produciendo la difusión y recombinación, en esta zona se va generando una carga global negativa. Esta carga ejerce una fuerza de repulsión sobre los electrones libres del semiconductor N, que va frenando el proceso de difusión.

En el semiconductor N hay un proceso análogo. Al principio, también era eléctricamente neutro, pero ha ido perdiendo electrones libres (cargas negativas) en la zona cercana a la unión. A medida que se va produciendo la difusión y recombinación, en las zonas cercanas a la unión del semiconductor N va quedando una carga global positiva que ejerce una fuerza de repulsión sobre los huecos del semiconductor P. Este hecho también frena el proceso mismo de difusión.

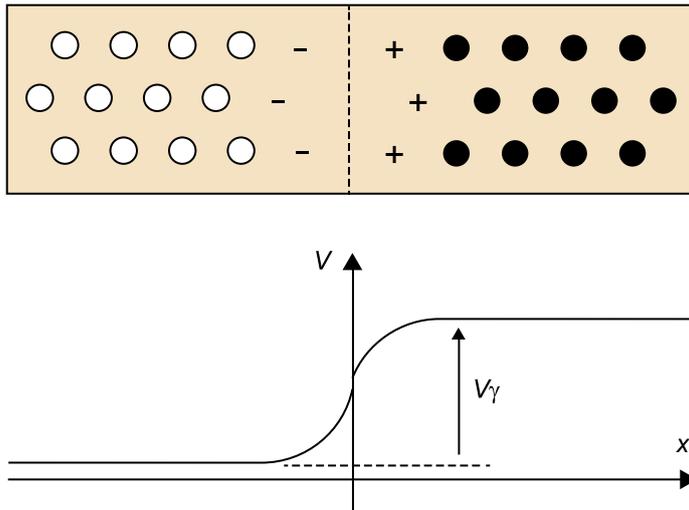
El proceso de difusión y recombinación continúa hasta que se llega a un estado de equilibrio, que podéis ver en la figura 8c. Fijaos en que en la zona central de la unión PN no hay portadores, pero ha quedado con carga eléctrica (negativa en el semiconductor P y positiva en el semiconductor N). Esta zona cargada pero libre de portadores se denomina **zona de carga espacial** (ZCE³).

⁽³⁾Abreviamos *zona de carga espacial* con la sigla ZCE.

Así pues, en esta zona de carga espacial aparece un campo eléctrico que provoca una diferencia de potencial entre la zona N y la zona P. Esta tensión que aparece, llamada **barrera de potencial**, es pequeña, generalmente de unas cuantas décimas de voltio, aunque su valor depende del material semiconductor utilizado. La barrera de potencial en ausencia de fuentes de tensión externas evita que los electrones libres de la zona N se muevan hacia la zona P, y que los huecos de la zona P se puedan mover hacia la zona N.

En la figura 9 podéis ver la unión PN una vez llegado al equilibrio, tal y como ya habíamos visto en la figura 8c. Bajo la unión PN, podéis observar la diferencia de potencial que forma la barrera de potencial (marcada como V_p) que se produce entre las zonas P y N de la unión PN. En esta figura el eje horizontal, denominado x , representa la separación física o distancia entre los dos materiales.

Figura 9. Unión PN en equilibrio con barrera de potencial formada



En este estado de equilibrio, la unión PN ya ha generado una barrera de potencial entre los extremos del material formado por dos zonas P y N. Vemos, pues, que disponemos ya de un bloque de material semiconductor que dejará pasar la corriente eléctrica entre sus extremos, siempre y cuando esta corriente eléctrica sea capaz de superar esta barrera de potencial, es decir, que vaya en la dirección correcta. Es un primer paso hacia nuestro objetivo de obtener un material semiconductor con el cual podamos regular la conductividad eléctrica de forma controlada.

Hay que destacar, además, que podríamos hacer crecer o disminuir esta barrera de potencial solo aplicando una tensión adicional entre los extremos del bloque PN. Si hacemos aumentar la barrera, la conductividad eléctrica será menor, y a la inversa. Disponemos, pues, de un **material apto para tener control sobre la conducción eléctrica**.

El hecho de disponer de un material semiconductor basado en una unión PN con el cual podamos regular la conductividad eléctrica de forma simple (con solo aplicar una tensión para aumentar o reducir la barrera de potencial) es la base de la electrónica actual. Los semiconductores permiten obtener dispositivos aptos para controlar a voluntad el paso de la corriente eléctrica, y por lo tanto, son la base de diodos, transistores, puertas lógicas, etc., y nos permiten, en definitiva, disfrutar de circuitos integrados muy complejos compuestos por millones de uniones PN con funciones muy avanzadas.

En este subapartado hemos visto qué son los materiales semiconductores. Estos materiales conducen la corriente eléctrica mediante el movimiento de electrones libres en la banda de conducción o el movimiento de huecos en la banda de valencia. Por medio del dopaje, podemos regular su conductividad. Finalmente, hemos visto que podemos unir dos materiales semiconductores, uno de tipo P y el otro de tipo N, para obtener la unión PN, que es la base de

los dispositivos electrónicos fundamentales en nuestros circuitos, como por ejemplo diodos y transistores. Es necesario destacar los conceptos de semiconductores y dopaje para poder entender el subapartado siguiente, orientado a explicar cómo se fabrican los circuitos integrados. En este proceso de fabricación, tanto los semiconductores como su proceso de dopaje juegan un papel destacado. También veremos cómo conseguir formar uniones PN durante estos procesos de fabricación.

2. Procesos de fabricación de circuitos integrados de silicio

La diversidad de tecnologías, los procesos involucrados y la cantidad de productos en el sector de la microelectrónica es inmensa. Continuamente se investiga en nuevos materiales y procesos que permitan mejorar las prestaciones de los CI. Pero una buena parte, por no decir una amplia mayoría, de los CI en producción actuales, están basados en el silicio.

Las principales razones del uso tan extensivo del silicio son la combinación de buenas propiedades semiconductoras y térmicas, junto con versatilidad química, estabilidad mecánica y una gran disponibilidad. Pero no siempre ha sido así; de hecho, los primeros dispositivos semiconductores comerciales fueron diodos y transistores de germanio (Ge). En realidad, el silicio no dispone de las mejores características entre los semiconductores. Por ejemplo, los portadores del Ge tienen mejor movilidad, de manera que los componentes fabricados sobre Ge responden mejor para trabajar a altas frecuencias. El Si adquirió su posición dominante, que perdura hoy en día, sobre todo gracias a la introducción de la **tecnología planar**, fundamental para la creación de CI. Tal como se desprende de los procesos que veremos a continuación, la generación de óxido protector sobre el sustrato semiconductor es fundamental para crear CI, y el Si es el semiconductor donde se puede crear este óxido más fácilmente. Se denomina tecnología planar en el sentido de que los CI se crean sobre el plano que forma la oblea, que, como veremos, es un disco de materiales semiconductores.

El proceso de fabricación de CI, así como de elementos semiconductores discretos, incluye un conjunto de procesos o procedimientos que han de ser ejecutados con mucho cuidado. Es un proceso complejo en el cual intervienen numerosas etapas de fotolitografía y procesamiento químico, durante las cuales los circuitos se generan sobre una oblea hecha de materiales puramente semiconductores.

Los procesos principales incluyen:

- Creación de obleas de silicio
- Oxidación
- Fotolitografía y máscara
- Grabado. Eliminación de capas no deseadas.
- Impurificación con materiales dopantes (difusión e implantación)
- Formación de capas delgadas (deposición, epitaxia y capas metálicas)

El silicio

El silicio es un elemento químico no metálico de la tabla periódica que tiene el símbolo Si con número atómico 14. Es el segundo elemento más abundante de la corteza terrestre (representa el 25,7% de su peso) después del oxígeno.

Estos procesos terminan fijando las características eléctricas y las restricciones geométricas y topologías de los elementos de base (transistores y conexiones).

A estos procesos hay que añadir otras operaciones tales como: montaje, pruebas, marcado, empaquetado, embalado y expedición; que no explicamos en este módulo puesto que no tienen relación directa con el proceso de fabricación.

El proceso de fabricación empieza con la creación de una oblea de silicio. Esta oblea se manipulará repetidamente durante la fabricación del CI aplicando en una secuencia determinada el resto de procesos comentados antes. Como resultado obtendremos una oblea procesada, que contendrá un conjunto de chips individuales que se separan físicamente de la oblea mediante cortes. Cada oblea, pues, permite fabricar tantos chips como veces quepa el área de un chip en el área de la oblea. Generalmente, decenas o centenares de chips por oblea.

Los componentes que se fabrican dentro de un CI son muy numerosos, y son tan pequeños que cualquier defecto en su producción puede derivar en un funcionamiento defectuoso. Para minimizar los CI defectuosos, y que deberán descartarse, todos los procesos de fabricación de la oblea se tienen que ejecutar en espacios cuanto menos contaminados mejor. Para alcanzar este objetivo, la fabricación de CI se realiza casi en su totalidad dentro de una sala limpia. Una **sala limpia** es una instalación donde hay un control de contaminación ambiental muy exhaustivo. El coste de la materia prima de los CI es relativamente bajo. En cambio, el coste de las instalaciones y el de la maquinaria de fabricación son en comparación muy altos.

Dentro de una sala limpia hay una exigencia de limpieza constante y exhaustiva que se logra mediante métodos químicos (ácidos, cáusticos, disolventes) puntualmente, y físicos (agua desionizada con detergente, filtros de aire, atmósfera a mayor presión que la del exterior) algunos de ellos continuos. También hace falta una indumentaria especial para el personal.

Solo si todos los procesos y las fases están bien caracterizados y parametrizados, y se ejecutan en salas limpias, obtendremos un **rendimiento de fabricación** aceptable, donde por rendimiento entendemos la relación entre el número de chips correctos con respecto al número de chips totales fabricados.

En los siguientes subapartados describiremos cada uno de los procesos.

2.1. Creación de la oblea

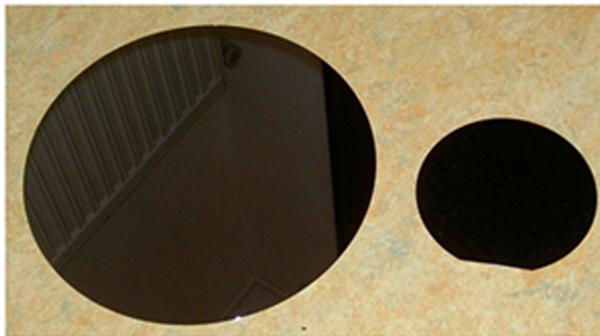
El material inicial para los circuitos integrados modernos es el silicio de muy alta pureza, que adquiere la forma de un cilindro sólido de color gris de entre unos 10 a 30 cm de diámetro y puede ser de 1 m a 2 m de longitud.

Durante la fabricación de estos lingotes de silicio es cuando se añaden las primeras impurezas necesarias para formar un semiconductor de tipo P o N, con el dopaje deseado. Normalmente la concentración de impurezas es de 10^{15} átomos en cada cm^3 . Para alcanzar esta concentración se incorpora al silicio fundido una pequeña cantidad de dopante, por ejemplo fósforo (para conseguir un semiconductor de tipo N) o boro (para uno de tipo P) antes de que se solidifique en un cilindro. En procesos posteriores también se añadirán dopantes según interese.

Una vez se dispone de los cilindros de silicio dopado en estado sólido, se cortan en rodajas para producir obleas circulares de $400\ \mu\text{m}$ a $600\ \mu\text{m}$ de grosor. Este proceso se basa típicamente en hacer un corte con láser o con una sierra con hilo de diamante, para poder tener la precisión suficiente y obtener obleas lo bastante finas.

Después, la pieza se pule hasta obtener un acabado de espejo a partir de técnicas de pulido químicas y mecánicas. Las propiedades eléctricas y mecánicas de la oblea dependen de la orientación de los planos cristalinos y de la concentración de átomos y de impurezas.

Figura 10. Obleas de silicio de 12" (30,8 cm) y 6" (15,4 cm) de diámetro



Cada una de las obleas permitirá obtener posteriormente decenas o centenares de “pastillas”, cada una de las cuales formará el cuerpo de un CI. El tamaño de cada pastilla puede ser variable, y generalmente varía entre el de un pin y el de un sello. Lo que se hace es repetir el patrón de un conjunto de pastillas para obtener diferentes tipos y tamaños en una misma oblea. Por ejemplo, podríamos pensar en cuatro tipos de pastillas diferentes, y de cada oblea saldrían centenares de cada tipo.

2.2. Oxidación

Además de las propiedades semiconductoras del silicio, la razón principal que ha llevado al silicio a ser el material más utilizado en la fabricación de circuitos integrados es la habilidad de formar sobre él una capa de óxido estable, de buena calidad y con magníficas propiedades aislantes.

⁽⁴⁾ $1\ \mu\text{m}$ es igual a $10^{-6}\ \text{m}$.

Reflexión

Aunque se podrían fabricar CI con obleas más finas, no se reduce su grosor para poder disponer de una superficie física mínima y que las obleas no sean demasiado quebradizas.

La oxidación se refiere al proceso químico de reacción del silicio con el oxígeno para formar dióxido de silicio (SiO_2).

Una vez limpia la superficie de la oblea con detergentes y disolventes (como último paso del proceso descrito en el subapartado anterior), se somete a un proceso de oxidación para crear una capa de óxido en su superficie.

En primer lugar, esta capa protege la superficie de posible contaminación, y además sirve de **máscara** en el proceso de difusión posterior, que veremos más adelante. La capa de óxido se puede formar bien por **oxidación térmica** o bien por **deposición**:

- En el caso de la **deposición**, los dos elementos, Si y O_2 , se dirigen a la superficie de la oblea y reaccionan formando una capa SiO_2 .
- En el proceso de **oxidación térmica**, se produce una reacción entre los átomos de silicio de la superficie de la oblea y oxígeno dentro de un horno a alta temperatura. El óxido creado por oxidación térmica tiene mucha más calidad que el que se obtiene mediante deposición.

Para acelerar la reacción de oxidación térmica se necesitan hornos especiales para alta temperatura muy limpios. El oxígeno que se utiliza en la reacción se introduce como un gas de alta pureza (proceso de **oxidación seca**) o como vapor (**oxidación húmeda**). La oxidación húmeda tiene una mayor tasa de crecimiento, aunque la oxidación seca produce mejores características eléctricas.

El rango de temperaturas en el cual se produce esta reacción de oxidación térmica está comprendido entre $850\text{ }^\circ\text{C}$ y $1.100\text{ }^\circ\text{C}$, mientras que el silicio no se funde hasta los $1.412\text{ }^\circ\text{C}$. La temperatura de oxidación se mantiene bastante por debajo, para evitar la generación de defectos en el cristal y el movimiento de los dopantes añadidos.

El dióxido de silicio resultante sobre la oblea es una película delgada transparente y aislante, cuya superficie es altamente reflectante. Durante este proceso de oxidación, parte de la capa original de silicio se consume, de manera que el dióxido se introduce en el silicio original. Por cada micra de dióxido crecido, se consumen $0,44$ micras de Si.

Si se ilumina con luz blanca una oblea oxidada, la interferencia constructiva y destructiva hará que ciertos colores se reflejen en ella. Sobre la base del color de la superficie de la oblea se puede deducir el grosor de la capa de óxido.

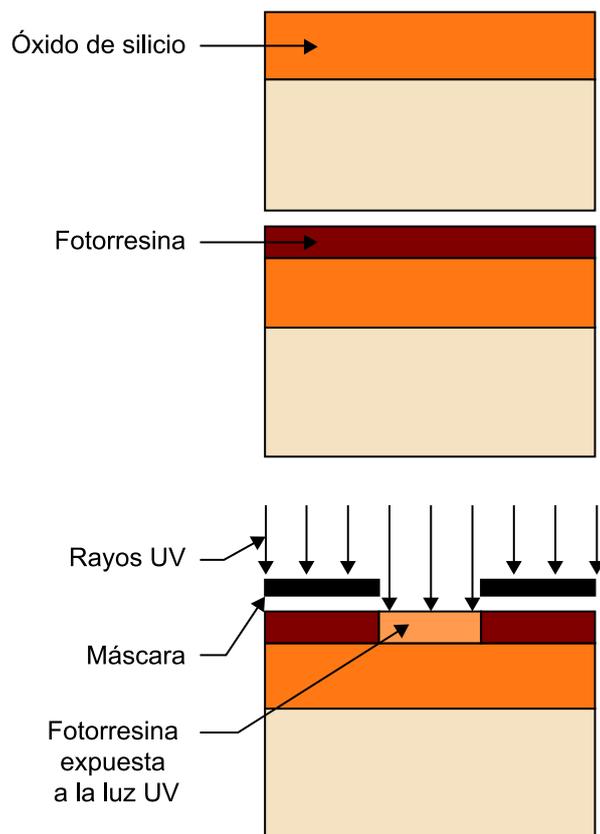
2.3. Fotolitografía y máscara

Una vez creada la capa de aislante de óxido sobre la oblea, parte de ella tiene que ser selectivamente eliminada en aquellos lugares donde hay que introducir átomos de dopante. Este grabado selectivo se realiza generalmente mediante el uso de un material sensible a la luz denominado **fotorresistencia** o **fotorresina**.

Así, en primer lugar, la oblea oxidada se cubre con una capa de fotorresistencia. Después, se utilizará una placa fotográfica con patrones dibujados, es decir, una **máscara**, para exponer de forma selectiva la capa fotosensible a la **iluminación ultravioleta (UV⁵)**. Las áreas expuestas se ablandarán y podrán ser eliminadas con un agente químico de iones reactivos, y de este modo producir geometrías de superficies muy finas de forma precisa. La capa fotosensible se utiliza pues para proteger por debajo los materiales contra el ataque químico. Este proceso exige una alta precisión geométrica, que impone requisitos mecánicos y ópticos muy rigurosos al equipo de fotolitografía.

⁽⁵⁾Abreviamos *ultravioleta* con la sigla UV.

Figura 11. Proceso de fotolitografía



La secuencia es la siguiente:

1) Deposición de resina fotosensible sobre toda la oblea.

2) Iluminación sobre ciertas áreas de la oblea. Se consigue colocando una máscara entre la fuente de iluminación y la oblea.

3) La resina que ha sido iluminada adquiere propiedades diferentes de la resina no iluminada. Esto permite que sea disuelta de forma selectiva con un revelador. Se puede eliminar la resina iluminada (resina positiva) o la resina no iluminada (resina negativa).

4) Ataque químico, que solo actúa sobre la resina positiva o bien la resina negativa (no ambas a la vez).

5) Eliminación de la resina restante.

Los tres primeros pasos se encuentran representados en la figura 11.

La iluminación ha sido históricamente siempre de tipo UV, pero desde hace tiempo los haces de electrones, los rayos X y los láseres van ganando terreno por sus mayores resoluciones.

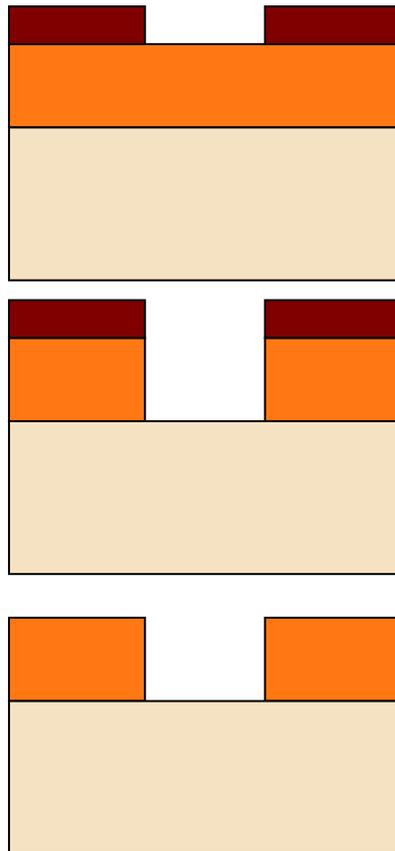
2.4. Grabado. Eliminación de capas no deseadas

El proceso siguiente se denomina grabado y consiste en eliminar la parte de óxido no protegida por la resina remanente. De este modo se abrirán las ventanas deseadas en el óxido, que dejarán a la vista el sustrato de silicio. Siguiendo con el esquema de la figura 11, en la figura 12 se ve el esquema de lo que se pretende conseguir: en primer lugar, eliminar el óxido no tapado por la resina, y a continuación, limpiar la superficie de la resina una vez esta ya ha hecho su función.

El **grabado húmedo** se utiliza más por ser más económico, y consiste en la utilización de líquidos (generalmente ácidos) que reaccionan con los materiales que se quieren eliminar. Se puede utilizar por ejemplo un baño de ácido fluorhídrico, que ataca el óxido de silicio no protegido, pero no ataca el silicio.

El **grabado seco** consiste en la aplicación de gases en una atmósfera vacía de alta energía. En primer lugar se ioniza un gas inerte, y se aceleran los iones mediante un campo eléctrico, para hacerlos colisionar contra la oblea y acabar eliminando la capa no deseada. Este primer paso trata pues de un proceso puramente físico y no químico. Por otro lado, al mismo tiempo se puede hacer que otros tipos de gases que son capaces de reaccionar químicamente con el sustrato se ionicen y creen radicales. Los radicales reaccionan químicamente con la superficie del material y eliminan el óxido.

Figura 12. Eliminación del óxido expuesto, y a continuación, de la resina



El grabado húmedo es más selectivo pero también más isotrópico, es decir, ataca por igual en todas las direcciones, de manera que también elimina material en los bordes exteriores de la fotorresina. Como consecuencia, se pierde resolución en la transferencia del dibujo de la máscara. En cambio, el grabado seco es un ataque en la dirección vertical, y por lo tanto, respeta los contornos de la geometría que se quiere obtener.

Después de haberse realizado el grabado, se elimina la capa de fotorresina con un disolvente orgánico, como por ejemplo, ácido sulfúrico.

2.5. Impurificación con materiales dopantes (difusión e implantación)

Una vez abiertas las ventanas en el óxido, se realiza el proceso de impurificación mediante la adición de dopantes. Con estas técnicas se puede dopar selectivamente el sustrato del semiconductor (silicio en este caso) y producir regiones de tipo P o de tipo N, según convenga. Hay dos métodos para impurificar el sustrato: difusión e implantación iónica.

Reflexión

Debemos aclarar que estamos intentando dopar el silicio para crear un material semiconductor con las propiedades vistas en el subapartado anterior, y que no pretendemos dopar el óxido, puesto que esto no nos aportaría ningún beneficio.

2.5.1. Difusión

Para introducir impurezas en el interior del semiconductor, se colocan las obleas en el interior de un horno por el cual se hace pasar un gas inerte que contenga el dopante deseado. El sistema es similar al utilizado en la oxidación térmica. Los rangos de temperatura van entre 800 °C y 1.200 °C. La penetración de material dopante en la oblea aumenta con la temperatura y el tiempo de exposición.

La idea principal es, pues, **difundir** las impurezas dentro del silicio, que se implantarán en las áreas no cubiertas por el óxido (que son las ventanas obtenidas con la fotolitografía).

La concentración de dopante resultante disminuye monótonamente a medida que nos adentramos en el silicio (en el eje vertical en la figura 12). Esta técnica tiene el inconveniente de que las impurezas no se difunden solo verticalmente (siguiendo la línea vertical marcada por la ventana en el óxido) sino también en cierta medida en sentido lateral hacia adentro del silicio.

2.5.2. Implantación iónica

En esta técnica, los átomos del dopante son implantados en el interior del semiconductor por medio de haces iónicos de alta energía. La zona dopada tiene un máximo de concentración de dopantes en el interior del semiconductor, y la posición de este máximo estará determinada por la masa de los iones y la energía con la cual inciden sobre la superficie semiconductor.

Las ventajas de la implantación iónica sobre la difusión son un control preciso de la cantidad de dopantes introducidos y su concentración en el material dopado, y una menor temperatura de proceso.

Al introducir impurezas, se producen daños en el cristal, y por eso después se somete la oblea a un proceso de calentamiento en un horno⁶ para reordenar el cristal con las nuevas impurezas.

⁶En inglés, *annealing*.

2.6. Formación de capas delgadas (deposición y epitaxia)

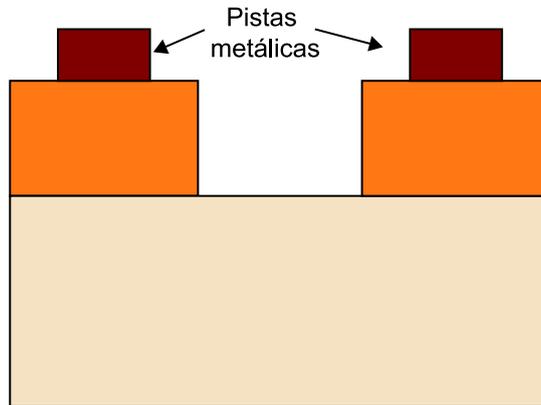
En el proceso de fabricación de los CI, es necesario que en algún momento se formen capas delgadas sobre el material semiconductor de base (silicio en el caso de estudio). Esto es así para poder disponer de pistas metálicas en los circuitos, por ejemplo, u otros tipos de capas que también vamos a describir.

Reflexión

La capa de dióxido vista hasta este momento se utiliza solo como capa "de apoyo" para conseguir dopar el material semiconductor. En este subapartado nos referimos a otras capas de material que sí que nos interesa que permanezcan en el proceso final, como pueden ser las pistas metálicas para crear conexiones.

Estas capas están formadas por diferentes materiales: óxidos, polisilicio, metales, silicio amorfo y silicio cristalino. Los conductores se usarán para realizar varios elementos como interconexiones y contactos. El polisilicio se usa como aislante en la puerta de los transistores MOS⁷, como veremos más adelante.

Figura 13. Pistas metálicas formadas sobre el óxido de la figura 11



Describiremos dos métodos orientados a la formación de capas delgadas de materiales sobre la oblea: la **deposición** y la **epitaxia**.

2.6.1. Deposición

El método más utilizado para efectuar deposiciones es la **deposición química de vapor (CVD)**⁸. Consiste en formar una película sólida sobre un sustrato mediante la reacción de reactivos químicos en estado de vapor, los cuales contienen el material constituyente. La reacción química de los gases puede producirse en la superficie de la oblea, o muy cerca de esta, con lo que estaríamos hablando de una **reacción heterogénea**; o alejada de la superficie, y tendríamos una **reacción homogénea**.

La reacción heterogénea es más aconsejable, puesto que crea películas de buena calidad. Por el contrario, las reacciones homogéneas son desaconsejables, ya que crean películas de poca adherencia con baja densidad y defectos.

2.6.2. Epitaxia

La epitaxia es uno de los procesos más importantes en la fabricación de un CI, y es la base sobre la cual se construyen muchos dispositivos. La epitaxia es el crecimiento de un cristal sobre un sustrato de estructura cristalina. Para ello, se calienta el nuevo cristal hasta casi su punto de fusión, y se pone en contacto con el material de base para que, una vez se enfríe, recristalice con la estructura adecuada, que queremos que sea la del material de base.

Se pueden distinguir dos tipos de epitaxia:

⁽⁷⁾MOS es la sigla de la expresión inglesa *metal oxide semiconductor*.

Ved también

Los transistores MOS se estudian en el subapartado 3.2 de este módulo didáctico.

⁽⁸⁾CVD es la sigla de la expresión inglesa *chemical vapour deposition*.

- La **homoepitaxia**, cuando el sustrato y la capa que crece sobre este son del mismo material.
- La **heteroepitaxia**, cuando el sustrato y la capa son de diferente material.

En principio, es mucho más sencillo obtener capas epitaxiales de buena calidad si el sustrato y la capa son iguales (homoepitaxia), puesto que en este caso no habrá diferencias en la forma de la red cristalina, como sucede en los procesos de heteroepitaxia.

Por ejemplo, en el caso del crecimiento de silicio sobre sustrato de silicio, el proceso de epitaxia aporta grandes ventajas respecto a otros procesos de fabricación. El proceso de epitaxia es un método mucho más sencillo para controlar el grosor, la concentración y el perfil de dopaje de la capa. Se podría crear, por ejemplo, una capa de silicio poco dopado sobre un sustrato altamente dopado.

2.6.3. Capas metálicas

Al hablar de la fabricación de dispositivos microelectrónicos, es necesario destacar que en todos los casos hay que dotar de contactos metálicos a los semiconductores, por ejemplo, para disponer de pistas conductoras. Además, las metalizaciones juegan un papel muy importante en la reducción de las dimensiones de los CI, puesto que ocupan más de la mitad del espacio disponible para el circuito. De hecho, la mayor parte de los fallos en los CI son debidos a defectos que surgen en las metalizaciones.

En todos los casos, la metalización se lleva a cabo depositando una capa fina de metal sobre las regiones del sustrato donde se pretenden realizar los contactos o pistas conductoras. Se trata pues de un caso particular de deposición. Hay varios métodos para la formación de películas delgadas conductoras, y uno de los más utilizados es el que ya hemos comentado: la deposición química en fase vapor (CVD).

Esto nos permitiría formar una capa de pistas conductoras metálicas, y por lo tanto, crear interconexiones allá donde se necesite. Si en un caso concreto pretendiéramos tener varios planos metálicos de conexión, lo podríamos hacer con la creación de capas de óxido (que actúen de aislante) entre los diversos planos metálicos. Solo habría que ir repitiendo los procesos vistos hasta este momento, en la secuencia correcta, y con las máscaras apropiadas en cada caso.

Hasta aquí hemos visto cómo dopar el semiconductor de base para ir creando encima de este diferentes capas de materiales, según nos interese, dada la fabricación de circuitos integrados que pretendemos hacer. Cada uno de estos procesos nos permitirá hacer crecer la capa adecuada, y se tendrán que aplicar en el orden correcto, teniendo en cuenta qué capas se van generando en

Ved también

La deposición química en fase vapor se estudia en el subapartado 2.6.1 de este módulo didáctico.

los pasos anteriores. Para ser más específicos, deberemos abordar tecnologías concretas como son la bipolar o la tecnología CMOS, que vamos a ver en el subapartado siguiente.

3. Introducción a la estructura física de circuitos integrados bipolares y CMOS

En el apartado anterior se han descrito los procesos y conceptos principales para la fabricación de los CI. En concreto, se ha visto cómo se utiliza la tecnología planar, en el sentido de que todos los circuitos se fabrican sobre un plano que marca la oblea, una pieza de semiconductor de entre 10 cm y 30 cm de diámetro y entre 400 μm y 600 μm de grosor. Hemos sido capaces de describir los procesos involucrados para conseguir depositar diferentes tipos de materiales, y que se dispongan en la forma y la orientación que necesitamos. Lo conseguimos mediante máscaras, la fotolitografía, y diferentes tipos de procesos de dopaje y depósito de capas de material, incluyendo las pistas conductoras.

En este apartado describiremos dos tipos de tecnología de CI, según los tipos de transistor que se fabriquen con la creación de la oblea.

Por un lado, tenemos la tecnología **bipolar**, que se orienta a fabricar transistores con dos uniones de semiconductores P y N; en concreto, pueden ser NPN o PNP. El **transistor de unión bipolar (BJT)**⁹⁾ es un dispositivo que consiste en dos uniones PN muy cercanas entre sí y que permite controlar el paso de la corriente eléctrica por sus terminales. La denominación *bipolar* se debe al hecho de que la conducción tiene lugar gracias al desplazamiento de portadores de dos polaridades (huecos positivos y electrones negativos), y son de utilidad en gran número de aplicaciones.

Por otro lado tenemos la tecnología **CMOS**, donde se pretende conseguir que la pieza fundamental del CI sea un transistor MOS (*metal oxide semiconductor*). Su funcionamiento no se basa en uniones PN, como el BJT, sino que el movimiento de carga se produce exclusivamente por la existencia de campos eléctricos en el interior del dispositivo; los transistores de este tipo se conocen como transistores de efecto de campo.

En los subapartados siguientes nos centraremos en qué características especiales presentan las dos tecnologías en cuanto a su disposición física, siempre teniendo en cuenta que se utilizan los procesos de producción generales descritos en el apartado anterior. No se trata de analizar el principio de funcionamiento de cada tipo de transistor, sino de ver qué implica su fabricación desde el punto de vista tecnológico o de disposición física. Más adelante, estudiaremos con detalle el funcionamiento eléctrico del transistor MOS.

⁹⁾BJT es la sigla de la expresión inglesa *bipolar junction transistor*.

Ved también

Véase la unión PN en el subapartado 1.3 de este módulo didáctico.

Los transistores BJT

Los transistores BJT se utilizan generalmente en electrónica analógica, aunque también en algunas aplicaciones de electrónica digital, como la tecnología TTL o BiCMOS.

Ved también

El funcionamiento eléctrico del transistor MOS se estudia con detalle en el módulo "Circuitos CMOS" de esta asignatura.

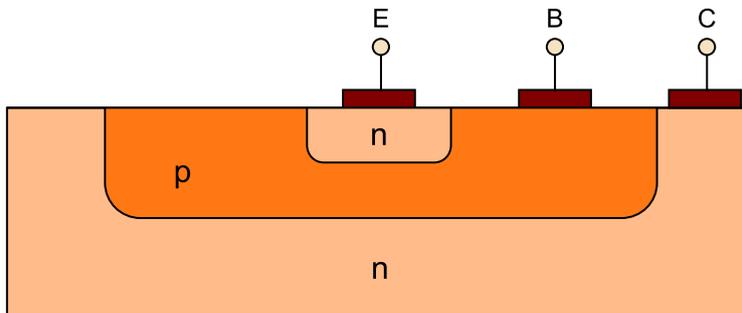
3.1. Tecnología bipolar

Un transistor de unión bipolar está formado por dos uniones PN en un mismo cristal semiconductor separadas por una región muy estrecha, como vemos en la figura 14. De este modo, quedan formadas tres zonas:

- **Emisor (E):** Se diferencia de las otras dos por estar fuertemente dopada, de forma que se comporta como un metal. Su nombre se debe al hecho de que este terminal funciona como emisor de portadores de carga.
- **Base (B):** Zona intermedia estrecha que separa el emisor del colector.
- **Colector (C):** Extensión mucho más grande.

La técnica de fabricación más común es la deposición epitaxial.

Figura 14. Estructura física de un transistor bipolar



Estas regiones son, respectivamente, de tipo P, de tipo N y de tipo P en un PNP; o de tipo N, de tipo P, y de tipo N en un transistor NPN. Cada región del semiconductor está conectada a un terminal metálico.

La base está físicamente localizada entre el emisor y el colector y está constituida por un material semiconductor ligeramente dopado y de alta resistividad. El colector rodea la región del emisor, y hace casi imposible que los electrones inyectados en la región de la base escapen de ser recolectados, lo cual maximiza las propiedades de amplificación de corriente del transistor.

El transistor de unión bipolar, a diferencia de otros transistores, no suele ser un dispositivo simétrico. Esto significa que intercambiando el colector y el emisor, hace que el transistor deje de funcionar de manera activa y empiece a funcionar de manera inversa. La falta de simetría se debe principalmente a las tasas de dopaje entre el emisor y el colector. El emisor está altamente dopado, mientras que el colector está ligeramente dopado.

Ved también

La técnica de la deposición epitaxial se estudia en el subapartado 2.6 de este módulo didáctico.

La razón por la cual el emisor está altamente dopado es para aumentar la eficiencia de inyección de portadores del emisor: la tasa de portadores inyectados por el emisor en relación con aquellos inyectados por la base. Para una gran ganancia de corriente, la mayoría de los portadores inyectados en la unión base-emisor tienen que provenir del emisor.

Gracias a esta construcción, pequeños cambios en la tensión aplicada entre los terminales base-emisor provoca que la corriente que circula entre el emisor y el colector cambie significativamente. Este efecto puede ser utilizado para amplificar la tensión o corriente de entrada. Los BJT pueden ser pensados como fuentes de corriente controladas por tensión, pero son caracterizados más simplemente como fuentes de corriente controladas por corriente, o por amplificadores de corriente, debido a la baja impedancia de la base.

Los primeros transistores fueron fabricados de germanio, pero la mayoría de los BJT modernos están compuestos de silicio. Actualmente, una pequeña parte de estos (los transistores bipolares de heterojuntura) están hechos de arseniuro de galio, especialmente utilizados en aplicaciones de alta velocidad.

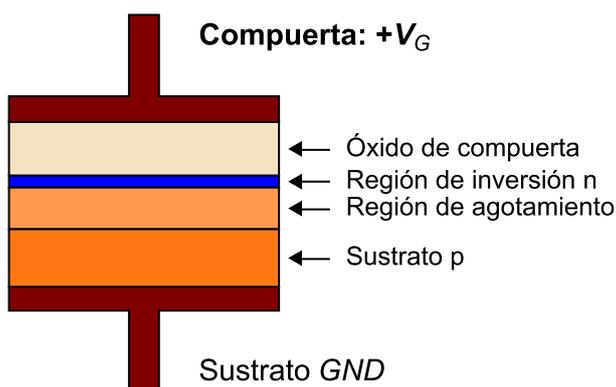
3.2. Tecnología CMOS

La tecnología CMOS¹⁰ se compone de una unidad fundamental que se conoce como estructura MOS.

⁽¹⁰⁾CMOS es la sigla de la expresión inglesa *complementary metal oxide semiconductor*.

Una **estructura MOS** es un dispositivo electrónico formado por un sustrato de silicio dopado sobre el cual se hace crecer una capa de óxido (SiO_2). En la figura 15 podemos ver su disposición física.

Figura 15. Estructura MOS con un sustrato de tipo P



Los semiconductores hacen contacto con dos terminales metálicos denominados *sustrato* y *compuerta*. La estructura es comparable a un condensador de placas paralelas, en las que una de las placas se sustituye por el silicio semiconductor del sustrato, y la otra por el óxido de la compuerta. Sin embargo, a la práctica, como óxido de compuerta se utiliza un policristal de silicio.

De hecho, en la figura 15 podemos ver que el sustrato es de tipo P, y por lo tanto nos encontramos con una **estructura PMOS**. La estructura PMOS está formada por un sustrato de silicio dopado con huecos. Al aplicar un potencial de compuerta positivo, los electrones presentes en el sustrato (portadores minoritarios) son atraídos hacia la capa de óxido de compuerta. Al mismo tiempo, los huecos son repelidos de la capa de óxido de compuerta a causa de que el potencial positivo los aleja. Esto ocasiona una acumulación de electrones en la proximidad del óxido en la que el silicio presenta un exceso de electrones, y por lo tanto, se convierte en tipo N (zona azul). La inversión del dopaje en el silicio (que antes era de tipo P) es la que da origen al nombre de esta región. También se produce una región de agotamiento de portadores en las cercanías del óxido porque los huecos del sustrato se recombinan con los electrones atraídos.

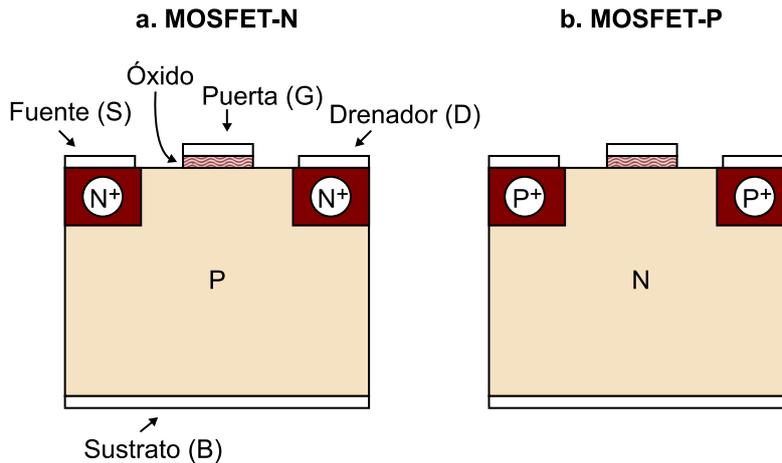
De manera análoga, una estructura NMOS está formada por un sustrato de silicio dopado con electrones. Al aplicar un potencial de compuerta negativo, los huecos presentes en el sustrato (portadores minoritarios) son atraídos hacia la capa de óxido de compuerta. Los electrones son repelidos por el óxido de compuerta porque el potencial negativo los aleja. Los huecos se acumulan en la proximidad del óxido en que el silicio acumula un exceso de huecos y por lo tanto se comporta como un material de tipo P. La recombinación de huecos y electrones produce una región de agotamiento.

Estas estructuras MOS también reciben el nombre de condensadores MOS, por su principio de funcionamiento. Son la base de los transistores denominados MOSFET¹¹, o transistores de efecto de campo. En la figura 16 podemos ver la estructura física, que, como podemos comprobar, se basa precisamente en condensadores MOS. Vemos que se trata de la estructura MOS a la cual añadimos dos zonas dopadas a cada lado del condensador MOS para formar la fuente (S) y el drenador (D).

Comparado con el condensador MOS, el MOSFET incluye dos terminales adicionales (fuente y drenador), cada uno conectado a regiones altamente dopadas que están separadas por la región del sustrato. Estas regiones pueden ser de tipo P o N, pero ambas tienen que ser del mismo tipo, y del tipo opuesto al del sustrato. La fuente y el drenador (de forma diferente al sustrato) están fuertemente dopadas y en la notación se indica con un signo ⁺ después del tipo de dopaje.

⁽¹¹⁾MOSFET es la sigla de la expresión inglesa *metal oxide semiconductor field effect transistor*.

Figura 16. Estructura física de los transistores MOS, N y P



El dopaje del sustrato es opuesto al tipo de portador que origina la corriente. Así, para un transistor de tipo N (electrones en conducción), el dopaje del sustrato es de tipo P, y a la inversa.

Cuando se aplica una tensión positiva al terminal de puerta de un MOSFET de tipo N, se crea un campo eléctrico bajo la capa de óxido que incide perpendicularmente sobre la superficie del semiconductor. Este campo atrae a los electrones hacia la superficie bajo el óxido, repeliendo los huecos hacia el sustrato. Si el campo eléctrico es muy intenso, se consigue crear en esta superficie una región muy rica en electrones, denominada canal N, que es precisamente el mecanismo que hemos visto para una estructura MOS. Se trata de la zona de inversión (zona azul en la figura 15). Este canal permite el paso de corriente de la fuente al drenador; cuanto mayor sea la tensión de puerta, mayor será el campo eléctrico y, por lo tanto, la carga en el canal. Una vez creado el canal, la corriente se origina aplicando en el drenador una tensión positiva con respecto a la de la fuente.

En un MOSFET de tipo P el funcionamiento es a la inversa, puesto que los portadores son huecos (cargas positivas cuyo valor es el módulo de la carga del electrón). En este caso, para que haya conducción, el campo eléctrico perpendicular a la superficie ha de tener sentido opuesto al del MOSFET de tipo N, por lo que la tensión aplicada debe ser negativa. Ahora, los huecos son atraídos hacia la superficie bajo el óxido, y los electrones repelidos hacia el sustrato. Si la superficie es muy rica en huecos, se forma el canal P. Cuanto más negativa sea la tensión de puerta, más intensa puede ser la corriente (más huecos en el canal P) que se establece al aplicar en el terminal del drenador una tensión negativa con respecto a la de la fuente. La corriente tiene sentido opuesto a la de un MOSFET de tipo N.

Las regiones de dopaje elevado del mismo tipo que los portadores del canal (regiones N^+ y P^+) se fabrican con objeto de facilitar la corriente a través de los terminales fuente y drenador. El transistor MOS es simétrico: los terminales

fuente y drenador son intercambiables entre sí. En el MOSFET de tipo N el terminal de tensión más alta actúa de drenador (recoge los electrones), mientras que el de menor tensión se encuentra en el de tipo P (recoge los huecos).

Ya hemos estudiado qué es una estructura MOS y cómo se pueden crear transistores (nMOS y pMOS) a partir de esta estructura. En este punto ya podemos definir qué es un CMOS: *complementary metal-oxide-semiconductor* es una de las familias lógicas empleadas en la fabricación de circuitos integrados. Su principal característica consiste en la utilización conjunta de transistores de tipo pMOS y tipo nMOS, configurados de tal forma que, en estado de reposo, el consumo de energía es únicamente el causado por las corrientes parásitas.

Así pues, en las obleas CMOS encontraremos que se fabrican tanto transistores MOS de tipo N como MOS de tipo P, de manera que se puedan combinar en una escala superior (puerta lógica) y proporcionar una serie de ventajas.

En la actualidad, la mayoría de los circuitos integrados que se fabrican utilizan la tecnología CMOS. Esto incluye microprocesadores, memorias, procesadores digitales de señales y otros muchos tipos de circuitos integrados digitales, cuyo consumo es considerablemente bajo.

Ved también

Las ventajas de los transistores CMOS se estudian con detalle en el módulo "Circuitos CMOS", cuando estemos en disposición de analizar el funcionamiento de los transistores y de su configuración de puerta lógica CMOS.

Resumen

En este módulo hemos aprendido que los **materiales semiconductores** son la base de la electrónica actual. En ellos, la corriente circula por medio de electrones libres y huecos. Con un dopaje adecuado, podemos modificar las características eléctricas de estos materiales para regular su conductividad.

A continuación hemos hablado de cómo somos capaces de crear estos circuitos físicamente, de qué **procesos de fabricación** se encuentran involucrados en su producción. Hemos aprendido que a partir de una oblea de silicio, hacemos uso de una serie de procesos físicos y químicos para crear semiconductores dispuestos de tal forma que tengamos transistores y otros tipos de dispositivos que son la base de los circuitos integrados. Veíamos que todo se basa en el hecho de poder depositar controladamente capas de materiales sobre el sustrato de silicio de la oblea.

Finalmente, hemos profundizado algo más en la disposición física que presentan los transistores de dos tecnologías ampliamente utilizadas, como son: la tecnología **bipolar** y la **CMOS**.

Glosario

aislante *m* Material que no conduce corriente eléctrica.

banda *f* Intervalo de energía donde se disponen los electrones en los materiales de estado sólido.

banda de conducción *f* Banda con energía superior a la de valencia, en que los electrones se mueven libremente.

banda de valencia *f* La última banda que contiene electrones en un material sólido sin la aplicación de ningún campo eléctrico.

banda prohibida *f* Banda de energía donde los electrones del cristal no pueden estar situados.

CI Sigla de circuito integrado.

CMOS Sigla de *complementary metal oxide semiconductor* (semiconductor óxido-metal complementario); es una de las tecnologías empleadas en la fabricación de circuitos integrados.

conductor eléctrico *m* Material que conduce corriente eléctrica.

dopaje *m* Proceso que consiste en inyectar impurezas en un semiconductor para modificar sus características eléctricas.

electrón libre *m* Electrón en banda de conducción, que se puede mover libremente.

hueco *m* Posición libre que deja un electrón cuando abandona la banda de valencia. Eléctricamente, se considera una partícula con la misma carga que un electrón, pero de signo positivo.

impurezas *f* Átomos que se añaden al semiconductor en el dopaje, para añadirle huecos o electrones libres.

semiconductor *m* Material en el que podemos regular la corriente que circula por él.

unión PN *f* Unión de dos materiales semiconductores, uno de tipo P y otro de tipo N.

Bibliografía

Baker, R. J. (2010). "CMOS circuito design, layout, and simulation" (3.^a ed.). *IEEE press series*. John Wiley & Sons, Inc.

Ghandhy, S. K. (1994). *VLSI fabrication principles* (2.^a ed.). John Wiley & Sons, Inc.

Streetman, B.; Banerjee, S. (2005). *Solid state electronics devices* (6.^a ed.). Prentice Hall.

Sze, S. M.; Chang, C. Y. (1996). *ULSI technology*. McGraw Hill.