CAPITULO III

ALGUNOS MODELOS DE PROCESOS ALEATORIOS

§ 1. ALGUNAS DEFINICIONES Y EJEMPLOS

1. Definición general del proceso aleatorio. Hablando sobre el proceso aleatorio, como regla, se toma en consideración una determinada magnitud aleatoria $\xi(t)$, que varía en el transcurso del tiempo t. Llamaremos *proceso aleatorio* $\xi = \xi(t)$ *a la* función de parámetro real $t \in T$, cuyos valores $\xi(t)$ para cada t son magnitudes aleatorias.

Las leyes del proceso aleatorio $\xi(t)$, $t \in T$ se determinan por las distribuciones conjuntas de probabilidades de sus valores $\xi(t_1), \ldots$ $\ldots, \xi(t_n)$ para los distintos $t_1, \ldots, t_n$ (ellas se llaman *distribuciones de dimensión finita* del proceso aleatorio dado).

Por ejemplo, éstas pueden ser distribuciones de Gauss con la densidad de probabilidad (véase (4.10), cap. II)

$$p_{t_1, \ldots, t_n}(x_1, \ldots, x_n) =$$

$$= \frac{1}{(2\pi)^{n/2} |B|^{1/2}} \exp\left\{-\frac{1}{2} \sum_{h,j=}^{n} b_{hj} [x_h - A(t_h)] [(x_j - A(t_j)]\right\},$$

donde $A(t_1), \ldots A(t_n)$ es el valor de la función

$$A(t) = \mathbf{M}\xi(t), \quad t \in T,$$

llamada *valor medio* del proceso aleatorio $\xi(t)$, $t \in T$ y la matriz de correlación $B = \{B(t_h, t_j)\}$ está compuesta de los valores de la función

$$B(s, t)\ \mathbf{M}[\xi(s) - A(s)][\xi(t) - A(t)], \quad s, t \in T$$

llamada *función de correlación* de este proceso (recordemos que los coeficientes $b_{hj}k, j = 1, \ldots, n$ bajo el signo del exponente forman la matriz $\{b_{hj}\}$, inversa a la matriz de correlación $\{B(t_h, t_j)\}$ de las magnitudes aleatorias $\xi(t_1), \ldots, \xi(t_n)$; el propio proceso aleatorio $\xi(t)$, $t \in T$ con distribuciones de Gauss de dimensiones finitas, también se llama *distribución de Gauss*.

Cada valor $\xi(t)$ del proceso aleatorio, siendo una magnitud aleatoria, depende formalmente del resultado elemental ω: $\xi(t) = \xi(\omega, t)$. Examinando el proceso aleatorio $\xi(t)$, $t \in T$, para cada resultado aleatorio aislado ω, nosotros tratamos la función correspondiente $\xi(\omega, \cdot) = \xi(\omega, t)$ del parámetro $t \in T$, que se denomina *trayectoria* o *función seleccionada* del proceso aleatorio dado. Observando realmente el proceso aleatorio, observamos de hecho una de sus posibles trayectorias $x = x(t)$, $t \in T$.

Podemos imaginarnos que se tiene un conjunto determinado X de todas las trayectorias posibles $x = x(t)$, $t \in T$, y un «mecanismo aleatorio» elige una de estas funciones $x \in X$. Con ello la elección aleatoria de tal o cual trayectoria $x = x(t)$, $t \in T$, de X se puede considerar como un resultado elemental. Recordemos que tal punto de vista ya fue expresado antes por nosotros en relación a las magnitudes aleatorias (la magnitud aleatoria n-dimensional se puede considerar como una función aleatoria $\xi(t)$ del parámetro t, que recorre un número finito de valores $t = 1, \ldots, n$). Se indicó precisamente, que examinando las magnitudes $[\xi(1), \ldots, \xi(n)]$, se puede considerar espacio de resultados elementales al espacio X de n dimensiones de vectores $x = [x(1), \ldots, x(n)]$, tomando como resultado elemental aquel valor de $x = [x(1), \ldots, x(n)]$ que toma, según el caso, la magnitud observada $\xi = [\xi(1), \ldots, \xi(n)]$. Con ello el mecanismo aleatorio correspondiente, se da de hecho, por la distribución conjunta de probabilidades de las magnitudes $[\xi(1), \ldots, \xi(n)]$.

2. Los procesos aleatorios de Markov. Se denomina proceso aleatorio *de Markov* $\xi = \xi(t)$ si las magnitudes aleatorias $\xi(t)$, $t \geqslant u$, no dependen de $\xi(s)$, $s \leqslant u$, en cualquier momento de tiempo u para el valor fijado $\xi(u) = x$ (cualquiera que fuese x). Si acordamos considerar $\xi(t)$ como el estado fásico de un sistema físico determinado en el momento de tiempo t, entonces el proceso de Markov $\xi = \xi(t)$, que describe la evolución del sistema examinado, se puede caracterizar más demostrativamente de la siguiente forma: el comportamiento del sistema después de un determinado momento de tiempo u en el estado conocido $x = \xi(u)$ no depende de su comportamiento hasta este momento.

Una propiedad característica de los procesos aleatorios de Markov se ve claramente en un ejemplo tan sencillo como es el juego infantil de «quien va despacio va lejos». En este juego la ficha del jugador deberá pasar un número determinado de puntos 1, 2, . . . El paso de un punto al otro se determina cada vez por el resultado del lanzamiento del dado de juego. Precisamente, si en el paso dado n, la ficha se encuentra en el punto $\xi(n) = i$, por las reglas del juego se establece el punto $\xi(n+1)$ de su desplazamiento en el paso siguiente, en dependencia del número de puntos con que cayó el dado de juego. Desde cualquier punto i la ficha pasa en su paso siguiente al punto j correspondiente con una determinada probabilidad p_{ij}. Se comprende fácilmente que el proceso aleatorio del cambio de estado $\xi(n)$ durante el «tiempo» $n = 0, 1, 2, \ldots$ es un proceso de Markov. (Aquí figura, en lugar del tiempo real, el número n de pasos realizados: se puede considerar convencionalmente, que cada paso se realiza en la unidad de tiempo).

Ejemplo (*fluctuación aleatoria*). Examinemos la fluctuación aleatoria de la partícula por los puntos de números enteros de la recta real, durante la cual la partícula se desplaza en cada paso en una unidad a la derecha con la probabilidad p y en una unidad a la izquierda con la probabilidad $q = 1 - p$. Sea $\xi(n)$ la posición de la partícula después de n pasos. La sucesión $\xi(0), \xi(1), \xi(2), \ldots$ es de Markov: si la partícula se encuentra en un determinado punto i, entonces su futuro comportamiento no depende de las circunstancias, precedentes a la caída en el punto i, y en los n pasos siguientes la partícula pasa al estado correspondiente j con la probabilidad $p_{ji}(n)$.

Está claro, que el paso de i a j, cuando $|i - j| > n$, es imposible y en este caso $p_{ij}(n) = 0$. También está claro, que la partícula durante n pasos, sólo puede pasar a aquellos estados j, para los cuales la diferencia $|i - j|$ tiene la misma paridad que n, es decir, el número

$$m = \frac{n + |j - i}{2}$$

deberá ser entero. Para $j \geqslant i$ se puede caer en el estado j entonces y solamente entonces cuando entre todos los n pasos se realizan justamente $m = \frac{n + (j - i)}{2}$ pasos en dirección positiva. La probabilidad de esto es (véase (3.1), cap. II)

$$p_{ij}(n) = C_n^m p^m q^{n-m}, \quad j \geqslant i. \tag{1.1}$$

Análogamente se expresa la probabilidad del paso de i a j, para $j \leqslant i$:

$$p_{ij}(n) = C_n^m p^{n-m} q^m, \quad j \leqslant i.$$

Ejemplo (*desintegración radiactiva*) (véanse las págs. 87, 90). Consideraremos como anteriormente, que cada átomo Ra independiente de las circunstancias precedentes se transforma en el tiempo t en átomo Rn con la probabilidad $p(t) = 1 - e^{-\lambda t}$ $\left(\lambda = \frac{\log 2}{T}\right.$, T es la constante de semidesintegración$\left.\right)$. Entonces, el número total $\nu(t)$ de átomos de Ra desintegrados en el tiempo t, igual al número de partículas α emitidas en este período de tiempo, está distribuido según la ley de Poisson:

$$\mathbf{P}\{\nu(t) = k\} = \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, 2, \ldots,$$

donde $a = np(t)$, n es el número inicial de átomos de Ra. El número de átomos restantes de Ra será $\xi(t) = n - \nu(t)$. Si es conocida la cantidad de radio en un determinado momento s: $\xi(s) = i$, entonces, independientemente del carácter del proceso de desintegración hasta el momento s, se emiten k partículas α en intervalo desde s hasta t con la probabilidad $\frac{a^k}{k!}e^{-a}$, $a = ip(t-s)$. De tal modo, $\xi(t)$ es un proceso aleatorio de Markov, con ello la probabilidad de paso del estado $\xi(s) = i$ al estado $\xi(t) = j$ $(t > s)$ es

$$p_{ij}(t-s) = \frac{a^{i-j}}{(i-j)!}e^{-a}, \quad a = ip(t-s), \quad j \leqslant i \tag{1.2}$$

(para $j > i$, es imposible el paso de i a j).

Ejemplo (*proceso de Poisson*). Examinemos la corriente de sucesos de Poisson, en la cual el número de sucesos durante un determinado intervalo de tiempo (s, t) no depende del número de sucesos ocurridos hasta el momento s (véase (2.5) y más adelante, cap. II). Está claro, que el proceso aleatorio de la variación $\xi(t)$, del número de sucesos ocurridos hasta el momento actual t, es proceso de Markov, con ello la probabilidad del paso del estado $\xi(s) = i$ al estado $\xi(t) = j$ $(t > s)$ es

$$p_{ij}(t-s) = \frac{(\lambda(t-s))^{j-i}}{(j-i)!}e^{-\lambda(t-s)} \quad j \leqslant i \tag{1.3}$$

(para $j < i$, es imposible el paso de i a j).

Ejemplo (*movimiento browniano*). El proceso del movimiento browniano fue descrito en el p. 2 del § 3, cap. II). Como se ve de esta descripción (especialmente, si nos volvemos al modelo discreto), el movimiento de la partícula browniana desde cualquier estado $\xi(s) = x$ no depende de su comportamiento hasta el momento s (en el modelo discreto ella se traslada en el siguiente paso a uno de los estados $x + \Delta x$, $x - \Delta x$ con la misma probabilidad), con ello, la distribución de probabilidades de la magnitud $\xi(t)$, o sea, las coordenadas de la partícula browniana en el momento de tiempo $t > s$, para la condición de que $\xi(s) = x$, tiene la densidad

$$p(t-s, x, y) = \frac{1}{\sqrt{2\pi(t-s)}\,\sigma}e^{-(y-x)^2/2(t-s)\sigma^2}, \quad -\infty < y < \infty. \tag{1.4}$$

Examinemos el proceso aleatorio de Markov $\xi = \xi(t)$, considerando, como convenimos anteriormente, en que él describe el comportamiento de un determinado sistema físico, en el transcurso del tiempo t, y $\xi(t)$ significa el estado fásico de este sistema en el momento t.

Supongamos que se tiene un número finito o numerable de estados fásicos distintos. Si el sistema se encuentra en el estado $\xi(s) = x$, en el momento de tiempo s, entonces después del tiempo $t - s$ se encontrará en tal ó cual estado y con una probabilidad determinada, que designamos por $P(s, x, t, y)$; notemos, que para el proceso de Markov la probabilidad indicada no depende de su comportamiento hasta el momento s. Formalmente, $P(s, x, t, y)$ significa la probabilidad condicional de que $\xi(t) = y$ para la condición $\xi(s) = x$, $s \leqslant t$. Se puede decir que $P(s, x, t, y)$ es la probabilidad del paso desde el estado x, en que se encuentra el sistema en el momento s, al estado y durante el tiempo $t - s$.

En general, en el caso del espacio arbitrario de estados, se introducen las llamadas *probabilidades de paso* $P(s, x, t, B)$, o sea, las probabilidades del paso del estado inicial $x = \xi(s)$ a uno de los estados y del conjunto indicado de estados B durante el tiempo $t - s$ (formalmente, $P(s, x, t, B)$ significa la probabilidad condicional de que $\xi(t) \in B$ para la condición $\xi(s) = x$, $s \leqslant t$). Por ejemplo si sólo se tiene un conjunto finito o numerable de estados, entonces

$$P(s, x, t, B) = \sum_{y \in B} P(s, x, t, y);$$

si $\xi(t)$ es un vector aleatorio en el espacio de n dimensiones E^n, y su distribución de probabilidades en condiciones que $\xi(s) = x$ tiene la densidad $p(s, x, t, y)$, $y \in E^n$, entonces

$$P(s, x, t, B) = \int_B p(s, x, t, y)\, dy$$

(la densidad condicional de probabilidad $p(s, x, t, y)$, $y \in E^n$, la denominan corrientemente *densidad de paso* del proceso de Markov $\xi = \xi(t)$).

Las probabilidades de paso $P(s, x, t, B)$ del proceso de Markov en el espacio vectorial de n dimensiones E^n, dan de hecho las distribuciones de probabilidades del incremento $\xi(t) - \xi(s)$ para el valor fijado $\xi(s) = x$; si designamos, precisamente, por $B + x$ al conjunto de puntos $y \in E^n$, tales que $y - x \in B$, entonces para la condición $\xi(s) = x$, $\xi(t) - \xi(s) \in B$ con la probabilidad $P(s, x, t, B + x)$. Con respecto a esto señalemos una clase importante de *procesos aleatorios con incrementos independientes*, que poseen tales propiedades que los incrementos $\xi(t) - \xi(t_0)$ son independientes de las magnitudes $\xi(s)$, $s \leqslant t_0$ (cualesquiera que sean $s \leqslant t_0 \leqslant t$). Como ejemplo de procesos unidimensionales de tal tipo puede servir el proceso de Poisson o el movimiento browniano y conocidos por nosotros.

El proceso de Markov $\xi = \xi(t)$ se llama *homogéneo* si la regularidad de su comportamiento en cualquier intervalo (s, t) para el estado determinado $\xi(s) = x$, no depende de la situación del intervalo

(s, t) en el eje del tiempo; para las probabilidades de paso $P(s, x, t, B)$, esto significa de hecho que sólo depende de la diferencia $t - s$, y no de s y t, como en el caso general:

$$P(s, x, t, B) = P(t - s, x, B)$$

(todos los procesos de Markov expuestos anteriormente en los ejemplos, son homogéneos).

Naturalmente, la esencia de la definición no depende de las designaciones que, para comodidad de anotación, pueden ser cambiadas (compárense las probabilidades de paso en los ejemplos (1.1)—(1.3) o la densidad de paso en el ejemplo (1.4)).

En los ejemplos examinados anteriormente, los procesos de Markov, poseen la tal llamada propiedad rigurosa de Markov, que se puede describir en términos generales del siguiente modo: al caer en tal o cual estado x, el comportamiento ulterior del proceso (para el valor dado de x) no depende de su comportamiento en el pasado.

Precisemos este enunciado, utilizando el concepto del momento aleatorio de tiempo independiente del futuro Así se llama cualquier magnitud τ acerca de la cual se puede unívocamente decir, para cualquier t de la trayectoria $\xi(s)$, $s \leqslant t$, si es $\tau \leqslant t$ o es $\tau > t$ (τ también se llama *momento de tiempo de Markov*). Como ejemplo de magnitudes independientes del futuro puede servir el momento τ de la primera caída en tal o cual estado x, y en general, el momento de caída en un estado determinado x, después de algún momento τ_0 de Markov.

Para aclaración, exponemos el ejemplo del momento aleatorio τ, dependiente del futuro y que no es de Markov: $\tau = \tau_1$ si el momento τ_1 de la primera caída después de t_0, en un determinado estado x es menor que $t = t_0 + h$, y $\tau = \tau_1 - h$, si $\tau_1 > t$; está claro que en la trayectoria $\xi(s)$, $t_0 \leqslant s \leqslant t$, hablando en general, no se puede determinar si es $\tau \leqslant t$ (es decir, $\tau_1 \leqslant t + h$) o no. El proceso de Markov $\xi = \xi(t)$ se llama *proceso riguroso de Markov*, si para cualquier momento de tiempo de Markov τ, a condición de que $\xi(\tau) = x$, donde x es el estado conocido, el comportamiento del proceso para $t > \tau$ no depende de su comportamiento anterior al momento τ, con ello el paso desde $x = \xi(\tau)$ a tal o cual estado y durante el tiempo h se realiza con las mismas probabilidades que en el caso del momento fijado $\tau = s$; más exacto: la probabilidad condicional durante el tiempo h, resulta ser en el conjunto B igual a $P(\tau, x, \tau + h, B)$, donde $P(s, x, t, B)$ es la probabilidad de paso del proceso examinado de Markov.

Al hablar en lo sucesivo sobre el proceso de Markov $\xi = \xi(t)$ (véanse las cadenas de Markov, etc.) sobreentenderemos que se cumple la propiedad rigurosa de Markov descrita anteriormente.

§ 2. CADENAS DE MARKOV. CLASIFICACION DE LOS ESTADOS. DISTRIBUCIONES ESTACIONARIAS

1. Probabilidades de paso. Como convenimos anteriormente, al examinar el proceso aleatorio $\xi = \xi(t)$, hablaremos del sistema, cuyo estado fásico en el momento de tiempo t es $\xi(t)$.

Sea $\xi(t)$, $t \geqslant 0$ un proceso homogéneo de Markov con un número finito o numerable de estados para los cuales se pueden tomar los números naturales $i = 1, 2, \ldots$

Supongamos que el parámetro t toma sólo valores de números enteros $t = 0, 1, 2, \ldots$; entonces, tratamos sobre la cadena de pasos

$$\xi(0) \to \xi(1) \to \xi(2) \to \ldots$$

El proceso $\xi = \xi(t)$ del tipo descrito se llama corrientemente *cadena de Markov.*

Supongamos que están dadas las probabilidades de paso p_{ij}, o sea, las probabilidades de paso del sistema del estado i al estado j en un paso, en otras palabras, la probabilidad de que el sistema se encuentre en tal o cual estado $j = 1, 2, \ldots$ en condiciones de que el sistema estuviese en el paso anterior en el estado i:

$$p_{ij} = \mathbf{P}\{\xi(n) = j \mid \xi(n-1) = i\} \left(\sum_j p_{ij} = 1; \quad i = 1, 2, \ldots\right).$$

Supongamos también que el sistema se encuentra en el momento inicial en uno de los estados i con la probabilidad correspondiente p_i^0, $\sum_i p_i^0 = 1$ (en particular, si $p_i^0 = 1$, $p_k^0 = 0$ para $k \neq i$, entonces, esto significa que i es el estado inicial). ¿Cuál es la probabilidad de que el sistema se encuentre después de n pasos, en tal o cual estado $j = 1, 2, \ldots$?

Designemos a esta probabilidad por $p_j(n)$: $p_j(n) = \mathbf{P}\{\xi(n) = j\}$. Después de $n - 1$ pasos el sistema se encontrará obligatoriamente en uno de los estados $k = 1, 2, \ldots$, con esto se encontrará en k con la probabilidad designada por $p_k(n-1)$. Para la condición de que después de $n - 1$ pasos se encuentre el sistema en el estado k, la probabilidad de encontrarse después de n pasos en el estado j será igual a la probabilidad p_{kj} de paso k a j. Utilizando la fórmula de la probabilidad total, obtenemos

$$\mathbf{P}\{\xi(n) = j\} = \sum_k \mathbf{P}\{\xi(n) = j \mid \xi(n-1) = k\} \mathbf{P}\{\xi(n-1) = k\}.$$

Esto da la siguiente relación recurrente para las probabilidades $p_j(n)$, $j = 1, 2, \ldots$:

$$p_j(0) = p_j^0, \quad p_j(n) = \sum_k p_k(n-1) \cdot p_{kj}, \quad (n = 1, 2, \ldots). \quad (2.1)$$

Si el sistema se encuentra en el momento inicial en un estado determinado i, entonces la distribución inicial de probabilidades es $p_i^0 = 1$, $p_k^0 = 0$ para $k \neq i$ y la probabilidad $p_j(n)$ coincide con la probabilidad $p_{ij}(n)$ de que el sistema pase del estado i, en n pasos, al estado j:

$$p_{ij}(n) = \mathbf{P}\{\xi(n) = j \mid \xi(0) = i\}, \quad i, j = 1, 2, \ldots$$

Para la distribución inicial de la forma $p_i^0 = 1$, $p_k^0 = 0$, para $k \neq i$, la fórmula (2.1) da las siguientes relaciones entre *las probabilidades de paso* $p_{ij}(n)$; $i, j = 1, 2, \ldots$:

$$p_{ij}(0) = \begin{cases} 1 & \text{para } j = i \\ 0 & \text{para } j \neq i \end{cases} \qquad p_{ij}(n) = \sum_k p_{ik}(n-1)\, p_{kj}, \qquad (2.2)$$
$$(n = 1, 2, \ldots).$$

Es cómodo introducir la matriz $P(n) = \{p_{ij}(n)\}$. De acuerdo a la fórmula (2.2)

$$P(0) = I, \quad P(1) = P, \quad P(2) = P(1)\,P = P^2, \ldots,$$

donde I es la matriz unitaria, $P = \{p_{ij}\}$ es la matriz de las probabilidades de paso. Vemos que tiene lugar la siguiente igualdad

$$P(n) = P^n \quad (n = 0, 1, 2, \ldots).$$

Conociendo las probabilidades de paso $p_{ij}(n)$ (véanse (2.1), (2.2)), se puede determinar fácilmente, *la distribución de dimensiones finitas* de la cadena de Markòv examinada. Precisamente, para cualesquiera momentos de tiempo $n_1 \leqslant n_2 \leqslant \ldots \leqslant n_k$ y de estados $i_1, i_2, \ldots, i_k$ en condiciones de que $\xi(n_1) = i_1, \ldots, \xi(n_{k-1}) = i_{k-1}$, la probabilidad de encontrarse en el estado i_k en el momento n_k será igual a $p_{i_{k-1}i_k}(n_k - n_{k-1})$, o de otro modo

$$\mathbf{P}\{\xi(n_1) = i_1, \ldots, \xi(n_k) = i_k \mid \xi(n_1) = i_1, \ldots,$$
$$\xi(n_{k-1}) = i_{k-1})\} = p_{i_{k-1}i_k}(n_k - n_{k-1});$$

de donde

$$\mathbf{P}\{\xi(n_1) = i_1, \ldots, \xi(n_k) = i_k\} =$$
$$= \mathbf{P}\{\xi(n_1) = i_1, \ldots, \xi(n_{k-1}) = i_{k-1}\} \cdot p_{i_{k-1}i_k}(n_k - n_{k-1});$$

y, análogamente,

$$\mathbf{P}\{\xi(n_1) = i_1, \ldots, \xi(n_{k-1}) = i_{k-1}\} = \\ = \mathbf{P}\{\xi(n_1) = i_1, \ldots, \xi(n_{k-2}) = i_{k-2}\} p_{i_{k-2}i_{k-1}}(n_{k-1} - n_{k-2}),$$

. . . .

Como resultado obtenemos la fórmula siguiente:

$$\mathbf{P}\{\xi(n_1) = i_1, \xi(n_2) = i_2, \ldots, \xi(n_k) = i_k\} = \\ = p_{i_1}(n_1) \cdot p_{i_1 i_2}(n_2 - n_1) \cdot \ldots \cdot p_{i_{k-1} i_k}(n_k - n_{k-1}). \quad (2.3)$$

Examinemos el proceso homogéneo de Markov $\xi = \xi(t)$, cuyos estados varían durante el tiempo c o n t i n u o t, $t \geqslant 0$. Designemos por ξ_n el n-ésimo estado, según el recuento, y por τ_n, el momento de caída en el estado ξ_n (de otro modo, τ_n es el tiempo hasta el n-ésimo paso $\xi_{n-1} \to \xi_n$, $n = 1, 2, \ldots$); supongamos que $\tau_0 = 0$.

Está claro que el momento τ_n no depende del futuro, y se ve fácilmente, que gracias a la propiedad rigurosa de Markov del proceso inicial $\xi(t)$, el proceso de los pasos de estado en estado, descrito por la cadena

$$\xi_0 \to \xi_1 \to \xi_2 \to \ldots,$$

es una cadena de Markov. Hallemos las probabilidades de paso π_{ij} de esta cadena de Markov $\xi_n = \xi(\tau_n)$, $n = 0, 1, \ldots$

Hallemos al principio las distribuciones de probabilidades de las magnitudes $\tau_{n+1} - \tau_n$ para cada uno de los estados posibles $\xi_n = 1, 2, \ldots$ Debido a la homogeneidad del proceso $\xi = \xi(t)$, para la condición $\xi_n = i$, la distribución $\tau_{n+1} - \tau_n$ es igual que la distribución $\tau = \tau_1$ para un mismo estado inicial $i = \xi_0$ (τ no es otra cosa que el tiempo hasta el cambio del estado i).

Señalemos que, de acuerdo con la propiedad rigurosa de Markov, el comportamiento del proceso $\xi = \xi(t)$ después del momento τ_n, para cualquier $i = \xi(\tau_n)$, no depende de su comportamiento hasta este momento, y en particular, el tiempo $\tau_{n+1} - \tau_n$ de estancia en los estados $i = \xi(\tau_n)$ no depende de las magnitudes $\tau_{k+1} - \tau_k$, $k \leqslant n - 1$ (de donde, a propósito, no es difícil deducir que las diferencias indicadas $\tau_{n+1} - \tau_n$, $n = 0, 1, \ldots$, son magnitudes aleatorias independientes).

Supongamos que, para algún valor de s, el proceso se encuentra en un estado determinado i ($\xi(s) = i$). Examinemos el tiempo τ (después del momento s) hasta el momento de cambio de este estado i. Si $\tau > t$ y, en particular, el estado del proceso en el momento de tiempo $s_1 = s + t$ es el mismo: $\xi(s_1) = i$, entonces el tiempo τ_1 (después de s_1) hasta el momento de cambiar el estado, tiene una distribución de probabilidades igual que τ, ya que gracias a la homogeneidad del proceso, su comportamiento después del momento s_1 se somete a las mismas leyes, que después del momento s, (en ambos casos el estado inicial es i). Esto significa que

$$\mathbf{P}\{\tau > u + v \mid \xi(s) = i, \quad \tau > u\} = \\ = \mathbf{P}\{\tau_1 > v \mid \xi(s_1) = i\} = \mathbf{P}\{\tau > v \mid \xi(s) = i\},$$

por que el suceso $\{\xi(s) = i, \tau > u\}$ incluye en sí al suceso $\{\xi(s_1) = i\}$, y con la condición de que $\xi(s_1) = i$, el comportamiento del proceso de Markov $\xi(t)$ no depende de su comportamiento hasta el momento s_1. Aquí obtenemos, como al deducir la fórmula (2.8), cap. II, que *el tiempo τ de espera hasta el cambio de estado tiene la distribución exponencial de probabilidades*:

$$\mathbf{P}\{\tau > t \mid \xi(s) = i\} = e^{-\lambda_i t}, \quad t \geqslant 0, \tag{2.4}$$

donde λ_i es una constante determinada no negativa ($1/\lambda_i = \mathbf{M}\tau$ es el tiempo medio de espera); *la constante λ_i se llama densidad de salida* desde el estado correspondiente i.

Para $\lambda_i = 0$, el proceso $\xi(t)$ se queda para siempre en el estado i; tal estado se llama *absorbente*; para $0 < \lambda_i < \infty$ la probabilidad de que el estado i se cambie en el pequeño intervalo Δt es

$$\lambda_i \Delta t + o(\Delta t), \tag{2.5}$$

donde $o(\Delta t)$ es un infinitamente pequeño de orden superior en relación a Δt, $\Delta t \to 0$.

Volviendo a la cuestión sobre las probabilidades de paso π_{ij}, supongamos que el paso del proceso $\xi(t)$ del estado i al estado j ($j \neq i$) en el tiempo pequeño Δt es

$$\lambda_{ij} \Delta t + o(\Delta t),$$

donde $o(\Delta t)/\Delta t \to 0$ para $\Delta t \to 0$.

Si la densidad de salida λ_i es igual a cero (el estado i es absorbente), el sistema después de caer en i, se queda allí para siempre, es decir, $\pi_{ij} = 0$ para todos los $j \neq i$; y si $\lambda_i \neq 0$, el sistema sale del estado i con la probabilidad 1 después de un tiempo finito τ (recordemos, que $\mathbf{P}\{\tau < t\} = 1 - e^{-\lambda_i t}$). Para la condición $\tau > t$ tenemos $\xi(t) = i$ y la probabilidad de paso en el intervalo $(t, t + \Delta t)$ será, según la suposición, $\lambda_{ij}\Delta t + o(\Delta t)$; por consiguiente, la probabilidad incondicional $\pi_{ij}(t, t + \Delta t)$ de tal paso es

$$\pi_{ij}(t; t + \Delta t) = [\lambda_{ij}\Delta t + o(\Delta t)] \cdot \mathbf{P}\{\tau > t\} = \\ = [\lambda_{ij}\Delta t + o(\Delta t)]\, e^{-\lambda_i t}.$$

Dividiendo el semieje $0 \leqslant t < \infty$ con los puntos $t = k\Delta t$, $k = 0, 1, \ldots$, tenemos

$$\pi_{ij} = \sum_{k=0}^{\infty} \pi_{ij}(k\Delta t, (k+1)\Delta t) = \\ = \sum_{k=0}^{\infty} [\lambda_{ij}\Delta t + o(\Delta t)]\, e^{-\lambda_i k\Delta t} \sim \lambda_{ij} \int_0^{\infty} e^{-\lambda_i t}\, dt,$$

de donde obtenemos

$$\pi_{ij} = \lambda_{ij} \int_0^\infty e^{-\lambda_i t}\, dt = \frac{\lambda_{ij}}{\lambda_i}, \qquad j \neq i. \tag{2.6}$$

(Recordemos que examinamos el proceso $\xi_n = \xi(\tau_n)$, $n = 0, 1, \ldots$, de los pasos sucesivos, en el cual cada estado siguiente j es diferente del estado anterior i, de modo que $\pi_{ii} = 0$).

2. Estados reversibles e irreversibles. Examinemos la cadena de Markov con los estados 1, 2, . . . y con las probabilidades $p_{ij}(n)$ de paso en n pasos, del estado i al estado j; $i, j = 1, 2, \ldots$

Supongamos que en el momento inicial el sistema se encuentra en un estado determinado i; Designemos por v_n la probabilidad de que el sistema vuelva por primera vez al estado inicial i, exactamente después de n pasos. El valor

$$v = \sum_{n=1}^{\infty} v_n$$

es la probabilidad de que el sistema caiga en el estado inicial i en un paso cualquiera $n = 1, 2, \ldots$, de otro modo, v es *la probabilidad de reversibilidad* a i.

El estado i se llama *reversible* si la probabilidad de volver a él es igual a 1, e *irreversible* si esta probabilidad es menor de 1.

Teorema 1. *El estado i es reversible, entonces y sólo entonces, cuando*

$$\sum_{n=1}^{\infty} p_{ii}(n) = \infty. \tag{2.7}$$

Demostración. Tiene lugar la igualdad siguiente:

$$u_n = u_0 v_n + u_1 v_{n-1} + \ldots + u_{n-1} v_1 + u_n v_0, \qquad n = 1, 2, \ldots, \tag{2.8}$$

donde se supone que $u_n = p_{ii}(n)$ y se han introducido, complementariamente, $u_0 = 1$ y $v_0 = 0$. Esta es un corolario consecuencia de la fórmula general de la probabilidad total. Si introducimos, precisamente, los sucesos B_k, o sea, «el sistema vuelve al estado inicial i después de k pasos», $k = 1, \ldots, n$ y el suceso B_{n+1}, o sea, «el sistema no cae ni una vez en el estado i durante los n primeros pasos», entonces $B_1, \ldots, B_{n+1}$ será un sistema completo de sucesos, y la probabilidad del suceso A, o sea, «después de n pasos el sistema se encontrará en el estado inicial i», según la fórmula de la probabilidad total es

$$\mathbf{P}(A) = \sum_{i=1}^{n+1} \mathbf{P}(A \mid B_k) P(B_k),$$

donde

$$\mathbf{P}(B_k) = v_k; \; \mathbf{P}(A \mid B_k) = u_{n-k}, \; k = 1, \ldots, n; \; P(A \mid B_{n+1}) = 0$$

Consideremos las llamadas *funciones productoras*

$$U(z)=\sum_{k=0}^{\infty} u_k z^k, \quad V(z)=\sum_{k=0}^{\infty} v_k z^k, \quad |z| \leqslant 1.$$

Evidentemente, para $|z|<1$ éstas son funciones analíticas. La relación (2.8), justa para todos los valores $n = 1, 2, \ldots$ se puede escribir en la forma

$$U(z)-u_0 = U(z)\,V(z), \quad u_0 = 1,$$

de donde

$$U(z)=\frac{1}{1-V(z)}. \tag{2.9}$$

La reversibilidad del estado i significa que

$$v=\sum_{k=0}^{\infty} v_k = \lim_{z\to 1} V(z) = 1.$$

Como se ve de la igualdad (2.9) esta relación límite es equivalente a que

$$\lim_{z\to 1} U(z) = \lim_{z\to 1} \frac{1}{1-V(z)} = \infty.$$

Pero

$$\lim_{z\to 1} U(z) = \lim_{z\to 1} \sum_{n=0}^{\infty} u_n z^n = \sum_{n=0}^{\infty} u_n,$$

y de tal modo, la reversibilidad del estado i es equivalente a que la serie $\sum_{n=0}^{\infty} u_n$ sea divergente. Para terminar la demostración sólo queda por recordar que $u_n = p_{ii}(n)$.

Teorema 2. *Si el estado inicial i es reversible, el sistema vuelve a i un número infinito de veces, durante un número infinito de pasos con la probabilidad* 1. *Si este estado es irreversible, el sistema durante un número infinito de pasos estará en el estado i solamente un número finito de veces con la probabilidad* 1, *en otras palabras, después de cierto número finito de pasos el sistema no volverá nunca al estado i.*

D e m o s t r a c i ó n. Designemos por ν_1 el número de pasos hasta el primer regreso al estado i, por ν_2 el número de pasos hasta el segundo regreso, etc. Si durante un número infinito de pasos se realizan menos de k regresos, entonces suponemos que $\nu_k = \infty$. El suceso $\{\nu_k < \infty\}$ significa que por lo menos se realizaron k regresos. La probabilidad del regreso es $\mathbf{P}\{\nu_1 < \infty\} = v$. En condiciones de realización del suceso $\{\nu_1 < \infty\}$, el sistema después de un número finito ν_1 de pasos regresa al estado inicial i, después de lo cual su futuro comportamiento se somete a las mismas leyes, como si él sólo comenzara su movimiento. De tal forma, la probabilidad del suceso $\{\nu_2 < \infty\}$ en condiciones de que $\{\nu_1 < \infty\}$, también será igual a v:

$$\mathbf{P}\{\nu_2 < \infty \mid \nu_1 < \infty\} = v,$$

Evidentemente, si $\nu_1 = \infty$, también será $\nu_2 = \infty$. Por eso

$$\mathbf{P}\{\nu_2 < \infty\} = \mathbf{P}\{\nu_2 < \infty \mid \nu_1 < \infty\}\cdot\mathbf{P}\{\nu_1 < \infty\} = v^2.$$

De modo absolutamente análogo, para cualquier k,

$$\mathbf{P}\{\nu_k < \infty \mid \nu_{k-1} < \infty\} = v, \quad \mathbf{P}\{\nu_k < \infty\} = v^k.$$

La irreversibilidad del estado i significa, que $v < 1$. En este caso

$$\sum_{k=1}^{\infty} \mathbf{P}\{\nu_k < \infty\} = \sum_{k=1}^{\infty} v^k < \infty$$

y, de acuerdo al lema de Borel—Cantelli (véase la pag. 63), puede ocurrir sólo un número finito de sucesos del tipo $\{\nu_k < \infty\}$ con la probabilidad 1; es decir, durante un número infinito de pasos el sistema puede encontrarse en el estado i con la probabilidad 1, sólo un número finito de veces.

La reversibilidad del estado i significa que $v = 1$. En este caso, para cualquier k

$$\mathbf{P}\{\nu_k < \infty\} = 1.$$

Designemos por $\varkappa$ al número de regresos durante un número infinito de pasos. Evidentemente, el suceso $\{\varkappa \geqslant k\}$ es idéntico al suceso $\{\nu_k < \infty\}$, y $\{\varkappa = \infty\} = \bigcap\limits_k \{\nu_k < \infty\}$. Debido a la continuidad de la probabilidad

$$\mathbf{P}\{\varkappa = \infty\} = \lim_{k\to\infty} \mathbf{P}\{\nu_k < \infty\} = 1.$$

El teorema queda demostrado.

Ejemplo (fluctuación aleatoria). Examinemos la fluctuación aleatoria de una partícula por los puntos de números enteros de la recta real, por la cual la partícula se desplaza a 1 en dirección positiva con probabilidad p y en dirección negativa con probabilidad $q = (1 - p)$. Según la fórmula (1.1)

$$p_{ii}(2n) = \frac{(2n)!}{n!\,n!}[pq]^n \sim \frac{[4pq]^n}{\sqrt{\pi n}},$$

y en la fluctuación aleatoria simétrica, cuando $p = q = 1/2$, $p_{ii}(2n) \sim 1/\sqrt{\pi n}$ y la serie $\sum\limits_n p_{ii}(2n)$ es divergente, todos los estados son reversibles. Para $p \neq q$, cuando $4pq < 1$ y $\sum\limits_n p_{ii}(2n) < \infty$, todos los estados son irreversibles.

Hallemos la probabilidad del regreso. Supongamos que en el punto 0 se tiene la tal llamada pantalla absorbente: al caer la partícula en el punto 0 se queda allí para siempre. Evidentemente, desde cualquier estado $i > 0$ se puede caer en cualquier estado $j > 0$ con probabilidad positiva. Designemos por v_{ij} la probabilidad de caer alguna vez desde i en j. Aquí tiene lugar la siguiente relación:

$$v_{ij} = pv_{i+1,\,j} + qv_{i-1,\,j},$$

que nos da la fórmula de la probabilidad total. Esta relación representa en sí una ecuación en diferencias finitas para v_{ij} como función de $i = 1, 2, \ldots$. Con objeto de determinar la solución v_{ij} de esta ecuación, son necesarias además las «condiciones límites». Ellas se pueden hallar de las siguientes consideraciones. Las probabilidades v_{ij} para $i \neq j$ no se cambian si en el punto j ponemos también una pantalla absorbente. En este caso, evidentemente, tendremos $v_{0j} = 0$ y $v_{jj} = 1$. Respondiendo a estas «condiciones límites» la función v_{ij} de $i = 1, 2, \ldots, j-1$ para $p \neq q$ tiene la forma

$$v_{ij} = A_j + B_j\left(\frac{q}{p}\right)^i, \quad 0 < i < j,$$

donde

$$A_j = \frac{1}{1-\left(\frac{q}{p}\right)^j}, \quad B_j = -\frac{1}{1-\left(\frac{q}{p}\right)^j};$$

si es $p = q = 1/2$, entonces

$$v_{ij} = \frac{i}{j}, \quad 0 < i < j.$$

Está claro que tienen lugar fórmulas análogas cuando la pantalla absorbente está en un punto determinado k y $k < i < j$ (en las partes derechas de las fórmulas indicadas sólo es necesario sustituir i por $i - k$ y j por $j - k$). Si $k \to -\infty$, entonces la influencia de la pantalla absorbente desaparece y las fórmulas límites dan la expresión para la probabilidad v_{ij} de paso del punto i al punto $j > i$ en la fluctuación aleatoria corriente sin cualquier pantalla absorbente. Estas fórmulas tienen la forma

$$v_{ij} = \begin{cases} \left(\frac{p}{q}\right)^{j-i} & \text{para} \quad p < 1/2, \\ 1 & \text{para} \quad p \geqslant 1/2. \end{cases} \quad (i \leqslant j).$$

Sustituyendo aquí p por q, obtenemos las expresiones de las probabilidades v_{ij} para $i > j$, a saber:

$$v_{ij} = \begin{cases} \left(\frac{q}{p}\right)^{i-j} & \text{para} \quad p > 1/2, \\ 1 & \text{para} \quad p \leqslant 1/2. \end{cases} \quad (i \geqslant j).$$

Todas estas fórmulas han sido obtenidas con la suposición de que en el punto j está la pantalla absorbente, y naturalmente, para $i = j$, dan el valor $v_{ij} = 1$. La verdadera probabilidad v de regreso al estado inicial $i = j$ puede ser determinada según las probabilidades ya encontradas v_{ij} para $i \neq j$, como

$$v = pv_{i+1,\,i} + qv_{i-1,\,i},$$

lo que da la siguiente expresión para la probabilidad de regreso:

$$v = \begin{Bmatrix} 2p & \text{para} & p < 1/2 \\ 2q & \text{para} & q > 1/2 \\ 1 & \text{para} & p = q \end{Bmatrix} = 1 - |p - q|.$$

3. Tiempo medio de estancia en un estado. Clasificación de los estados. Volvamos a la cadena general de Markov $\xi(n)$, $n = 0, 1, \ldots$ *Supongamos que en el momento inicial de tiempo el sistema se encuentra* en el estado $\xi(0) = i$. Designemos por ν al número de pasos hasta el primer regreso a este estado (si consideramos que cada paso ocupa una unidad de tiempo, entonces ν será el tiempo de regreso a i). Para el estado reversible el tiempo de regreso con probabilidad 1 es naturalmente $\mathbf{P}\{\nu < \infty\} = 1$; para el estado irreversible

$$\mathbf{P}\{\nu < \infty\} = v < 1, \quad \mathbf{P}\{\nu = \infty\} = 1 - v > 0. \tag{2.10}$$

Introduzcamos el valor medio

$$\mu = \mathbf{M}\nu = \sum_{n=1}^{\infty} n\mathbf{P}\{\nu = n\}, \tag{2.11}$$

suponiendo naturalmente, que $\mu = \infty$ para el estado reversible i.

Examinemos al principio el caso cuando $\mu < \infty$. De acuerdo con el teorema 2, en el transcurso del tiempo el sistema regresará de nuevo y de nuevo al estado inicial i. Designemos por ν_1 $(=\nu)$ el momento del primer regreso, por ν_2 el momento del segundo regreso . . ., por ν_n el momento del regreso n al estado i. Como ya indicamos en la demostración del teorema 2, el comportamiento del sistema después del regreso, en uno de los momentos ν_k, al estado i, no depende (para $\xi(\nu_k) = i$) de su comportamiento hasta el momento ν_k, y se somete a las mismas leyes que rigen desde el principio, para el estado inicial $\xi(0) = i$. En particular, las magnitudes $\tau_1 = \nu_1$, o sea, el tiempo del primer regreso, $\tau_2 = (\nu_2 - \nu_1)$, el tiempo del segundo regreso . . ., $\tau_n = (\nu_n - \nu_{n-1})$, el tiempo del n-ésimo regreso al estado inicial i serán magnitudes aleatorias independientes con la misma distribución de *probabilidades, que tienen el valor medio* μ. *Según la ley refor-*

zada de los grandes números (véase (5.7), cap. I)),

$$\frac{1}{n}\sum_{k=1}^{n}\tau_k \to \mu \quad \text{para } n \to \infty \tag{2.12}$$

con probabilidad 1.

En esta relación $\nu_n = \sum_{k=1}^{n}\tau_k$ es el tiempo durante el cual se realizaron exactamente n regresos al estado i (recordemos que según las condiciones cada paso se prolonga en una unidad de tiempo), en otras palabras, n es el tiempo total que el sistema empleó en el estado i desde el intervalo de tiempo 1 hasta el ν_n. Designemos por n_N el tiempo total empleado por el sistema en el estado i durante el intervalo arbitrario desde 1 hasta N ($n = n_N$ es el número de caídas en el estado i durante el tiempo desde 1 hasta N). Ya que $n = n_N \to \infty$ para $N \to \infty$ (véase el teorema 2), entonces de la relación límite (2.12) se deduce que

$$\frac{\nu_n}{n} \to \mu, \quad \frac{n}{n+1} \to 1, \quad \frac{\nu_{n+1}}{n} \to \mu \text{ para } N \to \infty$$

con la probabilidad 1. Evidentemente, debido a la propia definición del número n_N, tenemos $\nu_{n_N} \leqslant N \leqslant \nu_{n_N+1}$ y, por consiguiente,

$$\frac{N}{n_N} \to \mu \quad \text{para } N \to \infty \tag{2.13}$$

con probabilidad 1.

Mostremos ahora que la relación límite (2.13) tiene lugar también en el caso, cuando el valor medio μ, determinado por la igualdad (2.11), es infinito ($\mu = +\infty$). Para el estado irreversible i esto se deduce inmediatamente de la segunda parte del teorema 2 (n_N queda limitado para $N \to \infty$ con probabilidad 1). Para el estado reversible i nos dirigimos de nuevo a las magnitudes independientes $\tau_k = \nu_k - \nu_{k-1}$, $k = 1, 2, \ldots$ Determinemos las magnitudes «cortadas» τ_k^m ($\tau_k^m \leqslant \tau_k$), poniendo

$$\tau_k^m = \begin{cases} \tau_k & \text{para } \tau_k \leqslant m, \\ 0 & \text{para } \tau_k > m. \end{cases}$$

El valor medio

$$\mu^m = \mathbf{M}\tau_k^m = \sum_{n=1}^{m} n\mathbf{P}(\tau_k^m = n)$$

de estas magnitudes (véase (2.11)) para los valores de m suficientemente grandes, se puede hacer tan grande como se quiera, ya que $\sum_{n=1}^{\infty} n\mathbf{P}(\tau_k = n) = \infty$. Según la ley reforzada de los grandes números,

$\lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{k=1}^{m} \tau_k^m = \mu^m$ con probabilidad 1, y por eso

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \tau_k \geqslant \lim_{n\to\infty} \sum_{k=1}^{n} \tau_k^m = \mu^m.$$

De tal modo, el límite inferior $\lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{k=1}^{n} \tau_k$ supera al número tan grande como se quiera μ^m, es decir

$$\frac{1}{n} \sum_{k=1}^{n} \tau_k \to \infty \text{ para } n \to \infty \tag{2.14}$$

con probabilidad 1. Exactamente igual que de (2.12), de la relación (2.14) se deduce (2.13) con $\mu = \infty$.

La relación n_N/N entre las magnitudes n_N y N que figuran en la relación límite (2.13) representa en sí la parte de tiempo, empleado por el sistema en el estado i durante el intervalo de tiempo desde 1 hasta N, y (2.13) es equivalente a la siguiente proposición:

Teorema 3. *Con la probabilidad* 1

$$\frac{n_N}{N} \to \frac{1}{\mu} \quad para \quad N \to \infty \tag{2.15}$$

donde μ *es el tiempo medio de regreso al estado* i.

Se comprende fácilmente que este resultado queda en vigor para cualquier estado i (que no es obligatoriamente inicial) con tal de que el sistema caiga tarde o pronto en este estado (después de lo cual su comportamiento se somete a las mismas leyes que rigen el estado inicial i).

De la relación (2.15) se deduce (véase el teorema 5 del § 4, cap. I) que $1/\mu$ coincide con el límite $\lim\limits_{N\to\infty} \mathbf{M} \left[\frac{n_N}{N}\right]$ de las esperanzas matemáticas de las magnitudes aleatorias limitadas n_N/N. Con objeto de hallar la expresión explícita de la esperanza matemática $\mathbf{M}n_N$ del número de caídas en el estado i durante el tiempo N, introducimos las magnitudes

$$\chi_k = \begin{cases} 1 & \text{para } \xi(k) = i, \\ 0 & \text{en los casos restantes.} \end{cases}$$

Evidentemente,

$$n_N = \sum_{k=1}^{N} \chi_k, \quad \mathbf{M}n_N = \sum_{k=1}^{N} \mathbf{M}\chi_k = \sum_{k=1}^{N} p_{ii}(k). \tag{2.16}$$

En resultado obtenemos que

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} p_{ii}(k) = \frac{1}{\mu}. \tag{2.17}$$

El estado i se llama *positivo* cuando $\frac{1}{\mu} > 0$, y nulo cuando $\frac{1}{\mu} = 0$ (recordemos que de acuerdo con (2.15) la magnitud $1/\mu$ es la parte de tiempo, empleado por el sistema en el estado i en un intervalo infinito de tiempo).

Se dice que el estado j es *accesible* desde i, si durante un determinado número de pasos el sistema pasa desde i hasta j con probabilidad positiva, es decir, $p_{ij}(M) > 0$ para un M dado. Si j es accesible desde i y a su vez i es accesible desde j, entonces estos estados i, j se llaman *comunicantes*.

Sea i un estado reversible y j accesible desde i. Entonces i es a la vez accesible desde j, ya que en caso contrario el sistema saliendo, desde i, después de M pasos, cae en el estado j con probabilidad positiva $p_{ij}(M) = \alpha > 0$, después de lo cual ya no puede regresar a i; de tal modo, la probabilidad del regreso a i no será mayor que $1 - \alpha$, y esto contradice a la reversibilidad del estado i. Así pues, si j es accesible del estado reversible i, entonces a su vez, i es accesible desde j, es decir $p_{ij}(N) = \beta > 0$ para un N dado, es decir, *los estados i y j son comunicantes*.

Para cualesquiera estados comunicantes i, j (tales que $p_{ij}(M) = \alpha > 0$, $p_{ji}(N) = \beta > 0$ para unos determinados M y N) de la relación

$$P(M + K + N) = P(M) P(K) P(N) = P(N) P(K) P(M)$$

(donde $P(n) = \{p_{kl}(n)\}$ significa la matriz de la probabilidad de paso $p_{kl}(n) \geqslant 0$, véase (2.2) y más adelante) deducimos que

$$p_{ii}(K + M + N) \geqslant p_{ij}(M) p_{jj}(K) p_{ji}(N) = \alpha\beta p_{jj}(K),$$
$$p_{jj}(K + M + N) \geqslant p_{ji}(N) p_{ii}(K) p_{ij}(M) = \alpha\beta p_{ii}(K).$$

Estas desigualdades muestran, primeramente que las series

$$\sum_k p_{ii}(k) \qquad \sum_k p_{jj}(k)$$

convergen o divergen al mismo tiempo y segundo, que los límites

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} p_{ii}(k), \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} p_{jj}(k)$$

bien ambos son iguales a cero, o bien ambos son positivos. Teniendo en cuenta el teorema 2, de aquí sacamos la conclusión de que *los estados accesibles desde cierto estado reversible, también son reversibles.*

Evidentemente (véanse (2.7) y (2.17)), los estados comunicantes pertenecen a un mismo tipo: son reversibles o irreversibles, positivos o nulos.

Sea i un estado reversible. Como fue indicado, todos los estados j accesibles desde i son reversibles y se comunican con el estado i. Todos los estados indicados j forman la llamada clase *cerrada* de estados (designémosla por E) tal que si el sistema en un paso cae en uno de los estados $j \in E$, entonces, después de esto, se queda para siempre en el conjunto de estados E (cualquier estado, accesible desde cualquier estado $j \in E$, también entra en E); con ello, para cualquier $j \in E$, se puede determinar la propia clase como un conjunto de todos los estados accesibles desde j.

Examinemos dos clases cerradas cualesquiera de estados reversibles designándolas por E_1 y E_2. Supongamos que un estado j entra simultáneamente en E_1 y en E_2. Entonces, estas clases E_1 y E_2 coinciden, ya que cada una de ellas coincide con el conjunto de todos los estados, accesibles desde j. Por consiguiente, bien coinciden las clases cerradas E_1 y E_2, o bien no se cortan (no tienen estados comunes j).

Separemos el conjunto E_0 de todos los estados irreversibles y dividamos los estados reversibles restantes en clases cerradas que no se cortan E_1, E_2, . . . (cada una de las clases representa en sí un conjunto determinado de estados reversibles que se comunican uno con otro).

Como muestra el ejemplo de la fluctuación aleatoria en la pág. 147, hablando en general, el sistema en el transcurso del tiempo puede marchar «al infinito» por una cadena de estados irreversibles en E_0 (si hay un número infinito de tales estados). Esta posibilidad se excluye, si sólo se tiene un número finito de estados irreversibles. En este caso, el sistema estará en cualesquiera de los estados irreversibles, sólo un número finito de veces con probabilidad 1 y tarde o temprano caerá en un determinado estado reversible i; en el futuro «circulará» en la clase cerrada de estados reversibles E, que contiene i (así, precisamente, ocurre con la cadena con un número finito de estados).

4. Teorema ergódico (convergencia hacia la distribución estacionaria). La distribución de probabilidades p_i^*, $i = 1, 2, \ldots$, se llama distribución *estacionaria* para las cadenas de Markov con probabilidades de paso p_{ij} si se cumple la condición

$$p_j^* = \sum_i p_i^* p_{ij} \quad j = 1, 2, \ldots \tag{2.18}$$

De las relaciones recurrentes generales (2.1) se ve que la condición (2.18) significa lo siguiente: *en la distribución $p_i(n_0) = p_i^*$, $i = 1, 2, \ldots$, las probabilidades $p_j(n)$ de que en el paso n el sistema se encuentre en el estado $j = 1, 2, \ldots$, quedan invariables para todos los valores $n \geqslant n_0$*:

$$p_j(n) = p_i^*, \quad j = 1, 2, \ldots \tag{2.19}$$

Supongamos que los estados de la cadena de Markov examinada son positivos, forman una clase cerrada y además de esto, para cierto n_0, es positivo *el coeficiente de ergodicidad* $\boldsymbol{k}(n_0)$ que se determina por la fórmula

$$\boldsymbol{k}(n_0)=1-\frac{1}{2}\sup_{i,j}\sum_m |p_{im}(n_0)-p_{jm}(n_0)|.$$

Señalemos que para el coeficiente de ergodicidad $\boldsymbol{k}(n_0)$ tiene lugar la siguiente valuación sencilla

$$\boldsymbol{k}(n_0)\geqslant\sup_j\inf_i p_{ij}(n_0), \tag{2.20}$$

de la cual se deduce que $\boldsymbol{k}(n_0)>0$, si existe aunque sea un sólo estado j, accesible desde cualquier estado i (para un mismo número de pasos n_0) con probabilidad $p_{ij}(n_0)$, $p_{ij}(n_0)\geqslant\delta>0$ para todos los valores $i=1, 2, \ldots$

Señalemos también que la condición $\boldsymbol{k}(n_0)>0$ no siempre se cumple. De ejemplo trivial puede servir la cadena de Markov con dos estados 1 y 2, cuando el sistema cambia periódicamente (a cada paso) su estado (en este caso $p_{11}(n)=1$, $p_{21}(n)=0$, $p_{12}(n)=0$, $p_{22}(n)=1$ para los valores pares de n, $p_{11}(n)=0$, $p_{21}(n)=1$, $p_{12}(n)=1$, $p_{22}(n)=0$ para los n impares y $\sum_k |p_{1k}(n)-p_{2k}(n)|\equiv 2$ para todos los n).

Teorema 4. Si $\boldsymbol{k}(n_0)>0$ para un determinado n_0, entonces existen los límites

$$\lim_{n\to\infty} p_j(n)=p_j^*,\quad j=1, 2, \ldots, \tag{2.21}$$

donde las probabilidades p_j^*, $j=1, 2, \ldots$, forman la distribución estacionaria. Para cualquier distribución inicial p_i^0, $i=1, 2, \ldots$, las probabilidades límites p_j^*, $j=1, 2, \ldots$, son las mismas, con ello tiene lugar la siguiente valuación de la velocidad de convergencia en la relación límite (2.21):

$$\sup_{p_i^0}|p_j(n)-p_j^*|\leqslant Ce^{-Dn} \tag{2.22}$$

$$\left(C=\frac{1}{1-\boldsymbol{k}(n_0)},\ D=\frac{1}{n_0}\ln\frac{1}{1-\boldsymbol{k}(n_0)}\right).$$

Demostración. Admitamos que

$$r_j(n)=\inf_i p_{ij}(n),\quad R_j(n)=\sup_i p_{ij}(n).$$

Tenemos

$$r_j(n+1)=\inf_i p_{ij}(n+1)=\inf_i\sum_k p_{ik}p_{kj}(n)\geqslant\inf_i\sum_k p_{ik}r_j(n)=r_j(n),$$

$$R_j(n+1)=\sup_i p_{ij}(n+1)=\sup_i\sum_k p_{ik}p_{kj}(n)\leqslant\sup_i\sum_k p_{ik}R_j(n)=R_j(n).$$

De tal modo, obtenemos la siguiente cadena de desigualdades:

$$r_j(1) \leqslant r_j(2) \leqslant \ldots \leqslant r_j(n) \leqslant \ldots \leqslant R_j(n) \leqslant \ldots \leqslant R_j(2) \leqslant R_j(1).$$

Para cualesquiera estados α y β

$$\sum_k p_{\alpha k}(n_0) - \sum_k p_{\beta k}(n_0) = \\ = \sum_k^{+} [p_{\alpha k}(n_0) - p_{\beta k}(n_0)] + \sum_k^{-} [p_{\alpha k}(n_0) - p_{\beta k}(n_0)] = 0,$$

donde $\sum_k^{+}$ significa la adición por todos los valores de k, para los cuales la diferencia $p_{\alpha k}(n_0) - p_{\beta k}(n_0)$ es positiva, y $\sum_k^{-}$ es la adición por aquellos valores de k, para los cuales la diferencia $p_{\alpha k}(n_0) - p_{\beta k}(n_0)$ es negativa. Evidentemente,

$$\sup_{\alpha,\beta} \sum_k^{+} [p_{\alpha k}(n_0) - p_{\beta k}(n_0)] = 1 - \boldsymbol{k}(n_0),$$

donde $\boldsymbol{k}(n_0)$ es el coeficiente de ergodicidad. Utilizando las relaciones obtenidas, valoremos la diferencia $R_j(N) - r_j(N)$. Tenemos

$$R_j(n_0) - r_j(n_0) = \sup_{\alpha} p_{\alpha j}(n_0) - \inf_{\beta} p_{\beta j}(n_0) = \\ = \sup_{\alpha,\beta} [p_{\alpha j}(n_0) - p_{\beta j}(n_0)] \leqslant \sup_{\alpha,\beta} \sum_k^{+} [p_{\alpha k}(n_0) - p_{\beta k}(n_0)] = [1 - \boldsymbol{k}(n_0)]$$

y

$$R_j(n_0+n) - r_j(n_0+n) = \sup_{\alpha,\beta} [p_{\alpha j}(n_0+n) - p_{\beta j}(n_0+n)] = \\ = \sup_{\alpha,\beta} \sum_k [p_{\alpha k}(n_0) - p_{\beta k}(n_0)]\, p_{kj}(n) \leqslant \\ \leqslant \sup_{\alpha,\beta} \{\sum_k^{+} [p_{\alpha k}(n_0) - p_{\beta k}(n_0)]\, R_j(n) + \sum_k^{-} [p_{\alpha k}(n_0) - p_{\beta k}(n_0)]\, r_j(n)\} = \\ = \sup_{\alpha,\beta} \{\sum_k^{+} [p_{\alpha k}(n_0) - p_{\beta k}(n_0)]\, [R_j(n) - r_j(n)]\} = \\ = [1 - \boldsymbol{k}(n_0)]\, [R_j(n) - r_j(n)].$$

De aquí se deduce que

$$R_j(Nn_0) - r_j(Nn_0) \leqslant [1 - \boldsymbol{k}(n)]^N, \quad N = 1, 2, \ldots$$

La sucesión $r_j(n)$, $n = 1, 2, \ldots$, crece monótonamente y la sucesión $R_j(n)$, $n = 1, 2 \ldots$ decrece monótonamente, siendo $r_j(n) \leqslant R_j(n)$. La valuación, obtenida anteriormente, de la diferencia $R_j(n) - r_j(n)$ muestra que estas sucesiones tienen un mismo límite p_j^*:

$$p_j^* = \lim_{N\to\infty} r_j(n) = \lim_{N\to\infty} R_j(n).$$

Evidentemente

$$|p_{ij}(n)-p_j^*|\leqslant R_j(n)-r_j(n)\leqslant(1-\boldsymbol{k}(n_0))^{\frac{n}{n_0}-1}.$$

Para cualquier distribución inicial p_j^0, $j=1, 2, \ldots$, tenemos

$$|p_j(n)-p_j^*|=|\sum_i p_i^0 p_{ij}(n)-p_j^*|\leqslant\sum_i p_i^0|p_{ij}(n)-p_j^*|\leqslant$$

$$\leqslant\sum_i p_i^0[R_j(n)-r_j(n)]=R_j(n)-r_j(n)\leqslant(1-\boldsymbol{k}(n_0))^{\frac{n}{n_0}-1}.$$

La valuación obtenida se puede escribir, naturalmente, en la forma (2.22) con las respectivas constantes C y D, suponiendo

$$C=\frac{1}{1-\boldsymbol{k}(n_0)},\quad D=\frac{1}{n_0}\ln\frac{1}{1-\boldsymbol{k}(n_0)}.$$

Mostremos que *las probabilidades límites p_j^* satisfacen al sistema de ecuaciones* (2.18). Para esto, señalemos al principio que la suma $\sum_j p_j^*$ es finita, ya que para cualquier m

$$\sum_{j\leqslant m}p_j^*=\lim_{n\to\infty}\sum_{j\leqslant m}p_j(n)\leqslant 1.$$

Luego, de la desigualdad $p_j(n+1)\geqslant\sum_{i\leqslant m}p_i(n)p_{ij}$ (véase (2.1)) deducimos que $p_j^*\geqslant\sum_{i\leqslant m}p_i^*p_{ij}$ para cualquier valor limitado de m, y, por consiguiente,

$$p_j^*\geqslant\sum_i p_i^*p_{ij},\quad j=1, 2, \ldots \tag{2.23}$$

Sumando todas estas desigualdades, obtenemos

$$\sum_j p_j^*\geqslant\sum_j\left(\sum_i p_i^*p_{ij}\right)=\sum_i p_i^*\left(\sum_j p_{ij}\right)=\sum_i p_i^*.$$

Aquí se ve que debe existir la igualdad rigurosa $\left(\sum_j p_j^*=\sum_j p_j^*\right)$, y por consiguiente, deberá existir la igualdad rigurosa en todas las relaciones (2.23), que es lo que se exigía demostrar.

Establezcamos luego que todas las probabilidades p_j^* son positivas y $\sum_j p_j^*=1$. Evidentemente,

$$p_j^*=\lim_{n\to\infty}p_{ij}(n)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}p_{ij}(n),$$

y, por consiguiente, las probabilidades finales p_j^* para todos los estados positivos serán tales que

$$p_j^*=\frac{1}{\mu_j}>0,\quad j=1, 2, \ldots, \tag{2.24}$$

donde μ_j es el tiempo medio de regreso al estado j (véase (2.17)). Luego, si tomamos en calidad de distribución inicial de probabilidades a

$$p_j^0 = p_i^* \frac{1}{\sum_j p_j^*}, \quad i = 1, 2, \ldots,$$

entonces, ya que las probabilidades p_i^0 se diferencian de p_i^* ($i = 1, 2, \ldots$) sólo en un factor constante $\left(\sum_j p_j^*\right)^{-1}$, éstas satisfacerán al sistema homogéneo de ecuaciones (2.18)

$$p_j^0 = \sum_i p_i^0 p_{ij}, \quad j = 1, 2, \ldots$$

Para la distribución estacionaria p_i^0, $i = 1, 2, \ldots$ tenemos

$$p_i(n) = p_j^0 \quad \text{y} \quad p_j^* = \lim_{n\to\infty} p_j(n) = p_j^0, \quad j = 1, 2, \ldots,$$

de donde se desprende que $\sum_j p_j^* = 1$. De tal modo, las probabilidades límites p_i^*, $i = 1, 2, \ldots$, forman la distribución de probabilidades estacionaria (que evidentemente es única).

Señalemos que las probabilidades límites p_j^*, como cualesquiera probabilidades estacionarias, pueden ser halladas del sistema de ecuaciones lineales (2.18).

El teorema queda demostrado.

Ejemplo (pila de libros). Sobre la mesa del escritorio descansa una pila de m libros. Si designamos a cada libro con un número correspondiente, entonces el orden de su disposición desde arriba abajo se puede describir por la permutación de m números $(i_1, i_2, \ldots, i_m)$, donde i_1 es el número del libro que descansa arriba, i_2 es el número del libro siguiente, ..., i_m es el número del último libro que descansa lo más abajo. Supongamos que cada libro se coge con una probabilidad determinada, digamos, el libro con número k se coge con la probabilidad p_k ($k = 1, \ldots, m$), con ello al devolverlo se coloca encima.

Tomemos un estado arbitrario $(i_1, i_2, \ldots, i_m)$. En el siguiente paso aquél se encuentra o bien invariable, lo que ocurre con probabilidad p_i, al elegir el libro que descansa desde arriba con el número i_1, o bien se cambia por uno de los $m - 1$ estados de la forma $(i_h, i_1 \ldots)$, lo que ocurre con probabilidad p_{i_h} al elegir el libro con el número $i_h \neq i_1$.

Ante nosotros tenemos la cadena de Markov con los estados, cada uno de los cuales se describe por la correspondiente permutación $(i_1, i_2, \ldots, i_m)$ y por las probabilidades de paso indicadas. Si un libro determinado i no se elige nunca ($p_i = 0$), entonces todos los estados $(i_1, \ldots, i_m)$, donde $i_1 = i$ (el libro con el número i descansa arriba), son irreversibles, ya que después del primer paso se elige un libro determinado j, $j \neq i$ y en lo sucesivo el i-ésimo libro que no se elige nunca, desciende hacia abajo. Si cada libro se elige con probabilidad positiva $p_i > 0$, entonces cada estado es accesible desde cualquier otro estado (en total se tienen $m!$ estados diferentes, o sea, las distintas permutaciones $(i_1, \ldots, i_m)$), y todos estos estados comunicantes son positivos, formando una clase cerrada. De cada estado $(i_1, \ldots, i_m)$ se puede pasar después de m pasos a cualquier estado $(j_1, \ldots, j_m)$ con probabilidad superior al producto $p_i \ldots p_m$ (este producto es igual a la probabilidad de paso de $(i_1, \ldots, i_m)$ a $(j_1, \ldots, j_m)$, cuando se elige en el primer paso el libro con el número j_m, en el segundo paso, el libro con el número j_{m-1} y así sucesivamente, en el último paso m se elige el libro con el número j_1). Por consiguiente, se tiene un coeficiente de ergodicidad positivo $\boldsymbol{k}(m)$ ($\boldsymbol{k}(m) \geqslant p_1, p_2, \ldots, p_m$; véase (2.20)) y en el transcurso del tiempo se establece la distribución de probabilidades estacionaria.

Examinemos al principio el caso $m = 2$. Entonces se tienen sólo dos estados (1,2) y (2,1). Las probabilidades de paso tienen la forma

$$p_{11} = p_{21} = p_1, \quad p_{12} = p_{22} = p_2,$$

y la matriz de las probabilidades de paso es

$$P = \begin{pmatrix} p_1 & p_2 \\ p_1 & p_2 \end{pmatrix}.$$

Las probabilidades de paso en dos pasos son

$$p_{11}(2) = p_{21}(2) = p_1p_1 + p_2p_1 = p_1(p_1 + p_2) = p_1,$$
$$p_{12}(2) = p_{22}(2) = p_1p_2 + p_2p_2 = p_2(p_1 + p_2) = p_2.$$

Se ve que $P^2 = P$ y en general, $P^n = P$. Para cualquier distribución inicial de probabilidades tenemos

$$p_1(n) = p_1^0 p_{11}(n) + p_2^0 p_{21}(n) = p_1(p_1^0 + p_2^0) = p_1,$$
$$p_2(n) = p_1^0 p_{12}(n) + p_2^0 p_{22}(n) = p_2(p_1^0 + p_2^0) = p_2.$$

Se ve que la distribución estacionaria se establece ya en el primer paso.

Examinemos el caso del valor arbitrario de m. Designemos por $p_{(i_1, \ldots, i_m), (j_1, \ldots, j_m)}$ la probabilidad del paso del estado $(i_1, \ldots, i_m)$ al estado $(j_1, \ldots, j_m)$. Como fue mostrado,

$$p_{(i_1, \ldots, i_m), (j_1, \ldots, j_m)} = \begin{cases} p_{i_k} & \text{para } (j_1, \ldots, j_m) = (i_k, \ldots), \\ 0 & \text{para los restantes } (j_1, \ldots, j_m), \end{cases}$$

donde la permutación $(i_k, \ldots)$ se obtiene de $(i_1, \ldots, i_m)$ con la elección de un determinado i_k y su permutación en el primer lugar. Las probabilidades estacionarias $p^*_{(j_1, \ldots, j_m)}$ son la solución del siguiente sistema de ecuaciones lineales (véase (2.18)):

$$p^*_{(i_1,\ldots,i_m)}=p_{i_1}\sum_{k=1}^{m}p^*_{(i_2,\ldots,i_{k-1},i_1,i_k,\ldots)}.$$

Es natural preguntar, ¿con qué probabilidad cada uno de los libros existentes resultará estar colocado arriba? ¿Cuándo se establecerá prácticamente, después de un número bastante grande de pasos, la distribución de probabilidades estacionaria (es decir, cuándo ocupará la pila de libros la posición correspondiente $(i_1, \ldots, i_m)$ con unas probabilidades invariables $p^*_{(i_1,\ldots,i_m)}$?

La probabilidad de que el libro con el número i descanse arriba, será evidentemente,

$$p^*_i=\sum_{i_2,\ldots,i_m}p^*_{(i,i_2,\ldots,i_m)},$$

donde la suma se toma según todos los estados, en los cuales está en primer lugar i. De las ecuaciones para las probabilidades estacionarias obtenemos que

$$p^*_i=\sum_{i_2,\ldots,i_m}p_i\sum_k p^*_{(i_2,\ldots,i_{k-1},i,i_k,\ldots,i_m)}=$$
$$=p_i\sum_{i_1,\ldots,i_m}p^*_{(i_1,\ldots,i_m)}=p_i,\quad i=1,\ \ldots,\ m,$$

es decir, la probabilidad p^*_i de que el libro con el número i se encuentre arriba, será igual a la probabilidad p_i, con la cual se elige este libro. De tal modo, cuanto más frecuentemente se tome tal o cual libro, tanto mayor será la probabilidad de que éste se encuentre arriba.

Ejemplo (la fluctuación aleatoria). Examinemos la fluctuación aleatoria, en la cual la partícula con probabilidad p_i pasa del punto i al punto vecino $j=i+1$, y con probabilidad $q_i=1-p_i$, al punto $j=0$. Consideraremos que $0<p_i<1$. Evidentemente, todos los estados son comunicantes y al mismo tiempo son reversibles o irreversibles, nulos o positivos.

Supongamos que la partícula se encuentra en el estado inicial $i=0$. La probabilidad de que después de los n pasos siguientes no regrese ni una vez a la posición inicial $i=0$, será igual al producto $p_0p_1\ldots p_{n-1}$, o sea, la probabilidad de que el sistema recorra sucesivamente la cadena de estados $0\to1\to\ldots\to n$. Se ve fácilmente, que la probabilidad de que éste no regrese ni una vez al estado inicial $i=0$, después de un número infinito de pasos, será igual al producto infinito

$$\prod_{k=0}^{\infty}p_k=\lim_{n\to\infty}p_0p_1\ldots p_n.$$

Si este producto infinito es convergente hacia cero: $\lim_{n\to\infty} p_0p_1 \dots$ $\dots p_n = 0$, entonces el estado $i = 0$ es reversible. En caso contrario la probabilidad de regreso es

$$v = 1 - \lim_{n\to\infty} p_0p_1 \dots p_n < 1,$$

y el estado $i = 0$ es irreversible.

A este resultado se puede llegar por otro camino, examinando directamente las probabilidades v_n de regresar por primera vez al *estado inicial* 0, *exactamente después de* n *pasos. Evidentemente*, la partícula se vuelve por primera vez al estado 0 en el paso n, si ella en los primeros $n - 1$ pasos pasa, sucesivamente, desde el estado $i - 1$ al i (con probabilidades p_{i-1}, $i = 1, \dots, n - 1$), y por eso

$$v_1 = 1 - p_0, \quad v_n = p_0 \dots p_{n-2}(1 - p_{n-1}), \quad n = 2, 3, \dots$$

La probabilidad de regreso al estado inicial 0, según la definición, será igual a la suma $\sum_{n=1}^{\infty} v_n$ y es

$$v = \sum_{n=1}^{\infty} v_n = 1 - p_0 + [p_0(1 - p_1)] + \dots = 1 - \lim_{n\to\infty} p_0p_1 \dots p_n.$$

Para los estados irreversibles, cuando $v < 1$, la partícula se alejará en dirección $+\infty$ para $n \to \infty$, con la probabilidad 1; para los estados reversibles regresará un número ilimitado de veces a cada estado. El tiempo medio de regreso al estado inicial $i = 0$ es

$$\begin{aligned}\mu = \sum_{n=1}^{\infty} nv_n &= (1 - p_0) + 2(1 - p_1)p_0 + 3(1 - p_2)p_0p_1 + \dots \\ &\dots + n(1 - p_{n-1})p_0 \dots p_{n-2} + \dots = \\ &= 1 + p_0 + p_0p_1 + \dots + p_0p_1 \dots p_{n-1} + \dots,\end{aligned}$$

y para la condición

$$\mu = 1 + \sum_{n=1}^{\infty} p_0 \dots p_{n-1} < \infty$$

todos los estados son positivos.

En el caso más sencillo, cuando las probabilidades $1 - p_i$ de paso de todos los estados i al estado inicial 0 son tales que $1 - p_i \geqslant \delta > 0$, $i = 0, 1, \dots$ se puede dar fácilmente la valuación del coeficiente de ergodicidad

$$\boldsymbol{k}(1) > \inf_i (1 - p_i) \geqslant \delta$$

(véase (2.20)), y junto con ésta la valuación de la velocidad de convergencia hacia las probabilidades límites p_j^*, $j = 0, 1, \dots$ (véase (2.22)), que forman la distribución estacionaria. Esta distribución satisface al sistema de ecuaciones (2.18) que en nuestro caso se representa del siguiente modo:

$$p_j^* = p_{j-1}^* p_{j-1}, \quad j = 1, 2, \ldots,$$

de donde

$$p_1^* = p_0^* p_0, \quad p_2^* = p_0^* p_0 p_1, \ldots, p_n^* = p_0^* p_0 \ldots p_{n-1}, \ldots$$

Ya que

$$1 = \sum_{n=0}^{\infty} p_n^* = p_0^* \left[1 + \sum_{n=1}^{\infty} p_0 \ldots p_{n-1} \right] = p_0^* \mu$$

entonces, como resultado tenemos

$$p_0^* = \frac{1}{\mu}, \quad p_1^* = \frac{p_0}{\mu}, \ldots, p_n^* = \frac{p_0 \ldots p_{n-1}}{\mu}, \ldots \tag{2.25}$$

§ 3. CADENAS DE MARKOV (TIEMPO CONTINUO)

1. Ecuaciones diferenciales para las probabilidades de paso. Examinemos el proceso homogéneo de Markov $\xi(t)$ con un número finito y numerable de estados 1, 2, ..., que se diferencia de las cadenas de Markov examinadas en el párrafo anterior, sólo, en que el parámetro t (tiempo) varía continuamente y el paso de un estado a otro es posible en cualquier momento de tiempo t. Igual que anteriormente, hablaremos convencionalmente sobre el sistema físico, cuyo estado fásico en el momento t es $\xi(t)$.

Designemos por $p_{ij}(t)$ la probabilidad de paso del estado i al estado j durante el tiempo t; precisemos que $p_{ij}(t)$ es, según la definición, la probabilidad de encontrarse en el estado j después del tiempo t, con la condición de que i fuese el estado inicial:

$$p_{ij}(t) = \mathbf{P}\{\xi(t+s) = j \mid \xi(s) = i\}, \quad i, j = 1, 2, \ldots$$

Sea p_i^0, $i = 1, 2, \ldots$, la distribución inicial de probabilidades. Análogamente a (2.1), para las probabilidades $p_j(t)$ de encontrarse en el estado correspondiente, $j = 1, 2, \ldots$, después del tiempo t tienen lugar las siguientes fórmulas:

$$p_j(t) = \sum_i p_i^0 p_{ij}(t), \quad j = 1, 2, \ldots, \tag{3.1}$$

$$p_{ij}(t+s) = \sum_k p_{ik}(t) p_{kj}(s), \quad i, j = 1, 2, \ldots,$$

para todos los valores de t y s, donde

$$p_{ij}(0) = \begin{cases} 1 & \text{para } i = j \\ 0 & \text{para } i \neq j. \end{cases} \tag{3.2}$$

Consideraremos que para todos los $i, j = 1, 2, \ldots$ las probabilidades de paso $p_{ij}(t)$ satisfacen las siguientes condiciones[1]) (compárese (2.5) y más adelante):

$$\begin{aligned} 1 - p_{ii}(\Delta t) &= \lambda_i \Delta t + o(\Delta t), \\ p_{ij}(\Delta t) &= \lambda_{ij}\Delta t + o(\Delta t), \quad j \neq i, \end{aligned} \tag{3.3}$$

donde $\frac{o(\Delta t)}{\Delta t} \to 0$ para $\Delta t \to 0$, λ_i es *la densidad de salida* del estado i; λ_{ij} es *la densidad de paso* de i al estado correspondiente j. Supongamos $\lambda_{ii} = -\lambda_i$, $i = 1, 2, \ldots$

Teorema 1. *Las probabilidades de paso p_{ij}, para el proceso de Markov con un número finito de estados para la condición* (3.3), *satisfacen a las ecuaciones diferenciales*

$$p'_{ij}(t) = \sum_k \lambda_{ik} p_{kj}(t), \quad i, j = 1, 2, \ldots, \tag{3.4}$$

y

$$p'_{ij}(t) = \sum_k p_{ik}(t) \lambda_{kj}, \quad i, j = 1, 2, \ldots, \tag{3.5}$$

con las condiciones iniciales (3.2).

Las ecuaciones (3.5) forman el llamado *sistema directo* y las ecuaciones (3.4), *el sistema inverso de ecuaciones diferenciales de Kolmogórov*.

D e m o s t r a c i ó n. De acuerdo con la fórmula general (3.1),

$$p_{ij}(t + \Delta t) = \sum_k p_{ik}(\Delta t) p_{kj}(t) = \sum_k p_{ik}(t) p_{kj}(\Delta t).$$

Utilizando las expresiones asintóticas (3.3) obtenemos que

$$\frac{p_{ij}(t+\Delta t) - p_{ij}(t)}{\Delta t} = \sum_k \left[\lambda_{ik} + \frac{o(\Delta t)}{\Delta t}\right] p_{kj}(t) = \sum_k p_{ik}(t) \left[\lambda_{kj} + \frac{o(\Delta t)}{\Delta t}\right].$$

Las partes derechas de estas igualdades tienen un límite completamente determinado para $\Delta t \to 0$:

$$\lim_{\Delta t \to 0} \sum_k \left[\lambda_{ik} + \frac{o(\Delta t)}{\Delta t}\right] p_{kj}(t) = \sum_k \lambda_{ik} p_{kj}(t),$$

$$\lim_{\Delta t \to 0} \sum_k p_{ik}(t) \left[\lambda_{kj} t + \frac{o(\Delta t)}{\Delta t}\right] = \sum_k p_{ik}(t) \lambda_{kj}.$$

Por consiguiente, también existe el límite

$$\lim_{\Delta t \to 0} \frac{p_{ij}(t+\Delta t) - p_{ij}(t)}{\Delta t} = p'_{ij}(t),$$

[1]) Señalemos que la relación (3.3) tendrá lugar, si (igual que en la deducción de la fórmula (2.6) para las probabilidades π_{ij} de paso desde i a j directamente) suponemos que la probabilidad de tal paso durante un tiempo pequeño Δt es $\lambda_{ij}\Delta t + o(\Delta t)$, y la probabilidad de más de un paso de estado a estado durante el tiempo Δt es $o(\Delta t)$.

donde deben cumplirse las igualdades (3.4) y (3.5). El teorema queda demostrado.

Se debe señalar que las ecuaciones diferenciales indicadas para las probabilidades de paso $p_{ij}(t)$ tienen lugar no sólo en el caso de un número finito de estados, sino también con algunas restricciones complementarias, y en el caso de número de estados numerables. Por ejemplo, el método anterior utilizado para la deducción del sistema directo de ecuaciones diferenciales queda en vigor, si en las expresiones asintóticas (3.3) los miembros restantes $o(\Delta t)$ son tales que

$$\frac{o(\Delta t)}{\Delta t} \to 0 \quad \text{para} \quad \Delta t \to 0$$

es u n i f o r m e por todas las i, con ello, las densidades de paso λ_{ij} están limitadas u n i f o r m e m e n t e para el valor fijado de j:

$$\lambda_{ij} \leqslant \Lambda_j < \infty, \qquad i = 1, 2, \ldots \tag{3.6}$$

Señalemos también que el sistema directo (3.5) tiene lugar en las probabilidades $p_j(t)$ para cualquier distribución inicial: precisamente de (3.1) y (3.5) se deduce inmediatamente que

$$p'_j(t) = \sum_k p_k(t)\lambda_{kj}, \qquad j = 1, 2, \ldots \tag{3.7}$$

Examinemos algunos ejemplos.

Ejemplo (*corriente de exigencias de Poisson*).

Supongamos que en un sistema de servicio determinado en el transcurso del tiempo se reciben exigencias de modo que el número de exigencias $\xi(t)$ durante el tiempo t, es un proceso homogéneo de Markov (con los estados $i = 0, 1, \ldots$).

Desde el estado i sólo se puede pasar directamente al siguiente estado $i+1$, $i = 0, 1, \ldots$ Supongamos que las densidades de salida desde cada estado i son las mismas e iguales a λ. Evidentemente, la probabilidad $p_{ii}(\Delta t)$ coincide, en nuestro caso, con la probabilidad de que el sistema se encuentre en el estado i en todo el intervalo de tiempo Δt, y de acuerdo con la fórmula general (1.7),

$$1 - p_{ii}(\Delta t) = \lambda\Delta t + o(\Delta t).$$

Está claro, que la densidad de salida λ coincide con la densidad de paso al siguiente estado, y de tal modo, en nuestro caso se cumplen las relaciones (3.3) en las cuales

$$\lambda_i = \lambda, \quad \lambda_{i,i+1} = \lambda, \quad \lambda_{ij} = 0 \quad \text{para} \quad j \neq i,\ i+1.$$

Para las probabilidades de paso $p_{ij}(t)$ tiene lugar el sistema de ecuaciones diferenciales (3.5). Evidentemente, $p_{ij}(t) = p_{0,j-i}(t)$. Supongamos que

$$p_j(t) = p_{0j}(t), \qquad j = 0, 1, \ldots$$

Las ecuaciones diferenciales indicadas para las funciones $p_j(t)$ tienen la siguiente forma:

$$p_0'(t) = -\lambda p_0(t), \quad p_k'(t) = \lambda p_{k-1}(t) - \lambda p_k(t), \quad k = 1, 2, \ldots$$

Si pasamos a las nuevas funciones $f_k(t) = e^{\lambda t} p_k(t)$, entonces obtenemos que

$$f_0'(t) = \lambda f_0(t) + e^{\lambda t} p_0'(t) = \lambda f_0(t) - \lambda e^{\lambda t} p_0(t) = 0$$

$$f_k'(t) = \lambda f_k(t) + e^{\lambda t} p_k'(t) = \lambda f_k(t) + \lambda e^{\lambda t} p_{k-1}(t) - \\ - \lambda e^{\lambda t} p_k(t) = \lambda f_{k-1}(t) \quad k = 1, 2, \ldots,$$

donde $f_0(0) = 1$, $f_k(0) = 0$, para $k = 1, 2, \ldots$ El sistema de ecuaciones diferenciales de la forma

$$f_1'(t) = 0, \quad f_k'(t) = \lambda f_{k-1}(t), \quad k = 1, 2, \ldots,$$

con las condiciones iniciales indicadas tiene evidentemente, la siguiente solución:

$$f_0(s) = 1, \quad f_1(t) = \lambda t, \; \ldots, \; f_n(t) = \frac{(\lambda t)^n}{n!}, \; \ldots$$

Volviéndonos a las funciones iniciales $p_k(t) = e^{-\lambda t} \cdot f_k(t)$ obtenemos que

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \ldots,$$

es decir, la corriente de exigencias examinada es una corriente *de Poisson.*

Ejemplo (*sistema con tiempo exponencial de servicio*). Supongamos que un sistema de servicio recibe una corriente de exigencias de Poisson con el parámetro λ_0 (las diferentes exigencias se reciben independientemente una de otra y la probabilidad de recibir una exigencia por separado durante el pequeño intervalo de tiempo Δt es $\lambda_0 \cdot \Delta t + o(\Delta t)$. Supongamos, que en el servicio de cada exigencia aislada se gasta un tiempo aleatorio τ, distribuido según una ley exponencial con el parámetro λ_1, es decir,

$$\mathbf{P}\{\tau > t\} = e^{-\lambda_1 t}$$

Examinemos dos estados del sistema de servicio; 0 significa el sistema libre, 1 significa el sistema ocupado. Consideraremos que si el sistema está ocupado, entonces las exigencias que se reciben en este tiempo, reciben la negativa y salen de la esfera de servicio.

Supongamos que en un momento t_0 el sistema se encuentra en el estado 0. Debido a la independencia del recibimiento de las exigencias, su comportamiento futuro no depende de las circunstancias anteriores, y en particular, durante el tiempo Δt el sistema pasa al estado 1 con la probabilidad $\lambda_0 \cdot \Delta t + o(\Delta t)$. Supongamos, que el sistema se encuentra en el estado 1 en el momento t_1. Si designamos por τ_1 el momento aleatorio de paso de 1 a 0 (momento de finalizar el servicio), entonces, de acuerdo a la propiedad de la ley de distribución exponencial (véase el punto 2 del § 2, cap. II), para el tiempo de servicio tiene lugar la siguiente igualdad:

$$\mathbf{P}\{\tau_1 > t \mid \tau_1 > t_1\} = e^{-\lambda_1 (t-t_1)}.$$

Se ve que el paso al estado 0 y el comportamiento ulterior del sistema no dependen de su comportamiento hasta el momento t_1. En particular, durante el intervalo Δt siguiente a t_1, el sistema pasa al estado 0 con probabilidad $\lambda_1 \cdot \Delta t + o(\Delta t)$. De tal modo, la evolución del sistema se describe por el proceso de Markov con dos estados 0 y 1, y con las correspondientes densidades de paso λ_0, λ_1.

Examinemos la probabilidad de paso $p_{ij}(t)$. En nuestro caso $p_{01}(t) = 1 - p_{00}(t)$; $p_{10}(t) = 1 - p_{11}(t)$ y las ecuaciones diferenciales (3.5) tienen la forma

$$p'_{00}(t) + (\lambda_0 + \lambda_1)\, p_{00}(t) = \lambda_1,$$
$$p'_{11}(t) + (\lambda_0 + \lambda_1)\, p_{11}(t) = \lambda_0.$$

Resolviéndolas, obtenemos que

$$p_{00}(t) = \left(1 - \frac{\lambda_1}{\lambda_0+\lambda_1}\right) e^{-(\lambda_0+\lambda_1)t} + \frac{\lambda_1}{\lambda_0+\lambda_1},$$
$$p_{11}(t) = \left(1 - \frac{\lambda_0}{\lambda_0+\lambda_1}\right) e^{-(\lambda_0+\lambda_1)t} + \frac{\lambda_0}{\lambda_0+\lambda_1}.$$

Como conclusión de este punto señalemos que el sistema de ecuaciones diferenciales (3.5) puede tener solución diferente de las probabilidades de paso $p_{ij}(t)$ del proceso de Markov con las correspondientes densidades de paso λ_{ij}, si se altera *la condición de estabilidad* siguiente: *durante un tiempo finito ocurre con probabilidad* 1 *sólo un número finito de pasos.*

Ejemplo (proceso de multiplicación pura). Supongamos que el proceso homogéneo de Markov $\xi(t)$ describe el proceso de multiplicación de ciertas partículas ($\xi(t)$ es el número de partículas en el momento de tiempo t), en el cual, en caso de existir i partículas en el siguiente intervalo de tiempo Δt con probabilidad $\lambda_i\, \Delta t + o(\Delta t)$ se puede añadir una partícula, y la probabilidad de añadirle un número mayor de partículas es $o(\Delta t)$. Las densidades de paso λ_{ij} tienen la siguiente forma:

$$\lambda_{ij} = \begin{cases} \lambda_i & \text{para } j = i+1, \\ 0 & \text{para los restantes } j \end{cases} \qquad (i = 1, 2, \ldots).$$

Evidentemente, la condición de estabilidad consiste en que el tiempo τ, que se exige para un número infinito de pasos $1 \to 2 \to \ldots$ (desde 1 hasta ∞), debe ser infinito: $\tau = \infty$. Si introducimos las magnitudes aleatorias τ_i, o sea, el tiempo de estancia en el estado i (τ_i coincide con el tiempo de espera del cambio de estado i), entonces tendremos

$$\tau = \tau_1 + \tau_2 + \ldots = \lim_{n\to\infty} \sum_{i=1}^{n} \tau_i,$$

donde τ_i, $i = 1, 2, \ldots$, son magnitudes aleatorias independientes, distribuidas según la ley exponencial con el parámetro correspondiente λ_i (véase (1.6)). La condición de estabilidad $\tau = \infty$ es equivalente a que

$$e^{-\tau} = \lim_{n\to\infty} \prod_{i=1}^{n} e^{-\tau_i} = 0.$$

Pero para la magnitud no negativa $e^{-\tau}$ la condición $e^{-\tau} = 0$ con probabilidad 1 es evidentemente equivalente, a que sea igual a cero la esperanza matemática $\mathbf{M}e^{-\tau}$, que es

$$\mathbf{M}e^{-\tau} = \lim_{n\to\infty} \mathbf{M}\left(\prod_{i=1}^{n} e^{-\tau_i}\right) = \lim_{n\to\infty} \prod_{i=1}^{n} \mathbf{M}e^{-\tau_i}$$

(véase el teorema 5 del § 4, cap. I). Tenemos

$$\mathbf{M}e^{-\tau_i} = \int_0^\infty e^{-t} (\lambda_i e^{-\lambda_i t})\, dt = \frac{\lambda_i}{1+\lambda_i} = 1 - \frac{1}{1+\lambda_i},$$

$$\mathbf{M}e^{-\tau} = \lim_{n\to\infty} \prod_{i=1}^{n} \left(1 - \frac{1}{1+\lambda_i}\right) = \prod_{i=1}^{\infty} \left(1 - \frac{1}{1+\lambda_i}\right),$$

y del tal modo, la condición de estabilidad examinada es equivalente a que

$$\prod_{i=1}^{\infty} \left(1 - \frac{1}{1+\lambda_i}\right) = 0.$$

Pero lo último significa que $\sum\limits_{i=1}^{\infty} \frac{1}{1+\lambda_i} = \infty$ y esto a su vez es equivalente a la condición $\sum \lambda_i^{-1} = \infty$. De tal modo, la condición de estabilidad del proceso examinado de la multiplicación pura consiste en que

$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i} = \infty.$$

Supongamos que se altera esta condición; entonces, durante un intervalo de tiempo t, se forma un número infinito de partículas con una probabilidad positiva: $\mathbf{P}\{\tau < t\} < 0$. Supongamos que en el momento inicial se tiene una sóla partícula. Introduzcamos un estado más, que significa la existencia de una cantidad infinita de partículas. El proceso pasa a este estado durante el tiempo t con la probabilidad positiva $p_{1\infty}(t) = \mathbf{P}\{\tau < t\}$, a pesar de que las densidades de paso correspondientes $\lambda_{i\infty}$ son todas iguales a cero y la ecuación diferencial (3.5) para $p_{i\infty}(t)$ tiene la forma

$$p'_{i\infty}(t) = 0, \quad p_{1\infty}(0) = 0;$$

su solución es la función $p_{1\infty}(t) \equiv 0$, mientras que la probabilidad $p_{1\infty}(t) > 0$.

2. Coeficiente de ergodicidad y convergencia hacia la distribución estacionaria. Del mismo modo que el caso del tiempo discreto, tiene lugar el resultado siguiente.

Teorema 2. *Si el coeficiente de ergodicidad*

$$\boldsymbol{k}(t_0) = 1 - \frac{1}{2} \sup_{i,j} \sum_k |p_{ik}(t_0) - p_{jk}(t_0)|$$

es positivo para un determinado $t_0 > 0$, las probabilidades $p_j(t)$ de que el sistema se encuentre, después de un tiempo t, en el estado correspondiente $j = 1, 2, \ldots$, para $t \to \infty$, tienen los valores límites

$$p_j^* = \lim_{t\to\infty} p_j(t), \quad j = 1, 2, \ldots \tag{3.8}$$

Las probabilidades límites p_j^* no dependen de la distribución inicial p_i^0, $i = 1, 2, \ldots$, con esto

$$\sup_{p_i^0} |p_j(t) - p_j^*| \leqslant Ce^{-Dt} \tag{3.9}$$

$$\left(C = \frac{1}{1-\boldsymbol{k}(t_0)}, \; D = \frac{1}{t_0} \ln \frac{1}{1-\boldsymbol{k}(t_0)}\right).$$

(La demostración es la misma que en el teorema 4 del párrafo anterior).

Respecto a la condición $\boldsymbol{k}(t_0) > 0$ de este teorema, en el caso del tiempo continuo, se puede añadir lo siguiente: *si un estado j es accesible desde otro estado i (es decir, $p_{ij}(t_0) > 0$ para un determinado t_0), también es accesible para cualquier $t > 0$ con una probabilidad positiva* $p_{ij}(t) > 0$ (para la cadena con un número finito de estados comunicantes, se deduce de aquí, que todas las probabilidades de paso y el coeficiente de ergodicidad $\boldsymbol{k}(t) \geqslant \max \min p_{ij}(t)$ son rigurosamente positivos para todos los $t > 0$).

Demostremos esto para el caso de un número finito de estados. Examinemos al principio la probabilidad $p_{ii}(t)$ que se supone continua como función de t y $p_{ii}(0) = 1$, ya que esta probabilidad es positiva para valores de t suficientemente pequeños. Ahora bien, de la relación (3.1) se ve que para cualesquiera s y t

$$p_{ii}(s + t) \geqslant p_{ii}(s)\, p_{ii}(t),$$

y en realidad la probabilidad $p_{ii}(t)$ es positiva para todos los valores de t. Examinemos ahora el estado j accesible desde el estado i, es decir, tal estado que $p_{ij}(s) > 0$ para un determinado s. Mostremos que tal probabilidad $p_{ij}(t)$ es positiva para cualquier $t > 0$. Tenemos

$$p_{ij}(t) \geqslant p_{ij}(u)\, p_{jj}(t-u), \quad u \leqslant t.$$

Como hemos visto, la probabilidad $p_{jj}(t-u)$ es siempre positiva, de modo que es suficiente demostrar que $p_{ij}(u) > 0$ para $u \leqslant t$. Según la suposición, $p_{ij}(s) > 0$. Examinemos la cadena de Markov con el tiempo discreto y con las probabilidades de paso $p_{ij} = p_{ij}\left(\frac{s}{n}\right)$, donde el número natural n satisface a la condición $m\frac{s}{n} \leqslant t$ (m es el número de estados). Ya que $p_{ij}\left(n\frac{s}{n}\right) > 0$, el estado j es accesible desde i. Se ve fácilmente, que es accesible no sólo después de n pasos, sino también en un determinado número de pasos n_0, no superior al número m de todos los estados existentes, es decir, $p_{ij}\left(n_0\frac{s}{n}\right) > 0$, donde $n_0\frac{s}{n} = u \leqslant t$. La demostración está terminada.

Es útil señalar la siguiente circunstancia. Eligiendo una unidad convencional de tiempo Δ, se puede considerar a $\xi(t)$, en el momento $t = k\Delta$, como una cadena de Markov $\tilde{\xi}(k)$, $k = 0, 1, \ldots$, con las probabilidades de paso $\tilde{p}_{ij}(k) = p_{ij}(k\Delta)$. Evidentemente, las probabilidades límites p_j^*, $j = 1, 2, \ldots$, serán las mismas que para el proceso inicial $\xi(t)$. Si referimos los estados $j = 1, 2, \ldots$, de este proceso al mismo tipo que en la cadena correspondiente de Markov $\tilde{\xi}(k)$ (cada estado j puede ser reversible o irreversible, positivo o nulo), entonces de la relación (3.8), en el caso de tiempo continuo t (lo mismo que de la relación (2.21) en el caso del tiempo discreto t) se deduce, que las probabilidades límites (3.8), *para la clase cerrada de estados positivos, forman la única distribución estacionaria*

$$p_j^* = \sum_j p_i^* p_{ij}(t), \quad j = 1, 2, \ldots, \tag{3.10}$$

para todos los valores de t.

Para las probabilidades estacionarias $p_j(t) = p_j^*$, obtenemos de las ecuaciones diferenciales (3.7) (para $p_j'(t) = 0$) el siguiente sistema lineal:

$$\sum_i p_i^* \lambda_{ij} = 0, \quad j = 1, 2, \ldots \tag{3.11}$$

Ejemplo (sistema de servicio. Fórmula de Erlang). Examinemos el sistema que puede atender a la vez a m demandas. Consideraremos que tenemos m líneas y la demanda siguiente se recibe en una de las líneas, si por lo menos está libre una de ellas; en caso contrario a la demanda recibida se le da la negativa y aquélla sale fuera de la esfera de servicio. Supongamos, que la corriente de demandas es la corriente de Poisson con el parámetro λ_0, y que las demandas son atendidas independientemente y que el tiempo de servicio en cada demanda (en cada una de las m líneas) está distribuido según la ley exponencial con el parámetro λ.

Examinemos los estados $k = 0, 1, \ldots, m$, donde el estado k significa que están ocupadas justamente k líneas. El paso del sistema de estado a estado durante el tiempo t, representa en sí un proceso de Markov, cuya densidad de paso tiene la forma

$$\lambda_{0j} = \begin{cases} \lambda_0 & \text{para } j = 1, \\ 0 & \text{para } j \neq 1 \end{cases} \qquad \lambda_{kj} = \begin{cases} \lambda_0 & \text{para } j = k+1, \\ k\lambda & \text{para } j = k-1, \\ 0 & \text{para } j \neq k-1,\ k+1. \end{cases}$$

En efecto, el paso desde k a $k+1$ se realiza al recibirse la demanda siguiente, que ocurre en el tiempo Δt con la probabilidad $\lambda_0 \Delta t + o(\Delta t)$. La probabilidad de que ninguna de las k líneas ocupadas se libere en el tiempo Δt es $[1 - \lambda\Delta t - o(\Delta t)]^k$ (ya que las líneas se sirven, independientemente una de otra) y la probabilidad de liberación de una de las líneas, es decir, el paso del estado k al $k-1$ es $1 - [1 - \lambda\Delta t - o(\Delta t)]^k = k\lambda\Delta t + o(\Delta t)$. La probabilidad de otros cambios en el sistema durante el intervalo de tiempo Δt es $o(\Delta t)$.

Como se ve directamente de la expresión para las probabilidades $p_{ij}(t)$, obtenidas antes en el caso $m = 1$ (véase la pág. 165), cuando $t \to \infty$ las probabilidades $p_{ij}(t)$ tienden rápido, exponencialmente, hacia sus valores límites. De acuerdo a la valuación general (3.9), también se observa lo mismo en el caso de m líneas. Las probabilidades estacionarias $p_j^*(t)$ pueden ser halladas de las ecuaciones (3.11), que en nuestro caso tienen la siguiente forma:

$$-\lambda_0 p_0^* + \lambda p_1^* = 0,$$

$$\lambda_0 p_{k-1}^* - (\lambda_0 + k\lambda)\, p_k^* + (k+1)\, \lambda p_{k+1}^* = 0, \quad 1 \leqslant k < m,$$

$$\lambda_0 p_{m-1}^* - m\lambda p_m^* = 0.$$

De estas ecuaciones obtenemos que

$$p_k = \frac{\frac{1}{k!}\left(\frac{\lambda_0}{\lambda}\right)^k}{\sum\limits_{l=0}^{m} \frac{1}{l!}\left(\frac{\lambda_0}{\lambda}\right)^l}, \quad k = 0, 1, \ldots, m.$$

(Las expresiones halladas para las probabilidades estacionarias se denominan *fórmulas de Erlang*).

§ 4. PROCESOS QUE SE RAMIFICAN

1. Ecuación diferencial de las funciones generatrices. Supongamos que se tiene un conjunto determinado de partículas, que en el transcurso del tiempo se transforman en partículas del mismo tipo, con ello este proceso de «multiplicación» posee la siguiente propiedad: cada una de las partículas iniciales, después de un tiempo t, independientemente de las otras partículas y de las circunstancias precedentes al momento inicial, genera un grupo de n partículas con la misma probabilidad $p_n(t)$ para todas las partículas.

Designemos por $\xi(t)$ el número de partículas existentes en el momento de tiempo t. Evidentemente, la evolución de la magnitud $\xi(t)$ representa en sí un proceso aleatorio de Markov. Un proceso de tal tipo se llama *proceso que se ramifica*[1]).

Supongamos que en un momento de tiempo s, digamos $s = 0$, se tienen justamente k partículas. Designemos por $\xi_i(t)$ el número de partículas generadas por la partícula i, $i = 1, \ldots, k$, después de un tiempo t. Entonces el número total de partículas después de un tiempo t será

$$\xi(t) = \xi_1(t) + \ldots + \xi_k(t). \tag{4.1}$$

Aquí, las magnitudes aleatorias $\xi_1(t), \ldots, \xi_k(t)$ son i n d e p e n d i e n t e s entre sí y tienen una misma distribución de probabilidades

$$\mathbf{P}\{\xi_i(t) = n\} = p_n(t), \quad n = 0, 1, \ldots$$

Supongamos que una partícula aislada, en el pequeño intervalo de tiempo Δt con la probabilidad

$$p_n(\Delta t) = \lambda_n \cdot \Delta t + o(\Delta t), \quad n \neq 1,$$

se transforma en n partículas nuevas, y se queda invariable, con la probabilidad

$$p_1(\Delta t) = 1 - \lambda \Delta t + o(\Delta t).$$

Supongamos luego, que $\lambda_1 = -\lambda$, $\sum\limits_k \lambda_k = 0$ y las probabilidades de paso $p_n(t) = p_{1n}(t)$ satisfacen a la ecuación diferencial de Kolmogórov (véase (3.4))

$$\frac{d}{dt} p_n(t) = \sum_k \lambda_k p_{kn}(t), \quad n = 0, 1, \ldots,$$

[1]) El modelo descrito puede servir para estudiar muchos procesos reales (reacciones fotoquímicas, procesos nucleares, etc.).

donde $p_{kn}(t)$ son las probabilidades de paso del proceso que se ramifica de Markov $\xi(t)$ ($p_{kn}(t)$ es la probabilidad de que, k partículas generen n partículas durante el tiempo t).

Introduzcamos las funciones generatrices

$$F(t, z) = \sum_{n=0}^{\infty} p_n(t) z^n,$$
$$F_k(t, z) = \sum_{n=0}^{\infty} p_{kn}(t) z^n. \tag{4.2}$$

Para cada z, $|z| < 1$, tenemos

$$\sum_{u=0}^{\infty} z^n \frac{d}{dt} p_n(t) = \frac{d}{dt} \sum_{n=0}^{\infty} p_n(t) z^n = \sum_k \lambda_k \sum_{n=0}^{\infty} p_{kn}(t) z^n,$$

lo que da la siguiente ecuación diferencial para la función generatriz $F(t, z)$:

$$\frac{d}{dt} F(t, z) = \sum_k \lambda_k F_k(t, z). \tag{4.3}$$

Las funciones $F(t, z)$ y $F_k(t, z)$ determinadas por las fórmulas (4.2) son tales que para un valor fijado de z representan en sí las esperanzas matemáticas

$$F(t, z) = \mathbf{M} z^{\xi_i(t)}, \quad i = 1, \ldots, k,$$
$$F_k(t, z) = \mathbf{M} z^{\xi(t)},$$

donde $\xi_i(t)$ es el número de partículas generadas por la partícula i durante el tiempo t y $\xi(t) = \sum_{i=1}^{k} \xi_i(t)$. Ya que todas las magnitudes $\xi_i(t)$, $i = 1, \ldots, k$ son independientes, entonces

$$\mathbf{M} z^{\xi_1(t) + \ldots + \xi_k(t)} = \mathbf{M} z^{\xi_1(t)} \ldots \mathbf{M} z^{\xi_k(t)},$$

lo que da la siguiente relación para las funciones generatrices examinadas

$$F_k(t, z) = [F(t, z)]^k, \quad k = 1, 2, \ldots \tag{4.4}$$

Ya que $F_0(t, z) \equiv 1$, la ecuación diferencial para la función productora $F(t, z)$ se puede escribir en la forma

$$\frac{d}{dt} F(t, z) = \sum_k \lambda_k F^k(t, z). \tag{4.5}$$

Consideraremos que las densidades de paso λ_k, $k = 0, 1, \ldots$ son parámetros dados del proceso que se ramifica $\xi(t)$. Introduzcamos la función

$$f(x)=\sum_{k=0}^{\infty} \lambda_k x^k, \tag{4.6}$$

la cual es analítica para $0 < x < 1$ y la fórmula (4.6) da su desarrollo en serie de potencias. De acuerdo a la igualdad (4.5) la función generatriz $F(t, z)$ es la solución de la ecuación diferencial de la forma

$$\frac{dx}{dt}=f(x). \tag{4.7}$$

Ya que $F(0, z) = z$, la función generatriz $F(t, z)$ para cada z, $0 \leqslant z \leqslant 1$, coincide con la solución $x = x(t)$ de esta ecuación que satisface a la condición inicial $x(0) = z$.

En lugar de la ecuación (4.7), es cómodo examinar su ecuación diferencial equivalente para la función $t = t(x)$ inversa a $x = x(t)$

$$\frac{dt}{dx}=\frac{1}{f(x)}.$$

La función $t = t(x)$, que es la solución de esta ecuación, tiene la forma

$$t=\int_z^x \frac{du}{f(u)}, \quad 0 \leqslant x \leqslant 1. \tag{4.8}$$

Ejemplo. Supongamos que las densidades de paso son

$$\lambda_0 = \lambda, \quad \lambda_1 = -\lambda, \quad \lambda_k = 0 \quad \text{para} \quad k = 2, 3, \ldots$$

En este caso $f(x) = \lambda(1 - x)$ y

$$t=\int_z^x \frac{du}{f(u)}=-\frac{1}{\lambda}[\ln(1-x)-\ln(1-z)].$$

De esta relación se obtiene fácilmente la función $F = F(t, z)$. Precisamente,

$$\ln(1 - F) = -\lambda t + \ln(1 - z)$$

y

$$F(t, z) = 1 - e^{-\lambda t}(1 - z).$$

Las probabilidades $p_n(t)$ determinadas del desarrollo $F(t, z) = \sum_{n=0}^{\infty} p_n(t) z^n$ en el caso examinado son

$$p_0(t) = 1 - e^{-\lambda t}, \quad p_1(t) = e^{-\lambda t}, \quad p_n(t) = 0 \quad \text{para} \quad n = 2, 3, \ldots$$

Ejemplo. Imaginémonos que algunas partículas se multiplican mediante su fisión por la mitad; en el proceso que se ramifica, correspondiente a este fenómeno, hay que tomar

$$\lambda_0 = 0, \quad \lambda_1 = -\lambda, \quad \lambda_2 = \lambda \quad (\lambda_k = 0 \quad \text{para} \quad k < 2).$$

En este caso

$$f(x) = \lambda x (x-1)$$

y

$$t = \int_z^x \frac{du}{f(u)} = \frac{1}{\lambda} \int_z^x \left[\frac{1}{u-1} - \frac{1}{u} \right] du = \frac{1}{\lambda} \left[\ln\left(1 - \frac{1}{x}\right) - \ln\left(1 - \frac{1}{z}\right) \right].$$

De tal modo, la función $F(t, z) = \sum_{n=0}^{\infty} p_n(t) z^n$ puede ser determinada de la relación

$$t = \frac{1}{\lambda} \left[\ln\left(1 - \frac{1}{F}\right) - \ln\left(1 - \frac{1}{z}\right) \right].$$

Hallamos fácilmente que

$$F(t, z) = e^{-\lambda t} \frac{z}{1 - z(1 - e^{-\lambda t})} = e^{-\lambda t} z \sum_{k=0}^{\infty} (1 - e^{-\lambda t})^k z^k.$$

Como resultado obtenemos que

$$p_0(t) = 0, \quad p_n(t) = e^{-\lambda t} (1 - e^{-\lambda t})^{n-1}, \quad n = 1, 2, \ldots$$

Examinemos la ecuación diferencial (4.7) en la cual la función $f(x)$ se determina por la fórmula (4.6). De esta fórmula se ve que

$$f''(x) = \sum_{k=2}^{\infty} k(k-1) \lambda_k x^{k-2} \geqslant 0$$

$$\text{para } 0 \leqslant x < 1,$$

de modo que la función $f(x)$ es convexa y su derivada $f'(x)$ crece monótonamente en el intervalo $0 \leqslant x \leqslant 1$. El valor de $x = 1$ es una raíz de la ecuación $f(x) = 0$, ya que $\sum_k \lambda_k = 0$. Es posible que sólo exista todavía otra raíz $x = \alpha$ de esta ecuación, y en correspondencia con esto, el gráfico de la función $f(x)$ tiene la forma indicada en la fig. 20.

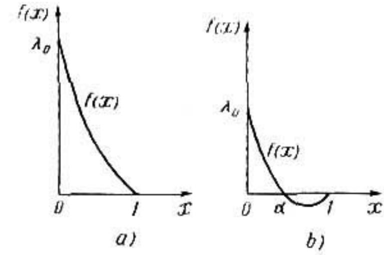

Fig. 20.

Supongamos que se tiene la raíz $x = \alpha$, $0 < \alpha \leqslant 1$ que determina una curva integral especial $x(t) \equiv \alpha$ de las ecuaciones diferenciales examinadas. Tomemos la curva integral (4.8) que pasa por el punto $t = 0$, $x = z$, $0 \leqslant z < \alpha$:

$$t = \int_z^x \frac{du}{f(u)}.$$

Ya que la derivada $f'(\alpha)$ es finita y para $x \sim \alpha$ la función $f(x)$ tiene la forma $f(x) \sim f'(\alpha) \cdot (x - \alpha)$, entonces a lo largo de la curva integral crece ilimitadamente el valor $t = \int_z^x \frac{dx}{f(u)}$ para $x \to \alpha$, con ello la propia curva no corta en ningún lugar a otra curva integral $x(t) \equiv \alpha$. En el intervalo $0 \leqslant x < \alpha$ la función $f(x)$ es positiva y, por consiguiente, la curva integral $x = x(t)$ es positiva y, por consiguiente, la curva integral $x = x(t)$ crece monótonamente para $t \to \infty$, quedando limitada por el valor $x = \alpha$.

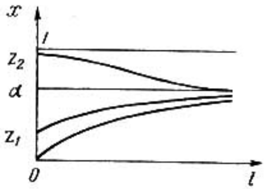

Fig. 21.

Como función monótona limitada, la función $x(t)$ tiene un límite determinado $\beta = \lim_{t\to\infty} x(t)$, $z \leqslant \beta \leqslant \alpha$. Pero cuando $x \to \beta$ la función $f(x)$ tiene su límite $f(\beta)$:

$$f(\beta) = \lim_{t\to\infty} f[x(t)] = \lim_{t\to\infty} x'(t).$$

Está claro que el valor $f(\beta)$ deberá ser igual a cero, ya que en caso contrario la función

$$x(t) = z + \int_0^t f[x(s)]\, ds$$

crecerá ilimitadamente para $t \to \infty$. Por consiguiente, β es raíz de la ecuación $f(x) = 0$, y coincide con α: $\beta = \alpha$. De tal modo, todas las curvas integrales $x = x(t)$ para $t = 0$ que pasan por el punto $x = z$, $0 \leqslant z < \alpha$ crecen monótonamente para $t \to \infty$ y

$$\lim_{t\to\infty} x(t) = \alpha. \tag{4.9}$$

Es completamente análogo el comportamiento de las curvas integrales que pasan por los puntos $x = z$, $\alpha < z < 1$ $(0 \leqslant \alpha < 1)$ para $t = 0$. La diferencia sólo consiste en que $x(t)$ decrece monótonamente, ya que la derivada $x'(t) = f[x(t)]$ es negativa ($f(x) \leqslant 0$ para $\alpha < x < 1$). El cuadro general de las curvas integrales correspondientes a los valores del parámetro z en el intervalo $0 \leqslant z < 1$, está expuesta en la fig. 21.

El caso $z = 1$ exige un examen especial. A él siempre le corresponde la curva integral de la forma $x(t) \equiv 1$.

Si para un determinado x_0, $\alpha < x_0 < 1$,

$$\int_{x_0}^{1} \frac{dx}{f(x)} = -\infty \tag{4.10}$$

(que siempre es así, cuando $f'(1) < \infty$), entonces la curva integral arbitraria de la forma

$$t = t_0 + \int_{x_0}^{x} \frac{du}{f(u)}, \quad 0 \leqslant x < 1,$$

que pasa por un punto determinado (t_0, x_0), para $x \to 1$, decrece ilimitadamente:

$$t = t_0 + \int_{x_0}^{x} \frac{du}{f(u)} \to -\infty.$$

Esto quiere decir, que cualquiera que fuese $t_0 > 0$, para un determinado $x = z$, $0 \leqslant z < 1$ tiene lugar la igualdad

$$t(z) = t_0 + \int_{x_0}^{z} \frac{du}{f(u)} = 0.$$

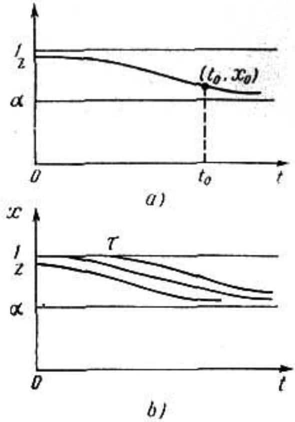

Fig. 22.

Vemos (véase la fig. 22,*a*) que todas las curvas integrales cortan al eje $t = 0$ en un punto $(0, z)$, donde $0 \leqslant z < 1$ y, por consiguiente, $x(t) \equiv 1$ es la única curva integral que pasa por el punto $(0, 1)$.

Si

$$\int_{x_0}^{1} \frac{dx}{f(x)} > -\infty, \tag{4.11}$$

entonces para valores suficientemente grandes de $t_0 > 0$, la curva integral $t = t_0 + \int_{x_0}^{x} \frac{dx}{f(x)}$ corta a la curva integral $x(t) \equiv 1$ siendo tangente a ésta en un punto determinado $(\tau, 1)$, donde

$$\tau = t_0 + \int_{x_0}^{1} \frac{dx}{f(x)}$$

(véase la fig. 23, *b*). En este caso, por el punto $(0, 1)$ pasa toda una familia de curvas integrales $x_\tau(t)$, cada una de las cuales corresponde a su valor $\tau \geqslant 0$. Entre ellas se encuentra la curva integral $x_0(t)$ correspondiente al valor $\tau = 0$ y teniendo las mismas propiedades que la curva $x_0(t)$ se halla por debajo de todas las curvas integrales restantes $x_\tau(t)$:

$$x_0(t) \leqslant x_\tau(t), \quad 0 \leqslant t < \infty.$$

Esto se explica, porque dentro de la zona $0 \leqslant x < 1$, $0 < t < \infty$, la solución de la ecuación diferencial examinada es única y en esta zona las curvas integrales no se cortan entre sí. También se ve fácilmente que la curva integral $x_0(t)$ sirve de límite para las otras curvas integrales $x(t, z)$ que descansan debajo de ella y pasan por el punto correspondiente $(0, z)$, donde $0 \leqslant z < 1$

$$x_0(t) = \lim_{z \to 1} x(t, z). \tag{4.12}$$

2. **Efectos de degeneración y explosión.** Recordemos que la función generatriz $F(t, z) = \sum_{k=0}^{\infty} p_k(t) z^k$ del proceso que se ramifica (donde $p_k(t) = \mathbf{P}\{\xi(t) = k\}$ a condición de que $\xi(0) = 1$), como función de t, coincide con la solución $x(t, z)$ de la ecuación diferencial (4.7), con la condición inicial $x(0, z) = z$. El análisis de esta ecuación diferencial expuesto anteriormente, permite hacer las siguientes conclusiones referentes al propio proceso que se ramifica $\xi(t)$.

Hablando en general, existe la probabilidad positiva de que después de un tiempo determinado t no quede ni una sola partícula. Naturalmente, esto no puede ocurrir, si $\lambda_0 = 0$ (es decir, si las partículas no pueden desaparecer sino solamente multiplicarse). Si en el momento inicial $t = 0$ se tiene una partícula, entonces esta probabilidad es $p_0(t) = F(t, 0)$. Si al principio se tienen k partículas, esta probabilidad será $p_{k0}(t) = F^k(t, 0) = p_0^k(t)$.

La probabilidad $p_0(t)$, como función de t, es una solución de la ecuación diferencial (4.7) correspondiente al parámetro $z = 0$:

$$p_0'(t) = f[p(t)], \quad p_0(0) = 0.$$

Anteriormente fue mostrado que esta solución, cuando $t \to \infty$, se aproxima asintóticamente a un valor determinado $p_0 = \alpha$ que es la raíz menor de la ecuación $f(x) = 0$ (véase la relación (4.9)), es decir,

$$\lim_{t \to \infty} p_0(t) = \alpha. \tag{4.13}$$

De tal modo, $p_0 = \alpha$ *es la probabilidad de degeneración del proceso que se ramifica* $\xi(t)$ (la probabilidad de que en un cierto momento de tiempo no quede ni una sola partícula).

Si la función $f(x)$ es positiva en el intervalo $0 \leqslant x < 1$, la probabilidad de degeneración del proceso que se ramifica $\xi(t)$ será igual a 1.

Examinemos el llamado *fenómeno de explosión*, cuando se forma una cantidad infinita de partículas.

La probabilidad de que ocurra la explosión antes del momento t (si al principio había una partícula) es

$$p_\infty(t) = 1 - \mathbf{P}\{\xi(t) < \infty\} =$$
$$= 1 - \sum_{n=0}^{\infty} \mathbf{P}(\xi(t) = n\} = 1 - \sum_{n=0}^{\infty} p_n(t) = 1 - \lim_{z\to 1} F(t, z).$$

En el caso cuando $x(t) \equiv 1$ es la única curva integral de la ecuación diferencial (4.7) que pasa por el punto (0, 1), el límite $\lim\limits_{z\to 1} F(t, z)$ evidentemente, será igual a 1. Por consiguiente, con la condición (4.10), $p_\infty(t) = 0$ para cualquier t, de modo que la probabilidad de la explosión queda excluida. Para la condición (4.11), las curvas integrales $x(t, z)$ cuando $z \to 1$, convergen monótonamente hacia la función $x_0(t)$ descrita anteriormente (véase la relación (4.12)) de modo que *la probabilidad de explosión* es

$$p_\infty(t) = 1 - x_0(t) > 0; \tag{4.14}$$

Se ve fácilmente, que si al principio había k partículas, entonces la probabilidad de explosión después del tiempo t para la condición (4.10) también será igual a cero, y para la condición (4.11) será

$$p_{k\infty}(t) = 1 - x_0^k(t). \tag{4.15}$$

§ 5. ALGUNOS PROCESOS DE SERVICIOS EN MASA Y FLUCTUACIONES ALEATORIAS

1. Procesos de restablecimiento. Sea $\xi_1, \xi_2, \ldots$ una sucesión de magnitudes aleatorias positivas independientes, que tienen la misma distribución de probabilidades, y

$$S_n = \xi_1 + \xi_2 + \ldots + \xi_n.$$

Consideraremos convencionalmente, que se tiene un cierto aparato con el plazo de servicio ξ_1; después de salir fuera de uso (después del tiempo aleatorio ξ_1) se cambia por otro aparato nuevo, que a su vez sale fuera de uso después del tiempo aleatorio ξ_2, después de lo cual se cambia por el siguiente aparato nuevo, etc. Con tal interpretación, las magnitudes S_n, $n = 1, 2, \ldots$, se denominan naturalmente *momentos de restablecimiento*.

En *el proceso de restablecimiento* examinado S_n, $n = 1, 2, \ldots$, nos ocuparemos, fundamentalmente, de las magnitudes $\nu(t)$ que es el número de restablecimientos en el intervalo de tiempo $[0, t]$ y $\eta(t)$, que es el plazo restante de servicio del aparato sucesivo que trabaja en el momento t.

Supondremos que el restablecimiento se realiza en momentos discretos de tiempo $t = kh$ múltiplos de un cierto $h > 0$ (digamos $h = 1$). En relación con esto debemos precisar en que momento se reemplaza el aparato estropeado. Examinando sólo los valores de números enteros $t \geqslant 0$, consideraremos que si el aparato se estropeó en el intervalo $[t, t+1]$, se cambia por otro aparato nuevo en el momento de tiempo $t+1$ (el plazo de servicio de cada aparato puede ser igual a $1, 2, \ldots$).

Así pues, supongamos que las magnitudes examinadas ξ_k, $k = 1, 2, \ldots$, *toman los valores de números enteros* $x = 1, 2, \ldots$ con las probabilidades correspondientes $P(x)$, $\sum_1^\infty P(x) = 1$.

Designemos por $\xi(t)$ el tiempo de trabajo del *aparato siguiente* en el momento t:

$$\xi(t) = t - S_{\nu(t)}. \tag{5.1}$$

En el siguiente momento de tiempo $t+1$ el aparato bien sale fuera de trabajo y se cambia por un aparato nuevo, en este caso $\xi(t+1) = 0$, o bien continúa trabajando, en este caso $\xi(t+1) = \xi(t) + 1$. Con la condición de que el aparato que trabajó el tiempo x, sale fuera de trabajo después de la unidad de tiempo y se cambia por un aparato nuevo con la probabilidad condicional

$$q(x) = \mathbf{P}\{\xi_1 = x+1 \mid \xi_1 > x\} = \frac{P(x+1)}{\sum\limits_{y=x+1}^{\infty} P(y)}, \qquad x = 0, 1, \ldots \tag{5.2}$$

El proceso aleatorio

$$\xi(0) \to \xi(1) \to \xi(2) \ldots$$

representa en sí la cadena de Markov del tipo examinado anteriormente (véase, el ejemplo de la pág. 159) cuando es posible el paso del estado x, al estado siguiente $x+1$ que se realiza con la probabilidad

$$p(x) = 1 - q(x) = \frac{G(x+1)}{G(x)}, \qquad x = 0, 1, \ldots, \tag{5.3}$$

o bien al estado «inicial» 0, que se realiza con la probabilidad $q(x)$ determinada por la fórmula (5.2); en la expresión (5.3) $G(x) = \sum\limits_{y=x+1}^{\infty} P(y)$ significa la probabilidad de que el plazo de servicio del aparato sea mayor que x Evidentemente, la caída en el estado 0 durante t ($\xi(t) = 0$) significa el restablecimiento en el momento t; el tiempo de regreso al estado 0 no es otra cosa que el tiempo de servicio de un aparato; el número de caídas en el estado 0 durante el intervalo $(s, t]$ es igual al número de restablecimientos en el intervalo de tiempo desde s hasta t, etc.

Como fue establecido para nuestra cadena de Markov $\xi(t)$ con los estados $x = 0, 1, \ldots$, con la condición de que el tiempo medio de regreso μ al estado 0 sea finito se tiene *la distribución estacionaria de probabilidades* (2.25):

$$p^*(0) = \frac{1}{\mu}; \quad p^*(x) = \frac{p(0)\,p(1)\ldots p(x-4)}{\mu}, \quad x = 1, 2, \ldots,$$

que para las probabilidades de paso $p(x) = \frac{G(x+1)}{G(x)}$, $x = 0, 1, \ldots$, tiene la forma

$$p^*(x) = \frac{G(x)}{\mu}, \quad x = 0, 1, \ldots; \tag{5.4}$$

con ello *el tiempo medio de regreso* $\mu = \sum_{x=0}^{\infty} G(x)$ *coincide con el tiempo medio de servicio del aparato aislado*:

$$\sum_{x=1}^{\infty} xP(x) = \sum_{x=1}^{\infty} x[G(x-1) - G(x)] =$$
$$= [G(0) - G(1)] + 2[G(1) - G(2)] + 3[G(2) - G(3)] + \ldots =$$
$$= \sum_{x=0}^{\infty} G(x) = \mu. \tag{5.5}$$

Recordemos que $G(x)$ es la probabilidad de que el plazo de servicio del aparato sea mayor que x; supongamos que $G(x)$ decrece cuando $x \to \infty$ lo suficientemente rápido, a saber:

$$G(x+1) \leqslant \alpha G(x), \quad \text{donde} \quad \alpha < 1. \tag{5.6}$$

En este caso todas las probabilidades de paso $q(x) = 1 - \frac{G(x+1)}{G(x)}$ no son menores que $1 - \alpha > 0$, y por consiguiente, *el coeficiente de ergodicidad* de tal cadena de Markov será positivo:

$$\boldsymbol{k}(1) \geqslant 1 - \alpha.$$

Con esta condición las probabilidades $p_t(x) = \mathbf{P}\{\xi(t) = x\}$ para $t \to \infty$, convergen hacia la probabilidad estacionaria $p^*(x)$:

$$\lim_{t\to\infty} p_t(x) \frac{G(x)}{\mu}, \quad x = 0, 1, \ldots, \tag{5.7}$$

*y aún más, tiene lugar la siguiente valuación de la velocidad de con*vergencia (véase el teorema 4 del § 2 de este capítulo)

$$\left| p_t(x) - \frac{G(x)}{\mu} \right| \leqslant \frac{1}{\alpha} e^{-\alpha t}. \tag{5.8}$$

Examinemos el número medio de restablecimientos durante el tiempo desde t hasta $t+s$ (de otro modo, el número medio de caídas en el estado 0 durante ese tiempo), que es

$$n(t+s) - n(t) = \mathbf{M}[\nu(t+s) - \nu(t)] = \sum_{u=t+1}^{t+s} p_u(0)$$

(anteriormente designamos por $\nu(t)$ al número de restablecimientos en el intervalo de tiempo de 0 a t). Está claro, que para cualquier intervalo limitado $(t, t+s]$

$$n(t+s) - n(t) \to \frac{s}{\mu} \quad \text{para} \quad t \to \infty, \tag{5.9}$$

ya que $p_u(0) \to \frac{1}{\mu}$ y, además, como se deduce de la desigualdad (5.8),

$$\left| [n(t+s) - n(t)] - \frac{s}{\mu} \right| \leqslant \frac{s}{\alpha} e^{-\alpha t}. \tag{5.10}$$

La relación (5.9) muestra que *para un proceso prolongado de restablecimiento el número medio de restablecimientos durante el tiempo s es aproximadamente igual a la magnitud* $\frac{s}{\mu}$, *donde* μ *es el tiempo medio de servicio del aparato aislado.*

Hemos convenido en considerar que si el aparato trabajó en el momento t (múltiplo de h) entonces él se cambia no antes de un tiempo h ($h = 1$). Como tiempo durante el cual todavía trabaja el aparato después del momento t, tomamos la magnitud

$$\eta(t) = S_{\nu(t)+1} - h - t, \tag{5.11}$$

donde $S_{\nu(t)+1}$ es el momento de restablecimiento siguiente a t (momento de cambio del aparato examinado).

Está intuitivamente claro, que durante el proceso prolongado de restablecimiento ($t \to \infty$) la distribución de probabilidades de la magnitud $\eta(t)$ deberá ser aproximadamente igual a la de la magnitud $\xi(t) = t - S_{\nu(t)}$. De base para ello sirve la «reversibilidad» durante el tiempo del proceso de restablecimiento; los momentos de restablecimiento se pueden contar en orden contrario (el número de restablecimientos $n = \nu(t)$ crece ilimitadamente, y las magnitudes

$$S_1 = \xi_1,\ S_2 = \xi_1 + \xi_2,\ \ldots,\ S_n = \xi_1 + \ldots + \xi_n$$

y

$$S'_1 = \xi_n,\ S'_2 = \xi_n + \xi_{n+1},\ \ldots,\ S'_n = \xi_n + \ldots + \xi_1$$

tienen la misma distribución conjunta).

Señalemos, que en el caso cuando el tiempo de espera de la rotura del aparato tiene la distribución exponencial de probabilidades

$$G(x) = e^{-\lambda x}, \quad x = 0,\ 1,\ \ldots,$$

ya desde el principio

$$\mathbf{P}\{\eta(t) = x\} = \frac{G(x)}{\mu}, \quad x = 0,\ 1,\ \ldots \left(\mu = \frac{1}{1-e^{-\lambda}}\right)$$

(véase el punto 2 del § 2, cap. II).

Sin limitarnos a las observaciones hechas anteriormente, damos un fundamento riguroso al hecho de que *la magnitud* $\eta(t)$ *para* $t \to \infty$, *tiene la distribución límite de probabilidades*

$$\lim_{t\to\infty} \mathbf{P}\{\eta(t) = x\} = \frac{G(x)}{\mu}, \quad x = 0,\ 1,\ \ldots \tag{5.12}$$

Suponiendo que $S_0 = 0$, tenemos

$$\mathbf{P}\{\eta(t) = x\} = \sum_{n=0}^{\infty} \mathbf{P}\{S_n \leqslant t,\ S_{n+1} = t + 1 + x\},$$

donde (según la fórmula de la probabilidad total)

$$\begin{aligned}
\mathbf{P}\{S_n \leqslant t,\ S_{n+1} = t+1+x\} &= \\
&= \sum_{0\leqslant s\leqslant t} \mathbf{P}\{S_{n+1} = t+1+x \mid S_n = s\}\,\mathbf{P}\{S_n = s\} = \\
&= \sum_{0\leqslant s\leqslant t} \mathbf{P}\{\xi_{n+1} = t+1+x-s\}\,\mathbf{P}\{S_n = s\} = \\
&= \sum_{0\leqslant s\leqslant t} P(t+1+x-s)\,\mathbf{P}\{S_n = s\}.
\end{aligned}$$

Se ve que

$$\begin{aligned}
\mathbf{P}\{\eta(t) = x\} &= \sum_{0\leqslant s\leqslant t} P(t-s+1+x)\,p_s(0) = \\
&= \sum_{0\leqslant u\leqslant t} P(u+1+x)\,p_{t-u}(0),
\end{aligned} \tag{5.13}$$

ya que

$$\sum_{n=0}^{\infty} \mathbf{P}\{S_n = s\} = \mathbf{P}\{\xi(s) = 0\} = p_s(0)$$

es la probabilidad de que el momento s sea el momento de restablecimiento. Tomando en consideración que para $t \to \infty$

$$\sum_{u=0}^{t} P(u+1+x) \to G(x) \quad \text{y} \quad p_{t-u}(0) \to \frac{1}{\mu}$$

de la igualdad (5.13) obtenemos la fórmula (5.12).

Consideremos ahora el número de restablecimientos $\nu(t+s)-\nu(t)$ durante el intervalo de tiempo $(t, t+s]$. Evidentemente,

$$\mathbf{P}\{\nu(t+s)-\nu(t)=0\}=\mathbf{P}\{\eta(t)\geqslant s\}, \tag{5.14}$$

y para $n=1, 2, \ldots$ la probabilidad condicional de que durante el intervalo de tiempo $(t, t+s]$ ocurran n restablecimientos con la condición de que el primer restablecimiento se realiza en el momento $t+u+h$, es

$$\mathbf{P}\{\nu(t+s)-\nu(t)=n \mid \eta(t)=u\}=$$
$$=\begin{cases} 0, & u\geqslant s \\ \mathbf{P}\{\nu(t+s)-\nu(t+u+1)=n-1\}, & u<s. \end{cases}$$

Utilizando la fórmula de la probabilidad total, obtenemos

$$\mathbf{P}\{\nu(t+s)-\nu(t)=n\}=$$
$$=\sum_{0\leqslant u<s}\mathbf{P}\{\nu(t+s)-\nu(t+u+1)=n-1\}\,\mathbf{P}\{\eta(t)=u\},$$

que junto con la fórmula (5.14) nos da la relación recurrente para la distribución de probabilidades del número de restablecimientos durante cualquier intervalo $(t, t+s]$. En particular, *para las probabilidades límites*

$$P_n(s)=\lim_{t\to\infty}\mathbf{P}\{\nu(t+s)-\nu(t)=n\} \tag{5.15}$$

tenemos *las relaciones recurrentes siguientes*:

$$P_0(s)=\frac{1}{\mu}\sum_{u\geqslant s}G(u),$$

$$P_n(s)=\frac{1}{\mu}\sum_{0\leqslant u<s}P_{n-1}(s-u-1)\,G(u),\quad n=1, 2, \ldots \tag{5.16}$$

Señalemos que las distribuciones de probabilidades de las magnitudes

$$\xi(t)=t-S_{\nu(t)},\quad \eta(t)=S_{\nu(t)+1}-t-h,$$
$$\nu(t+s)-\nu(t) \tag{5.17}$$

coincidirán con las distribuciones estacionarias límites (descritas por las fórmulas (5.7), (5.12), (5.16)), si en el momento inicial $t=0$, el primer aparato ya ha trabajado un tiempo aleatorio determinado $\xi(0)$, en donde la magnitud aleatoria $\xi(0)$ tiene la distribución estacionaria $\mathbf{P}\{\xi(0)=x\}=G(x)/\mu$, $x=0, 1, \ldots$

Como conclusión de este punto es necesario decir que los resultados obtenidos anteriormente, también tienen lugar en el caso general¹). En particular, para los sumandos $\xi_1, \xi_2, \ldots$ distribuidos continuamente con la densidad de probabilidad $p(x)$, las distribuciones de *probabilidades de las magnitudes correspondientes*

$$\xi(t) = t - S_{\nu(t)}, \quad \eta(t) = S_{\nu(t)+1} - t$$

convergen débilmente hacia la distribución límite con la densidad

$$p^*(x) = \frac{G(x)}{\mu}, \quad 0 \leqslant x < \infty, \tag{5.18}$$

donde

$$G(x) = \int_x^\infty p(y)\,dy, \quad \mu = \int_0^\infty xp(x)\,dx = \int_0^\infty G(x)\,dx. \tag{5.19}$$

2. Sucesión de sumas de las magnitudes independientes. Distribución del máximo. En este punto examinemos algunas propiedades generales de la sucesión de sumas

$$S_n = \sum_{k=1}^{n} \xi_k, \quad n = 1, 2, \ldots, \tag{5.20}$$

de *las magnitudes aleatorias independientes* ξ_k *con igual distribución* de probabilidades. Para mayor claridad nos podemos imaginar que una determinada partícula fluctúa aleatoriamente por la recta real, desplazándose a cada paso en una magnitud ξ_n, entonces $S_n = \sum_{k=1}^{n} \xi_k$, $n = 1, 2, \ldots$ $(S_0 = 0)$ será la trayectoria de esta fluctuación aleatoria. Supondremos que las magnitudes $\xi_1, \xi_2, \ldots$ tienen una esperanza matemática a distinta de cero, considerando que

$$a = \mathbf{M}\xi_1 < 0.$$

Según la ley reforzada de los grandes números $\frac{S_n}{n} \to a < 0$; por consiguiente, con la probabilidad 1

$$S_n \to -\infty \quad \text{para} \quad n \to \infty,$$

y existe tal ν dependiente del caso, en el cual $S_n < 0$, para todos los valores $n \geqslant \nu$. Esto indica que el máximo

$$\zeta = \text{máx}\,(S_0, S_1, \ldots) \tag{5.21}$$

de la trayectoria de fluctuación aleatoria es una magnitud finita que coincide con el valor máximo de la sucesión $S_0, S_1, \ldots, S_\nu$.

¹) Referente a esta cuestión (y también en relación con los puntos 2, 3 siguientes) véase por ejemplo, el libro de V. Feller «*Introducción a la teoría de probabilidades y sus aplicaciones*» («An Introduction to the Probability Theory and its Applications» N. Y., 1950, 1954), tomo II, IL, Moscú, 1967.

Más adelante hallaremos el enlace de la distribución de probabilidades de esta magnitud ζ con un determinado proceso de restablecimiento S_0^*, S_1^*, ..., construido según la fluctuación aleatoria S_0, S_1, ... (véase más adelante (5.31)), lo que permite deducir las fórmulas explícitas (5.37), (5.40) así como también la fórmula (5.41), que se utilizarán al estudiar los procesos del servicio en masa.

Supongamos que

$$\tilde{S}_0 = 0,\ \tilde{S}_1 = \xi_2,\ \tilde{S}_2 = \xi_2 + \xi_3,\ \ldots$$

y

$$\tilde{\zeta} = \text{máx}\,(\tilde{S}_0,\ \tilde{S}_1,\ \ldots).$$

Claro está, que la magnitud aleatoria $\tilde{\zeta}$ tiene la misma distribución de probabilidades que la magnitud aleatoria ζ. Evidentemente, para cualquier $z \geqslant 0$, el suceso $\{\zeta \leqslant z\}$ se realiza entonces y sólo entonces, cuando $\xi_1 \leqslant z$ y $\tilde{\zeta} \leqslant z - \xi_1$. Ya que la magnitud $\tilde{\zeta}$ no depende de ξ_1, para cada valor fijado $\xi_1 = x$ $(x \leqslant z)$, tenemos que

$$\mathbf{P}\,\{\tilde{\zeta} \leqslant z - \xi_1 \mid \xi_1\} = F_\zeta\,(z - \xi_1),$$

donde $F_\zeta\,(z) = \mathbf{P}\,\{\zeta \leqslant z\}$ es la función de distribución de la magnitud ζ, y de tal modo

$$\mathbf{P}\{\zeta \leqslant z \mid \xi_1\} = \begin{cases} 0, & \text{si } \xi_1 > z, \\ F_\zeta(z-\xi_1), & \text{si } \xi_1 \leqslant z. \end{cases}$$

Teniendo en cuenta que para el valor ζ no negativo, la función de distribución $F_\zeta\,(z)$ es igual a cero para $z < 0$, la relación obtenida anteriormente se puede escribir en la forma

$$\mathbf{P}\,\{\zeta \leqslant z \mid \xi_1\} = F_\zeta\,(z - \xi_1).$$

Utilizando la fórmula de la probabilidad total (véase (4.20), cap. I) obtenemos como resultado la siguiente ecuación para la función $F_\zeta\,(z)$:

$$F_\zeta\,(z) = \mathbf{M}F_\zeta\,(z - \xi_1), \quad z \geqslant 0. \tag{5.22}$$

Esta ecuación en el caso de magnitudes discretas ξ_1, ξ_2, ... significa que

$$F_\zeta(z) = \sum_{x \leqslant z} F_\zeta\,(z-x)\,P_{\xi_1}(x) = \sum_{y \geqslant 0} P_{\xi_1}(z-y)\,F_\zeta\,(y), \tag{5.23}$$

y en el caso cuando las magnitudes ξ_1, ξ_2, ... tienen la densidad de probabilidad $p_{\xi_1}(x)$,

$$F_\zeta(z) = \int\limits_{-\infty}^{z} F_\zeta\,(z-x)\,p_{\xi_1}(x)\,dx = \int\limits_{0}^{\infty} p_{\xi_1}(z-y)\,F_\zeta\,(y)\,dy. \tag{5.24}$$

Mostremos que el suceso $\{\zeta = 0\}$ se realiza con la probabilidad positiva $q > 0$:

$$q = F_{\zeta}(0) > 0 \tag{5.25}$$

(es decir, al salir la partícula desde el punto $x = 0$ con probabilidad $q \neq 0$, nunca cae en el semieje positivo $x > 0$).

En efecto, supongamos lo contrario. Designemos por z_0 la cara inferior de aquellos valores de z, para los cuales $F_{\zeta}(z) > 0$; de acuerdo con nuestra suposición, será $z_0 > 0$, o bien $z_0 = 0$ y $F_{\zeta}(0) = 0$. Si $z_0 > 0$, para c u a l q u i e r z, $0 \leqslant z \leqslant z_0$ de la ecuación (5.22) obtenemos

$$F_{\zeta}(z) = \mathbf{M}F_{\zeta}(z - \xi_1) = 0.$$

Evidentemente, en el caso de la magnitud no negativa $F_{\zeta}(r - \xi_1)$, cuando $F(z - \xi_1) > 0$ para $z - \xi_1 > z_0$, la igualdad indicada puede existir sólo entonces, cuando $z - \xi_1 \leqslant z_0$, $\xi_1 \geqslant z - z_0$ con la probabilidad 1, de donde deducimos, que para $z_0 > 0$ es $\xi_1 \geqslant 0$ con la probabilidad 1. Pero esto contradice nuestra condición inicial $a = \mathbf{M}\xi_1 < 0$, y por consiguiente, deberá ser $z_0 = 0$. Exactamente igual conduce a una contradicción la suposición de que $F_{\zeta}(0) = 0$ con la condición $\mathbf{M}\xi_1 < 0$, en el caso de $z_0 = 0$.

La magnitud ζ toma rigurosamente un valor positivo con la probabilidad $p = 1 - q$. Esto quiere decir que en el caso de fluctuación aleatoria considerado, la partícula cae tarde o temprano en la parte positiva de la recta real con la probabilidad p; designemos por ξ^* su posición en la primera caída en el intervalo $(0, \infty)$. La magnitud ξ_1^* sólo existe con la probabilidad p:

$$\mathbf{P}\{\xi_1^* \in (0, \infty)\} = p \quad (= 1 - q)$$

(consideraremos convencionalmente, que la magnitud ξ_1^* «desaparece» con la probabilidad $q = 1 - p$).

Designemos por τ_1 el momento de caída de la partícula en el intervalo $(0, \infty)$; $\tau_1 = \nu_1^*$, donde ν_1^* es el número de pasos hasta que el desplazamiento de la partícula se haga positiva por primera vez: $S_n - S_0 \leqslant 0$ para $n < \tau_1$, $S_{\tau_1} - S_0 = \xi_1^* > 0$.

Supongamos que $S_1^* = S_{\tau_1}$. El desplazamiento de la partícula se realiza, después del momento τ_1, según la misma ley como después del momento inicial $\tau_0 = 0$. En particular, con la misma probabilidad $p = 1 - q$ que antes, la partícula cae, en un momento determinado τ_2 (después de un número aleatorio de pasos $\nu_2^* = \tau_2 - \tau_1$), en el punto $S_2^* = S_{\tau_2}$, m á s a l a d e r e c h a del punto inicial S_1^* (ν_2^* es el número de pasos después del momento τ_1 antes de que el desplazamiento de la partícula se haga de nuevo positivo: $S_n - S_1^* \leqslant 0$ para $n < \tau_2$, $S_{\tau_2} - S_1^* = \xi_2^* < 0$). Empezando desde el momento $\tau_2 = \nu_1^* + \nu_2^*$ se observa el mismo cuadro: con la probabilidad p, en un momento determinado τ_3 (después de un número aleatorio de pasos $\nu_3^* = \tau_3 - \tau_2$) la partícula cae en el punto $S_3^* = S_{\tau_3}$, más a la derecha del punto inicial S_2^* (ν_2^* es el número de pasos después del momento τ_2 antes de que el desplazamiento de la partícula se haga de nuevo positivo: $S_n - S_2^* \leqslant 0$ para $n \leqslant \tau_3$, $S_{\tau_3} - S_2^* = \xi_3^* > 0$). Continuando este proceso más adelante llegamos a la sucesión de sumas $S_n^* = S_{\tau_n}$,

$$S_n^* = \sum_{k=1}^{n} \xi_k^*, \quad n = 1, 2, \ldots \tag{5.26}$$

de las magnitudes aleatorias *positivas* ξ_k^*, $k = 1, 2, \ldots$, donde ξ_{k+1}^* representa en sí el desplazamiento de la partícula (después de su salida desde el punto S_k^*) en el momento τ_{k+1}, cuando este desplazamiento se hace positivo *por primera vez*: $S_n - S_k^* \leqslant 0$, para $n < \tau_{k+1}$, $S_{\tau_{k+1}} - S_k^* = \xi_{k+1}^* > 0$.

Ya que las magnitudes $\xi_1, \xi_2, \ldots$, son independientes y tienen una distribución igual de probabilidades, el desplazamiento de la partícula después de su salida del punto inicial S_k^* se realiza según la misma ley, que después de la salida del punto $S_0^* = 0$ (independientemente de su comportamiento antes del momento τ_k); por eso *las magnitudes* $\xi_1^*, \xi_2^*, \ldots$ definidas por nosotros, *son independientes y tienen una misma distribución de probabilidades*.

En la terminología del punto 1 anterior, la sucesión S_n^*, $n = 1, 2, \ldots$, *es un proceso de restablecimiento. Con ello, el primer* momento S_1^* de restablecimiento sólo existe con la probabilidad $\mathbf{P}\{S_1^* \in (0, \infty)\} = p$, y hablando en general, al existir los momentos $S_1^*, \ldots, S_{n-1}^*$, el momento S_n^* de restablecimiento siguiente, sólo existe con la misma probabilidad p:

$$\mathbf{P}\{S_{n+1}^* \in (0, \infty) \mid S_n^* \in (0, \infty)\} = p.$$

De aquí obtenemos, para la probabilidad $\mathbf{P}\{S_{n+1}^* \in (0, \infty)$ de existencia del restablecimiento S_{n+1}^* en el momento $n+1$, que

$$\mathbf{P}\{S_{n+1}^* \in (0, \infty)\} = \mathbf{P}\{S_{n+1}^* \in (0, \infty) \mid S_n^* \in (0, \infty)\}\, \mathbf{P}\{S_n^* \in (0, \infty)\} = p \cdot \mathbf{P}\}S_n^* \in (0, \infty)\}, \quad n = 1, 2, \ldots,$$

y como resultado, tenemos que

$$\mathbf{P}\{S_n^* \in (0, \infty)\} = p^n, \quad n = 1, 2, \ldots \tag{5.27}$$

Puesto que la serie $\sum_{n=1}^{\infty} \mathbf{P}\{S_n^* \in (0, \infty)\}$ es convergente, según el lema de Borel—Cantelli se realiza con la probabilidad 1, sólo un número finito de sucesos $\{S_n^* \in (0, \infty)\}$, es decir, *el proceso de restablecimiento* S_n^*, $n = 1, 2, \ldots$, *se interrumpe en un determinado paso.* Se ve fácilmente que no habrá ningún restablecimiento con la probabilidad $q = 1 - p$, y sólo habrá un restablecimiento en todo el intervalo infinito $[0, \infty)$ con la probabilidad $p(1-p)$, y en general, el proceso de restablecimiento S_n^*, $n = 1, 2, \ldots$, se interrumpe después de n restablecimientos con la probabilidad $p^n(1-p)$. Suponiendo que $S_0^* = 0$ y considerando también a S_0^*, como momento de restablecimiento obtenemos, para el número total N de restablecimientos (en todo el intervalo infinito $[0, \infty)$) la distribución geométrica de probabilidades:

$$\mathbf{P}\{N \geqslant n\} = \mathbf{P}\{S_{n-1}^* \in (0, \infty)\} = p^{n-1} \tag{5.28}$$
$$\mathbf{P}\{N = n\} = qp^{n-1}, \quad n = 1, 2, \ldots$$

Designemos por $N(x)$ al número de restablecimientos en el intervalo $[0, x]$, $0 \leqslant x < \infty$. Evidentemente,

$$\mathbf{P}\{N(x) \geqslant n\} = \mathbf{P}\{S_{n-1}^* \in [0, x]\}$$

y el valor medio $M(x) = \mathbf{M}[N(x)]$ es (compárese con (5.5))

$$M(x) = \sum_{n=0}^{\infty} \mathbf{P}\{N(x) > n\} = \sum_{n=0}^{\infty} \mathbf{P}\{S_n^* \in [0, x]\}. \tag{5.29}$$

Para el número total N de restablecimientos

$$\mathbf{M}[N] = \sum_{n=0}^{\infty} \mathbf{P}\{N > n\} = \sum_{n=0}^{\infty} p^n = \frac{1}{q} \tag{5.30}$$

Evidentemente, la magnitud $\zeta = \text{máx}(0, S_1, S_2, \ldots)$ que nos interesa coincide con el valor máximo de la sucesión monótonamente creciente $0, S_1^*, S_2^*, \ldots$, en otras palabras, con el último momento S_{N-1}^* de restablecimiento. Por eso para cualquier $x > 0$

$$\mathbf{P}\{0 < \zeta \leqslant x\} = \sum_{n>1} \mathbf{P}\{N = n, \quad S_{n-1}^* \in [0, x].$$

Tenemos que

$$\mathbf{P}\{N = n, S_{n-1}^* \in [0, x]\} = \mathbf{P}\{S_{n-1}^* \in [0, x] \mid N = n\} \mathbf{P}(N = n),$$

donde el suceso $\{N = n\}$ significa que $\{S_{n-1}^* \in (0, \infty)\}$, y la magnitud ξ_n^* «desaparece». Pero $S_{n-1}^* = \sum_{k=1}^{n-1} \xi_k^*$ no depende de ξ_n^* ($\xi_1^*, \xi_2^*, \ldots$ es la sucesión de magnitudes independientes) y se ve fácilmente que

$$\mathbf{P}\{S_{n-1}^* \in [0, x] \mid N = n\} = \mathbf{P}\{S_{n-1}^* \in [0, x] \mid S_{n-1}^* (0, \infty)\}.$$

De las fórmulas obtenidas anteriormente (5.27), (5.28) deducimos luego que

$$\mathbf{P}\{S_{n-1}^* \in [0, x] \mid S_{n-1}^* \in (0, \infty)\} \mathbf{P}\{N = n\} =$$
$$= \frac{\mathbf{P}\{S_{n-1}^* \in [0, x]\}}{\mathbf{P}\{S_{n-1}^* \in (0, \infty)\}} \mathbf{P}\{N = n\} = q\mathbf{P}\{S_{n-1}^* \in [0, x]\}, \quad n > 1.$$

Teniendo en cuenta la igualdad (5.29), como resultado obtenemos la siguiente relación:

$$F_\zeta(x) = qM(x) \quad (q = F_\zeta(0)). \tag{5.31}$$

Para la función $M(x)$, o sea, el número medio de restablecimientos en el intervalo $[0, x]$, obtenemos la fórmula (5.29) en la cual

$$\mathbf{P}\{S_0^* \in [0, x]\} = F^{0*}(x) \equiv 1,$$

y para $n > 0$

$$\mathbf{P}\{S_n^* \in [0, x]\} = F^{n*}(x), \quad 0 \leqslant x < \infty,$$

es la función de distribución de la suma de las magnitudes independientes $\xi_1^*, \ldots, \xi_n^*$, de igual distribución. Conociendo la función de distribución $F^*(x) = \mathbf{P}\{\xi_1^* \leqslant x\}$, en principio se puede determinar también que

$$M(x) = \sum_{n=0}^{\infty} F^{n*}(x), \quad 0 \leqslant x < \infty \tag{5.32}$$

($F^{n*}(x) = 0$ cuando $x < 0$ para todos los $n = 0, 1, \ldots$).

Deduzcamos una relación útil que, en el caso de la distribución exponencial de las magnitudes iniciales $\xi_1, \xi_2, \ldots$ permitirá obtener expresiones explícitas para la distribución de probabilidades de las magnitudes ξ_i^* (véase (5.35) y más adelante).

Precisamente, en la fluctuación aleatoria inicial (5.20), la magnitud ξ_1^* significa la posición de la partícula en la primera caída en el intervalo $(0, \infty)$: $\xi_1^* = S_{\tau_1}$, donde τ_1 es el número de pasos hasta la primera caída en el intervalo $(0, \infty)$. Evidentemente, para cualquier $x \geqslant 0$

$$\mathbf{P}\ \tau_1 = n,\ \xi_1^* > x\} = \mathbf{P}\{S_1 \leqslant 0, \ldots, S_{n-1} \leqslant 0,\ S_n > x\},$$

y para los valores fijados $S_1 \leqslant 0, \ldots, S_{n-1} \leqslant 0$, el suceso $\{\tau_1 = n,\ \xi_1^* < x\}$ significa que $\xi_n = S_n - S_{n-1} > x - S_{n-1}$. Por consiguiente, $\mathbf{P}\{\tau_1 = n,\ \xi_1^* > x \mid S_1, \ldots, S_{n-1}\} = G_{\xi_1}(x - S_{n-1}) \times \chi_{(S_1 \leqslant 0, \ldots, S_{n-1} \leqslant 0)}$, donde $G_{\xi_1}(x) = 1 - F_{\xi_1}(x)$, $F_{\xi_1}(x)$ es la función de distribucíon de las magnitudes $\xi_1, \xi_2, \ldots$, y $\chi_{(s_1 \leqslant 0, \ldots, s_{n-1} \leqslant 0)}$ es el indicador del suceso[1] $\{S_1 \leqslant 0, \ldots, S_{n-1} \leqslant 0$. Según la fórmula de la probabilidad total

$$\mathbf{P}\{\tau_1 = n,\ \xi_1^* > x\} = \mathbf{M}\,[G_{\xi_1}(x - S_{n-1}) \cdot \chi_{(S_1 \leqslant 0, \ldots, S_{n-1} \leqslant 0)}], \quad n = 1, 2, \ldots;\ x \geqslant 0. \tag{5.33}$$

[1]) Recordemos que el indicador del suceso A se determina como

$$\chi_A = \begin{cases} 1 & \text{para la aparición de } A, \\ 0 & \text{en caso contrario.} \end{cases}$$

Supongamos que cada una de las magnitudes $\xi_1, \xi_2, \ldots$, está distribuida en el intervalo $(0, \infty)$ según la ley exponencial:

$$G_{\xi_1}(x) = \mathbf{P}\{\xi_1 > x\} = Ae^{-\lambda x}, \quad x \geqslant 0, \tag{5.34}$$

donde $A = \mathbf{P}\{\xi_1 > 0\}$. En este caso la relación (5.33), en la cual S_{n-1} figura solamente, para la condición $S_{n-1} \leqslant 0$, nos da las siguientes igualdades:

$$\mathbf{P}\{\tau_1 = n, \ \xi_1^* > x\} =$$
$$= \mathbf{M}\,[Ae^{\lambda x - \lambda S_{n-1}}\chi_{\{S_1 \leqslant 0, \ldots, S_{n-1} \leqslant 0\}}] = B_n e^{-\lambda x},$$

donde

$$B_n = \mathbf{M}\,\{Ae^{-\lambda S_{n-1}}\chi_{\{S_1 \leqslant 0, \ldots, S_{n-1} \leqslant 0\}}\};$$

sumando estas igualdades por todos los valores de $n = 1, 2, \ldots$ obtenemos que

$$\mathbf{P}\{\xi_1^* > x\} = pe^{-\lambda x}, \quad x \geqslant 0, \tag{5.35}$$

donde la constante $p \sum_{n=1}^{\infty} B_n$ es, evidentemente,

$$p = \mathbf{P}\{\xi_1^* > 0\} = 1 - q \quad (q = F_\zeta(0)).$$

Examinemos más detalladamente, el caso de las magnitudes distribuidas continuamente $\xi_1, \xi_2, \ldots$ (para las magnitudes discretas los resultados son completamente análogos). De acuerdo con la fórmula (5.35) para la densidad de probabilidad

$$p_{\xi_1}(x) = A\lambda e^{-\lambda x}, \quad x > 0,$$

las magnitudes $\xi_1^*, \xi_2^*, \ldots$ estarán distribuidas según la ley exponencial con la densidad

$$p^*(x) = p\lambda e^{-\lambda x}, \quad x > 0, \tag{5.36}$$

que se diferencia de la densidad exponencial corriente, sólo en el factor p, que indica que la magnitud ξ_1^* «desaparece» con la probabilidad $q = 1 - p$. Como sabemos (véase (2.16), cap. II), la suma de las magnitudes independientes $\xi_1^*, \ldots, \xi_n^*$ con la misma distribución exponencial (5.36) tiene la densidad de probabilidad

$$p^{n*}(x) = p^n \lambda \frac{(\lambda x)^{n-1}}{(n-1)} e^{-\lambda x}, \quad x > 0.$$

La densidad $p^{n*}(x)$ es la derivada de $F^{n*}(x)$ y, como se ve fácilmente,

$$m(x) = \sum_{n=1}^{\infty} p^{n*}(x) = p\lambda e^{-\lambda x} \sum_{n=0}^{\infty} \frac{(p\lambda x)^n}{n!} = p\lambda e^{-\lambda(1-p)x}, \quad x > 0$$

es la derivada de la función $M(x) = \sum_{n=0}^{\infty} F^{n*}(x)\,(M(0)=1)$, y, por consiguiente,

$$M(x) = 1 + \int_0^x m(x)\,dx = \frac{1}{q}(1 - pe^{-\lambda q x}), \quad q = 1 - p.$$

Como resultado, de (5.31) obtenemos que *la función de distribución del máximo* $\zeta = \text{máx}\,(0, S_1, S_2, \ldots)$ *tiene la forma*

$$F_\zeta(x) = 1 - (1-q)\,e^{-\lambda q x}, \quad x \geqslant 0, \tag{5.37}$$

donde la probabilidad $q = F_\zeta(0)$ *puede ser determinada de la ecuación* (5.24):

$$q = \int_{-\infty}^{0} [1 - (1-q)\,e^{\lambda q x}]\,p_{\xi_1}(x)\,dx. \tag{5.38}$$

En el caso cuando la densidad de probabilidad para $x \leqslant 0$ tiene el mismo tipo exponencial que para $x > 0$, a saber:

$$p_{\xi_1}(x) = B\mu e^{\mu x}, \quad x > 0, \tag{5.39}$$

de la ecuación (5.38) se deduce fácilmente que $q = B\frac{\lambda+\mu}{\lambda} - \frac{\mu}{\lambda}$; teniendo en cuenta la igualdad $B = \mathbf{P}\{\xi_1 \leqslant 0\} = 1 - A$, tenemos

$$q = B - A\frac{\mu}{\lambda}. \tag{5.40}$$

Señalemos que el caso examinado $\mathbf{M}\xi_1 = -\frac{B}{\mu} + \frac{A}{\lambda}$, y se ve directamente, que para $\mathbf{M}\xi_1 < 0$ el valor indicado $q = -\mu\mathbf{M}\xi_1$ es rigurosamente positivo.

Volvamos a examinar el caso general de fluctuación aleatoria (5.20), suponiendo como antes, que $\mathbf{M}\xi_1 < 0$.

Mostremos que *la magnitud* ν_0 *es el número de pasos hasta la primera caída (después de salir de* $S_0 = 0$*) en la zona* $x \leqslant 0$ *y tiene una esperanza matemática finita* $\mathbf{M}\nu_0 < \infty$; *y además,*

$$\mathbf{M}\nu_0 = \frac{1}{q} \quad (q = F_\zeta(0)). \tag{5.41}$$

El hecho de que la magnitud ν_0 es finita con la probabilidad 1, fue señalado desde el mismo comienzo (recuérdese que $S_n \to -\infty$ para $n \to \infty$). Para la deducción de la igualdad (5.41) examinemos los sucesos

$$A_n = \{S_n > 0,\ S_n > S_1,\ \ldots,\ S_n > S_{n-1}\}, \quad n = 1, 2, \ldots$$

Evidentemente cada uno de ellos significa que S_n es el momento de restablecimiento en el proceso de restablecimiento (5.26). Teniendo en cuenta que cada momento de restablecimiento en (5.26) coincide con un valor determinado S_n, e introduciendo el suceso cierto A_0 que significa que $S_0^* = S_0$ es también un momento de restablecimiento, para el número total N de restablecimientos y el valor medio $\mathbf{M}[N]$ tendremos que $N = \sum_{n=0}^{\infty} \chi_{A_n}$, donde χ_{A_n} es el indicador del suceso A_n y

$$\mathbf{M}[N] = \sum_{n=0}^{\infty} \mathbf{M}\chi_{A_n} = \sum_{n=0}^{\infty} \mathbf{P}(A_n) = \frac{1}{q} \qquad (5.42)$$

(véanse (5.30) y (4.39), cap. I). Refiriéndonos a las magnitudes $S_k' = S_n - S_{n-k}$, $k = 1, \ldots, n$:

$$S_1' = \xi_1,\ S_2' = \xi_n + \xi_{n-1},\ \ldots,\ S_n' = \xi_n + \ldots + \xi_1,$$

que tienen la misma distribución de probabilidades que las magnitudes

$$S_1 = \xi_1,\ S_2 = \xi_1 + \xi_2,\ \ldots,\ S_n = \xi_1 + \ldots + \xi_n,$$

se ve fácilmente que cada uno de los sucesos A_n coincide con el suceso correspondiente $\{S_1' > 0, \ldots, S_n' > 0\}$, que tiene la misma probabilidad que el suceso

$$B_n = \{S_1 > 0, \ldots, S_n > 0\}, \quad n = 1, 2, \ldots$$

Pero cada uno de los sucesos B_n significa que la partícula no cayó en la zona $x \leqslant 0$ en los primeros n pasos, es decir, $B_n = \{\nu_0 > n\}$. Teniendo en cuenta que $\mathbf{P}(B_n) = \mathbf{P}(A_n)$, $n = 0, 1, \ldots$, de la igualdad (5.42) obtenemos para la esperanza matemática de la magnitud positiva ν_0, que

$$\mathbf{M}\nu_0 = \sum_{n=0}^{\infty} \mathbf{P}\{\nu_0 > n\} = \sum_{n=0}^{\infty} \mathbf{P}(A_n) = \frac{1}{q}.$$

La fórmula (5.41) queda demostrada.

3. Procesos aleatorios en los sistemas con una línea de servicio. Supongamos que en un sistema de servicio determinado se reciben demandas en los momentos aleatorios de tiempo $\tau_1, \tau_2, \ldots$ Supongamos que es imposible la recepción simultánea de distintas demandas y que los intervalos $\xi_1 = \tau_2 - \tau_1$, $\xi_2 = \tau_3 - \tau_2, \ldots$ entre los momentos $\tau_1, \tau_2, \ldots$ son magnitudes aleatorias independientes, que tienen una misma distribución de probabilidades. Supongamos, además de esto, que para el servicio de la demanda n se invierte el tiempo η_n y $\eta_1, \eta_2, \ldots$, son magnitudes aleatorias independientes con una misma distribución de probabilidades, también independientes de los momentos de tiempo $\tau_1, \tau_2, \ldots$ (tenemos en cuenta que η_n es el tiempo «neto» para el servicio de la demanda n, sin contar el tiempo aleatorio de espera $\mathscr{H}_n$ desde el momento de recepción τ_n hasta el comienzo del servicio).

El tiempo total invertido por la demanda n en el sistema de servicio, en nuestras designaciones $\mathscr{H}_n + \eta_b$. Supongamos que si la siguiente demanda $(n+1)$ se recibe después del tiempo $\xi_n \geqslant \mathscr{H}_n + \eta_n$, entonces ésta encuentra libre al sistema de servicio el cual *inmediatamente empieza a servirla*, es decir, $\mathscr{H}_{n+1} = 0$; y si $\xi_n < \mathscr{H}_n + \eta_n$, entonces en el momento $\tau_{n+1} = \tau_n + \xi_n$ el sistema todavía está ocupado por el servicio de las demandas anteriores, y hasta el comienzo del servicio, la demanda $n+1$ debe esperar un tiempo $\mathscr{H}_{n+1} = \mathscr{H}_n + \eta_n - \xi_n$.

Ahora nos ocuparemos con la ley del proceso de servicio prolongado y ante todo con la distribución de probabilidades de la magnitud $\mathscr{H}_n$, o sea, con el tiempo de espera del comienzo de servicio para la demanda n (para $n \to \infty$).

Supongamos que

$$\Delta_n = \eta_n - \xi_n; \; S_n = \sum_{k=1}^{n} \Delta_k, \quad n = 1, 2, \ldots \tag{5.43}$$

Como fue indicado, las magnitudes $\mathscr{H}_n$ están ligadas con las magnitudes aleatorias i n d e p e n d i e n t e s Δ_n $(n = 1, 2, \ldots)$ por las siguientes relaciones:

$$\mathscr{H}_1 = 0, \; \mathscr{H}_{n+1} = \begin{cases} 0 & \text{para } \mathscr{H}_n + \Delta_n \leqslant 0, \\ \mathscr{H}_n + \Delta_n & \text{para } \mathscr{H}_n + \Delta_n > 0. \end{cases} \tag{5.44}$$

Comparemos la sucesión $\mathscr{H}_1, \mathscr{H}_2, \ldots$ con la sucesión $S_1, S_2, \ldots$

Designemos por $\nu_0, \nu_1, \ldots$ los valores sucesivos n, para los cuales $\mathscr{H}_{n+1} = 0$, $n \geqslant 1$. El suceso $\nu_0 = 1$ significa $\{S_1 \leqslant 0\}$, y se ve de la relación (5.44) que $\mathscr{H}_{n+1} = S_n$, para $1 \leqslant n < \nu_0 - 1$ y

$$\{\nu_0 = m\} = \{S_1 > 0, \ldots, S_{m-1} > 0, \; S_m \leqslant 0\}, \quad m > 1. \tag{5.45}$$

Por consiguiente, para $1 \leqslant n \leqslant m$, $m = \nu_0$ el enlace entre $\mathscr{H}_{n+1}$ y S_n se puede expresar formalmente por la siguiente relación:

$$\mathscr{H}_{n+1} = S_n - \min(0, S_1, \ldots, S_n). \tag{5.46}$$

Para $m \leqslant n < \nu_1$, $m = \nu_0$ las magnitudes $\mathscr{H}_{n+1}$ recibieron en cada paso n el mismo incremento Δ_n que las magnitudes S_n: $\mathscr{H}_{n+1} - \mathscr{H}_{m+1} = S_n - S_m > 0$, donde

$$\mathscr{H}_{m+1} = 0, \quad S_m = \min(0, S_1, \ldots, S_n).$$

Del mismo modo que en (5.46), $\mathscr{H}_{n+1} = S_n - \min(0, S_1, \ldots, S_n)$. Para $n = \nu_2$ tenemos $\mathscr{H}_n + \Delta_n = (S_{n-1} - S_m) + \Delta_n = S_n - S_m \leqslant 0$ *tal que la fórmula (5.46) que nos da $\mathscr{H}_{n+1} = 0$*, es también válida como anteriormente. Ahora ya debe estar claro que esta fórmula tiene lugar para todos los $n = 1, 2, \ldots$

Examinemos las sumas sucesivas $S'_1 = \Delta_n$, $S'_2 = \Delta_n + \Delta_{n-1}, \ldots$ $\ldots, S'_n = \Delta_n + \ldots + \Delta_1$ de las mismas magnitudes aleatorias independientes $\Delta_1, \ldots, \Delta_n$ que en (5.43), pero tomadas en orden contrario. Tenemos

$$S'_k = S_n - S_{n-k}, \quad k = 1, \ldots, n, \tag{5.47}$$

y evidentemente,

$$\text{máx}\,(0, S'_1, \ldots, S'_n) = S_n - \text{mín}\,(0, S_1, \ldots, S_n) = \mathscr{H}_n. \tag{5.48}$$

La distribución conjunta $(S'_1, \ldots, S'_n)$ es igual a la distribución $(S_1, \ldots, S_n)$ y de tal modo, ha sido obtenido el resultado siguiente.

Lema 1. *La distribución de probabilidades de la magnitud $\mathscr{H}_n$ es la misma que la de la magnitud $\zeta_n = \text{máx}\,(0, S_1, \ldots, S_n)$.*

Sea a el intervalo medio de tiempo entre las demandas que se reciben sucesivamente ($a = \mathbf{M}\xi_1$), y b el tiempo medio de servicio de una demanda aislada ($b = \mathbf{M}\eta_1$).

El valor medio de las magnitudes $\Delta_n = \eta_n - \xi_n$, $n = 1, 2, \ldots$, es igual a $b - a$. Para la condición

$$\mathbf{M}\Delta_1 = b - a < 0,$$

como fue mostrado en el punto 2 anterior, la magnitud $\zeta = \text{máx}\,(0, S_1, S_2, \ldots)$ es finita con la probabilidad 1.

La sucesión de magnitudes $\zeta_n = \text{máx}\,(0, S_1, \ldots, S_n)$, $n = 1, 2, \ldots$, crece monótonamente y converge hacia la magnitud $\zeta = \text{máx}\,(0, S_1, S_2, \ldots)$. Evidentemente, la desigualdad $\zeta \leqslant x$ es equivalente a que $\zeta_n \leqslant x$ para todos los $n = 1, 2, \ldots$ Además de esto, los sucesos $A_n = \{\zeta_n \leqslant x\}$ forman la sucesión monótona: $A_1 \supseteq A_2 \supseteq \ldots$ y debido a la continuidad de la probabilidad para el suceso $A = \{\zeta \leqslant x\}$ que coinciden con la intersección $\bigcap_n A_n$, tenemos $\mathbf{P}(A) = \lim_{n\to\infty} \mathbf{P}(A_n)$.

Esto junto con el lema 1 da el siguiente resultado.

Teorema 1. *Para la condición*

$$\mathbf{M}\xi_1 = a > b = \mathbf{M}\mu_1 \tag{5.49}$$

la distribución de probabilidades de las magnitudes $\mathscr{H}_n$ converge hacia la distribución del máximo $\zeta = (0, S_1, S_2, \ldots)$: para cualquier x

$$\mathbf{P}\,\{\mathscr{H}_n \leqslant x\} \to F_\zeta(x) \quad \text{para} \quad n \to \infty. \tag{5.50}$$

(Recordemos que la función de distribución $F_\zeta(x)$ del máximo de la sucesión $0, S_1, S_2, \ldots$ de las sumas de las magnitudes aleatorias independientes fue examinada por nosotros en el punto 2 anterior).

Como complemento del teorema 1 señalemos que para la condición

$$\mathbf{M}\Delta_1 = b - a > 0$$

según la ley reforzada de los grandes números $\frac{S_n}{n} \to \mathbf{M}\Delta_1$, $S_n \to \infty$ con la probabilidad 1, y, por consiguiente,

$$\zeta_n = \text{máx}\,(0, S_1, \ldots, S_n) \to \infty. \tag{5.51}$$

Para las magnitudes $\mathscr{H}_n$, que tienen la misma distribución de probabilidades que las magnitudes ζ_n, se deduce de la relación (5.51) que

$$\mathbf{P}\,\{\mathscr{H}_n \geqslant x\} \to 1 \quad \text{para} \quad n \to \infty \tag{5.52}$$

para cualquier x tan grande como se quiera. Hablando a grueso modo, esto significa que las demandas $\mathscr{H}_n n \to \infty$, recibidas lo suficientemente tarde, con la probabilidad próxima a 1, esperarán su servicio un tiempo infinitamente largo.

Ejemplo. Supongamos que el tiempo de servicio de la demanda de turno tiene una distribución exponencial de probabilidades con la densidad

$$p_{\eta_1}(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

y sea $p_{\xi_1}(y)$ la densidad de probabilidad de la magnitud ξ_1 ($p_{\xi_1}(y) = 0$ para $y < 0$). Entonces la suma $\Delta_1 = \eta_1 - \xi_1$ de las magnitudes independientes η_1 y $-\xi_1$ tiene la densidad

$$p(x) = \int_{-\infty}^{0} p_{\eta_1}(x-y)\, p_{-\xi_1}(y)\, dy = \int_{0}^{\infty} p_{\eta_1}(x+y)\, p_{\xi_1}(y)\, dy,$$

que para $x > 0$ se da por la fórmula

$$p(x) = \int_{0}^{\infty} \lambda e^{-\lambda(x+y)} p_{\xi_1}(y)\, dy = A\lambda e^{-\lambda x}, \quad x > 0,$$

donde

$$A = \int_{0}^{\infty} e^{-\lambda y} p_{\xi_1}(y)\, dy = \mathbf{P}\,\{\Delta_1 > 0\}.$$

Como establecimos anteriormente, en este caso la función de distribución $F_\zeta(x)$ se describe por la fórmula (5.37), a saber:

$$F_\zeta(x) = 1 - (1-q)\, e^{-\lambda q x}, \quad x \geqslant 0.$$

Si el tiempo de espera de la demanda siguiente también tiene la ley exponencial de distribución

$$p_{\xi_1}(x) = \begin{cases} \mu e^{-\mu x}, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

entonces la densidad de probabilidad $p(x)$ para $x < 0$ será

$$p(x) = B\mu e^{\mu x},$$

donde $B = \mathbf{P}\{\Delta_1 \leqslant 0\} = 1 - A$, siendo en la densidad indicada $p(x)$

$$A = \frac{\mu}{\lambda+\mu}, \quad B = \frac{\lambda}{\lambda+\mu}.$$

Según la fórmula (5.40), la probabilidad $q = F_\zeta(0)$, para grandes valores de n, es aproximadamente igual a la probabilidad de que la demanda n encuentre libre al sistema, será

$$q = B - A\frac{\mu}{\lambda} = 1 - \frac{\mu}{\lambda}.$$

(Recordemos, que el intervalo medio de tiempo entre los momentos en que se reciben las demandas es $a = \frac{1}{\mu}$, el tiempo medio de servicio de cada demanda aislada es $b = \frac{1}{\lambda}$ y la condición (5.49) en el teorema 1 significa que $\lambda > \mu$.)

Volvamos al esquema general, suponiendo que sólo se satisface la condición (5.49), y examinemos *el período de ocupación* del sistema de servicio, o sea, el intervalo de tiempo desde el momento de recibir la primera demanda hasta el momento, cuando el sistema queda de nuevo libre.

Si ν_0 es el número de todas las demandas, que se sirven en el período de ocupación, entonces la duración de este período será $T = \sum_{k=1}^{\nu_0} \eta_k$. De la relación (5.45) se ve que ν_0 coincide con el número de pasos hasta la primera entrada de la sucesión $S_n = \sum_{k=1}^{n} \Delta_k$, $n = 1, 2, \ldots$ en la zona $x \leqslant 0$. Como fue mostrado en el punto 2 anterior, para la condición (5.49) la magnitud aleatoria ν_0 tiene una esperanza matemática finita $\mathbf{M}\nu_0$ igual a $\frac{1}{q}$ ($q = F_\zeta(0)$, véase la fórmula (5.41)).

Calculemos la esperanza matemática de la magnitud aleatoria

$$T = \sum_{k=1}^{\nu_0} \eta_k.$$

Lema 2. *Tiene lugar la siguiente igualdad:*

$$\mathbf{M}\sum_{k=1}^{\nu_0} \eta_k = \mathbf{M}\eta_1 \mathbf{M}\nu_0 \tag{5.53}$$

se denomina corrientemente *identidad de Wald*).

Demostración. Aquí tratamos sobre la esperanza matemática de la suma de un número aleatorio de magnitudes $\eta_1, \eta_2, \ldots$ igualmente distribuidas. La igualdad (5.53) es evidente para las magnitudes $\eta_1, \eta_2, \ldots$, independientes del número ν_0, ya que para la condición $\nu_0 = n$ $\mathbf{M}\left(\sum_{k=1}^{\nu_0} \eta_k \mid \nu_0 = n\right) = n\mathbf{M}\eta_1$ y según la fórmula de la esperanza matemática total (véase (4.19), cap. I),

$$\mathbf{M}\sum_{k=1}^{\nu_0} \eta_k = \mathbf{M}\eta_1 \sum_{n=1}^{\infty} n\mathbf{P}\{\nu_0 = n\} = \mathbf{M}\eta_1 \cdot \mathbf{M}\nu_0.$$

Utilicemos un método análogo para la deducción de la fórmula (5.53), y en nuestro caso, cuando las magnitudes $\eta_1, \eta_2, \ldots$ y el número ν_0 están ligados con una misma sucesión S_n, $n = 1, 2, \ldots$, y son magnitudes independientes. Tenemos

$$\begin{aligned}\mathbf{M}\left(\sum_{k=1}^{\nu_0} \eta_k\right) &= \mathbf{M}(\eta_1 \mid \nu_0 \leqslant 1)\,\mathbf{P}(\nu_0 \leqslant 1) + \\ &\quad + \mathbf{M}\left(\eta_1 + \sum_{k=2}^{\nu_0} \eta_k \mid \nu_0 > 1\right)\mathbf{P}(\nu_0 > 1) = \\ &= [\mathbf{M}(\eta_1 \mid \nu_0 \leqslant 1)\,\mathbf{P}(\nu_0 \leqslant 1) + \mathbf{M}(\eta_1 \mid \nu_0 > 1)\,\mathbf{P}(\nu_0 > 1)] + \\ &\quad + \mathbf{M}\left(\sum_{k=2}^{\nu_0} \eta_k \mid \nu_0 > 1\right)\mathbf{P}(\nu_0 > 1) = \\ &= \mathbf{M}\eta_1 + \mathbf{M}\left(\sum_{k=2}^{\nu_0} \eta_k \mid \nu_0 > 1\right)\mathbf{P}(\nu_0 > 1).\end{aligned}$$

El suceso $\{\nu_0 > 1\}$ se determina completamente por las magnitudes η_1, ξ_1, y, por consiguiente, la magnitud η_2 no depende de este suceso (recordemos que las magnitudes aleatorias $\eta_1, \eta_2, \ldots$ son independientes entre sí y de las magnitudes $\xi_1, \xi_2, \ldots$), en particular, $\mathbf{M}(\eta_2 \mid \nu_0 > 1) = \mathbf{M}\eta_2$. Por eso, teniendo en cuenta las distribuciones de probabilidades para la condición $\nu_0 > 1$, así como anteriormente, obtenemos

$$\begin{aligned}\mathbf{M}\left(\sum_{k=2}^{\nu_0} \eta_k \mid \nu_0 > 1\right) &= \mathbf{M}(\eta_2 \mid \nu_0 > 1) + \\ &\quad + \mathbf{M}\left(\sum_{k=3}^{\nu_0} \eta_k \mid \nu_0 > 2\right)\mathbf{P}\{\nu_0 > 2 \mid \nu_0 > 1\} = \\ &= \mathbf{M}\eta_2 + \mathbf{M}\left(\sum_{k=3}^{\nu_0} \eta_k \mid \nu_0 > 2\right)\frac{\mathbf{P}(\nu_0 > 2)}{\mathbf{P}(\nu_0 > 1)}\end{aligned}$$

y, por consiguiente,

$$\mathbf{M}\Big(\sum_{k=2}^{\nu_0}\eta_k\,|\,\nu_0>1\Big)\mathbf{P}(\nu_0>1)=$$
$$=\mathbf{M}\eta_2\mathbf{P}(\nu_0>1)+\mathbf{M}\Big(\sum_{k=3}^{\nu_0}\eta_k\,|\,\nu_0>2\Big)\mathbf{P}(\nu_0>2).$$

Utilizando sucesivamente este método para los restantes valores de $n=3,\ 4,\ \ldots$ obtenemos que

$$\mathbf{M}\Big(\sum_{k=n}^{\nu_0}\eta_k\,|\,\nu_0>n-1\Big)=$$
$$=\mathbf{M}\eta_n+\mathbf{M}\Big(\sum_{k=n+1}^{\nu_0}\eta_k\,|\,\nu_0>n\Big)\frac{\mathbf{P}(\nu_0>n)}{\mathbf{P}(\nu_0>n-1)},\quad n=3,\ 4,\ \ldots,$$

y en resumen,

$$\mathbf{M}\Big(\sum_{k=1}^{\nu_0}\eta_k\Big)=\mathbf{M}\eta_1+\mathbf{M}\eta_2\mathbf{P}\{\nu_0>1\}+\ldots+$$
$$+\mathbf{M}\eta_n\mathbf{P}\{\nu_0>n-1\}+\ldots=\mathbf{M}\eta_1\sum_{n=0}^{\infty}\mathbf{P}\{\nu_0>n\}=\mathbf{M}\eta_1\mathbf{M}\nu_0,$$

que es lo que se exigía demostrar.

La identidad de Wald de la siguiente expresión para el valor medio del período T de ocupación:

$$\mathbf{M}T=\frac{b}{q}, \tag{5.54}$$

donde $q=F_\zeta(0)$ para grandes valores de n es, aproximadamente, la probabilidad de que la demanda n encuentre libre al sistema $\left(\frac{1}{q}=\mathbf{M}\nu_0\right)$ y $b=\mathbf{M}\eta_1$ es el tiempo medio de servicio de una sola demanda.

Señalemos que la fórmula (5.54) no sólo da el valor medio del primer período de ocupación (desde el momento de recibir la primera demanda), sino también, en general, de cualquier período de ocupación

$$T_1=\sum_{k=\nu_0+1}^{\nu_1}\eta_k,\qquad T_2=\sum_{k=\nu_1+1}^{\nu_2}\eta_k,\ \ldots,$$

donde $\nu_0,\ \nu_1,\ \ldots$ son los valores sucesivos de n cuando la demanda $n+1$ encuentre libre al sistema.

Examinemos «la duración de turno» $x(t)$ después del tiempo t desde el comienzo del servicio de la primera demanda, que se recibe en un momento τ_1 (más exacto, $x(t)$ es el número de demandas que se recibieron en el sistema y todavía no han sido servidas). Pongamos

$$\mathbf{P}_k(t) = \mathbf{P}\{x(t) = k\}, \quad k = 0, 1, \ldots;$$

$P_0(t)$ es la probabilidad de que el sistema esté libre en el momento de tiempo $\tau_1 + t$, $P_1(t)$ es la probabilidad de que el sistema tiene una sola demanda que se sirve directamente y así sucesivamente.

Antes observamos que las magnitudes $\mathcal{H}_n$, o sea, el tiempo de espera hasta el comienzo del servicio de la demanda n, tienen una distribución límite de probabilidades cuando $n \to \infty$. Mostremos que la distribución límite para $t \to \infty$ tiene «la duración de turno» $x(t)$.

Designemos por

$$\tau_0^* = \tau_1, \quad \tau_1^* = \tau_1 + \sum_{k=1}^{\nu_0} \xi_k, \quad \tau_2^* = \tau_1^* + \sum_{k=\nu_0+1}^{\nu_1} \xi_k, \ldots$$

a los momentos sucesivos de tiempo, en los cuales las demandas recibidas encuentran libre al sistema (aquí $\nu_0, \nu_1, \ldots$ son los mismos valores que antes, véase la pág. 192). Introduzcamos el siguiente proceso de restablecimiento S_n^*, $n = 1, 2, \ldots$:

$$S_0^* = 0, \quad S_1^* = \tau_1^* - \tau_0^*, \quad S_2^* = \tau_2^* - \tau_1^*, \ldots$$

donde $S_n^* = \sum_{k=1}^{n} \xi_k^*$ *son las sumas de las magnitudes positivas independientes igualmente distribuidas* $\xi_k^* = \tau_k^* - \tau_{k-1}^*$, $k = 1, 2, \ldots$ De acuerdo con la identidad de Wald existe la esperanza matemática $\mathbf{M}\xi_1^*$:

$$\mathbf{M}\xi_1^* = \mathbf{M} \sum_{k=1}^{\nu_0} \xi_k = \mathbf{M}\xi_1 \cdot \mathbf{M}\nu_0 = \frac{a}{q} \tag{5.55}$$

(compárese (5.54)).

Como sabemos, para las magnitudes $\xi^*(t) = t - S^*_{\nu^*(t)}$, donde $\nu^*(t)$ es el número de restablecimientos en el tiempo t, tiene lugar el teorema del restablecimiento: *la distribución de probabilidades de la magnitud $\xi^*(t)$ para $t \to \infty$ converge débilmente hacia la distribución de cierta magnitud ξ^*.*

Teorema 2. *Supongamos que las probabilidades $P_k(t)$, $k = 0, 1, \ldots$, son funciones continuas de t*[1]). *Entonces existen los valores límites*

$$P_k = \lim_{t \to \infty} P_k(t), \quad k = 0, 1, \ldots \tag{5.56}$$

D e m o s t r a c i ó n. Evidentemente, después de cada momento τ_n^*, cuando la demanda que se recibe encuentra libre al sistema, el proceso de formación del turno empieza de nuevo, y por eso, la pro-

[1]) Señalemos que la función de los valores discretos $t = 0, 1, 2, \ldots$ siempre es formalmente continua.

babilidad de que el turno sea igual a k durante el tiempo u, después del momento $\tau_0^* + S_n^*$, es la misma que para $n = 0$, es decir, es igual a $P_k(u)$. Por consiguiente, para el valor fijado $S^*_{\nu^*(t)} = s$, la probabilidad de que $x(t) = k$, será

$$\mathbf{P}\{x(t) = k \mid S^*_{\nu^*(t)}\} = P_k(t - S^*_{\nu^*(t)}) = P_k(\xi^*(t))$$

y

$$P_k(t) = \mathbf{M}P_k(\xi^*(t)).$$

Ahora bien, según la suposición, las distribuciones de las magnitudes aleatorias $\xi^*(t)$ para $t \to \infty$ convergen débilmente hacia la distribución de probabilidades de una magnitud aleatoria ξ^* y de acuerdo al teorema 1 del § 5, cap. II,

$$\mathbf{M}u(\xi^*(t)) \to \mathbf{M}u(\xi^*)$$

para cualquier función continua limitada $u(x)$; para $u(x) = P_k(x)$ obtenemos la relación límite (5.55) en la cual $P_k = \mathbf{M}P_k(\xi^*)$.

§ 6. PROCESOS ALEATORIOS EN LOS SISTEMAS LINEALES

1. Algunas observaciones auxiliares. A un determinado sistema físico lo llamemos lineal, si bajo la acción de las perturbaciones externas $y(s)$, $0 \leqslant s \leqslant t$, su estado fásico en el momento de tiempo t (para las condiciones iniciales nulas) es

$$x(t) = \int_0^t w(t, s)\, y(s)\, ds, \tag{6.1}$$

donde $w(t, s)$, $0 \leqslant s \leqslant t$, es una función continua a trozos: $w(t, s)$, $t \geqslant s$, que corrientemente se llama *función de peso*. Esta describe en el momento de tiempo s, el comportamiento del sistema sacado del estado de reposo por «el impulso unitario» $\delta(t - s)$. (Aquí y en adelante $\delta(t)$ significa «la función delta» por todos conocida).

Por ejemplo, un sistema es lineal, si se describe por la diferencial lineal

$$x^{(n)}(t) + a_1 x^{(n-1)}(t) + \ldots + a_n x(t) = y(t), \tag{6.2}$$

donde los coeficientes $a_k = a_k(t)$, $k = 1, \ldots, n$, pueden cambiar en el transcurso del tiempo t. En este caso la función de peso $w(t, s)$, $t \geqslant s$, para cada valor fijado de s es una solución de la ecuación homogénea

$$\frac{d^n}{dt^n} w(t, s) + a_1 \frac{d^{n-1}}{dt^{n-1}} w(t, s) + \ldots + a_n w(t, s) = 0$$

con las condiciones iniciales

$$\frac{d^k}{dt^k} w(t, s)\Big|_{t=s} = 0, \quad k = 0, 1, \ldots, n-2; \quad \frac{d^{n-1}}{dt^{n-1}} w(t, s)\Big|_{t=s} = 1.$$

(Se puede comprobar directamente, que $x(t) = \int\limits_0^t w(t, s)\, y(s)\, ds$ es una solución de la ecuación diferencial (6.2) que satisface a las condiciones iniciales nulas $x^{(k)}(0) = 0$ para $k = 0, 1, \ldots, n-2$ y $x^{(n-1)}(0) = 1$.)

Examinemos el comportamiento del sistema lineal con la función de peso $w(t, s)$ bajo la acción de las alteraciones caóticas que cambian rápidamente.

Consideremos que sobre el sistema actúan los impulsos aleatorios independientes de la forma $\Delta\eta(t)\,\delta(t - t_k)$, que se originan en los momentos de tiempo indicados t_k, donde $\Delta\eta(t_k)$, $k = 1, 2, \ldots$, son magnitudes aleatorias independientes, que tienen esperanzas matemáticas y dispersiones

$$\mathbf{M}\Delta\eta(t_k) = a_k\Delta t_k, \quad \mathbf{D}\,\Delta\eta(t_k) = b_k\Delta t_k \tag{6.3}$$

($\Delta t_k = t_k - t_{k-1}$ significa el intervalo de tiempo hasta la aparición del impulso de turno). Como resultado de tal acción «a la salida» del sistema se tendrá el proceso aleatorio $\xi = \xi(t)$, $t \geqslant t_0$ ($t_0 = 0$) de la forma

$$\xi(t) = \sum_{0 \leqslant t_k \leqslant t} w(t, t_k)\,\Delta\eta(t_k), \tag{6.4}$$

para el cual

$$\mathbf{M}\xi(t) = \sum_{0 \leqslant t_k \leqslant t} w(t, t_k)\,\mathbf{M}\Delta\eta(t_k) = \sum_{0 \leqslant t_k \leqslant t} w(t, t_k)\, a_k\Delta t_k,$$

$$\mathbf{D}\xi(t) = \sum_{0 \leqslant t_k \leqslant t} w^2(t, t_k)\,\mathbf{D}\Delta\eta(t_k) = \sum_{0 \leqslant t_k \leqslant t} w^2(t, t_k)\, b_k\Delta t_k.$$

Imaginémonos ¿qué ocurrirá si los intervalos entre los distintos impulsos son muy pequeños? o más exactamente ¿qué ocurrirá cuando máx $\Delta t_k \to 0$?

Para subrayar el enlace de los coeficientes a_k y b_k en la fórmula (6.3) con el intervalo de tiempo correspondiente (t_{k-1}, t_k), introduzcamos las funciones $a(t)$ y $b(t)$ de la forma $a(t) = a_k$, $b(t) = b_k$ para $t_{k-1} < t \leqslant t_k$, $k = 1, 2, \ldots$. Supongamos que estas funciones tienen límite cuando máx $\Delta t_k \to 0$; para sencillez podemos considerar desde el mismo comienzo que

$$a_k = a(t_k), \quad b_k = b(t_k),$$

donde $a(t)$ y $b(t)$ son unas funciones continuas a trozos. Entonces el estado de nuestro sistema en el límite para máx $\Delta t_k \to 0$, en el momento de tiempo t, será una magnitud aleatoria determinada $\xi(t)$ con la esperanza matemática

$$\mathbf{M}\xi(t) = \lim \sum w(t, t_k)\, a(t_k)\, \Delta t_k = \int_0^t w(t, s)\, a(s)\, ds$$

y con la dispersión

$$\mathbf{D}\xi(t) = \lim \sum w^2(t, t_k)\, b(t_k)\, \Delta t_k = \int_0^t w^2(t, s)\, b(s)\, ds.$$

De las suposiciones que hicimos, se deduce, que las magnitudes $\Delta\eta(t_k)$ son infinitamente pequeñas para máx $\Delta t_k \to 0$, o más exactamente,

$$\mathbf{M}\,[\Delta\eta(t_k)]^2 = b(t_k)\,\Delta t_k + [a(t_k)\,\Delta t_k]^2 \to 0$$

y con ello

$$\sum_{0\leqslant t_k\leqslant t} \mathbf{D}\Delta\eta(t_k) = \sum b(t_k)\,\Delta t_k \to \int_0^t b(s)\, ds.$$

Si suponemos complementariamente que $[\Delta\eta(t_k)]^3$ tienen en determinado sentido un orden infinitesimal más elevado, más exacto, suponemos el cumplimiento de la condición

$$\sum_{0\leqslant t_k\leqslant t} \mathbf{M}\,|\Delta\eta(t_k)|^3 \to 0, \tag{6.5}$$

entonces, de acuerdo al teorema límite central, se puede afirmar que en el límite (para máx $\Delta t_k \to 0$) la magnitud $\xi(t)$ tendrá la *distribución de probabilidades de Gauss*:

$$\mathbf{P}\{x' \leqslant \xi(t) \leqslant x''\} = \frac{1}{\sqrt{2\pi B}} \int_{x'}^{x''} e^{-\frac{(x-A)^2}{2B}}\, dx, \tag{6.6}$$

donde

$$A = \int_0^t w(t, s)\, a(s)\, ds, \quad B = \int_0^t w^2(t, s)\, b(s)\, ds.$$

En efecto, como se ve fácilmente, para la suma de los sumandos independientes $\xi(t) = \sum\limits_{0\leqslant t_k\leqslant t} w(t, t_k)\,\Delta\eta(t_k)$ se cumplirá la condición conocida de Liapunov:

$$\sum_{0\leqslant t_k\leqslant t} \mathbf{M}\,|w(t, t_k)\,\Delta\eta(t_k) - \mathbf{M}\,w(t, t_k)\,\Delta\eta(t_k)|^3 \leqslant$$

$$\leqslant \max_{0\leqslant s\leqslant t} |w(t, s)|^3 \sum \mathbf{M}\,|\Delta\eta(t_k) - a(t_k)\,\Delta t_k|^3 \to 0$$

(recordemos aquí que $\mathbf{M}\xi(t) \to A$, $\mathbf{D}\xi(t) \to B$).

La acción exterior

$$\dot{\eta}(t) = \sum_{0 \leq t_k \leq t} \Delta\eta(t_k)\,\delta(t - t_k) \tag{6.7}$$

se describe cómodamente, recurriendo al proceso casual

$$\eta(t) = \int_0^t \dot{\eta}(s)\,ds = \sum_{0 \leq t_k \leq t} \Delta\eta(t_k), \tag{6.8}$$

que hablando a grueso modo, caracteriza la acción total sobre el sistema durante el tiempo t. Partiendo del proceso aleatorio $\eta(t)$, $t \geqslant 0$, se podría representar la magnitud $\Delta\eta(t_k)$ como

$$\Delta\eta(t_k) = \eta(t_k) - \eta(t_{k-1}), \quad k = 1, 2, \ldots,$$

tomando en calidad de $\eta(t)$, $t \geqslant 0$ el proceso correspondiente con incrementos independientes, es decir, tal proceso cuyos incrementos $\Delta\eta(t_k) = \eta(t_k) - \eta(t_{k-1})$, $k = 1, 2, \ldots$, son magnitudes aleatorias independientes (cualesquiera que fuesen $t_0 \leqslant t_1 \leqslant t_2 \leqslant \ldots$). Esta propiedad la poseen, por ejemplo, el proceso del movimiento browniano y el proceso de Poisson ya conocidos.

Para máx $\Delta t_k \to 0$ tratamos de hecho hablando a grueso modo, de la acción continua sobre el sistema con pequeñas perturbaciones independientes, y en esta situación es natural pasar del modelo «discreto» *(6.4) al modelo «continuo» correspondiente para el cual puede* servir la integral estocástica

$$\xi(t) = \int_0^t w(t, s)\,d\eta(s) \tag{6.9}$$

(más adelante serán dadas su definición y propiedades).

2. Integral estocástica. El proceso aleatorio $\eta = \eta(t)$ se denomina *proceso con incrementos no correlacionados*, si las magnitudes $\eta_1 = \eta(t_1) - \eta(s_1)$ y $\eta_2 = \eta(t_2) - \eta(s_2)$ son no correlacionadas, para cualesquiera $s_1 \leqslant t_1 \leqslant s_2 \leqslant t_2$:

$$\mathbf{M}(\eta_1 - \mathbf{M}\eta_1)(\eta_2 - \mathbf{M}\eta_2) = 0.$$

A este tipo se refiere, naturalmente, cualquier proceso con incrementos independientes (que tienen esperanzas matemáticas y dispersiones finitas) y en particular, el proceso del movimiento browniano y el proceso de Poisson.

En lo sucesivo (véase el § 7) tendremos que examinar los procesos aleatorios c o m p l e j o s.

Señalemos, que para la magnitud compleja η, la *dispersión* se determina como

$$\mathbf{D}\eta = \mathbf{M}\,|\,\eta - \mathbf{M}\eta\,|^2,$$

y las magnitudes aleatorias η_1 y η_2 que toman valores complejos se denominan *no correlacionadas*, si

$$\mathbf{M}\,[\eta_1 - \mathbf{M}\eta_1]\,\overline{[\eta_2 - \mathbf{M}\eta_2]} = 0. \tag{6.10}$$

Correspondientemente a esto, el proceso aleatorio $\eta(t)$ se llama *proceso con incrementos no correlacionados*, si para cualesquiera incrementos $\eta_1 = \eta(t_1) - \eta(s_1)$, $\eta_2 = \eta(t_2) - \eta(s_2)$ donde $s_1 \leqslant t_1 \leqslant \leqslant s_2 \leqslant t_2$, se cumple la condición (6.10).

Examinemos tal proceso aleatorio $\eta(t)$, $c_1 \leqslant t \leqslant c_2$, con incrementos no correlacionados, que

$$\mathbf{M}\,[\eta(t) - \eta(s)] = \int_s^t a(u)\,du, \quad \mathbf{D}\,[\eta(t) - \eta(s)] = \int_s^t b(u)\,du \tag{6.11}$$

(la función no negativa $b(t)$, $c_1 \leqslant t \leqslant c_2$, en la relación (6.11) la llamamos *densidad de estructura* del proceso aleatorio $\eta(t)$).

Examinemos al principio el caso, cuando $a(t) \equiv 0$.

Para cualquier función constante a trozos $\varphi(t)$, $c_1 \leqslant t \leqslant c_2$, que conserva los valores constantes $y_k = \varphi(t)$, para $t_{k-1} < t \leqslant t_k$, (donde los puntos $c_1 = t_0, t_1, \ldots, t_n = c_2$ dividen el segmento $[c_1, c_2]$ en un número finito de intervalos), la *integral estocástica* se determina por la fórmula

$$\int_{c_1}^{c_2} \varphi(t)\,d\eta(t) = \sum_{k=1}^{n} y_k \Delta\eta(t_k), \tag{6.12}$$

poniendo $\Delta\eta(t_k) = \eta(t_k) - \eta(t_{k-1})$, $k = 1, \ldots, n$. Está claro, que la magnitud $\int_{c_1}^{c_2} \varphi(t)\,d\eta(t)$ no depende del método elegido de división en "intervalos de constancia" $(t_{k-1}t_k)$, $k = 1, \ldots, n$.

Evidentemente, para cualesquiera funciones constante a trozos $\varphi(t)$, $\psi(t)$ y de las constantes λ_1, λ_2,

$$\int_{c_1}^{c_2} [\lambda_1\varphi(t) + \lambda_2\psi(t)]\,d\eta(t) = \lambda_1 \int_{c_1}^{c_2} \varphi(t)\,d\eta(t) + \lambda_2 \int_{c_1}^{c_2} \psi(t)\,d\eta(t) \tag{6.13}$$

Se comprueba fácilmente que

$$\mathbf{M} \int_{c_1}^{c_2} \varphi(t)\,d\eta(t) = 0$$

y como se deduce de la condición (6.10)

$$\mathbf{M}\left[\int_{c_1}^{c_2} \varphi(t)\,d\eta(t)\,\overline{\int_{c_1}^{c_2} \psi(t)\,d\eta(t)}\right] = \int_{c_1}^{c_2} \varphi(t)\,\overline{\psi(t)}\,b(t)\,dt \tag{6.14}$$

(se comprueban lo más sencillamente (6.13) y (6.14), eligiendo para φ y ψ la división común $c_1 = t_0, t_1, \ldots, t_n = c_2$).

En particular,

$$\mathbf{M}\left|\int_{c_1}^{c_2} \varphi(t)\, d\eta(t)\right|^2 = \left\|\int_{c_1}^{c_2} \varphi(t)\, d\eta(t)\right\|^2 = \int_{c_1}^{c_2} |\varphi(t)|^2 b(t)\, dt, \qquad (6.15)$$

y para la distancia media cuadrática entre las magnitudes

$$\xi = \int_{c_1}^{c_2} \varphi(t)\, d\eta(t) \quad \text{y} \quad \eta = \int_{c_1}^{c_2} \psi(t)\, d\eta(t)$$

obtenemos la siguiente expresión:

$$\|\xi - \eta\|^2 = \int_{c_1}^{c_2} |\varphi(t) - \psi(t)|^2 b(t)\, dt. \qquad (6.16)$$

Sea $\varphi(t)$, $c_1 \leqslant t \leqslant c_2$, una función que satisface la condición

$$\int_{c_1}^{c_2} |\varphi(t)|^2 b(t)\, dt < \infty, \qquad (6.17)$$

tal que se puede hallar la sucesión de funciones constantes a trozos $\varphi_n(t)$, $n = 1, 2, \ldots$, que convergen hacia $\varphi(t)$ en su media cuadrática:

$$\int_{c_1}^{c_2} |\varphi_n(t) - \varphi(t)|^2 b(t)\, dt \to 0 \text{ para } n \to \infty \qquad (6.18)$$

(siendo también $\int_{c_1}^{c_2} |\varphi_n(t) - \varphi_m(t)|^2 b(t)\, dt \to 0$, cuando $n, \ m \to \infty$).

La sucesión correspondiente de integrales estocásticas

$$\xi_n = \int_{c_1}^{c_2} \varphi_n(t)\, d\eta(t), \quad n = 1, 2, \ldots,$$

satisfacerá la condición

$$\|\xi_n - \xi_m\|^2 = \int_{c_1}^{c_2} |\varphi_n(t) - \varphi_m(t)|^2 b(t)\, dt \to 0 \text{ para } n, m \to \infty,$$

y, por consiguiente, existe la magnitud límite ξ de esta sucesión (véase el teorema 2 del § 4, cap. I):

$$\|\xi_n - \xi\| \to 0 \quad \text{para} \quad n \to \infty.$$

A este límite lo designamos por $\xi = \int_{c_1}^{c_2} \varphi(t)\, d\eta(t)$ y lo llamamos *integral estocástica*[1]) (de la función $\varphi(t)$). Está claro, que la magnitud ξ está determinada unívocamente con la probabilidad 1 y no depende de la elección de la sucesión $\varphi_n(t)$, $n = 1, 2, \ldots$ que converge hacia $\varphi(t)$.

Se ve fácilmete, que para la integral estocástica definida por nosotros se cumplen las relaciones (6.13)—(6.16) indicadas anteriormente, en el caso de las funciones constantes, a trozos $\varphi(t)$ y $\psi(t)$, ya que estas relaciones se conservan en el paso límite utilizado antes, desde las funciones constantes a trozos a las funciones de tipo general que satisfacen la condición (6.17).

Es útil señalar que para las funciones continuas a trozos

$$\int_{c_1}^{c_2} \varphi(t)\, d\eta(t) = \underset{n\to\infty}{\mathrm{l.i.m}} \sum_{k=1}^{n} \varphi(t_{kn})\, \Delta\eta(t_{kn}), \tag{6.19}$$

donde en la parte derecha se toman «las sumas integrales» para una división cada vez más pequeña del intervalo $[c_1, c_2]$ con los puntos t_{kn}, $k = 1, \ldots, n$, lo que corresponde a la definición general de la integral estocástica dada anteriormente, al elegir la sucesión de funciones constantes a trozos $\varphi_n(t)$, $n = 1, 2, \ldots$, del tipo

$$\varphi_n(t) = \varphi(t_{kn}), \quad t_{k-1,n} < t < t_{k,n}; \quad k = 1, \ldots, n.$$

Extendamos la definición de integral estocástica al intervalo infinito $[c_1, c_2]$.

Para cualquier función $\varphi(t)$, $c_1 \leqslant t \leqslant c_2$, que satisface la condición (6.17), eligiendo la sucesión monótona decreciente $c_{1n} \to c_1$ y la sucesión monótona creciente $c_{2n} \to c_2$, $n = 1, 2, \ldots$, para $n, m \to \infty$, tenemos

$$\left\| \int_{c_{1n}}^{c_{2n}} \varphi(t)\, d\eta(t) - \int_{c_{1m}}^{c_{2m}} \varphi(t)\, d\eta(t) \right\|^2 =$$

$$= \int_{{}_1c_n}^{c_{1m}} |\varphi(n)|^2\, b(t)\, dt + \int_{c_{2m}}^{c_{2n}} |\varphi(t)|^2\, b(t)\, dt \to 0$$

[1]) Llamemos la atención sobre la circunstancia de que la «diferencial» $d\eta(t)$ no se puede interpretar en el mismo sentido que para las funciones diferenciables corrientes (para las cuales $d\eta(t) = \eta'(t)\, dt$), ya que el incremento $\Delta\eta(t) = \eta(t + \Delta t) - \eta(t)$ tiene como media el orden $\sqrt{\Delta t}$ (y no Δt como debía ser para la función diferenciable $\eta(t)$; a saber:

$$\mathbf{M}\,|\Delta\eta(t)|^2 \sim b(t)\,\Delta t \quad \text{y} \quad \frac{|\Delta\eta(t)|}{\Delta t} \to \infty \quad \text{para} \quad \Delta t \to 0.$$

(donde para concretizar consideremos $n \geqslant m$), de aquí se deduce que existe el límite

$$\underset{c_{1n} \to c_1,\ c_{2n} \to c_2}{\text{l.i.m}} \int_{c_{1n}}^{c_{2n}} \varphi(t)\, d\eta(t) = \int_{c_1}^{c_2} \varphi(t)\, d\eta(t). \tag{6.20}$$

Extendamos ahora el concepto de integral estocástica al caso en que el proceso con incrementos no correlacionados $\eta(t)$, $c_1 \leqslant t \leqslant c_2$, tiene una esperanza matemática distinta de cero

$$\mathbf{M}\eta(t) = \int_{c_1}^{t} a(u)\, du.$$

Para cualquier función $\varphi(t)$, $c_1 \leqslant t \leqslant c_2$ que satisface a la condición

$$\int_{c_1}^{c_2} |\varphi(t)\, a(t)|\, dt < \infty, \quad \int_{c_1}^{c_2} |\varphi(t)|^2\, b(t)\, dt < \infty \tag{6.21}$$

pongamos que

$$\int_{c_1}^{c_2} \varphi(t)\, d\eta(t) = \int_{c_1}^{c_2} \varphi(t)\, d\eta^0(t) + \int_{c_1}^{c_2} \varphi(t)\, a(t)\, dt, \tag{6.22}$$

donde

$$\eta^0(t) = \eta(t) - \mathbf{M}\eta(t), \quad c_1 \leqslant t \leqslant c_2,$$

es un proceso aleatorio con incrementos no correlacionados, del tipo examinado anteriormente (que tiene una esperanza matemática nula) con la densidad de estructura $b(t)$, $c_1 \leqslant t \leqslant c_2$.

Las fórmulas fundamentales (véase (6.14)) para las integrales definidas anteriormente variarán evidentemente de la siguiente forma:

$$\left.\begin{aligned}
&\mathbf{M}\left[\int_{c_1}^{c_2} \varphi(t)\, d\eta(t)\right] = \int_{c_1}^{c_2} \varphi(t)\, a(t)\, dt,\\
&\mathbf{M}\left[\int_{c_1}^{c_2} \varphi(t)\, d\eta(t) \cdot \overline{\int_{c_1}^{c_2} \psi(t)\, d\eta(t)}\right] =\\
&= \int_{c_1}^{c_2} \varphi(t)\, \overline{\psi(t)}\, b(t)\, dt + \int_{c_1}^{c_2} \varphi(t)\, a(t)\, dt \cdot \overline{\int_{c_1}^{c_2} \psi(t)\, a(t)\, dt};
\end{aligned}\right\} \tag{6.23}$$

en lugar de (6.15) tenemos

$$\mathbf{M}\left|\int_{c_1}^{c_2} \varphi(t)\, d\eta(t)\right|^2 = \int_{c_1}^{c_2} |\varphi(t)|^2\, b(t)\, dt + \left|\int_{c_1}^{c_2} \varphi(t)\, a(t)\, dt\right|^2. \tag{6.24}$$

Lo mismo que antes, si la sucesión de funciones $\varphi_n(t)$, $n=1, 2, \ldots$, es tal que $\int\limits_{c_1}^{c_2} |\varphi_n(t)-\varphi(t)|^2 b(t) \to 0$, y, además de esto

$$\int\limits_{c_1}^{c_2} \varphi_n(t)\, a(t)\, dt \to \int\limits_{c_1}^{c_2} \varphi(t)\, a(t)\, dt,$$

entonces tendremos

$$\int\limits_{c_1}^{c_2} \varphi(t)\, d\eta(t) = \underset{n\to\infty}{\mathrm{l.i.m}} \int\limits_{c_1}^{c_2} \varphi_n(t)\, d\eta(t) \tag{6.25}$$

en el sentido de convergencia de la media cuadrática; en particular, para la función continua a trozos $\varphi(t)$, la integral estocástica $\int\limits_{c_1}^{c_2} \varphi(t)\, d\eta(t)$ es el límite de las «sumas integrales» (6.19).

3. Convergencia hacia el proceso estacionario. Examinemos el sistema lineal con la función de peso

$$w(t, s) = w(t-s), \quad t \geqslant s, \tag{6.26}$$

que sólo depende de la diferencia $(t-s)$ (la función $w(t)$, $t \geqslant 0$, también la llaman *función de peso*). Llamamos *estable* al sistema lineal si son cumplidas las siguientes condiciones:

$$\int\limits_{0}^{\infty} |w(t)|\, dt < \infty, \quad \int\limits_{0}^{\infty} |w(t)|^2\, dt < \infty. \tag{6.27}$$

Supongamos que sobre este sistema actúan las perturbaciones externas $\dot{\eta}(t)$ de tal tipo, que el proceso aleatorio $\xi(t)$ a la salida del sistema lo representamos por la integral estocástica

$$\xi(t) = \int\limits_{0}^{t} w(t-s)\, d\eta(s), \tag{6.28}$$

donde el proceso aleatorio $\eta(t)$ con incrementos no correlacionados (que caracteriza la acción exterior integral, véase (6.8)) es *homogéneo* según el tiempo, en el sentido de que las funciones $a(t)$ y $b(t)$ de las relaciones (6.11) son constantes:

$$a(t) \equiv a, \quad b(t) \equiv b. \tag{6.29}$$

Examinemos el proceso aleatorio

$$\xi^0(t) = \int\limits_{t_0}^{t} w(t-s)\, d\eta(s), \quad t_0 \leqslant t < \infty,$$

que empieza en el momento de tiempo t_0, y el proceso límite

$$\xi^*(t)=\int_{-\infty}^{t} w(t-s)\,d\eta(s)=\lim_{t_0\to-\infty}\int_{t_0}^{t} w(t-s)\,d\eta(s),\quad -\infty<t<\infty \tag{6.30}$$

(que empezó en el pasado infinitamente lejano). Utilizando las fórmulas (6.23), (6.24), obtenemos

$$\|\xi(t)-\xi^*(t)\|^2=\left\|\int_{-\infty}^{0} w(t-s)\,d\eta(s)\right\|^2=$$

$$=b\int_{-\infty}^{0}|w(t-s)|^2\,ds+\left|a\int_{-\infty}^{0} w(t-s)\,ds\right|^2=$$

$$=b\int_{t}^{\infty}|w(s)|^2\,ds+\left|a\int_{t}^{\infty} w(s)\,ds\right|^2;$$

se ve que

$$\|\xi(t)-\xi^*(t)\|\to 0 \quad \text{para} \quad t\to\infty. \tag{6.31}$$

Suponiendo que $w(t)=0$ para $t<0$, el proceso aleatorio $\xi^*(t)$ se puede representar en la forma

$$\xi^*(t)=\int_{-\infty}^{\infty} w(t-s)\,d\eta(s). \tag{6.32}$$

El valor medio de tal proceso es constante:

$$A(t)=\mathbf{M}\xi^*(t)=a\int_{-\infty}^{\infty} w(u)\,du; \tag{6.33}$$

la función de correlación

$$B(t,s)=\mathbf{M}\xi^*(t)\,\xi^*(s)=b\int_{-\infty}^{\infty} w(t-u)\,w(s-u)\,du=$$

$$=b\int_{-\infty}^{\infty} w(t-s-u)\,w(u)\,du=B(t-s) \tag{6.34}$$

depende sólo de la diferencia $t-s$. Tal proceso $\xi^*(t)$ se llama *estacionario* (*en un amplio sentido*).

Nosotros hemos obtenido el siguiente resultado.

Teorema. *El proceso aleatorio $\xi(t)$ a la salida de un sistema estable, representado por la integral estocástica* (6.28), *cuando $t \to \infty$, converge en una media cuadrática hacia el proceso estacionario $\xi^*(t)$ determinado por la fórmula* (6.30).

Como complemento a esto señalamos además lo siguiente.

Supongamos que el proceso aleatorio $\eta(t)$ es un proceso *homogéneo con incrementos independientes*, es decir, las distribuciones de probabilidades de los incrementos $\eta(t) - \eta(s)$ son las mismas para todos los integrales (s, t) de igual longitud $t - s$, siendo las magnitudes $\eta(t_k) - \eta(s_k)$, $k = 1, \ldots, n$, mutuamente independientes, para cualesquiera $s_1 \leqslant t_1 \leqslant \ldots \leqslant s_n \leqslant t_n$.

Entonces, como se comprende fácilmente, el proceso $\xi^*(t) = \int\limits_{-\infty}^{\infty} w(t-s)\,d\eta(s)$ será *estacionario* en el sentido de que la distribución conjunta de probabilidades de los valores $\xi(t_1 + s), \ldots, \xi(t_n + s)$ para cualesquiera valores fijados $t_1, \ldots, t_n$, es la misma para todos los valores de s. En particular, la distribución de probabilidades de los distintos valores $\xi^*(t)$, para todos los t, es la misma que para magnitud

$$\xi^* = \int\limits_{-\infty}^{\infty} w(-u)\,d\eta(u). \tag{6.35}$$

De la convergencia de la media cuadrática $\|\xi(t) - \xi^*(t)\| \to 0$ se deduce (véase el teorema 2 del § 5, cap. I) que las distribuciones de probabilidades de las magnitudes $\xi(t)$ converge débilmente, para $t \to \infty$, hacia la distribución de probabilidades de la magnitud ξ^* determinada por la fórmula (6.35).

4. Procesos de efecto fraccionario. Examinemos el sistema lineal (6.1) al actuar sobre él los impulsos aleatorios $\eta_k\delta(t - \tau_k)$, $k = 1, 2, \ldots$, originados en los momentos aleatorios de tiempo τ_k, $k = 1, 2, \ldots$, Tal acción representa en sí al proceso aleatorio del tipo

$$\dot{\eta}(t) = \sum_{0 \leqslant \tau_k \leqslant t} \eta_k\delta(t - \tau_k), \quad 0 \leqslant t \leqslant \infty; \tag{6.36}$$

el estado del sistema $\xi(t)$ se cambiará durante el tiempo t de la forma siguiente:

$$\xi(t) = \sum_{0 \leqslant \tau_k \leqslant t} \eta_k w(t - \tau_k), \quad 0 \leqslant t \leqslant \infty \tag{6.37}$$

(compárese (6.4), (6.7)). El proceso aleatorio $\xi = \xi(t)$ de tal tipo se llama corrientemente *proceso de efecto fraccionario*.

Supongamos que los impulsos aleatorios $\eta_k\delta(t - \tau_k)$ surgen según la ley de Poisson con el parámetro λ, es decir, en el intervalo de tiempo (s, t) con la probabilidad

$$\frac{[\lambda(t-s)]^n}{n!}e^{-\lambda(t-s)}$$

habrá justamente n impulsos (independientemente del número de impulsos fuera del intervalo dado desde s hasta t). Supongamos que las magnitudes aleatorias η_k, $k = 1, 2, \ldots$, que caracterizan la dirección y la intensidad de los impulsos correspondientes, tienen una misma distribución de probabilidades, siendo estas magnitudes tanto independientes entre sí, como de los momentos de tiempo τ_k, $k = 1, 2, \ldots$

La acción externa del tipo descrito se puede caracterizar, recurriendo al proceso aleatorio

$$\eta(t) = \int_0^t \dot{\eta}(s)\,ds = \sum_{k=1}^{\nu(t)} \eta_k, \quad 0 \leqslant t < \infty, \tag{6.38}$$

donde $\nu(t)$, $0 \leqslant t < \infty$, es el proceso corriente de Poisson y η_1, η_2, ... es la sucesión correspondiente de magnitudes aleatorias independientes con la misma distribución de probabilidades, también independientes de $\nu(t)$, $0 \leqslant t < \infty$. Tal proceso aleatorio $\eta = \eta(t)$ se denomina *proceso complejo de Poisson*.

Mostremos que con la condición de que las magnitudes correspondientes η_k tengan esperanza matemática y el segundo momento finito:

$$a = \mathbf{M}\eta_k, \quad b = \mathbf{M}\eta_k^2$$

se cumplirán las relaciones (6.11) siendo

$$a(t) = a\lambda t, \quad b(t) = b\lambda t, \tag{6.39}$$

(donde λ es el parámetro del proceso correspondiente de Poisson $\nu(t)$).

En efecto, si designamos por $\nu(s, t)$ el número de momentos τ_k en el intervalo $[s, t]$, entonces, para el valor fijado de $\nu(s, t) = n$, la magnitud $\eta = \eta(t) - \eta(s)$ reperesenta en sí la suma de n sumandos independientes η_k con un valor medio a y una dispersión $\sigma^2 = b - a^2$. Por consiguiente,

$$\mathbf{M}(\eta \mid \nu(s, t)) = a\nu(s, t),$$

y según la fórmula de la esperanza matemática total

$$\mathbf{M}\eta = \mathbf{M}\{\mathbf{M}(\eta \mid \nu(s, t))\} = a\cdot\mathbf{M}\nu(s, t) = a\lambda(t-s).$$

Luego,

$$\mathbf{M}(\eta^2 \mid \nu(s, t)) = \sigma^2\nu(s, t) + (a\nu(s, t))^2,$$

$$\mathbf{M}\eta^2 = \mathbf{M}\{\mathbf{M}(\eta^2 \mid \nu(s, t))\} = \sigma^2\lambda(t-s) + a^2\mathbf{M}[\nu(s, t)]^2$$

y

$$\mathbf{D}\eta = \mathbf{M}\eta^2 - (\mathbf{M}\eta)^2 = \sigma^2\lambda(t-s) + a^2[\mathbf{M}\nu^2(s, t) - (\mathbf{M}\nu(s, t))^2] =$$
$$= \sigma^2\lambda(t-s) + a^2\mathbf{D}\nu(s, t) = \sigma^2\lambda(t-s) + a^2\lambda(t-s) =$$
$$= b\lambda(t-s).$$

Examinemos la integral estocástica $\int_{c_1}^{c_2} \varphi(t)\, d\eta(t)$, donde $\eta(t)$ es un proceso complejo de Poisson del tipo descrito anteriormente.

Mostremos, que para cualquier función continua a trozos $\varphi(t)$, $c_1 \leqslant t \leqslant c_2$, con la probabilidad 1

$$\int_{c_1}^{c_2} \varphi(t)\, d\eta(t) = \sum_{c_1 \leqslant \tau_j \leqslant c_2} \varphi(\tau_j)\, \eta_j. \tag{6.40}$$

(Recordemos que la propia integral estocástica $\xi = \int_{c_1}^{c_2} \varphi(t)\, d\eta(t)$ como límite medio cuadrático de una determinada sucesión de magnitudes aleatorias $\xi_n = \int_{c_1}^{c_2} \varphi_n(t)\, d\eta(t)$, $n = 1, 2, \ldots$, está solamente determinada con la probabilidad 1, a saber: cada magnitud que se diferencia de ξ con la probabilidad 0, también será límite de la misma sucesión ξ_n, $n = 1, 2, \ldots$).

Para la función constante a trozos $\varphi(t)$, $c_1 \leqslant t \leqslant c_2$, que conserva constantes los valores $y_k = \varphi(t)$ para $t_k < t < t_{k+1}$, donde los puntos $t_0 = c_1, \ldots, t_n = c_2$ dividen el segmento $[c_1, c_2]$ en un número finito de intervalos, la probabilidad de coincidencia de los momentos aleatorios τ_j con los puntos t_k, $k = 0, \ldots, n-1$, es igual a cero, y con la probabilidad 1

$$\sum_{c_1 \leqslant \tau_j \leqslant c_2} \varphi(\tau_j)\, \eta_j = \sum_{k=0}^{n-1} \Big[\sum_{t_k < \tau_j < t_{k+1}} \varphi(t_j)\, \eta_j \Big] =$$
$$= \sum_{k=0}^{n-1} \Big[y_k \sum_{t_k \leqslant \tau_j \leqslant t_{k+1}} \eta_j \Big] =$$
$$= \sum_{k=0}^{n-1} y_k [\eta(t_{k+1}) - \eta(t_k)] = \int_{c_1}^{c_2} \varphi(t)\, d\eta(t).$$

Para la función continua a trozos $\varphi(t)$, que es el límite de una sucesión u n i f o r m e m e n t e convergente de funciones continuas a trozos $\varphi_n(t)$, $n = 1, 2, \ldots$, la fórmula (6.40) se puede obtener por un paso al límite:

$$\int_{c_1}^{c_2} \varphi_n(t)\, d\eta(t) = \sum_{c_1 \leqslant \tau_j \leqslant c_2} \varphi_n(\tau_j)\, \eta_j \to \int_{c_1}^{c_2} \varphi(t)\, d\eta(t),$$

y suponiendo, por otro lado, que

$$\Delta_n(t) = \varphi_n(t) - \varphi(t) \quad \text{y} \quad \varepsilon_n = \max_{c_1 \leqslant t \leqslant c_2} |\Delta_n(t)|,$$

tenemos

$$\Big\| \sum_{c_1 \leqslant \tau_j \leqslant c_2} \varphi_n(\tau_j)\, \eta_j - \sum_{c_1 \leqslant \tau_j \leqslant c_2} \varphi(\tau_j)\, \eta_j \Big\|^2 = \mathbf{M} \Big| \sum_{c_1 \leqslant \tau_j \leqslant c_2} \Delta_n(\tau_j)\, \eta_j \Big|^2 \leqslant \varepsilon_n^2 \mathbf{M} \Big[\sum_{c_1 \leqslant \tau_j \leqslant c_2} |\eta_j| \Big]^2 \leqslant C\varepsilon_n^2 \to 0,$$

donde

$$C = \mathbf{M} \Big[\sum_{c_1 \leqslant \tau_j \leqslant c_2} |\eta_j| \Big]^2 = [\mathbf{M} |\eta_j|^2 + (\mathbf{M} |\eta_j|)^2]\, \lambda (c_2 - c_1).$$

La fórmula (6.40) que hemos obtenido nos permite representar el proceso de efecto fraccionario de la forma (6.37) por la integral estocástica del tipo (6.8):

$$\xi(t) = \int_0^t w(t, s)\, d\eta(s). \tag{6.41}$$

En particular, para la función de peso $w(t, s) = w(t - s)$, que satisface las condiciones (6.27), al proceso $\xi(t)$ de efecto fraccionario se le puede aplicar el teorema sobre la convergencia hacia el proceso estacionario (véase la pág. 209) y también su complemento, ya que el proceso complejo de Poisson $\eta(t)$, $t \geqslant 0$, en (6.41) es un proceso homogéneo con incrementos independientes.

¿Con qué probabilidad estará el sistema después del tiempo t, en tal o cual estado fásico? o más exactamente: ¿cuál será la distribución de probabilidades de la magnitud aleatoria $\xi(t)$?

Está claro, que esta distribución de probabilidades depende de la distribución de las magnitudes aleatorias η_k, $k = 1, 2, \ldots$, a la entrada del sistema. Sea $g(u)$ la función característica de las magnitudes η_k de igual distribución; hallemos la función característica $f_t(u) = \mathbf{M} e^{iu\xi(t)}$, $-\infty < u < \infty$, de la magnitud $\xi(t)$ que nos interesa.

Designemos por $\nu(t)$ el número de impulsos durante el tiempo t. Recordemos (véase el § 2, cap. II) que para el valor fijado $\nu(t) = n$, los momentos $\tau_1, \ldots, \tau_n$ de aparición de los impulsos están distribuidos en el segmento $[0, t]$ de tal modo como si los puntos $\tau_1, \ldots, \tau_n$ hubiesen sido lanzados casualmente, independientes uno del otro con la distribución uniforme de probabilidades de modo que para el valor fijado $\nu(t) = n$, la magnitud $\xi(t)$ es la suma de n sumandos independientes $\eta_k w(t\tau_k)$, $k = 1, \ldots, n$ igualmente distribuidos (los τ_k están distribuidos uniformemente en el segmento $[0, t]$ y los η_k son independientes de $\nu(t)$ y $\tau_1, \tau_2, \ldots$). Por consiguiente, para el valor fijado $\tau_k = s$, el valor medio de la magnitud $e^{iu\,w(t,\tau_k)\,\eta_k}$ será

$$\mathbf{M}e^{iuw(t,\,s)\,\eta_k} = g(uw(t, s)),$$

y para el valor fijado $\nu(t) = n$

$$\mathbf{M}\{e^{iuw(t,\,\tau_k)\,\eta_k} \mid \nu(t) = n\} =$$
$$= \mathbf{M}\{g[uw(t, \tau_k)] \mid \nu(t) = n\} = \frac{1}{t}\int_0^t g[uw[t, s)]\,ds$$

y

$$\mathbf{M}\{e^{iu\xi(t)} \mid \nu(t) = n\} = \left[\frac{1}{t}\int_0^t g(uw(t, s))\,ds\right]^n.$$

Según la fórmula de la esperanza matemática total

$$\mathbf{M}e^{iu\xi(t)} = \sum_{n=0}^{\infty} \mathbf{M}\{e^{iu\xi(t)} \mid \nu(t) = n\}\,\mathbf{P}\{\nu(t) = n\} =$$
$$= \sum_{n=0}^{\infty}\left[\frac{1}{t}\int_0^t g(uw(t, s))\,ds\right]^n \frac{(\lambda t)^n}{n!}e^{-\lambda t} = e^{-\lambda t}\left[e^{\lambda\int_0^t g(uw(t,\,s))\,ds}\right].$$

Como resultado, *para la función característica* $f_t(u) = \mathbf{M}e^{iu\,\xi(t)}$ *se tiene la siguiente fórmula*:

$$f_t(u) = \exp\left\{\lambda\int_0^t [g(uw(t, s)) - 1]\,ds\right\}, \quad -\infty < u < \infty, \tag{6.42}$$

donde $g(u) = \mathbf{M}e^{iu\eta_k}$ *es la función característica de las magnitudes* η_k, $k = 1, 2, \ldots$, y λ *es el parámetro de la ley de Poisson correspondiente.*

En particular, para $w(t) \equiv 1$, $0 \leqslant t < \infty$, cuando tratamos con *el proceso complejo de Poisson* $\xi(t) = \eta(t)$ (véase (6.38)), la función característica correspondiente $f_t(u)$ será

$$f_t(u) = e^{\lambda t\,[g(u) - 1]}, \quad -\infty < u < \infty \tag{6.43}$$

(para el proceso corriente de Poisson $\eta_k \equiv 1$ y $g(u) = e^{iu}$).

En resumen señalemos, que para el sistema lineal estable con la función de peso $w(t)$, cuando el proceso aleatorio $\xi(t)$ converge hacia el proceso estacionario para $t \to \infty$, (véase (6.31)), la distribución de probabilidades de las magnitudes $\xi(t)$ converge débilmente hacia la distribución de la magnitud ξ^*, determinada por la fórmula (6.35), y por consiguiente, las funciones características de las magnitudes $\xi(t)$, que según (6.42) tienen la forma

$$f_t(u) = \exp\left\{\lambda \int_0^t [g(uw(s)) - 1]\, ds\right\}, \quad -\infty < u < \infty,$$

convergen hacia la función característica de la magnitud ξ^* para $t \to \infty$

$$f(u) = \exp\left\{\lambda \int_0^\infty [g(uw(s)) - 1]\, ds\right\}, \quad -\infty < u < \infty.$$

§ 7. PROCESOS ALEATORIOS ESTACIONARIOS

1. Representación espectral de los procesos estacionarios y transformación de Fourier. Recordemos que el proceso aleatorio $\xi = \xi(t)$, $-\infty < t < \infty$, se llama estacionario, en amplio sentido, si el valor medio $A(t) = \mathbf{M}\xi(t)$ es uno mismo para todos los valores de t, y la función de correlación $B(t, s) = \mathbf{M}\xi(t)\,\xi(s)$, sólo depende de la diferencia $t - s$. En adelante consideraremos que $A(t) = 0$.

Ampliemos el concepto de estacionaridad, extendiéndolo a los procesos aleatorios con valores complejos $\xi(t)$, $\mathbf{M}\xi(t) = 0$ del modo siguiente: el proceso aleatorio $\xi(t)$, $-\infty < t < \infty$, se llama *estacionario en amplio sentido*, si $\mathbf{M}\xi(t)\,\overline{\xi(s)}$ sólo depende de la diferencia $t - s$

$$\mathbf{M}\xi(t)\,\overline{\xi(s)} = B(t - s); \tag{7.1}$$

la función $B(t)$, $-\infty < t < \infty$, se llama *función de correlación* del proceso estacionario $\xi(t)$.

Ejemplo (*proceso estacionario con espectro discreto*). Examinemos el proceso aleatorio de la forma

$$\xi(t) = \sum_{k=-n}^{n} e^{i\lambda_k t}\,\Phi_k, \quad -\infty < t < \infty \tag{7.2}$$

sonde $\lambda_{-n}, \ldots, \lambda_n$ son determinados números reales, y $\Phi_{-n}, \ldots, \Phi_n$ don magnitudes aleatorias n o c o r r e l a c i o n a d a s con valores medios nulos:

$$\mathbf{M}\Phi_k\overline{\Phi}_j = 0 \quad \text{para} \quad k \neq j.$$

Evidentemente,

$$\mathbf{M}\xi(t) = \sum_{k=-n}^{n} e^{i\lambda_k t} \mathbf{M}\Phi_k = 0$$

y

$$\mathbf{M}\xi(t)\overline{\xi(s)} = \sum_{k=-n}^{n} \sum_{j=-n}^{n} e^{i(\lambda_k t - \lambda_j s)} \mathbf{M}\Phi_k \overline{\Phi}_j = \sum_{k=-n}^{n} e^{i\lambda_k (t-s)} F_k, \quad (7.3)$$

donde $F_k = \mathbf{M} \mid \Phi_k \mid^2$, $k = -n, \ldots, n$, y de tal modo, el proceso aleatorio $\xi(t)$ es estacionario.

Este proceso da un ejemplo de oscilaciones aleatorias que tienen como componentes suyas a las oscilaciones armónicas $\xi_k(t) = e^{i\lambda_k t}\Phi_k$ con las frecuencias $\omega_k = \mid \lambda_k \mid$, $k = -n, \ldots, n$, cuya face y amplitud son magnitudes aleatorias. Si convenimos en llamar al valor medio cuadrático $\mathbf{M} \mid \xi(t) \mid^2$ la energía media del proceso estacionario $\xi(t)$, entonces de la igualdad (7.3), para $s = t$, tenemos

$$\mathbf{M} \mid \xi(t) \mid^2 = \sum_{k=-n}^{n} F_k,$$

donde $F_k = \mathbf{M} \mid \Phi_k \mid^2 = \mathbf{M} \mid \xi_k(t) \mid^2$, y se ve que la energía del proceso estacionario $\xi(t)$ se compone de las energías de las componentes separadas $\xi_k(t) = e^{i\lambda_k t}\Phi_k$, $k = -n, \ldots, n$.

Los valores λ_k, $k = -n, \ldots, n$, se llaman corrientemente, *frecuencias* (incluso para los valores negativos de λ_k); en su conjunto forman *el espectro de frecuencias* del proceso estacionario $\xi(t)$.

Ejemplo (proceso estacionario con espectro continuo). Examinemos el proceso aleatorio $\xi(t)$ que admite la llamada *representación espectral*:

$$\xi(t) = \int_{-\infty}^{\infty} e^{\lambda t} d\Phi(\lambda), \quad (7.4)$$

donde $\Phi(\lambda)$, $-\infty < \lambda < \infty$ es el proceso aleatorio con incrementos no correlacionados y con un valor medio nulo tal que para cualquier intervalo (λ_1, λ_2)

$$\mathbf{M} \mid \Phi(\lambda_1) - \Phi(\lambda_2) \mid^2 = \int_{\lambda_1}^{\lambda_2} f(\lambda)\, d\lambda; \quad (7.5)$$

la densidad de estructura $f(\lambda)$ de este proceso se supone que es integrable.

En la fórmula (7.4) figura la integral estocástica de la función compleja $\varphi(\lambda) = e^{i\lambda t}$, $-\infty < \lambda < \infty$.

Recordemos, que la integral estocástica $\int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda)$ definida por nosotros en el § 6, posee las siguientes propiedades:

$$\mathbf{M} \int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda) = 0,$$

$$\mathbf{M} \left| \int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda) \right|^2 = \int_{-\infty}^{\infty} |\varphi(\lambda)|^2 f(\lambda)\, d\lambda, \tag{7.6}$$

y en general,

$$\mathbf{M} \left[\int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda) \right] \overline{\left[\int_{-\infty}^{\infty} \psi(\lambda)\, d\Phi(\lambda) \right]} = \int_{-\infty}^{\infty} \varphi(\lambda) \overline{\psi(\lambda)} f(\lambda)\, d\lambda, \tag{7.7}$$

donde la función $f(\lambda) \geqslant 0$ es la densidad de estructura del proceso $\Phi(\lambda)$ que figura en la relación (7.5).

Así pues, examinemos el proceso aleatorio $\xi(t)$ determinado por la igualdad (7.4). Para $\varphi(\lambda) = e^{i\lambda t}$, $\psi(\lambda) = e^{i\lambda s}$ obtenemos de las fórmulas (7.6) y (7.7) que

$$\mathbf{M}\xi(t) = 0$$

y

$$\mathbf{M}\xi(t)\overline{\xi(s)} = \int_{-\infty}^{\infty} e^{i\lambda(t-s)} f(\lambda)\, d\lambda.$$

Se ve que $\xi(t)$ es **un proceso estacionario** con la función de correlación

$$B(t) = \int_{-\infty}^{\infty} e^{i\lambda t} f(\lambda)\, d\lambda, \tag{7.8}$$

que coincide con la transformación de Fourier de la densidad de estructura $f(\lambda)$ del proceso $\Phi(\lambda)$; $f(\lambda)$ también se llama *densidad espectral* del proceso estacionario $\xi(t)$. A su vez, la densidad espectral $f(\lambda)$ se determina unívocamente, por su transformación de Fourier $B(t)$, y, en particular, cuando la función de correlación $B(t)$, $-\infty < t < \infty$, es integrable

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda t} B(t)\, dt. \tag{7.9}$$

La fórmula (7.4) da la descomposición del proceso estacionario $\xi(t)$ en oscilaciones armónicas elementales, más exacto,

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\, d\Phi(\lambda) \sim \sum_{k=-n}^{n} e^{i\lambda_k t}\, \Phi_k,$$

donde $\Phi_k = \Phi(\lambda_{k+1}) - \Phi(\lambda_k)$, $k = -n, \ldots, n$, y el signo de equivalencia $\sim$ significa, que para la división cada vez más pequeña, de las bandas de frecuencias $-\Lambda_n \leqslant \lambda \leqslant \Lambda_n$ en intervalos $\Delta_k = (\lambda_k, \lambda_{k+1})$ $k = -n, \ldots, n$,

$$\xi(t) = \underset{n\to\infty}{\text{l.i.m.}} \sum_{k=-n}^{n} e^{i\lambda_k t}\, \Phi_k \tag{7.10}$$

(para $\max\limits_k |\lambda_{k+1} - \lambda_k| \to 0$, $\Lambda_n \to \infty$).

Señalemos que para cualquier intervalo de frecuencias $\lambda_1 \leqslant \lambda \leqslant \lambda_2$, el proceso aleatorio estacionario

$$\xi_\Delta(t) = \int_{\lambda_1}^{\lambda_2} e^{i\lambda t}\, d\Phi(\lambda), \tag{7.11}$$

en un pequeño intervalo $\Delta = (\lambda_1, \lambda_2)$ representa en sí aproximadamente, una oscilación armónica de frecuencia λ, $\lambda_1 \leqslant \lambda \leqslant \lambda_2$ (tanto más exacto en el intervalo de tiempo fijado $t_0 \leqslant t \leqslant t_1$, cuanto menor sea Δ) y su energía media es

$$\mathbf{M}|\xi_\Delta(t)|^2 = \int_{\lambda_1}^{\lambda_2} f(\lambda)\, d\lambda. \tag{7.12}$$

La energía total del proceso estacionario $\xi(t)$ es

$$\mathbf{M}|\xi(t)|^2 = \int_{-\infty}^{\infty} f(\lambda)\, d\lambda.$$

y de este modo, como indica la fórmula (7.12), la densidad espectral $f(\lambda)$ caracteriza la distribución de la energía del proceso examinado $\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\, d\Phi(\lambda)$, según las componentes de la forma $\xi_\Delta(t) = \int_{\lambda_1}^{\lambda_2} e^{i\lambda t} d\Phi(\lambda)$ en dependencia del intervalo $\Delta = (\lambda_1, \lambda_2)$ de aquellas frecuencias λ, que entran en el espectro de frecuencias de las oscilaciones de la componente correspondiente $\xi_\Delta(t)$.

Examinemos el proceso aleatorio estacionario de la forma

$$\xi(t) = \int_{\infty}^{t} w(t-s)\, d\eta(s), \tag{7.13}$$

que se establece en el transcurso del tiempo a la salida del sistema lineal estable con la función de peso $w(t)$, al actuar sobre ella las perturbaciones aleatorias h o m o g é n e a s $\dot{\eta}(t)$ (véase el teorema del § 6 anterior).

Recordemos que hemos llegado al proceso estacionario del tipo (7.13) con ayuda del paso límite:

$$\xi(t) = \underset{t_0 \to -\infty}{\text{l.i.m.}} \int_{t_0}^{t} w(t-s)\, d\eta(s),$$

desplazando el momento inicial t_0 del funcionamiento del sistema examinado, al pasado infinitamente lejano. Para cualesquiera valores de $t_0 \leqslant s \leqslant t$, los incrementos

$$\eta(t) - \eta(s) = \int_{s}^{t} \dot{\eta}(u)\, du,$$

que caracterizan la acción total exterior sobre el sistema en el intervalo de tiempo desde s hasta t, no depende de la elección del origen de lectura t_0. A pesar de que los propios valores $\eta(t)$ pierden su sentido, para $t_0 \to -\infty$ de hecho tratamos sólo con las magnitudes $\Delta\eta = \eta(t) - \eta(s)$, determinadas para cualquier intervalo $\Delta = (s, t]$, que poseen la misma propiedad que para la división del intervalo $\Delta = (s, t]$ en intervalos que no se cortan $\Delta_k = (s_k, t_k]$, $k = 1, \ldots, n$ $(\Delta = \bigcup_{k=1}^{n} \Delta_k)$, tiene lugar la igualdad

$$\Delta\eta = \sum_{k=1}^{n} \Delta_k \eta.$$

Para mayor claridad, hablaremos sobre el proceso aleatorio h o m o g é n e o $\eta(t)$ con incrementos no correlacionados $\Delta\eta = \eta(t) - \eta(s)$, y en el caso $t_0 = -\infty$, suponiendo que

$$\left.\begin{aligned} \mathbf{M}\Delta\eta = 0, \quad \mathbf{M}\,|\Delta\eta|^2 = \sigma^2 (t-s), \\ \mathbf{M}(\Delta_1\eta \cdot \overline{\Delta_2\eta}) = 0 \end{aligned}\right\} \tag{7.14}$$

para cualesquiera intervalos Δ_1, Δ_2 que no se cortan; la homogeneidad del proceso $\eta(t)$, mencionada anteriormente, significa que su densidad de estructura es igual a una constante (en nuestro caso esta constante se ha designado por σ^2).

Examinemos ahora el proceso aleatorio $\psi(\lambda)$ determinado para todos los $\lambda \geqslant \lambda_0$ como

$$\psi(\lambda_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\lambda t} - e^{-i\lambda_0 t}}{-it}\, d\eta(t).$$

Para cualesquiera $\lambda_0 \leqslant \lambda_1 \leqslant \lambda_2$ los incrementos $\Delta\psi = \psi(\lambda_2) - \psi(\lambda_1)$ en el intervalo $\Delta = (\lambda_1, \lambda_2)$ representados en la forma

$$\Delta\psi = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\lambda_2 t} - e^{-i\lambda_1 t}}{-it}\, d\eta(t). \tag{7.15}$$

poseen las siguientes propiedades:

$$\left.\begin{aligned} &\mathbf{M}\Delta\psi = 0, \quad \mathbf{M}\,|\Delta\psi|^2 = \frac{\sigma^2}{2\pi}(\lambda_2 - \lambda_2). \\ &\mathbf{M}\,(\Delta_1\psi \cdot \overline{\Delta_2\psi}) = 0 \end{aligned}\right\} \tag{7.16}$$

para los intervalos Δ_1, Δ_2 que no se cortan. Las dos últimas relaciones que necesitan su demostración se obtienen fácilmente al utilizar la igualdad de Parceval; precisamente, la función subintegral

$$e_\Delta(t) = \frac{1}{2\pi} \frac{e^{-i\lambda_2 t} - e^{-i\lambda_1 t}}{-it}, \qquad \int_{-\infty}^{\infty} |e_\Delta(t)|^2\, dt < \infty,$$

es la transformación de Fourier de la forma

$$e_\Delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda t} \chi_\Delta(\lambda)\, d\lambda, \tag{7.17}$$

donde $\chi_\Delta(\lambda)$ es el indicador del intervalo $\Delta = (\lambda_1, \lambda_2]$ ($\chi_\Delta(\lambda) = 1$ para $\lambda \in \Delta$ y $\chi_\Delta(\lambda) = 0$ para $\lambda \notin \Delta$) y

$$\mathbf{M}\,|\Delta\psi|^2 = \sigma^2 \int_{-\infty}^{\infty} |e_\Delta(t)|^2\, dt = \frac{\sigma^2}{2\pi} \int_{-\infty}^{\infty} |\chi_\Delta(\lambda)|^2\, d\lambda = \frac{\sigma^2}{2\pi}(\lambda_2 - \lambda_1).$$

$$\mathbf{M}\,(\Delta_1\psi \cdot \overline{\Delta_2\psi}) = \sigma^2 \int_{-\infty}^{\infty} e_{\Delta_1}(t)\, \overline{e_{\Delta_2}(t)}\, dt = \frac{\sigma^2}{2\pi} \int_{-\infty}^{\infty} \chi_{\Delta_1}(\lambda)\, \overline{\chi_{\Delta_2}(\lambda)}\, d\lambda = 0$$

para los cualesquiera intervalos Δ_1, Δ_2 que no se cortan.

La fórmula (7.15) indica que los incrementos $\Delta\psi = \psi(\lambda_2) - \psi(\lambda_1)$ son los mismos para cualquier valor inicial λ_0. Por eso, para las funciones $\varphi(\lambda)$ que satisfagan a la condición

$$\int_{-\infty}^{\infty} |\varphi(\lambda)|^2\, d\lambda < \infty,$$

está determinada la integral estocástica

$$\int_{-\infty}^{\infty} \varphi(\lambda)\, d\psi(\lambda) = \underset{\lambda_0 \to -\infty}{\mathrm{l.i.m.}} \int_{\lambda_0}^{\infty} \varphi(\lambda)\, d\psi(\lambda).$$

En el mismo sentido que sobre el proceso $\eta(t)$, $-\infty < t < \infty$, hablaremos del proceso aleatorio $\psi(\lambda)$, $-\infty < \lambda < \infty$, con incrementos no correlacionados, que tiene un valor medio nulo y una densidad de estructura constante $g(\lambda) = \sigma^2/2\pi$, aunque tienen sólo sentido real los «incrementos» $\Delta\psi = \psi(\lambda_2) - \psi(\lambda_1)$ determinados por la fórmula (7.15) para cualquier intervalo $\Delta = (\lambda_1, \lambda_2]$ (a este proceso aleatorio generalizado $\psi(\lambda)$ lo llaman corrientemente, *transformación de Fourier* del proceso aleatorio $\eta(t)$.

Consideremos la integral estocástica $\int\limits_{-\infty}^{\infty} \varphi(\lambda)\, d\Psi(\lambda)$.

Para la función constante a trozos de la forma $\varphi(\lambda) = \sum\limits_{k=1}^{n} y_k \chi_{\Delta_k}(\lambda)$, donde $\Delta_k = (\lambda_k', \lambda_k'')$, $k = 1, \ldots, n$ son los intervalos que no se cortan de longitud finita y su transformación de Fourier

$$c(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-i\lambda t} \varphi(\lambda)\, d\lambda \tag{7.18}$$

tenemos

$$\int\limits_{-\infty}^{\infty} \varphi(\lambda)\, d\Psi(\lambda) = \int\limits_{-\infty}^{\infty} c(t)\, d\eta(t), \tag{7.19}$$

lo que es un corolario simple de las fórmulas (7.15), (7.17), a saber:

$$\int\limits_{-\infty}^{\infty} \varphi(\lambda)\, d\Psi(\lambda) = \sum_{k=1}^{n} y_k \Delta_k \Psi = \int\limits_{-\infty}^{\infty} \left[\sum_{k=1}^{n} y_k e_{\Delta_k}(t) \right] d\eta(t) = \int\limits_{-\infty}^{\infty} c(t)\, d\eta(t).$$

La distancia media cuadrática entre las integrales estocásticas del tipo $\int\limits_{-\infty}^{\infty} \varphi(\lambda)\, d\Psi(\lambda)$ es

$$\mathbf{M} \left| \int\limits_{-\infty}^{\infty} \varphi_1(\lambda)\, d\Psi(\lambda) - \int\limits_{-\infty}^{\infty} \varphi_2(\lambda)\, d\Psi(\lambda) \right|^2 = \frac{\sigma^2}{2\pi} \int\limits_{-\infty}^{\infty} |\varphi_1(\lambda) - \varphi_2(\lambda)|^2\, d\lambda.$$

Análogamente

$$\mathbf{M} \left| \int\limits_{-\infty}^{\infty} c_1(t)\, d\eta(t) - \int\limits_{-\infty}^{\infty} c_2(t)\, d\eta(t) \right|^2 = \sigma^2 \int\limits_{-\infty}^{\infty} |c_1(t) - c_2(t)|^2\, dt.$$

Recordemos que para las funciones $c(t)$, $\int_{-\infty}^{\infty} |c(t)|^2 dt < \infty$, que son transformaciones de Fourier de la forma (7.18) de las funciones $\varphi(\lambda)$ integrables absolutamente, según la igualdad de Parceval

$$\int_{-\infty}^{\infty} |c(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\varphi(\lambda)|^2 d\lambda;$$

por eso

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |\varphi_1(\lambda) - \varphi_2(\lambda)|^2 d\lambda = \int_{-\infty}^{\infty} |c_1(t) - c_2(t)|^2 dt$$

y

$$\mathbf{M} \left| \int_{-\infty}^{\infty} \varphi_1(\lambda) d\Psi(\lambda) - \int_{-\infty}^{\infty} \varphi_2(\lambda) d\Psi(\lambda) \right|^2 = \\ = \mathbf{M} \left| \int_{-\infty}^{\infty} c_1(t) d\eta(t) - \int_{-\infty}^{\infty} c_2(t) d\eta(t) \right|^2,$$

donde $c_1(t)$ y $c_2(t)$ son transformaciones de Fourier del tipo (7.18) de las funciones $\varphi_1(\lambda)$ y $\varphi_2(\lambda)$.

Está claro, que si la función $\varphi(\lambda)$ es el límite medio cuadrático de la sucesión de funciones constantes a trozos $\varphi_n(\lambda)$, $n = 1, 2, \ldots$, entonces su transformación de Fourier $c(t)$ coincide con el límite medio cuadrático de la sucesión correspondiente $c_n(t)$, $n = 1, 2, \ldots$, siendo

$$\int_{-\infty}^{\infty} \varphi(\lambda) d\Psi(\lambda) = \underset{n\to\infty}{\text{l.i.m.}} \int_{-\infty}^{\infty} \varphi_n(\lambda) d\Psi(\lambda),$$

$$\int_{-\infty}^{\infty} c(t) d\eta(t) = \underset{n\to\infty}{\text{l.i.m.}} \int_{-\infty}^{\infty} c_n(t) d\eta(t).$$

Sin embargo, para las funciones constantes a trozos $\varphi_n(\lambda)$ tiene lugar la igualdad (7.19)

$$\int_{-\infty}^{\infty} \varphi_n(\lambda) d\Psi(\lambda) = \int_{-\infty}^{\infty} c_n(t) d\eta(t), \quad n = 1, 2, \ldots,$$

y, por consiguiente,

$$\int_{-\infty}^{\infty} \varphi(\lambda) d\Psi(\lambda) = \underset{n\to\infty}{\text{l.i.m.}} \int_{-\infty}^{\infty} \varphi_n(\lambda) d\Psi(\lambda) = \\ = \underset{n\to\infty}{\text{l.i.m.}} \int_{-\infty}^{\infty} c_n(t) d\eta(t) = \int_{-\infty}^{\infty} c(t) d\eta(t).$$

De tal modo, la igualdad (7.19) tiene lugar para cualesquiera pares $\varphi(\lambda)$ y $c(t)$ que satisfagan las condiciones

$$\int_{-\infty}^{\infty} |\varphi(\lambda)|^2 d\lambda < \infty, \quad \int_{-\infty}^{\infty} |c(t)|^2 dt < \infty$$

(donde $c(t)$ es la transformación de Fourier de la forma (7.18) de la función $\varphi(\lambda)$).

Partiendo de aquellas funciones $\varphi(\lambda)$ y $c(t)$ que satisfacen las condiciones complementarias

$$\int_{-\infty}^{\infty} |\varphi(\lambda)|^2 d\lambda < \infty, \quad \int_{-\infty}^{\infty} |c(t)|\, dt < \infty \tag{7.20}$$

(en este caso la transformación inversa de Fourier nos da $\varphi(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda t} c(t)\, dt$), con ayuda del paso límite, la igualdad (7.19) se puede extender a *cualesquiera funciones* $c(t)$, *absolutamente integrables* y a sus transformaciones de Fourier

$$\varphi(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda t} c(t)\, dt.$$

Consideremos ahora directamente al proceso estacionario $\xi(t)$, representado por la integral estocástica (7.13):

$$\xi(t) = \int_{-\infty}^{\infty} w(t-s)\, d\eta(s).$$

Llamemos *característica espectral*[1]) del sistema lineal con la función de peso $w(t)$, ($w(t)$, $-\infty < t < \infty$ es absolutamente inte-

[1]) En las aplicaciones el sistema lineal se caracteriza, corrientemente, por *la tal llamada función de transmisión* $\varphi(p)$ que está ligada a la función de peso $w(t)$ correspondiente ($w(t) = 0$ para $t < 0$), con la transformación de Laplace:

$$\varphi(p) = \int_{0}^{\infty} e^{-pt} w(t)\, dt$$

(*véase, por ejemplo, el libro* de J. Lanning y R. Battin «Procesos aleatorios en los sistemas de dirección automática» (trad. al ruso del inglés), IL, Moscú, 1958.

grable), a la función

$$\varphi(i\lambda)=\int_{-\infty}^{\infty} e^{-i\lambda t} w(t)\,dt. \qquad (7.21)$$

Supongamos que la característica espectral $\varphi(i\lambda)$ es tal que $\int_{-\infty}^{\infty} |\varphi(i\lambda)|^2\, d\lambda<\infty$. Evidentemente, la función $c(s)=w(t-s)$, $-\infty<s<\infty$, satisface a las condiciones (7.20) siendo

$$\int_{-\infty}^{\infty} e^{-i\lambda s} c(s)\,ds=e^{i\lambda t}\int_{-\infty}^{\infty} e^{-i\lambda(t-s)} w(t-s)\,ds=e^{i\lambda t}\varphi(i\lambda).$$

Por consiguiente, de acuerdo con la fórmula general (7.19), tenemos

$$\xi(t)=\int_{-\infty}^{\infty} w(t-s)\,d\eta(s)=\int_{-\infty}^{\infty} e^{i\lambda t}\varphi(i\lambda)\,d\Psi(\lambda). \qquad (7.22)$$

La igualdad

$$\Phi(\lambda)=\int_{-\infty}^{\lambda} \varphi(i\lambda)\,d\Psi(\lambda) \qquad (7.23)$$

determina el proceso aleatorio complejo con incrementos no correlacionados, que tiene el valor medio nulo y la densidad de estructura

$$f(\lambda)=\frac{\sigma^2}{2\pi}|\varphi(i\lambda)|^2. \qquad (7.24)$$

En efecto, designando por $\Delta\Phi=\Phi(\lambda_2)-\Phi(\lambda_1)$ el incremento en el intervalo correspondiente $\Delta=(\lambda_1, \lambda_2)$ tenemos

$$\mathbf{M}\Delta\Phi=0,$$

$$\mathbf{M}|\Delta\Phi|^2=\mathbf{M}\left|\int_{-\infty}^{\infty} \varphi(i\lambda)\,\chi_\Delta(\lambda)\,d\Psi(\lambda)\right|^2=\frac{\sigma^2}{2\pi}\int_{\lambda_1}^{\lambda_2} |\varphi(i\lambda)|^2\,d\lambda,$$

$$\mathbf{M}(\Delta_1\Phi\overline{\Delta_2\Phi})=\mathbf{M}\left[\int_{-\infty}^{\infty} \varphi(i\lambda)\,\varkappa_{\Delta_1}(\lambda)\,d\Psi(\lambda)\,\overline{\int_{-\infty}^{\infty} \varphi(i\lambda)\,\varkappa_{\Delta_2}\,d\Psi(\lambda)}\right]=$$

$$=\frac{\sigma^2}{2\pi}\int_{-\infty}^{\infty} |\varphi(i\lambda)|^2\,\chi_{\Delta_1}(\lambda)\,\overline{\chi_{\Delta_2}(\lambda)}\,d\lambda=0$$

para cualesquiera intervalos que no se cortan Δ_1, Δ_2.

De la definición general de integral estocástica se deduce fácilmente que

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\varphi(i\lambda)\, d\Psi(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda t}\, d\Phi(\lambda). \tag{7.25}$$

Hemos obtenido la representación espectral del proceso estacionario $\xi(t)$ (véase (7.4) y más adelante).

Así pues, tiene lugar el siguiente resultado.

Teorema. *Cualquier proceso estacionario aleatorio* $\xi(t)$ *del tipo* (7.13) *admite la representación espectral* (7.4); *con ello su densidad espectral* $f(\lambda)$ *se expresa por la fórmula* (7.24).

2. Transformaciones lineales. Ejemplos. Sea $\xi(t)$, $-\infty < t < \infty$, un proceso estacionario aleatorio, que admite la representación espectral

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\, d\Phi(\lambda).$$

Sea $T = [a, b]$ un segmento finito o infinito sobre el eje del tiempo $-\infty < t < \infty$. Designemos por $L_T(f)$ al conjunto de todas las funciones $\varphi(\lambda)$, $-\infty < \lambda < \infty$, que tienen bien la forma $\sum_{k=1}^{n} c_k e^{i\lambda t_k}$ (donde $t_1, \ldots, t_n \in T$), o bien son límites para tales funciones:

$$\varphi(\lambda) = \text{l.i.m.} \sum_k c_k e^{i\lambda t_k} \tag{7.26}$$

en el sentido de *la convergencia media cuadrática* con relación a la distancia

$$\| \varphi_1(\lambda) - \varphi_2(\lambda) \| = \left(\int_{-\infty}^{\infty} | \varphi_1(\lambda) - \varphi_2(\lambda) |^2 f(\lambda)\, d\lambda \right)^{1/2},$$

donde $f(\lambda)$ es la densidad espectral del proceso estacionario examinado $\xi(t)$.

Designemos por $H(T)$ al conjunto de todas las magnitudes de la forma

$$\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda), \tag{7.27}$$

donde $\varphi(\lambda) \in L_T(f)$. Para $\varphi(\lambda) = \sum_k c_k e^{i\lambda t_k}$ las magnitudes indicadas representan en sí las combinaciones lineales de los valores correspondientes $\xi(t_k)$, $t_k \in T$:

$$\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda) = \int_{-\infty}^{\infty} \left[\sum_k c_k e^{i\lambda t_k}\right] d\Phi(\lambda) =$$

$$= \sum_k c_k \int_{-\infty}^{\infty} e^{i\lambda t_k}\, d\Phi(\lambda) = \sum_k c_k \xi(t_k)$$

y en caso general son límites (con relación a la convergencia media cuadrática) para tales combinaciones lineales:

$$\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda) = \text{l.i.m.} \sum_k c_k \xi(t_k). \tag{7.28}$$

Ejemplo (*integración*). Sea $c(t)$, $-\infty < t < \infty$, una función continua a trozos, igual a cero fuera del segmento finito $T = [a, b]$ y sea $\varphi(\lambda)$ su transformación de Fourier:

$$\varphi(\lambda) = \int_a^b e^{i\lambda t} c(t)\, dt = \lim \sum_{k=0}^{n-1} e^{i\lambda t_k} c(t_k) \Delta t_k.$$

donde $\Delta t_k = t_{k+1} - t_k$, $k = 0, 1, \ldots, n-1$, y el límite «de las sumas integrales» $\varphi_n(\lambda) = \sum_{k=0}^{n-1} e^{i\lambda t_k} c(t_k)\Delta t_k$ se toma, como corrientemente, para la división cada vez más pequeña del segmento $[a, b]$ por los puntos $t_0 = a, t_1, \ldots, t_n = b$ (cuando $\max_k \Delta t_k \to 0$). Se ve fácilmente, que la relación límite indicada también se cumple en el sentido de la convergencia media cuadrática:

$$\text{l.i.m.} \int_{-\infty}^{\infty} |\varphi(\lambda) - \varphi_n(\lambda)|^2 f(\lambda)\, d\lambda = 0. \tag{7.29}$$

En efecto, para cualquiera $\varepsilon > 0$ se puede elegir tal a que $\int_{|\lambda| > a} f(\lambda)\, d\lambda \leqslant \varepsilon$ (recordemos, que la densidad espectral $f(\lambda)$ es una función integrable); la sucesión $\varphi_n(\lambda) = \sum_{k=1}^{n} e^{i\lambda t_k} c(t_k)\Delta t_k$ está limitada y converge hacia $\varphi(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda t} c(t)\, dt$ uniformemente por λ en cada intervalo finito:

$$|\varphi(\lambda) - \varphi_n(\lambda)| \leqslant \varepsilon \quad \text{para} \quad |\lambda| \leqslant a;$$

para todos los valores de n suficientemente grandes (digamos, para $n \geqslant n_\varepsilon$) tenemos

$$\int_{-\infty}^{\infty} |\varphi(\lambda)-\varphi_n(\lambda)|^2 f(\lambda)\,d\lambda = \int_{|\lambda|>a} |\varphi(\lambda)-\varphi_n(\lambda)|^2 f(\lambda) +$$

$$+ \int_{-a}^{a} |\varphi(\lambda)-\varphi_n(\lambda)|^2 f(\lambda)\,d\lambda \leqslant C \int_{|\lambda|>a} f(\lambda)\,d\lambda + \varepsilon \int_{-a}^{a} f(\lambda)\,d\lambda \leqslant C_1\varepsilon,$$

cuando $n \geqslant n_\varepsilon$ para cualquier valor de $\varepsilon > 0$ dado de antemano (donde C, C_1 son determinados constantes).

Examinemos la magnitud correspondiente $\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\,d\Phi(\lambda)$ del espacio $H(T)$:

$$\eta = \text{l.i.m.} \int_{-\infty}^{\infty} \varphi_n(\lambda)\,d\Phi(\lambda) = \text{l.i.m.} \sum_{k=1}^{n} c(t_k)\,\xi(t_k)\,\Delta_n t_k.$$

Llamemos *integral* al límite medio cuadrático de "las sumas integrales" $\sum_{k=1}^{n} c(t_k)\,\xi(t_k)\,\Delta t_k$ suponiendo[1]) que

$$\int_{-\infty}^{\infty} c(t)\,\xi(t)\,dt = \text{l.i.m.} \sum_{k=1}^{n} c(t_k)\,\xi(t_k)\,\Delta t_k.$$

En estas designaciones

$$\int_{-\infty}^{\infty} \varphi(\lambda)\,d\Phi(\lambda) = \int_{-\infty}^{\infty} c(t)\,\xi(t)\,dt. \tag{7.30}$$

Por un paso límite la fórmula (7.30) se extiende a las funciones arbitrarias absolutamente integrables $c(t)$, $-\infty < t < \infty$, y a sus transformaciones de Fourier $\varphi(\lambda)$, $-\infty < \lambda < \infty$ (véase (7.19)).

[1]) Señalemos, que para el proceso aleatorio $\eta(t)$, $a \leqslant t \leqslant b$, en el cual todas trayectorias posibles son absolutamente integrables (para nosotros $\eta(t) = c(t)\,\xi(t)$), el límite medio cuadrático de «las sumas integrales» $\text{l.i.m.} \sum_{k=1}^{n} \eta(t_k)\,\Delta t_k = \int_a^b \eta(t)\,dt$ coincide, como magnitud aleatoria con probabilidad 1, con la integral corriente $\int_a^b \eta(t)\,dt$ de la trayectoria existente $\eta(t)$, $a \leqslant t \leqslant b$ (véase, por ejemplo el cap. IV del libro citado anteriormente, de I. I. Guijman y A. V. Skorojod).

Ejemplo (*diferenciación*). Examinemos *la derivada media cuadrática*

$$\xi'(t) = \underset{\Delta t \to \infty}{\text{l.i.m.}} \frac{\xi(t+\Delta t) - \xi(t)}{\Delta t}$$

del proceso aleatorio $\xi(t)$ (se tiene en cuenta el límite medio cuadrático). Para el proceso estacionario $\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\, d\Phi(\lambda)$ en el caso de existencia de tal derivada obtenemos

$$\xi'(t) = \int_{-\infty}^{\infty} i\lambda e^{i\lambda t}\, d\Phi(\lambda). \qquad (7.31)$$

Naturalmente, el límite medio cuadrático indicado

$$\underset{\Delta t \to 0}{\text{l.i.m.}} \frac{\xi(t+\Delta t) - \xi(t)}{\Delta t}$$

y la integral estocástica $\int_{-\infty}^{\infty} i\lambda e^{i\lambda t}\, d\Phi(\lambda)$ no siempre existen y, precisamente, entonces y sólo entonces, cuando

$$\int_{-\infty}^{\infty} \lambda^2 f(\lambda)\, d\lambda < \infty,$$

En efecto, para esta condición $\varphi(\lambda) = i\lambda e^{i\lambda t}$ es el límite de la función $\frac{e^{i\lambda(t+\Delta t)} - e^{i\lambda t}}{\Delta t}$, $\Delta t \to 0$, en el sentido de la convergencia media cuadrática y

$$\frac{\xi(t+\Delta t) - \xi(t)}{\Delta t} = \int_{-\infty}^{\infty} \left[\frac{e^{i\lambda(t+\Delta t)} - e^{i\lambda t}}{\Delta t} \right] d\Phi(\lambda) \to \int_{-\infty}^{\infty} \varphi(\lambda)\, d\Phi(\lambda).$$

Ejemplo (*fórmula de inversión*). Examinemos la fórmula (7.30) para

$$\varphi(\lambda) = \int_{-\infty}^{\infty} \chi_\Delta(\mu)\, \delta(\lambda - \mu)\, d\mu, \quad -\infty < \lambda < \infty,$$

donde $\Delta = (\lambda_1, \lambda_2]$ es un intervalo finito, $\chi_\Delta(\lambda)$ es su indicador ($\chi_\Delta(\lambda) = 1$ para $\lambda \in \Delta$ y $\chi_\Delta(\lambda) = 0$ para $\lambda \notin \Delta$) y

$$\delta(\lambda) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\lambda^2/2\sigma^2};$$

la función correspondiente $c(t)$ es

$$c(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\lambda t}\varphi(\lambda)\,d\lambda=\frac{1}{2\pi}\,\frac{e^{-i\lambda_2 t}-e^{-i\lambda_1 t}}{-it}\,e^{-\sigma^2 t/2},$$

$-\infty<t<\infty.$

Se ve fácilmente, que para $\sigma\to 0$ las funciones $\varphi(\lambda)$ convergen en la media cuadrática, hacia la función $\chi_\Delta(\lambda)$ (esto se puede fundamentar lo mismo que en la relación límite (7.29), utilizando el hecho de que las funciones $\varphi(\lambda)$ están limitadas y convergen uniformemente hacia $\chi_\Delta(\lambda)$ según λ, para $-a\leqslant\lambda\leqslant\lambda_1-\varepsilon$, $\lambda_1+\varepsilon\leqslant\lambda\leqslant\lambda_2-\varepsilon$, $\lambda_2+\varepsilon\leqslant\lambda\leqslant a$ para cualesquiera valores de a tan grandes como se quieran y de ε tan pequeños como se deseen. Ya que las magnitudes

$$\int_{-\infty}^{\infty}\varphi(\lambda)\,d\Phi(\lambda)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-i\lambda_2 t}-e^{-i\lambda_1 t}}{-it}\,e^{-\sigma^2t^2/2}\xi(t)\,dt$$

entran en el espacio $H(T)$ para $T=(-\infty,\infty)$, entonces en este espacio se contienen también las magnitudes $\Delta\Phi=\Phi(\lambda_2)-\Phi(\lambda_1)$:

$$\Delta\Phi=\int_{-\infty}^{\infty}\chi_\Delta(\lambda)\,d\Phi(\lambda)=\underset{\sigma\to 0}{\text{l.i.m.}}\,\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-i\lambda_2 t}-e^{-i\lambda_1 t}}{-it}\,e^{-\sigma^2t^2/2}\,\xi(t)\,dt$$

(7.32)

Esta es la llamada *fórmula de inversión*, que permite determinar el proceso aleatorio $\Phi(\lambda)$, $-\infty<\lambda<\infty$, con incrementos no correlacionados que figuran en la representación espectral del proceso estacionario $\xi(t)$, según su trayectoria $\xi(t)$, $-\infty<t<\infty$.

Examinemos el sistema lineal con la función de peso $w(t)$ absolutamente integrable

$$w(t)=0\ \text{para}\ t<0,\quad\int_0^{\infty}|w(t)|^2\,dt<\infty,$$

y con la característica espectral $\varphi(i\lambda)$, $\displaystyle\int_{-\infty}^{\infty}|\varphi(i\lambda)|^2\,d\lambda<\infty.$

Supongamos que en la entrada de este sistema aparece el proceso aleatorio estacionario en amplio sentido

$$\eta(t)=\int_{-\infty}^{\infty}e^{i\lambda t}\,d\Psi(\lambda)\tag{7.33}$$

con la densidad espectral $g(\lambda)$. Para tal acción exterior durante el tiempo desde 0 hasta t, el sistema pasa del estado inicial (digamos

$\xi(0) = 0$) al estado

$$\xi(t) = \int_0^t w(t-s)\,\dot{\eta}(s)\,ds. \tag{7.34}$$

De acuerdo con la igualdad general (7.30) el proceso aleatorio $\xi(t)$ se puede representar en la forma

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\varphi_t(\lambda)\,d\Psi(\lambda),$$

donde la función subintegral $\varphi_t(\lambda) = \int_0^t e^{-i\lambda s} w(s)\,ds$ converge hacia la característica espectral $\varphi(i\lambda)$ para $t \to \infty$. Se ve fácilmente que

el proceso aleatorio $\xi(t)$ converge en su media cuadrática hacia el proceso estacionario de la forma[1])

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\varphi(i\lambda)\,d\Psi(\lambda). \tag{7.35}$$

La fórmula (7.35) nos da de hecho la representación espectral de este proceso estacionario $\xi(t)$, que se establece en el curso del tiempo ($t \to \infty$) a la salida del sistema lineal con la característica espectral $\varphi(i\lambda)$ o sea en la representación espectral

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\,d\Phi(\lambda)$$

es necesario tomar

$$\Phi(\lambda) = \int_{-\infty}^{\lambda} \varphi(i\mu)\,d\Psi(\mu), \quad -\infty < \lambda < \infty.$$

De la relación

$$\mathbf{M}\,|\Phi(\lambda_2) - \Phi(\lambda_1)|^2 = \int_{\lambda_1}^{\lambda_2} |\varphi(i\lambda)|^2 g(\lambda)\,d\lambda$$

se ve directamente que la densidad espectral $f(\lambda)$ del proceso estacionario $\xi(t)$ es

$$f(\lambda) = |\varphi(i\lambda)|^2 g(\lambda). \tag{7.36}$$

---

[1]) Del mismo modo, que en el punto 3 del § 6, se podría examinar el paso límite (para $t_0 \to -\infty$) del proceso $\xi^0(t) = \int_{t_0}^{t} w(t-s)\,\dot{\eta}(s)\,ds$ al proceso estacionario $\xi(t)$ determinado por la fórmula (7.35).

Recordemos que la densidad espectral $f(\lambda)$ caracteriza la distribución de la energía total del proceso estacionario $\xi(t)=\int\limits_{-\infty}^{\infty} e^{i\lambda t}\,d\Phi(\lambda)$ por las componentes de la forma $\xi_\Delta(t)=\int\limits_{\lambda_1}^{\lambda_2} e^{i\lambda t}\,d\Phi(\lambda)$ en dependencia del intervalo $\Delta=(\lambda_1, \lambda_2)$ de aquellas frecuencias λ que entran en e espectro de frecuencias de la componente correspondiente $\xi_\Delta(t)$ $\left(\text{a saber, } \xi_\Delta(t) \text{ tiene una energía igual a } \int\limits_{\lambda_1}^{\lambda_2} f(\lambda)\,d\lambda\right)$. La relación (7.36) indica que al elegir la característica espectral adecuada $\varphi(i\lambda)$ del sistema lineal, se puede reforzar (o debilitar) las "oscilaciones de entrada" $\eta_\Delta=\int\limits_{\lambda_1}^{\lambda_2} e^{i\lambda t}\,d\Psi(\lambda)$ cuya energía $\int\limits_{\lambda_1}^{\lambda_2} g(\lambda)\,d\lambda$ se transforma en $\xi_\Delta(t)=\int\limits_{\lambda_1}^{\lambda_2} e^{i\lambda t}\varphi(i\lambda)\,d\Psi(\lambda)$; como ya se indicó la energía de las "oscilaciones de salida" $\xi_\Delta(t)=\int\limits_{\lambda_1}^{\lambda_2} e^{i\lambda t}\,d\Phi(\lambda)$ será igual a $\int\limits_{\lambda_1}^{\lambda_2} f(\lambda)\,d\lambda$, donde $f(\lambda)=|\varphi(i\lambda)|^2 g(\lambda)$.

Anteriormente, en el § 6 examinamos el comportamiento del sistema bajo la acción de perturbaciones aleatorias homogéneas $\dot{\eta}(t)$ con la suposición de que $\eta(t)=\int\limits_0^t \dot{\eta}(s)\,ds$ representa en sí el proceso con incrementos no correlacionados, que satisface a las condiciones (7.14); en estas condiciones el proceso de salida

$$\xi(t)=\int\limits_0^t w(t-s)\,d\eta(s),$$

que se puede representar simbólicamente por la fórmula (7.34) con el cambio de $d\eta(s)$ por $\dot{\eta}(s)\,ds$, será convergente hacia el proceso estacionario para $t\to\infty$. Con respecto a este proceso estacionario

$$\xi(t)=\int\limits_{-\infty}^{\infty} e^{i\lambda t}\varphi(i\lambda)\,d\Psi(\lambda)$$

fue establecido que él admite la representación espectral y tiene una densidad espectral $f(\lambda)=\frac{\sigma^2}{2\pi}|\varphi(i\lambda)|^2$ (véase (7.22)—(7.25)).

Tal proceso estacionario entra formalmente en el esquema general descrito anteriormente cuando $g(\lambda)=\frac{\sigma^2}{2\pi}$. Naturalmente, en el sentido de la definición dada antes, la función $g(\lambda)=\frac{\sigma^2}{2\pi}$ no puede servir como densidad espectral del proceso estacionario corriente $\left(\text{ya que para ella } \int_{-\infty}^{\infty} g(\lambda)\,d\lambda=\infty\right)$. Pero es cómodo, de un modo puramente simbólico, introducir el proceso estacionario correspondiente $\dot{\eta}(t)$ con la densidad espectral $g(\lambda)=\frac{\sigma^2}{2\pi}$, utilizándolo en el sentido de que (véase (7.19))

$$\int_{-\infty}^{\infty} c(t)\,\dot{\eta}(t)\,dt=\left(\int_{-\infty}^{\infty} c(t)\,d\eta(t)\right)=\int_{-\infty}^{\infty} \varphi(\lambda)\,d\Psi(\lambda).$$

Este proceso estacionario generalizado $\eta(t)$ se llama, corrientemente, *ruido blanco.*

Ejemplo (*movimiento perturbado del péndulo*). En conclusión, examinemos un ejemplo en el cual se puede fácilmente observar algunas particularidades del comportamiento de los sistemas lineales bajo acción de las oscilaciones aleatorias. Examinemos precisamente el movimiento del péndulo (para la existencia del rozamiento) con un gran período de oscilaciones propias, que se describen por la ecuación

$$x''(t)+2hx'(t)+\omega_0^2x(t)=0$$

y se dan en forma explícita por la función

$$x(t)=Ae^{-ht}\operatorname{sen}(\omega t+\theta),\quad \omega=\sqrt{\omega_0^2-h^2}.$$

La función de transmisión de tal sistema es

$$\varphi(p)=\frac{1}{p^2+2hp+\omega_0^2}$$

(la característica espectral es igual a $\varphi(i\lambda)$).

Supongamos que este sistema se encuentra en un barco. Y supongamos que sea la frecuencia ω de oscilaciones propias del péndulo mucho menor que la frecuencia Ω del balanceo del barco: $\omega\ll\Omega$. Si consideramos que en el péndulo actúan los golpes aleatorios como resultado del balanceo, que se originan aproximadamente después de pequeños intervalos de tiempo $\Delta t\sim\frac{\pi}{\Omega}$, siendo esta perturbación externa $\dot{\eta}(t)$ homogénea según el tiempo y se puede considerar formalmente que $\dot{\eta}(t)$ es «un ruido blanco»[1]), entonces el movimiento establecido del péndulo representará en sí un proceso estacionario con la densidad espectral de la forma $f(\lambda) = \frac{\sigma^2}{2\pi} \mid \varphi(i\lambda) \mid^2$.

Recordemos que la densidad espectral caracteriza la distribución de energía del proceso aleatorio $\xi(t) = \int\limits_{\infty}^{\infty} e^{i\lambda t}\, d\Phi(\lambda)$ según las componentes de «las oscilaciones elementales» en dependencia de la frecuencia λ, a saber: la amplitud media de las oscilaciones aleatorias, dadas por la integral estocástica $\int\limits_{\lambda_1}^{\lambda_2} e^{i\lambda t}\, d\Phi(\lambda)$ es igual a $\int\limits_{\lambda_1}^{\lambda_2} f(\lambda)\, d\lambda$. En nuestro caso la densidad espectral es

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{(\lambda^2 - \omega_0^2) + 4h^2\lambda^2}$$

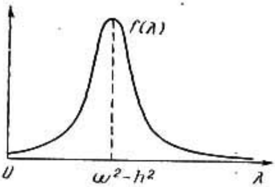

Fig. 23. Vista general de la densidad espectral

$$f(\lambda) = \frac{\sigma^2}{\pi} \cdot \frac{1}{(\lambda^2 - \omega_0^2)^2 + 4h^2\lambda^2}.$$

y tiene un máximo para $\lambda^2 = \omega^2 - h^2$ (fuertemente expresado para «un coeficiente de rozamiento» h pequeño) y decrece suficientemente rápido al alejarse del punto máximo (fig. 23). Esto quiere decir que en el proceso aleatorio $\xi(t)$ predominan, fuertemente, «las oscilaciones elementales» con frecuencias próximas a la *frecuencia propia* ω del sistema examinado.

Para la comparación indiquemos que si se considera la acción externa (del balanceo del barco) como oscilación armónica de frecuencia Ω, entonces el péndulo durante el movimiento establecido realizará oscilaciones forzadas con esta frecuencia Ω, lo que se diferencia cualitativamente del resultado obtenido anteriormente.

8. PROCESOS DE DIFUSION

1. Procesos aleatorios representados por la integral estocástica Itô. Para mayor claridad, imaginémonos un sistema determinado, que con la acción exterior $y(t)$, $t \geqslant t_0$, cambia su estado fásico $x(t)$ de tal modo que

$$x'(t) = a(t) + \sigma(t)\, y(t),$$

[1]) Referente a la acción de tal tipo de proceso aleatorio sobre el sistema lineal, además de lo dicho anteriormente sobre «el ruido blanco» véase también el punto 1 del § 6.

o sea en forma integral

$$x(t)=x(t_0)+\int_{t_0}^{t} a(s)\,ds+\int_{t_0}^{t} \sigma(s)\,y(s)\,ds,$$

donde los coeficientes $a(t)$ y $\delta(t)$, que son parámetros del sistema, pueden depender no sólo del tiempo t, sino también de la acción $y(s)$, $t_0 \leqslant s \leqslant t$. Al actuar sobre tal sistema las perturbaciones caóticas aleatorias $\dot{\eta}(t)$, $t \geqslant t_0$, como modelo adecuado que describe su comportamiento, puede servir el proceso aleatorio

$$\xi(t)=\xi(t_0)+\int_{t_0}^{t} a(s)\,ds+\int_{t_0}^{t} \sigma(s)\,d\eta(s), \tag{8.1}$$

donde $\eta(t)=\int_{t_0}^{t} \dot{\eta}(t)\,dt$ (compárese el punto 1 del § 6).

Más adelante examinaremos los procesos aleatorios $\xi(t)$ del tipo indicado, suponiendo que $\eta(t)$, $t \geqslant t_0$, es el proceso «standard» del movimiento browniano[1]) con una media nula y el coeficiente unitario de difusión ($\mathbf{M}[\eta(t)-\eta(t_0)]^2=t-t_0$), y los coeficientes $a(t)$, $\sigma(t)$ dependen, hablando en general, de $\eta(s)$, $t_0 \leqslant s \leqslant t$ (más exacto, $a(t)$ y $\delta(t)$ para cada t son magnitudes aleatorias, cuyos valores en el momento t se determinan unívocamente por las trayectorias del proceso aleatorio $\eta(s)$, $t_0 \leqslant s \leqslant t$).

Pasemos a la definición exacta de las integrales estocásticas existentes en la fórmula (8.1).

Sea $\varphi(t)$ una función real constante a trozos en el segmento $c_1 \leqslant t \leqslant c_2$, que para una división de este segmento por los puntos $c_1=t_0, t_1, \ldots, t_n=c_2$ toma los valores aleatorios $\varphi(t_k)$ en cada uno de los intervalos $\Delta_k=(t_k, t_{k+1})$, $k=0, \ldots, n-1$. Subrayamos una vez más que los valores de la función $\varphi(t)$ son magnitudes aleatorias. Supongamos que

$$\int_{c_1}^{c_2} \varphi(t)\,dt=\sum_{k=0}^{n-1} \varphi(t_k)\,\Delta t_k,$$

donde $\Delta t_k=t_{k+1}-t_k$, $k=0, \ldots, n-1$. Está claro, que la integral definida de este modo posee las propiedades corrientes de aditividad y homogeneidad:

$$\int_{c_1}^{c_2} [\lambda_1\varphi_1(t)+\lambda_2\varphi_2(t)]\;dt=\lambda_1\int_{c_1}^{c_2} \varphi_1(t)\,dt+\lambda_2\int_{c_1}^{c_2} \varphi_2(t)\,dt \tag{8.2}$$

---

[1]) En otras palabras, $\dot{\eta}(t)$ es el «ruido blanco» de Gauss.

y ademas de esto;

$$\mathbf{M}\int_{c_1}^{c_2}\varphi(t)\,dt=\int_{c_1}^{c_2}[\mathbf{M}\varphi(t)]\,dt \tag{8.3}$$

y

$$\left\|\int_{c_1}^{c_2}\varphi(t)\,dt\right\|\leqslant\int_{c_1}^{c_2}\|\varphi(t)\|\,dt \tag{8.4}$$

(donde $\|\eta\|=\sqrt{\mathbf{M}\eta^2}$ significa el valor medio cuadrático de la magnitud aleatoria η).

Sea ahora $\varphi(t)$ una función aleatoria en el segmento $c_1\leqslant t\leqslant c_2$, que puede ser representada como límite medio cuadrático de las funciones constantes a trozos $\varphi_n(t)$, $n=1, 2, \ldots$, o más exáctamente,

$$\|\varphi(t)-\varphi_n(t)\|\to 0 \quad \text{para} \quad n\to\infty \tag{8.5}$$

uniformemente según t, $c_1\leqslant t\leqslant c_2$. Evidentemente,

$$\|\varphi_n(t)-\varphi_m(t)\|\leqslant\|\varphi_n(t)-\varphi(t)\|+\|\varphi_m(t)-\varphi(t)\|\to 0$$

para n, $m\to\infty$ también uniformemente según t y, por consiguiente,

$$\left\|\int_{c_1}^{c_2}\varphi_n(t)\,dt-\int_{c_1}^{c_2}\varphi_m(t)\,dt\right\|\leqslant\int_{c_1}^{c_2}\|\varphi_n(t)-\varphi_m(t)\|\,dt\to 0.$$

Esto demuestra que existe el límite medio cuadrático $\underset{n\to\infty}{\mathrm{l.i.m.}}\int_{c_1}^{c_2}\varphi_n(t)\,dt$; evidentemente, este límite es el mismo también para cualquiera sucesión $\varphi_n(t)$, $n=1, 2, \ldots$, del tipo indicado anteriormente. Supongamos que

$$\int_{c_1}^{c_2}\varphi(t)\,dt=\underset{n\to\infty}{\mathrm{l.i.m.}}\int_{c_1}^{c_2}\varphi_n(t)\,dt. \tag{8.6}$$

Se ve fácilmente, que en tal paso límite se conservan las relaciones (8.2)—(8.4).

Señalemos que para la función aleatoria $\varphi(t)$, *uniformemente continua en la media cuadrática*:

$$\|\varphi(t+\Delta t)-\varphi(t)\|\to 0 \tag{8.7}$$

para $\Delta t\to 0$ uniformemente según t, $c_1\leqslant t\leqslant c_2$; la *integral estocástica* determinada por la igualdad (8.6) es el límite medio cuadrático de las sumas integrales corrientes:

$$\int_{c_1}^{c_2}\varphi(t)\,dt=\mathrm{l.i.m.}\sum_{k=0}^{n-1}\varphi(t_k)\,\Delta t_k$$

para una división cada vez más pequeña del segmento $[c_1, c_2]$ por los puntos $c_1 = t_0, t_1, \ldots, t_n = c_2$ en intervalos de longitud $\Delta t_k = t_{k+1} - t_k$, cuando máx $\Delta t_k \to 0$.

Pasemos a la definición de integral estocástica del tipo $\int_{c_1}^{c_2} \varphi(t) \times d\eta(t)$, donde $\varphi(t)$ es una función aleatoria, y $\eta(t)$ un proceso aleatorio con incrementos independientes; con esto nos limitamos solamente al caso cuando el proceso $\eta(t)$ es el movimiento browniano (con una media nula y con el coeficiente unitario de difusión).

Para las funciones corrientes $\varphi(t)$ (no aleatorias), la integral estocástica de tal tipo fue determinada antes en el § 6. Para las funciones aleatorias $\varphi(t)$ es fundamental la suposición de que

para cualquier u los incrementos $\eta(t) - \eta(u)$, $t \geqslant u$, no dependen de las magnitudes aleatorias $\varphi(s)$, $s \leqslant u$.

Sea $\varphi(t)$ una función constante a trozos real, que toma los valores aleatorios $\varphi(t_k)$ en los intervalos correspondientes $\Delta_k = (t_k, t_{k+1})$, $k = 0, \ldots, n-1$. Pongamos

$$\int_{c_1}^{c_2} \varphi(t)\, d\eta(t) = \sum_{k=0}^{n-1} \varphi(t_k)\, \Delta\eta(t_k),$$

donde $\Delta\eta(t_k) = \eta(t_{k+1}) - \eta(t_k)$; $k = 0, \ldots, n-1$. Está claro, que la integral así determinada posee las propiedades corrientes de aditividad y homogeneidad (compárese (8.2)).

Supondremos que $\mathbf{M}[\varphi(t)]^2 < \infty$ para todos los t, $c_1 \leqslant t \leqslant c_2$. De la condición de independencia de los incrementos $\eta(t) - \eta(u)$, $t \geqslant u$, de los valores $\varphi(s)$, $s \leqslant u$, se deduce que

$$\mathbf{M}[\varphi(t_k) \cdot \Delta\eta(t_k)] = \mathbf{M}\varphi(t_k) \cdot \mathbf{M}\Delta\eta(t_k) = 0,$$
$$\mathbf{M}[\varphi(t_k) \cdot \Delta\eta(t_k)]^2 = \mathbf{M}\varphi^2(t_k) \cdot \mathbf{M}[\Delta\eta(t_k)]^2 = \mathbf{M}\varphi^2(t_k) \cdot \Delta(t_k),$$
$$k = 0, \ldots, n-1;$$

además de esto, para todos los $k \neq j$ (considerando para concretizar que $k < j$) tenemos

$$\mathbf{M}[\varphi(t_k)\, \Delta\eta(t_k) \cdot \varphi(t_j)\, \Delta\eta(t_j)] = \mathbf{M}[\varphi(t_k)\, \Delta\eta(t_k)\, \varphi(t_j)] \times \mathbf{M}\Delta\eta(t_j) = 0.$$

Evidentemente,

$$\mathbf{M} \int_{c_1}^{c_2} \varphi(t)\, d\eta(t) = 0, \tag{8.8}$$

$$\left\| \int_{c_1}^{c_2} \varphi(t)\, d\eta(t) \right\|^2 = \int_{c_1}^{c_2} \| \varphi(t) \|^2\, dt$$

y además de esto,

$$\mathbf{M}\left[\int_{c_1}^{c_2} \varphi(t)\,d\eta(t)\cdot\int_{c_1}^{c_2} \psi(t)\,d\eta(t)\right]=\int_{c_1}^{c_2} [\mathbf{M}\,\varphi(t)\,\psi(t)]\,dt. \qquad (8.9)$$

Sea ahora $\varphi(t)$ una función aleatoria en el segmento $c_1 \leqslant t \leqslant c_2$, que puede ser representada como el límite medio cuadrático de las funciones constantes a trozos $\varphi_n(t)$, $n = 1, 2, \ldots$; más exactamente, supongamos cumplida la condición (8.5). Entonces

$$\left\|\int_{c_1}^{c_2} \varphi_n(t)\,d\eta(t)-\int_{c_1}^{c_2} \varphi_m(t)\,d\eta(t)\right\|^2=\int_{c_1}^{c_2} \|\varphi_n(t)-\varphi_m(t)\|^2 dt \to 0$$

para $n, m \to \infty$ y, por consiguiente, existe el límite medio cuadrático $\underset{n\to\infty}{\mathrm{l.i.m.}} \int_{c_1}^{c_2} \varphi_n(t)\,d\eta(t)$. Se ve fácilmente, que este límite no depende de la elección de la sucesión $\varphi_n(t)$, $n = 1, 2, \ldots$, que satisfaga la condición (8.5). Pongamos

$$\int_{c_1}^{c_2} \varphi(t)\,d\eta(t)=\underset{n\to\infty}{\mathrm{l.i.m.}} \int_{c_1}^{c_2} \varphi_n(t)\,d\eta(t). \qquad (8.10)$$

Evidentemente, para tal paso límite, las relaciones (8.8), (8.9) se extienden a las funciones límites $\varphi(t)$.

Señalemos, que para las funciones uniformemente continuas $\varphi(t)$ (en la media cuadrática) la integral estocástica definida por la igualdad (8.10) es el limite medio cuadrático de las sumas integrales correspondientes:

$$\int_{c_1}^{c_2} \varphi(t)\,d\eta(t)=\mathrm{l.i.m.}\sum_{k=0}^{n-1} \varphi(t_k)\,\Delta\eta(t_k).$$

Después de introducir las integrales estocásticas del tipo (8.6), (8.10) también se puede hablar con conocimiento de causa sobre el proceso aleatorio $\xi(t)$ representado por la fórmula (8.1).

Suponemos que las funciones aleatorias $a(t)$ y $\sigma(t)$ cuyos valores para cada t se determinan unívocamente por el proceso aleatorio $\eta(s)$, $t_0 \leqslant s \leqslant t$, son *uniformemente continuas en la media cuadrática* en cada segmento limitado. En esta suposición existe la integral $\int_{t_0}^{t} a(s)\,ds$, y ya que para cualquier u, los incrementos $\eta(t)-\eta(u)$, $t \geqslant u$, no dependen de $\eta(s)$, $t_0 \leqslant s \leqslant u$, ellos tampoco dependen de las magnitudes $\sigma(s)$, $s \leqslant u$

de modo que existe la integral estocástica $\int_{t_0}^{t} \sigma(t)\, d\eta(t)$ determinada anteriormente.

Examinemos el caso en que el proceso aleatorio $\xi(t)$ es representable por la integral estocástica (8.1). Aquí consideramos que existen los procesos aleatorios $a(t)$, $\sigma(t)$ y $\eta(t)$ (y también el proceso aleatorio $\xi(t)$) y la cuestión se concierne sólo al enlace de estos procesos, expresado por la fórmula (8.1).

Claro está que los valores $\xi(t)$ dados por esta fórmula se determinan unívocamente por el valor inicial $\xi(t_0)$ y por el proceso aleatorio $\eta(s)$, $t_0 \leqslant s \leqslant t$; examinaremos el proceso aleatorio $\xi(t)$, $t \geqslant t_0$ de este tipo. *Se indicará que la relación integral (8.1) es un corolario* de la relación local

$$\Delta\xi(t) \sim a(t)\,\Delta t + \sigma(t)\,\Delta\eta(t) \tag{8.11}$$

(donde $\Delta\xi(t) = \xi(t+\Delta t) - \xi(t)$, $\Delta\eta(t) = \eta(t+\Delta t) - \eta(t)$ y $\Delta t \to 0$), que comprenderemos en el sentido[1]) de que

$$\mathbf{M}\,\{\Delta\xi(t) - [a(t)\,\Delta t + \sigma(t)\,\Delta\eta(t)] \mid \eta(s),\ s \leqslant t\} = o(\Delta t), \tag{8.12}$$
$$\mathbf{M}\,\{\Delta\xi(t) - [a(t)\,\Delta t + \sigma(t)\,\Delta\eta(t)]\}^2 = o(\Delta t),$$

donde las magnitudes $o(\Delta t)$ son infinitamente pequeñas de orden superior en comparación con Δt, o más exacto,

$$\frac{\| o(\Delta t) \|}{\Delta t} \to 0 \quad \text{para} \quad \Delta t \to 0$$

uniformemente por t en cada intervalo finito $t_0 \leqslant t \leqslant t_1$.

Detengámonos detalladamente en estas relaciones. Las magnitudes $a(t)$ y $\sigma(t)$ en la primera de éstas se determinan unívocamente por $\eta(s)$, $s \leqslant t$, y para l a t r a y e c t o r i a f i j a d a $\eta(s)$, $s \leqslant t$, son sencillamente constantes, mientras que el incremento $\Delta\xi(t)$ es aleatorio y tiene una distribución de probabilidades que depende de la trayectoria correspondiente $\eta(s)$, $s \leqslant t$, y el incremento $\Delta\eta(t)$ no depende de $\eta(s)$, $s \leqslant t$. Por eso, para la esperanza matemática condicional (para los valores fijados de $\eta(s)$, $s \leqslant t$) tenemos

$$\mathbf{M}\,\{\sigma(t)\,\Delta\eta(t) \mid \eta(s),\ s \leqslant t\} = \sigma(t)\,\mathbf{M}\Delta\eta(t) = 0,$$
$$\mathbf{M}\,\{\Delta\xi(t) - [a(t)\,\Delta t + \sigma(t)\,\Delta\eta(t)] \mid \eta(s),\ s \leqslant t\} =$$
$$= \mathbf{M}\,\{\Delta\xi(t) \mid \eta(s),\ s \leqslant t\} - a(t)\,\Delta t.$$

---

[1]) Véase el § 2 del siguiente cap. IV. Más adelante, en este párrafo utilizaremos, esencialmente, el aparato de las esperanzas matemáticas condicionales, incluyendo la fórmula de la esperanza matemática repetida del tipo

$$\mathbf{M}\,\{\mathbf{M}\,[\xi \mid \eta] \mid \zeta\} = \mathbf{M}\,\{\xi \mid \zeta\},$$

donde, hablando a grueso modo, el complejo de condiciones «ζ» se determina unívocamente para cada una de las condiciones fijas «η»; se puede leer más detalladamente sobre este particular, por ejemplo, en el libro citado antes de G u i j m a n y S k o r o j o d.

Luego, ya que se supone que la función aleatoria $a(t)$ es continua en la media cuadrática, la magnitud $\| a(t) \|$ está limitada en cada segmento $t_0 \leqslant t \leqslant t_1$ y, se ve fácilmente, que

$$\| \Delta\xi(t) - [a(t)\Delta t + \sigma(t)\Delta\eta(t)] \|^2 = \\ = \| \Delta\xi(t) - \sigma(t)\Delta\eta(t) \|^2 + o(\Delta t).$$

De tal modo, las relaciones (8.12) también se pueden comprender en el sentido de que

$$\mathbf{M}\{\Delta\xi(t) \mid \eta(s),\ s \leqslant t\} = a(t)\Delta t + o(\Delta t), \tag{8.13}$$

$$\| \Delta\xi(t) - \sigma(t)\Delta\eta(t) \|^2 = o(\Delta t). \tag{8.14}$$

Lema 1. *Con la condición* (8.13) *para cualesquiera* $t_1 \leqslant t_2$ *tiene lugar la siguiente igualdad:*

$$\mathbf{M}\{\xi(t_2) - \xi_1(t_1) \mid \eta(s),\ s \leqslant t_1\} = \mathbf{M}\left\{\int_{t_1}^{t_2} a(t)\,dt \mid \eta(s),\ s < t_1\right\}. \tag{8.15}$$

Demostración. Dividamos el segmento $[t_1, t_2]$ con los puntos $t_1 = s_0, s_1, \ldots, s_n = t_2$ en intervalos de longitud $\Delta s_k = s_{k+1} - s_k$, $k = 0, \ldots, n-1$. Supongamos que

$$o(\Delta s_k) = \mathbf{M}\{\Delta\xi(s_k) - a(s_k)\Delta s_k \mid \eta(s),\ s \leqslant s_k\},$$

donde $\Delta\xi(s_k) = \xi(s_{k+1}) - \xi(s_k)$; $k = 0, \ldots, n-1$. Según la suposición, $\frac{\| o(\Delta s_k) \|}{\Delta s_k} \to 0$ uniformemente por k para máx $\Delta s_k \to 0$. Tomando por segunda vez la esperanza matemática condicional (pero ahora con la condición $\eta(s)$, $s \leqslant t_1$, donde $t_1 \leqslant s_k$) tenemos

$$\mathbf{M}\{o(\Delta s_k) \mid \eta(s),\ s \leqslant t_1\} = \mathbf{M}\{\Delta\xi(s_k) - \\ - a(s_k)\Delta s_k \mid \eta(s),\ s \leqslant t_1\},$$

de donde obtenemos que

$$\| \mathbf{M}\{\Delta\xi(s_k) - a(s_k)\Delta s_k \mid \eta(s),\ s \leqslant t_1\} \| = \\ = \sqrt{\mathbf{M}[\mathbf{M}\{o(\Delta s_k) \mid \eta(s),\ s \leqslant t_1\}]^2} \leqslant \\ \leqslant \sqrt{\mathbf{M}[\mathbf{M}\{o(\Delta s_k)^2 \mid \eta(s),\ s \leqslant t_1\}]} = \\ = \| o(\Delta s_k) \|,\quad k = 0, \ldots, n-1.$$

Por esto, como se ve fácilmente,

$$\left\| \mathbf{M}\{\xi(t_2) - \xi(t_1) \mid \eta(s),\ s < t_1\} - \mathbf{M}\left\{\int_{t_1}^{t_2} a(t)\,dt \mid \eta(s),\ s \leqslant t_1\right\} \right\| \leqslant \\ \leqslant \sum_{k=0}^{n-1} \mathbf{M}\{\Delta\xi(s_k) - a(s_k)\Delta s_k \mid \eta(s),\ s \leqslant t_1\} \| +$$

$$+\left\|\mathbf{M}\left\{\int_{t_1}^{t_2} a(t)\,dt-\sum_{k=0}^{n-1} a(s_k)\,\Delta s_k \mid \eta(s),\ s\leqslant t_1\right\}\right\|\leqslant$$

$$\leqslant \sum_{k=0}^{n-1} \|o(\Delta s_k)\|+\left\|\int_{t_1}^{t_2} a(t)\,dt-\sum_{k=0}^{n-1} a(s_k)\,\Delta s_k\right\|\to 0$$

para máx $\Delta s_k \to 0$ y, por consiguiente,

$$\|\mathbf{M}\{\xi(t_2)-\xi(t_1)\mid\eta(s),\ s\leqslant t_1\}-\mathbf{M}\left\{\int_{t_1}^{t_2} a(t)\,dt\mid\eta(s),\ s\leqslant t_1\right\}\right\|=0,$$

es decir, tiene lugar la igualdad (8.15).

Pasemos del proceso inicial $\xi(t)$ al proceso aleatorio

$$\tilde{\xi}(t)=\xi(t)-\xi(t_0)-\int_{t_0}^{t} a(s)\,ds. \tag{8.16}$$

De la igualdad (8.15) se deduce que

$$\mathbf{M}\{\Delta\tilde{\xi}(t)\mid\eta(s),\ s\leqslant t\}=0. \tag{8.17}$$

Además de esto, junto con la condición (8.14) para los incremento

$$\Delta\tilde{\xi}(t)=\Delta\xi(t)-\int_{t}^{t+\Delta t} a(s)\,ds$$

será evidentemente

$$\|\Delta\tilde{\xi}(t)-\sigma(t)\,\Delta\eta(t)\|^2=o(\Delta t), \tag{8.18}$$

donde $o(\Delta t)$ es un infinitamente pequeño del mismo tipo que en (8.14), es decir, $\frac{o(\Delta t)}{\Delta t}\to 0$ uniformemente en cada intervalo finito de $t_0\leqslant t\leqslant t_1$.

Lema 2. *Con la condición* (8.14) *tiene lugar la siguiente igualdad*:

$$\tilde{\xi}(t)=\int_{t_0}^{t} \sigma(s)\,d\eta(s). \tag{8.19}$$

D e m o s t r a c i ó n. Elijamos una división del segmento $[t_0, t$ y valoremos la diferencia correspondiente

$$\tilde{\xi}(t)-\sum_{k=0}^{n-1}\sigma(t_k)\,\Delta\eta(t_k)=\sum_{k=0}^{n-1}[\Delta\tilde{\xi}(t_k)-\sigma(t_k)\,\Delta\eta(t_k)].$$

De la condición (8.17) se deduce que para $t_k \neq t_j$ (digamos, $t_k < t_j$)

$$\mathbf{M}\Delta\tilde{\xi}(t_k)\,\Delta\tilde{\xi}(t_j) = \mathbf{M}\,[\mathbf{M}\,\{\Delta\tilde{\xi}(t_k)\,\Delta\tilde{\xi}(t_j) \mid \eta(s),$$
$$s \leqslant t_j\}] = \mathbf{M}\,[\Delta\tilde{\xi}(t_k)\,\mathbf{M}\,\{\Delta\tilde{\xi}(t_j) \mid \eta(s),\ s \leqslant t_j\}] = 0,$$

ya que para los valores fijados $\eta(s)$, $s \leqslant t_j$, las magnitudes $\Delta\tilde{\xi}(t_k)$ son constantes, análogamente:

$$\mathbf{M}\,[\Delta\tilde{\xi}(t_k)\,\sigma(t_j)\,\Delta\eta(t_j)] = 0,$$
$$\mathbf{M}\,[\sigma(t_k)\,\Delta\eta(t_k)\,\sigma(t_j)\,\Delta\eta(t_j)] = 0.$$

De aquí, tomando en consideración la condición (8.18), obtenemos que

$$\left\| \tilde{\xi}(t) - \sum_{k=0}^{n-1} \sigma(t_k)\,\Delta\eta(t_k) \right\|^2 = \sum_{k=0}^{n-1} \| \Delta\tilde{\xi}(t_k) - \sigma(t_k)\,\Delta\eta(t_k) \|^2 =$$
$$= \sum_{k=1}^{n-1} o(\Delta t_k) \to 0$$

para máx $\Delta t_k \to 0$, es decir, "las sumas integrales" $\sum\limits_{k=0}^{n-1} \sigma(t_k)\,\Delta\eta(t_k)$ tienen su límite $\tilde{\xi}(t)$ y, por consiguiente, $\tilde{\xi}(t) = \int\limits_0^t \sigma(s)\,d\eta(s)$ que es, lo que se quería demostrar.

La expresión

$$d\xi(t) = a(t)\,dt + \sigma(t)\,d\eta(t) \tag{8.20}$$

se llama *diferencial estocástica* del proceso aleatorio $\xi(t)$, si

$$\xi(t) = \xi(t_0) + \int_{t_0}^{t} a(s)\,ds + \int_{t_0}^{t} \sigma(s)\,d\eta(s).$$

Uniendo los lemas 1, 2 obtenemos el resultado siguiente.

Teorema 1. *Con la condición* (8.12) (*equivalente a* (8.13), (8.14)) *el proceso aleatorio* $\xi(t)$ *tiene la diferencial estocástica, expresada por la fórmula* (8.20).

Hallemos la diferencial estocástica de la función aleatoria

$$\xi(t) = \varphi(t,\ \eta(t)), \tag{8.21}$$

donde $\eta(t)$ es el proceso del movimiento browniano, y $\varphi(t, x)$ es una función real (no aleatoria) de las variables t y x, que tiene derivadas continuas $\frac{\partial\varphi}{\partial t}$, $\frac{\partial\varphi}{\partial x}$ y $\frac{\partial^2\varphi}{\partial x^2}$. Para que sea más claro el objeto de nuestros cálculos, exponemos inmediatamente la solución:

$$d\xi(t) = \left[\frac{\partial\varphi(t,\ \eta(t))}{\partial t} + \frac{1}{2}\,\frac{\partial^2\varphi(t,\ \eta(t))}{\partial x^2}\right] dt + \frac{\partial\varphi(t,\ \eta(t))}{\partial x}\,d\eta(t). \tag{8.22}$$

Subrayemos aquí particularidad en comparación con la diferencial corriente para las funciones (no aleatorias) diferenciables $\xi(t)$ y $\eta(t)$, que tiene la forma $d\xi(t)=\frac{\partial\varphi(t,\eta(t))}{\partial t}dt+\frac{\partial\varphi(t,\eta(t))}{\partial x}d\eta(t)$; precisamente, en la fórmula (8.22) se contiene el miembro complementario $\frac{1}{2}\frac{\partial^2\varphi(t,\eta(t))}{\partial x^2}dt$, que aparece, como podemos cerciorarnos más adelante, gracias a que, para el movimiento browniano $\eta(t)$ examinado con el coeficiente unitario de difusión $\mathbf{M}[\Delta\eta(t)]^2=\Delta t$. En general[1]),

$$\mathbf{M}[\Delta\eta(t)]^{2k}=\frac{(2k)!}{k!\,2^k}(\Delta t)^k,\quad k=1,\ 2,\ \ldots,\tag{8.23}$$

y esto aporta su específica al valorar el miembro residual

$$o(\Delta t)=\left[\frac{\partial\varphi(t+\theta_1\Delta t,\ \eta(t)+\Delta\eta(t)}{\partial t}-\frac{\partial\varphi(t,\ \eta(t))}{\partial t}\right]+\\+\frac{1}{2}\left[\frac{\partial^2\varphi(t,\ \eta(t)+\theta_2\Delta\eta(t))}{\partial x^2}-\frac{\partial^2\varphi(t,\ \eta(t))}{\partial x^2}\right]\tag{8.24}$$

(donde $|\theta_1|,|\theta_2|\leqslant 1$) en la relación

$$\Delta\xi(t)=\frac{\partial\varphi(t,\ \eta(t))}{\partial t}\Delta t+\frac{\partial\varphi(t,\ \eta(t))}{\partial x}\Delta\eta(t)+\\+\frac{1}{2}\frac{\partial^2\varphi(t,\ \eta(t))}{\partial x^2}[\Delta\eta(t)]^2+o(\Delta t),\tag{8.25}$$

que se obtiene en la fórmula de Taylor para las funciones $\varphi(t, x)$.

Supongamos complementariamente, que las derivadas $\frac{\partial\varphi}{\partial t}$, $\frac{\partial\varphi}{\partial x}$ y $\frac{\partial^2\varphi}{\partial x^2}$ satisfacen la tal llamada *condición de Lipshitz* (para la función $y(t, x)$ ésta significa que

$$|y(t+\Delta t,\ x+\Delta x)-y(t,\ x)|\leqslant C_1|\Delta t|+C_2|\Delta x|,$$

donde C_1, C_2 son determinadas constantes). Entonces las funciones aleatorias $\frac{\partial\varphi(t,\ \eta(t))}{\partial t}$, $\frac{\partial\varphi(t,\ \eta(t))}{\partial x}$ y $\frac{\partial^2\varphi(t,\ \eta(t))}{\partial x^2}$ son uniformemente continuas en la media cuadrática y, por consiguiente, las funciones $\left\|\frac{\partial\varphi}{\partial t}\right\|$, $\left\|\frac{\partial\varphi}{\partial x}\right\|$ y $\left\|\frac{\partial^2\varphi}{\partial x^2}\right\|$ están uniformemente limitadas en cada segmento finito $t_0\leqslant t\leqslant t_1$. Teniendo en cuenta esto, y también

---

[1]) Para la magnitud de Gauss η con la media nula y la dispersión σ^2, que tiene la función característica $\varphi(u)=e^{-\sigma^2u^2/2}$, $-\infty<u<\infty$, la fórmula $\varphi^{(k)}(0)=i^{-k}\mathbf{M}\eta^k$, $k=0,\ 1,\ \ldots$, nos da

$$\mathbf{M}\eta^{2k}=\frac{(2k)!}{k!\,2^k}\sigma^{2k},\quad k=0,\ 1,\ \ldots$$

circunstancia de que los incrementos $\Delta\eta(t)$ no dependen de $\eta(s)$, $s \leqslant t$, se puede deducir fácilmente, de las relaciones (8.23)—(8.25), que las funciones

$$a(t) = \frac{\partial\varphi(t, \eta(t))}{\partial t} + \frac{1}{2}\frac{\partial^2\varphi(t, \eta(t))}{\partial x^2} \quad \text{y} \quad \sigma(t) = \frac{\partial\varphi(t, \eta(t))}{\partial t}$$

satisfacen a las condiciones (8.12) siendo justa, para las magnitudes $o(\Delta t)$ que figuran en estas condiciones en cada segmento limitado $t_0 \leqslant t \leqslant t_1$, la valuación $|o(\Delta t)| \leqslant C(\Delta t)^{3/2}$, donde C es una constante determinada. Según el teorema 1, el proceso aleatorio examinado $\xi(t) = \varphi(t, \eta(t))$ tiene la diferencial estocástica $d\xi(t) = a(t)\,dt + \sigma(t)\,d\eta(t)$.

Ejemplo. Calculemos la integral $\int_{t_0}^{t} \eta(s)\,d\eta(s)$, donde $\eta(t)$, $t \geqslant t_0$, es el proceso del movimiento browniano. Intentemos hallar el proceso aleatorio de la forma $\xi(t) = \varphi(t, \eta(t))$ que tiene la diferencial estocástica $d\xi(t) = \eta(t)\,d\eta(t)$; para tal proceso $\int_{t_0}^{t} \eta(s)\,d\eta(s) = \xi(t) - \xi(t_0)$.

De acuerdo con la fórmula general (8.22), la función incógnita $\varphi(t, x)$ deberá satisfacer las siguientes condiciones:

$$\frac{\partial\varphi}{\partial t} + \frac{1}{2}\frac{\partial^2\varphi}{\partial x^2} = 0, \quad \frac{\partial\varphi}{\partial x} = x.$$

Hallamos fácilmente que

$$\varphi(t, x) = \frac{1}{2}x^2 + c(t) \quad \text{y} \quad c'(t) + \frac{1}{2} = 0,$$

de donde $\varphi(t, x) = \frac{1}{2}(x^2 - t)$. Como resultado obtenemos

$$\int_{t_0}^{t} \eta(s)\,d\eta(s) = \frac{1}{2}[\eta^2(t) - \eta^2(t_0) - (t - t_0)].$$

Señalemos de modo aleccionador que para la función diferenciable corriente $\eta(t)$, la integral correspondiente es

$$\int_{t_0}^{t} \eta(t)\,d\eta(t) = \frac{1}{2}[\eta^2(t) - \eta^2(t_0)].$$

Examinemos el proceso aleatorio de Márkov $\xi(t)$, que tiene la diferencial estocástica $a(t)\,dt + \sigma(t)\,d\eta(t)$.

Consideraremos que los valores de las funciones aleatorias $a(t)$ y $\sigma(t)$ dependen explícitamente, sólo de t y del estado variable $\xi(t)$, de modo que $a(t) = a(t, \xi(t))$ y $\sigma(t) = \sigma(t, \xi(t))$, donde $a(t, x)$ y $\sigma(t, x)$ son unas funciones (no aleatorias) del par de variables t, x. En términos generales, esto significa que durante el tiempo Δt el proceso pasa desde cada estado $\xi(t) = x$ al estado $x + \Delta\xi(t)$, donde

$$\Delta\xi(t) \sim a(t, \xi(t))\,\Delta t + \sigma(t, \xi(t))\,\Delta\eta(t).$$

Antes interpretábamos esta relación en el sentido de que se cumpliesen las condiciones (8.12). Reforcemos algo a estas condiciones, sustituyéndolas por las siguientes:

$$\mathbf{M}\{\Delta\xi(t) - [a(t, \xi(t))\,\Delta t + \sigma(t, \xi(t))\,\Delta\eta(t)] \mid \xi(t)\} = o(\Delta t), \quad (8.26)$$
$$\mathbf{M}\{[\Delta\xi(t) - a(t, \xi(t))\,\Delta t - \sigma(t, \xi(t))\,\Delta\eta(t)]^2 \mid \xi(t)\} = o(\Delta t),$$

donde $o(\Delta t)$, igual que anteriormente, significa también para cualquier valor fijado $\xi(t) = x$, una magnitud infinitamente pequeña de orden superior en comparación con $\Delta t \to 0$, uniformemente según t en cada intervalo limitado $t_0 \leqslant t \leqslant t_1$

$$|a(t, x)| \leqslant C, \qquad C_1 \leqslant |\sigma(t, x)| \leqslant C_2$$

para las constantes determinadas C, C_1, C_2.

Mostremos que *para la condición* (8.26) *el proceso aleatorio de Márkov* $\xi(t)$ *entra en la clase de los llamados procesos de difusión*, cuyas distribuciones de probabilidades están íntimamente ligadas con las ecuaciones diferenciales, que describen los procesos físicos de difusión[1]) (las ecuaciones mencionadas serán examinadas en el punto siguiente).

Demos una difinición formal del proceso de difusión. Para esto nos será cómodo introducir las siguientes designaciones:

$$\eta_\varepsilon = \begin{cases} \eta & \text{para } |\eta| \leqslant \varepsilon \\ 0 & \text{para } |\eta| > \varepsilon. \end{cases}$$

Al proceso aleatorio de Márkov $\xi(t)$ se le llama proceso *de difusión*, si para cualquier $\varepsilon > 0$, los incrementos $\Delta\xi(t)$ satisfacen las condiciones

$$\left.\begin{aligned} &\mathbf{P}\{|\Delta\xi(t)| > \varepsilon \mid \xi(t) = x\} = o(\Delta t), \\ &\mathbf{M}\{(\Delta\xi(t))_\varepsilon \mid \xi(t) = x\} = a(t, x)\,\Delta t + o(\Delta t), \\ &\mathbf{M}\{(\Delta\xi(t))_\varepsilon^2 \mid \xi(t) = x\} = \sigma^2(t, x)\,\Delta t + o(\Delta t), \end{aligned}\right\} \quad (8.27)$$

donde la magnitud $o(\Delta t)$ tiene el mismo sentido que en las condiciones (8.26).

Señalemos, que si se tienen en cuenta no las magnitudes $(\Delta\xi(t))_\varepsilon$ «cortadas hasta ε», sino los propios incrementos $\Delta\xi(t)$, entonces de las relaciones (8.26) se deduce inmediatamente que

$$\mathbf{M}\{\Delta\xi(t) \mid \xi(t) = x\} = a(t, x)\,\Delta t + o(\Delta t)$$

[1]) Con respecto a esta cuestión véase por ejemplo, el libro de (A. Ya. Jinchin, «Leyes asimptóticas de la teoría de probabilidades»).

y

$$\mathbf{M}\,\{\Delta\xi^2(t) \mid \xi(t) = x\} = \sigma^2(t,\ x)\,\Delta t + o\,(\Delta t).$$

Para pasar a las magnitudes «cortadas» $(\Delta\xi(t))_\varepsilon$ que figuran en las condiciones (8.27), tenemos que demostrar algunas proposiciones auxiliares.

Más adelante se tratará de las distribuciones condicionales de probabilidades y de las esperanzas matemáticas para los valores fijados $\xi(t) = x$.

Examinemos la magnitud $\eta = a\,(t,\ x)\,\Delta t + \sigma\,(t,\ x)\,\Delta\eta\,(t)$, donde $\Delta\eta\,(t)$ es el incremento del movimiento browniano correspondiente durante el tiempo Δt. La densidad de probabilidad de esta magnitud es $p\,(y) = \frac{1}{2\sqrt{2\pi\sigma}}\,e^{-(y-a)^2/2\sigma^2}$, donde para abreviar suponemos $a = a\,(t,\ x)\,\Delta t$ y $\sigma^2 = \sigma^2(t,\ x)\,\Delta t$. Evidentemente,

$$\mathbf{M}\,[\eta - \eta_\varepsilon]^2 = \int\limits_{|y|>\varepsilon} y^2 p\,(y)\,dy =$$

$$= \frac{\Delta t}{\sqrt{2\pi}\,\sigma\,(t,\ x)} \int\limits_{|z|>\frac{\varepsilon}{\sqrt{\Delta t}}} z^2 e^{-(z-a)^2/2\sigma(t,\ x)^2}\,dz = o\,(\Delta t)$$

uniformemente según t y x para $|\,a\,(t,\ x)\,| \leqslant C$, $C_1 \leqslant \sigma\,(t,\ x) \leqslant C_2$. Aquí nos interesa la magnitud $\Delta\xi\,(t)$, para la cual

$$\mathbf{M}\,[\Delta\xi\,(t) - \eta]^2 = o\,(\Delta t).$$

De la desigualdad $\|\,\Delta\xi\,(t) - \eta_\varepsilon\,\| \leqslant \|\,\Delta\xi - \eta\,\| + \|\,\eta - \eta_\varepsilon\,\|$ obtenemos también que

$$\mathbf{M}\,[\Delta\xi\,(t) - \eta_\varepsilon]^2 = o\,(\Delta t).$$

Ya que $|\,\Delta\xi\,(t)\,|^2 \leqslant 2\,[\,|\,\eta_\varepsilon|^2 + |\,\Delta\xi\,(t) - \eta_\varepsilon\,|^2]$, entonces

$$\chi_A\,|\,\Delta\xi\,(t)\,|^2 \leqslant 2[\,\chi_A\varepsilon^2 + |\Delta\xi\,(t) - \eta_\varepsilon\,|^2],$$

donde $\chi_A = 1$, al aparecer el suceso A, en caso contrario $\chi_A = 0$. Para el suceso $A = \{|\,\Delta\xi\,(t)\,| > 2\varepsilon\}$ que entra en el suceso $B = \{|\,\Delta\xi\,(t) - \eta_\varepsilon\,| > \varepsilon\}$ tenemos

$$\mathbf{M}\,[\chi_A\,\Delta\xi\,(t)]^2 \leqslant 2\,[\varepsilon^2\mathbf{P}\,(A) + \mathbf{M}\,[\Delta\xi\,(t) - \eta_\varepsilon]^2]$$

y

$$\mathbf{P}\,(A) \leqslant \mathbf{P}\,(B) \leqslant \frac{1}{\varepsilon^2}\,\mathbf{M}\,[\Delta\xi\,(t) - \eta_\varepsilon]^2.$$

Como resultado obtenemos

$$\mathbf{M}\,[\chi_A\,\Delta\xi\,(t)]^2 = \mathbf{M}\,[\Delta\xi\,(t) - (\Delta\xi\,(t))_{2\varepsilon}]^2 \leqslant$$

$$\leqslant 4\mathbf{M}\,[\Delta\xi\,(t) - \eta_\varepsilon]^2 = o\,(\Delta t)$$

para cualquier $\varepsilon > 0$. Ahora ya es fácil deducir que

$$\mathbf{P}\{|\Delta\xi(t)| > 2\varepsilon\} \leqslant \frac{1}{4\varepsilon^2}\mathbf{M}[\Delta\xi(t) - (\Delta\xi(t))_{2\varepsilon}]^2 = o(\Delta t),$$

$$|\mathbf{M}\Delta\xi(t) - \mathbf{M}(\Delta\xi(t))_{2\varepsilon}| \leqslant \frac{1}{2\varepsilon}\mathbf{M}[\Delta\xi(t) - (\Delta\xi(t))_{2\varepsilon}]^2 = o(\Delta t)$$

y, ya que $\|\Delta\xi(t)\|^2 = \sigma^2(t, x)\Delta t + o(\Delta t)$,

$$|\mathbf{M}(\Delta\xi(t))^2 - \mathbf{M}(\Delta\xi(t))^2_{2\varepsilon}| = \|\Delta\xi(t)\|^2 - \|(\Delta\xi(t))_{2\varepsilon}\|^2 \leqslant$$
$$\leqslant 2\|\Delta\xi(t)\| \cdot \|(\Delta\xi(t)) - (\Delta\xi(t))_{2\varepsilon}\| = o(\Delta t).$$

de donde, si sustituimos 2ε por ε, se deducen directamente las condiciones (8.27).

2. Ecuaciones diferenciales de Kolmogórov. Examinemos el proceso aleatorio de Márkov $\xi(t)$. Consideraremos que para cualesquiera $s \leqslant t$ con la condición de que $\xi(s) = x$, la magnitud aleatoria $\xi(t)$ tiene la densidad condicional de probabilidad $p(s, x, t, y)$, $-\infty < y < \infty$, de modo que en el tiempo desde s hasta t, el proceso pasa del estado x a un estado determinado y, $y' \leqslant y \leqslant y''$, con la probabilidad

$$\mathbf{P}\{y' \leqslant \xi(t) \leqslant y'' \mid \xi(s) = x\} = \int_{y'}^{y''} p(s, x, t, y)\,dy.$$

De hecho $p(s, x, t, y)$ nos da la densidad condicional de probabilidad de la magnitud $\xi(t)$ para cualquier condición sobre la magnitud $\xi(u)$, $u \leqslant s$, incluyendo en sí $\xi(s) = x$, ya que para el proceso aleatorio de Márkov, la magnitud $\xi(t)$ no depende de $\xi(u)$, $u \leqslant s$, para el estado conocido $\xi(s) = x$.

Demostremos que la densidad de paso de la probabilidad $p(s, x, t, y)$ satisfase la llamada *ecuación de Kolmogórov-Chapman*:

$$p(s, x, t, y) = \int_{-\infty}^{\infty} p(s, x, u, z)\,p(u, z, t, y)\,dz \tag{8.28}$$

para cualesquiera $s \leqslant u \leqslant t$.

En efecto, para las condiciones $\xi(s) = x$, $\xi(u) = z$, la magnitud $\xi(t)$ no depende de x y tiene una densidad condicional de probabilidad, igual a $p(u, z, t, y)$; examinando la distribución de probabilidades de las magnitudes $\xi(u)$ y $\xi(t)$ para el valor fijado $\xi(s) = x$, hallamos que

$$p(s, x, u, z)\,p(u, z, t, y), \quad -\infty < z,\ y < \infty,$$

es la densidad conjunta de probabilidad de las magnitudes $\xi(u)$, $\xi(t)$[1]; integrando esta densidad conjunta por z, $-\infty < z < \infty$, obte-

[1]) Análogamente, para cualesquiera $s \leqslant t_1 \leqslant \ldots \leqslant t_n$, $p(s, x, t_1, y_1) \times p(t_1, y_1, t_2, y_2) \ldots p(t_{n-1}, y_{n-1}, t_n, y_n)$, $-\infty < y_1, \ldots, y_n < \infty$, es la densidad conjunta de probabilidad de las magnitudes aleatorias $\xi(t_1), \xi(t_2), \ldots, \xi(t_n)$ para el valor fijado $\xi(s) = x$.

nemos la densidad de probabilidad $p(s, x, t, y)$ de la magnitud $\xi(t)$ (para el valor fijado $\xi(s) = x$).

Sea $\xi(t)$ un proceso de difusión. Evidentemente las condiciones (8.27) se pueden expresar directamente por su densidad de paso $p(s, x, t, y)$ del modo siguiente:

$$\int\limits_{|y|>\varepsilon} p(t, x, t+\Delta t, y)\, dy = o(\Delta t),$$

$$\int\limits_{|y-x|\leqslant\varepsilon} (y-x)\, p(t, x, t+\Delta t, y)\, dy = a(t, x)\,\Delta t + o(\Delta t), \quad (8.29)$$

$$\int\limits_{|y-x|\leqslant\varepsilon} (y-x)^2\, p(t, x, t+\Delta t, y)\, dy = b(t, x)\,\Delta t + o(\Delta t),$$

donde, recordamos una vez más que $o(\Delta t)$ significa una magnitud infinitamente pequeña de orden superior en relación a $\Delta t \to 0$, uniformemente según t en cada intervalo limitado $t_0 \leqslant t \leqslant t_1$ y $b(t, x) = \sigma^2(t, x)$.

Teorema 2. *Supongamos que la densidad de paso de la probabilidad $p(s, x, t, y)$ tiene las derivadas continuas $\frac{\partial p}{\partial s}$, $\frac{\partial p}{\partial x}$ y $\frac{\partial^2 p}{\partial x^2}$ uniformemente continuas según y en cada intervalo limitada $y' \leqslant y \leqslant y''$. Entonces ésta satisface a la ecuación diferencial*

$$-\frac{\partial p}{\partial s} = a(s, x)\frac{\partial p}{\partial x} + \frac{1}{2} b(s, x)\frac{\partial^2 p}{\partial x^2}. \quad (8.30)$$

D e m o s t r a c i ó n. Tomemos una función continua arbitraria $\varphi(x)$, que sea igual a cero fuera de un intervalo finita y supongamos que

$$\varphi(s, x) = \int\limits_{-\infty}^{\infty} \varphi(y)\, p(s, x, t, y)\, dy.$$

De la ecuación de Kolmogórov—Chapman se deduce que para cualesquiera $t_0 \leqslant s \leqslant u \leqslant t$

$$\varphi(s, x) = \int\limits_{-\infty}^{\infty} \varphi(y) \int\limits_{-\infty}^{\infty} p(s, x, u, z)\, p(u, z, t, y)\, dy =$$

$$= \int\limits_{-\infty}^{\infty} \varphi(u, z)\, p(s, x, u, z)\, dz.$$

Evidentemente, la función $\varphi(s, x)$ tiene las derivadas continuas $\frac{\partial \varphi}{\partial s}$, $\frac{\partial \varphi}{\partial x}$ y $\frac{\partial^2 \varphi}{\partial x^2}$. Desarrollemos la función $\varphi(u, z)$ en los alrededores del punto x (para el valor fijado u) según la fórmula de Taylor

$$\varphi(u, z) - \varphi(u, x) = \frac{\partial \varphi(u, x)}{\partial x}(z-x) + \frac{1}{2}\left[\frac{\partial^2 \varphi(u, x)}{\partial x^2} + O(\delta_\varepsilon)\right](z-x)^2,$$

donde

$$\delta_\varepsilon = \sup_{|z-x|\leq\varepsilon} \left| \frac{\partial^2\varphi(u, z)}{\partial x^2} - \frac{\partial^2\varphi(u, x)}{\partial x^2} \right| \to 0$$

para $\varepsilon \to 0$. De las relaciones (8.29) obtenemos

$$\varphi(s, x) - \varphi(u, x) = \int_{-\infty}^{\infty} [\varphi(u, z) - \varphi(u, x)]\, p(s, x, u, z)\, dz =$$

$$= \int_{|z-x|\leq\varepsilon} [\varphi(u, z) - \varphi(u, x)]\, p(s, x, u, z)\, dz +$$

$$+ o(u-s) = \frac{\partial\varphi(u, x)}{\partial x} \int_{|z-x|\leq\varepsilon} (z-x)\, p(s, x, u, z)\, dz +$$

$$+ \frac{1}{2}\left[\frac{\partial^2\varphi(u, x)}{\partial x^2} + O(\delta_\varepsilon)\right] \int_{|z-x|\leq\varepsilon} (z-x)^2\, p(s, x, u, z)\, dz +$$

$$+ o(u-s) = \left\{a\frac{\partial\varphi}{\partial x} + \frac{1}{2} b\left[\frac{\partial^2\varphi}{\partial x^2} + O(\delta_\varepsilon)\right]\right\}(u-s) + o(u-s),$$

donde $O(\delta_\varepsilon) \to 0$ para $\varepsilon \to 0$, de donde se ve que

$$\lim_{u\to s} \frac{\varphi(s, u) - \varphi(u, x)}{u-s} = a(s, x)\frac{\partial\varphi}{\partial x} + \frac{1}{2} b(s, x)\frac{\partial^2\varphi}{\partial x^2},$$

y, de tal modo,

$$-\frac{\partial\varphi}{\partial s} = a(s, x)\frac{\partial\varphi}{\partial x} + \frac{1}{2} b(s, x)\frac{\partial^2\varphi}{\partial x^2}.$$

Teniendo en cuenta la definición de la función $\varphi(s, x)$, esta ecuación se puede escribir de la siguiente forma:

$$\int_{-\infty}^{\infty} \varphi(y)\left[\frac{\partial p}{\partial s} + a(s, x)\frac{\partial p}{\partial x} + \frac{1}{2} b(s, x)\frac{\partial^2 p}{\partial x^2}\right] dy = 0,$$

donde, recordemos $\varphi(y)$ es una función continua arbitraria, igual a cero fuera del intervalo finito y, por consiguiente, se deberá cumplir la igualdad

$$\frac{\partial p}{\partial s} + a(s, x)\frac{\partial p}{\partial x} + \frac{1}{2} b(s, x)\frac{\partial^2 p}{\partial x^2} = 0.$$

El teorema queda demostrado.

Señalemos que la densidad de paso de la probabilidad $p(s, x, t, y)$ coincide con la tal llamada solución fundamental de la ecuación (8.30), que se determina por la misma condición que para cualquier función continua limitada $\varphi(x)$

$$\lim_{t\to s} \int_{-\infty}^{\infty} \varphi(y)\, p(s, x, t, y)\, dy = \varphi(x).$$

Teorema 3. *Supongamos que se tienen las derivadas continuas*

$$\frac{\partial}{\partial t} p(s, x, t, y), \quad \frac{\partial}{\partial y}[a(t, y) p(s, x, t, y)], \quad \frac{\partial^2}{\partial y^2}[b(t, y) p(s, x, t, y)].$$

Entonces la densidad de paso $p(s, x, t, y)$ satisface a la ecuación diferencial

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial y}[a(t, y) p(s, x, t, y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[b(t, y) p(s, x, t, y)]. \qquad (8.31)$$

Demostración. Exactamente igual que en la demostración del teorema 2, se puede obtener fácilmente, que para cualquier función $\varphi(x)$ dos veces diferenciable continuamente, que sea igual a cero fuera de un intervalo finito, existe el límite

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\int_{-\infty}^{\infty} \varphi(y) p(s, x, t+\Delta t, y)\, dy - \varphi(x)\right] =$$
$$= a(t, x)\varphi'(x) + \frac{1}{2} b(t, x)\varphi''(x).$$

Tenemos

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} p(s, x, t, y)\varphi(y)\, dy =$$
$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\int_{-\infty}^{\infty} p(s, x, t+\Delta t, y)\varphi(y)\, dy - \int_{-\infty}^{\infty} p(s, x, t, z)\varphi(z)\, dz\right] =$$
$$= \int_{-\infty}^{\infty} p(s, x, t, z) \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\int_{-\infty}^{\infty} p(s, z, t+\Delta t, y)\varphi(y)\, dy - \varphi(z)\right] dz =$$
$$= \int_{-\infty}^{\infty} p(s, x, t, z)\left[a(t, z)\varphi'(z) + \frac{1}{2} b(t, z)\varphi''(z)\right] dz.$$

Integrando por partes la última expresión, obtenemos

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} p(s, x, t, y)\varphi(y)\, dy = \int_{-\infty}^{\infty}\left[\frac{\partial}{\partial t} p(s, x, t, y)\right]\varphi(y)\, dy =$$
$$= \int_{-\infty}^{\infty}\left\{-\frac{\partial}{\partial y}[a(t, y) p(s, x, t, y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[b(t, y) p(s, x, t, y)]\right\}\varphi(y)\, dy,$$

de donde, se deduce la igualdad (8.31) debido a la arbitrariedad de la función $\varphi(y)$. El teorema queda demostrado.

La ecuación (8.30) se llama *inversa*, y la ecuación (8.31) es *la ecuación directa de Kolmogórov*.

El método de las ecuaciones diferenciales descrito anteriormente es utilizable al examinar, no sólo, la densidad de probabilidad $p(s, x, t, y)$ de la magnitud $\xi(t)$, sino también para otras magnitudes aleatorias, ligadas con el comportamiento del proceso de difusión.

Examinemos, por ejemplo, el momento τ de la primera salida del proceso aleatorio $\xi(t)$, $t \geqslant t_0$ del intervalo (c_1, c_2) (consideramos, que la trayectoria $\xi = \xi(t)$ del proceso de difusión es una función continua y por eso la magnitud τ es el momento en que por primera vez esta trayectoria alcanza los valores $x = c_1$ o $x = c_2$).

Designemos por $\mathbf{P}_{s,x}(A)$ la probabilidad del suceso A y por $\mathbf{M}_{s,x}\eta$, la esperanza matemática de la magnitud aleatoria η en condiciones que $\xi(s) = x$. Según la suposición, para el proceso de difusión $\xi = \xi(t)$ tenemos

$$\mathbf{P}_{s,x}\{|\xi(s+\Delta s) - x| > \varepsilon\} = o(\Delta s) \tag{8.32}$$

uniformemente por s en cada intervalo finito (cualquiera que sea el número fijado $\varepsilon > 0$), de donde se deduce que para la condición $\xi(s) = x$, la probabilidad de encontrarse durante el intervalo de tiempo de s a $s + \Delta s$ en un determinado punto y, $|y - x| > \varepsilon$, es la magnitud $o(\Delta s)$ de un orden infinitesimal superior del mismo tipo que en la relación (8.32); por consiguiente, también

$$\mathbf{P}_{s,x}\{\tau \leqslant s + \Delta s\} = o(\Delta s) \tag{8.33}$$

uniformemente por s en cada intervalo finito.

Supongamos que para la condición $\xi(s) = x$ la magnitud aleatoria τ tiene la densidad de probabilidad $p_{s,x}(t)$. Examinemos la función

$$u(s, x) = \mathbf{M}_{s,x} \exp\left\{\int_s^\tau U(t)\,dt\right\} = \int_{-\infty}^{\infty} \exp\left\{\int_s^t U(u)\,du\right\} p_{s,x}(t)\,dt,$$

donde la función continua $U(t)$ tiene una parte real $\operatorname{Re} U(t) \leqslant 0$ de modo que $\exp\left\{\int_s^\tau U(t)\,dt\right\}$ es una magnitud limitada. Para $U(t) = i\lambda t$ tenemos, por ejemlo,

$$u(s, x) = \int_{-\infty}^{\infty} e^{i\lambda t} p_{s,x}(t)\,dt$$

(donde $p_{s,x}(t) = 0$ para $t < s$), es decir, $u(s, x)$ nos da la función característica de la magnitud τ en el punto λ, $-\infty < \lambda < \infty$.

Para cualquier magnitud limitada η, evidentemente,

$$\begin{aligned}\mathbf{M}_{s,x}\eta = \mathbf{M}_{s,x}\{\eta \mid \tau > s + \Delta s\}\,\mathbf{P}\{\tau > s + \Delta s\} + \\ + \mathbf{M}_{s,x}\{\eta \mid \tau \leqslant s + \Delta s\}\,\mathbf{P}\{\tau \leqslant s + \Delta s\} = \\ = \mathbf{M}_{s,x}\{\eta \mid \tau > s + \Delta s\}\,\mathbf{P}\{\tau > s + \Delta s\} + o(\Delta s),\end{aligned}$$

donde $o(\Delta s)$ es una magnitud del mismo tipo que en la relación (8.33). Tenemos

$$\mathbf{M}_{s,x}\left\{\exp\int_s^{\tau} U(t)\,dt \mid \tau > s+\Delta s\right\} =$$

$$= \mathbf{M}_{s,x}\left\{\mathbf{M}\left[\exp\int_s^{\tau} U(t)\,dt \mid \tau > s+\Delta s,\ \xi(s+\Delta s)=\eta\right] \mid \tau > s+\Delta s\right\} =$$

$$= \mathbf{M}_{s,x}\left\{\exp\left[\int_s^{s+\Delta s} U(t)\,dt\right]\mathbf{M}_{s+\Delta s,\,\eta}\exp\left[\int_{s+\Delta s}^{\tau} U(t)\,dt\right]\right\},$$

ya que las relaciones $\tau > s+\Delta s$, $\xi(s+\Delta s)=\eta$ representan en sí las condiciones sobre el comportamiento de la trayectoria $\xi(t)$, $t \leqslant s+\Delta s$, y para el proceso de Márkov $\xi = \xi(t)$ pueden ser sustituidas (como fue hecho anteriormente) sencillamente por la condición $\xi(s+\Delta s) = \eta$. Teniendo en cuenta que

$$\exp\left[\int_s^{s+\Delta s} U(t)\,dt\right] = 1 + U(s)\,\Delta s + o(\Delta s) = \frac{1}{1-U(s)\,\Delta s} + o(\Delta s),$$

obtenemos

$$u(s, x) = \mathbf{M}_{s,x}\left\{\frac{1}{1-U(s)\,\Delta s}\,u[s+\Delta s,\ \xi(s+\Delta s)]\right\} + o(\Delta s),$$

o bien,

$$\int_{-\infty}^{\infty}[u(s+\Delta s, y) - u(s, x)]\,p(s, x, s+\Delta s, y)\,dy +$$
$$+ U(s)\,u(s, x)\,\Delta s + o(\Delta s) = 0,$$

lo pue para el proceso de difusión es equivalente a la relación

$$\int_{|y-x|\leqslant\varepsilon}[u(s+\Delta s, y) - u(s, x)]\,p(s, x, s+\Delta s, y)\,dy +$$
$$+ U(s)\,u(s, x)\,\Delta s + o(\Delta s) = 0.$$

Supongamos ahora, que la función $u(s, x)$ tiene las derivadas continuas $\frac{\partial u}{\partial s}$, $\frac{\partial u}{\partial x}$ y $\frac{\partial^2 u}{\partial x^2}$. Entonces, igualmente que al deducir la ecuación diferencial (8.30), obtenemos

$$\int_{|y-x|\leqslant\varepsilon}\left[\frac{\partial u(s, x)}{\partial s}\Delta s + \frac{\partial u(s, x)}{\partial x}\Delta x + \frac{1}{2}\,\frac{\partial^2 u(s, x)}{\partial x^2}(1+O(\delta))\,\Delta x^2\right]\times$$
$$\times\, p(s, x, s+\Delta s, y)\,dy + U(s)\,u(s, x)\,\Delta s + o(\Delta s) =$$

$$= \left[\frac{\partial u(s, x)}{\partial s} + a(s, x)\frac{\partial u(s, x)}{\partial x} + \right.$$

$$\left. + \frac{1}{2} b(s, x)\frac{\partial^2 u(s, x)}{\partial x^2}(1 + O(\delta)) + U(s)\,u(s, x)\right]\Delta s + o(\Delta s) = 0,$$

donde $O(\delta) \to 0$ para $\varepsilon \to 0$, de donde se ve que la función $u(s, x)$ satisface a la ecuación diferencial

$$\frac{\partial u}{\partial s} + a\frac{\partial u}{\partial x} + \frac{1}{2} b\frac{\partial^2 u}{\partial x^2} + Uu = 0. \tag{8.34}$$

Adémas de esto, es evidente que para $x = c_1$ y $x = c_2$ (cuando $\tau = s$), según la definición de la función $u(s, x)$ tenemos

$$u(s, c_1) = u(s, c_2) \equiv 1.$$

CAPITULO IV

ALGUNAS TAREAS DE PRONOSTICACIÓN, FILTRADO. Y REGULACIÓN DE LOS PROCESOS ALEATORIOS

§ 1. TAREA GENERAL SOBRE LA APROXIMACION OPTIMA. EJEMPLOS

La mayoría de las tareas de pronosticación, filtrado y regulación de los procesos aleatorios descansan sobre la base del problema fundamental de hallar una valuación suficientemente buena de una determinada magnitud desconocida ξ según los valores existentes $\eta(t)$ de un determinado proceso aleatorio, en tal o cual intervalo de tiempo $a \leqslant t \leqslant b$; con ello corrientemente $\xi = \xi(t)$ es el valor en el momento variable de tiempo t (o relacionado al futuro) de otro proceso aleatorio, ligado de alguna forma con el proceso «observado» $\eta(t)$, $a \leqslant t \leqslant b$.

El problema más sencillo de tal género fue examinado por nosotros ya en el § 4 del cap. I, cuando tratamos la aproximación óptima de la magnitud ξ por combinaciones lineales de la forma $\sum\limits_{k=1}^{n} c_k\eta_k$, donde η_k, $k=1, \ldots, n$ son determinadas magnitudes dadas. Precisamente se exigía hallar la magnitud $\hat{\xi} = \sum\limits_{k=1}^{n} c_k^0\eta_k$ de tal modo que

$$\|\xi-\hat{\xi}\| = \min_{c_k}\left\|\xi-\sum_{k=1}^{n} c_k\eta_k\right\|. \tag{1.1}$$

Como se indicó, si introducimos el espacio lineal H de $(n+1)$ dimensiones de todas las magnitudes $\eta = \sum\limits_{k=0}^{n} c_k\eta_k$ (donde $\eta_0 = \xi$ y $c_0, \ldots, c_n$ son coeficientes reales arbitrarios) con el producto escalar

$$\langle\eta_1, \eta_2\rangle = \mathbf{M}(\eta_1\cdot\eta_2) \tag{1.2}$$

y con la distancia correspondiente

$$\|\eta_1-\eta_2\| = \sqrt{\mathbf{M}(\eta_1-\eta_2)^2}, \tag{1.3}$$

entonces la magnitud $\hat{\xi} = \sum_{k=1}^{n} c_k^0 \eta_k$ que satisface a la condición (1.1) significa geométricamente la base de la perpendicular bajada desde el punto $\xi \in H$ al subespacio L de todas las magnitudes $\eta = \sum_{k=1}^{n} c_k \eta_k$ y se determina, unívocamente, por la condición de ortogonalidad de la diferencia $\xi - \hat{\xi}$ hacia el subespacio L

$$\langle \xi - \hat{\xi}, \eta \rangle = 0, \ \eta \in L, \tag{1.4}$$

esto es equivalente al siguiente sistema lineal de ecuaciones con relación a $c_1^0, \ldots, c_n^0$

$$\sum_{k=1}^{n} c_k^0 \langle \eta_k, \eta_j \rangle = \langle \xi, \eta_j \rangle, \quad j = 1, \ldots, n. \tag{1.5}$$

Examinemos la tarea general sobre la aproximación de una magnitud ξ por las magnitudes η de un conjunto L determinado, que es como se dice, *el hiperplano*, y que significa lo siguiente: para cualquier elemento $\eta_0 \in L$ el conjunto de magnitudes $\Delta = \eta - \eta_0$, $\eta \in L$ forma *un espacio lineal*, o sea, contiene junto con cualquier Δ_1, Δ_2 también su combinación lineal $c_1\Delta_1 + c_2\Delta_2$ (donde c_1, c_2 son coeficientes reales arbitrarios).

La magnitud $\hat{\xi} \in L$ la denominamos *aproximación óptima* de ξ, si

$$\| \xi - \hat{\xi} \| = \min_{\eta \in L} \| \xi - \eta \|. \tag{1.6}$$

Lema sobre la perpendicular. *La condición* (1.6) *es equivalente a lo siguiente*:

$$\langle \xi - \hat{\xi}, \eta - \hat{\xi} \rangle = 0 \ \textit{para todos los} \ \eta \in L. \tag{1.7}$$

Antes de demostrar esta afirmación, demos a la relación (1.7) una interpretación geométrica sencilla. Si designamos, precisamente, por $\hat{L}$ el espacio lineal de todos los elementos $\eta - \hat{\xi}$, $\eta \in L$, entonces la relación (1.7) significará que el producto escalar de los elementos $\xi - \hat{\xi}$ y $\Delta = \eta - \hat{\xi}$ será igual a cero para todos los $\Delta \in L$, en otras palabras, la diferencia $\xi - \hat{\xi}$ es p e r p e n d i c u l a r al subespacio $\hat{L}$ (fig. 24).

La magnitud $\hat{\xi}$ se llama *proyección* de la magnitud ξ en el hiperplano L y la diferencia $\xi - \hat{\xi}$, perpendicular (desde el punto ξ).

Señalemos también que en el caso cuando el hiperplano L por sí mismo es un espacio lineal ($L = \hat{L}$), la condición (1.7) es equivalente a la condición (1.4).

D e m o s t r a c i ó n d e l l e m a. Supongamos que $\hat{\xi} \in L$ es la aproximación óptima de ξ. Evidentemente,

$$\| \xi - \hat{\xi} \| = \min \| \xi - \hat{\xi} - \Delta \|,$$

donde el mínimo se toma por todas las diferencias $\Delta = \eta - \hat{\xi}$, $\eta \in L$. Pero el conjunto de tales diferencias forma el espacio lineal $\hat{L}$ y en particular, en $\hat{L}$ se contiene junto con el elemento Δ el elemento $\lambda\Delta$, donde λ es un número real. Fijando cualquier valor $\Delta \neq 0$ y suponiendo que

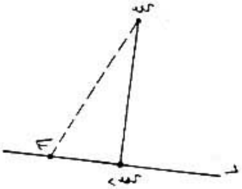

Fig. 24.

$$A^2 = \| \xi - \hat{\xi} \|^2,\ B = \langle \xi - \hat{\xi},\ \Delta \rangle,\ C = \| \Delta \|^2,$$

tendremos

$$\min_{\lambda} \| \xi - \hat{\xi} - \lambda\Delta \|^2 = \min_{\lambda} \times (A^2 + 2B\lambda + \lambda^2 C^2) = A^2.$$

Se ve que el mínimo de la forma cuadrática $A^2 + 2B\lambda + \lambda C^2$ se alcanza cuando $\lambda = 0$ y, por consiguiente, el coeficiente B es igual a cero, es decir, se cumple la condición (1.7). A su vez, si está cumplida esta condición, entonces para cualquier magnitud $\eta \in L$ se tiene

$$\| \xi - \eta \|^2 = \| (\xi - \hat{\xi}) - (\eta - \hat{\xi}) \|^2 = \| \xi - \hat{\xi} \|^2 + 2 \langle \xi - \hat{\xi},\ \eta - \hat{\xi} \rangle + \| \eta - \hat{\xi} \|^2 = \| \xi - \hat{\xi} \|^2 + \| \eta - \hat{\xi} \|^2 \geqslant \| \xi - \hat{\xi} \|^2,$$

de donde se ve, que $\hat{\xi}$ es la aproximación óptima de la magnitud ξ. El lema queda demostrado.

Señalemos, que *sólo existe la magnitud única* $\hat{\xi} \in L$, *que satisface la condición* (1.7) (*es decir, también la condición* (1.6)).

En efecto, si $\hat{\xi}_1$, $\hat{\xi}_2$ satisfacen las relaciones

$$\langle \xi - \hat{\xi}_1,\ \eta - \hat{\xi}_1 \rangle = 0, \quad \langle \xi - \hat{\xi}_2,\ \eta - \hat{\xi}_2 \rangle = 0$$

para todos los $\eta \in L$, entonces, en particular

$$\langle \xi - \hat{\xi}_1,\ \hat{\xi}_2 - \hat{\xi}_1 \rangle = 0,\ \langle \xi - \hat{\xi}_2,\ \hat{\xi}_2 - \hat{\xi}_1 \rangle = 0,$$

y tomando la diferencia de estas expresiones, obtenemos

$$\| \hat{\xi}_2 - \hat{\xi}_1 \| = 0, \text{ es decir, } \hat{\xi}_2 = \hat{\xi}_1.$$

Anteriormente hemos recordado como hallar la aproximación óptima (lineal) de la magnitud ξ con las combinaciones lineales $\sum_{k=1}^{n} c_k \eta_k$ de ciertas magnitudes dadas, η_k, $k = 1, \ldots, n$ (véase (1.1) — (1.5)).

¿Existe la aproximación óptima absoluta $\hat{\xi} = \varphi_0(\eta_1, \ldots, \eta_n)$ que satisfaga la condición (1.6) en la cual se toma el mínimo por t o d a s l a s m a g n i t u d e s p o s i b l e s $\eta = \varphi(\eta_1, \ldots, \eta_n)$, que son funciones de los valores dados $\eta_1, \ldots, \eta_n$?

Teorema. *La aproximación óptima absoluta de la magnitud ξ por las magnitudes de la forma $\eta = \varphi(\eta_1, \ldots, \eta_n)$ se da por la fórmula*

$$\hat{\xi} = \mathbf{M}(\xi \mid \eta_1, \ldots, \eta_n), \tag{1.8}$$

donde $\mathbf{M}(\xi \mid \eta_1, \ldots, \eta_n)$ significa la esperanza matemática condicional de la magnitud ξ para los valores fijados $\eta_1, \ldots, \eta_n$.

D e m o s t r a c i ó n. Evidentemente, el conjunto L de todas las magnitudes de la forma $\eta = \varphi(\eta_1, \ldots, \eta_n)$ para los cuales $\|\eta\|^2 = \mathbf{M}\eta^2 < \infty$ es un espacio lineal. Ya que $\mathbf{M}\xi^2 < \infty$, entonces a este espacio le pertenece también la magnitud $\hat{\xi}$ indicada en (1.8), que es una función determinada de los valores $\eta_1, \ldots, \eta_n$:

$$\hat{\xi} = \varphi_0(\eta_1, \ldots, \eta_n) \in L,$$

$$\hat{\xi}^2 = [\mathbf{M}(\xi|\eta_1, \ldots \eta_n)]^2 \leqslant \mathbf{M}(\xi^2 \mid \eta_1, \ldots, \eta_n),$$

y

$$\mathbf{M}\,\hat{\xi}^2 \leqslant \mathbf{M}[\mathbf{M}(\xi^2 \mid \eta_1, \ldots, \eta_n)] = \mathbf{M}\xi^2 < \infty.$$

Para la demostración de que la magnitud $\hat{\xi} = \varphi_0(\eta_1, \ldots, \eta_n)$ da la aproximación óptima absoluta, es suficiente comprobar la condición correspondiente (1.7) que en nuestro caso significa que

$$\mathbf{M}[(\xi - \hat{\xi})\cdot\varphi(\eta_1, \ldots, \eta_n)] = 0$$

para cualquier magnitud $\varphi(\eta_1, \ldots, \eta_n) \in L$. Sin embargo, para los valores fijados $\eta_1, \ldots, \eta_n$

$$\mathbf{M}[(\xi - \hat{\xi})\cdot\varphi(\eta_1, \ldots, \eta_n) \mid \eta_1, \ldots, \eta_n] = \\ = \varphi(\eta_1, \ldots, \eta_n)[\mathbf{M}(\xi \mid \eta_1, \ldots, \eta_n) - \hat{\varepsilon}] = 0,$$

y por consiguiente,

$$\mathbf{M}\{(\xi - \hat{\xi})\cdot\varphi(\eta_1, \ldots, \eta_n)\} = \\ = \mathbf{M}\{\mathbf{M}[(\xi - \hat{\xi})\cdot\varphi(\eta_1, \ldots, \eta_n) \mid \eta_1, \ldots, \eta_n]\} = 0.$$

El teorema queda demostrado.

Ejemplo (*tarea sobre un complejo de aparatos*). Imaginémonos que existen n aparatos destinados a la medición de una magnitud desconocida a, con ello cada aparato admite (en dependencia del caso) un error en las mediciones indicando correspondientemente $X_k = a + \Delta_k$ en lugar del valor verdadero a ($k = 0, \ldots, n - 1$). Surge la tarea de «unir» estos n aparatos en un complejo de tal forma que las indicaciones del «aparato en complejo» $X = a + \Delta$ den el error mínimo: $\|\Delta\| = \text{mín}$.

Está claro, que durante la elaboración de los errores de medición $\Delta_1, \ldots, \Delta_n$ al principio es necesario excluir la propia magnitud a desconocida de los datos existentes $X_1, \ldots, X_n$. Para esto se puede pasar a las diferencias

$$X_k - X_0 = \Delta_k - \Delta_0 = \xi_k, \quad k = 1, \ldots, n-1,$$

intentando luego hallar la aproximación óptima para el error desconocido Δ_0 $(= \xi_0)$ por las *magnitudes conocidas* $\xi_1, \ldots, \xi_{n-1}$.

Teniendo tal aproximación óptima

$$\hat{\Delta}_0 = \varphi_0(\xi_1, \ldots, \xi_{n-1}) = \varphi_0(X_1 - X_0, \ldots, X_{n-1} - X_0),$$

se puede realizar «el aparato en complejo» del siguiente modo:

$$X = X_0 - \varphi_0(X_1 - X_0, \ldots, X_{n-1} - X_0)$$

en el cual el error $\Delta = X - a$ será

$$\Delta = \Delta_0 - \varphi_0(\xi_1, \ldots, \xi_{n-1}).$$

Es natural considerar que los errores $\Delta_0, \Delta_1, \ldots, \Delta_{n-1}$ de los distintos aparatos son magnitudes aleatorias *independientes*. Supondremos que éstas tienen una misma distribución de probabilidades (con el valor medio nulo). Entonces la matriz de correlación $\{\mathbf{R}_{kj}\}$ de las magnitudes $\xi_0, \xi_1, \ldots, \xi_{n-1}$ tiene la forma siguiente:

$$R_{00} = \mathbf{M}\Delta_0^2 = \sigma^2, \quad R_{k0} = -\sigma^2, \quad R_{kk} = 2\sigma^2,$$
$$k = 1, \ldots, n-1.$$
$$R_{kj} = \sigma^2, \; k \neq j; \; k, j = 1, \ldots, n-1.$$

De las ecuaciones (1.5) obtenemos que la aproximación lineal óptima es

$$-\frac{1}{n}\sum_{k=1}^{n-1} \xi_k = \frac{1}{n}\sum_{k=1}^{n-1} [X_0 - X_k],$$

y el error correspondiente del «aparato en complejo» será

$$\Delta = \Delta_0 - \frac{1}{n}\sum_{k=1}^{n-1} [\Delta_0 - \Delta_k] = \frac{1}{n}\sum_{k=1}^{n} \Delta_k.$$

El "aparato en complejo" se realiza del modo siguiente:

$$X = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

Supongamos ahora que el error de cada aparato por separado tiene *una distribución uniforme de probabilidades* (en un segmento $[-d, d]$).

Hallemos la aproximación óptima absoluta

$$\varphi_0(\xi_1, \ldots, \xi_{n-1}) = \mathbf{M}(\Delta_0 \mid \xi_1, \ldots, \xi_{n-1}).$$

La densidad conjunta de probabilidad para las magnitudes iniciales $\Delta_0, \ldots, \Delta_{n-1}$ es

$$p_{\Delta_0, \ldots, \Delta_{n-1}}(y_0, \ldots, y_{n-1}) = \begin{cases} \frac{1}{(2d)^n} & \text{para } -d \leqslant y_k \leqslant d, \ k=0, \ldots, n-1, \\ 0 & \text{en caso contrario.} \end{cases}$$

Por consiguiente, la densidad conjunta de probabilidad para las magnitudes $\xi_0, \ldots, \xi_{n-1}$, ligadas con $\Delta_0, \ldots, \Delta_{n-1}$ por la transformación $\xi_0 = \Delta_0$, $\xi_k = \Delta_k - \Delta_0$, $k = 1, \ldots, n-1$, será

$$p_{\xi_0, \ldots, \xi_{n-1}}(x_0, \ldots, x_{n-1}) = \begin{cases} \frac{1}{(2d)^n} & \text{para } \alpha \leqslant x_0 \leqslant \beta, \\ 0 & \text{en caso contrario,} \end{cases}$$

donde

$$\alpha = \text{máx}\,(-d, -d-x_1, \ldots, -d, -x_{n-1}),$$
$$\beta = \text{mín}\,(d, d-x_1, \ldots, d-x_{n-1}),$$

ya que el jacobiano de la transformación indicada es igual a 1, y la desigualdad $\alpha \leqslant x_0 \leqslant \beta$ significa exactamente que los valores correspondientes $y_0, \ldots, y_{n-1}$ se encuentran entre los límites $-d \leqslant y_k \leqslant \leqslant d$, $k = 0, \ldots, n-1$. Luego, la densidad de probabilidad de las magnitudes $\xi_1, \ldots, \xi_{n-1}$ es

$$p_{\xi_1, \ldots, \xi_{n-1}}(x_1, \ldots, x_{n-1}) = \int_{-\infty}^{\infty} p_{\xi_0\xi_1, \ldots \xi_{n-1}}(x_0, x_1, \ldots, x_{n-1})\,dx_0 = \begin{cases} \frac{\beta-\alpha}{(2d)^n} & \text{para } \alpha \leqslant \beta, \\ 0 & \text{para } \alpha > \beta, \end{cases}$$

de modo que para la densidad condicional de probabilidad $p_{\xi_0}(x_0 \mid x_1, \ldots, x_{n-1})$ obtenemos la expresión

$$p_{\xi_0}(x_0 \mid x_1, \ldots, x_{n-1}) = \begin{cases} \frac{1}{\beta-\alpha} & \text{para } \alpha \leqslant \beta \\ 0 & \text{para } \alpha > \beta. \end{cases}$$

La función buscada $\varphi_0(x_1, \ldots, x_{n-1})$ es

$$\varphi_0(x_1, \ldots, x_{n-1}) = \int_{-\infty}^{\infty} x_0 p_{\xi_0}(x_0 \mid x_1, \ldots, x_{n-1})\,dx_0 = \frac{\alpha+\beta}{2}.$$

No es difícil ver que el «aparato en complejo» óptimo se realiza del modo siguiente:

$$X = \frac{X_* + X^*}{2},$$

donde X_* y X^* significan los valores mínimo y máximo en las indicaciones de diferentes aparatos $X_0, \ldots, X_{n-1}$ (*X es la semisuma de los valores extremos*).

En efecto, como se ve fácilmente,

$$\alpha = -d - (X_* - X_0), \quad \beta = d - (X^* - X_0)$$

y

$$X = X_0 - \varphi_0(\xi_1, \ldots, \xi_{n-1}) = X_0 - \frac{\alpha+\beta}{2} = \frac{X_* + X^*}{2},$$

donde

$$X_* = \text{mín}\,(X_0, \ldots, X_{n-1}), \quad X^* = \text{máx}\,(X_0, \ldots, X_{n-1}).$$

Es interesante comparar los errores correspondientes

$$\delta_0 = \left\| \frac{1}{n} \sum_{k=0}^{n-1} \Delta_k \right\| \quad \text{y} \quad \delta = \left\| \frac{1}{2}(\Delta_* + \Delta^*) \right\|$$

para las aproximaciones óptimas lineales y absolutas; aquí

$$\Delta_* = \text{mín}\,(\Delta_1, \ldots, \Delta_n), \quad \Delta^* = \text{máx}\,(\Delta_1, \ldots, \Delta_n).$$

Suponiendo para concretar que $d = 1/2$, tenemos

$$\delta_0^2 = \frac{1}{n} \int_{-1/2}^{1/2} x^2\,dx = \frac{1}{12n}$$

Luego, ya que

$$\Delta_* = \Delta_i, \ \Delta^* = \Delta_j \ \text{para} \ \Delta_i \leqslant \Delta_k \leqslant \Delta_j \quad (k = 0, \ldots n-1),$$

entonces

$$\delta^2 = \frac{1}{4} \mathbf{M} (\Delta_* + \Delta^*)^2 =$$

$$= \frac{1}{4} \sum_{i,\,j\,(i \neq j)} \int_{-1/2}^{1/2} \int_{y_i}^{1/2} (y_i + y_j)^2 \underbrace{\int_{y_i}^{y_j} \ldots \int_{y_i}^{y_j}}_{n-2 \text{ veces}} dy_0 \ldots dy_n =$$

$$= \frac{1}{4} n(n-1) \int_{-1/2}^{1/2} \int_{y_0}^{1/2} \int_{y_0}^{y_1} \ldots \int_{y_0}^{y_1} (y_0 + y_1)^2\,dy_0 \ldots dy_{n-1} = \frac{1}{2(n+1)(n+2)}$$

Vemos que para valores grandes de n la aproximación óptima absoluta en nuestro caso tiene más preponderancia que la aproximación óptima lineal.

§. 2 PRONOSTICACION Y FILTRADO DE LOS PROCESOS ALEATORIOS ESTACIONARIOS

1. Tarea de pronosticación lineal. Sea $\xi(t)$ un proceso aleatorio estacionario real en un ámplio sentido que admite su representación espectral

$$\xi(t)=\int_{-\infty}^{\infty} e^{i\lambda t}\,d\Phi(\lambda) \tag{2.1}$$

con la densidad espectral $f(\lambda)$. Designemos por $H_{(-\infty,\,t)}$ el conjunto de todas las magnitudes

$$\eta=\int_{-\infty}^{\infty} \varphi(\lambda)\,d\Phi(\lambda), \tag{2.2}$$

donde $\varphi(\lambda)$ entra en el espacio correspondiente $L_{(-\infty,\,t)}(f)$ de todas las funciones $\varphi(\lambda)$, $\int_{-\infty}^{\infty} |\varphi(\lambda)|^2 f(\lambda)\,d\lambda$, que bien tienen la forma

$$\varphi(\lambda)=\sum_{k=1}^{n} c_k e^{i\lambda t_k}$$

(donde $c_1, \ldots, c_n$ son coeficientes reales y $t_1, \ldots, t_n \leqslant t$), o bien son los límites de las funciones de tal forma (en otras palabras, $H_{(-\infty,\,t)}$ es el espacio de las magnitudes, que bien son combinaciones lineales $\eta=\sum_{k=1}^{n} c_k\xi(t_k)$ de los valores del proceso estacionario examinado en los momentos $t_1 \ldots t_n \leqslant t$, o bien son los límites en la media cuadrática de tales combinaciones lineales — véase el punto 2 del § 7).

La tarea de pronosticación lineal consiste en hallar la aproximación óptima del valor «futuro» $\xi(t+\tau)$, $\tau>0$, según las magnitudes $\eta \in H_{(-\infty,\,t)}$.

Demos la solución de esta tarea en el caso en que la densidad espectral es una función racional de λ y es representable en la forma

$$f(\lambda)=\frac{|P(i\lambda)|^2}{|Q(i\lambda)|^2}, \quad -\infty<\lambda<\infty, \tag{2.3}$$

(forma general de la densidad espectral racional) donde los polinomios $P(z)$ y $Q(z)$ tienen coeficientes reales, todas sus raíces se hallan en el semiplano izquierdo $\operatorname{Re} z<0$ de la variable compleja z y el polinomio $Q(z)$ no tiene raíces múltiples:

$$Q(z) = q_0 (z - q_1) \ldots (z - q_n). \tag{2.4}$$

Supongamos de momento, que $t = 0$. Abreviemos nuestras designaciones poniendo $H_0 = H_{(-\infty,\, 0)}$ y $L_0 = L_{(-\infty,\, 0)}(f)$. Buscaremos la función $\varphi_0(i\lambda) \in L_0$, con ayuda de la cual la aproximación óptima $\hat{\xi}$ para el valor futuro $\xi(\tau)$ puede ser dada como

$$\hat{\xi} = \int_{-\infty}^{\infty} \varphi_0(i\lambda)\, d\Phi(\lambda). \tag{2.5}$$

La condición general (1.7)

$$\langle \xi(\tau) - \hat{\xi},\ \eta \rangle = 0, \quad \eta \in L, \tag{2.6}$$

que determina unívocamente la aproximación óptima $\hat{\xi}$, en nuestro caso es equivalente a la condición

$$\langle \xi(\tau) - \hat{\xi},\ \xi(t) \rangle = 0 \quad \text{para } t \leqslant 0 \tag{2.7}$$

ya que de la igualdad (2.7) se deduce la igualdad (2.6) para todas las combinaciones lineales posibles $\eta = \sum_k c_k \xi(t_k)$, de las cuales se puede pasar a las magnitudes arbitrarias $\eta \in L$ por un paso límite.

La condición significa analíticamente (2.7) que

$$\int_{-\infty}^{\infty} e^{-i\lambda t} [e^{i\lambda\tau} - \varphi_0(i\lambda)]\, f(\lambda)\, d\lambda = 0 \quad \text{para } t \leqslant 0, \tag{2.8}$$

es decir, que la *transformación de Fourier de la función*

$$\psi(\lambda) = [e^{i\lambda\tau} - \varphi_0(i\lambda)]\, f(\lambda)] \tag{2.9}$$

se convierte en cero en el semieje negativo $t \leqslant 0$.

Con objeto de hallar la función $\varphi_0(i\lambda)$ de la condición (2.8) nos hacen falta las siguientes proposiciones.

Lema 1. *Si la función* $\psi(\lambda)$ *es analítica en el semiplano superior de la variable compleja* $\operatorname{Im}\lambda \geqslant 0$ *y decrece allí para* $\lambda \to \infty$ *de modo que* $|\psi(\lambda)| \leqslant C|\lambda|^{-1-\varepsilon}$, *donde* $\varepsilon > 0$, *entonces*

$$\int_{-\infty}^{\infty} e^{-i\lambda t} \psi(\lambda)\, d\lambda = 0 \ \text{ para } t \leqslant 0 \tag{2.10}$$

D e m o s t r a c i ó n. Para la función analítica $e^{i\lambda t}\psi(\lambda)$, $t \leqslant 0$, en el semiplano superior, la integral $\oint e^{i\lambda t}\psi(\lambda)\, d\lambda$ por el contorno limitado con el segmento $[-R, R]$ del eje real y con la semicircunferencia Γ de radio R (fig. 25), según el teorema de Cauchy es igual a cero.

Evidentemente, para $R \to \infty$

$$\left| \oint_{\Gamma} e^{i\lambda t} \psi(\lambda)\, d\lambda \right| \leqslant \max_{|\lambda|=R} |\psi(\lambda)| \pi R \leqslant C |R|^{-\varepsilon} \to 0,$$

y por esto para cualquier $t \leqslant 0$

$$\int_{-\infty}^{\infty} e^{i\lambda t} \psi(\lambda)\, d\lambda = \lim_{R\to\infty} \int_{-R}^{R} e^{i\lambda t} \psi(\lambda)\, d\lambda = \lim_{R\to\infty} \oint e^{i\lambda t} \psi(\lambda)\, d\lambda = 0,$$

que es lo que se exigía demostrar.

Lema 2. *La función racional*

$$\varphi_0(i\lambda) = \frac{\sum\limits_{k=0}^{n-1} a_k (i\lambda)^k}{P(i\lambda)}, \tag{2.11}$$

en la que todas las raíces del denominador, o sea, del polinomio $P(z)$, *de grado* m, *descansan en el semiplano izquierdo de la variable compleja* $\operatorname{Re} z < 0$, *representable en la forma*

$$\varphi_0(i\lambda) = \sum_{k=0}^{n-m-1} c_k (i\lambda)^k + \int_0^{\infty} e^{-i\lambda t} c(t)\, dt, \tag{2.12}$$

donde $c(t)$ *es una función integrable.*

D e m o s t r a c i ó n. Designemos por $\sum\limits_{k=0}^{n-m-1} c_k (i\lambda)^k$ polinomio que se obtiene de la división de $\sum\limits_{k=0}^{n} a_k (i\lambda)^k$ por $P(i\lambda)$ y sea $P_1(z)$ el resto correspondiente. Después, el quebrado $P_1(i\lambda)/P(i\lambda)$ se puede representar como una combinación lineal de "fracciones ordinarias" de la forma $1/(i\lambda - p)^k$ donde $p = \alpha + i\beta$, $\alpha < 0$, es la raíz del polinomio $P(z)$, y el grado k no supera la multiplicidad de esta raíz. Ya que

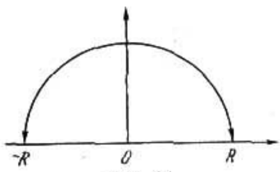

Fig. 25.

$$\frac{1}{(i\lambda - p)^k} = \int_0^{\infty} \frac{(-1)^k t^{k-1}}{(k-1)!} e^{-(i\lambda - p)t}\, dt,$$

se ve fácilmente, que cualquier combinación lineal de expresiones de la forma $1/(i\lambda - p)^k$ puede ser representada por la integral $\int_0^{\infty} e^{-i\lambda t} c(t)\, dt$ en la cual $c(t)$ es una combinación lineal de las funciones correspondientes de la forma

$$\frac{(-1)^k t^{k-1}}{(k-1)!} e^{pt} \quad 0 \leqslant t < \infty.$$

El lema queda demostrado.

Buscaremos una función $\varphi_0(i\lambda)$ en forma de la expresión (2.11) determinando los coeficientes a_k, $k=0, \ldots, n-1$ de modo que la función $\psi(\lambda)$, ligada con $\varphi_0(i\lambda)$ por la igualdad (2.9) sea analítica en el semiplano superior de la variable compleja $\operatorname{Im}\lambda \gg 0$.

Debido a que la densidad espectral $f(\lambda)=\frac{|P(i\lambda)|^2}{|Q(i\lambda)|^2}$ tiene de polos en el semiplano superior las raíces $-iq_1, \ldots, -iq_n$ del polinomio $Q(i\lambda)$, entonces la función

$$\psi(\lambda)=\frac{e^{i\lambda\tau}P(i\lambda)-\sum_{k=0}^{n-1} a_k (i\lambda)^k}{Q(i\lambda)}\,\frac{P(-i\lambda)}{Q(-i\lambda)}$$

para cualesquiera coeficientes a_k, $k=0, \ldots, n$, será analítica en el semiplano superior excluyendo puede ser los polos simples en los puntos $\lambda=-iq_1, \ldots, -iq_n$. Si determinamos los coeficientes a_k, $k=0, \ldots$ $\ldots, n-1$, del sistema de ecuaciones diferenciales

$$\sum_{k=0}^{n-1} a_k q_j^k = e^{q_j\tau} P(q_j), \qquad j=1, \ldots, n, \tag{2.13}$$

que, para los diferentes $q_1, \ldots, q_n$, tiene un determinante distinto de cero

$$\begin{vmatrix} 1 & q_1 & q_1^2 & \ldots & q_1^{n-1} \\ 1 & q_2 & q_2^2 & \ldots & q_2^{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & q_n & q_n^2 & \ldots & q_n^{n-1} \end{vmatrix},$$

entonces la función correspondiente $\psi(\lambda)$ será analítica en el semiplano superior. Ya que $|\psi(\lambda)|$ decrece para $|\lambda|\to\infty$ como $|\lambda|^{-n+m-1}$, entonces según el lema 1 la transformación de Fourier de la función $\psi(\lambda)$ se convierte en cero en el semieje negativo $t \leqslant 0$.

Evidentemente, la función $\varphi_0(i\lambda)$ determinada por la igualdad (2.11) satisface la condición $\int\limits_{-\infty}^{\infty} |\varphi_0(i\lambda)|^2 f(\lambda)\, d\lambda < \infty$. Según el lema 2, esta función está representada por la fórmula (2.12) de la cual se ve directamente, que $\varphi_0(i\lambda)$ pertenece al espacio L_0.

En efecto, la densidad espectral $f(\lambda)$ decrece para $|\lambda|\to\infty$ como $|\lambda|^{-2(n-m)}$, y las funciones determinadas sucesivamente por el paso límite

$$\varphi(\lambda)=\lim_{h\to -0}(i\lambda)^{k-1}\frac{e^{i\lambda(t+h)}-e^{i\lambda t}}{h}=(i\lambda)^k e^{i\lambda t}, \qquad k=0, \ldots, n-m-1,$$

todas pertenecen a L_0, cuando $t \leqslant 0$ (para $t = 0$ tenemos $\varphi(\lambda) = (i\lambda)^k$, $k = 0, \ldots, n-m-1$); también pertenecen al espacio L_0 las funciones del tipo

$$\varphi(\lambda) = \int_0^\infty e^{-i\lambda t} c(t)\, dt = \int_{-\infty}^0 e^{i\lambda t} c(-t)\, dt,$$

donde la función $c(t)$ es continua e integrable absolutamente, en la semirecta $0 \leqslant t < \infty$ (véase el punto 2 del § 7).

Así, pues, hemos hallado que la función $\varphi_0(i\lambda) \in L_0$ satisface la condición (2.8) y, *por consiguiente, la aproximación óptima para el valor* $\xi(\tau)$ con las magnitudes $\eta \in L_0$ es

$$\hat{\xi} = \int_{-\infty}^\infty \varphi_0(i\lambda)\, d\Phi(\lambda) = \sum_{k=0}^{n-m-1} c_k \xi^{(k)}(0) + \int_{-\infty}^0 c(-s)\, \xi(s)\, ds.$$

Se comprueba fácilmente que

$$\hat{\xi}(t) = \int_{-\infty}^\infty e^{i\lambda t} \varphi_0(i\lambda)\, d\Phi(\lambda) = \sum_{k=0}^{n-m-1} c_k \xi^{(k)}(t) + \int_{-\infty}^t c(t-s)\, \xi(s)\, ds \quad (2.14)$$

es la aproximación óptima del futuro valor $\xi(t+\tau)$ con las magnitudes $\eta \in H_{(-\infty, t)}$. En resumen hemos obtenido el siguiente resultado.

Teorema 1. *La aproximación óptima* $\xi(t+\tau)$ *con las magnitudes* $\eta \in H_{(-\infty, t)}$ *se da por la fórmula* (2.14) *en la cual la función* $\varphi_0(i\lambda)$ *está determinada por la igualdad* (2.11) *donde los coeficientes* $a_0, \ldots, a_{n-1}$ *satisfacen al sistema de ecuaciones lineales* (2.13).

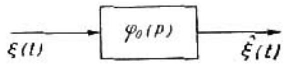

Fig. 26.

De tal modo el pronóstico lineal óptimo del proceso aleatorio estacionario $\xi(t)$, $-\infty < t < \infty$, en el tiempo anterior τ para el futuro puede realizarse por el sistema lineal con la función de transmisión

$$\varphi_0(p) = \sum_{k=0}^{n-1} a_k p^k / P(p).$$

2. Filtrado lineal (valuación del valor medio). Examinemos el proceso aleatorio $\xi = \xi(t)$ de la forma

$$\xi(t) = \theta(t) + \Delta(t), \quad (2.15)$$

donde $\theta = \theta(t)$ es una función de t (que en adelante se llama « señal ») y $\Delta = \Delta(t)$ es un proceso aleatorio con un valor medio nulo $\mathbf{M}\Delta(t) = 0$ (llamado en adelante «ruido aleatorio»). Examinemos la tarea sobre la separación de la señal $\theta = \theta(t)$ sobre el fondo del ruido aleatorio

$\Delta = \Delta(t)$, más exactamente, la tarea sobre la valuación de la función $\theta = \theta(t)$ según la trayectoria correspondiente del proceso aleatorio $\xi(t) = \theta(t) + \Delta(t)$, $t \in T$, en un segmento T, considerando que las señales posibles $\theta = \theta(t)$ tienen la forma

$$\theta(t) = \sum_{k=1}^{m} \alpha_k \theta_k(t), \tag{2.16}$$

donde $\theta_1(t), \ldots, \theta_m(t)$ son ciertas funciones dadas linealmente independientes, y los coeficientes $\alpha_1, \ldots, \alpha_m$ pueden tomar cualesquiera valores reales.

Para la solución de esta tarea valuemos los coeficientes desconocidos $\alpha_1, \ldots, \alpha_m$ en la expresión (2.16).

Designemos por $H(T)$ el conjunto de todas las combinaciones lineales $\eta = \sum_k c_k \xi(t_k)$ y sus límites medios cuadráticos

$$(\eta = \text{l.i.m.} \sum c_k \xi(t_k), \text{ donde } t_k \in T.$$

Denominemos a las magnitudes $\eta_1, \ldots, \eta_m \in H(T)$ *valuaciones no desplazadas linealmente* de los coeficientes $\alpha_1, \ldots, \alpha_m$ si los valores medios $\mathbf{M}\eta_k$, $k = 1, \ldots, m$, para cualesquiera valores posibles $\alpha_1, \ldots, \alpha_m$ satisfacen la condición

$$\mathbf{M}\eta_k = \alpha_k, \; k = 1, \ldots, m. \tag{2.17}$$

Ejemplo (*valuación de los cuadrados mínimos*).

Supongamos que la función desconocida $\theta = \theta(t)$ y la función aleatoria $\xi = \xi(t)$ considerada son integrables en el cuadrado sobre el segmento $T = [a, b]$. Para tales funciones cualesquiera $x_1 = x_1(t)$ y $x_2 = x_2(t)$ pongamos

$$(x_1, x_2) = \int_a^b x_1(t) x_2(t)\, dt.$$

Según la trayectoria existente $\xi(t)$, $a \leqslant t \leqslant b$, determinemos las magnitudes $\eta_1, \ldots, \eta_m$ de la condición

$$\int_a^b \Big[\xi(t) - \sum_{k=1}^{n} \eta_k \theta_k(t)\Big]^2 dt = \min_{y_1, \ldots, y_m} \int_a^b \Big[\xi(t) - \sum_{k=1}^{n} y_k \theta_k(t)\Big]^2 dt,$$

donde el número mínimo se toma por todos los valores reales posibles $y_1, \ldots, y_m$. Se trata del mínimo de la forma cuadrática no negativa

$$R(y_1, \ldots, y_m) = (\xi, \xi) - 2\sum_{k=1}^{m} y_k (\theta_k \xi) + \sum_{k,\, j=1}^{m} y_k y_j (\theta_k \theta_j)$$

de las variables $(y_1, \ldots, y_m)$. Este mínimo se alcanza en el punto $(\eta_1, \ldots, \eta_m)$ en el cual

$$\frac{\partial}{\partial y_k} R(\eta_1, \ldots, \eta_m) = -2\left[(\theta_k, \xi) + \sum_{j=1}^{m} \eta_j (\theta_k, \theta_j)\right] = 0, \quad k = 1, \ldots, m,$$

lo que nos da el sistema de ecuaciones lineales con relación a $\eta_1, \ldots, \eta_m$

$$\sum_{j=1}^{m} \eta_j (\theta_k, \theta_j) = (\xi, \theta_k), \quad k = 1, \ldots, m. \tag{2.18}$$

Para las funciones linealmente independientes $\theta_1(t), \ldots, \theta_m(t)$, la matriz $D = \{(\theta_k, \theta_j)\}$ de los coeficientes

$$(\theta_k \theta_j) = \int_a^b \theta_k(t)\, \theta_j(t)\, dt, \quad k, j = 1, \ldots, m,$$

del sistema de ecuaciones (2.18) es no degenerado y la solución de este sistema es

$$\eta_k = \sum_{j=1}^{m} \sigma_{kj} (\theta_j, \xi), \quad k = 1, \ldots, m, \tag{2.19}$$

donde σ_{kj} son los elementos de la matriz $\{\sigma_{kj}\}$ inversa nos da la matriz D.

Para cualquier función $\theta_j(t)$ la integral $(\theta_j, \xi) = \int_a^b \theta_j(t)\, \xi(t)\, dt$ es el límite medio cuadrático de las sumas integrales correspondientes, que representan en sí las combinaciones lineales de la forma $\sum_k c_k \xi(t_k)$. Por consiguiente, todas las magnitudes (θ_j, ξ) en la expresión (2.19) y junto con ellas, también las magnitudes $\eta_1, \ldots, \eta_m$ entran en el espacio lineal $H(T)$.

Además de esto, ya que el valor medio $\mathbf{M}\xi(t)$ es $\theta(t) = \sum_{k=1}^{m} \alpha_k \theta_k(t)$, tenemos que

$$\mathbf{M}(\theta_j, \xi) = \mathbf{M} \int_a^b \theta_j(t)\, \xi(t)\, dt = \int_a^b \theta_j(t)\, [\mathbf{M}\xi(t)]\, dt =$$

$$= \int_a^b \theta_j(t)\, \theta(t)\, dt = (\theta_j, \theta),$$

y se ve que los valores medios

$$\mathbf{M}\eta_k = \sum_{k=1}^{m} \sigma_{kj} (\theta_j, \theta), \quad k = 1, \ldots, m,$$

se obtienen de la fórmula (2.19) sustituyendo la función $\xi = \xi(t)$ por la función $\theta = \theta(t)$. Por consiguiente, estos valores dan el punto del mínimo de la expresión

$$\int_a^b \left|\theta(t) - \sum_{k=1}^m y_k\theta_k(t)\right|^2 dt \text{ y, ya que para } \theta(t) = \sum_{k=1}^m \alpha_k\theta_k(t)$$

el mínimo igual a cero, se alcanza para $y_k = \alpha_k$, $k = 1, \ldots, m$, de aquí deducimos que

$$\mathbf{M}\eta_k = \alpha_k, \quad k = 1, \ldots, m.$$

De tal modo, las magnitudes $\eta_1, \ldots, \eta_m$ determinadas por la fórmula (2.19) son las valuaciones lineales no desplazadas de los coeficientes desconocidos $\alpha_1, \ldots, \alpha_m$ (corrientemente se las llama *valuaciones de los cuadrados mínimos*).

Llamemos valuaciones lineales no desplazadas $\hat{\alpha}_1, \ldots, \hat{\alpha}_m$ de los coeficientes $\alpha_1, \ldots, \alpha_m$ a las *valuaciones óptimas*, si

$$\| \alpha_k - \hat{\alpha}_k \| = \min_{\eta \in L_k} \| \alpha_k - \eta \|, \quad k = 1, \ldots, m, \tag{2.20}$$

donde el mínimo se toma por todas las valuaciones lineales no desplazadas η para el coeficiente correspondiente α_k (cuyo conjunto lo designamos por L_k).

Por definición L_k es el hiperplano (en el subespacio lineal $H(T)$) de todas las magnitudes $\eta \in H(T)$ que satisfacen la condición $\mathbf{M}\eta = \alpha_k$. Por eso (véase el lema sobre la perpendicular, pág. 253), las valuaciones óptimas no desplazadas $\hat{\alpha}_k \in L_k$ se determinan unívocamente por la condición

$$\langle \alpha_k - \hat{\alpha}_k, \eta - \hat{\alpha}_k \rangle = 0 \text{ para todos los } \eta \in L_k, \quad k = 1, \ldots, m. \tag{2.21}$$

Intentemos describir más explícitamente el espacio $H(T)$ y los hiperplanos L_k, $k = 1, \ldots, m$ suponiendo que en la expresión (2.15) $\Delta = \Delta(t)$ es un proceso aleatorio estacionario con la densidad espectral $f(\lambda)$ que admite la representación espectral

$$\Delta(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\, d\Phi_0(\lambda), \tag{2.22}$$

y la función $\theta = \theta(t)$ es la transformación de Fourier de $\tilde{\theta} = \tilde{\theta}(\lambda)$:

$$\theta(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\tilde{\theta}(\lambda)\, d\lambda, \tag{2.23}$$

donde la función $\tilde{\theta}(\lambda)$ satisface la condición

$$\int_{-\infty}^{\infty} \frac{|\tilde{\theta}(\lambda)|^2}{f(\lambda)} d\lambda < \infty. \tag{2.24}$$

Las fórmulas (2.22), (2.23) nos permiten representar el proceso aleatorio examinado $\xi(t) = \theta(t) + \Delta(t)$ por la integral estocástica

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\lambda t} d\Phi(\lambda) \tag{2.25}$$

(compárese con la representación espectral (7.4), cap. III), donde

$$\Phi(\lambda) = \int_{-\infty}^{\lambda} \tilde{\theta}(\lambda) d\lambda + \Phi_0(\lambda)$$

es un proceso aleatorio con los incrementos no correlacionados, que tienen el valor medio

$$\mathbf{M}\Phi(\lambda) = \int_{-\infty}^{\lambda} \tilde{\theta}(\lambda) d\lambda$$

y la densidad de estructura $f(\lambda)$.

Igualmente que para los procesos estacionarios, examinaremos el espacio lineal $H(T)$ de todas las magnitudes η, representadas por la integral estocástica

$$\eta = \int_{-\infty}^{\infty} \varphi(\lambda) d\Phi(\lambda), \tag{2.26}$$

donde las funciones $\varphi(\lambda)$ entran en el espacio correspondiente $L_T(f)$, con ello existe la integral $\int_{-\infty}^{\infty} \varphi(\lambda)\tilde{\theta}(\lambda) d\lambda$ (véase el punto 2, del § 6).

Hablando en general, la existencia de la integral $\int_{-\infty}^{\infty} \varphi(\lambda)\tilde{\theta}(\lambda) d\lambda$ es la condición complementaria para las funciones $\varphi(\lambda) \in L_T(f)$, ahora bien, para las suposiciones hechas anteriormente, ella se deduce automáticamente de la condición general $\int_{-\infty}^{\infty} |\varphi(\lambda)|^2 f(\lambda) d\lambda <$

$< \infty$ ya que

$$\int_{-\infty}^{\infty} |\varphi(\lambda)\,\tilde{\theta}(\lambda)|\,d\lambda = \int_{-\infty}^{\infty} \left|\varphi(\lambda)\sqrt{f(\lambda)}\,\frac{\tilde{\theta}(\lambda)}{\sqrt{f(\lambda)}}\right| d\lambda \leqslant$$
$$\leqslant \sqrt{\int_{-\infty}^{\infty} |\varphi(\lambda)|^2 f(\lambda)\,d\lambda}\,\sqrt{\int_{-\infty}^{\infty} \frac{|\tilde{\theta}(\lambda)|^2}{f(\lambda)}\,d\lambda} < \infty.$$

De tal modo, *el espacio lineal $H(T)$ está compuesto por todas las magnitudes*

$$\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\,d\Phi(\lambda), \quad \varphi(\lambda) \in L_T(f).$$

Recordemos que para las integrales estocásticas (2.26)

$$\mathbf{M}\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\,\tilde{\theta}(\lambda)\,d\lambda, \tag{2.27}$$

y

$$\langle \eta_1 - \mathbf{M}\eta_1,\ \eta_2 - \mathbf{M}\eta_2 \rangle = \langle \varphi_1,\ \varphi_2 \rangle, \tag{2.28}$$

donde φ_1, φ_2 son las funciones correspondientes del espacio $L_T(f)$ y

$$\langle \varphi_1,\ \varphi_2 \rangle = \int_{-\infty}^{\infty} \varphi_1(\lambda)\,\overline{\varphi_2(\lambda)}\,f(\lambda)\,d\lambda.$$

Recurramos a las funciones

$$\psi_1(\lambda),\ \ldots,\ \psi_m(\lambda) \in L_T(f),$$

que satisfacen las ecuaciones integrales[1])

$$\int_{-\infty}^{\infty} e^{-i\lambda t}\psi_k(\lambda)\,f(\lambda)\,d\lambda = \theta_k(t), \quad t \in T, \quad k = 1, \ldots, m. \tag{2.29}$$

Como se deduce de la fórmula (2.27) que determina el hiperplano L_k la condición $\mathbf{M}\eta = \alpha_k$ para las magnitudes $\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\,d\Phi(\lambda)$

---

[1]) Señalemos que tales funciones $\psi_1, \ldots, \psi_m$ existen siempre, teniendo para cada $k = 1, \ldots, m$ la única solución $\psi_k(\lambda)$ de la ecuación (2.29), que pertenece al espacio $L_T(f)$ (con relación a estas ecuaciones véase, por ejemplo, el libro de Yu. A. Rozanov, «Procesos aleatorios estacionarios», Fismatguiz, Moscú, 1963, pág. 190).

significa que

$$\int_{-\infty}^{\infty} \varphi(\lambda)\,\overline{\theta}(\lambda)\,d\lambda = \int_{-\infty}^{\infty} \varphi(\lambda)\left[\sum_{j=1}^{m} \alpha_j \overline{\psi_j(\lambda)}\right] f(\lambda)\,d\lambda = \sum_{j=1}^{m} \alpha_j \langle \varphi, \psi_j \rangle \equiv \alpha_k \tag{2.30}$$

para todos los $\alpha_1, \ldots, \alpha_m$, lo que, en vista de la arbitrariedad de estos coeficientes, es equivalente a la condición

$$\langle \varphi, \psi_j \rangle = \begin{cases} 1 \text{ para } j = k, \\ 0 \text{ para } j \neq k. \end{cases} \tag{2.31}$$

Hemos hallado que *cada uno de los hiperplanos* L_k *(de las valuaciones lineales no desplazadas* $\eta = \int_{-\infty}^{\infty} \varphi(\lambda)\,d\Phi(\lambda)$ *para los coeficientes incógnitos* α_k*) se determina por el sistema de ecuaciones (2.31) en el cual las funciones* $\psi_1(\lambda), \ldots, \psi_m(\lambda)$ *satisfacen las ecuaciones integrales* (2.29).

Volvamos a la condición (2.21) que determina la valuación óptima $\hat{\alpha}_k = \int_{-\infty}^{\infty} \varphi_{0k}(\lambda)\,d\Phi(\lambda)$ para cada $k = 1, \ldots, m$. De la fórmula general (2.28) se desprende que esta condición para la función correspondiente $\varphi_{0k}(\lambda) \in L_k$ significa lo siguiente:

$$\langle \varphi_{0k}, \varphi \dot{-} \varphi_{0k} \rangle = 0 \tag{2.32}$$

para todos los $\varphi(\lambda) \in L_T(f)$ que satisfacen las ecuaciones (2.31).

Introduzcamos la matriz $D = \{d_{kj}\}$ con los elementos

$$d_{kj} = \langle \psi_k, \psi_j \rangle = \int_{-\infty}^{\infty} \psi(\lambda)\,\overline{\psi_j(\lambda)}\,f(\lambda)\,d\lambda. \tag{2.33}$$

Ya que para las funciones linealmente independientes $\theta_1(t), \ldots, \theta_m(t)$ las funciones $\psi_1(\lambda), \ldots, \psi_m(\lambda)$ determinades por las ecuaciones (2.29), también son linealmente independientes, la matriz D no está degenerada y existe la matriz inversa $D^{-1} = \{\sigma_{kj}\}$:

$$\sum_{j=1}^{m} \sigma_{ij} d_{jk} = \begin{cases} 1 \text{ para } k = i, \\ 0 \text{ para } k \neq i. \end{cases}$$

Mostremos que las funciones desconocidas $\varphi_{0k}(\lambda)$ son

$$\varphi_{0k}(\lambda) = \sum_{j=1}^{m} \sigma_{kj} \psi_j(\lambda), \qquad k = 1, \ldots, m. \tag{2.34}$$

En efecto, ya que todas las funciones $\psi_j(\lambda)$, $j = 1, \ldots, m$, pertenecen al espacio lineal $L_T(f)$, también pertenecen a él las funciones

$\varphi_{0k}(\lambda)$, $k = 1, \ldots, m$. Cada una de éstas satisface la condición respectivo (2.31) a saber:

$$\langle \varphi_{0k}, \psi_j \rangle = \sum_{i=1}^{m} \sigma_{ki} \langle \psi_i, \psi_j \rangle = \sum_{i=1}^{m} \sigma_{ki} d_{ij} = \begin{cases} 1 & \text{para } j = k, \\ 0 & \text{para } j \neq k. \end{cases}$$

Además de esto, para cada $k = 1, \ldots, m$ tiene lugar la igualdad (2.32), ya que para cualquier función $\varphi(\lambda)$ que satisface la condición (2.31).

$$\langle \varphi_{0k}, \varphi \rangle = \sum_{j=1}^{m} \sigma_{kj} \langle \psi_j, \varphi \rangle = \sigma_{kk}$$

y al mismo tiempo

$$\langle \varphi_{0k}, \varphi_{0k} \rangle = \sum_{i,\, j=1}^{m} \sigma_{ki} d_{ij} \sigma_{jk} = \sigma_{kk}.$$

Señalemos que la matriz $D^{-1} = \{\sigma_{kj}\}$ es la matriz correlativa de los errores $\hat{\alpha}_k - \alpha_k$, $k = 1, \ldots, m$ en las valuaciones óptimas de los coeficientes $\alpha_1, \ldots, \alpha_m$, a saber:

$$\sigma_{kj} = \langle \varphi_{0k}, \varphi_{0j} \rangle = \mathbf{M}(\hat{\alpha}_k - \alpha_k)(\hat{\alpha}_j - \alpha_j),\ k,\ j = 1, \ldots, m. \tag{2.35}$$

Enunciemos el resultado que hemos obtenido.

Teorema 2. *Las valuaciones óptimas lineales $\hat{\alpha}_1, \ldots, \hat{\alpha}_m$ de los coeficientes desconocidos $\alpha_1, \ldots, \alpha_m$ se dan por la expresión*

$$\alpha_k = \int_{-\infty}^{\infty} \varphi_{0k}(\lambda)\, d\Phi(\lambda), \quad k = 1, \ldots, m,$$

donde las funciones $\varphi_{0k}(\lambda)$, $k = 1, \ldots, m$, están ligadas por las fórmulas (2.34) con la solución $\psi_k(\lambda) \in L_T(f)$, $k = 1, \ldots, m$ de las ecuaciones integrales (2.29).

Señalemos que algunas veces se puede valuar sin error[1]) los coeficientes desconocidos α_k, $k = 1, \ldots, m$.

Ejemplo (ley de los grandes números). Sea

$$\theta(t) = \alpha,\ 0 \leqslant t < \infty,$$

donde α es una constante desconocida. Entonces

$$\alpha = \underset{n \to \infty}{\text{l.i.m.}} \frac{1}{n} \int_0^n \xi(t)\, dt.$$

[1]) Así ocurrirá siempre al infringir la condición (2.24). Véase, por ejemplo, el libro de I. A. Ibraguimov, Yu. A. Rozanov «Procesos aleatorios de Gauss», Naúka, Moscú, 1970.

En efecto, $\frac{1}{n}\int_0^n \xi(t)\,dt-\alpha=\frac{1}{n}\int_0^n \Delta(t)\,dt$, donde para el proceso estacionario $\Delta(t)=\int_{-\infty}^{\infty} e^{i\lambda t}\,d\Phi_0(\lambda)$ tenemos

$$\frac{1}{n}\int_0^n \Delta(t)\,dt=\int_{-\infty}^{\infty}\left[\frac{1}{n}\int_0^n e^{i\lambda t}\,dt\right]d\Phi_0(\lambda)=\int_{-\infty}^{\infty}\frac{e^{i\lambda n}-1}{i\lambda n}\,d\Phi_0(\lambda),$$

y se ve fácilmente, que para $n\to\infty$

$$\left\|\frac{1}{n}\int_0^n \xi(t)\,dt-\alpha\right\|^2=\int_{-\infty}^{\infty}\left|\frac{e^{i\lambda n}-1}{i\lambda n}\right|^2 f(\lambda)\,d\lambda=\int_{-\infty}^{\infty}\frac{4\operatorname{sen}^2\lambda n}{\lambda^2 n^2}f(\lambda)\,d\lambda\to 0.$$

§ 3. ESPERANZAS MATEMATICAS CONDICIONALES Y ALGUNAS TAREAS DE PRONOSTICACIÓN Y DE FILTRADO

1. Una vez más sobre las esperanzas matemáticas condicionales. Examinemos una familia arbitraria de magnitudes aleatorias $\eta(t)$, $t\in T$, dependientes del parámetro t y la magnitud aleatoria ξ, que de algún modo está ligada con las magnitudes $\eta(t)$, $t\in T$.

Para mayor claridad nos podemos imaginar una prueba compleja en la que al principio se observan las magnitudes $\eta(t)$ $t\in T$, y luego, en dependencia de los valores $\eta(t)$, $t\in T$, se realiza una prueba complementaria, en la cual se observa la magnitud aleatoria ξ; en otras palabras, «el mecanismo de aleatoriedad» actúa sucesivamente: al principio, las magnitudes $\eta(t)$, $t\in T$, toman tales o cuales valores con cierta distribución de probabilidades y, luego, la magnitud ξ toma tal o cual valor con cierta distribución (condicional) de probabilidades, en dependencia de aquellas magnitudes.

En estas condiciones la esperanza matemática de la magnitud ξ para los valores fijados de $\eta(t)$, $t\in T$ (designémosla por $\mathbf{M}\,\{|\,\xi\mid\eta\})$ depende de las magnitudes aleatorias $\eta(t)$, $t\in T$, y en este sentido es una magnitud aleatoria. Parece intuitivamente que se puede determinar el valor medio $\mathbf{M}\xi$ tomando al principio $M\,\{\xi\mid\eta\}$ para los valores fijados $\eta(t)$, $t\in T$, y luego tomando la esperanza matemática $\mathbf{M}\,[M\{\xi\mid\eta\}]$:

$$\mathbf{M}\xi=\mathbf{M}\,[\mathbf{M}\{\xi\mid\eta\}]. \tag{3.1}$$

Además de esto, para cualquier magnitud aleatoria $\varphi(\eta)$ determinada unívocamente por los valores $\eta(t)$, $t \in T$, se deberá cumplir la igualdad

$$\mathbf{M}\{\varphi(\eta)\cdot\xi \mid \eta\} = \varphi(\eta)\cdot\mathbf{M}\{\xi \mid \eta\}, \tag{3.2}$$

ya que para los valores fijados $\eta(t)$, $t \in T$, la magnitud $\varphi(\eta)$ es constante.

Designemos por L el conjunto de tales magnitudes $\varphi(\eta)$ para las cuales $\mathbf{M}\varphi^2(\eta) < \infty$. Supongamos que para cada uno de los valores fijados $\eta(t)$, $t \in T$, se tiene la distribución condicional de probabilidades correspondiente, con ello la esperanza matemática condicional $\mathbf{M}\{\xi \mid \eta\}$ posee las propiedades (3.1) y (3.2).

Examinemos la magnitud ξ, $\mathbf{M}\xi^2 < \infty$. La magnitud correspondiente $\hat{\xi} = \mathbf{M}(\xi \mid \eta)$ entra en L, ya que según la desigualdad general para los momentos

$$\hat{\xi}^2 = [\mathbf{M}(\xi \mid \eta)]^2 \leqslant \mathbf{M}(\xi^2 \mid \eta),$$

donde de acuerdo con la fórmula (3.1) aplicada a la magnitud ξ^2,

$$\mathbf{M}[\mathbf{M}(\xi^2 \mid \eta)] = \mathbf{M}\xi^2$$

y, por consiguiente, $\mathbf{M}\hat{\xi}^2 \leqslant \mathbf{M}\xi^2 < \infty$. Según la fórmula (3.2) $\mathbf{M}\{\varphi(\eta)\cdot\xi \mid \eta\} = \varphi(\eta)\cdot\hat{\xi}$, de modo que la igualdad (3.1) para $\varphi(\eta)\,\xi$ da

$$\mathbf{M}[\mathbf{M}\{\varphi(\eta)\cdot\xi \mid \eta\}] = \mathbf{M}[\varphi(\eta)\cdot\xi] = \langle\varphi(\eta), \xi\rangle$$

y al mismo tiempo

$$\mathbf{M}[\mathbf{M}\{\varphi(\eta)\cdot\xi \mid \eta\}] = \mathbf{M}[\varphi(\eta)\cdot\hat{\xi}] = \langle\varphi(\eta), \hat{\xi}\rangle,$$

de donde obtenemos que

$$\langle\xi - \hat{\xi}, \varphi(\eta)\rangle = 0 \text{ para todos los valores } \varphi(\eta) \in L. \tag{3.3}$$

Pero antes, en el § 1 fue mostrado que existía, una sola magnitud $\hat{\xi} \in L$ que satisfacía la condición (3.3): $\hat{\xi} = \mathbf{M}\{\xi \mid \eta\}$ *esto significa geométricamente la proyección de la magnitud* ξ *sobre el espacio lineal* L *de todas las magnitudes* $\varphi(\eta)$, $\mathbf{M}\varphi^2(\eta) < \infty$ *que son funciones de los valores* $\eta(t)$, $t \in T$.

Partiendo de las esperanzas matemáticas condicionales, se puede determinar la distribución condicional de probabilidades ya que

$$P\{x' \leqslant \xi \leqslant x'' \mid \eta\} = \mathbf{M}\{\varphi(\xi) \mid \eta\}, \tag{3.4}$$

donde la función $\varphi(x)$ tiene la forma

$$\varphi(x) = \begin{cases} 1 \text{ para } x' \leqslant x \leqslant x'' \\ 0 \text{ para los restantes } x. \end{cases}$$

Ejemplo. Sean ξ y η magnitudes aleatorias discretas, que tienen una distribución conjunta de probabilidades $P_{\xi\eta}(x, y)$, $-\infty < x, y < \infty$. Hallemos la distribución condicional de probabilidades $P_{\xi}(x \mid \eta)$ de la magnitud ξ para el valor fijado de η.

Para la magnitud aleatoria $\varphi(\xi)$ que es función de ξ, la esperanza matemática con relación a la distribución de probabilidades $P_{\xi}(x \mid \eta)$ para el valor fijado de η es

$$\mathbf{M}\{\varphi(\xi) \mid \eta\} = \sum_{-\infty}^{\infty} \varphi(x) P_{\xi}(x \mid \eta).$$

Para la función limitada $\psi = \psi(y)$ tendremos

$$\mathbf{M}\{\varphi(\xi)\psi(\eta) \mid \eta\} = \sum_{-\infty}^{\infty} \varphi(x)\psi(\eta) P_{\xi}(x \mid \eta).$$

Tomando las esperanzas matemáticas de la parte izquierda y derecha, por una parte obtenemos

$$\mathbf{M}[\mathbf{M}\{\varphi(\xi)\psi(\eta) \mid \eta\}] = \mathbf{M}\varphi(\xi)\psi(\eta) = \sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} \varphi(x)\psi(y) P_{\xi\eta}(x, y),$$

y de otra parte,

$$\mathbf{M}[\mathbf{M}\{\varphi(\xi)\psi(\eta) \mid \eta\}] = \mathbf{M}\left[\sum_{-\infty}^{\infty} \varphi(x)\psi(\eta) P_{\xi}(x \mid \eta)\right] =$$

$$= \sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} \varphi(x)\psi(y) P_{\xi}(x \mid y) P_{\eta}(y)$$

(donde P_{η}, (y) es la distribución de la magnitud η), lo que da la siguiente igualdad:

$$\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} \varphi(x)\psi(y) P_{\xi}(x \mid y) P_{\eta}(y) = \sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} \varphi(x)\psi(y) P_{\xi\eta}(x, y).$$

De aquí, suponiendo que

$$\varphi(x) = \begin{cases} 1 \text{ para } x = x_0, \\ 0 \text{ para } x \neq x_0; \end{cases}$$

$$\psi(y) = \begin{cases} 1 \text{ para } y = y_0, \\ 0 \text{ para } y \neq y_0 \end{cases}$$

para cualesquiera x_0, y_0, obtenemos, que

$$P_{\xi}(x_0 \mid y_0) P_{\eta}(y_0) = P_{\xi\eta}(x_0, y_0).$$

De tal modo, para cualesquiera valores de y, que toma la magnitud η con probabilidad positiva.

$$P_{\xi}(x \mid y) = \frac{P_{\xi\eta}(x, y)}{P_{\eta}(y)} \tag{3.5}$$

(compárese con (3.16), cap. I).

Ejemplo. Sea η una magnitud aleatoria discreta con distribución de probabilidades $P_\eta(y)$. Supongamos que la magnitud ξ para cada valor fijado η tiene la densidad condicional de probabilidades $p_\xi(x \mid \eta)$. Para

$$\varphi(x) = \begin{cases} 1 \text{ para } x' \leqslant x \leqslant x'', \\ 0 \text{ para } x < x', \ x > x''; \end{cases}$$

$$\psi(y) = \begin{cases} 1 \text{ para } y = y_0, \\ 0 \text{ para } y \neq y_0 \end{cases}$$

tenemos

$$\mathbf{M}\left[\mathbf{M}\{\psi(\eta)\varphi(\xi) \mid \eta\}\right] = \mathbf{M}\left[\psi(\eta)\int_{x'}^{x''} p_\xi(x \mid \eta)\,dx\right] =$$

$$= \sum_y \left[\psi(y)\int_{x'}^{x''} p_\xi(x \mid y)\,dx\right] P_\eta(y) = \int_{x'}^{x''} p_\xi(x \mid y_0)\,dx \cdot P_\eta(y_0).$$

Ahora bien, para cualquier y_0 la fórmula (3.1) aplicada a la magnitud $\psi(\eta)\varphi(\xi)$ nos da a la vez:

$$\mathbf{M}\left[\mathbf{M}\{\psi(\eta)\varphi(\xi) \mid \eta\}\right] = \mathbf{M}\psi(\eta)\varphi(\xi) = \\ = \mathbf{P}\{x' \leqslant \xi \leqslant x'', \ \eta = y_0\},$$

de modo que en total obtenemos

$$P\{x' \leqslant \xi \leqslant x'', \eta = y\} = \int_{x'}^{x''} p_\xi(x \mid y)\,dx P_\eta(y) \tag{3.6}$$

para todas los x', x'' e y.

Ejemplo. Sea ξ una magnitud discreta con distribución de probabilidades $P_\xi(x)$, $-\infty < x < \infty$ y la magnitud aleatoria η para cada valor fijado de ξ tiene la densidad de probabilidad $p_\eta(y \mid \xi)$, $-\infty < y < \infty$. Hallemos la distribución conjunta de probabilidades de estas magnitudes y la distribución condicional $P_\xi(x \mid \eta)$ de la magnitud ξ respecto a η.

La probabilidad del suceso $\{y' \leqslant \eta \leqslant y''\}$ para la condición $\xi = x$ es

$$P\{y' \leqslant \eta \leqslant y'' \mid \xi = x\} = \int_{y'}^{y''} p_\eta(y \mid x)\,dy.$$

Según la fórmula (3.6)

$$P\{\xi = x, \ y' \leqslant \eta \leqslant y''\} = P_\xi(x)\int_{y'}^{y''} p_\eta(y \mid x)\,dy,$$

y por, consiguiente, la distribución conjunta se da como

$$\mathbf{P}\{x' \leqslant \xi \leqslant x'', y' \leqslant \eta \leqslant y''\} = \int_{y'}^{y''} \left[\sum_{x'}^{x''} P_\xi(x)\, p_\eta(y \mid x)\right] dy.$$

De aquí se deduce, en particular, que

$$\mathbf{P}\{y' \leqslant \eta \leqslant y''\} = \int_{y'}^{y''} \left[\sum_{-\infty}^{\infty} P_\xi(x)\, p_\eta(y \mid x)\right] dy,$$

y se ve que la magnitud aleatoria η tiene la densidad de probabilidad

$$p_\eta(y) = \sum_{-\infty}^{\infty} p_\eta(y \mid x)\, P_\xi(x).$$

Mostremos que la distribución condicional $P_\xi(x \mid y)$ de la magnitud aleatoria ξ para el valor fijo de $\eta = y$ se da por la llamada *fórmula de Bayes*:

$$P_\xi(x \mid y) = P_\xi(x)\, \frac{p_\eta(y \mid x)}{p_\eta(y)}. \tag{3.7}$$

En realidad, la fórmula (3.7) da la distribución de probabilidades, con relación a la cual el valor medio

$$\mathbf{M}\{\varphi(\xi)\, \eta\} = \sum_{-\infty}^{\infty} \varphi(x)\, P_\xi(x \mid \eta)$$

de las magnitudes de la forma $\varphi(\xi)$ satisface la condición (3.1):

$$\mathbf{M}[\mathbf{M}\{\varphi(\xi) \mid \eta\}] = \int_{-\infty}^{\infty} \mathbf{M}\{\varphi(\xi) \mid y\}\, p_\eta(y)\, dy =$$

$$= \sum_{-\infty}^{\infty} \varphi(x)\, P_\xi(x) \left[\int_{-\infty}^{\infty} p_\eta(y \mid x)\, dy\right] = \sum_{-\infty}^{\infty} \varphi(x)\, P_\xi(x) = \mathbf{M}\varphi(\xi),$$

y también la condición (3.2) y, como ya sabemos, esta condición determina unívocamente la magnitud $\mathbf{M}\{\varphi(\xi) \mid \eta\}$; en su conjunto (para distintos $\varphi(\xi)$ los valores medios $\mathbf{M}\{\varphi(\xi) | \eta\}$ determinan unívocamente la distribución de probabilidades correspondiente $P_\xi(x \mid \eta)$ (digamos, $\mathbf{M}\{\varphi(\xi) | \eta\} = \mathbf{P}_\xi(x_0 \mid \eta)$ para la función $\varphi(x)$ que es igual a 1 para $x = x_0$ e igual a cero para los restantes x).

Ejemplo. Sean ξ y η magnitudes aleatorias que tienen la densidad conjunta de probabilidad $p_{\xi\eta}(x, y)$.

Mostremos que la densidad condicional $p_\xi(x \mid \eta)$ de la magnitud ξ con relación a η tiene la forma

$$p_\xi(x \mid y) = \frac{p_{\xi\eta}(x, y)}{p_\eta(y)}, \quad -\infty < x < \infty, \tag{3.8}$$

donde

$$p_\eta(y) = \int_{-\infty}^{\infty} p_{\xi\eta}(x, y)\,dx, \quad -\infty < y < \infty,$$

es la densidad de probabilidad de la magnitud η (compárese con (3.17), cap. I).

En efecto, la esperanza matemática de la magnitud del tipo $\varphi(\xi)$ con respecto a la distribución con la densidad de probabilidades indicada $p_\xi(x \mid \eta)$:

$$\mathbf{M}\{\varphi(\xi) \mid \eta\} = \int_{-\infty}^{\infty} \varphi(x)\, p_\xi(x \mid \eta)\,dx,$$

satisface la condición (3.1) a saber:

$$\mathbf{M}[\mathbf{M}\{\varphi(\xi) \mid \eta\}] = \int_{-\infty}^{\infty} [\varphi(x)\, p_\xi(x \mid y)\,dx]\, p_\eta(y)\,dy =$$

$$= \int_{-\infty}^{\infty} \varphi(x) \left[\int_{-\infty}^{\infty} p_{\xi\eta}(x, y)\,dy\right] dx = \int_{-\infty}^{\infty} \varphi(x)\, p_\xi(x)\,dx = \mathbf{M}\varphi(\xi),$$

de tal modo que $\mathbf{M}\{\varphi(\xi) \mid \eta\}$ da realmente la esperanza matemática condicional de las magnitudes $\varphi(\xi)$ con relación a η. Pero en su conjunto (para distintos $\varphi(\xi)$ los valores medios $\mathbf{M}\{\varphi(\xi) \mid \eta\}$ determinan unívocamente la distribución de probabilidades, en particular, según la fórmula (3.4).

$$\mathbf{P}\{x' \leqslant \xi \leqslant x'' \mid \eta\} = \int_{x'}^{x''} p_\xi(x \mid \eta)\,dx,$$

y de tal modo, la función $p_\xi(x \mid \eta)$, $-\infty < x < \infty$ dada por la igualdad (3.8) es la densidad condicional de probabilidad de la magnitud ξ para el valor fijo η.

Señalemos que la densidad condicional $p_\eta(y \mid \xi)$ se determina análogamente a (3.8), con ello tiene lugar la siguiente igualdad (compárese con la fórmula de Bayes (3.7)

$$p_\xi(x \mid y) = \frac{p_\eta(y \mid x)}{p_\eta(y)}\, p_\xi(x). \tag{3.9}$$

2. Papel de las probabilidades a posteriori en algunas tareas de pronosticación y filtrado.

Como hemos visto, las propiedades fundamentales de las esperanzas matemáticas condicionales expresadas por las relaciones (3.1) y (3.2) permiten dar la siguiente interpretación geométrica del valor medio condicional $\hat{\xi} = \mathbf{M}(\xi \mid \eta)$ de la magnitud ξ, $\mathbf{M}\xi^2 < \infty$, en relación a las magnitudes $\eta(t)$, $t \in T$: $\hat{\xi} = \mathbf{M}(\xi \mid \eta)$ *es la proyección de* ξ *en el espacio* L *de todas las magnitudes de la forma* $\varphi(\eta)$ *que son funciones de los valores* $\eta(t)$, $t \in T$. Con esto *la magnitud* $\hat{\xi}$ *da la aproximación óptima de* ξ con las magnitudes $\varphi(\eta) \in L$:

$$\|\xi - \hat{\xi}\| = \min_{\varphi(\eta)} \|\xi - \varphi(\eta)\|.$$

Examinemos la tarea sobre la valuación del valor de ξ por las magnitudes de $\eta(t)$, $t \in T$, en el caso cuando sólo se tiene un número f i n i t o de valores posibles ξ. Para mayor claridad se puede suponer que ξ describe el estado del sistema que nos interesa, sobre el que sólo podemos juzgar por la observación de un proceso aleatorio $\eta = \eta(t)$.

En ejemplos sencillos nos podemos convencer fácilmente de que, hablando en general, la esperanza matemática $\mathbf{M}(\xi \mid \eta)$ no toma obligatoriamente uno de los valores posibles de ξ (digamos, $\xi = 0, 1, \ldots, N$) de modo que la magnitud $\mathbf{M}(\xi \mid \eta)$ no nos da directamente una indicación respecto a los estados posibles $0, 1, \ldots, N$ en que se encuentra precisamente el sistema.

En adelante, llamaremos a la magnitud $\varphi_0(\eta) \in L$ *la valuación óptima* del estado desconocido ξ, si la probabilidad de los errores, es decir, $\mathbf{P}\{\varphi_0(\eta) \neq \xi\}$ es mínima:

$$\mathbf{P}\{\varphi_0(\eta) \neq \xi\} = \min_{\varphi(\eta)} \mathbf{P}\{\varphi(\eta) \neq \xi\}. \tag{3.10}$$

Está claro que para hallar la valuación óptima es suficiente limitarse a las magnitudes $\varphi(\eta) \in L$, cuyos valores posibles son los mismos que para las magnitudes ξ (es decir, según la suposición hecha son $0, 1, \ldots, N$).

Suponiendo como anteriormente, que existe la distribución condicional de probabilidades [1] $P_\xi(k \mid \eta)$, $k = 0, \ldots, N$ de la magnitud ξ para cualesquiera valores fijos $\eta(t)$, $t \in T$, *admitamos que* $\hat{\xi} = \varphi_0(\eta)$ *es igual al valor de* $k = 0, 1, \ldots, N$ *para el cual la probabilidad* $P_\xi(k \mid \eta)$ *es máxima*:

$$P_\xi(\hat{\xi} \mid \eta) = \max_{0 \leqslant k \leqslant N} P_\xi(k \mid \eta). \tag{3.11}$$

T e o r e m a. *La magnitud* $\hat{\xi} = \varphi_0(\eta)$ *que satisface la condición* (3.11) *es la valuación óptima de* ξ.

D e m o s t r a c i ó n. Introduzcamos las funciones $\chi_k(x)$, $k = 0, 1, \ldots, N$ que son iguales a 1 para $x = k$ e iguales a cero para los restantes valores de la variable x. Tomemos cualquiera magnitud $\varphi = \varphi(\eta)$ que tome uno de los valores $0, 1, \ldots, N$. Se ve fácilmente que

$$\mathbf{P}\{\varphi(\eta) \neq \xi\} = \mathbf{M}\zeta = \mathbf{M}[\mathbf{M}(\zeta \mid \eta)],$$

[1]) La probabilidad $P_\xi(k \mid \eta)$ algunas veces la llaman *a posteriori* (después de la «observación» η).

donde la magnitud $\zeta=\frac{1}{2}\sum_{k=0}^{N}[\chi_k(\xi)-\chi_k(\varphi)]^2$ es igual a 1 para $\varphi(\eta)\neq\xi$ y es igual a cero en los casos restantes (para $\varphi(\eta)=\xi$). Para los valores fijos $\eta(t)$, $t\in T$, la magnitud $\chi_k(\varphi)$ es constante y $\mathbf{M}\{\chi_k(\xi)\mid\eta\}=P_\xi(k\mid\eta)$ de modo que

$$\mathbf{M}(\zeta\mid\eta)=\frac{1}{2}\sum_{k=0}^{N}\mathbf{M}\{[\chi_k(\xi)-\chi_k(\varphi)]^2\mid\eta\}=$$

$$=\frac{1}{2}\sum_{k=0}^{N}\mathbf{M}\{[\chi_k(\xi)-P_\xi(k\mid\eta)]^2\mid\eta\}+\frac{1}{2}\sum_{k=0}^{N}[P_\xi(k\mid\eta)-\chi_k(\varphi)]^2,$$

donde para $\varphi(\eta)=j$

$$\frac{1}{2}\sum_{k=0}^{N}[P_\xi(k\mid\eta)-\chi_k(\varphi)]^2=\frac{1}{2}\sum_{k=0}^{N}P_\xi(k\mid\eta)^2-P_\xi(j\mid\eta)+\frac{1}{2}.$$

Supongamos que $\zeta_0=\frac{1}{2}\sum_{k=0}^{N}[\chi_k(\xi)-\chi_k(\varphi_0)]^2$ y examinemos la diferencia

$$\mathbf{M}(\zeta\mid\eta)-\mathbf{M}(\zeta_0\mid\eta)=P_\xi(j_0\mid\eta)-P_\xi(j\mid\eta),$$

donde $j_0=\varphi_0(\eta)$ es aquel valor para el cual la probabilidad $P_\xi(j\mid\eta)$ es máxima (véase (3.11)). Está claro que para cualquier $j=\varphi(\eta)$

$$\mathbf{M}(\zeta\mid\eta)-\mathbf{M}(\zeta_0\mid\eta)=\mathbf{M}\{\zeta-\zeta_0\mid\eta\}\geqslant 0$$

y por consiguiente,

$$\mathbf{M}\zeta-\mathbf{M}\zeta_0=\mathbf{M}(\zeta-\zeta_0)=\mathbf{M}[\mathbf{M}\{\zeta-\zeta_0\mid\eta\}]\geqslant 0,$$

es decir

$$\mathbf{P}\{\varphi(\eta)\neq\xi\}\geqslant\mathbf{P}\{\varphi_0(\eta)\neq\xi\},$$

que es lo que se quería demostrar.

Ejemplo (*tarea sobre el desajuste*). Supongamos que nos interesa el estado del sistema en el momento variable de tiempo t, sobre el cual se podría juzgar sin error según la señal determinada $\theta(t)$, que caracteriza el paso de un estado a otro, pero de hecho «observamos» sólo el proceso aleatorio $\xi=\xi(t)$, $t\geqslant 0$, de tal modo que

$$d\xi(t)=\theta(t)\,dt+d\eta(t),\tag{3.12}$$

donde en la diferencial estocástica $d\xi(t)$ figura, como corrientemente, el proceso del movimiento browniano standard $\eta(t)$. Se exige dar la valuación del estado del sistema en el momento dado de tiempo t

según la trayectoria $\xi(s)$, $s \leqslant t$, más exactamente, dar la valuación de la magnitud $\theta(t)$.

Supongamos para mayor sencillez que el sistema se encuentre en el momento inicial en el estado 0 y en el transcurso del tiempo puede pasar a algún otro estado 1 (el paso puede significar, por ejemplo, el desajuste o rotura del sistema examinado).

Consideraremos que

$$\theta(t) = \begin{cases} 0 \text{ para } t < \tau, \\ 1 \text{ para } t \geqslant \tau, \end{cases}$$

donde el momento τ del paso de 0 a 1, es una magnitud aleatoria con la distribución exponencial de probabilidades:

$$\mathbf{P}\{\tau > t\} = e^{-\lambda t}.$$

Como fue indicado anteriormente, la valuación óptima del valor desconocido $\theta(t)$ por las magnitudes $\xi(s)$, $s \leqslant t$ (cuyo conjunto designemos, para abreviar, por ξ_t), se halla con ayuda de *las probabilidades a posteriori* [1])

$$\pi_0(t) = \pi(t), \quad \pi_1(t) = 1 - \pi(t),$$

donde

$$\pi(t) = \mathbf{P}\{\theta(t) = 0 \mid \xi_t\}.$$

Supongamos que el proceso aleatorio $\xi = \xi(t)$ se observa en los momentos discretos de tiempo t, múltiplos de Δt, y sigamos la evolución de las probabilidades a posteriori $\pi(t)$ en el transcurso del tiempo t.

Examinemos preliminarmente la densidad condicional de probabilidades $p\{x \mid \theta(t) = 0, \xi_t\}$ de la magnitud $\xi(t + \Delta t)$ para la condición $\theta(t) = 0$ y para la trayectoria fija $\xi_t = \{\xi(s), s \leqslant t\}$. Como se ve de la fórmula (3.12)

$$\xi(t + \Delta t) = \xi(t) + \int_t^{t+\Delta t} \theta(s)\, ds + \Delta\eta(t),$$

donde $\Delta\eta(t)$ es una magnitud de Gauss independiente con la media nula y dispersión Δt. Para la función fija $\theta(s)$, $t \leqslant s \leqslant t + \Delta t$, la densidad buscada es

$$\frac{1}{\sqrt{2\pi\Delta t}} \exp\left\{-\frac{1}{2\Delta t}\left[x - \xi(t) - \int_t^{t+\Delta t} \theta(s)\, ds\right]^2\right\}, \qquad (3.13)$$

$$-\infty < x < \infty,$$

y ya que

$$\int_t^{t+\Delta t} \theta(s)\, ds = \begin{cases} \Delta t - u & \text{para } \tau - t = u,\ 0 \leqslant u \leqslant \Delta t, \\ 0 & \text{para } \tau - t > \Delta t \end{cases}$$

[1]) *A posteriori* (del latín), después de la prueba.

y la magnitud aleatoria $\tau - t$ para la condición $\theta(t) = 0$ está distribuida en la semirecta $u \geqslant 0$ con la densidad de probabilidad $\lambda e^{-\lambda u}$, entonces obtenemos como resultado que

$$p\{x \mid \theta(t)=0, \xi_t\} = \int_0^{\Delta t} \frac{1}{\sqrt{2\pi\Delta t}} \exp\left\{-\frac{1}{2\Delta t}[x-\xi(t)-(\Delta t-u)]^2\right\} \lambda e^{-\lambda u}\,du + \frac{1}{\sqrt{2\pi\Delta t}} \exp\left\{-\frac{1}{2\Delta t}[x-\xi(t)]^2\right\} e^{-\lambda\Delta t}.$$

Para $\theta(t) = 0$, $\theta(t+\Delta t) = 0$ (es decir, de hecho para $\theta(s) \equiv 0$, $t \leqslant s \leqslant t+\Delta t$) la expresión (3,13) nos da la densidad condicional de probabilidad

$$p\{x \mid \theta(t)=0,\ \xi_t,\ \theta(t+\Delta t)=0\} = \frac{1}{\sqrt{2\pi\Delta t}} \exp\left\{-\frac{1}{2\Delta t}[x-\xi(t)]^2\right\}$$

Se comprueba fácilmente que tiene lugar la igualdad siguiente, análoga a la fórmula de Bayes (3.7):

$$\mathbf{P}\{\theta(t+\Delta t)=0 \mid \theta(t)=0,\ \xi_t,\ \xi(t+\Delta t)\} = \mathbf{P}\{\theta(t+\Delta t)=0 \mid \theta(t)=0, \xi_t\} \times \frac{p\{\xi(t+\Delta t) \mid \theta(t)=0,\ \xi_t,\ \theta(t+\Delta t)=0\}}{p\{\xi(t+\Delta t) \mid \theta(t)=0,\ \xi_t\}}, \qquad (3.14)$$

en la cual

$$\frac{p\{\xi(t+\Delta t) \mid \theta(t)=0,\ \xi_t\}}{p\{\xi(t+\Delta t) \mid \theta(t)=0,\ \xi_t,\ \theta(t+\Delta t)=0\}} = \int_0^{\Delta t} \exp\left\{\frac{1}{2\Delta t}[2\Delta\xi(t)\ (\Delta t-u)-(\Delta t-u)^2]\right\} \lambda e^{-\lambda u}\,du + e^{-\lambda\Delta t} = [1+o(1)]\,\lambda\Delta t + e^{-\lambda\Delta t} = 1 + o(\Delta t),$$

donde la magnitud $o(\Delta t)$, dependiente del incremento $\Delta\xi(t) = \xi(t+\Delta t) - \xi(t)$, es una magnitud infinitesimal de orden superior en comparación con Δt para $\Delta t \to 0$ (cuando $\Delta\xi(t) \to 0$). Teniendo en cuenta que $\mathbf{P}\{\theta(t+\Delta t) = 0 \mid \theta(t) = 0, \xi_t\}$ coincide con la probabilidad de quedarse en el estado 0 durante un intervalo de tiempo Δt (esta probabilidad es igual a $e^{-\lambda\Delta t}$), de la igualdad (3.14) obtenemos la relación siguiente:

$$\mathbf{P}\{\theta(t+\Delta t) = 0 \mid \theta(t) = 0,\ \ \xi_{t+\Delta t}\} = 1 - \lambda\,\Delta t + o(\Delta t),$$

donde la magnitud $o(\Delta t)$ tiene el mismo sentido que anteriormente. Luego, análogamente a (3.14) tiene lugar la igualdad

$$\mathbf{P}\{\theta(t)=0 \mid \xi_t, \xi(t+\Delta t)\} = \mathbf{P}\{\theta(t)=0 \mid \xi_t\} \frac{p\{\xi(t+\Delta t) \mid \theta(t)=0,\ \xi_t\}}{p\{\xi(t+\Delta t) \mid \xi_t\}}. \qquad (3.15)$$

Evidentemente,

$$p\,\{\xi(t+\Delta t)\mid \xi_t\} = p\,\{\xi(t+\Delta t)\mid \theta(t) = 0,\ \xi_t\}\,\pi_0(t) + p\,\{\xi(t+\Delta t)\mid \theta(t)=1,\ \xi_t\}\,\pi_1(t),$$

donde la condición $\theta(t) = 1$ de hecho significa que $\theta(s) = 1$ para todos los $s \geqslant t$; de la fórmula general (3.13) obtenemos

$$\frac{p\,\{\xi(t+\Delta t)\mid\theta(t)=1,\ \xi_t\}}{p\,\{\xi(t+\Delta t)\mid\theta(t)=0,\ \xi_t\}} = \exp\left\{\Delta\xi(t) - \frac{\Delta t}{2}\right\}(1+o(\Delta t)),$$

y la igualdad (3.15) permite establecer que

$$\begin{aligned}
\mathbf{P}\{\theta(t)=0\mid\xi_{t+\Delta t}\} &= \mathbf{P}\{\theta(t)=0\mid\xi_t,\ \xi(t+\Delta t)\} = \\
&= \pi_0(t)\,\frac{1}{\pi_0(t)+\pi_1(t)+\exp\left\{\Delta\xi(t)-\frac{\Delta t}{2}\right\}(1+o(\Delta t))} = \\
&= \pi_0(t)\,\frac{1}{1+\pi_1(t)\left(\Delta\xi(t)-\frac{\Delta t}{2}+\frac{1}{2}\Delta\xi(t)^2\right)+o(\Delta t)} = \\
&= \pi_0(t)\left\{1-\pi_1(t)\left(\Delta\xi(t)-\frac{\Delta t}{2}+\frac{1}{2}\Delta\xi(t)^2\right)+\right. \\
&\quad \left.+\left[\pi_1(t)\left(\Delta\xi(t)-\frac{\Delta t}{2}+\frac{1}{2}\Delta\xi(t)^2\right)\right]^2+o(\Delta t)\right\} = \\
&= \pi_0(t)\left\{1-\pi_1(t)\left(\Delta\xi(t)-\frac{\Delta t}{2}+\frac{1}{2}\Delta\xi(t)^2\right)+\pi_1(t)^2\Delta\xi(t)^2+o(\Delta t)\right\},
\end{aligned}$$

donde $o(\Delta t)$ ahora significa una magnitud de un orden infinitesimal superior en comparación con $\Delta t + \Delta\xi(t)^2$.

Por último, utilizando la fórmula

$$\mathbf{P}\{\theta(t+\Delta t) = 0\mid\xi_{t+\Delta t}\} = \mathbf{P}\{\theta(t+\Delta t)=0\mid\theta(t)=0,\ \xi_{t+\Delta t}\}\,\mathbf{P}\{\theta(t)=0\mid\xi_{t+\Delta t}\},$$

de las relaciones obtenidas anteriormente, obtenemos para $\pi(t) = \pi_0(t) = 1-\pi_1(t)$:

$$\begin{aligned}
\pi(t+\Delta t) &= (1-\lambda\Delta t)\,\pi(t)\left[1-(1-\pi(t))\left(\Delta\xi(t)-\frac{\Delta t}{2}+\frac{1}{2}\Delta\xi(t)^2\right)+\right. \\
&\quad \left.+(1-\pi(t))^2\Delta\xi(t)^2\right]+o(\Delta t) = \pi(t)-\pi(t)\times \\
&\quad \times\left\{\lambda\Delta t+\frac{1}{2}(1-\pi(t))\left[\Delta\xi(t)^2-\Delta t\right]-(1-\pi(t))^2\Delta\xi(t)^2\right\}+ \\
&\quad +\pi(t)(1-\pi(t))\,\Delta\xi(t)+o(\Delta t).
\end{aligned}$$

Si despreciamos aquí la magnitud $o(\Delta t)$, que tiene un orden superior en comparación con $\Delta t + \Delta\xi(t)^2$, y también el miembro $\frac{1}{2}\times$ $\times(1-\pi(t))\,[\Delta\xi(t))^2-\Delta t]$, que contiene la magnitud $\Delta\xi(t)^2-\Delta t$

para la cual

$$\mathbf{M}\,\{\,\Delta\xi\,(t)^2 - \Delta t \mid \xi_t\} = o\,(\Delta t)$$

y

$$\mathbf{M}\,\{[\Delta\xi\,(t)^2 - \Delta t]^2 \mid \xi_t\} = o\,(\Delta t)$$

(compárese con las condiciones (8.13), (8.14)), entonces obtenemos

$$\Delta\pi\,(t) \sim -\pi\,(t)\,(\lambda - (1 - \pi\,(t)^2)\,\Delta t + \pi\,(t)\,(1 - \pi\,(t))\,\Delta\xi\,(t).$$

Examinando el caso cuando t es continua se puede indicar que $\pi\,(t) = \mathbf{P}\,\{\theta\,(t) = 0 \mid \xi_t\}$ como función del proceso aleatorio $\xi\,(s)$, $s \leqslant t$, varía continuamente con el transcurso del tiempo t de tal modo que existe la diferencial estocástica

$$d\pi\,(t) = -\,\pi\,(t)\,(\lambda - (1 - \pi\,(t)^2)\,dt + \pi\,(t)\,(1 - \pi\,(t))\,d\xi\,(t),$$

donde la diferencial estocástica $d\xi\,(t)$ está determinada por la fórmula (3.12).

INDICE DE MATERIAS

E

F

H

I

J

L

M

N

P

R

S

T

U

V

A NUESTROS LECTORES:

« MIR » edita libros soviéticos traducidos al español, inglés, francés y árabe. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica; manuales para los centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia ficción.

Dirijan sus opiniones a Editorial « MIR » I Rizhski per. 2, GSP I-110, Moscú, 129820, URSS.

EN 1974 LA EDITORIAL MIR PUBLICARA:

GMURMAN V.

La teoría de las probabilidades y la estadística matemática

Este tratado ha sido escrito de conformidad con el nuevo curso de la teoría de las probabilidades y de la estadística matemática. El material de esta obra se presenta en tres partes fundamentales: las dos primeras se dedican a la teoría de las probabilidades y la última, a la estadística matemática. En este manual se estudian los temas siguientes: probabilidad de las hipótesis; fórmulas de Bayes; distribución de Poisson; nociones de la estadística matemática y noción de correlación.

En el libro se presta gran atención a los métodos estadísticos de elaboración de los datos experimentales; las tablas que se exponen para los cálculos son muy cómodas. Cada capítulo contiene problemas y sus respuestas que han sido elegidos adecuadamente. Además, todo capítulo va acompañado del análisis de las soluciones de los problemas del material correspondiente.

Esta obra contiene diecisiete capítulos, varias tablas de números, veintidós figuras y un gran número de ejemplos teóricos y técnicos; se recomienda para los estudiantes de las facultades ingeniero—técnicas y económicas.

En 1975 se editará:

GMURMAN V.

Obra de referencia para resolver los problemas de la teoría de las probabilidades y de la estadística matemática

Por su contenido, esta obra corresponde al curso de la teoría de las probabilidades y de estadística matemática y contiene más de 500 problemas para todas las partes del curso. Al principio de cada párrafo el autor expone los conocimientos teóricos necesarios, las fórmulas requeridas y muestra los ejemplos cómo resolver problemas típicos. Los problemas destinados para resolverse independientemente se dan según sus dificultades crecientes. Además de los problemas clásicos acerca de lanzamientos de monedas o dados, este libro contiene un gran número de problemas que ilustran la aplicación de la teoría de las probabilidades en la técnica (fiabilidad, control estadístico de la calidad de producción y tratamiento de los datos experimentales). El mérito fundamental de este libro consiste en que el autor, en todos los problemas, pone en primer plano su esencia probabilística, reduciendo al mínimo el cálculo numérico.

Esta obra puede servir para enseñar a resolver los problemas de la teoría de las probabilidades, de estadística matemática y para los estudiantes de las especialidades ingenierotécnicas y económicas.