

Apéndice A
Semiconductores, junturas
y dispositivos sólidos activos

A.1. Semiconductores

Introducción

Para poder explicar la física de los semiconductores (germanio, silicio o materiales compuestos, como el arseniuro de galio, fosfuro de galio, carburo de silicio, etc.) podemos utilizar dos modelos diferentes:

El primer modelo es de análisis cualitativo y se lo conoce como *modelo de enlaces*. En la estructura diamante del germanio y del silicio, cada átomo forma cuatro enlaces covalentes con los cuatro átomos que lo rodean. Para romper una unión y que surja un electrón libre, debemos comunicar energía al sólido, lo que implica subir la temperatura. A $T > 0^\circ\text{K}$, al romperse una unión, surge un electrón libre animado de energía de agitación térmica que lo hace moverse al azar. El lugar que deja el electrón se lo denomina laguna o hueco. Este hecho provoca una no neutralidad eléctrica, por lo que un electrón ligado de un enlace cercano puede introducirse en el lugar vacante, dejando a su vez un nuevo hueco. Es decir, todo sucede como si una partícula de carga positiva se hubiese desplazado en sentido contrario al movimiento del electrón. Podemos decir entonces que la conducción en un semiconductor se debe tanto a una corriente de electrones libres como de lagunas.

El segundo modelo, es de análisis cuantitativo y se lo conoce como *modelo de bandas*. Si tengo dos átomos separados con un mismo nivel de energía permitido, puedo tener dos electrones en cada átomo, debido al spin. Cuando se van acercando, por el principio de exclusión de Pauli, dicho nivel deberá desdoblarse en dos subniveles. En el caso de tener N átomos con un mismo nivel de energía permitido, dicho nivel se desdoblará en N niveles, formando una banda de energía cuasi continua (dada la gran cantidad de átomos que interactúan), donde se ubicarán los $2N$ electrones ⁽¹⁾.

La densidad de corriente J , será $J = q \sum v_i$, donde se sumarán todos los electrones de la banda por unidad de volumen. Tanto en una banda vacía, como en una llena, $J = 0$. En una banda casi vacía, la corriente estará dada por los electrones que pasaron a los niveles permitidos superiores por ser más energéticos. Mientras que en una banda casi llena, la corriente puede ser expresada por la producida por los estados vacantes, tratados como partículas de carga positiva (huecos).

En los semiconductores, entre la última banda llena (cuyo nivel superior se denomina energía de valencia, E_v) y la primera vacía (cuyo nivel inferior se lo denomina energía de conducción, E_c), habrá una zona de energía prohibida o "*gap*", $E_g = E_c - E_v$. Los semiconductores poseen valores de E_g entre $0,5 \text{ eV}$ y 5 eV . Valores superiores de E_g corresponden a materiales aisladores.

(1) Dado un sólido de distancia interatómica " a ", su diagrama de bandas es equivalente a hacer un corte en el diagrama de distribución de niveles de energía a una distancia " a ".

A.1.1. Semiconductores en equilibrio

Un semiconductor se encuentra en *equilibrio térmico*, cuando se encuentra a temperatura constante en el tiempo y uniforme en el espacio y ningún agente físico o químico externo actúa sobre él, salvo el intercambio térmico con el medio ambiente. En ese caso, si cada proceso que ocurre en el material se encuentra contrarrestado por otro proceso opuesto, se dice que se encuentra en *equilibrio termodinámico*. En este caso, la generación térmica de pares electrón –hueco por unidad de tiempo y de volumen (excitación de un electrón desde la banda de valencia a la de conducción) resulta igual a la recombinación térmica de pares por unidad de tiempo y de volumen (desexcitación de un electrón desde la banda de conducción a la de valencia).

La generación dependerá de la temperatura, mientras que la recombinación dependerá también de la concentración de electrones " n " en la banda de conducción y de la concentración " p " de huecos en la banda de valencia, dado que ambos deben interactuar para que la recombinación ocurra:

$$g = f_1(T) = r = n.p.f_2(T) \Rightarrow n.p = f_3(T)$$

Es decir, el producto de las concentraciones de electrones y huecos en un dado semiconductor en equilibrio térmico es sólo función de la temperatura.

Si aplicamos esta afirmación para un semiconductor puro o *intrínseco* ⁽¹⁾, donde $n = p = n_i$, entonces $n_i^2 = f_3(T)$, y como la concentración intrínseca es constante para un semiconductor a una dada temperatura, es posible reemplazar n_i para obtener $n.p = n_i^2$ (ecuación válida para semiconductores intrínsecos o extrínsecos en equilibrio termodinámico).

A.1.2. Semiconductores extrínsecos

La forma más usada de controlar el número de portadores en un semiconductor, es incorporando impurezas que ocupan en la red cristalina el mismo lugar del átomo substituído.

Supongamos que agregamos impurezas correspondientes al grupo V de la tabla periódica (5 electrones de valencia). El quinto electrón no estará unido por unión covalente como los otros cuatro con átomos vecinos del semiconductor. Sólo se encuentra unido al átomo de impurezas por un exceso de carga positiva del núcleo. Con una pequeña cantidad de energía ($T > 0^\circ\text{K}$) este quinto electrón puede romper su enlace electrostático y convertirse en electrón libre, dejando un ión positivo. Este tipo de impurezas se denominan *donoras*.

(1) Se considera intrínseco cuando tiene hasta 1 átomo de impureza por cada 10^{10} átomos de semiconductor. Cada electrón excitado a la banda de conducción deja un hueco en la banda de valencia. Es decir, los electrones y huecos se crean de a pares.

Si agregamos impurezas correspondientes al grupo III (3 electrones de valencia), los tres electrones de valencia se unirán por enlaces covalentes con tres átomos vecinos de semiconductor, dejando un enlace vacante. Con una pequeña cantidad de energía ($T > 0^\circ\text{K}$), un electrón de un enlace covalente vecino puede ocupar dicho enlace vacante y crear una laguna, dejando negativo al ión de impureza. Este tipo de impurezas se denominan *aceptoras*.

Un semiconductor con impurezas donoras o aceptoras, se lo denomina *extrínseco*.

Cuando el número de átomos donores por unidad de volumen, N_D es mayor que el de aceptores, N_A , se dice que el semiconductor es de *tipo N*. En el caso opuesto, se dice que es de *tipo P*. Ambas cantidades deberán ser mucho menores que el número total de átomos del semiconductor (10^{22} átomos/cm³), en caso contrario se dice que el material se encuentra degenerado y pierde sus propiedades de monocristal para transformarse en cristal amorfo.

Entonces, existen dos formas para obtener portadores libres en un semiconductor: por generación térmica en átomos propios o por ionización de átomos de impurezas.

Supongamos un semiconductor tipo N, fuertemente contaminado o *dopado mayoritariamente* con $N_D \gg n_{\text{generación intrínseca}}$ y $N_A = 0$. Podemos decir que por encima de los 200°K , ya todos los átomos de impureza están ionizados (los componentes semiconductores trabajarán normalmente entre 200 y 450°K). Entonces, la cantidad de electrones libres estará dado por:

$$n = n_{\text{átomos de impurezas}} + n_{\text{generación intrínseca}} \cong N_D$$

Mientras que la cantidad de huecos corresponderá sólo a la generada térmicamente (pues la generación intrínseca produce de a pares electrón - hueco):

$$p = p_{\text{generación intrínseca}} = n_{\text{generación intrínseca}} \leq n_i$$

Podemos concluir que, bajo estas condiciones típicas, la concentración de portadores mayoritarios es prácticamente constante, mientras que la de minoritarios depende fuertemente de la temperatura.

A.1.3. Conductividad

En un cristal, a temperaturas mayores que 0°K , los electrones y los huecos son partículas casi libres en el sentido que no están asociados a ningún sitio en particular de la red y se mueven aleatoriamente colisionando con la red (con un tiempo libre medio entre colisiones) a la cual le entregan energía en cada proceso de choque.

Cuando aplicamos un campo eléctrico E sobre el cristal, en promedio los electrones se deslizarán en sentido contrario al campo con una velocidad de corrimiento v_c .

La *densidad de corriente* que fluye en la dirección del campo será:

$$J_n = n \cdot q \cdot \mu_n \cdot E.$$

La *movilidad* μ_n describe la facilidad con la que un electrón se mueve bajo la acción de un campo aplicado.

Haciendo un análisis similar con los huecos, se obtiene que la densidad de corriente total será la suma de la de huecos y electrones:

$$J = q \cdot (n \cdot \mu_n + p \cdot \mu_p) \cdot E = \sigma \cdot E$$

La *conductividad* σ representa la facilidad con que el campo genera una densidad de corriente en un material.

En un semiconductor la movilidad de los electrones es mayor que la de los huecos (3 veces más en Si), ya que resulta más fácil comunicarle movimiento a un electrón libre que a uno ligado, cuyo desplazamiento provocara el desplazamiento de la laguna. En Si , $\mu_n = 1350 \text{ cm}^2/\text{Vseg}$ y en $GaAs$, $\mu_n = 8500 \text{ cm}^2/\text{Vseg}$.

La movilidad disminuye al aumentar la temperatura, ya que el aumento de la vibración de los átomos de la red cristalina, aumenta la superficie efectiva de choque, disminuyendo el tiempo entre colisiones. También disminuye al aumentar la concentración, ya que los átomos de impurezas agregados al cristal producen distorsiones locales en la red, lo cual produce un aumento de la dispersión de los portadores libres.

Normalmente, la movilidad de electrones libres en un semiconductor es algo mayor que la existente en un metal (unas 10 veces), ya que las distancias interatómicas son menores que en el semiconductor, aumentando la probabilidad de choque (menor tiempo libre medio entre colisiones). Por otro lado, como el número de portadores libres en un metal es muy superior a los de un semiconductor (unas 1000 veces), la conductividad en un metal será siempre mucho mayor que en un semiconductor.

Para analizar la variación de la conductividad con la temperatura, podemos hacer una comparación entre el metal y el semiconductor. En el conductor, existen un solo tipo de portadores: electrones. Por lo tanto, $\sigma_M = q \cdot \mu_n \cdot n$. La densidad de electrones libres permanecerá constante con la temperatura, por lo que la conductividad disminuirá como lo hace la movilidad.

Por otro lado, en un semiconductor extrínseco:

$$\sigma_S = q \cdot (n \cdot \mu_n + p \cdot \mu_p)$$

Obteniéndose, para un Si tipo N, la siguiente gráfica de variación de la concentración con la temperatura:

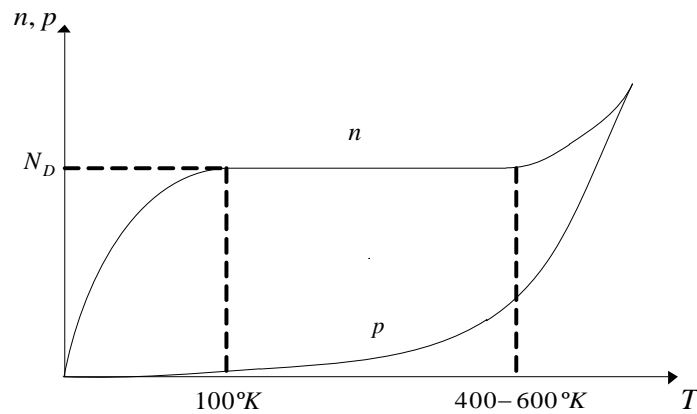


Fig. A.1

Con lo que se puede inferir que la conductividad disminuirá con la temperatura igual que la movilidad mientras la concentración permanezca constante, aumentando luego a partir de 400°K a 600°K , cuando la generación térmica deje de ser despreciable.

A.1.4. Ecuación de continuidad

Mediante el planteo de esta ecuación se obtendrán todas las expresiones relacionadas con la corriente en semiconductores. Es una ecuación general que plantea el principio de conservación de la carga eléctrica y se escribe en base a las siguientes consideraciones:

1) En un semiconductor extrínseco siempre se planteará para los portadores minoritarios, pues éstos son los que ponen de manifiesto los fenómenos de generación y recombinación.

2) Consideramos válida la hipótesis de cuasi neutralidad por la cual, si existe una densidad de carga neta en el material, es despreciable frente a la concentración de portadores mayoritarios y el material puede considerarse prácticamente neutro. Es decir que podemos admitir que los excesos de portadores mayoritarios y minoritarios son iguales. Sin embargo, el exceso de minoritarios en relación con la concentración en equilibrio es mucho mayor que el exceso de mayoritarios en relación con su concentración en equilibrio. Es decir, la concentración de mayoritarios no se ve afectada: *condición de bajo nivel de inyección*.

A.2. Juntura P-N: Diodos

A.2.1. Juntura P-N en equilibrio termodinámico

Existen dos técnicas principales para la construcción de junturas P-N:

Epitaxial: se parte de un semiconductor base P o N y se construye el otro material del tipo opuesto por una deposición de átomos. De esta forma se logra que la concentración de ambos materiales sea uniforme y que la juntura obtenida sea abrupta (es decir, la contaminación sufre

un salto de muy alta pendiente de la concentración donora a la aceptora).

Compensación: consiste en difundir un tipo de impureza (donora o aceptora) sobre un material de tipo opuesto. En este caso no se obtiene una juntura abrupta, sino gradual.

Para simplificar el análisis, tomaremos el caso de juntura abrupta. Supongamos que tenemos dos materiales de contaminación uniforme, tipos N y P. Admitamos que el material P está más contaminado que el N, es decir $N_A > N_D$, (se suele escribir como materiales de tipo P^+ y N).

Al unir estos dos materiales, por la diferencia de concentraciones existente en la zona de unión, los electrones del lado N se difundirán hacia el lado P^+ y los huecos del lado P^+ se difundirán hacia el N, dejando en la zona cercana al plano de unión, iones positivos y negativos, respectivamente, sin neutralizar. Dichos portadores mayoritarios, al difundirse a través de la juntura, se convertirán en minoritarios en exceso del lado opuesto y por lo tanto se recombinarán rápidamente con los mayoritarios, quienes al recombinarse dejan a su vez iones de átomos de impurezas sin neutralizar. Se forma entonces una *zona desierta* de portadores libres en los alrededores del plano de unión.

El proceso seguirá hasta que la carga de los iones fijos sin compensar a ambos lados de la unión (positiva del lado N y negativa del lado P) provoque un campo eléctrico tal que cause una tendencia al corrimiento exactamente igual y opuesta a la tendencia a la difusión.

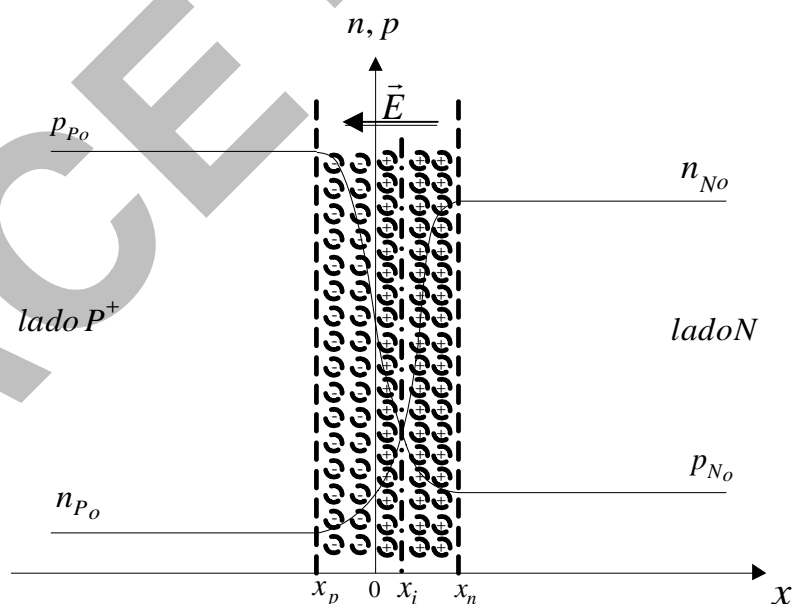


Fig. A.2

Obviamente, por el principio de conservación de la carga eléctrica, el nuevo material así unido sigue siendo eléctricamente neutro. Y como las cargas de los iones de ambos lados deben igualarse, la zona desierta será más extensa del lado del material menos dopado.

Como puede verse en la Fig. A.2, la hipótesis de la existencia de una zona de deserción total resulta válida, pues la concentración de mayoritarios en los alrededores de la unión desciende bruscamente. Fuera de esta zona desierta, los materiales se suponen perfectamente neutros.

A.2.2. Juntura P-N con polarización

Para polarizar una juntura, es decir, aplicarle una diferencia de potencial, es necesario construir contactos del tipo metal-semiconductor, tanto del lado P como del N. Dichos contactos se clasifican en dos tipos ⁽¹⁾:

Rectificantes: Permite conducción de corriente en un sentido y no en el otro, tal como una juntura P-N.

Óhmicos: La diferencia de potencial en el contacto es independiente de la corriente que la atraviesa en uno u otro sentido.

Para el caso de polarizar una juntura P-N, admitimos que los contactos metal-semiconductor que se construyen en los extremos de los materiales, serán de tipo óhmico. Obviamente, si cortocircuitamos los contactos, las caídas en las junturas óhmicas compensarán la diferencia de potencial de la unión P-N, manteniendo el sistema en equilibrio, es decir: $V_{jo} + V_A + V_B = 0$.

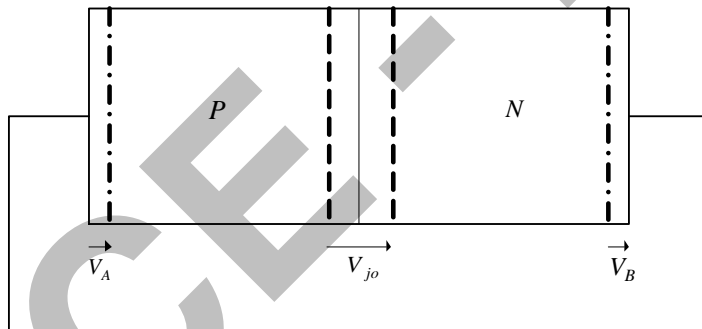


Fig. A.3

Si a esta juntura le aplicamos ahora una diferencia de potencial externo V_D , los potenciales V_A y V_B no variarán con la corriente circulante pues corresponden a contactos óhmicos, por lo que dicho potencial V_D caerá en la juntura, modificando el valor de V_{jo} (en las zonas neutras no habrá caída, precisamente por considerarlas eléctricamente neutras, es decir se admitirá conductividad infinita en dichas zonas).

Si polarizamos la juntura de modo tal que el borne positivo de la batería se conecte al contacto adyacente al lado P y el borne negativo de la batería se conecte al contacto adyacente al lado N, diremos que la juntura está polarizada en directa y el potencial externo aplicado V_D se considera positivo: $V_D > 0$.

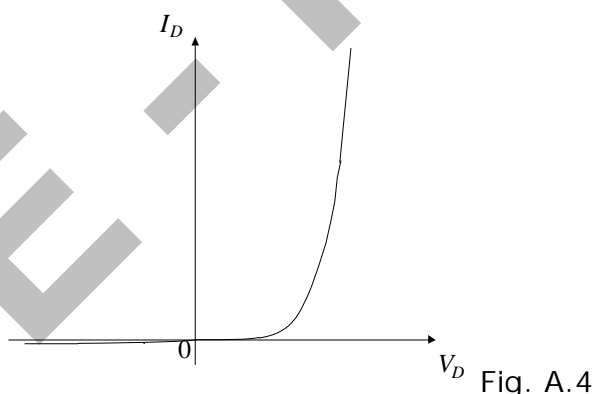
(1) Para una explicación más detallada sobre contactos M-S, referirse al apéndice B.

Así polarizada la juntura, la barrera de potencial de la unión P-N, disminuirá respecto de su valor de equilibrio, V_{jo} . Esto trae como consecuencia una reducción del campo eléctrico E (disminuye la zona desierta), permitiendo la difusión de mayoritarios desde cada zona neutra hacia el lado opuesto de la juntura (se convierten en minoritarios del lado opuesto de la zona desierta y dado que provocan un exceso de portadores, comienzan a difundirse y recombinarse con los mayoritarios de esa zona ⁽¹⁾). Es decir que en polarización directa la corriente a nivel de juntura se debe a los mayoritarios.

Si polarizamos la juntura de forma opuesta ($V_D < 0$), aumentará la barrera de potencial de la unión P-N, respecto a su valor de equilibrio, V_{jo} . Esto trae como consecuencia un aumento del campo eléctrico E (aumenta la zona desierta). El campo mayor que el necesario para contrarrestar la tendencia a la difusión en equilibrio provocará que los minoritarios cercanos a la juntura se vean arrastrados hacia el lado opuesto, convirtiéndose en mayoritarios, desplazándose hacia el respectivo contacto metálico. Es decir que con polarización inversa, la corriente a nivel de juntura es de minoritarios.

Teniendo en cuenta todo este análisis y a partir de aplicar la ecuación de continuidad, se obtiene la ecuación del diodo ideal:

$$I_D = I_S \cdot (e^{V_D / V_T} - 1)$$



donde:

I_S resulta la corriente en inversa del diodo ideal (corriente de minoritarios, es decir, la debida a la generación térmica en las zonas neutras), cuyo valor para silicio es de 1 a 10 fA ($10^{-15}A$) y para germanio de alrededor de 1 μA .

$V_T = kT/q$ (k : constante de Boltzman, T : temperatura de Kelvin y q : la carga del electrón). A temperatura ambiente ($T \cong 25^\circ C$), $V_T \cong 25mV$.

(1) El tiempo promedio en que el exceso de minoritarios se recombina, se lo denomina tiempo de vida. Si la zona neutra es tan corta que los portadores la recorren en un tiempo menor al de vida - tiempo de tránsito -, podemos decir que casi no hay recombinación en la zona neutra y los portadores minoritarios se recombinaran en los contactos metálicos.

A esta curva $I_D = f(V_D)$, se la denomina característica estática del diodo, pues se supone que las mediciones de tensión y corriente se realizan una vez pasados los transitorios. Construyendo el siguiente circuito (donde R_L se coloca como limitador de la corriente):

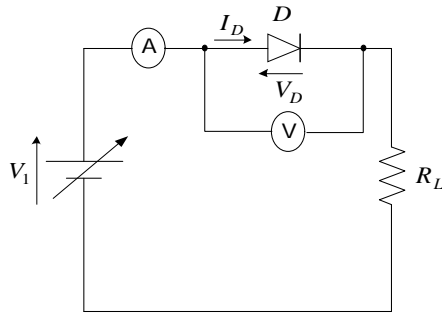


Fig. A.5

Haciendo variar el valor de la fuente V_1 y midiendo los valores de I_D y V_D , obtendremos la característica estática del diodo.

En muchas aplicaciones de los diodos de dispositivos semiconductores que involucren junturas P-N, es importante la dependencia de la temperatura de la característica del diodo. La principal dependencia con la temperatura proviene de la corriente de saturación inversa I_S (por sobre el V_T de la expresión exponencial), que siendo una corriente de minoritarios aumenta con la temperatura por generación térmica. Se observa que I_S se duplica aproximadamente cada 6°C en silicio (cada 10°C en germanio) obteniéndose, a corriente I_D constante, una disminución de la tensión en directa aproximada de $2\text{mV}/^\circ\text{C}$.

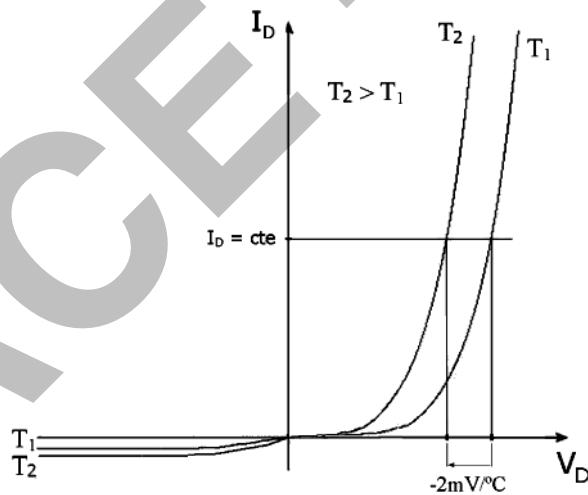


Fig. A.6

A.2.3. Características del diodo real

La forma de la característica del diodo real en directa es aproximadamente la indicada en la Fig. A.7:

Para bajas tensiones o zona de muy bajas corrientes (primera zona), la recombinación en las zonas neutras es muy pequeña por no haber aumentado mucho las concentraciones respecto de sus valores de

equilibrio. De esta manera no es despreciable la recombinación en la zona desierta frente a la recombinación en las zonas neutras. Puede aproximarse la curva a la forma: $I_D = I_S \cdot e^{VD/2VT}$

En la segunda zona, denominada de corrientes medias (para un diodo de bajo nivel de potencia, es decir menor a 1W), se puede admitir que las condiciones ideales se acercan al comportamiento del diodo real: $I_D = I_S \cdot e^{VD/VT}$

En altas corrientes (tercera zona o alto nivel de inyección), ya no puede despreciarse las caídas de tensión en zonas neutras y deberá considerarse el efecto resistivo que presentan (las zonas neutras dejan de ser neutras, es decir en el diodo ideal se consideraba conductividad infinita y campo eléctrico nulo). Por lo tanto, el potencial aplicado V_D será mayor a la caída de tensión en la zona desierta de la juntura (despreciando la caída en los contactos óhmicos), cayendo en las zonas neutras una tensión $I_D \cdot R_{serie}$ o $I_D \cdot R_{ZN}$, que será apreciable cuando la corriente por el diodo alcance valores elevados. La R_{serie} puede estimarse en algunos ohms.

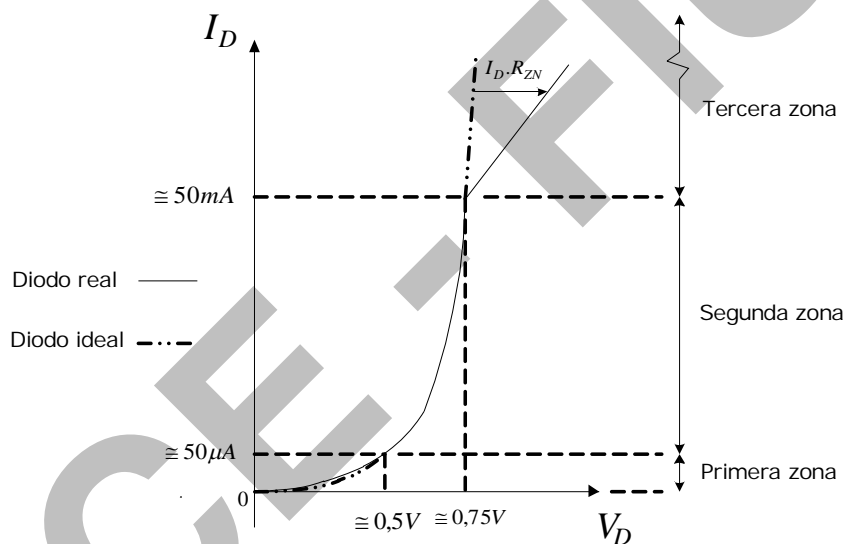


Fig. A.7

En inversa, para obtener la característica real del diodo debemos tener en cuenta la generación en las zonas desiertas que tanto en silicio como en germanio produce una corriente del orden de los nA . Por lo tanto, en el silicio la corriente debida a la generación en la zona desierta es mucho mayor que la debida a la generación en la zona neutra (que compone a I_S , que es del orden de los fA). Esto no es así en el germanio (cuya I_S es del orden de los μA), donde la curva real se acerca a la forma ideal. Por lo tanto, en silicio $|I_{Dinv}| \gg I_S$.

En las características del diodo real, se observa que llegado un valor de tensión inversa $V_D = -V_R$, la curva comienza a caer casi verticalmente (la corriente aumenta en valor absoluto con una pendiente muy elevada, muy similar a la del diodo en directa. La pendiente puede aproximarse a un valor de conductancia $1/r_z$, donde r_z es un valor dado por el fabricante y que es del orden de los ohms.

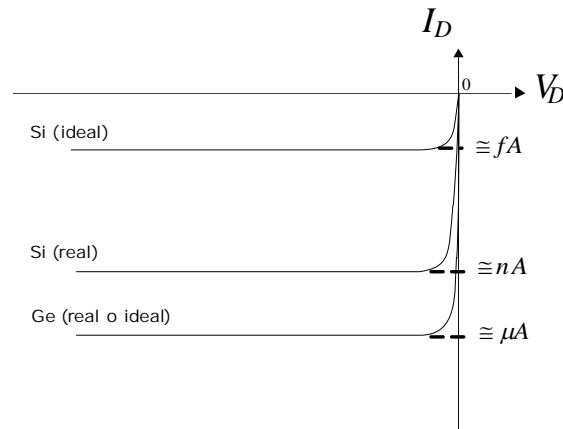


Fig. A.8

Este fenómeno se conoce como *ruptura* y puede deberse a dos efectos:

Efecto Avalancha: A medida que aumenta la tensión aplicada V_D en inversa, la longitud de la zona desierta aumenta y también el campo máximo en ella. Los portadores minoritarios que el campo inyecta dentro de la juntura tienen un camino libre medio entre choques. Si el campo eléctrico tiene suficiente intensidad como para comunicarle al electrón en ese camino libre medio una energía cinética suficiente como para que este en su próximo choque genere un par electrón-hueco, dicho par de portadores a su vez ganarán energía cinética y generarán otro par en su próximo choque, comunicándole movimiento a las dos nuevas partículas. Este efecto se denomina *multiplicación por avalancha*. Es decir, cuando se alcanza el campo necesario para producirlo, el incremento en la generación de portadores por choque es enorme y se produce una especie de reacción en cadena.

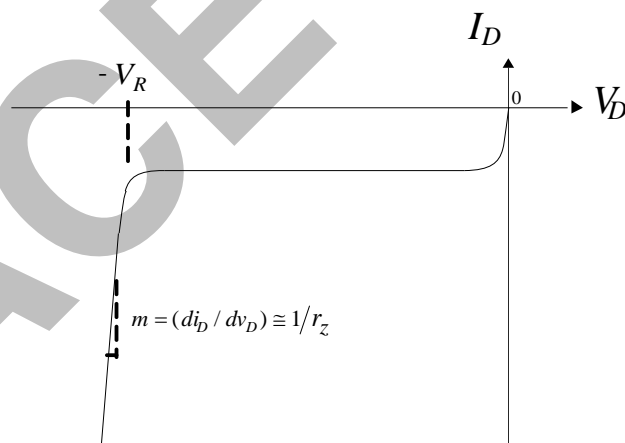


Fig. A.9

El efecto avalancha depende de la magnitud del campo eléctrico, de la longitud de la zona desierta y de la temperatura. O sea que, si aumenta la temperatura, la mayor vibración de la red cristalina (se le entrega energía a la red cristalina) provoca la reducción del camino libre medio, por lo que se necesitará un campo mayor (la tensión de ruptura aumenta, en valor absoluto) para poder entregar al electrón la misma energía cinética en un camino libre medio más corto. Por otro lado, a mayor contaminación de las zonas P y N, la longitud de la zona desierta

será menor (mayor densidad de carga) y se necesitará un campo mayor para poder generar la avalancha. Sin embargo, dado que al aumentar la contaminación, aumenta el campo empotrado (aumenta V_{jo}), la tensión externa aplicada para provocar la ruptura disminuirá.

Efecto Túnel (también llamado efecto Zener): Para elevadas contaminaciones de las zonas P y N, la zona desierta será tan angosta que habrá pasaje de electrones por efecto túnel. La tensión necesaria para producir este efecto es baja (aunque el campo eléctrico es muy elevado), y a medida que se reduce la contaminación se hace menos probable, necesitándose un valor de tensión más elevada. Alrededor de $V_R = 5V$ se encuentra la zona límite donde predomina para valores de V_R superiores la ruptura por avalancha y para valores inferiores la ruptura por túnel. El aumento de la temperatura hace que los portadores atraviesen la barrera de potencial por túnel más fácilmente, por lo que la tensión de ruptura disminuye. Esto significa, que en los alrededores de los 5V, el coeficiente de variación térmica de V_R está compensado entre ambos efectos de ruptura.

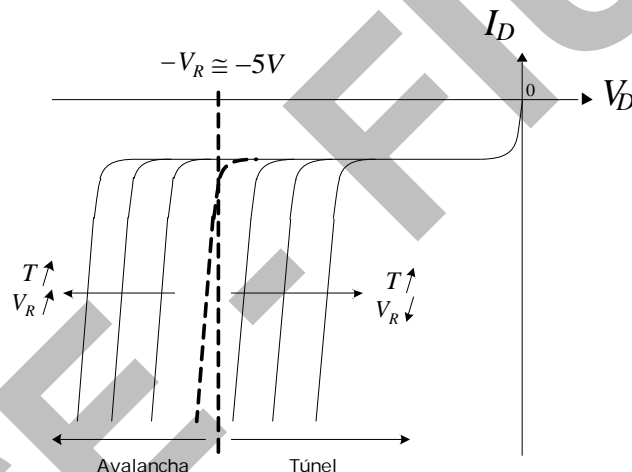


Fig. A.10

Si bien para todos los diodos se produce ruptura (y ello no implica destrucción del dispositivo, salvo que se supere el valor de potencia máxima que pueda disipar), existen algunos diodos especialmente preparados para trabajar en inversa, llamados *Zeners*. Los Zeners pueden trabajar por avalancha o túnel (en alguna bibliografía se denomina como "efecto zener" al efecto túnel).

Cuando se trabaja en inversa, mientras no se alcance la ruptura, la potencia que disipa el diodo es casi nula, pues la corriente que lo atraviesa es muy pequeña. En cambio, cuando se alcanza la ruptura, deberá disipar una potencia importante y para ello estos dispositivos deben ser preparados especialmente.

A los diodos que trabajan en la zona de ruptura se los utiliza como diodos reguladores de tensión, es decir como diodos que pueden conducir diferentes valores de corriente limitados por el circuito externo, manteniendo la tensión constante entre sus bornes (ver Fig. A.11).

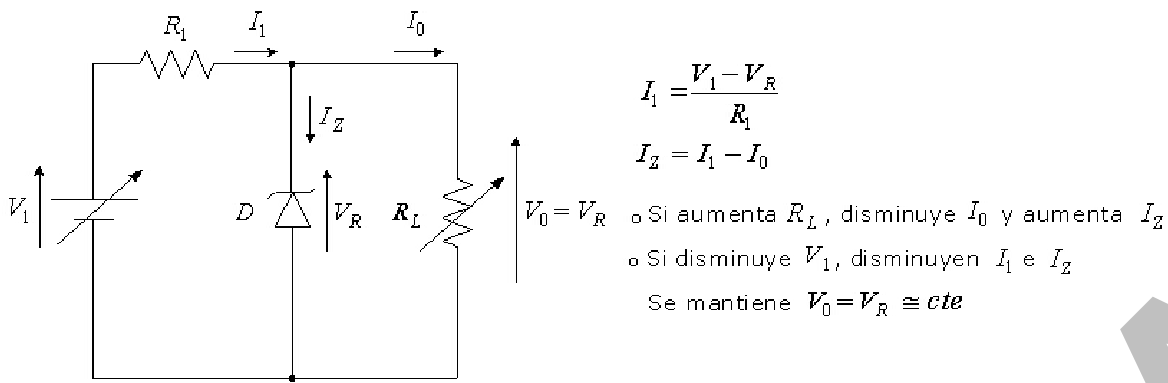


Fig. A.11

A.2.4. Modelo del diodo para pequeña señal

Resistencia dinámica del diodo

Si se producen pequeñas variaciones de la corriente (y tensión) en un diodo, alrededor de un dado punto de reposo Q , para el análisis de tales variaciones resulta adecuado aproximar la característica estática del diodo mediante la recta tangente en ese punto. Surge de este modo el concepto de *linealización* de la curva del diodo.

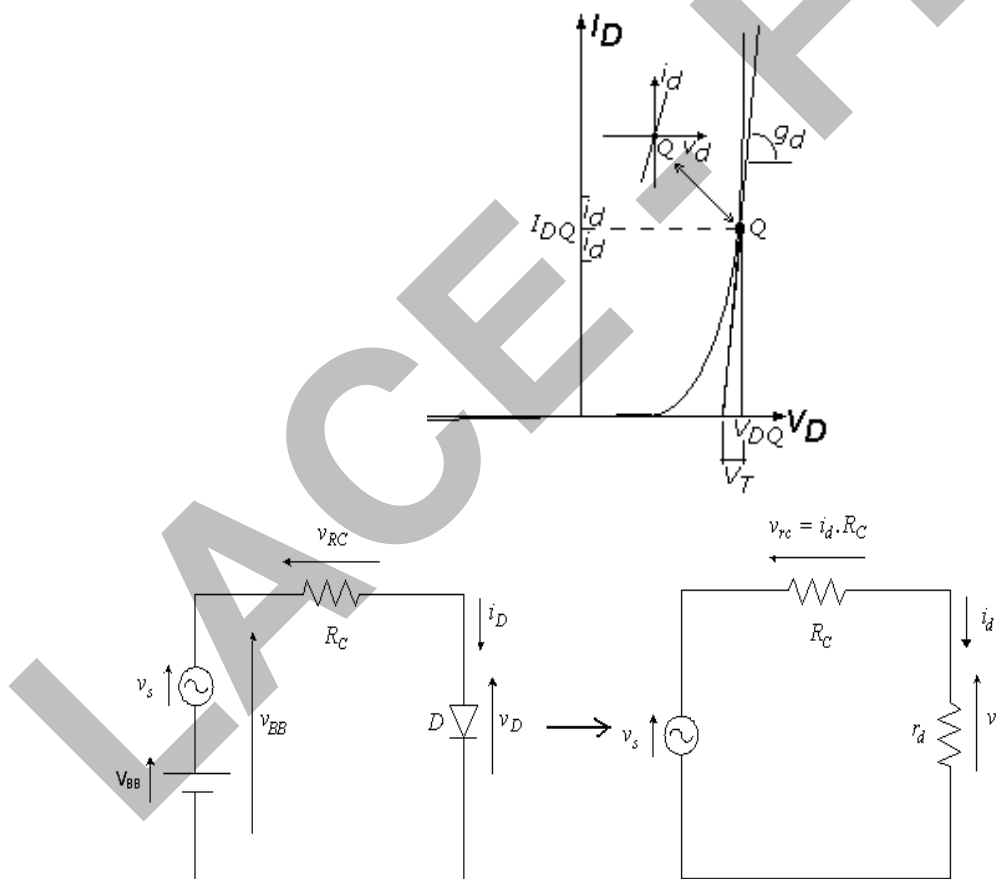


Fig. A.12

Definimos como *conductancia dinámica* g_d del diodo para un punto de reposo Q determinado, a la relación que existe entre una pequeña

variación de corriente entre sus bornes debida a la variación de la tensión entre sus terminales.

$$g_d = \Delta I_D / \Delta V_D |_Q = 1/r_d = I_{DQ} / V_T$$

La resistencia dinámica, r_d , se obtendrá como V_T / I_{DQ} y su valor para corrientes medias (zona de validez del modelo ideal del diodo) será del orden de los ohms (25Ω para $I_{DQ} = 1mA$).

De igual modo, puede hallarse una resistencia dinámica para el diodo polarizado en inversa, r_{inv} , que será del orden de los $10M\Omega$ (pendiente de la tangente de muy bajo valor). En este caso, su valor no es deducible en base a la ecuación del diodo ideal, ya que como se vió la corriente en inversa no corresponde con la del modelo ideal, especialmente en silicio.

Este modelo del diodo a los efectos de la señal aplicada por sobre el valor de continua de reposo, resulta válido mientras pueda aproximarse la característica del diodo a su recta tangente al punto de reposo con bajo error (o menor que un error prefijado). De acuerdo con la definición de r_d , un nivel de señal pequeña correspondería a valores muy inferiores a V_T (ver gráfico anterior). Por otro lado, la aproximación del modelo de pequeña señal del diodo a una resistencia, será válido para un rango de frecuencias en que no tengan influencia efectos reactivos propios del funcionamiento de la juntura P-N (rango de frecuencias bajas y medias).

Efectos reactivos en una juntura P-N

Capacitancia de juntura: Supongamos que a un diodo polarizado en directa se le aplica un incremento positivo. Dicho incremento pondrá más en directa la juntura, por lo que disminuirá el ancho de la zona desierta, l_d . Obviamente, la cantidad de electrones que entran al material N deberá tener la misma carga total que la cantidad de lagunas que entran al material P, por neutralidad eléctrica del conjunto, es decir: $\Delta Q_N = \Delta Q_P$.

Si bien los electrones y lagunas se movieron para neutralizar parte de la zona desierta de sus respectivos materiales, a efectos de la fuente se puede considerar a este sistema como si se comportase como un capacitor cuyo dieléctrico resulta ser el silicio de la zona desierta. Las variaciones de carga relacionadas con la modificación de la zona desierta de los lados N y P, se corresponderían con las cargas acumuladas en las placas de ese capacitor. De este modo, para el cálculo de la capacidad, es equivalente trabajar con la carga que entra a neutralizar a los iones fijos no compensados de la zona desierta o con la disminución de la carga de iones fijos no compensados en dicha zona. A este efecto capacitivo se lo denomina *capacitancia de juntura o transición*, C_j y resulta del orden de los pF .

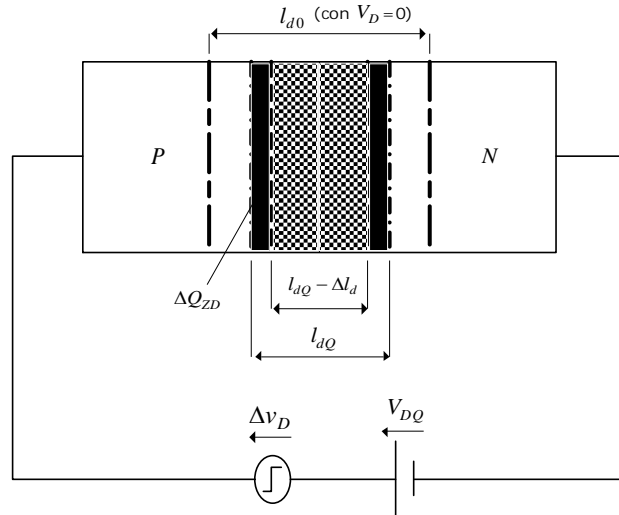


Fig. A.13

$$C_j = \Delta Q_{ZD} / \Delta V_D \big|_Q$$

C_j se define como una derivada, pues interesa su valor para incrementos de tensión superpuestos a la continua de polarización. El valor hallado depende de la tensión V_{DQ} a partir de la que se realiza el incremento y no puede obtenerse como Q_{ZD}/V_{DQ} dado que la relación entre ambos parámetros no sería lineal.

Si $\Delta l_d \ll l_{dQ}$, puede considerarse que las variaciones de carga de la zona desierta se realizarán en una zona de ancho despreciable, que puede admitirse como laminar, por lo que puede considerarse que este efecto capacitivo se debe a un capacitor de placas planas y paralelas separadas por una distancia l_{dQ} , es decir:

$$C_j = \varepsilon \cdot A / l_{dQ}$$

donde: $\varepsilon = \varepsilon_0 \cdot \varepsilon_r$ es la permitividad dieléctrica del silicio.

A : la sección transversal de la juntura.

l_{dQ} : la longitud de la zona desierta para V_{DQ} .

La expresión de C_j hallada resulta evidentemente válida para una señal de pequeña amplitud ya que esta puede considerarse como un incremento variable en el tiempo, de modo análogo al realizado al analizar el concepto de resistencia dinámica.

Admitiendo una juntura abrupta, la distribución de carga espacial de los iones no compensados será constante, por lo que el campo eléctrico variará linealmente a través de la zona desierta, y el potencial lo hará cuadráticamente. De este modo, puede admitirse que la longitud de la zona desierta será función de la raíz cuadrada de la tensión aplicada y por lo tanto la capacitancia de juntura variará de forma inversamente proporcional con la raíz de la tensión aplicada al diodo.

Obviamente, para polarización inversa, la expresión hallada para C_j resulta válida, disminuyendo a medida que el diodo esté más en inversa ya que aumenta la longitud de la zona desierta.

Si bien a medida que el diodo esté más en directa, la capacitancia de difusión será mayor, a medida que la polarización se acerque a la zona de altas corrientes (o valores de V_D cercanos al voltaje de barrera del contacto de la unión, V_{jo}), el pasaje de portadores a través de la zona desierta será tan elevado que no existirá variación de carga no compensada, por lo que $C_j \rightarrow 0$.

Capacitancia de difusión: Bajo las mismas condiciones iniciales en que se planteó la explicación de C_j , con un incremento ΔV_D , aumentará la cantidad de portadores mayoritarios que se inyectan hacia el lado opuesto de la juntura, ocasionando un exceso de las concentraciones de equilibrio de los minoritarios de dicha zona. Esta acumulación de carga, permanecerá hasta que todo el exceso desaparezca o por recombinación o por haberse trasladado los portadores hasta los contactos metálicos, donde se recombinan. En el primer caso estamos ante un diodo largo, donde la acumulación estará caracterizada por el *tiempo de vida medio* de los portadores, mientras que en el segundo caso, donde la recombinación es muy baja (correspondiente a un diodo corto), el tiempo necesario para que desaparezca la acumulación será el *tiempo de tránsito* a través de las zonas neutras, que será menor al de vida medio.

Este efecto de acumulación de carga móvil en las zonas neutras puede asemejarse a un efecto capacitivo, que se pondrá más en evidencia a medida que los tiempos de la señal incremental (período de una señal alterna, por ejemplo) sean del orden de alguno de los tiempos característicos. Esta capacitancia parásita se denomina capacitancia de difusión, C_d y se define como:

$$\begin{aligned} C_d &= \Delta Q_{ZN} / \Delta V_D \big|_Q = (\Delta Q_{ZN} / \Delta V_D) \big|_Q \Delta t / \Delta t \\ &= \Delta I_D / \Delta V_D \big|_Q \Delta t = g_d \cdot T_d = (I_{DQ} / V_T) \cdot T_d \end{aligned}$$

donde :

T_d : corresponderá al tiempo de vida o de tránsito según sea el diodo largo o corto. Ambos son en general del orden de los nseg.

Se observa en la expresión de C_d , que depende directamente de la corriente de reposo que circula por el diodo, lo cual es lógico, dado que el exceso de carga estará dado por la mayor o menor corriente que circule. Esto hace que, a corrientes medias, C_d sea del orden de decenas a cientos de pF .

En inversa, no habrá acumulación en zona neutra (corrientes casi nulas respecto a las que circulan en directa), por lo que $C_d \rightarrow 0$.

En resumen, si graficamos la variación de C_d y C_j con V_{DQ} , observamos que en directa predomina C_d y en inversa C_j , quedando simplificados los modelos de pequeña señal para altas frecuencias como se indica en la Fig. A.14.

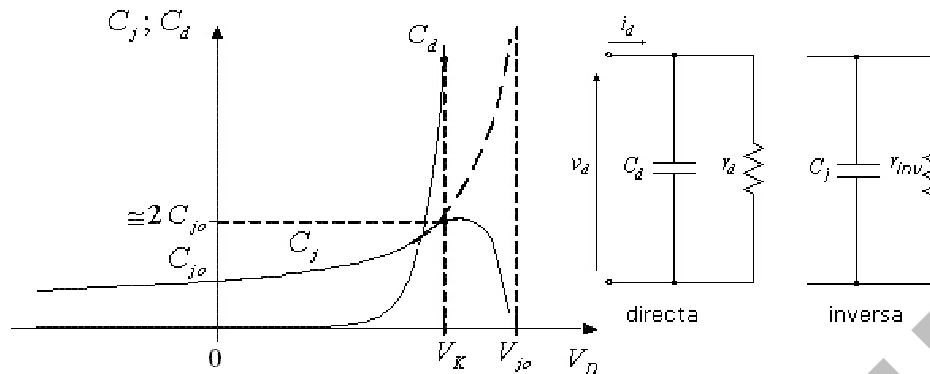


Fig. A.14

A.3. Dispositivos sólidos activos o de control de potencia

Generalidades

Un dispositivo de control de potencia (o control de señal) distribuye la potencia entregada a una carga R_C por una fuente de alimentación V_{CC} mediante el control de la corriente a través de la carga, por medio de una tensión (o corriente) que no forma parte de la malla donde se encuentra la carga. Este dispositivo posee dos pares de terminales: 1-1', llamados de entrada o control (donde se encuentra la tensión o corriente de control) y 2-2', de salida, donde se encuentra la carga y circula la corriente I_O , controlada por la entrada.

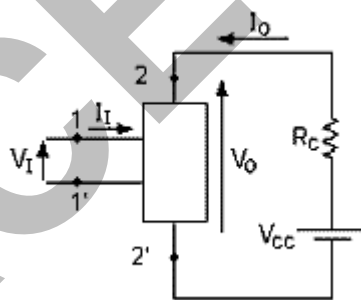


Fig. A.15

Para describir a un dispositivo de control de potencia particular, se necesitan en general al menos dos curvas características (Fig. A.16). Una debe relacionar el parámetro controlado con el de control (por ejemplo $I_O = f(V_I)$), llamada *característica de transferencia*. Otra debe relacionar su tensión y corriente de salida para distintos valores de la variable de control (por ejemplo, $I_O = f(V_O)$ para $V_I = cte$), llamada *característica de salida*. También puede ser de interés la característica de entrada ($I_I = f(V_I)$).

El control de I_O se hará dentro de una zona de valores de V_O . El límite inferior estará dado por la forma de funcionamiento del

dispositivo y el superior lo limitará o bien V_{CC} o bien la máxima potencia que pueda disipar el dispositivo entre sus terminales de salida.

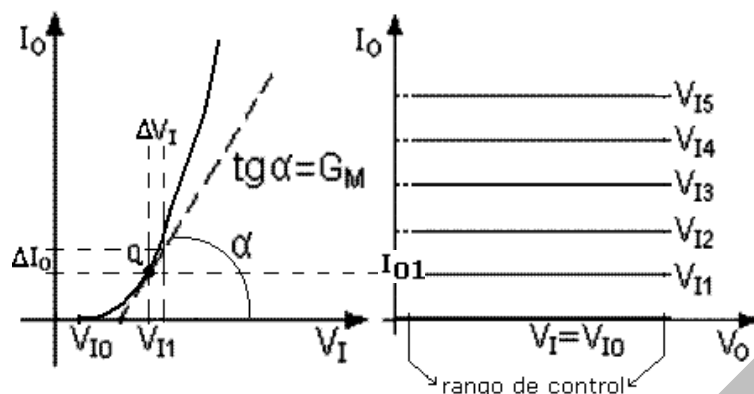


Fig. A.16

Otra forma de analizarlo es ver que, variando la tensión del control V_I , se está variando la relación V_O/I_O , que es la resistencia estática de salida que presenta el dispositivo. Esto equivale a que desde el circuito de entrada, se transfiera a la malla de salida, una resistencia variable a partir de las variaciones del parámetro de control, que es lo que se denomina *efecto transistor*.

Por otro lado, las pequeñas variaciones de V_I sobre I_O en los alrededores de V_{I1} (o V_{I0}), pueden analizarse sobre la recta tangente a la característica de transferencia (linealización de la característica para señal pequeña), obteniéndose $\Delta I_O = G_M \cdot \Delta V_I$, que muestra el efecto de control de la corriente de señal de salida debido a una tensión de señal de entrada. El parámetro G_M , que relaciona una corriente de una malla con una tensión en otra, se lo conoce como conductancia mutua o transconductancia del dispositivo. Así definida, G_M relaciona parámetros incrementales o de señal, pero su valor depende de la corriente de reposo del dispositivo (o corriente continua de funcionamiento o polarización $I_{O1} = I_{O0}$).

De acuerdo a su principio físico de funcionamiento, existen dos tipos de transistores, los transistores bipolares de juntura (TBJ) y los unipolares o transistores de efecto de campo (FET). Dentro de estos últimos podemos subdividirlos en, de efecto de campo de juntura (de juntura P-N o JFET y de juntura M-S o MESFET) y de efecto de campo metal-óxido semiconductor (MOSFET).

A.3.1. Transistor bipolar de juntura o TBJ

El TBJ tiene una estructura basada en dos junturas PN: una N+P y otra PN-, con una zona P corta compartida por ambas junturas.

Si se lo polariza como se indica en la Fig. A.17 y se admite un valor de V_{CC} tal que $V_{CE} = V_O \gg V_I \cong 0,7V$, la juntura N+P se encontrará en directa y la PN- en inversa. La inyección de electrones mayoritarios del lado N+ al P será la predominante. Una vez en la zona P, los portadores (ahora minoritarios) casi no se recombinarán con los

mayoritarios que entren por el terminal positivo de la batería V_I (por ser una zona corta) y se encontrarán con un campo elevado en la juntura PN- en inversa, que los arrastrará hacia la zona N- y serán colectados por la batería V_{CC} . Por lo tanto, la corriente del diodo P+N en directa (cuyo sentido convencional es opuesta a la circulación de los electrones) circulará por la malla que contiene a V_{CC} y la carga R_C .

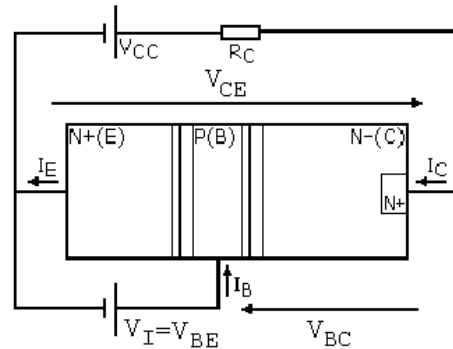


Fig. A.17

Si aumenta V_I la juntura N+P estará más en directa, aumentando la corriente que circula por la malla de salida, poniéndose en evidencia el efecto de control de potencia.

La zona N+ que inyecta los electrones mayoritarios se la llama *emisor* y la N- que los recibe, *colector*. La zona intermedia P se la conoce como *base*, término que proviene de la forma de construcción de los primeros transistores, donde se partía de un material base P y se difundían las zonas N de ambas caras del material base.

Esta estructura se denomina *transistor bipolar NPN*.

De esta forma, como la corriente entrante por el terminal de colector I_C , no es otra que prácticamente la del diodo en directa base-emisor (salvo el pequeño porcentaje de portadores que se recombinan en la base), la ecuación de transferencia que pone en evidencia el efecto transistor será:

$$I_C = I_S \cdot (e^{V_{BE}/V_T})$$

donde: I_S es la corriente de saturación inversa del diodo base-emisor.

Esta forma de funcionamiento (B-E en directa y B-C en inversa) se denomina *modo activo directo (MAD)* y es la forma normal de funcionamiento del transistor como control de señal. La forma de polarización opuesta (B-E en inversa y B-C en directa), se denomina *modo activo inverso (MAI)*, donde el control de potencia está degradado ya que el colector, que es el inyector, tiene baja contaminación.

Por otro lado, si ambas junturas están en inversa, la corriente que circula por el transistor será prácticamente nula y se dice que el dispositivo se encuentra en modo de *corte*. Si ambas junturas están en directa, la corriente I_C del transistor la fijará el circuito exterior de la

mallla de salida (en este caso, será prácticamente V_{CC}/R_C , ya que la tensión entre el terminal de colector y el de emisor, V_{CE} , será nula o casi nula, dependiendo si las junturas están polarizadas en directa a igual tensión o no). Este modo de funcionamiento se denomina *saturación*. En estos últimos dos modos de trabajo, el transistor no ejerce ningún control sobre la corriente de salida y funciona como una llave abierta (corte) o cerrada (saturación).

Las características de transferencia $I_C = f(V_{BE})$ y salida $I_C = f(V_{CE})|_{V_{BE} = \text{cte}}$, tendrán la forma indicada en la Fig. A.18. Se indican las zonas de las distintas formas de trabajo en la característica de salida. El lugar geométrico que limita las zonas de *MAD* y saturación resulta de la condición $V_{CE} = V_{BE}$, es decir, con la juntura base-colector sin polarizar. Esta condición se conoce como *saturación incipiente*. Cabe aclarar que este límite es teórico, pues mientras la juntura BC esté débilmente polarizada en directa, se mantendrá prácticamente el control de I_C con V_{BE} . Para valores de $V_{CE} \rightarrow 0$, la juntura colector base estará más en directa, circulando una corriente de base hacia el colector, que se opondrá a la que circula por arrastre de los electrones que provenía del emisor. Por lo tanto I_C , disminuirá hasta anularse. Del mismo modo se justifican las zonas de saturación y *MAI* en el tercer cuadrante del gráfico de la característica de salida.

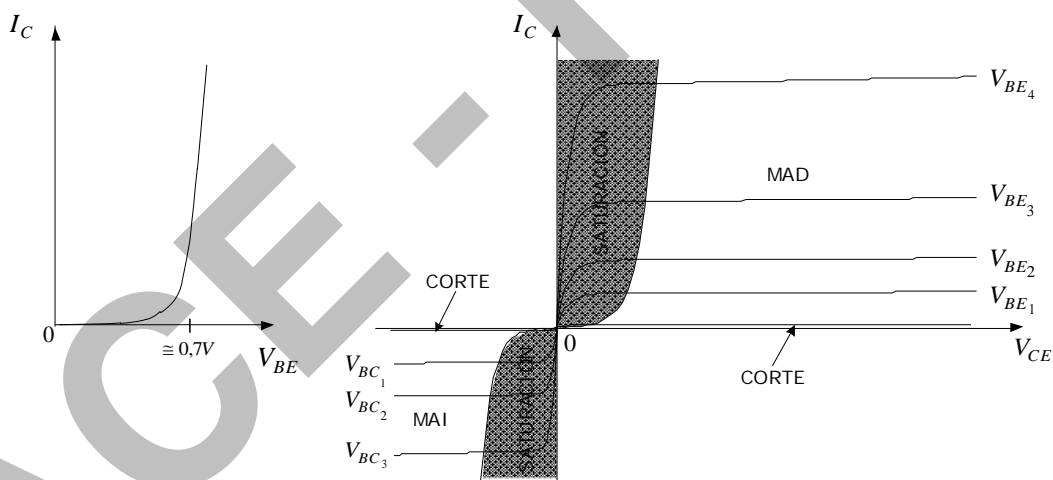


Fig. A.18

El funcionamiento del TBJ, como dispositivo de control de potencia, puede modelizarse mediante un circuito eléctrico equivalente, donde se muestran los dos diodos (BE y BC) y el efecto de control de la corriente I_C con V_{BE} , con la fuente de corriente $\alpha_F I_E = I_S \cdot (e^{V_{BE}/V_T})$.

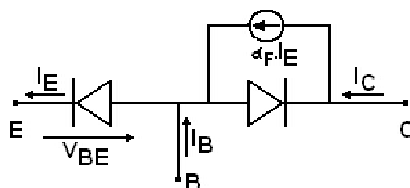


Fig. A.19

La constante $\alpha_F = I_C / I_E$ se conoce como *factor de transporte* y su valor resulta muy cercano a la unidad (por ejemplo, 0,995) y pone en evidencia la pequeña porción de portadores que se recombinan en la base. Si se busca relacionar la corriente de colector con la de base, teniendo en cuenta que $I_E = I_C + I_B$, se llega a:

$$I_C = I_B \cdot \alpha_F / (1 - \alpha_F) = I_B \cdot \beta_F$$

Donde β_F se denomina *ganancia de corriente en directa* (o *MAD*) y es un parámetro común en las hojas de datos de transistores bipolares. Su valor estará en el orden de 100 en los TBJ de baja potencia (menor a 1 watt). Este modelo eléctrico para el TBJ se lo conoce como modelo de Ebers-Moll ⁽¹⁾. Puede verse que en *corte*, con ambos diodos en inversa, el generador de corriente estará abierto, siendo I_C prácticamente nula, mientras que en saturación, la corriente por el diodo base-colector en directa se opone al generador controlado, disminuyendo el valor de I_C , tal como se obtiene en las características de salida.

A.3.2. Transistor bipolar real

Las desviaciones en el funcionamiento del transistor real son debidas a efectos similares a los indicados para el diodo y afectarán tanto la característica de transferencia como la de salida:

- 1) Recombinación en la zona desierta BE para muy bajas corrientes.
- 2) Caída de tensión en las zonas neutras de B y E en altas corrientes.
- 3) Ruptura en la juntura B-E en inversa para tensiones bajas (del orden de algunos volts).
- 4) Ruptura en la juntura B-C para V_{CE} elevadas (del orden de las decenas de volts).

Sin embargo, para valores de V_{CE} menores al de ruptura, existe otro fenómeno propio de los TBJ llamado *modulación del ancho de la base o efecto Early*. Al aumentar V_{CE} , la juntura BC estará más en inversa y su zona desierta aumentará, angostándose el ancho de la base. Dado que la corriente de colector se debe al tránsito durante un tiempo T_{tr} de la carga en exceso de minoritarios en la base Q_B inyectada desde emisor, podemos escribir que: $I_C = Q_B / T_{tr}$. Dado que la base se angostó debido al aumento de V_{CE} , disminuyó T_{tr} , aumentando I_C . Este aumento, si bien muy pequeño frente al control exponencial con V_{BE} , hace que las características de salida tengan una leve pendiente positiva.

(1) Se ha despreciado en esta explicación el efecto de *MAI*. Para tenerlo en cuenta, se debería agregar al modelo indicado una fuente de corriente similar pero en paralelo con el diodo BE, afectada por una constante α_R (*factor de transporte en inversa*), de valor muy inferior al de directa (por ejemplo $\alpha_R \approx 0,2$).

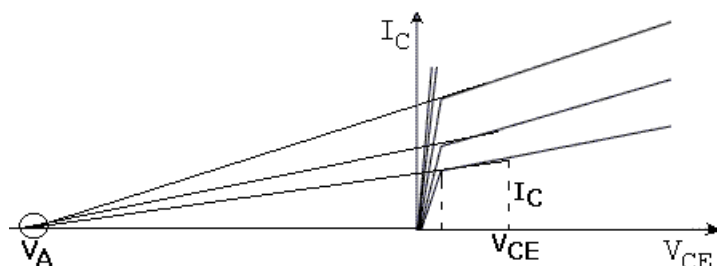


Fig. A.20

Si se extiende esa pendiente hasta cortar el eje de abscisas, las rectas de todas las características cortan al eje V_{CE} en una zona reducida, cuyo valor promedio se denomina V_A o *tensión de Early*. Este valor está en el orden de los 100 V.

Para obtener la expresión que corrige el valor de I_C por efecto Early, simplemente se obtiene la pendiente de las características de salida en forma gráfica. Es decir, admitiendo $V_{CE} \ll |V_A|$, la pendiente será $I_C / |V_A|$, por lo que I_C se puede corregir mediante una aproximación lineal, del siguiente modo:

$$I_C = I_S \cdot (e^{V_{BE}/V_T}) \cdot (1 + V_{CE} / |V_A|)$$

El transistor de estructura complementaria, es decir *PNP*, tendrá un modelo similar al planteado hasta aquí, con la salvedad de que las tensiones y corrientes tendrán signo opuesto a las indicadas para el *NPN*. (es decir, al ser el emisor un inyector de huecos, las corrientes tendrán sentido opuesto).

Los símbolos con que se presentan circuitalmente ambos tipos de TBJ, son los indicados en la Fig. A.21, donde se muestran los sentidos de referencia positivos para la continua.

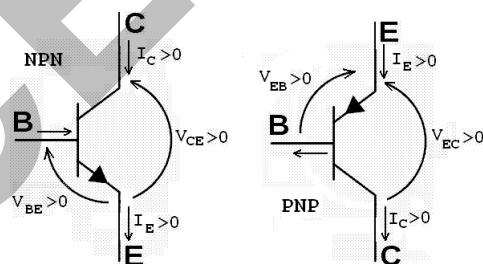


Fig. A.21

A.3.3. Transistores de efecto de campo o FET

Los FET son conocidos como transistores unipolares, ya que su funcionamiento se basa en el control de la conducción de un tipo de portador a través de una zona de semiconductor llamada *canal*. Es decir, el efecto transistor está dado por el control de la conductancia de dicho canal. Siendo la conductancia de un material proporcional a la conductividad y a su área transversal (sección), si para controlar la

corriente se varía la conductividad estamos en presencia de un MOSFET y si se varía la sección, de un JFET o MESFET.

Transistor de efecto de campo de juntura o JFET

La estructura de este transistor es la indicada en la Fig. A.22, donde puede verse que se basa en una juntura P+N en inversa, donde la corriente a controlar circula a través de la zona N o *canal*, cerrándose por la malla que forma la fuente V_{CC} y la carga R_D . Así conectada, V_{CC} entrega los electrones que circulan desde el terminal de *source* (fuente) al de *drain* (drenaje).

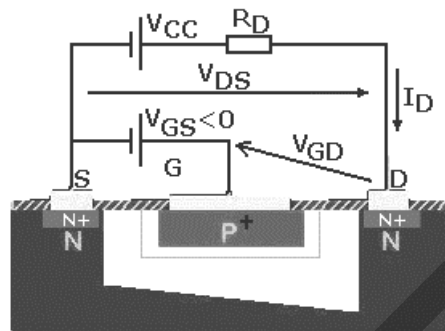


Fig. A.22

En forma general, puede escribirse:

$$I_D = G_{canal} \cdot V_{DS}$$

La conductancia del canal es $G_c = \sigma_c \cdot S_c / L_c$, donde:

σ_c : conductividad del material N.

S_c : sección transversal del canal

L_c : longitud del canal (distancia entre los terminales S y D)

Esta estructura se denomina JFET canal N.

Al conectar la fuente entre el terminal de *gate* (compuerta) y *source*, $V_{GS} < 0$, aumentará la zona desierta de la juntura P+N, ocupando parte del canal y por lo tanto variando su conductancia al modificarse S_c . Si V_{GS} se hace suficientemente negativa, la zona desierta podría ocupar todo el canal, anulando la corriente. Este valor de $V_{GS} = V_P$ para el cual la zona desierta ocupa completamente el canal, se lo conoce como *tensión de pinch-off* y es del orden de algunos volts.

Hasta aquí observamos que I_D dependerá tanto de V_{DS} como de V_{GS} . Sin embargo, si V_{CC} posee un valor tal que V_{DS} se hace superior a $|V_{GS}|$ la juntura cercana al *drain* se encontrará más en inversa que la próxima al *source*. De esta forma, la zona desierta será más amplia del lado del terminal D. Por lo tanto, cuando V_{DS} sea suficientemente grande como para que la tensión V_{GD} alcance el valor de V_P , la zona desierta ocupará completamente el canal en los alrededores de dicho terminal (*punto de estrangulamiento*).

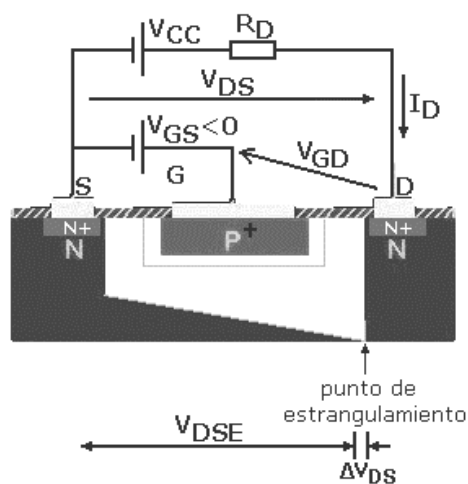


Fig. A.23

Bajo esas condiciones, en el canal caerá una tensión $V_{DSE} = V_{GS} - V_P$ (conocida como *tensión drain-source de estrangulamiento*) y todo aumento de V_{DS} por encima de V_{DSE} caerá en el punto de estrangulamiento, dado que al ser una zona desierta presentará conductividad nula.

Entonces, la corriente I_D a través del canal no dependerá ya de V_{DS} y circulará a través de la zona delgada de estrangulamiento porque el campo $V_{DS}/l_{\text{zona de estrang.}}$ en dicha zona será muy elevado por ser $l_{\text{zona de estrang.}} \rightarrow 0$ y arrastrará a los electrones que se acerquen al drain provenientes del source.

$$I_D = G_c \cdot V_{DSE} = G_c \cdot (V_{GS} - V_P)$$

Si a esta expresión le agregamos que G_c también dependerá de V_{GS} , ya que la variación de la sección del canal es debido al ensanchamiento de la zona desierta P+N, la ecuación de transferencia del JFET se puede escribir finalmente como ⁽¹⁾:

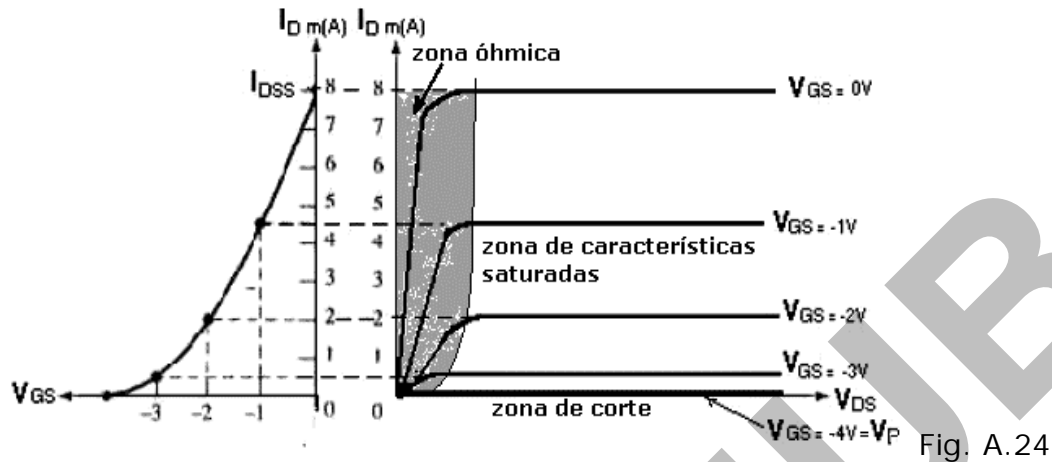
$$I_D = BETA \cdot (V_{GS} - V_P)^2 = I_{DSS} \cdot (1 - V_{GS}/V_P)^2$$

donde $BETA = I_{DSS} / V_P^2$, siendo I_{DSS} la corriente máxima por el canal (para $V_{GS} = 0$) cuyo valor ronda los 10 mA. (Para $V_{GS} > 0$ la juntura GS se encontrará en directa y circulará una corriente importante por estos terminales, no habiendo ya control de la corriente I_D).

Las características de transferencia $I_D = f(V_{GS})$ y salida $I_D = f(V_{DS})/V_{GS} = cte$, se muestran en la Fig. A.24. En la característica de salida se indican las distintas zonas trabajo. El lugar geométrico que limita las zonas de *modo de control de potencia* o *características saturadas* y *zona óhmica*, es $V_{DS} = V_{DSE}$ (condición de *estrangulamiento incipiente*). Para valores de $V_{DS} \rightarrow 0$, la zona desierta mantendrá un ancho casi constante a lo largo del canal sin estrangularlo.

(1) Como la longitud de la zona desierta resulta proporcional a la raíz de la tensión aplicada sobre la juntura, la ecuación de transferencia sería función de $V_{GS}^{3/2}$, pero se la aproxima a una función cuadrática, teniendo en cuenta que la conductividad, que se supuso constante a lo largo del canal, no lo será debido a la influencia del campo transversal de la zona desierta.

Bajo esta condición, puede verse que el JFET posee una característica de salida lineal equivalente al de una resistencia, cuya pendiente varía al variar V_{GS} . Es decir, su comportamiento en esta zona es la de una resistencia variable con V_{GS} , de allí su nombre de zona óhmica. Por ello, al JFET también se lo usa en esta zona ⁽¹⁾.



En el transistor JFET real, existirá una V_{DS} máxima dada por la ruptura de la juntura GD y también en la zona saturada tendrán una leve pendiente positiva debido a la *modulación del largo del canal*, ya que el punto de estrangulamiento se irá corriendo hacia el source a medida que aumenta V_{DS} , creándose una zona de estrangulamiento que irá acortando la longitud efectiva de canal. Haciendo una aproximación similar a la del TBJ, se puede corregir este efecto añadiendo un término lineal a la ecuación de transferencia.

$$I_D = I_{DSS} \cdot (1 - V_{GS}/V_P)^2 \cdot (1 + \lambda V_{DS})$$

donde λ , del orden de $0,01 \text{ V}^{-1}$, es equivalente a $1/V_A$ del TBJ.

El transistor de estructura complementaria, es decir *canal P*, tendrá un modelo similar al planteado hasta aquí, con la salvedad de que las tensiones y corrientes tendrán signo opuesto a las indicadas para el canal N. (es decir, al circular huecos por el canal, las corrientes tendrán sentido opuesto).

Los símbolos con que se presentan circuitalmente ambos tipos de JFET, son los indicados en la Fig. A.25, donde se muestran los sentidos de referencia para la continua.

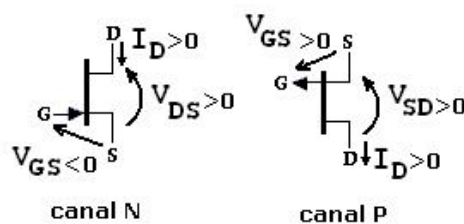


Fig. A.25

(1) Cabe aclarar que este comportamiento también lo tiene el TBJ, pero dado que para $V_{CE} \rightarrow 0$ la variación de la pendiente de la característica de salida es exponencial con V_{BE} , no resulta de utilidad por la dificultad en controlar su variación.

Si se reemplaza la juntura PN por una Metal-Semiconductor rectificante, el funcionamiento será similar al explicado. Este tipo de JFET se conoce como MESFET y se lo utiliza en muy altas frecuencias, ya que en general se construye con arseniuro de galio, semiconductor que posee una movilidad ocho veces superior al silicio.

Transistor de efecto de campo metal, óxido, semiconductor o MOSFET

La estructura de este transistor es la indicada en la Fig. A.26. El MOSFET (ó MOS) es simplemente un capacitor en el cual se controla el desplazamiento de su carga por una de sus placas. La tensión $V_{GS} > 0$ carga positivamente al metal del gate (aluminio o silicio policristalino), lo que provoca que, para neutralizar la carga, los electrones (minoritarios) del material P, llamado *sustrato o bulk*, se acerquen a la cara inferior del óxido (en el caso de silicio, óxido de silicio) invirtiendo la concentración de contaminante en dicha zona superficial (pasa de P a N). Obviamente no habrá circulación de carga a través del gate al debido a la presencia del dieléctrico aislante. Por otro lado, el terminal de *sustrato o bulk* se supondrá en principio a igual potencial que el source para que la juntura N+P formada no quede polarizada, evitando que circule corriente a través del sustrato ⁽¹⁾.

Cuando la tensión V_{GS} alcanza un valor V_{Th} (ó V_T), la concentración superficial N resulta igual a la del sustrato P y se dice que queda formado o *inducido* un *canal* de tipo N que conecta el source y el drain. A dicho valor de V_{Th} se lo llama *tensión de umbral*, y es del orden de algunos volts.

La corriente a controlar circula a través de la zona superficial N o *canal*, cerrándose por la malla que forma la fuente V_{CC} y la carga R_D . Así conectada, V_{CC} entrega los electrones que circulan desde el terminal de *source* (fuente) al de *drain* (drenaje). Esta estructura se denomina MOSFET canal N.

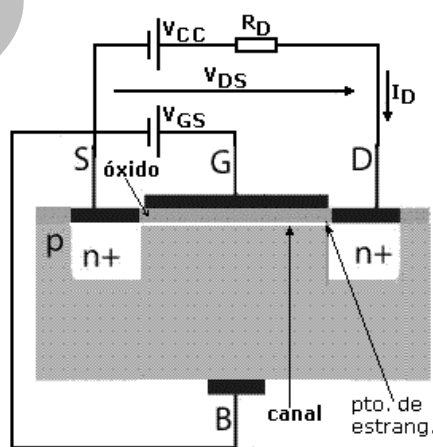


Fig. A.26

(1) En un circuito integrado, el sustrato P se encontrará al menor potencial posible, asegurando que la juntura N+P se encuentre en inversa o a lo sumo no polarizada, para cualquier MOSFET del circuito.

Se puede plantear entonces una ecuación equivalente a la del JFET: $I_D = G_{canal} \cdot V_{DS}$

Para $V_{GS} > V_{Th}$, aumenta la cantidad de carga en el canal y por lo tanto aumenta su conductancia al modificarse σ_c .

Siguiendo el mismo razonamiento que al explicar el funcionamiento del JFET, si V_{CC} posee un valor tal que V_{DS} se hace superior a V_{GS} , en la zona cercana al drain habrá menor densidad de carga superficial que en la próxima al *source*. Por lo tanto, cuando V_{DS} sea suficientemente grande como para que la tensión V_{GD} alcance el valor de V_{Th} , en los alrededores del terminal D, no habrá canal (de acuerdo con la definición de canal) ⁽¹⁾ y la concentración de carga superficial móvil será nula en dicha zona (*punto de estrangulamiento*).

Bajo esas condiciones, en el canal caerá una tensión $V_{DSE} = V_{GS} - V_{Th}$ y todo aumento de V_{DS} caerá en el punto de estrangulamiento, dado que al ser una zona desierta presentará conductividad nula.

Entonces, la corriente I_D a través del canal no dependerá ya de V_{DS} y circulará a través de la zona delgada de estrangulamiento porque el campo $V_{DS}/l_{zona\ de\ estrang.}$ en dicha zona será muy elevado por ser $l_{zona\ de\ estrang.} \rightarrow 0$ y arrastrará a los electrones que se acerquen al drain provenientes del source.

$$I_D = G_c \cdot V_{DSE} = G_c \cdot (V_{GS} - V_{Th}) = (\sigma_c \cdot S_c / L_c) \cdot (V_{GS} - V_{Th})$$

Donde la sección S_c del canal resulta del producto de su ancho W_c por su espesor "e". En este caso esta última dimensión es de capas atómicas, dado que el canal es una concentración superficial de carga. Por otro lado, su conductividad $\sigma_c = Q_c \cdot \mu_n$, dependerá de V_{GS} , ya que dicha tensión controla la carga superficial en las placas del capacitor de placas planas C_{ox} , es decir $Q_c = C_{ox} \cdot (V_{GS} - V_{Th})$, Por lo que la ecuación de transferencia del MOS se puede escribir como:

$$I_D = k' (W_c / L_c) \cdot (V_{GS} - V_{Th})^2 = K \cdot (V_{GS} - V_{Th})^2$$

donde la constante K depende fuertemente de parámetros constructivos del MOS y su valor estará en general entre los 0,1 y 10 mA/V².

Las características de transferencia $I_D = f(V_{GS})$ y salida $I_D = f(V_{DS})/V_{GS} = cte$, tendrán una forma similar a la del JFET. La ubicación de las distintas formas de trabajo también serán las mismas (Fig. A.27).

Se observa que se agregó una característica de transferencia con $V_{Th} < 0$. Este tipo de MOSFET se denomina de *canal preformado* y se logra implantando iones positivos dentro del óxido, por lo que el canal puede llegar a formarse sin tensión V_{GS} aplicada y para anular la corriente habrá que aplicar un $V_{GS} < 0$.

(1) En realidad, para valores de $V_{GS} < V_{Th}$, existirá un canal con baja densidad de portadores libres, por lo que la corriente circulante (corriente de subumbral) deberá ser mucho menor a las obtenidas por encima del umbral.

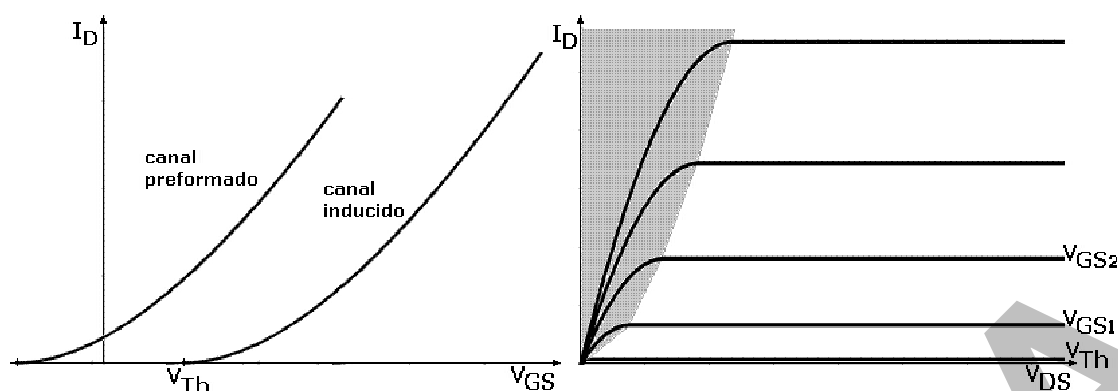


Fig. A.27

En el transistor MOSFET real, existirá una V_{DS} máxima dada por la ruptura de la juntura BD y también en la zona saturada las características de salida tendrán una leve pendiente positiva debido a la *modulación del largo del canal*, al igual que el JFET.

$$I_D = K \cdot (V_{GS} - V_{Th})^2 \cdot (1 + \lambda V_{DS})$$

El transistor de estructura complementaria, es decir *canal P*, tendrá un modelo similar al planteado hasta aquí, con la salvedad de que las tensiones y corrientes tendrán signo opuesto a las indicadas para el canal N.

Los símbolos con que se presentan circuitualmente ambos tipos de MOSFET, son los indicados en la Fig. A.28, donde se muestran los sentidos de referencia positivos para la continua. La flecha indica el tipo de sustrato. En este caso se lo ha conectado al source, pero como se dijo antes, en general no es así en un circuito integrado.

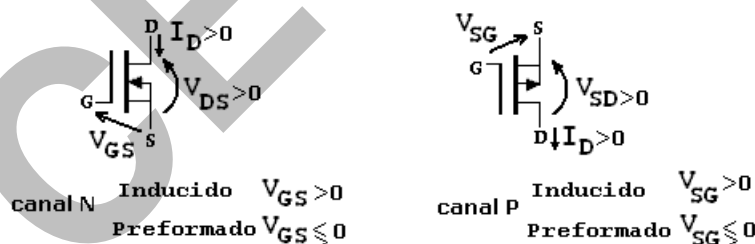


Fig. A.28

El efecto del potencial de sustrato, cuando la juntura BS y BD están en inversa, es el de empobrecer el canal, ya que la zonas desiertas de dichas junturas aumentan (con la raíz de V_{SB}). Esto hace que, para lograr volver a la condición de canal formado, se deba aumentar el valor de V_{GS} , o sea que V_{Th} aumenta con la tensión V_{SB} (es decir a medida que se pone más en inversa la juntura con el sustrato). La relación de V_{Th} con el voltaje del sustrato está dada aproximadamente por:

$$V_{Th} \cong V_{Th (V_{SB} = 0)} + \gamma \cdot (V_{SB})^{1/2}$$

donde la constante γ es del orden de $0,1 \text{ V}^{1/2}$.

A.3.4. Modelos para pequeña señal de los transistores

Cuando a alguno de los transistores descriptos y trabajando como dispositivos de control de potencia, se le aplica una señal de entrada de pequeña amplitud, como la indicada en la Fig. A.29, su comportamiento a los efectos de dicha señal puede describirse mediante un modelo lineal, tal como se analizó para el diodo.

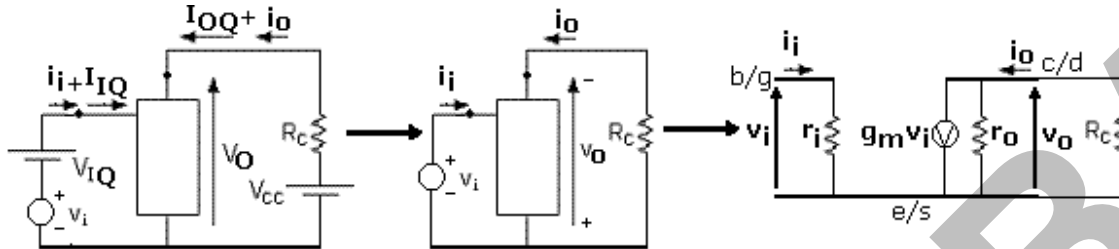


Fig. A.29

Los componentes lineales de dicho modelo representarán los distintos fenómenos físicos propios de cada tipo de transistor:

$g_m \cdot v_i$		Efecto de control de señal o efecto transistor. Variación de la tensión de control que modifica la corriente de salida. El factor de proporcionalidad es la transconductancia (pendiente de la recta tangente a la característica de transferencia del transistor en los alrededores del punto de reposo.	<i>TBJ</i> g_m	$= dI_C / dV_{BE} _Q = i_c / v_{be} _Q$ $= I_{CQ} / V_T$
			<i>JFET</i> g_m	$= dI_D / dV_{GS} _Q = i_d / v_{gs} _Q$ $= 2(I_{DQ} \cdot I_{DSS})^{1/2} / V_P$
			<i>MOSFET</i> g_m	$= dI_D / dV_{GS} _Q = i_d / v_{gs} _Q$ $= 2(K \cdot I_{DQ})^{1/2}$
r_i	<i>TBJ</i> r_π	Variación de la recombinación en la base al estar más en directa (o menos) la juntura BE	$= v_{be} / i_b _Q = \beta \cdot v_{be} / i_c = \beta / g_m$	
	<i>JFET</i> r_{gs}	Resistencia dinámica de la juntura en inversa GS	$= v_{gs} / i_g _Q = r_{gs}$	
	<i>MOSFET</i> r_{gs}	Resistencia del óxido aislante (que puede considerarse infinita)	$= v_{gs} / i_g _Q = r_{gs} \rightarrow \infty$	
r_o	<i>TBJ</i> r_o	Efecto Early ⁽¹⁾	$= v_{ce} / i_c _{Q, v_{be}=0} = V_A / I_{CQ}$	
	<i>JFET</i> r_{ds}	Modulación del largo del canal	$= v_{ds} / i_d _{Q, v_{gs}=0} = 1/\lambda I_{DQ}$	
	<i>MOSFET</i> r_{ds}	Modulación del largo del canal	$= v_{ds} / i_d _{Q, v_{gs}=0} = 1/\lambda I_{DQ}$	

(1) La variación del ancho de la base también afecta a la corriente de base, al variar la recombinación, por lo que en el modelo del transistor bipolar se agrega una resistencia, denominada r_μ entre los terminales de base y colector que muestra un aumento de la corriente de base que se suma a la que circula por r_π . Sin embargo como esta variación es muy pequeña esta resistencia resulta más de β veces superior a r_o , por lo que en general se termina despreciando.

En cuanto al límite de señal v_{imax} admitido como pequeña amplitud, suele aceptarse el segmento definido por la prolongación de la tangente a la curva de transferencia hasta el eje de v_i y el valor de V_{IQ} :

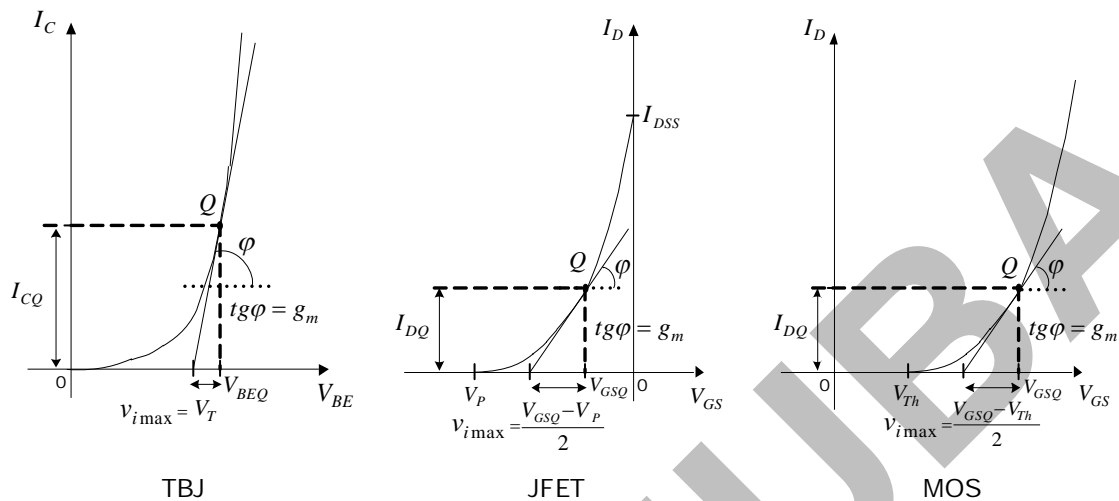


Fig. A.30

Al modelo de señal del TBJ, debe agregársele un elemento más que resulta muy influyente especialmente en altas frecuencias. La estructura de un TBJ real es por ejemplo la que se indica en la Fig. A.31 (existen también otras disposiciones de las distintas zonas especialmente en CI). La zona rectangular en línea de trazos es la zona activa del dispositivo (es decir el esquema básico sobre el que se analizó su funcionamiento). Puede observarse que la corriente de base debe atravesar un largo bloque semiconductor hasta alcanzar la zona activa. La resistencia que presenta este bloque semiconductor hace que la caída de tensión distribuida a través de la circulación de dicha corriente provoque que la juntura BE en la zona activa esté menos en directa que lo que se supone, de acuerdo con la V_{BE} aplicada desde los terminales externos. Esta resistencia se conoce como *del cuerpo de la base* (r_x ó r_b) y su valor ronda los 100Ω . Es el único elemento del modelo del TBJ que representa una resistencia concreta ya que el resto son equivalentes eléctricos que describen distintos fenómenos físicos. Al considerar r_x , la tensión de control en el modelo de pequeña señal se reduce a v_{be}' , debido al divisor resistivo con r_π .

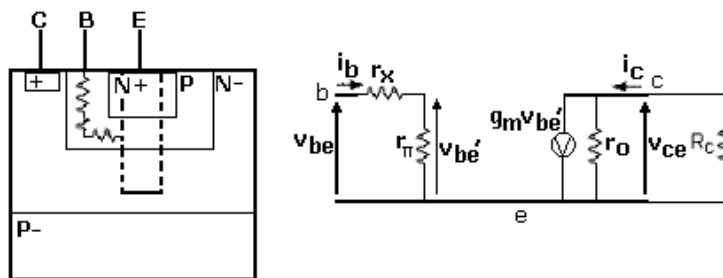


Fig. A.31

El sentido del generador controlado puede justificarse de manera simple analizando el comportamiento del dispositivo ante un incremento de la tensión de control.

Por ejemplo, para un NPN, a un incremento positivo de V_{BE} ($v_{be} > 0$), la juntura BE estará más en directa y por lo tanto habrá un aumento en el valor de I_C . Con corrientes de reposo positivas entrantes por el terminal de colector, este aumento es equivalente a considerar un incremento positivo de I_C ($i_c > 0$) también entrante por el terminal de colector, que se suma con el de reposo.

En el caso de un PNP, a un incremento positivo de V_{BE} ($v_{be} > 0$), la juntura BE estará menos en directa y por lo tanto habrá una disminución de I_C . Con corrientes de reposo positivas salientes por el terminal de colector, la disminución en el valor de I_C es equivalente a considerar un incremento positivo entrante que se resta con el de reposo.

El mismo tipo de análisis puede realizarse con los FET.

De esta manera concluimos que el modelo de pequeña señal es el mismo para un tipo de transistor y su complementario.

En el caso del MOSFET, si el terminal de sustrato se encuentra a un potencial diferente al de source, un aumento en la tensión V_{SB} , causará un aumento en el valor de V_{Th} y como consecuencia una disminución en el valor de I_D . Este efecto se pondrá en evidencia en el modelo de pequeña señal mediante un generador controlado extra $g_{mb} \cdot v_{sb}$, como el indicado en la Fig. A.32.

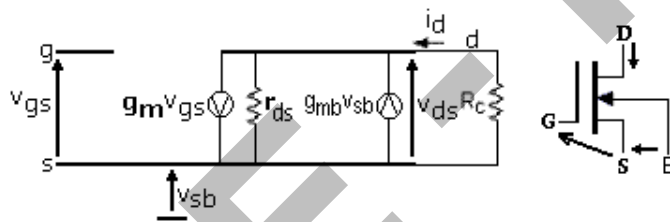


Fig. A.32

El circuito de pequeña señal analizado corresponde a un modelo de bajas frecuencias. El modelo para altas frecuencia contempla los efectos reactivos parásitos del dispositivo. Como mínimo se pueden considerar una capacitancia de entrada C_i y una de realimentación o transferencia C_f .

C_i	TBJ C_π	Capacitancia de difusión de la juntura BE en directa
	JFET C_{gs}	Capacitancia de juntura de la juntura GS en inversa
	MOSFET C_{gs}	Capacitancia C_{ox} del sistema MOS (metal-óxido-canal), que describe su funcionamiento y es del orden de algunos pF
C_f	TBJ C_μ	Capacitancia de juntura de la juntura BC en inversa
	JFET C_{gd}	Capacitancia de juntura de la juntura GD en inversa. Al ser la zona desierta más ancha del lado del drain, $C_{gs} > C_{gd}$, (ambas del orden de los pF)

MOSFET C_{gd}	Capacitancia constructiva debida al solapamiento de las zonas N+ del drain y el óxido C_{gdov} , (difusión lateral de la zona N+ más contaminada que el sustrato P). También existe del lado del source, pero su valor es despreciable frente a C_{ox} . Sin solapamiento, no hay otro efecto capacitivo ya que, con canal estrangulado, no hay carga acumulada cerca del drain ⁽¹⁾
-----------------	--

Todas las capacitancias son dato del fabricante, salvo C_π que se calcula a partir de la *frecuencia de transición* f_T .

La expresión para obtener C_π resulta de *calcular la respuesta en frecuencia de* $\beta = i_c / i_b \mid_{v_{ce}=0}$.

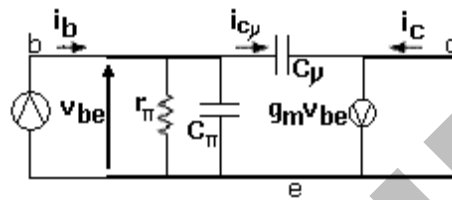


Fig. A.33

Para un amplio rango de frecuencias se cumple que $i_{c\mu} \ll g_m v_{be} \Rightarrow i_c \cong g_m v_{be}$.

Por otro lado, $i_b = v_{be} [s(C_\pi + C_\mu) + 1/r_\pi]$, entonces:

$$\beta \cong g_m r_\pi / [s(C_\pi + C_\mu)r_\pi + 1]$$

Y como en general $C_\pi \gg C_\mu$, para $s = j\omega$, queda:

$$\beta \cong \beta_{(0)} / (j\omega.C_\pi r_\pi + 1)$$

La *frecuencia de corte de -3dB* del $|\beta|$ será: $f_{-3dB} = f_\beta \cong 1 / 2\pi C_\pi r_\pi$

Y la *frecuencia para la cual* $|\beta| = 1$ será: $f_T \cong g_m / 2\pi C_\pi$.

Reemplazando la expresión de la capacitancia de difusión de un diodo, queda:

$$f_T \cong 1 / 2\pi T_{trB}$$

Por lo tanto, al igual que el *tiempo de tránsito en la base* T_{trB} , f_T determina el límite de validez del modelo de pequeña señal utilizado (tal como se explicó para el diodo). Este valor de frecuencia está en el orden de los cientos de MHz.

También podría definirse una f_T para los FET, considerando un tiempo de tránsito a través del canal, aunque en general el fabricante especifica valores de las capacitancias para distintos valores de continua de reposo.

(1) En el MOSFET podrían tenerse en cuenta además, las capacitancias de juntura de source y drain con el sustrato, C_{gb} y C_{db} , respectivamente.

LACE - FLUBA