

Modulación en amplitud

Alfonso Cuesta Hernández

17 de abril de 2001

0.1 Portadora suprimida.

Podemos escribir la expresión de una onda senoidal de la siguiente manera

$$\phi(t) = a(t) \cos [\omega_c t + \gamma(t)] \quad (1)$$

Se supondrá que $a(t)$ y $\gamma(t)$ varían lentamente comparados con $(\omega_c t)$. El término $a(t)$ se llama envolvente de la señal y el término ω_c , frecuencia portadora; $\gamma(t)$ es la modulación en fase de $\phi(t)$

En la modulación en amplitud, el término de fase $\gamma(t)$ de la ecuación (1) es constante y la envolvente se hace proporcional a una señal $f(t)$ dada. Haciendo la constante de proporcionalidad igual a uno se tiene

$$\phi(t) = f(t) \cos \omega_c t \quad (2)$$

El término $\cos \omega_c t$ se denomina portadora y el término $f(t)$ señal moduladora. La señal resultante $\phi(t)$ se denomina señal modulada.

Aplicando la propiedad de modulación de la transformada de Fourier a la ecuación (2) la densidad espectral de $\phi(t)$ es

$$\Phi(\omega) = \frac{1}{2} F(\omega + \omega_c) + \frac{1}{2} F(\omega - \omega_c) \quad (3)$$

Por lo tanto, la modulación en amplitud traslada el espectro de frecuencia de la señal en $\pm \omega_c$ rad/s pero deja inalterada su forma. Este tipo de modulación de amplitud se llama *portadora suprimida* porque la densidad espectral de $\phi(t)$ no presenta una portadora identificable, aunque el espectro se centre en ω_c .

Si observamos el espectro de $\phi(t)$ podemos notar que todo el contenido de frecuencia, tanto negativo como positivo, aparece como frecuencia positiva; por lo que el ancho de banda de la señal $f(t)$ se duplica. Del mismo modo, se pueden reconocer dos zonas muy marcadas tomando como referencia a ω_c : la *banda lateral superior* por encima de esta frecuencia y la *banda lateral inferior* por debajo de ella.

De lo anterior se concluye que la modulación en amplitud con portadora suprimida es un medio conveniente de observar el espectro de frecuencias completo de una señal $f(t)$. Todo lo que hay que hacer es trasladar la señal por medio de una frecuencia portadora ω_c , mayor que las cotas espectrales de la señal. Este principio se utiliza ampliamente en el análisis espectral.

El nombre de *doble banda lateral con portadora suprimida (DSB-SC)* proviene del hecho de que existan dos bandas laterales y que no aparezca ningún rastro de la portadora en el espectro de la modulada.

La recuperación de la señal modulada requiere de otro desplazamiento en la frecuencia. A este proceso de traslación se le llama demodulación o detección.

Al utilizar la propiedad de modulación de la transformada de Fourier probaremos su utilidad en la demodulación. En este caso $\phi(t) = f(t) \cos \omega_c t$, por lo tanto

$$\begin{aligned} \phi(t) \cos \omega_c t &= f(t) \cos^2 \omega_c t \\ &= \frac{1}{2} f(t) + \frac{1}{2} f(t) \cos 2\omega_c t \end{aligned} \quad (4)$$

Tomando la transformada de Fourier de ambos miembros se obtiene

$$\mathcal{F}\{\phi(t) \cos \omega_c t\} = \frac{1}{2} F(\omega) + \frac{1}{4} F(\omega + 2\omega_c) + \frac{1}{4} F(\omega - 2\omega_c) \quad (5)$$

El espectro que nos muestra la ecuación (5), nos indica que debemos utilizar un filtro pasa bajas para separar los términos de doble frecuencia de los componentes espectrales originales. Obsérvese también, que se requiere que $|W| < \omega_c$ donde W es la cota de frecuencias para $f(t)$

0.2 Gran portadora (AM).

Existen diversos problemas en la utilización a nivel práctico de la modulación con portadora suprimida. Además de lo complejo de los circuitos existen problemas debido a las altas exigencias de estos sistemas. Por esto es que se debe hallar alguna alternativa a estos problemas, sobre todo al último, aun a expensas de la eficiencia.

Con este objeto, supóngase que se incorpora la información de la portadora como parte de la onda transmitida y en el mismo ancho espectral de la señal deseada. De hecho, es conveniente hacer que la amplitud de este término portador sea mayor que cualquier otra porción de la densidad espectral de la señal. Pero, si se agrega esta portadora, la respuesta de baja frecuencia del sistema se deteriora; aunque en el caso del audio no es realmente necesaria una buena respuesta a frecuencias bajas. Para diferenciar este caso del anterior, se designará como *gran portadora de doble banda lateral (DSB-LC)*. Como este método se utiliza en las estaciones de radio se le conoce también como *modulación en amplitud*.

La onda modulada de una señal DSB-LC puede escribirse matemáticamente agregando simplemente el término portador, $A \cos \omega_c t$ a una señal DSB-SC

$$\phi_{AM}(t) = f(t) \cos \omega_c t + A \cos \omega_c t \quad (6)$$

La densidad espectral de $\phi_{AM}(t)$ es

$$\begin{aligned} \Phi_{AM}(\omega) &= \frac{1}{2}F(\omega + \omega_c) + \frac{1}{2}F(\omega - \omega_c) \\ &\quad + \pi A \delta(\omega + \omega_c) + \pi A \delta(\omega - \omega_c) \end{aligned} \quad (7)$$

Esta densidad es la misma que la de la señal DSB-SC con la adición de impulsos en $\pm \omega_c$.

La señal de amplitud modulada $\phi_{AM}(t)$ descrita por la ecuación (6) puede describirse de la forma.

$$\phi_{AM}(t) = [A + f(t)] \cos \omega_c t \quad (8)$$

De esta forma podemos identificar a la portadora separada de la señal moduladora $[A + f(t)]$. De la misma manera, si A es lo suficientemente grande, la magnitud de la envolvente será proporcional a $f(t)$. La demodulación, en este caso, se reduce a la detección de la envolvente de una senoide, que no depende de la fase exacta o la frecuencia de ésta. En el caso contrario, la envolvente de $\phi_{AM}(t)$ puede no ser proporcional a $f(t)$. Esta condición se expresa de manera más formal como sigue

$$A \geq |\min\{f(t)\}| \quad (9)$$

Si no se satisface dicha condición la señal no puede ser recuperada por simple detección de la envolvente; sin embargo, la detección síncrona demodulará satisfactoriamente a la señal.

Como las magnitudes relativas de la banda lateral y la porción portadora de la señal son variables, se define un factor de escala adimensional, m , para controlar la razón entre las bandas laterales y la portadora:

$$m = \frac{\text{amplitud pico DSB-SC}}{\text{amplitud pico de la portadora}} \quad (10)$$

de forma que

$$\begin{aligned} \phi_{AM}(t) &= A \cos \omega_c t + mA \cos \omega_m t \cos \omega_c t \\ &= A(1 + m \cos \omega_m t) \cos \omega_c t \end{aligned} \quad (11)$$

donde el término $\cos \omega_m t$ corresponde a un tono senoidal de prueba.

Los máximos de la envolvente son $(1 + m)A$ y los mínimos para $m \leq 1$ son $(1 - m)A$. En el caso de que $m = 0$, la envolvente se reduce a la constante A , tal como debería ocurrir.

A menudo es conveniente definir un porcentaje de modulación para una señal DSB-LC con una modulación senoidal como

$$\begin{aligned} \% \text{mod} &= \frac{(\text{máx. magnitud}) - (\text{mín. magnitud})}{(\text{máx. magnitud}) + (\text{mín. magnitud})} \times 100\% \\ &= \frac{(1 + m)A - (1 - m)A}{(1 + m)A + (1 - m)A} \times 100\% = m \times 100\% \end{aligned} \quad (12)$$

El parámetro m que controla las proporciones relativas entre la banda lateral y la portadora se llama índice de modulación de la onda de AM. De la ecuación (11) puede verse fácilmente que para que ocurra una detección de la envolvente sin fuerte distorsión, hace falta que $m \leq 1$. Si $m > 1$ se dice que la onda está sobremodulada.

0.2.1 Potencia de la portadora y la banda lateral en AM

En las señales de AM, el término portador no contiene información alguna, por lo tanto, la potencia invertida en la portadora se desperdicia para cualquier transferencia de información.

Una onda señal de AM, en general, se puede escribir como en la ecuación (6)

$$\phi_{AM}(t) = f(t) \cos \omega_c t + A \cos \omega_c t$$

La potencia media total para una carga de un ohm sería

$$\overline{\phi_{AM}^2(t)} = A^2 \overline{\cos^2 \omega_c t} + \overline{f^2(t) \cos^2 \omega_c t} + 2A \overline{f(t) \cos^2 \omega_c t} \quad (13)$$

Se supondrá que $f(t)$ varía lentamente con relación a $\cos \omega_c t$. Si además se supone que el valor medio de $f(t)$ es cero, como usualmente sucede, el último término de la ecuación (13) es cero y queda

$$\begin{aligned} \overline{\phi_{AM}^2(t)} &= A^2 \overline{\cos^2 \omega_c t} + \overline{f^2(t) \cos^2 \omega_c t} \\ &= A^2/2 + \overline{f^2(t)}/2 \end{aligned} \quad (14)$$

Entonces, la potencia total P_t puede expresarse como la suma de una potencia portadora P_c y una de banda lateral P_s :

$$P_t = \frac{1}{2}A^2 + \frac{1}{2}\overline{f^2(t)} = P_c + P_s \quad (15)$$

La fracción de potencia total contenida en las bandas laterales, μ , está dada por

$$\mu = \frac{P_s}{P_t} = \frac{\overline{f^2(t)}}{A^2 + \overline{f^2(t)}} \quad (16)$$

Volviendo al caso en el que $f(t)$ es un tono simple se tiene que

$$\overline{\phi_{AM}^2(t)} = \frac{1}{2}A^2 + \frac{1}{4}m^2A^2 \quad (17)$$

y

$$\mu = \frac{m^2}{2 + m^2} \quad (18)$$

Como $m \leq 1$, se ve en la ecuación (18) que la eficiencia en la transmisión de un sistema de AM no supera el 33%. Por el contrario, la eficiencia en un sistema de transmisión DSB-SC es de 100%.

0.3 Modulación de simple banda lateral (SSB)

En las secciones anteriores se vio que la modulación de doble banda lateral duplica el ancho de banda de la señal. Ésta, puede ser una desventaja en el caso en que alguna de las bandas esté saturada en frecuencias.

Como recordamos, la densidad espectral de cualquier función real $f(t)$ cumple la siguiente condición de simetría

$$F(-\omega) = F^*(\omega) \quad (19)$$

Después de multiplicarse por una senoide de frecuencia ω_c rad/s, la densidad espectral se traslada de frecuencia y se centra en ω_c y la otra mitad se centra en $-\omega_c$.

Del gráfico de densidades espectrales es fácil observar que lo que era originalmente el contenido de frecuencias positivo de la señal $F(\omega)$ se convierte en la banda lateral superior para $\omega > 0$ y la inferior para $\omega < 0$. De la misma manera, lo que era el contenido negativo de frecuencias se convierte en la banda lateral inferior para $\omega > 0$ y en la superior para $\omega < 0$. Cada par de bandas laterales contiene la información completa sobre la señal original.

Ahora, ¿por qué no transmitir el par de bandas laterales, superior e inferior, dado que cada par contiene toda la información acerca de la señal?. Se advierte que la ecuación (19) se cumple durante todo el proceso; además la señal original podrá recuperarse de nuevo a partir del par superior o inferior de bandas laterales por una traslación adecuada de frecuencias. A éste tipo de modulación se le conoce como *simple banda lateral* o *SSB*, por sus siglas en inglés. La eficiencia de la modulación SSB radica en que no se requiere más ancho de banda que el de la señal original y sólo la mitad de la señal DSB correspondiente.

0.3.1 Generación de señales SSB

Conceptualmente, la forma más sencilla de generar señales SSB es generando, primero, una señal DSB y luego aplicar un filtro para suprimir la banda no deseada. En la práctica esta operación no es tan sencilla, debido a los estrictos requerimientos en el corte del filtro. Dado que es imposible crear un filtro ideal, se deben dar algunas concesiones. Primero, si la señal no tiene componentes importantes a baja frecuencia –como es el caso de las señales de audio–, se puede reducir la necesidad de una pendiente muy pronunciada en la región de corte del filtro. Segundo, en la práctica, es más sencillo crear filtros de banda lateral a las frecuencias de los componentes de las bandas individuales que a la frecuencia de transmisión. Puede usarse el heterodinizado para trasladar el espectro cuanto se desee, después del filtrado. Aun con estas previsiones, el diseño de filtros de banda lateral no es sencillo.

Es posible generar señales SSB por una adecuada puesta en fase de las señales. Como este método provee ciertos elementos sobre la naturaleza de las señales SSB, se examinará con algún detenimiento.

Considérese una señal exponencial compleja de frecuencia simple. Esta señal tiene una densidad espectral unilateral y solamente tomando la parte real se convierte en bilateral. Específicamente, sea la señal moduladora $f(t) = e^{j\omega_m t}$ y la portadora $e^{j\omega_c t}$. Multiplicando ambas, se obtiene $f(t)e^{j\omega_c t} = e^{j\omega_m t}e^{j\omega_c t}$, desplazando la densidad espectral de acuerdo a la propiedad de traslación de frecuencia de la transformada de Fourier. Tomando ahora la parte real

$$\begin{aligned}\Re\{e^{j\omega_m t}e^{j\omega_c t}\} &= \Re\{e^{j\omega_m t}\}\Re\{e^{j\omega_c t}\} - \Im\{e^{j\omega_m t}\}\Im\{e^{j\omega_c t}\} \\ &= \cos \omega_m t \cos \omega_c t - \sin \omega_m t \sin \omega_c t\end{aligned}$$

El procedimiento de tomar la parte real restaura la característica bilateral a la densidad espectral, por lo que este resultado puede obtenerse usando señales reales. Como esto representa la banda lateral superior, puede escribirse

$$\phi_{SSB_+} = \cos \omega_m t \cos \omega_c t - \sin \omega_m t \sin \omega_c t \quad (20)$$

Aunque este resultado proviene del caso senoidal, son válidos para un caso más general, por lo que puede escribirse

$$\phi_{SSB_{\mp}} = f(t) \cos \omega_c t \pm \hat{f}(t) \sin \omega_c t \quad (21)$$

donde $\hat{f}(t)$ es la señal obtenida al correr la fase de $f(t)$ en 90° en cada frecuencia.

La señal descrita por la ecuación 20 puede producirse por moduladores balanceados, siempre que la señal moduladora como la portadora se desplacen de fase en 90° para formar el segundo término de dicha ecuación. Este método se emplea en la generación de baja frecuencia y digital de señales SSB.

0.4 Modulación de banda lateral vestigial (VSB)

La generación de señales SSB puede resultar difícil cuando la banda de la señal moduladora es ancha o no se pueden despreciar sus componentes de baja frecuencia. Para conservar el espacio espectral se puede hacer un compromiso entre DSB y SSB

en lo que se conoce como *banda lateral vestigial* o *VSB* por sus siglas en inglés. En ésta, sólo se transmite una porción de una banda lateral, de forma tal que el proceso de demodulación reproduce la señal original. La supresión parcial de una banda lateral reduce el ancho de banda con respecto a DSB, pero no alcanza la eficiencia de espectro de SSB. Si se transmite además una gran portadora, la señal deseada puede recobrase con un detector de envolvente. Si éste no es el caso, se puede recobrar usando un detector síncrono o el método de portadora inyectada.

La operación de filtrado puede representarse por un filtro $H_V(\omega)$ que pasa algo de la banda lateral inferior (o superior) y la mayor parte de la superior (o inferior). La densidad espectral de la salida de la banda lateral vestigial resultante es

$$\Phi_{VSB} = \left[\frac{1}{2}F(\omega - \omega_c) + \frac{1}{2}F(\omega + \omega_c) \right] H_V(\omega) \quad (22)$$

Se desea determinar los requisitos del filtro para que el detector síncrono restituya $f(t)$ sin distorsión.

La salida del detector síncrono es

$$e_0(t) = [\phi_{VSB}(t) \cos \omega_c t]_{LP}$$

en el dominio de Fourier esto es

$$E_0(\omega) = \frac{1}{4}F(\omega)H_V(\omega + \omega_c) + \frac{1}{4}F(\omega)H_V(\omega - \omega_c) \quad (23)$$

Para una reproducción confiable se requiere que

$$[H_V(\omega - \omega_c) + H_V(\omega + \omega_c)]_{LP} = \text{constante} \quad |\omega| < \omega_m \quad (24)$$

Nótese que, al menos en lo que respecta a la magnitud, la ecuación (24) se satisface si $|H_V(\omega)|$ es asimétrica con relación con la frecuencia portadora ω_c . Por esto es que la constante de la ecuación (24) sea $2H_V(\omega_c)$. Bajo estas condiciones, la ecuación se convierte en

$$[H_V(\omega - \omega_c) - H_V(\omega_c)] = -[H_V(\omega + \omega_c) - H_V(\omega_c)] \quad (25)$$

con la requerida asimetría con respecto a la frecuencia portadora. Por tanto, el necesario filtrado de una señal VSB es una simetría complementaria del filtro con respecto a la frecuencia portadora. Las relaciones de fase necesarias no son tan fáciles de satisfacer y la VSB se utiliza generalmente en los casos en que la respuestada correcta de fase no es de importancia primordial. Los mismos principios utilizados hasta aquí son válidos también cuando existe una portadora grande y se utiliza un detector de envolvente.

0.5 Modulación por multiplexión de cuadratura.

Como sabemos las señales DSB ocupan el doble del espectro de la señal del mensaje. Este problema se puede resolver transmitiendo dos señales DSB utilizando dos portadoras de la misma frecuencia pero desfasadas en cuadratura de fase.

Si las dos señales a transmitir son $m_1(t)$ y $m_2(t)$ entonces la señal QAM resultante es $\varphi_{QAM}(t)$, la suma de las dos señales DSB es:

$$\varphi_{\text{QAM}}(t) = m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t$$

De esta forma dos señales de mensaje pueden ser separadas en el receptor utilizando detección síncrona con dos portadoras locales en cuadratura de fase. Esto puede mostrarse considerando la salida del multiplicador $x_1(t)$ de la parte rama superior de la figura.

$$\begin{aligned} x_1(t) &= 2\varphi_{\text{QAM}}(t) \cos \omega_c t = 2[m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t] \cos \omega_c t \\ &= m_1(t) + m_1(t) \cos 2\omega_c t + m_2(t) \sin 2\omega_c t \end{aligned}$$

Los dos últimos términos son suprimidos por un filtro pasa-bajas. Este esquema se conoce como *modulación de amplitud por cuadratura* o *QAM*, por sus siglas en inglés. De esta manera, dos señales cada una con un ancho de banda de B Hz, pueden ser transmitidas simultáneamente en un ancho de banda $2B$ utilizando señales DSB y multiplexión por cuadratura.

La QAM es un esquema un tanto riguroso. Un pequeño error en la fase o la frecuencia de la portadora en el demodulador no sólo resulta en pérdida o distorsión de las señales sino en interferencia entre ambos canales. Para mostrar esto, sea la portadora en el demodulador $2 \cos(\omega_c t + \theta)$.

$$\begin{aligned} x_1(t) &= 2[m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t] \cos(\omega_c t + \theta) \\ &= m_1(t) \cos \theta + m_1(t) \cos(2\omega_c t + \theta) - m_2(t) \sin \theta + m_2(t) \sin(2\omega_c t + \theta) \end{aligned}$$

El filtro pasabanda suprime las dos señales a frecuencia $2\omega_c$, resultando en la salida $m_1(t) \cos \theta - m_2(t) \sin \theta$. Por lo tanto, a parte de la señal $m_1(t)$ recibimos a la señal $m_2(t)$, ambas multiplicadas por un factor senoidal o cosenoidal, en la rama superior.

La QAM es utilizada en las transmisiones de televisión en color para multiplexar las señales de crominancia, que portan la información acerca del color. En este caso, la sincronización se logra por la inserción periódica de una muestra de la señal portadora.