Capítulo 2

Filtros digitales

Uno de los problemas más habituales a la hora de trabajar con sistemas lineales, tanto en variable continua como en variable discreta, es la necesidad de un procedimiento de trabajo sistemático que permita huir de la necesidad de resolver empleando un procedimiento *ad hoc* cada problema particular. Por ese motivo, lo más habitual es trabajar con sistemas que constan de una estructura predefinida, muy concreta, capaz de adaptarse a cualquier funcionamiento que se pueda requerir, pero que proporcione, por medio de dicha estructura, unas características comunes que puedan aprovecharse para facilitar su manejo. Las estructuras por excelencia para este fin son los denominados *filtros digitales*.

2.1. Definición

Una clase importante de sistemas lineales e invariantes son aquellos en los que la entrada y la salida se relacionan por medio de una *ecuación lineal en diferencias de coeficientes constantes* de la forma

$$\sum_{k=-\infty}^{\infty} a_k y[n-k] = \sum_{k=-\infty}^{\infty} b_k x[n-k] \tag{2.1}$$

La salida en cada instante es una combinación lineal de las entradas y de las demás salidas. Este tipo de estructura, con algunas modificaciones necesarias que se indicarán a continuación, se denomina *filtro digital*.

Generalmente el número de coeficientes no nulos (orden) es finito, y en un caso real se tratará de un sistema causal, de forma que la ecuación en diferencias puede reescribirse como

$$\sum_{k=0}^{M} a_k y[n-k] = \sum_{k=0}^{N} b_k x[n-k] \tag{2.2}$$

que es la expresión general más habitual para un filtro digital, dado que constituye una versión realista de (2.1) al representar un sistema causal y con un número finito de coeficientes. En este caso, la salida en un determinado instante (suele decirse que *en el instante*

actual) es una combinación lineal de la entrada en dicho instante, las entradas en los N instantes anteriores y las últimas M salidas:

$$a_0 y[n] = \sum_{k=0}^{N} b_k x[n-k] - \sum_{k=1}^{M} a_k y[n-k] \tag{2.3}$$

La ecuación en diferencias, tal y como está planteada, es recurrente y por tanto no está completamente definida la salida del sistema. Siempre se necesitarían valores anteriores de entrada y salida, hasta $n = -\infty$, cuyo conocimiento en un sistema real es imposible. Por ello, para caracterizarlo completamente hacen falta unas *condiciones iniciales*, normalmente expresadas como el valor de la salida para ciertos valores de n de forma que pueda construirse la recurrencia a partir de ellos.

Por último, para evitar ciertas ambigüedades, y dado que al multiplicar ambos lados de la igualdad por una misma constante dicha igualdad se mantiene, es posible normalizar la ecuación en diferencias por cualquier valor sin alterar el sistema que representa. Se suele seleccionar escalar la ecuación de tal manera que se tenga siempre que $a_0 = 1$, de forma que pueda despejarse de una manera más cómoda el valor de la salida. En lo sucesivo, todas las representaciones asumirán dicho escalado.

En el caso particular en el que $M = 0$, la ecuación en diferencias es

$$y[n] = \sum_{k=0}^{N} b_k x[n-k] \tag{2.4}$$

En este caso, la salida actual es una suma ponderada de la entrada actual y las N anteriores, sin dependencia con el valor de salidas anteriores. Como no hay recurrencia, conocida la entrada se conoce la salida sin necesidad de acudir a condiciones iniciales. Además, se puede obtener la respuesta al impuso fácilmente haciendo $x[n] = \delta[n]$ y obteniendo la salida, que se puede comprobar que es finita de duración N+1:

$$h[n] = \sum_{k=0}^{N} b_k h[n-k] \tag{2.5}$$

Por este motivo, este tipo de filtro sin dependencia de las salidas anteriores se denomina filtro de Respuesta al Impulso Finita, FIR. En caso contrario, el general, filtro de Respuesta al Impulso Infinita, IIR[1].

Observando la representación de la respuesta de un filtro FIR causal con un número finito de coeficientes, dada por (2.4), y la respuesta del sistema obtenida mediante la convolución de la entrada con la respuesta al impulso del sistema:

$$y[n] = \sum_{k=0}^{N} h[k] x[n-k] \tag{2.6}$$

[1]En un filtro FIR está garantizada la longitud finita de la respuesta al impulso. En un IIR potencialmente puede ser infinita, dado que al terminar las entradas (la única muestra que representa el impulso) siguen excitando el sistema las salidas anteriores. Aún así, es posible que la salida se amortigüe y vaya desapareciendo hasta hacerse nula tras un número finito de muestras, pero la estructura del filtro no lo garantiza, potencialmente puede ser una respuesta infinita, y por eso se denomina IIR sea realmente infinita o no.

Comparando ambas expresiones se puede comprobar que se cumple que en un filtro FIR (y sólo en un FIR, no en un IIR) los coeficientes de la ecuación en diferencias coinciden con las muestras de la respuesta al impuso del sistema:

$$h[n] = b_n, \forall n \tag{2.7}$$

Si se consideran las representaciones vectoriales, $\mathbf{h} = [h[0], h[1], \ldots, h[N]]^T$, y se define el vector de *run-time* como

$$\mathbf{x}[n] = \mathbf{x}_n = \begin{bmatrix} x[n] \\ x[n-1] \\ x[n-2] \\ \vdots \\ x[n-N] \end{bmatrix} \tag{2.8}$$

que consta de las últimas N muestras de la entrada, siendo la primera de ellas la más reciente, entonces la salida del filtro FIR es

$$y[n] = \mathbf{h}^T \mathbf{x}_n \tag{2.9}$$

Si $h[n]$ toma valores reales, la salida de un filtro FIR coincide con el producto escalar entre su respuesta al impulso y el vector de run-time de la entrada.

También es posible expresar vectorialmente el caso de un filtro IIR, para lo que se definen los vectores

$$\mathbf{y}[n] = \mathbf{y}_n = \begin{bmatrix} y[n] \\ y[n-1] \\ y[n-2] \\ \vdots \\ y[n-M] \end{bmatrix} \tag{2.10}$$

$$\mathbf{b} = [b_0, b_1, \ldots, b_N]^T \tag{2.11}$$

$$\mathbf{a} = [a_0, a_1, \ldots, a_M]^T \tag{2.12}$$

quedando la ecuación en diferencias como

$$\mathbf{a}^T \mathbf{y}_n = \mathbf{b}^T \mathbf{x}_n \tag{2.13}$$

Se ha indicado que para que la definición de un filtro IIR sea completa son necesarias unas condiciones iniciales. Supongamos que la entrada al sistema es nula hasta el instante n_0 y que tenemos unas condiciones iniciales dadas por $y[n_0 - 1] = \alpha, \alpha \neq 0$. En un sistema lineal, invariante y causal, si la entrada es nula la salida también debe serlo, pero la condición inicial lo impide ya que está produciendo una salida antes de que se produzca entrada no nula. Por ello, el sistema no puede ser lineal, invariante y causal al mismo tiempo. Se puede establecer que un sistema IIR definido por unas ecuaciones en diferencias de coeficientes constantes es lineal, invariante y causal si, y sólo si, presenta condiciones iniciales nulas (también llamadas condiciones de *reposo inicial*): si la entrada es nula hasta un determinado instante, la salida también ha de serlo. La propia estructura elegida para la ecuación en diferencias garantiza que al tener condiciones de reposo inicial se aseguran la linealidad, la invarianza y la causalidad.

2.2. Representación de sistemas en el plano z

Hasta ahora hemos caracterizado los sistemas lineales e invariantes por medio de dos respuestas equivalentes: la respuesta al impulso, $h[n]$, y la respuesta en frecuencia, $H(\omega) = TF\{h[n]\}$. En muchos casos es interesante emplear una tercera representación que además va a facilitar el trabajo sistemático con sistemas que responden a la definición de filtros digitales: la función de sistema, $H(z)$. Ésta se define como la Transformada z de la respuesta al impulso:

$$H(z) = TZ\{h[n]\} = \sum_{n=-\infty}^{\infty} h[z]z^{-n}, \qquad z \in \mathbb{C} \tag{2.14}$$

Como en todos los casos de transformación, la Transformada z consiste en la representación de la señal en términos de unas funciones que actúan como base[2], $z^n, n \in \mathbb{Z}, z \in \mathbb{C}$, siendo la función transformada el conjunto de coeficientes de la citada representación o expansión, y que se obtienen como el producto escalar entre la señal y las funciones base (2.14). En este caso, la expansión, que se denotaría como transformada inversa, viene dada por

$$x[n] = \frac{1}{2\pi j} \oint_C X(z)z^{n-1}dz \tag{2.15}$$

donde C es un círculo cerrado definido en el plano z (Re$\{z\}$,Im$\{z\}$) que envuelve el origen y la región de convergencia (ROC), conteniendo todos los polos de $X(z)$. Algunos de estos términos requieren una breve explicación. Se denomina *plano z* al plano definido por las partes real e imaginaria de la variable z y que será de utilidad para representar funciones de sistema (figura 2.1). Al ser z una variable compleja, podemos encontrar que la función de sistema (o cualquier función definida en el plazo z) puede no estar definida para todos los valores de la variable; aquellos en los que sí está definida se denominan *región de convergencia* o ROC por sus siglas en inglés. Aquellos valores de z para los que se anula la función se denominan *ceros* (del sistema, en caso de estar estudiando uno) mientras que aquellos que anulan el denominador de la función de sistema se denominan *polos*, y para ellos no está definida la función, tendiendo su valor a ∞ o $-\infty$ en un entorno arbitrariamente cercano.

Al tratarse de una representación en una variable compleja, se suele hacer en el denominado plano z formado por sus partes real e imaginaria. En esta representación juega un papel importante la denominada *circunferencia unidad*, espacio geométrico definido por los puntos cuya distancia al origen es unitaria, y por tanto están dados por $z = e^{j\omega}$, siendo ω el valor de la fase de dicho punto. Como principal particularidad, los puntos de dicha circunferencia cumplen que

$$H(z = e^{j\omega}) = H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h[n]e^{-j\omega n} = TF\{h[n]\} \tag{2.16}$$

de forma que el valor de la función de sistema evaluada en los puntos que forman la circunferencia unidad coincide con la respuesta en frecuencia. Este resultado es extensible a

[2]El caso de la Transformada z tiene particularidades que hacen que se ponga en duda el comportamiento de las funciones como base, pero son ajenas al alcance de este texto.

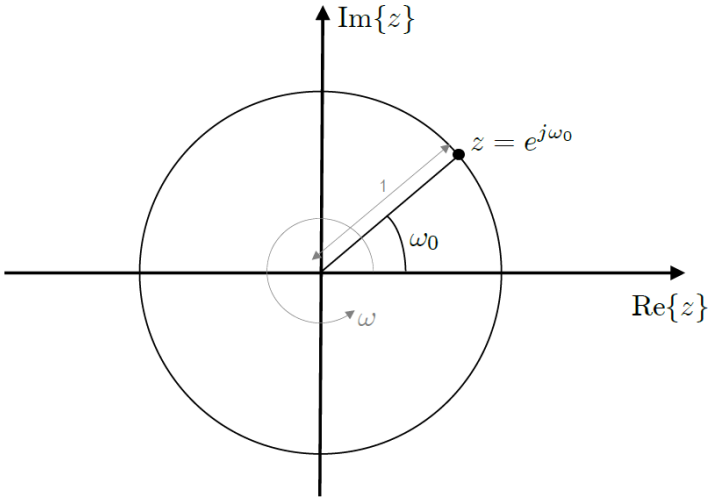

Figura 2.1: Representación del plano z y de la *circunferencia unidad*.

la relación entre la Transformada z y la de Fourier. Aprovechando este resultado se puede asimilar la fase con la pulsación o frecuencia, y se puede comprobar algún aspecto ya estudiado al tratar la Transformada de Fourier: el espectro de una señal, al corresponder con el valor de la Transformada z en un círculo, es periódico con periodo 2π, que corresponde con la fase de una vuelta completa al círculo.

Al igual que la Transformada de Fourier, la Trasnformada z cumple:

- Linealidad
$$TZ\{\alpha h_1[n] + \beta h_2[n]\} = \alpha H_1(z) + \beta H_2(z)$$
- Convolución
$$TZ\{h_1[n] * h_2[n]\} = H_1(z)H_2(z)$$
- Desplazamiento
$$TZ\{h[n - n_0]\} = z^{-n_0}X(z)$$

Supongamos, sin pérdida de generalidad, un filtro digital dado por

$$y[n] = x[n] + b_1x[n-1] + b_2x[n-2] + a_1y[n-1] + a_2y[n-2]$$

Si estudiamos este sistema, la Transformada dz conduce a

$$Y(z) = X(z) + b_1z^{-1}X(z) + b_2z^{-2}X(z) + a_1z^{-1}Y(z) + a_2z^{-2}Y(z)$$

Agrupando:

$$Y(z)(1 - a_1z^{-1} - a_2z^{-2}) = X(z)(1 + b_1z^{-1} + b_2z^{-2})$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1 + b_1z^{-1} + b_2z^{-2}}{1 - a_1z^{-1} - a_2z^{-2}} \times \frac{z^2}{z^2} = \frac{z^2 + b_1z + b2}{z^2 - a_1 - a_2}$$

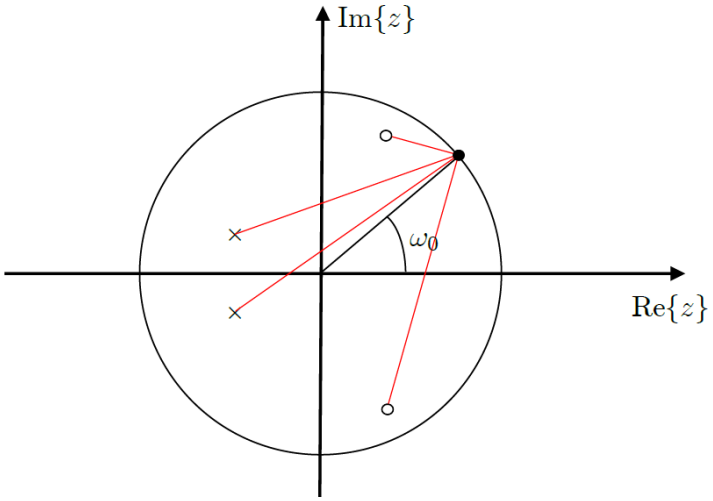

Figura 2.2: Representación de los ceros y polos en el plano z y sus distancias a la *circunferencia unidad*.

donde en el último paso se ha multiplicado por z^2 el numerador y el denominador para que no varíe la expresión final. Si se calculan las raíces de los dos polinomios, numerador y denominador, en ambos casos serán complejos conjugados por tratarse de dos polinomios con coeficientes reales (al ser los de una ecuación en diferencias como la hemos definido). Si denominamos z_c y z_c^* a los ceros (raíces del numerador) y z_p y z_p^* a los polos (raíces del denominador), podemos expresar:

$$H(z) = \frac{(z - z_c)(z - z_c^*)}{(z - z_p)(z - z_p^*)} \tag{2.17}$$

La respuesta en frecuencia se puede obtener sustituyendo $z = e^{j\omega}$ y, dado que es la amplitud de dicha respuesta la que suele ser de interés, considerando el valor absoluto:

$$|H(\omega)| = \frac{|e^{j\omega} - z_c| \cdot |e^{j\omega} - z_c^*|}{|e^{j\omega} - z_p| \cdot |e^{j\omega} - z_p^*)} = \frac{d_1 \cdot d_2}{d_3 \cdot d_4} \tag{2.18}$$

donde d_1, d_2, d_3 y d_4 pueden interpretarse como las distancias entre los polos y los ceros a la circunferencia unidad para un valor de ω dado. Si los ceros y polos son los representados en la figura 2.2 (normalmente se emplea un círculo para representar las posiciones de los ceros y un aspa para representar las de los polos), las distancias a un punto dado por la fase ω_0 (que igualmente representa una pulsación) serán las indicadas en rojo.

Con este diagrama presente, se puede aproximar el valor de la respuesta en frecuencia del sistema. Si se hace coincidir cada pulsación con el mismo valor de fase, pueden evaluarse las cuatro distancias entre el punto representado por dicha fase y los polos y ceros del sistema. Si, por ejemplo, la pulsación estudiada corresponde a un punto de la circunferencia unidad muy cercano a un cero, una de las distancias d_1 o d_2 se hará muy pequeña, reduciendo el valor de la expresión (2.18), y por tanto reduciendo el valor de la respuesta en frecuencia para esa pulsación. Si por el contrario el punto está cerca de un polo, será una de las distancias d_3 o d_4 la que se haga pequeña, aumentando el valor de (2.18) y

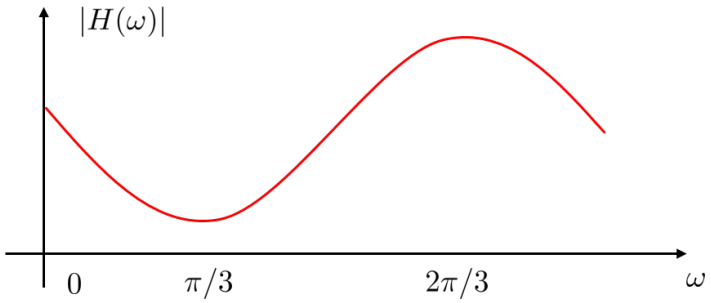

Figura 2.3: Boceto de la respuesta en frecuencia del sistema ejemplo con un cero situado en el entorno de $\omega = \pi/3$ y un polo en el entorno de $\omega = 2\pi/3$.

de la respuesta en frecuencia para la pulsación bajo estudio. De esta manera, la posición relativa de polos y ceros respecto a la circunferencia unidad determina la respuesta en frecuencia de un sistema, y puede ser utilizada directamente para diseñarlo. La figura 2.3 representa un boceto de la respuesta en frecuencia correspondiente al diagrama de polos y ceros representado en la figura 2.2.

También puede extenderse aquí que, al tener polos y ceros simétricos debido a que los coeficientes de la ecuación en diferencias son reales, la respuesta en frecuencia resulta ser simétrica respecto a $\omega = 0$, puesto que la Transformada z también resulta serlo respecto al eje $\mathrm{Re}\{z\}$.

Ejemplo 2.1. *Diseño de un filtro paso bajo con una pulsación de corte dada.*

Se desea diseñar un filtro paso bajo con una pulsación de corte dada por $\omega_c = pi/3$. Para ello, el sistema en que se va a implementar el filtro, la capacidad de computación disponible, etc, aconsejan emplear un máximo de 20 coeficientes (más coeficientes implican más necesidad de recursos computacionales y de almacenamiento, por ejemplo).

El uso de 20 coeficientes quiere decir que se dispondrá de un total de 20 polos y/o ceros cuya ubicación en el plazo z es tarea del diseñador. Por ejemplo, pueden colocarse 5 polos inmediatamente dentro de la circunferencia unidad y repartidos entre $\omega = 0$ y $\omega = \omega_c$, y otros tantos repartidos uniformemente entre ω_c y $\omega = \pi$, igualmente cerca de la circunferencia unidad. Al estar obligados a emplear coeficientes reales en el filtro, los restantes 10 valores estarán colocados de forma simétrica en la parte inferior del plano z.

De la estrategia para colocar más o menos elementos como polos o ceros, o para distribuirlos de manera uniforme o no dentro de cada rango, dependen factores como el rizado del filtro, la pendiente de la zona de transición entre la banda de paso y la de rechazo, etc. Evidentemente, a mayor número de coeficientes mejor respuesta puede obtenerse, reduciendo todos estos efectos no deseados.

Una vez establecidos los polos y ceros sobre el plazo z, su posición se introduce en la ecuación (2.17), adaptada al número de polos y ceros que se consideren, y de ella se puede obtener la ecuación en diferencias que representa al sistema diseñado.

2.2.1. Sistema inverso

Dado un sistema definido por su respuesta al impulso $h[n]$ y su función de sistema $H(z)$, se denomina *sistema inverso* a aquel que cumple que al recibir como entrada la salida del

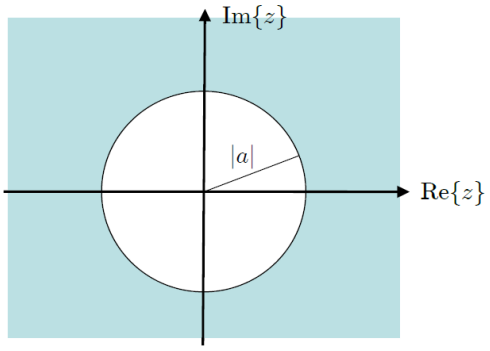

Figura 2.4: Región de convergencia $|z| > |a|$.

sistema original, proporciona como salida la entrada de éste: si $y[n] = x[n] * h[n]$, el inverso $h_i[n]$ (a veces $h^{-1}[n]$) cumple que $x[n] = y[n] * h_i[n]$. Esta expresión es equivalente a que se cumpla $h[n] * h_i[n] = \delta[n]$, o en el dominio transformado que $H(z)H_i(z) = 1$. Por tanto, se cumple que:

$$H_i(z) = \frac{1}{H(z)} = H^{-1}(z) \tag{2.19}$$

Este resultado indica que el sistema inverso tiene como polos los ceros del original o *directo*, y como ceros los polos de éste.

2.2.2. Región de convergencia

Consideremos, a modo de ejemplo, la secuencia $x[n] = a^n\mu[n]$, siendo a una constante real. Queremos evaluar la Transformada z de dicha secuencia, para lo que aplicamos directamente su definición:

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n} = \sum_{n=-\infty}^{\infty} a^n\mu[n]z^{-n} = \sum_{n=0}^{\infty} a^n z^{-n} = \sum_{n=0}^{\infty} (az^{-1})^n$$

Empleando expresiones conocidas[3] esta expresión puede comprobarse que únicamente está definida cuando $|az^{-1}| < 1$, tomando valor indefinido en otro caso. Si asumimos que se cumple esa cota, tenemos

$$X(z) = \frac{1}{1 - az^{-1}} = \frac{z}{z - a} \tag{2.20}$$

pero recordando que únicamente está definido para $|az^{-1}| < 1$, que equivale a $|z| > |a|$. Este conjunto de valores de z para los que está definida esta transformada constituyen la ROC (figura 2.4).

[3]Se emplea la siguiente relación:

$$\sum_{k=a}^{b} \gamma^k = \frac{\gamma^a - \gamma^{b+1}}{1 - \gamma}$$

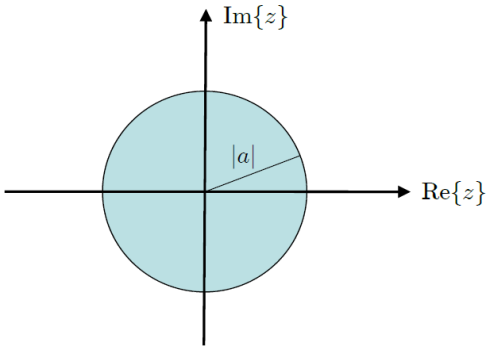

Figura 2.5: Región de convergencia $|z| < |a|$.

Consideremos un segundo ejemplo, con la secuencia $x[n] = -a^n\mu[-n-1]$. AL igual que en el caso anterior, aplicamos la definición de la Transformada z:

$$X(z) = \sum_{n=-\infty}^{\infty} -a^n\mu[-n-1]z^{-n} = \sum_{n'=1}^{\infty} -a^{n'}z^{n'} = -\sum_{n'=1}^{\infty}(a^{-1}z)^{n'}$$

donde se ha aplicado un cambio de variable $n' = -n$. De forma idéntica al caso anterior, considerando que $|a^{-1}z| < 1$ para que esté definida la solución, se puede obtener que

$$X(z) = -\frac{a^{-1}z}{1-a^{-1}z} = -\frac{z}{a-z} = \frac{z}{z-a} \tag{2.21}$$

En este caso, se ha obtenido la misma solución que en el caso anterior, pero definida únicamente para $|a^{-1}z| < 1$, o equivalentemente para $|z| < |a|$ (figura 2.5). Ambas soluciones tienen la misma expresión, pero difieren en la ROC, lo que indica que ésta es estrictamente necesaria para definir completamente la función en el dominio z.

Avanzando un poco más, podemos considerar un nuevo ejemplo donde $x[n] = \left(-\frac{1}{3}\right)^n\mu[n] - \left(\frac{1}{2}\right)^n\mu[-n-1]$. Es inmediato comprobar que esta secuencia se compone de la suma de dos secuencias más sencillas, y por tanto su transformada será igual a la suma de las transformadas de ambas. A su vez, las dos secuencias sumadas encajan con las empleadas en los ejemplos anteriores, particularizando en cada caso la variable a como $a = -1/3$ y $a = 1/2$, con lo que ya hemos calculado sus transformadas. Así, tendremos que

$$X(z) = \frac{z}{z+\frac{1}{3}} + \frac{z}{z-\frac{1}{2}}$$

En el primero de los términos conocemos que la ROC estará definida por la región $|z| > 1/3$, mientras que en el segundo será $|z| < 1/2$. El conjunto estará definido únicamente en donde estén definidos sus dos términos, y por tanto en la región intersección dada por $1/3 < |z| < 1/2$. Esta ROC, mostrada en la figura 2.6, tiene forma de anillo.

Estos tres ejemplos ilustran los tipos de ROC que podemos encontrar en el estudio de filtros digitales: regiones externas a una circunferencia, internas a ella, o anillos delimitados igualmente por circunferencias. Dado que los polos de un sistema no pertenecen en ningún caso a la ROC por definición (son puntos donde no está definida la respuesta del sistema),

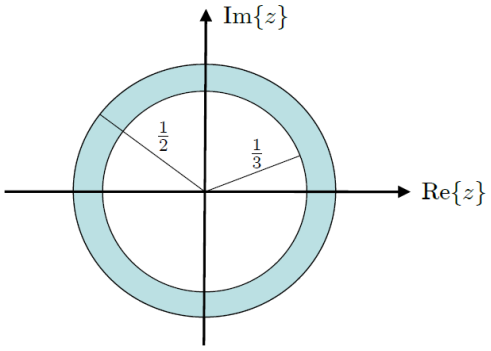

Figura 2.6: Región de convergencia $1/3 < |z| < 1/2$.

normalmente son dichos polos los que delimitan las circunferencias que definen la región de convergencia: hacia afuera del polo más alejado del origen, hacia dentro del más cercano al origen, o el anillo entre dos polos.

Un estudio detallado de la definición de un sistema como filtro digital y su representación en el dominio z conduce a las siguientes conclusiones:

- Si una sistema es lineal, invariante y causal, su ROC siempre responde a la definición $|z| > \max|d_k|$, donde d_k representa a los polos del sistema. Es decir, la región de convergencia siempre será hacia afuera del polo más alejado del origen.

- Un sistema estable, por definición, siempre debe presentar una ROC que contenga la circunferencia unidad: $\|z| = 1 \in$ ROC, dado que en caso contrario su respuesta frecuencial no estaría definida, lo que implica un comportamiento no estable para cualquier señal de entrada.

- Un sistema lineal, invariante, causal y estable, por tanto, debe tener todos sus polos contenidos dentro de la circunferencia unidad ($\max|d_k| < 1$, de forma que su ROC sea hacia afuera de ellos y además contenga a dicha circunferencia.

- Para que un sistema lineal, invariante, causal y estable tenga inverso lineal, invariante, causal y estable, debe cumplir que todos sus polos y ceros estén contenidos dentro de la circunferencia unidad (así los polos del inverso, que son los ceros del directo, estarán contenidos en ella). Los sistemas que cumplen esto se denominan *sistemas de fase mínima.*

2.3. Operación con filtros digitales

Consideremos un filtro digital genérico dado por su ecuación en diferencias

$$\sum_{k=0}^{M} a_k y[n-k] = \sum_{k=0}^{N} b_k x[n-k]$$

Para poder estudiar su función de sistema, simplemente aplicamos la transformada z obteniendo:

$$\sum_{k=0}^{M} a_k z^{-k} Y(z) = \sum_{k=0}^{N} b_k z^{-k} X(z)$$

y de aquí, la función de sistema viene dada por

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{N} b_k z^{-k}}{\sum_{k=0}^{M} a_k z^{-k}} = \frac{b_0}{a_0} \frac{\prod_{k=1}^{N}(a - c_k z^{-1})}{\prod_{k=1}^{M}(1 - d_k z^{-1})} \tag{2.22}$$

donde se han factorizado los polinomios del numerador y del denominador para obtener la expresión explícita de los ceros (c_k) y los polos (d_k). Se puede observar que la dependencia del sistema con la entrada es la que determina los ceros del sistema, mientras que la dependencia con la salida es la que determina lo polos.

Además, este procedimiento ha convertido la ecuación en diferencias en la función de sistema equivalente. Si se aplica sobre ella la Transformada z inversa, podemos obtener la respuesta al impuso del sistema. Inversamente, si se dispone de la respuesta al impulso y se calcula a partir de ella la función de sistema, es posible obtener la ecuación en diferencias equivalente simplemente deshaciendo los pasos que se acaban de seguir.

2.3.1. Transformada z inversa

La Transformada z inversa es, en general, una operación complicada. Sin embargo, la elección sistemática de sistemas que responden a la definición de filtros digitales y que están definidos por ecuaciones en diferencias de estructura fija permite emplear versiones simplificadas de esta operación, adaptadas a esta estructura.

En condiciones de trabajo normales con filtros digitales, podremos resolver las transformaciones inversas empleando alguno de los siguientes procedimientos:

- Por inspección. En muchos casos se trabajará con expresiones ya conocidas que pueden encontrarse en tablas de transformadas frecuentes. En general, muchas de las que se emplean al trabajar con filtros digitales responden exactamente a las expresiones (2.20) y (2.21).
- Por aplicación de la definición de Transformada z. Supongamos un sistema dado por $H(z) = z^2 - z + 5 - 2z^{-1} + z^{-2}$. De acuerdo a la expresión de la Transformada z, los términos no nulos de esta función deben venir dados por $H(z) = h[-2]z^2 + h[-1]z + h[0] + h[1]z^{-1} + h[2]z^{-2}$, de forma que podemos identificar que $h[n] = [1, -1, 5, -2, 1], n = -2, \ldots, 2$.
- Por descomposición en factores simples. La expresión de función de sistema dada por (2.22) puede representarse alternativamente, dependiendo de los valores de M y N según las siguientes expresiones:

  Si $M < N$, se puede expresar:

$$H(z) = \sum_{k=1}^{N} \frac{A_k}{1 - d_k z^{-1}} \tag{2.23}$$

donde $A_k = (1 - d_k z^{-1}) H(z)|_{z=d_k}$.

Si $M \geq N$, la función toma la forma:

$$H(z) = \sum_{r=0}^{M-N} B_r z^{-r} + \sum_{k=1}^{N} \frac{A_k}{1 - d_k z^{-1}} \tag{2.24}$$

Ejemplo 2.2. *Transformada z inversa mediante descomposición.*

Sea el sistema definido por

$$H(z) = \frac{1 + 2z^{-1} + z^{-2}}{1 - \frac{3}{2} z^{-1} + \frac{1}{2} z^{-2}}, \qquad \text{ROC:} |z| > 1$$

En este caso tenemos que $M = N$, con lo que debemos acudir a la expresión (2.24). Para ello, comprobamos que dividiendo el numerador entre el denominador podemos obtener como cociente el valor 2, quedando como resto la expresión $5z^{-1} - 1$. Entonces, se puede expresar:

$$H(z) = 2 + \frac{5z^{-1} - 1}{1 - \frac{3}{2} z^{-1} + \frac{1}{2} z^{-2}}$$

El primer término corresponde con el de (2.24), y el segundo ahora encaja con la expresión descrita en (2.23), con $M < N$, y puede ser representado entonces como:

$$H(z) = 2 + \frac{A_1}{1 - \frac{1}{2} z^{-1}} + \frac{A_2}{1 - z^{-1}} = 2 + H_1(z)$$

donde se han obtenido ya los polos del sistema obteniendo las raíces del polinomio denominador, $z = 1/2$ y $z = 1$. Observando únicamente la parte $H_1(z)$ de la expresión, tenemos:

$$A_1 = H_1(z)(1 - \frac{1}{2} z^{-1})|_{z=1/2} = -9$$

$$A_2 = H_1(z)(1 - z^{-1})|_{z=1} = 8$$

con lo que queda:

$$H(z) = 2 + \frac{-9}{1 - \frac{1}{2} z^{-1}} + \frac{8}{1 - z^{-1}}$$

Todos los términos tienen Transformada z inversa conocida, a falta de establecer la ROC. La conjunta está dada por $|z| > 1$, y debe ser la intersección de las correspondientes a los tres términos en que se ha descompuesto. El primero de los términos está definido para todo z; el segundo puede estar definido para $|z| > 1/2$ o $|z| < 1/2$, con lo que la única posibilidad que conduce a la ROC global con una intersección es que sea $|z| > 1/2$; y el tercer término puede estar definido para $|z| > 1$ o $|z| < 1$, de forma que de nuevo la única posibilidad que conduce a la intersección conocida es $|z| > 1$. Ahora, conocidas las ROC de cada término, se puede aplicar la Transformada z inversa a cada uno obteniendo:

$$h[n] = 2\delta[n] - 9 \left(\frac{1}{2}\right)^n \mu[n] + 8\mu[n]$$

2.4. Estructuras para filtros digitales

De cara a la implementación de un sistema definido conforme a la estructura de un filtro digital, es habitual acudir a representaciones estándar, derivadas en gran medida del uso de la Teoría de Grafos, que proporciona una herramienta muy interesante para la optimización de las estructuras implementadas.

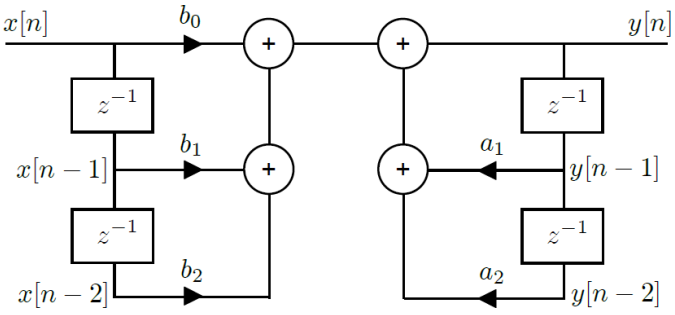

Figura 2.7: Implementación directa del sistema.

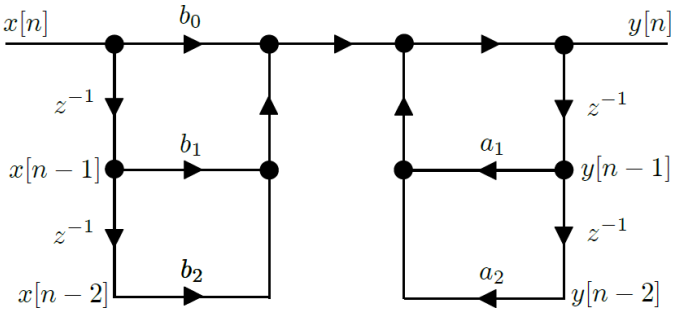

Figura 2.8: Grafo de la implementación directa del sistema.

2.4.1. Formas directas

Sin pérdida de generalidad, partiremos de un sistema especificado por medio de su ecuación en diferencias, dada por

$$y[n] = b_0x[n] + b_1x[n-1] + b_2x[n-2] + a_1y[n-1] + a_2y[n-2]$$

Se trata de un sistema con $N = M = 2$, pero veremos que la generalización a otros valores es inmediata. Las condiciones iniciales se considerarán de reposo inicial para trabajar con sistemas lineales, invariantes y causales, características de interés en la inmensa mayoría de diseños.

Una implementación posible del sistema se corresponde con el diagrama de la figura 2.7, donde los bloques denotados como z^{-1} indican un retardo de una unidad en n, y se indican mediante flechas los multiplicadores por constantes.

Apoyándonos en la Teoría de Grafos, definimos un *nodo* como un punto del esquema que suma todas sus entradas para proporcionar una salida (lo representaremos mediante un círculo) y una *rama* como una de las líneas del esquema, que multiplica por un factor constante o un desplazamiento (se representa como una flecha). De ese modo, el esquema queda como se representa en la figura 2.8. Esta implementación se denomina *forma directa I*.

Si el sistema fuese un filtro FIR, todo el bloque de la derecha se suprimiría (figura 2.9). Si se empleasen diferentes valores de M y N simplemente se modificaría el número de componentes del bloque correspondiente.

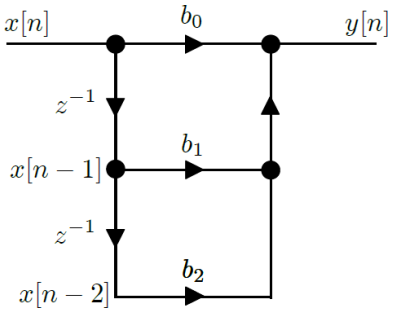

Figura 2.9: Grafo de la implementación directa del sistema, caso FIR.

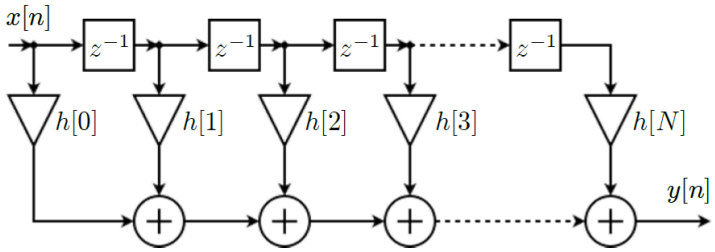

Figura 2.10: Representación habitual de un FIR.

Ejemplo 2.3. *Representación de un filtro FIR.*

En el caso particular de un filtro FIR es muy habitual encontrar representaciones que, siendo idénticas a las presentadas, suelen mostrarse de un modo muy particular, en un formato horizontal como el mostrado en la figura 2.10.

Ello es debido al mismo motivo por el que se emplea la denominación *filtro digital* para referirse a este tipo de sistemas. Observando la figura, es posible establecer un paralelismo entre un sistema como éste y un filtro tradicional analógico, asumiendo que una multiplicación por una constante tiene una relativa equivalencia con la atenuación que origina una resistencia, mientras que un retardo tiene similitud con el efecto de un condensador. La estrucutra de un filtro digital imita, por tanto, el comportamiento de un filtro RC (resistencia-condensador) electrónico analógico. De ahí su denominación.

La figura 2.11 recoge la representación en forma de grafo de esta estructura representada en el formato horizontal habitual.

A partir de la estructura de grafos de la forma directa I, se puede hacer uso de la propiedad conmutativa de la convolución para hacer modificaciones: la estructura realmente consta de dos partes diferenciadas, una para el lado FIR y otra para el lado IIR. Se pueden

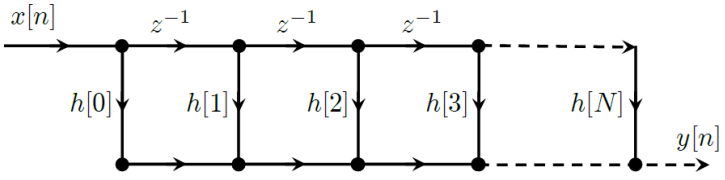

Figura 2.11: Grafo de la representación habitual de un FIR.

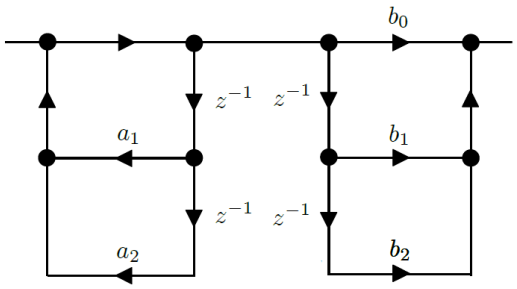

Figura 2.12: Grafo de la implementación directa modificando el orden de sus partes.

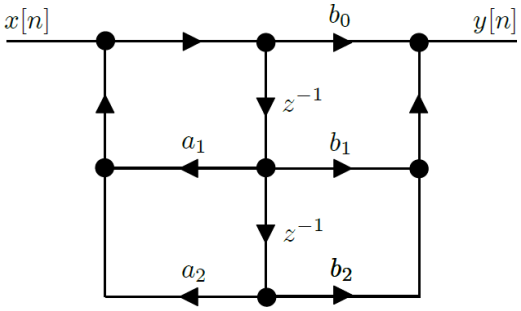

Figura 2.13: Grafo de la implementación en forma directa II.

considerar dos subsistemas, siendo la respuesta global la convolución de sus respuestas. Como la convolución tiene propiedad conmutativa, es posible cambiar el orden de los dos subsistemas poniendo en primer lugar la parte IIR y después la FIR (figura 2.12).

Hecha esta operación, se puede comprobar que las dos ramas centrales llevan exactamente la misma información a través de los dos retardadores y del nodo central. Puesto que ambas llevan el mismo dato y hacen la misma operación, pueden simplificarse con una única rama tal y como muestra la figura 2.13. en la que se denomina *forma directa II*.

Esta estructura es de tipo *canónico*, lo que significa que tiene el menor número posible de elementos (retardos, multiplicadores), siendo por tanto una estructura más económica.

2.4.2. Formas traspuestas

Otra forma de estructurar la implementación de un filtro digital viene dada por las llamadas *formas traspuestas*. Éstas se obtienen a partir de las directas siguiendo un procedimiento muy concreto:

1. Se mantiene la misma estructura de nodos y ramas (sin dirección).

2. Se cambia el sentido de todas las ramas manteniendo la operación que realizan sobre los datos.

3. Se intercambian entrada y salida.

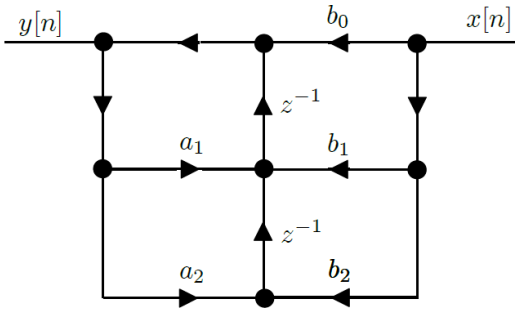

Figura 2.14: Grafo de la implementación en forma traspuesta II.

Aplicando este procedimiento a la forma directa I se obtiene la *forma traspuesta I*, mientras que aplicándolo a la forma firecta II se obtiene la forma traspuesta II. Por ejemplo, la figura 2.14 muestra la implementación mediante forma traspuesta II.

Capítulo 3

Revisión de conceptos básicos en tratamiento estadístico de señales

Las señales pueden clasificarse en dos grandes tipos: deterministas y aleatorias. Una señal *determinista* es aquella que puede puede reproducirse exactamente al repetir su medida o mediante el uso de alguna expresión matemática. Un ejemplo de señal determinista es la respuesta al impulso de un sistema lineal e invariante. Por otra parte, una señal *aleatoria*, o proceso aleatorio, es aquella que no es repetible de una forma predecible. Los ejemplos de señal aleatoria son innumerables: cualquier tipo de ruido, la señal de voz, etc. Cuando se realiza una medición de una señal aleatoria, se dispone de un conjunto "determinista"de valores, pero esos valores generalmente no representarán todos los posibles valores que puede tomar la señal en cada momento. A la hora de diseñar sistemas para procesar señales aleatorias deben tenerse en cuenta de alguna manera las propiedades que puedan presentar éstas en cualquier *realización*, y para ello será necesario acudir a parámetros que indiquen qué posibles situaciones pueden producirse, es decir, parámetros que caractericen en la mayor medida posible el comportamiento de estas señales o procesos. Estos parámetros vendrán determinados por las propiedades estadísticas de los procesos, y determinarán el diseño de los sistemas, que también se realizará mediante herramientas estadísticas.

En este capítulo se repasan las principales herramientas estadísticas necesarias para el tratamiento de señales aleatorias, así como para la adecuada caracterización de éstas.

3.1. Probabilidad

Un *experimento aleatorio*, ε, es aquél en el que son posibles varios *resultados* previamente identificados. Dados un experimento aleatorio y un determinado suceso A de entre los posibles resultados de dicho experimento, se define una función P denominada *probabilidad* del suceso A, $P(A)$, que cumple una serie de axiomas:

- $P(A) \in \mathbb{R}$
- $P(A) \geq 0$ (es una función no negativa)

- Está normalizada: $P(\mathcal{S}) = 1$, siendo S el espacio muestral, es decir, el conjunto de todos los posibles resultados del experimento.
- Si se considera una secuencia de eventos mutuamente excluyentes $A_1, A_2, \ldots, A_N$ $(A_i \cap A_j = \varnothing, \quad i \neq j)$, entonces

$$P\left(\bigcup_{i=1}^{N} A_i\right) = \sum_{i=1}^{N} P(A_i) \tag{3.1}$$

De forma intuitiva, el valor $P(A)$ puede interpretarse como una medida de la posibilidad de que ocurra el evento A en una realización de un experimento aleatorio. Para la definición de la probabilidad de un determinado suceso existen varias alternativas:

- *Definición de Laplace*: La probabilidad de un suceso A se puede calcular como el número de posibles resultados *favorables*, n_A, (sucesos diferentes del espacio muestral que están incluidos en A) partido el número total de resultados posibles, N (número de sucesos en el espacio muestral). Sólo es aplicable para sucesos de igual probabilidad (*equiprobables*[1]).

$$P(A) = \frac{n_A}{N} \tag{3.2}$$

Ejemplo 3.1. *Lanzamiento de un dado.*

ε = lanzar el dado
$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$
A = resultado par

$$P(A) = \frac{3}{6} = 0{,}5$$

- *Definición basada en frecuencia relativa*: Si r es el número de veces que se repite el experimento, la probabilidad del suceso A se define como

$$P = \lim_{n \to \infty} \frac{n_A}{r} \tag{3.3}$$

siendo n_A en este caso el número de veces que el resultado del experimento es A. El inconveniente de esta definición es que requiere infinitos experimentos para caracterizar exactamente la probabilidad.

- *Definición axiomática*: Un experimento aleatorio ε se modela como un *espacio de probabilidad* definido por $\langle \mathcal{S}, \mathcal{F}, P \rangle$, donde $\mathcal{S}$ es el espacio muestral, $\mathcal{F}$ es el conjunto de posibles eventos, denominado *álgebra de eventos*, y P es la función probabilidad. Entonces, P es una función que mapea eventos $A \in \mathcal{F}$ en el espacio de los números reales:

$$P : A \in \mathcal{F} \to P(A) \in \mathbb{R} \tag{3.4}$$

[1] Dos eventos A y B son equiprobables si $P(A) = P(B)$

3.1.1. Espacio muestral, eventos y álgebra de eventos

Como ya se han mencionado, se denomina *espacio muestral*, $\mathcal{S}$, al conjunto de todos los posibles resultados de un determinado experimento aleatorio. Por ejemplo, si el experimento aleatorio es lanzar un dado, el espacio muestral es

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Si el experimento consiste en lanzar una moneda, entonces el espacio muestral es

$$\mathcal{S} = \{\text{cara}, \text{cruz}\}$$

Si se trata de lanzar una moneda dos veces,

$$\mathcal{S} = \{\text{cara-cara}, \text{cara-cruz}, \text{cruz-cara}, \text{cruz-cruz}\}$$

Si se realiza un experimento aleatorio se obtiene un resultado s que es un elemento de $\mathcal{S}$ ($s \in \mathcal{S}$). Si se considera un evento A (que en teoría de la probabilidad se define como un subconjunto de $\mathcal{S}$, $A \subseteq \mathcal{S}$), diremos que *ocurre* A si habiendo realizado el experimento se ha obtenido un resultado $s \in \mathcal{S}$ que cumple $s \in A$. De lo anterior es evidente que para cualquier resultado s siempre ocurre $\mathcal{S}$, por lo que al espacio muestral también se le denomina *evento seguro*. Igualmente, para cualquier resultado s se cumple que $s \notin \varnothing$, y por ello el conjunto vacío $\varnothing$ nunca ocurre y se denomina *evento imposible*.

Ejemplo 3.2. *Lanzamiento de una moneda.*

ε = lanzar una moneda dos veces
$\mathcal{S} = \{\text{cara-cara}, \text{cara-cruz}, \text{cruz-cara}, \text{cruz-cruz}\}$
Consideraremos el evento A = obtener cara la primera vez. Entonces el evento A puede representarse como el siguiente subconjunto de $\mathcal{S}$:

$$A = \{\text{cara-cara}, \text{cara-cruz}\} \subset \mathcal{S}$$

Si se realiza el experimento obteniendo s:

Si s={cara-cara} o s={cara-cruz} $\Rightarrow$ ocurre A
Si s={cruz-cara} o s={cruz-cruz} $\Rightarrow$ no ocurre A

El conjunto de todos los eventos considerados en un experimento aleatorio dado se denomina *álgebra de eventos*, $\mathcal{F}$, y es un conjunto de subconjuntos de $\mathcal{S}$. No deben confundirse los elementos de $\mathcal{S}$, denominados *eventos elementales*, con los de $\mathcal{F}$. Los eventos contenidos en $\mathcal{F}$ pueden estar dados por combinaciones de eventos elementales, y además deben incluir todos los eventos elementales ($\mathcal{S}$) y el conjunto vacío. El álgebra de eventos debe ser lo suficientemente grande como para contener todos los eventos interesantes, pero no tanto como para que sea inmanejable.

Ejemplo 3.3. *Partido Real Madrid CF vs. FC Barcelona.*

$\mathcal{S}$ = {Gana el Madrid, Gana el Barça, Empate}
Por ejemplo, el evento 'El Real Madrid no pierde' está formado por la unión de los eventos 'Gana el Madrid' y 'Empate'.
Un álgebra de eventos $\mathcal{F}$ adecuado podría ser:

$\mathcal{F} = \{\{\varnothing\}, \{S\}, \{\text{Gana el Madrid}\}, \{\text{Gana el Barça}\}, \{\text{El Madrid no pierde}\}, \{\text{El Barça no pierde}\}\}$

Por ejemplo, el evento 'Gana el Madrid'∪'Gana el Barça' no se incluye en el álgebra de eventos, puesto que no tiene interés práctico (no pueden ganar los dos equipos a la vez).

Puede no haber un único álgebra de eventos asociado a cada experimento, pero debe elegirse adecuadamente en función de lo que se pretenda estudiar.

3.1.2. Probabilidad condicional

Se define la *probabilidad del suceso A condicionada al suceso B* como

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{3.5}$$

y representa la probabilidad del suceso A una vez que ha tenido lugar el suceso B. De alguna manera, el conocimiento de que ya ha tenido lugar un suceso determinado B influye en la probabilidad de que suceda el suceso A que se estudia reduciendo los posibles resultados (el espacio muestral) a aquellos eventos que cumplen B.

Ejemplo 3.4. *Lanzamiento de un dado.*

La probabilidad de que suceda A = '4' es $P(\text{'4'}) = \frac{1}{6}$. Si se conoce que ha sucedido B = 'El resultado de la tirada es un número par', entonces la probabilidad de A ya no es la misma. Sólo pueden suceder '2', '4' y '6', de forma que $P(\text{'4'}|\text{'par'}) = \frac{1}{3}$.

Una forma sencilla de razonar el concepto de probabilidad condicionada o *condicional* es mediante su representación equivalente como superficies (figura 3.1). Si se considera que la probabilidad de un suceso es equivalente a la cantidad de superficie que ocupa respecto de la del conjunto de eventos posibles (el espacio muestral $\mathcal{S}$, en principio), la probabilidad de un suceso A condicionada a un suceso B es el área que ocupa A respecto al que ocupa B (una vez que el evento B ha sucedido, se convierte en el espacio de sucesos posibles). En la figura 3.1 el espacio muestral $\mathcal{S}$ tiene una probabilidad unidad ($superficie\{\mathcal{S}\}/superficie\{\mathcal{S}\} = 1$). El suceso B, lógicamente, tiene una superficie menor. Puesto que se asume como condición que el suceso B ha tenido lugar, la probabilidad del suceso A será, necesariamente, el área ocupada por aquella parte del mismo que también sea parte de B, es decir, la intersección entre las dos superficies. De ahí se puede extraer directamente la expresión (3.5), donde también se tiene en cuenta que $P(B|B) = 1$, es decir, que una vez ha sucedido el suceso B, este suceso es seguro, y las probabilidades condicionales a él se normalizarán a su probabilidad.

De este razonamiento y de (3.5) pueden extraerse varias consecuencias:

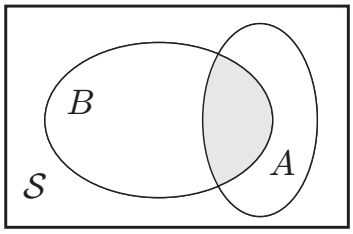

Figura 3.1: Representación de la probabilidad condicional como una superficie.

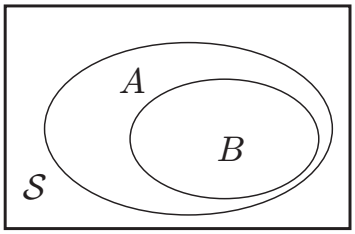

Figura 3.2: Representación de la probabilidad condicional como una superficie.

- Si $B \subset A \Rightarrow P(A|B) = 1$. Si B es un subconjunto de A, y ha ocurrido B, entonces necesariamente ha ocurrido A (figura 3.2).

- Si $A \subset B \Rightarrow P(A|B) = P(A)/P(B) \geq P(A)$. Cuando A es un subconjunto de B, y ha sucedido B, hay más probabilidad de que suceda A que a priori, es decir, se reduce el posible espacio muestral (figura 3.3).

- Si $A \cap B = \varnothing \Rightarrow P(A|B) = 0$. Si A y B son disjuntos y ha ocurrido B, no puede haber ocurrido A (figura 3.4).

Además de todo esto, si se cumple que $P(A|B) = P(A)$, entonces el suceso B no aporta ninguna información. Se dice en ese caso que los sucesos A y B son *sucesos independientes.*

3.1.3. Probabilidad total y Teorema de Bayes

Supongamos un conjunto de N sucesos disjuntos B_i, con $i = 1 \ldots N$, y que cumplen que $\cup_{i=1}^{N} B_i = \mathcal{S}$, es decir, que provienen de dividir el espacio muestral en dichos sucesos.

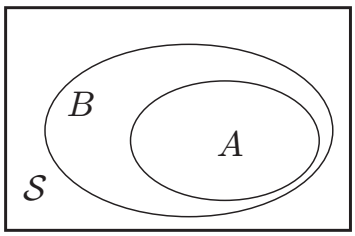

Figura 3.3: Representación de la probabilidad condicional como una superficie.

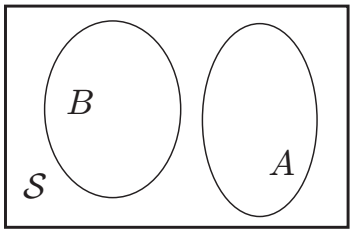

Figura 3.4: Representación de la probabilidad condicional como una superficie.

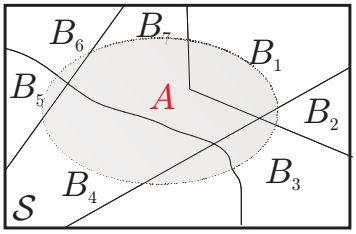

Figura 3.5: Representación del Teorema de Probabilidad Total.

El **Teorema de Probabilidad Total** establece que la probabilidad de un suceso A se puede calcular como:

$$P(A) = \sum_{i=1}^{N} P(B_i)P(A|B_i) \tag{3.6}$$

Esto significa que la probabilidad del suceso A es la suma de su intersección con todos los segmentos B_i que componen el espacio muestral (figura 3.5). Según la definición de probabilidad condicional (3.5), la probabilidad del suceso $A \cup B_i$ es

$$P(A \cup B_i) = P(B_i)P(A|B_i) \tag{3.7}$$

de donde se obtiene directamente la expresión (3.6).

3.2. Variables aleatorias

Una variable aleatoria es una función que asigna un valor real a cualquier posible resultado de un experimento. Si consideramos un experimento aleatorio definido por $\langle \mathcal{S}, \mathcal{F}, P \rangle$, la variable aleatoria X es una función de la forma

$$X : \omega \in \mathcal{S} \rightarrow X(\omega) \in \mathbb{R} \tag{3.8}$$

es decir, mapea cada resultado ω del experimento en un valor numérico (figura 3.6) de forma que permite convertir cualquier fenómeno aleatorio es otro numérico, independientemente de su naturaleza.

Ejemplo 3.5. *Colores de una ruleta.*

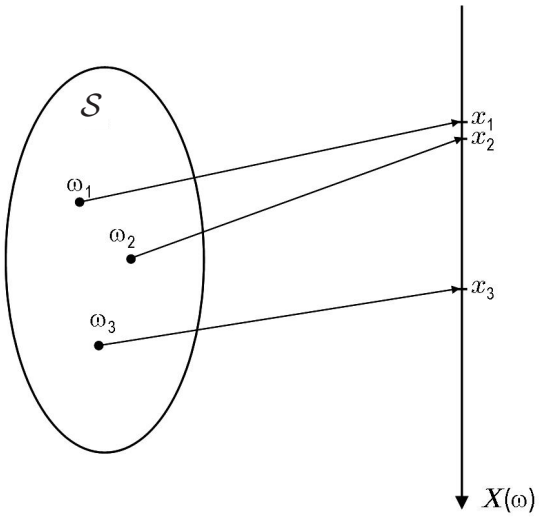

Figura 3.6: Mapeo de eventos elementales en puntos de la recta real.

Los posibles resultados del experimento "Color resultante en una ruleta" son 'rojo' y 'negro'. Se define:

$$X(\text{'rojo'}) = 0$$
$$X(\text{'negro'}) = 1$$

De este modo, $P(\text{'rojo'}) = P(0) = 0,5$, y $P(\text{'negro'}) = P(1) = 0,5$.

Ejemplo 3.6. *Lanzamiento de un dado.*

Los posibles resultados del experimento ya proporcionan un valor numérico de 1 a 6. La variable aleatoria va implícita en la naturaleza del experimento.

Una variable aleatoria sólo puede asignar a cada posible resultado en $\mathcal{S}$ un único número real $X(\omega) \in \mathbb{R}$. Se denomina *alcance* de X, $\mathcal{R}_X$, al conjunto de todos los números reales $X(\omega)$, mientras que el espacio muestral $\mathcal{S}$ es el *dominio* de X.

$$\mathcal{R}_X = \{X(\omega) : \omega \in \mathcal{S}\} \subseteq \mathbb{R} \tag{3.9}$$

Existen dos grandes tipos de variables aleatorias en función de las propiedades del espacio muestral sobre el que se realiza el experimento. Cuando el espacio muestral consiste en un conjunto finito de eventos se dice que la variable aleatoria es *discreta*. Cuando el conjunto de posibles resultados del experimento es infinito la variable aleatoria es *continua*. Por ejemplo, el resultado de lanzar una ruleta es una variable aleatoria discreta, puesto que el número de posibles resultados es finito. Del mismo modo, si se considera un punto de referencia en un punto de la ruleta, la fase con la que se detiene el giro de la ruleta puede tomar infinitos valores entre 0^o y 360^o, de forma que el valor de dicha fase constituye una variable aleatoria continua.

Cuando se trabaja con una variable aleatoria continua no tiene sentido definir una probabilidad asociada a un determinado resultado, dado que el número de posibles resultados del experimento estudiado es infinito, y cualquier probabilidad sería nula.

Ejemplo 3.7. *Temperatura en un punto geográfico.*

Supongamos que se desea conocer la probabilidad de que la temperatura en un determinado punto geográfico sea exactamente de 18^oC. El valor de la temperatura en ese punto es una variable aleatoria continua. Asumiendo que son ciertas las condiciones que permiten aplicar la definición de Laplace, la probabilidad deseada es:

$$P(\text{Temperatura} = 18^oC) = \frac{\text{casos favorables}}{\text{casos posibles}} = \frac{1}{\infty} = 0$$

Una probabilidad nula implicaría que no es posible que la temperatura sea de 18^oC, lo que evidentemente es falso.

La forma de trabajar con probabilidades asociadas a variables aleatorias continuas pasa por definir intervalos de valores y nuevos eventos consistentes en la pertenencia a dichos intervalos. Si se define un intervalo $I = (\alpha_1, \alpha_2]$, la probabilidad del evento $\omega \in I$ viene dada por

$$P(\omega \in I) = P(\alpha_1 < \omega \leq \alpha_2) \tag{3.10}$$

En aplicaciones de procesado de señal suele ser más interesante la descripción probabilística de la variable aleatoria que la caracterización estadística de los eventos del espacio muestral. Por ello, es más conveniente disponer de una ley de probabilidad que se aplique directamente sobre la variable aleatoria y no sobre los eventos. Para variables aleatorias reales una opción para la caracterización es la *función de distribución de probabilidad* (o, simplemente, función de distribución).

3.2.1. Función de distribución y función densidad de probabilidad

Sea una variable aleatoria definida sobre $\langle \mathcal{S}, \mathcal{F}, P \rangle$. Si se define un intervalo $I = (-\infty, x]$, con $x \in \mathbb{R}$, la probabilidad $P(X \in I)$ está bien definida y se puede expresar como

$$P(X \in I) = P(X \leq x) = F_X(x) \tag{3.11}$$

La función $F_X(x)$ es la función de distribución de la variable aleatoria X. Se trata de una función definible tanto para variables aleatorias continuas como discretas, y que cumple:

$$F_X : \mathbb{R} \to [0, 1] \tag{3.12}$$
$$F_X(-\infty) = 0 \tag{3.13}$$
$$F_X(\infty) = 1 \tag{3.14}$$
$$x_1 < x_2 \Rightarrow F_X(x_1) \leq F_X(x_2) \tag{3.15}$$

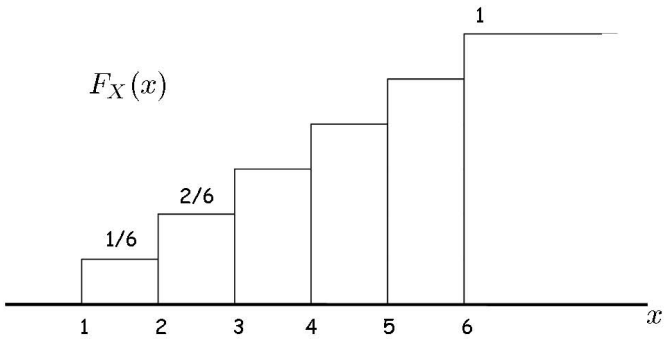

Figura 3.7: Función de distribución de la variable aleatoria 'Resultado de lanzar un dado'.

Ejemplo 3.8. *Lanzamiento de un dado.*

La función de distribución de la variable aleatoria 'resultado del lanzamiento de un dado sigue la forma de la figura 3.7.

En el ejemplo anterior se distinguen seis discontinuidades en la función de distribución para los valores de x que coinciden con los valores del espacio muestral. Estas discontinuidades se deben a *masas de probabilidad* en estos puntos, es decir, a la presencia de eventos con probabilidad no nula por tratarse de una variable aleatoria discreta. Formalmente se dice que una variable aleatoria X es continua si su función de distribución es *absolutamente continua*, es decir:

- $F_X(x)$ es continua $\forall x \in \mathbb{R}$
- La función $F_X(x)$ es derivable para todo x excepto quizá para un conjunto numerable de puntos

En caso contrario se dice que la variable X es discreta.

Otra caracterización estadística especialmente útil de una variable aleatoria es la *función de densidad de probabilidad*, $f_X(x)$, $f : \mathbb{R} \to \mathbb{R}$, que se define como

$$f_X(x) = \frac{d}{dx} F_X(x) \tag{3.16}$$

La función densidad de probabilidad cumple

$$f_X(x) \geq 0 \tag{3.17}$$

$$\int_{-\infty}^{\infty} f_X(x)dx = 1 \tag{3.18}$$

$$\int_{-\infty}^{x} f_X(x)dx = F_X(x) \tag{3.19}$$

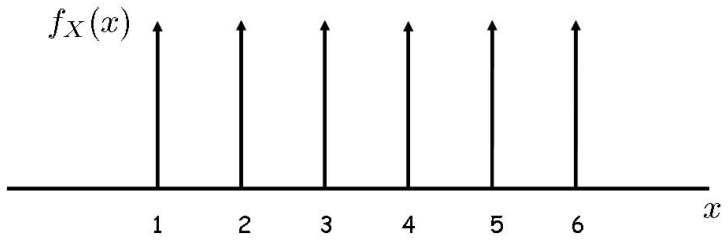

Figura 3.8: Función densidad de probabilidad de la variable aleatoria 'Resultado de lanzar un dado'.

De la ecuación (3.19) se obtiene que $F_X(x) = P(X \leq x)$ es el área bajo la función densidad de probabilidad desde $-\infty$ hasta x. Si consideramos dos valores x_1 y x_2, con $x_1 < x_2$, entonces

$$\int_{x_1}^{x_2} f_X(x)dx = \int_{-\infty}^{x_2} f_X(x)dx - \int_{-\infty}^{x_1} f_X(x)dx = P(X \leq x_1) - P(X \leq x_2)$$

$$\int_{x_1}^{x_2} f_X(x)dx = P(x_1 < x \leq x_2) \tag{3.20}$$

Es decir, el área bajo la función densidad de probabilidad en un determinado intervalo proporciona la probabilidad de dicho intervalo.

Ejemplo 3.9. *Lanzamiento de un dado.*

La función densidad de probabilidad de la variable aleatoria 'resultado del lanzamiento de un dado' sigue la forma de la figura 3.8.

En el ejemplo anterior la función densidad de probabilidad está formada por seis impulsos unitarios multiplicados por la probabilidad de cada uno de los eventos. En el caso de variables aleatorias discretas es típico encontrar funciones densidad de probabilidad formadas por impulsos debido a la presencia de las masas de probabilidad asociadas a cada uno de los eventos. Estas masas de probabilidad generaban las discontinuidades que se observaron en la función de distribución, que, al ser derivadas, corresponden a los impulsos que aparecen en la función densidad de probabilidad. Esta función es consistente con la definición de probabilidad para los eventos asociados a una variable aleatoria discreta, puesto que la integral de la función densidad de probabilidad en un intervalo arbitrariamente pequeño en torno a cada uno de tales eventos proporciona su probabilidad.

También pueden definirse una función de distribución de probabilidad condicional y una función de densidad de probabilidad condicional (condicionadas a un determinado suceso A) como

$$F_X(x|A) = P(X \leq x|A) \tag{3.21}$$

$$f_X(x|A) = \frac{d}{dx} F_X(x|A) \tag{3.22}$$

3.2.1.1. Distribuciones conjuntas

Cuando se trabaja con más de una variable aleatoria es necesario considerar las dependencias estadísticas que puedan existir entre las variables. Para ello se definen las funciones de distribución y de densidad de probabilidad conjuntas de la siguiente manera. Dadas dos variables aleatorias X e Y, la función de distribución conjunta es

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y) \tag{3.23}$$

que define la probabilidad de que X sea menor o igual que un valor x y de que Y sea menor o igual que y al mismo tiempo. Entonces, la función densidad de probabilidad conjunta es la derivada de la función de distribución, dada por

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) \tag{3.24}$$

Las funciones conjuntas también se emplean para la caracterización estadística de variables aleatorias complejas, considerando como variables diferentes las partes real e imaginaria.

Para más de dos variables aleatorias, las funciones conjuntas se definen de modo similar. Por ejemplo, para un conjunto de n variables aleatorias la función de distribución conjunta es

$$F_{X_1,X_2 \ldots X_n}(x_1, x_2 \ldots x_n) = P(X_1 \leq x_1, X_2 \leq x_2 \ldots X_n \leq x_n) \tag{3.25}$$

mientras que la función densidad de probabilidad conjunta será

$$f_{X_1,X_2 \ldots X_n}(x_1, x_2 \ldots x_n) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F_{X_1,X_2 \ldots X_n}(x_1, x_2 \ldots x_n) \tag{3.26}$$

3.2.2. Estadísticos de una variable aleatoria

3.2.2.1. Esperanza matemática

Una caracterización estadística completa de una variable aleatoria requiere que sea posible determinar la probabilidad de cualquier evento definido en el espacio muestral. Sin embargo, en muchas aplicaciones no es necesaria una caracterización completa si se conoce el comportamiento promedio de la variable aleatoria. Éste suele ser el caso de muchos problemas de tratamiento digital de señales, donde generalmente interesan promedios estadísticos. Una herramienta clave para este tipo de estudio es el operador *esperanza matemática* o *valor esperado*[2].

Consideremos el experimento "lanzar un dado". Si el dado se lanza N_T veces y el valor k aparece n_k veces, la *media muestral* o *valor promedio* es

$$\langle X \rangle_{N_T} = \frac{n_1 + 2n_2 + 3n_3 + 4n_4 + 5n_5 + 6n_6}{N_T}$$

[2]En ocasiones también es habitual referirse a la esperanza matemática como *media*, aunque esta expresión puede dar lugar a equívoco dado que algunos textos se emplea como equivalente de *promedio*. En este texto, el término *media* siempre se considerará sinónimo de *esperanza matemática*

Si N_T es muy grande, entonces $n_k/N_T \approx P(X = k)$, y la expresión anterior se puede expresar como

$$\langle X \rangle_{N_T} \approx \sum_{k=1}^{6} kP(X = k) \tag{3.27}$$

La esperanza matemática o valor esperado de una variable aleatoria X en la que se asume que la probabilidad de que una realización tome el valor x es $P(x)$ es

$$E[X] = \sum_i x_i P(x_i) \tag{3.28}$$

Lógicamente, esta expresión sólo es válida para variables aleatorias discretas. Sin embargo, puede expresarse en términos de la función densidad de probabilidad como

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx \tag{3.29}$$

Esta expresión es igualmente válida para variables aleatorias continuas y discretas.

Cuando una variable aleatoria presenta una función densidad de probabilidad simétrica con respecto a un determinado valor μ, se cumple que $f_X(\mu - \alpha) = f_X(\mu + \alpha), \forall \alpha \in \mathbb{R}$, y se puede comprobar que $E[X] = \mu$.

La esperanza matemática es un operador lineal, es decir, cumple que

$$E\left[\sum_{k=a_0}^{a_1} \alpha_k h_k(X)\right] = \sum_{k=a_0}^{a_1} \alpha_k E[h_k(X)]$$

En muchos problemas será necesario obtener la esperanza matemática de una función de una variable aleatoria. Si consideramos una función $h(X)$, $h : \mathbb{R} \to \mathbb{R}$, el valor medio de $h(X)$ es

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx \tag{3.30}$$

Por otra parte, podemos definir una nueva variable aleatoria $Y = h(X)$, cuya función densidad de probabilidad es $F_Y(y)$. La esperanza matemática de la variable Y será

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy = E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx \tag{3.31}$$

es decir, no es necesario conocer $f_Y(y)$ explícitamente para calcular la esperanza matemática de Y, sino que basta con conocer la función $h(X)$.

3.2.2.2. Varianza

Un parámetro estadístico importante que se puede expresar como un caso particular de esperanza matemática de una función de una variable aleatoria es la *varianza*. La varianza

es una medida de la dispersión de los valores que puede tomar la variable aleatoria respecto a su media. Suele denotarse como $\mathrm{Var}[X]$ o σ_X^2 y se define como

$$\sigma_X^2 = E\left[(X - E[X])^2\right] = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x) dx \tag{3.32}$$

La raíz cuadrada de la varianza, σ_X, se denomina *desviación típica*.

Empleando la linealidad del operador esperanza matemática, la varianza también se puede expresar como

$$\sigma_X^2 = E[X^2] - E^2[X] \tag{3.33}$$

Algunas propiedades adicionales de la varianza son:

- $\sigma_X^2 \geq 0$
- $\mathrm{Var}[aX + b] = a^2 \mathrm{Var}[X]$
- Si se observa la desviación estándar: $\sigma_{aX+b} = |a|\sigma_X$

3.2.2.3. Momentos conjuntos. Independencia, incorrelación y ortogonalidad

Del mismo modo que con variables aleatorias individuales, los promedios estadísticos proporcionan una útil e importante caracterización de variables aleatorias distribuidas conjuntamente. Los dos principales promedios estadísticos son la *correlación* y la *covarianza*. La correlación, denotada como R_{XY} es el momento de segundo orden dado por[3]

$$R_{XY} = E[X \cdot Y] \tag{3.34}$$

La covarianza se denota como C_{XY} o $\mathrm{Cov}[X, Y]$ y se define como

$$C_{XY} = E\left[(X - E[X])(Y - E[Y])\right] = E[X \cdot Y] - E[X]E[Y] \tag{3.35}$$

Se dice que dos variables aleatorias X e Y son *estadísticamente independientes* si su función de densidad de probabilidad conjunta es separable de la forma

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \tag{3.36}$$

La independencia estadística indica que no existe relación entre ambas variables y que una no ofrece información acerca de la otra. La definición de independencia estadística puede extenderse a cualquier número de variables.

Un forma más suave de independencia ocurre cuando el momento de segundo orden $E[XY]$ es separable:

$$R_{XY} = E[XY] = E[X]E[Y] \tag{3.37}$$

Dos variables aleatorias que satisfacen (3.37) son *incorreladas*. Es inmediato comprobar que dos variables aleatorias incorreladas presentan covarianza nula. Dos variables aleatorias

[3] Se está asumiendo que se trabaja con variables aleatorias reales.

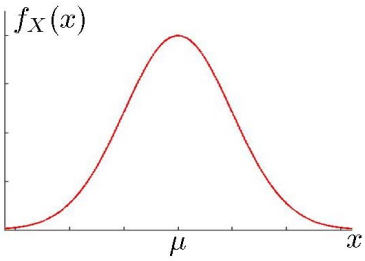

Figura 3.9: Función densidad de probabilidad de una variable aleatoria Gaussiana.

estadísticamente independientes están siempre incorreladas, pero la afirmación inversa no es cierta en general. Una propiedad útil de las variables aleatorias incorreladas es que la varianza de la suma de ambas variables es igual a la suma de sus varianzas:

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] \tag{3.38}$$

Una propiedad relacionada con la incorrelación es la ortogonalidad. Dos variables aleatorias son ortogonales si su correlación es nula. Aunque dos variables ortogonales no tienen que ser necesariamente incorreladas, las variables aleatorias incorreladas de media nula también son ortogonales.

Como ya se ha mencionado, el principal interés en el estudio de fenómenos aleatorios en tratamiento digital de señales es el comportamiento promedio frente a la descripción probabilística de las variables aleatorias. Por ello, en general resultará de mayor interés conocer si dos variables están o no incorreladas que si son estadísticamente independientes.

3.2.3. Algunos tipos de variable aleatoria

3.2.3.1. Variables aleatorias Gaussianas

Se dice que una variable aleatoria X es Gaussiana o normal, de parámetros μ y σ, y se representa de la forma $X \sim N(\mu, \sigma^2)$, si su función densidad de probabilidad es de la siguiente forma (figura 3.9):

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad x \in \mathbb{R} \tag{3.39}$$

El parámetro μ se suele denominar *media* de la variable aleatoria, σ es su *desviación típica* y σ^2 es su *varianza*.

La distribución gaussiana es la más común en la naturaleza y juega un papel primordial en la teoría de la probabilidad, así como en en el procesado de señal para comunicaciones. Por ejemplo, el valor del ruido de comunicaciones en un determinado instante suele ser modelado como una realización de una variable aleatoria Gaussiana.

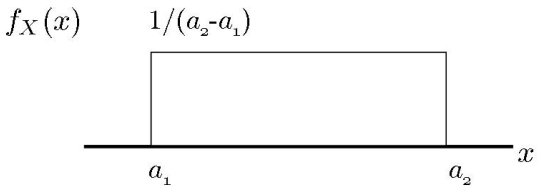

Figura 3.10: Función densidad de probabilidad de una variable aleatoria uniforme.

Supongamos una variable aleatoria Y definida como una combinación lineal de N variables aleatorias Gaussianas incorreladas:

$$Y = \sum_{i=1}^{N} \alpha_i X_i, \quad X_i \sim N(\mu_i, \sigma_i^2) \tag{3.40}$$

La variable Y es también Gaussiana con las siguientes propiedades:

$$E[Y] = \textstyle\sum_{i=1}^{N} \alpha_i \mu_i \tag{3.41}$$
$$\sigma_Y^2 = \textstyle\sum_{i=1}^{N} \alpha_i^2 \sigma_i^2 \tag{3.42}$$

3.2.3.2. Variables aleatorias uniformes

Se dice que una variable aleatoria X es uniforme entre a_1 y a_2 si su función densidad de probabilidad es de la siguiente forma (figura 3.10):

$$\begin{array}{ll} f_X(x) = \frac{1}{a_2 - a_1}, & x \in [a_1, a_2] \\ f_X(x) = 0, & x \notin [a_1, a_2] \end{array} \tag{3.43}$$

Aunque la distribución uniforme es menos común que la normal, resulta igualmente de gran utilidad en muchos problemas de tratamiento de señales. Sus principales propiedades son:

$$\begin{array}{ll} \text{Media:} & \frac{a_1 + a_2}{2} \\ \text{Varianza:} & \frac{(a_2 - a_1)^2}{12} \end{array} \tag{3.44}$$

3.3. Procesos estocásticos

Las señales que se manejan en problemas de comunicaciones, o de tratamiento digital de señales en general, suelen ser de naturaleza aleatoria. El valor de las señales en un determinado instante n es una variable aleatoria $X[n]$ que toma el valor $x[n]$. Esto responde perfectamente a la definición de *proceso estocástico*. Las señales aleatorias se modelan como procesos estocásticos, lo que permite utilizar herramientas estadísticas para su tratamiento.

Un proceso estocástico es una secuencia de variables aleatorias indexadas. Considerando un espacio de probabilidad $\langle \mathcal{S}, \mathcal{F}, P \rangle$, un proceso estocástico es una secuencia de

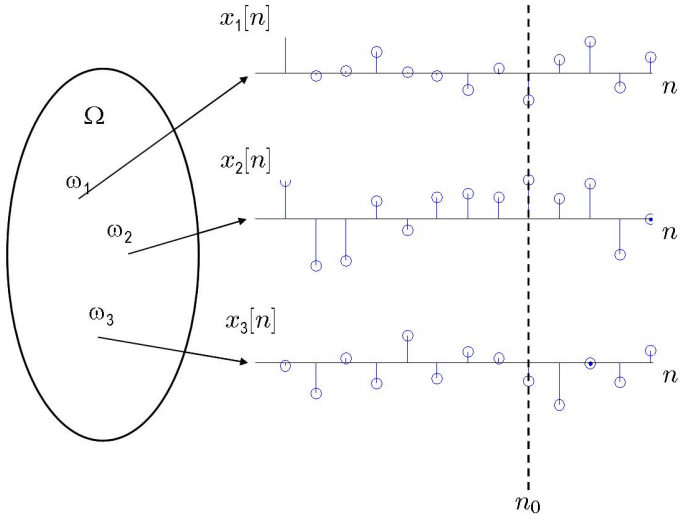

Figura 3.11: Mapeo de eventos elementales en señales de tiempo discreto.

variables aleatorias indexadas de la forma $X(t)$, definidas sobre $\mathcal{S}$, y con $t \in \mathcal{T} \subseteq \mathbb{R}$. Dependiendo de si el índice t es continuo o discreto, el proceso estocástico se denomina de *tiempo continuo* o de *tiempo discreto*. En este texto siempre se considerarán procesos de tiempo discreto que se denotarán de la forma $X[n]$, $n \in \mathbb{Z}$.

Del mismo modo que una variable aleatoria puede verse como un mapeo desde el espacio muestral de un experimento al conjunto de los números reales (o complejos), un proceso estocástico en tiempo discreto es un mapeo desde el espacio muestral $\mathcal{S}$ al conjunto de las señales de tiempo discreto (figura 3.11). Así, un proceso estocástico en tiempo discreto es un conjunto de señales de tiempo discreto. Si el número de posibles eventos en el espacio muestral es finito, el proceso estocástico es discreto, mientras que si el número de eventos posibles es infinito, el proceso estocástico es continuo.

Un proceso estocástico puede entenderse también como una función de dos argumentos, $[n, s] \in \mathbb{Z} \times \mathcal{S} \rightarrow X[n, s] \in \mathbb{R}$. Para un índice $n = n_0$, $X[n_0, s]$ es una variable aleatoria, mientras que para un suceso determinado $s = s_0$, $X[n, s_0]$ es una realización del proceso estocástico y una señal/función de n (figura 3.12).

3.3.1. Propiedades de un proceso estocástico

Un proceso estocástico puede expresarse como un vector de variables aleatorias, $X[n] = [X_1 X_2 X_3 X_4 \ldots] = [X[1]X[2]X[3]X[4]\ldots]$. Esas variables aleatorias no tienen por qué ser iguales, pero en cualquier caso cada una de ellas tendrá las propiedades de cualquier variable aleatoria. Por ejemplo, cada una de las variables aleatorias tiene una función de distribución de probabilidad, de forma que tendremos

$$F_{X[n]}(x) = P(X[n] \leq x) \tag{3.45}$$

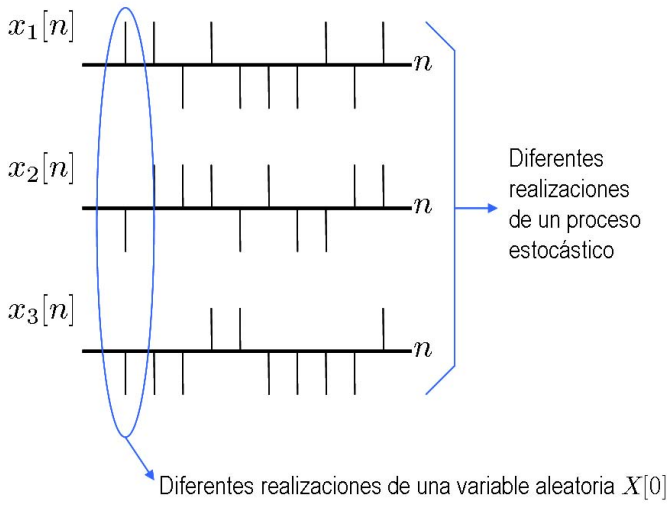

Figura 3.12: Realizaciones de un proceso estocástico como conjunto de variables aleatorias indexadas.

donde el índice n se incluye para hacer referencia a la variable aleatoria n-ésima dentro del proceso. Igualmente, cada una tendrá una función densidad de probabilidad:

$$f_{X[n]} = \frac{d}{dx} F_{X[n]}(x) \tag{3.46}$$

formando un conjunto de funciones de distribución y de densidad de probabilidad que caracterizan al proceso estocástico.

Igualmente, cada variable tendrá una esperanza matemática y una varianza, con lo que podemos definir un vector de esperanzas y otro de varianzas de la forma

$$\mu[n] = E[X[n]] = [E[X[1]], E[X[2]], \ldots] \tag{3.47}$$

$$\sigma_X^2[n] = E[(X[n] - \mu[n])^2] \tag{3.48}$$

También se puede caracterizar el proceso estocástico mediante otros estadísticos importantes como son la autocorrelación y la autocovarianza. La autocorrelación es

$$R_{XX}[l, m] = E[X[l]X[m]] \tag{3.49}$$

y la autocovarianza

$$C_{XX}[l, m] = E[(X[l] - \mu[l])(X[m] - \mu[m])] \tag{3.50}$$

donde se puede observar que si $l = m$ la autocovarianza coincide con la varianza del proceso para índice m. Al autocovarianza y la autocorrelación de un proceso están relacionadas por

$$C_{XX}[l, m] = R_{XX}[l, m] - \mu[l]\mu[m] \tag{3.51}$$

y coinciden en el caso de procesos de medias nulas. Tanto la autocorrelación como la autocovarianza proporcionan información sobre la relación estadística entre dos variables aleatorias que forman parte del mismo proceso.

Paralelamente, en aplicaciones donde intervienen dos procesos $X[n]$ e $Y[n]$, se pueden definir la correlación cruzada

$$R_{XY}[l,m] = E[X[l]Y[m]] \tag{3.52}$$

y la covarianza cruzada

$$C_{XY}[l,m] = E[(X[l] - \mu_X[l])(Y[m] - \mu_Y[m])] \tag{3.53}$$

Se dice que dos procesos estocásticos están *incorrelados* si $C_{XY}[l,m] = 0, \forall l,m$, y que son ortogonales si $R_{XY}[l,m] = 0, \forall l,m$. Obviamente, dos procesos estocásticos incorrelados de media nula son ortogonales.

3.3.2. Estacionariedad

La estacionariedad es una propiedad de los procesos estocásticos que aporta información sobre la evolución de sus propiedades con el tiempo, o más estrictamente, con la variable que se utiliza para indexar las variables aleatorias que contiene, n. En un proceso estacionario no es necesario estudiar la dependencia con dicha variable puesto que se cumple que su comportamiento general no cambia.

La estacionariedad de un proceso se estudia a través de un análisis exhaustivo de sus propiedades estadísticas. Sin embargo, en aplicaciones de tratamiento digital de señales no suele ser necesario, siendo suficiente el estudio de la *estacionariedad en sentido amplio*, que es una restricción mucho más débil que la estacionariedad *en sentido estricto*.

Un proceso estocástico es *estacionario en sentido amplio* si cumple que:

- Su media no depende de n:
$$\mu[n] = \mu, \forall n$$
- Su autocorrelación sólo depende de la diferencia de índices:
$$R_{XX}[n, n+m] = R_{XX}[m], \forall n, m$$
- La varianza del proceso es finita.

En el caso de dos procesos se define la estacionariedad conjunta. Dos procesos $X[n]$ e $Y[n]$ son conjuntamente estacionarios en sentido amplio si ambos procesos son estacionarios en sentido amplio y su correlación cruzada sólo depende de la diferencia entre índices:

$$R_{XY}[n, n+m] = R_{XY}[n+l, n+l+m] = R_{XY}[m]$$

La autocorrelación de un proceso real estacionario en sentido amplio es simétrica respecto del origen de la variable *diferencia de índices*, m. Además, la autocorrelación de un proceso real estacionario en sentido amplio tiene como valor máximo $R_{XX}[0] = E[|X[n]|^2] = \mu^2 + \sigma^2$, donde μ es la esperanza matemática del proceso y σ^2 su varianza.

3.3.3. Ergodicidad

Consideremos un proceso estocástico estacionario. Si $x[n]$ es una realización cualquiera (una señal), definimos su media muestral (o temporal) como

$$\langle x[n] \rangle = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x[n] \tag{3.54}$$

y la correlación muestral (o temporal) como

$$\phi_{xx}[m] = \langle x[n]x[n+m] \rangle = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x[n]x[n+m] \tag{3.55}$$

Un proceso estocástico estacionario es *ergódico* si la media muestral coincide con la media estadística, y la correlación muestral también lo hace con la correlación estadística. De alguna manera, la ergodicidad de un proceso viene a significar que es equivalente estudiar una realización de un proceso en sus infinitas muestras que estudiar estadísticamente cualquiera de las variables que componen el proceso estacionario. Esta propiedad abrirá la puerta a muchas aplicaciones, dado que el estudio de una señal aleatoria a través de sus muestras permitirá caracterizar señales sin necesidad de acudir directamente a un estudio estadístico a veces imposible. Afortunadamente, la inmensa mayoría de señales modeladas como procesos estocásticos estacionarios cumplen que también son ergódicas.

3.4. Densidad Espectral de Potencia

Los procesos estocásticos se emplean para modelar conjuntos de señales, o una señal de naturaleza aleatoria, pero al no ser señales deterministas no tienen Transformada de Fourier que permita estudiar directamente su comportamiento en el dominio de la frecuencia. Aunque el operador *Transformada de Fourier* puede aplicarse a una de sus realizaciones, eso no representa al conjunto de señales o al proceso aleatorio, y por tanto no puede emplearse como presentación con ningún sentido físico. Sin embargo, esto no implica que no pueda estudiarse el comportamiento espectral de una señal aleatoria. Su comportamiento frecuencial puede estudiarse a partir de valores estadísticos, y especialmente empleando su Densidad Espectral de Potencia (DEP). Ésta se define como:

$$S_{XX}(\omega) = \sum_{m=-\infty}^{\infty} R_{XX}[m] e^{-j\omega m} \tag{3.56}$$

es decir, es la Transformada de Fourier de la autocorrelación del proceso, que sí es determinista. Inversamente:

$$R_{XX}[m] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{XX}(\omega) e^{j\omega m} d\omega \tag{3.57}$$

Si se trabaja con procesos reales, la autocorrelación del proceso es real y par, lo que implica que la densidad espectral de potencia (DEP) es también real y par. La densidad espectral de potencia de un proceso real estacionario en sentido amplio es no negativa.

También se cumple que la potencia de un proceso estacionario en sentido amplio es proporcional al área bajo la curva de la DEP. Esta propiedad se deriva de que se cumple que

$$R_{XX}[0] = E[|X[n]|^2] = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{XX}(\omega)e^{j\omega 0}d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{XX}(\omega)d\omega \tag{3.58}$$

3.5. Ruido blanco en variable discreta

Un proceso de especial importancia tanto en comunicaciones como en aplicaciones de tratamiento de señales en general es el ruido blanco. En el mundo analógico, se denomina ruido blanco a aquella señal que contiene todas las frecuencias del espectro por igual (de ahí su denominación *blanco*). Las implicaciones de un espectro como ese conducen a que su comportamiento sea impredecible, sus valores absolutamente independientes entre sí (incorrelados), y a que sea modelado siempre como una señal aleatoria. En el caso de señales de variable discreta, un proceso estocástico estacionario en sentido amplio se dice que es blanco si su autocovarianza es nula excepto para diferencias nulas de índices:

$$C_{XX}[m] = \sigma^2\delta[m] \tag{3.59}$$

Un proceso que tiene una autocovarianza como ésta tiene una autocorrelación que igualmente constará de una muestra en el origen, y un valor constante adicional en función de la media del proceso y que puede ser o no nulo. Su Tranformada de Fourier, y por tanto su Densidad Espectral de Potencia, constará de una constante para todas las pulsaciones debido a la presencia de la delta en el origen, y por tanto su espectro será plano o blanco. Asimismo, una covarianza como la indicada significa que se trata de un proceso sin relación entre sus muestras, o entre las variables que lo componen: dada una realización de una de esas variables, es imposible predecir lo que hará la siguiente. Ese comportamiento es lo que define la propia naturaleza del ruido.

Suele existir cierta confusión entre el significado de que un ruido sea blanco y que sea Gaussiano, hasta el punto de que en muchas ocasiones se considera erróneamente que ambas expresiones son equivalente. No es así. Dado que un proceso se define como blanco en función de un momento de segundo orden, hay una infinita variedad de procesos aleatorios calificables como ruido blanco. El hecho de que sea Gaussiano no corresponde con la relación entre sus muestras ni con el espectro que presenta, sino con el valor que puede tomar cada una de esas muestras (con una u otra probabilidad), es decir, con la función de densidad de probabilidad de sus variables aleatorias. Por ello, un *ruido blanco y Gaussiano* es un proceso blanco consistente en variables aleatorias reales Gaussianas e incorreladas.

3.6. Filtrado digital de procesos estocásticos

La figura 3.13 muestra un caso de señales aleatorias modeladas como procesos estocásticos que atraviesan un filtro digital, modelado mediante su respuesta al impulso,

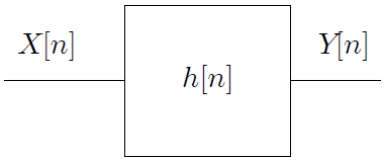

Figura 3.13: Filtrado de procesos estocásticos.

para producir una señal de salida. Necesariamente, la salida del sistema es también una señal aleatoria modelada como proceso estocástico.

Aunque se desconozca una expresión del proceso estocástico de entrada y no pueda representarse analíticamente, el comportamiento del sistema no es aleatorio y sigue respondiendo a los principios desarrollados al inicio de este texto. Por ello, la salida del sistema se puede expresar como

$$Y[n] = X[n] * h[n] \tag{3.60}$$

aunque no sea posible calcularla. Sin embargo, conocidos los parámetros estadísticos del proceso de entrada, sí es posible calcular los del proceso de salida. Por ejemplo, si el proceso de entrada tiene una esperanza matemática μ_x, se puede comprobar que la del proceso de salida es:

$$\mu_y = E[Y[n]] = E[X[n] * h[n]] = E\left[\sum_{k=-\infty}^{\infty} h[k]X[n-k]\right] = \sum_{k=-\infty}^{\infty} E[X[n-k]]h[k]$$

donde se puede aplicar que los valores de $h[n]$ no son aleatorios, que la esperanza matemática es un operador lineal, y que la esperanza del proceso estacionario $X[n]$ no depende de k para obtener

$$\mu_y = \mu_x \sum_{k=-\infty}^{\infty} h[k] \tag{3.61}$$

De modo idéntico se puede desarrollar la autocorrelación de la señal de salida en función de la autocorrelación de la entrada como

$$R_{YY}[m] = R_{XX}[m] * R_{hh}[m] \tag{3.62}$$

donde $R_{hh}[m]$ es la autocorrelación de la respuesta al impulso, que es determinista, y por tanto responde a la definición presentada en su momento para este tipo de señales.

Cabe señalar como conclusión que, a la vista de estos resultados, si el proceso de entrada es estacionario en sentido amplio, el de salida también lo es.