



201/46
Universidad Nacional Autónoma
de México

FACULTAD DE CIENCIAS

" ECUACIONES DIFERENCIALES ESTOCASTICAS:
DIFUSION Y LANGEVIN "

T E S I S
Que para obtener el titulo de
LICENCIADO FISICO
p r e s e n t a

MARTIN ROMERO CASTILLO

México, D. F.

1989

TESIS CON
FALLA DE ORIGEN



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO

INTRODUCCION

1

CAPITULO I. VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD

1. Introducción.	5
2. Distribuciones de una variable aleatoria.	5
3. Distribuciones de varias variables aleatorias.	13
4. Teorema de Limite Central.	20
5. Discusión.	23

CAPITULO II. PROCESOS ESTOCASTICOS

1. Introducción.	24
2. Que es un Proceso estocástico.	24
3. Teoría Formal de los Procesos Estocásticos.	27
4. Procesos de Markov.	31
5. Procesos de Wiener.	33
6. Funciones y tiempos de correlación.	40
7. Discusión.	45

CAPITULO III. ECUACIONES DIFERENCIALES ESTOCASTICAS

1. Introducción.	47
2. Ecuaciones diferenciales estocásticas y la ecuación de difusión.	48
3. La ecuación de Langevin.	59
4. Procesos de Ornstein-Uhlenbeck.	65
5. Discusión.	70

CAPITULO IV. ECUACION DE LANGEVIN GENERALIZADA

1. Introducción.	71
2. La ecuación de difusión generalizada.	72
3. Formulación microscópica de la ecuación de Langevin generalizada.	75
4. Ecuación de Langevin generalizada.	80
5. Discusión.	91

DISCUSION	93
-----------	----

REFERENCIAS	95
-------------	----

INTRODUCCION

En este trabajo se pretende discutir y analizar dos tipos de ecuaciones diferenciales estocásticas, a saber, las ecuaciones de difusión y de Langevin. Este tipo de ecuaciones diferenciales tiene un gran interés ya que describen una amplia variedad de problemas físicos que aparecen en la electrónica, la hidrodinámica, la química, la teoría de sólidos, etc. Sin embargo, para la discusión de estas ecuaciones se hará referencia a los problemas particulares de difusión y del movimiento Browniano de una partícula esférica inmersa en un fluido.

El movimiento Browniano fue observado por primera vez en 1828 por el botánico inglés Robert Brown^(1,2). En su estudio sobre el polen de plantas observó que al ponerlo en agua se dispersaba en un número muy grande de pequeñas partículas, las cuales realizaban un movimiento continuo, muy accidentado, en zigzag. El orden de magnitud de polen variaba entre 5 y 6 micras (1 micra = 10^{-6} m.).

En los años posteriores al descubrimiento de Brown aparecieron varios trabajos en los que se proponían diferentes mecanismos para explicar ese extraño comportamiento, pero no fue sino hasta principios de este siglo que se halló la explicación de dicho fenómeno.

En el año de 1905, Albert Einstein^(1,2) publicó un importante trabajo en el que propuso la explicación del movimiento Browniano.

Einstein predijo la distancia que debe recorrer una partícula suspendida en un fluido. Para ello se toman los cuadrados de estas distancias encontradas en una sucesión muy grande de experiencias

y se calcula el valor del desplazamiento cuadrático promedio de la partícula Browniana. Los resultados que se encuentran se muestran en la figura 1.

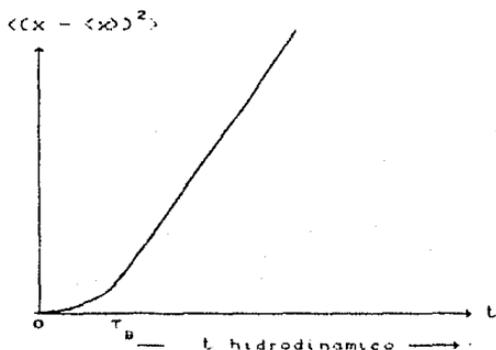


Figura 1. El desplazamiento cuadrático promedio de la partícula Browniana predicho por Einstein ($\tau_B \approx 10^{-8}$ seg.)

En mecánica el hecho de que la curva sea una parábola nos indica que en el intervalo entre 0 y τ_B el tiempo de relajación, la partícula suspendida en el fluido se comporta como una partícula libre, esto es, al colocar la partícula en un fluido, mientras no choque con ninguna de las partículas del fluido se comporta como si no lo sintiera, este intervalo de tiempo es muy corto ($\tau_B \approx 10^{-8}$ seg.) como se indica en la figura 1.

Una vez que la partícula empieza a chocar con las partículas del fluido, la curva cambia a una línea recta, este comportamiento es conocido como difusión. Einstein encontró, que la inclinación de la recta depende de varias cantidades; la temperatura del fluido, su viscosidad, las dimensiones de la partícula y el número

de Avogrado, que es el número de átomos que contiene un mol de una sustancia.

Jean Perrin⁽¹⁾ comprobó experimentalmente el trabajo de Einstein; midió el desplazamiento cuadrático promedio de una partícula Browniana y confirmó, con ello, (a través del cálculo de número de Avogrado) la realidad de la estructura atómica de la materia.

En años posteriores la teoría del movimiento Browniano ha sido introducida en otros campos de la física. En este trabajo se pretende hacer una revisión detallada de dos ecuaciones diferenciales estocásticas que aparecen en el estudio del movimiento Browniano (pero cuya aplicación es más general): las ecuaciones de difusión y Langevin.

En el capítulo I se introduce la teoría de la probabilidad, ya sus conceptos serán utilizados ampliamente a lo largo de esta tesis. En el capítulo II se explican los procesos estocásticos dando una breve discusión sobre el origen de estos, así como su importancia en el campo de la física, en particular, se introducen los procesos de Markov y de Wiener. Se definen las funciones y tiempos de correlación y se muestra la conexión de estas funciones con la densidad espectral. En los capítulos III y IV se presentan las ecuaciones diferenciales estocásticas de difusión y de Langevin que surgen de considerar que la dinámica del movimiento es generada por una fuerza estocástica. En el capítulo III se discute, para referencia posterior, la difusión a partir de la teoría clásica de transporte. Así mismo, se muestran como las ecuaciones de difusión y de Langevin, a distintos tiempos, describen el movimiento de una partícula inmersa en un fluido. Se hace ver la inconsistencia matemática de estas ecuaciones, surgida

al proponer un proceso de Markov ó "ruido blanco" a tiempos muy pequeños para las fuerzas estocásticas. En la última sección de capítulo III, la solución de la ecuación de Langevin es tratada con el proceso de Wiener (Ornstein-Uhlenbeck). Se indica cómo, análogo al caso de Markovianidad, la solución de Ornstein y Uhlenbeck no resuelve el problema de la inconsistencia, que aparece en el cálculo de la correlación de velocidades a tiempos cortos, utilizando los procesos de Wiener. En el capítulo IV se presentan las ecuaciones de difusión y de Langevin generalizadas y se discute la consistencia matemática presentada al proponer un proceso no-Markoviano ó "ruido de color". Finalmente, se dan las conclusiones de el trabajo presentado.

CAPITULO I

VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD

1. INTRODUCCIÓN

En este capítulo se establecen los conceptos fundamentales de la teoría de probabilidad; se discuten los conceptos de variable aleatoria y distribuciones de probabilidad, básicos para el estudio posterior de los procesos estocásticos y las ecuaciones diferenciales estocásticas.

En la sección siguiente se definen algunos conceptos de teoría de la probabilidad de una variable aleatoria. En la sección 3 se formulan a través de la generalización del caso de una variable, los conceptos de teoría de la probabilidad para varias variables aleatorias. La sección 4 es usada para discutir el teorema de límite central, este teorema, fundamental en mecánica estadística, nos asegura que cualquier función de distribución no Gaussiana con media cero y varianza σ^2 satisface, para un número muy grande de eventos una función de distribución Gaussiana y centrada. Finalmente, en la sección 5 se presenta una discusión relacionada con la utilidad, en posteriores capítulos, de la teoría de la probabilidad.

2. DISTRIBUCIONES DE UNA VARIABLE ALEATORIA^(3,4)

Un concepto fundamental en la teoría de probabilidades, que será usado a lo largo de esta revisión es el de variable

aleatoria. Esta se define^(3.3) como una función X contenida en un espacio de probabilidad A , en cada elemento del espacio. Por ejemplo: cuando un dado es lanzado hay seis posibles salidas, por lo que el espacio de probabilidad $A = \{1, 2, 3, 4, 5, 6\}$. La probabilidad de encontrar un número particular $X = 1, 2, 3, 4, 5, 6$ es $P(X) = 1/6$. Se supone que el dado es bueno, esto es, sus caras son simétricas.

Así también, para el caso del dado, para el subconjunto de los números $\{2, 4, 6\}$, la probabilidad de encontrar cada número al ser lanzado es $P(\{2, 4, 6\}) = 3/6 = 1/2$ o $P(X) = 1/2$ donde $X = \{2, 4, 6\}$.

En general, la probabilidad $P(X; A)$ es una función definida en A la cual satisface^(3.5):

$$i) \quad X \subset A, P(X; A) \geq 0.$$

$$ii) \quad P(X_1 \cup X_2; A) = P(X_1; A) + P(X_2; A), \text{ y } X_1 \cap X_2 = \emptyset.$$

$$iii) \quad \text{Si } X_n \subset A \text{ entonces } \sum_n P(X_n; A) = 1.$$

Si el conjunto A consiste de todos los números reales X del intervalo $-\infty$ a x , se define^(3.4) la función de distribución de X

$$F(X; x) = P[X \leq x] = \sum_n P_n(X_n \leq x), \quad (1.1)$$

que es la probabilidad de que el valor observado de X sea menor ó igual a un número real x .

Sea $f(X; x)$ una función de densidad de probabilidad, y X una variable aleatoria continua, entonces la función de distribución de X está dada por:

$$F(X; x) = P [X \leq x] = \int_{-\infty}^x f(X; x') dx', \quad (1.2)$$

donde $f(X; x)dx$ es la probabilidad de que el valor observado de x se encuentre en el intervalo $(x, x + dx)$. Un ejemplo de una función de distribución de una variable aleatoria continua es dado por la distribución Gaussiana ó Normal

$$F [X; x] = \int_{-\infty}^x \frac{\exp(-(x' - m)^2 / (2\sigma^2))}{(2\pi\sigma^2)^{1/2}} dx', \quad (1.3)$$

con m un parámetro y σ un factor de normalización tales que con $-\infty < m < \infty$, $\sigma > 0$, la función de densidad de probabilidad es

$$f(X; x') = \frac{\exp(-(x' - m)^2 / (2\sigma^2))}{(2\pi\sigma^2)^{1/2}}. \quad (1.4)$$

Esta función siempre puede centrarse en cero, para ello hagamos el cambio de variable de integración de x a t , donde $t = (x - m)/\sigma$, encontramos que (1.3) se transforma en

$$F(X; x) = \int_{-\infty}^x \frac{e^{-1/2 t^2}}{(2\pi)^{1/2}} dt, \quad (1.5)$$

esto es, la distribución normal de $(x - m)/\sigma$ está centrada. En general, siempre se podrá reducir cualquier distribución normal a aquella de $m = 0$ y $\sigma = 1$. Más adelante veremos que m es la media y σ la desviación estándar.

Sea X una variable aleatoria continua con una función de densidad de probabilidad $f(X; x)$, se define^(3,4) su valor medio denotado por $\langle X \rangle$ como

$$\langle X \rangle = \int_{-\infty}^{\infty} x f(X; x) dx, \quad (1.6)$$

y para X discreta con una función masiva de probabilidad^(*) $P(X; x)$ dada por

$$\langle X \rangle = \sum_x x P(X; x), \quad (1.7)$$

cubriendo ambos casos para x continua y x discreta, con

$$\langle X \rangle = \int_{-\infty}^{\infty} x dF(X; x), \quad (1.8)$$

la cual se entiende como (1.6) para X continua y (1.7) para X discreta. La varianza es un concepto ampliamente usado en teoría de probabilidad, esta cantidad mide el grado de dispersión de los datos y se define^(3,4) como

$$\begin{aligned} V(X) &= \langle (X - \langle X \rangle)^2 \rangle \\ &= \langle X^2 \rangle - \langle X \rangle^2. \end{aligned}$$

Cuando X es centrado, se tiene $\langle X \rangle = 0$, y $V(X)$ es justamente

(*) Esta definición es debida a su semejanza con la definición de centro de gravedad usada en mecánica (ver referencia d).

$\langle X^2 \rangle$. La desviación estándar σ de X es la raíz cuadrada positiva de su varianza $V(X)$. El momento n -ésimo de X es $\langle X^n \rangle$, es decir, el valor medio de una variable aleatoria es su primer momento y la varianza de una variable aleatoria es su segundo momento.

La función generadora de momentos^(3,4) $\theta(X; k)$ de orden n de la variable aleatoria X es el valor medio de e^{kx} .

$$\theta(X; k) = \langle \exp(kx) \rangle = \int_{-\infty}^{\infty} e^{kx} dF(X; x), \quad (1.9)$$

si todos los momentos s -ésimos existen, podemos derivar la ecuación anterior s veces, esto es:

$$\frac{d^s \theta(X; k)}{dk^s} = \int x^s e^{kx} dF(X; x), \quad s = 1, 2, \dots, s$$

la cual para $k = 0$ nos da $\langle X^s \rangle$. Esto facilita la expansión de MacClaurin de $\theta(X; x)$, dada por la siguiente expresión:

$$\theta(X; x) = 1 + k \langle X \rangle + \frac{k^2}{2!} \langle X^2 \rangle + \dots + \frac{k^s}{s!} \langle X^s \rangle + k \lambda(k), \quad (1.10)$$

donde $\lambda(k) \rightarrow 0$ con k , ya que la s -ésima derivada de $\theta(X; k)$ nos da $\langle X^s \rangle$ para $k = 0$. Las expresiones (1.9) y (1.10) expresan el hecho de que para conocer la función de densidad de probabilidad $f(X; x)$ es necesario conocer todos los momentos de alto orden, es decir, un número infinito. La función característica $\phi(X; u)$ para un parámetro real u , se expresa por^(3,4)

$$\phi(X; u) = \langle \exp(i u x) \rangle = \int_{-\infty}^{\infty} e^{i u x} dF(X; x). \quad (1.11)$$

tomando la transformada inversa de Fourier dada por la expresión:

$$f(X; x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i u x} \phi(X; u) du, \quad (1.12)$$

recuperamos la función de densidad de probabilidad, es decir, la función característica y la función de densidad de probabilidad son la transformada de Fourier una de otra.

El valor medio para la distribución Gaussiana $F(X; x)$ es:

$$\begin{aligned} \langle X \rangle &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} x \exp\left\{-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}\right\} dx \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} (z+m) \exp\left\{-\frac{1}{2} (z/\sigma)^2\right\} dz \\ &= m. \end{aligned}$$

donde se usó que $z = x - m$, y el segundo momento,

$$\langle X^2 \rangle = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} (z^2 + 2zm + m^2) \exp\{-1/2(z/\sigma)^2\} dz$$

por lo tanto,

$$VC(X) = \langle X^2 \rangle - \langle X \rangle^2$$

$$= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} z^2 \exp\left\{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2\right\} dz = \sigma^2.$$

Esto es, m es el valor medio, σ la desviación estandar y $m^2 + \sigma^2$ el segundo momento de X . Análogamente se puede encontrar

$$\langle X^3 \rangle = 3m\sigma^2 + m^3,$$

$$\langle X^4 \rangle = 3\sigma^2 + 6m^2\sigma^2 + m^4,$$

es claro que los momentos n -ésimos se representan como un polinomio homogéneo de orden n en m y σ . En este caso, solo los dos primeros momentos m y σ bastan para describir la distribución Normal. De las ecuaciones (1.4) y (1.11) se calcula la función característica de la distribución Gaussiana:

$$\begin{aligned} \phi(X; \omega) &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{i\omega x} \exp\left\{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right\} dx \\ &= \frac{e^{i\omega m}}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{i\omega z} \exp\{-1/2(z/\sigma)^2\} dz \\ &= \frac{2e^{i\omega m}}{(2\pi\sigma^2)^{1/2}} \int_0^{\infty} \cos \omega z \exp\{-1/2(z/\sigma)^2\} dz, \end{aligned}$$

empleando el resultado.

$$\int \exp(-a^2 x^2) \cos bx \, dx = \frac{\pi^2 \exp(-b^2/(4a^2))}{2a}, \quad (1.13)$$

se deduce;

$$\phi(X; u) = \exp \left\{ i\mu u - \frac{1}{2} \sigma^2 u^2 \right\}, \quad (1.14)$$

puesto que

$$\int_{-\infty}^{\infty} |\phi(X; u)| \, du = \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \sigma^2 u^2 \right\} \, du = \frac{(2\pi)^{1/2}}{\sigma}.$$

tiene un valor finito, podemos usar la inversa de Fourier de (1.14) para obtener la función Gaussiana (1.4) y así comprobarla, esto es

$$\begin{aligned} f(X; x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \exp \left\{ i\mu u - \frac{1}{2} \sigma^2 u^2 \right\} \, du \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i u(x-m)} \exp \left\{ -\frac{1}{2} \sigma^2 u^2 \right\} \, du \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos u(x-m) \exp \left\{ -\frac{1}{2} \sigma^2 u^2 \right\} \, du, \end{aligned}$$

y usando el resultado (1.13), se encuentra

$$f(X; x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{(x - m)^2}{2\sigma^2} \right\},$$

que es la expresión (1.4), la función de densidad de probabilidad Gaussiana. Una variable aleatoria la cual tiene una distribución Gaussiana se llama una variable aleatoria Gaussiana⁽³⁾. La condición necesaria y suficiente para que una variable aleatoria sea Gaussiana es que su función característica sea de la forma $\exp(i\mu u - 1/2\sigma^2 u^2)$. Cuando la variable Gaussiana es centrada, la función característica toma la forma $\exp(-1/2\sigma^2 u^2)$.

3. DISTRIBUCIONES DE VARIAS VARIABLES ALEATORIAS^(3,4)

Consideremos un sistema el cual involucra más de una variable aleatoria X_1, X_2, \dots, X_n , las cuales se definen como una función en un mismo espacio, por ejemplo, cada X_i ($i = 1, 2, 3, \dots$) pueden ser las componentes del vector de posición de las partículas de un gas en tres dimensiones \vec{r} en un contenedor. La función de distribución conjunta⁽³⁾ $F(X_1, \dots, X_n; x_1, \dots, x_n)$ es

$$F(X_1, X_2, \dots, X_n; x_1, x_2, \dots, x_n)$$

$$= P \{ s: X_1(s) \leq x_1, X_2(s) \leq x_2, \dots, X_n(s) \leq x_n \}, \quad (1.15)$$

donde s es un elemento del espacio y la probabilidad P se toma para valores de s que satisfacen las desigualdades en (1.15). Si las variables aleatorias son continuas, es posible escribir la relación anterior como^(3,4)

$$F(X_1, X_2, \dots, X_n; x_1, x_2, \dots, x_n) =$$

$$\int_{-\infty}^{x_1} dy_1 \int_{-\infty}^{x_2} dy_2 \dots \int_{-\infty}^{x_n} dy_n f(X_1, X_2, \dots, X_n; y_1, y_2, \dots, y_n). \quad (1.16)$$

La función conjunta generadora de momentos $\theta(X_1, \dots, X_n; k_1, \dots, k_n)$ y la función característica conjunta $\phi(X_1, \dots, X_n; u_1, \dots, u_n)$ son respectivamente,

$$\theta(X_1, \dots, X_n; k_1, \dots, k_n) = \langle \exp(k_1 X_1 + \dots + k_n X_n) \rangle, \quad (1.17)$$

y,

$$\phi(X_1, \dots, X_n; u_1, \dots, u_n) = \langle \exp \{ i(u_1 X_1 + u_2 X_2 + \dots + u_n X_n) \} \rangle.$$

El valor medio de una función conjunta $g(X_1, X_2, \dots, X_n)$ se define como

$$\langle g(X_1, X_2, \dots, X_n) \rangle =$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(X_1, X_2, \dots, X_n) dF(X_1, \dots, X_n; x_1, \dots, x_n),$$

donde dF denota (como ya se indicó) el caso discreto y continuo.

Si las variables aleatorias X_1, \dots, X_n son independientes, entonces

$$P\{s: X_1(s) \leq x_1, \dots, X_n(s) \leq x_n\}$$

$$= P\{s: X_1 \leq x_1\} P\{s: X_2 \leq x_2\} \dots P\{s: X_n \leq x_n\}, \quad (1.18)$$

esto implica que:

$$F(X_1, X_2, \dots, X_n; x_1, x_2, \dots, x_n) = F(X_1; x_1) F(X_2; x_2) \dots F(X_n; x_n).$$

Si las variables son continuas, la condición necesaria y suficiente para que sean independientes es que la función de densidad de probabilidad satisfaga

$$f(X_1, \dots, X_n; x_1, \dots, x_n) = f(X_1; x_1) f(X_2; x_2) \dots f(X_n; x_n), \quad (1.19)$$

esto es, podemos deducir la probabilidad de encontrar s en los elementos dx_1, dx_2, \dots, dx_n . La ecuación (1.18) también puede ser escrita en términos de $g_i(X_i)$ y $g_j(X_j)$ como

$$\begin{aligned} \langle g_i(X_i) g_j(X_j) \rangle &= \langle g_i(X_i) \rangle \langle g_j(X_j) \rangle \\ & \quad i \neq j, \quad (1.20) \\ &= \langle g_j(X_j) \rangle \langle g_i(X_i) \rangle, \end{aligned}$$

la cual se puede extender para el producto de tres o más funciones. En este caso la función característica cumple que

$$\phi(X_1, X_2, \dots, X_n; u_1, u_2, \dots, u_n) = \langle e^{iu_1 X_1} \rangle \langle e^{iu_2 X_2} \rangle \dots \langle e^{iu_n X_n} \rangle,$$

que puede expresarse como:

$$\begin{aligned}\phi(X_1, X_2, \dots, X_n; u_1, u_2, \dots, u_n) &= \\ &= \phi(X_1; u_1) \phi(X_2; u_2) \dots \phi(X_n; u_n).\end{aligned}\quad (1.21)$$

Similarmente, se obtiene para la función generadora

$$\begin{aligned}\theta(X_1, X_2, \dots, X_n; k_1, k_2, \dots, k_n) &= \\ &= \theta(X_1; k_1) \theta(X_2; k_2) \dots \theta(X_n; k_n).\end{aligned}\quad (1.22)$$

Si tomamos dos variables independientes X_1 con media m_1 y desviación estándar σ_1 , y X_2 con media m_2 y desviación estándar σ_2 entonces^(3,4)

$$\langle X_1 + X_2 \rangle = \langle X_1 \rangle + \langle X_2 \rangle = m_1 + m_2,$$

y utilizando el resultado (1.20) se obtiene,

$$\begin{aligned}\langle (X_1 + X_2)^2 \rangle &= \langle X_1^2 \rangle + \langle X_2^2 \rangle + 2 \langle X_1 \rangle \langle X_2 \rangle \\ &= \sigma_1^2 + m_1^2 + \sigma_2^2 + m_2^2 + 2 m_1 m_2,\end{aligned}$$

y entonces

$$V(X_1 + X_2) = \sigma_1^2 + \sigma_2^2, \quad (1.23)$$

lo que implica que $X_1 + X_2$ es una variable aleatoria con media $m_1 + m_2$ y desviación estandar $(\sigma_1^2 + \sigma_2^2)^{1/2}$. Además, empleando (1.20) se obtiene para la función característica,

$$\langle \exp[iu(X_1 + X_2)] \rangle = \langle e^{iuX_1} \rangle \langle e^{iuX_2} \rangle,$$

la cual claramente se puede extender para más variables aleatorias independientes,

$$\phi(X_1 + X_2 + \dots + X_n; u) = \phi(X_1; u) \phi(X_2; u) \dots \phi(X_n; u). \quad (1.24)$$

Similarmente, para la función generadora de momentos,

$$\theta(X_1 + X_2 + \dots + X_n; k) = \theta(X_1; k) \theta(X_2; k) \dots \theta(X_n; k). \quad (1.25)$$

Estas ecuaciones no deben ser confundidas con (1.21) y (1.22), las cuales se refieren a la distribución de variables aleatorias independientes X_1, X_2, \dots, X_n , mientras que las anteriores nos refieren a la distribución de una variable aleatoria particular $X_1 + X_2 + \dots + X_n$ que es la suma de las variables aleatorias independientes.

Se deduciran ahora algunas relaciones que también serán de utilidad en cálculos posteriores.

La covarianza de dos variables aleatorias se define^(3, 4) como:

$$\text{Cov}(X_1, X_2) = \langle (X_1 - \langle X_1 \rangle) (X_2 - \langle X_2 \rangle) \rangle,$$

y se tiene las siguientes identidades

$$\text{Cov}(X_2, X_1) = \text{Cov}(X_1, X_2).$$

$$\text{Cov}(X_1, X_2) = \langle X_1 X_2 \rangle - \langle X_1 \rangle \langle X_2 \rangle. \quad (1.26)$$

$$\text{Cov}(X_1, X_1) = \text{VC}(X_1).$$

Así mismo, se define^(3,4) la correlación entre dos variables aleatorias X_1 y X_2 , como

$$\rho(X_1, X_2) = \frac{\langle X_1 X_2 \rangle - \langle X_1 \rangle \langle X_2 \rangle}{\sigma_1 \sigma_2},$$

donde $\sigma_1 \neq 0$ y $\sigma_2 \neq 0$ son las desviaciones estándar de X_1 y X_2 respectivamente. La correlación nos da la relación entre dos variables aleatorias, esto es, qué tanta dependencia hay entre ellas. Si la correlación entre X_1 y X_2 es completa, entonces $\rho = 1$. En el caso $\rho = 0$ tenemos que X_1 y X_2 son independientes. El inverso no es necesariamente válido, esto es, la correlación ρ puede tomar cualquier valor sin importar que X_1 , X_2 sean o no independientes.

Si X_1 y X_2 no están correlacionadas se tiene que

$$\langle X_1 X_2 \rangle = \langle X_1 \rangle \langle X_2 \rangle, \quad (1.27)$$

y entonces

$$\begin{aligned} \text{VC}(X_1 + X_2) - \text{VC}(X_1) - \text{VC}(X_2) &= \\ &= \langle (X_1 + X_2)^2 \rangle - \langle X_1 + X_2 \rangle^2 - \langle X_1^2 \rangle + \langle X_1 \rangle^2 - \langle X_2^2 \rangle + \langle X_2 \rangle^2 \\ &= 2\langle X_1 \rangle \langle X_2 \rangle - (\langle X_1 \rangle + \langle X_2 \rangle)^2 + \langle X_1 \rangle + \langle X_2 \rangle^2 = 0. \end{aligned}$$

por tanto,

$$V(X_1 + X_2) = V(X_1) + V(X_2). \quad (1.28)$$

La matriz de covarianza de X_1 y X_2 se define^(3,4) por

$$\begin{bmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) \end{bmatrix}. \quad (1.29)$$

Similarmenre, la matriz de covarianza de las variables aleatorias X_1, X_2, \dots, X_n es una matriz $n \times n$, cuyo elemento ij es $\text{Cov}(X_i, X_j)$.

Si X_1 y X_2 son continuas, entonces $f(X_1; x_1)$ y $f(X_2; x_2)$ son las funciones de densidad de probabilidad y $f(X_1, X_2; x_1, x_2)$ su densidad de probabilidad conjunta.

Cuando $X_1 = x_1$, $\Pi_2(X_1, X_2; x_1, x_2) dx_2$ es la probabilidad de encontrar X_2 en $(x_2, x_2 + dx_2)$, Π_2 es la probabilidad condicional, de aquí que tengamos⁽³⁾

$$\Pi_2(X_1, X_2; x_1, x_2) = \frac{f(X_1, X_2; x_1, x_2)}{f(X_1; x_1)}. \quad (1.30)$$

En el próximo capítulo se discutirá ampliamente el caso de más variables aleatorias en Π_n . Esto conducirá, de manera natural, a definir los procesos estocásticos, que serán la base para tratar, en términos probabilísticos, las ecuaciones diferenciales estocásticas de difusión y Langevin.

4. TEOREMA DE LIMITE CENTRAL^(9,4)

El teorema del limite central establece las condiciones bajo las cuales sumas de variables aleatorias independientes están normalmente distribuidas. Es decir, si se tiene una distribución común $f(X; x)$ no Gaussiana, pero con media cero y varianza finita $\sigma^2 = 1$, conforme $n \rightarrow \infty$ (es decir, $X_1, X_2, \dots, X_{n \rightarrow \infty}$) la distribución de las sumas normalizadas $\sum_n X_n = (X_1 + \dots + X_n) / n^{1/2}$ tiende a la distribución normal $f(0,1)$ con densidad $f(X; x) = e^{-1/2 X^2} / (2\pi)^{1/2}$.

Mostraremos ahora este teorema. Sea un conjunto de variables aleatorias independientes X_1, X_2, \dots, X_n . Supongamos que el conjunto es distribuido idénticamente, esto es, cada miembro del conjunto tiene la misma densidad de probabilidad Gaussiana $f(X; x)$ con media m y varianza σ^2 finitas (ver ec. (1.3)). De acuerdo con (1.10) la función generadora de momentos de X_r , es

$$1 + m k + \frac{1}{2} (\sigma^2 + m^2) k^2 + k^2 \lambda(k),$$

donde $\lambda(k)$ es un residuo que se supone pequeño, λ y k son independientes de r puesto que se tiene una distribución idéntica. Para la variable $X_r - m$ la media es cero, entonces

$$\begin{aligned} V(X_r - m) &= \langle (X_r - m)^2 \rangle = \langle X_r^2 - m^2 \rangle \\ &= V(X_r) = \sigma^2, \end{aligned}$$

por lo tanto la función generadora de momentos de $X_r - m$ es:

$$1 + 1/2 \sigma^2 k^2 + k^2 \mu(k), \quad (1.31)$$

donde $\mu(k)$ es un residuo pequeño independiente de n . Por otro lado, la función generadora de momentos de $(X - \langle X \rangle)n^{-1/2}$ donde $X = X_1 + X_2 + \dots + X_n$ es:

$$\left[1 + \frac{\sigma^2 k^2}{2n} + k^2 \mu\left(\frac{k}{n^{1/2}}\right)^n \right]. \quad (1.32)$$

Si escogemos una cantidad positiva ϵ arbitrariamente pequeña, para encontrar un entero n tal que $|\mu(k / (n^{1/2}))| < \epsilon$, entonces (1.32) se expresa como

$$\left[1 + \frac{(\sigma^2 + \eta^2)k^2}{2n} \right]^n,$$

donde $|\eta| < 2\epsilon$. Cuando n aumenta indefinidamente, tomando el límite $n \rightarrow \infty$ en el resultado anterior, se obtiene

$$\lim_{n \rightarrow \infty} \left[1 + \frac{(\sigma^2 + \eta^2)k^2}{2n} \right]^n = \lim_{n \rightarrow \infty} \left[1 + \frac{\sigma^2 k^2}{2n} \right]^n = \exp\left(\frac{1}{2} \sigma^2 k^2\right).$$

Similarmente, la función característica es $\exp(-1/2 \sigma^2 u^2)$. Por tanto, en el límite la distribución de $\left[\sum_{r=1}^n X_r - \sum_{r=1}^n \langle X_r \rangle \right] n^{-1/2}$ es una Gaussiana con media cero y varianza σ^2 .

Puesto que las variables aleatorias X_j son independientes se encuentra para la media

$$\sum_{r=1}^n \langle X_r \rangle = n m,$$

y para la varianza,

$$V\left(\sum_{r=1}^n X_r\right) = n \sigma^2.$$

Si se divide la diferencia de $\left[\sum_{r=1}^n X_r - \sum_{r=1}^n \langle X_r \rangle\right]$ por la desviación estándar $\sigma n^{1/2}$, se obtiene una nueva variable aleatoria

$$Y_n = \frac{\sum_{r=1}^n X_r - \sum_{r=1}^n \langle X_r \rangle}{\left[V\left(\sum_{r=1}^n X_r\right)\right]^{1/2}},$$

esto es,

$$Y_n = \frac{\sum_{r=1}^n X_r - n m}{\sigma n^{1/2}}.$$

Finalmente, la función de distribución de probabilidad de Y_n en el límite cuando $n \rightarrow \infty$, es

$$\lim_{n \rightarrow \infty} P(Y_n) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^x e^{-1/2 s^2} ds,$$

es decir, en el límite cuando $n \rightarrow \infty$, la probabilidad de que Y_n se encuentre en el intervalo $(-\infty, x)$ está dada por una función de distribución Gaussiana. Con esto el teorema de límite central queda demostrado. Este teorema juega un importante papel en el campo de la estadística, en particular, se utilizará ampliamente en el análisis de las ecuaciones diferenciales estocásticas presentadas posteriormente.

5. DISCUSIÓN

El propósito de introducir la teoría de probabilidad en este capítulo, fué el de fundamentar el concepto y uso de las funciones de distribución, los mismos que serán utilizados para fundamentar la noción de procesos estocásticos o aleatorios y realizar el tratamiento de las ecuaciones diferenciales estocásticas de difusión y Langevin.

La dinámica temporal aquí no ha sido analizada, sin embargo, ésta es de suma importancia. Por ejemplo, al tratar la posición de las moléculas de un gas como variables aleatorias, se observa que éstas tienen una fuerte dependencia en el tiempo.

La dependencia en un parámetro temporal de las funciones de densidad de probabilidad es el estudio de los procesos estocásticos.

En el siguiente capítulo se verá como son introducidos éstos formalmente y se estudiarán algunos procesos que serán de particular importancia para el análisis de las ecuaciones diferenciales de difusión y Langevin.

CAPITULO II

PROCESOS ESTOCASTICOS

1. INTRODUCCIÓN

La teoría de los procesos estocásticos juega un papel importante en la descripción de sistemas que no se comportan de manera determinista, sino que contienen fluctuaciones estadísticas de las variables del sistema. Tales sistemas ocurren en muchos campos de la ciencia, en particular de la física.

En el presente capítulo se tratan con detalle los procesos estocásticos. Desde el punto de vista matemático, algunos resultados formales obtenidos en el capítulo anterior serán usados en este contexto.

En la siguiente sección se presenta una discusión física de los procesos estocásticos y su origen. En la sección 3 se introducen los conceptos fundamentales de los procesos estocásticos. Los procesos de Markov y Wiener son discutidos y se dan sus definiciones en las secciones 4 y 5 respectivamente. Se tratan las funciones y tiempos de correlación, las cuales, serán de utilidad. Finalmente se presenta una discusión sobre la importancia que tienen estos conceptos en los procesos estocásticos y en la interpretación física de cantidades medibles.

2. QUE ES UN PROCESO ESTOCASTICO⁽⁷⁾

Los procesos estocásticos y la teoría de probabilidad son en

general objeto de un gran número de estudios y aplicaciones dentro de la ciencia y en particular de la física.

La importancia de los procesos estocásticos en la física se debe a que de la gran cantidad de fenómenos que dependen del tiempo de manera extremadamente complicada, y más aun, lejos de cualquier posibilidad de cálculo y a menudo de observación, se pueden obtener algunos hechos promedios concretos que son observados y obedecen leyes simples. Por ejemplo; el valor instantáneo de la fuerza ejercida por las moléculas de un gas sobre un pistón varía en forma rápida e impredecible, pero cuando se integra sobre intervalos de tiempos largos (que es automáticamente hecho por la inercia del pistón) se obtiene una unción suave que obedece la ley de Boyle. En otro ejemplo similar, las fluctuaciones de la corriente instantánea en un circuito que contiene una resistencia Ohmica ó un tubo de vacío son muy complicadas⁽⁷⁾, pero si se toma el cuadrado y se integra sobre el tiempo se obtiene una cantidad que tiene conexión con leyes simples u otras cantidades físicas.

La experiencia nos enseña, que a pesar de nuestra ignorancia sobre el comportamiento preciso de las variables microscópicas es posible detectar regularidades en el comportamiento macroscópico y formular con ellas leyes generales. Ocurre entonces que los valores precisos de las variables microscópicas no son importantes y por lo tanto se puede promediar sobre estos. Para aclarar, consideremos ahora un gas monoatómico de N moléculas en una caja. El microestado del sistema se determina por las $6N$ coordenadas y momentos. Seleccionemos un instante fijo llamado "tiempo inicial"; los $6N$ valores de las coordenadas y momentos en este tiempo inicial se denotarán por x . El microestado inicial x determina

unívocamente el microestado en cualquier otro tiempo t a través de las ecuaciones de movimiento. Una cantidad física Y relacionada al sistema es una función de las $6N$ variables, por ejemplo, al considerar el número de partículas en un elemento de volumen dado, Y puede ser la fuerza sobre un pistón. El valor de Y al tiempo t es una función $Y_x(t)$ de x y de t . La idea básica de la mecánica estadística es que el sistema puede ser reemplazado por un ensemble de sistemas escogido adecuadamente, todos teniendo las mismas ecuaciones de movimiento pero con diferentes microestados iniciales x . La estructura del ensemble se especifica por la función de densidad $\rho(x)$ tal que $\rho(x)dx$ es el número de sistemas cuyo microestado inicial está en el elemento de volumen dx . Esta sustitución de un ensemble por un sistema simple tiene el efecto de tornar x en una variable estocástica X . El rango de X consiste entonces de todos los posibles microestados y la densidad de probabilidad es, aparte de la normalización, igual a ρ

$$P_x(x) = \frac{\rho(x)}{\int \rho(x') dx'}$$

una vez que esta idea básica ha sido aceptada ésta solo se mantiene para escoger la apropiada P_x , calcular los valores promedio por medio de esta, e interpretar los números resultantes como los valores observados de las cantidades físicas. "El ensemble sirve meramente como un medio para visualizar la distribución de probabilidad", ésto es, la probabilidad de que x esté en un cierto elemento de el ensemble. Por ejemplo, la presión observada sobre un pistón se identifica con el promedio sobre el ensemble de la fuerza ejercida por las moléculas, más que por su

promedio temporal. La pregunta de porqué y cuándo ambos promedios son el mismo es el problema fundamental a justificar la mecánica estadística en equilibrio (teorema ergódico⁽³⁾).

Esto es, la idea básica de la mecánica estadística es que para un proceso estacionario, se puede usar el promedio sobre el ensemble en lugar de el promedio temporal que está directamente conectado con las observaciones.

3. TEORIA FORMAL DE LOS PROCESOS ESTOCASTICOS^(3,4)

Un proceso aleatorio o estocástico es un conjunto de variables aleatorias que dependen del tiempo, $X(t_1)$, $X(t_2)$, ..., $X(t_n)$ con t_1, t_2, \dots, t_n ($t_1 < t_2 < \dots < t_n$) que pertenecen a un conjunto I . La función de distribución conjunta⁽³⁾ de las variables aleatorias $X(t_1), X(t_2), \dots, X(t_n)$ está dada como:

$$\begin{aligned}
 F[X(t_1), X(t_2), \dots, X(t_n); x_1, x_2 \dots x_n] &= \\
 &= P [X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n], \quad (2.1)
 \end{aligned}$$

acorde a la notación usada en el capítulo anterior. La función característica conjunta es:

$$\begin{aligned}
 \phi [X(t_1), X(t_2), \dots, X(t_n); u_1, u_2, \dots, u_n] &= \\
 &= \langle \exp \{ i(u_1 X(t_1) + \dots + u_n X(t_n)) \} \rangle. \quad (2.2)
 \end{aligned}$$

Un proceso estacionario satisface⁽³⁾ que

$$F[X(t_1+\tau), X(t_2+\tau), \dots, X(t_n+\tau); x_1, x_2, \dots, x_n] =$$

$$= F[X(t_1), X(t_2), \dots, X(t_n); x_1, x_2, \dots, x_n], \quad (2.3)$$

y se dice que F es invariante bajo una traslación en el tiempo τ .

Si el conjunto de variables aleatorias $X(t)$, esto es $X(t_1), X(t_2), \dots, X(t_n)$, es continuo, se puede definir la función de densidad de probabilidad. Sea $\omega_1(x_1, t_1)dx_1$ la probabilidad de encontrar $X(t_1)$ en $(x_1, x_1 + dx_1)$, $\omega_2(x_1, t_1; x_2, t_2)dx_1 dx_2$ la probabilidad de encontrar $X(t_1)$ en $(x_1, x_1 + dx_1)$, $X(t_2)$ en $(x_2, x_2 + dx_2), \dots, \omega_n(x_1, t_1; x_2, t_2; \dots, x_n, t_n)dx_1 dx_2 \dots dx_n$ la probabilidad conjunta de encontrar $X(t_1)$ en $(x_1, x_1 + dx_1)$, $X(t_2)$ en $(x_2, x_2 + dx_2), \dots, X(t_n)$ en $(x_n, x_n + dx_n)$. Entonces el conjunto $\omega_1, \omega_2, \dots, \omega_n$ son las funciones de densidad de probabilidad para un conjunto t_1, t_2, \dots, t_n .

De las definiciones anteriores se sigue directamente que la probabilidad de encontrar X_1 en $(x_1, x_1 + dx_1)$ es:

$$\omega_1(x_1, t_1) = \int \omega_2(x_1, t_1; x_2, t_2) dx_2, \quad (2.4)$$

de manera general, la probabilidad de encontrar X_1 en $(x_1, x_1 + dx_1)$ X_2 en $(x_2, x_2 + dx_2), \dots, X_k$ en $(x_k, x_k + dx_k)$ esta dada por

$$\omega_k(x_1, t_1; \dots, x_k, t_k) = \int \dots \int dx_{k+1} \dots dx_n \omega_n(x_1, t_1; \dots, x_n, t_n).$$

Si el conjunto de variables $X(t)$ son independientes, la densidad de probabilidad se factoriza en la forma:

$$\omega_2(x_1, t_1; x_2, t_2) dx_1 dx_2 = \omega_1(x_1, t_1) dx_1 \omega_1(x_2, t_2) dx_2,$$

ésto es,

$$\omega_2(x_1, t_1; x_2, t_2) = \omega_1(x_1, t_1) \omega_1(x_2, t_2), \quad (2.5)$$

generalizando, se tiene que

$$\omega_n(x_1, t_1; x_2, t_2, \dots, x_n, t_n) = \omega_1(x_1, t_1) \omega_1(x_2, t_2) \dots \omega_1(x_n, t_n),$$

para el caso de variables aleatorias independientes, la función característica dada por la ecuación (2.2), puede escribirse como

$$\begin{aligned} \phi [X(t_1), X(t_2), \dots, X(t_n); u_1, u_2, \dots, u_n] &= \\ &= \phi [X(t_1); u_1] \phi [X(t_2); u_2] \dots \phi [X(t_n); u_n]. \end{aligned}$$

Si el conjunto de variables aleatorias $X(t_1), X(t_2), \dots, X(t_n)$ no son independientes entre si, pero están distribuidas como variables aleatorias Gaussianas, el proceso se llama ⁽³⁾ proceso estocástico Gaussiano. Entonces $\omega_1, \omega_2, \dots, \omega_n$ son densidades de probabilidad Gaussianas. La densidad de probabilidad puede escribirse como,

$$\omega_k(x_1, t_1; \dots, x_k, t_k) = \frac{\exp\{-1/2(\vec{x}-\vec{m}) \cdot \vec{C}^{-1} \cdot (\vec{x}-\vec{m})\}}{[(2\pi)^k \det \vec{C}]^{1/2}}, \quad (2.6)$$

donde $\vec{x} = (x_1, x_2, \dots, x_k)$, $m_i = \langle X(t_i) \rangle$ con $i = 1, 2, \dots, k$ y C^{-1} es la inversa de la matriz de covarianza C_{kk} , definida en el capítulo anterior por la ecuación (1.29). De las definiciones de w_1 , w_2 y la densidad de probabilidad condicional Π_2 introducida en la sección 6 del capítulo 1, se tiene que

$$w_2(x_1, t_1; x_2, t_2) dx_1 dx_2 = w_1(x_1, t_1) dx_1 \Pi_2(x_1, t_1; x_2, t_2) dx_2,$$

esto es,

$$w_2(x_1, t_1; x_2, t_2) = w_1(x_1, t_1) \Pi_2(x_1, t_1; x_2, t_2). \quad (2.7)$$

Si $w_1 \geq 0$ y $w_2 \geq 0 \rightarrow \Pi_2 \geq 0$. Las ecuaciones (2.7) y (2.4) nos permiten escribir

$$\begin{aligned} w_1(x_1, t_1) \int \Pi_2(x_1, t_1; x_2, t_2) dx_2 &= \int w_2(x_1, t_1; x_2, t_2) dx_2 \\ &= w_1(x_1, t_1), \end{aligned}$$

finalmente,

$$\int \Pi_2(x_1, t_1; x_2, t_2) dx_2 = 1. \quad (2.8)$$

De la ec. (2.7) y de la simetría de w_2 en (x_1, t_1) y (x_2, t_2) , se puede obtener otra relación para w_1 , esto es

$$\int w_1(x_1, t_1) \Pi_2(x_1, t_1; x_2, t_2) dx_1 = \int w_2(x_1, t_1; x_2, t_2) dx_1$$

$$\begin{aligned}
 &= \int \omega_2(x_2, t_2; x_1, t_1) dx_1 \\
 &= \omega_1(x_2, t_2). \quad (2.9)
 \end{aligned}$$

En el caso del movimiento Browniano, puede suponerse que la probabilidad de encontrar $X(t_2)$ en $(x_2, x_2 + dx_2)$ no depende del valor de $X(t_1)$ en un tiempo anterior t_1 , si $t_2 - t_1$ es muy grande^(*). Por lo tanto, en el caso límite aceptamos la relación (2.5), es decir, la independencia entre eventos, combinando esto con (2.7) se encuentra que

$$\lim_{t_2 - t_1 \rightarrow \infty} \Pi_2(x_1, t_1; x_2, t_2) = \omega_1(x_2, t_2), \quad (2.10)$$

para el movimiento Browniano.

4. PROCESOS DE MARKOV^(3,4)

Una subclase de procesos estocásticos son los llamados procesos de Markov. Estos se utilizarán ampliamente a lo largo del resto del trabajo. Un proceso de Markov es un proceso aleatorio que cumple la siguiente relación⁽³⁾:

$$\Pi_n(x_1, t_1; \dots; x_{n-1}, t_{n-1}; x_n, t_n) = \Pi_2(x_{n-1}, t_{n-1}; x_n, t_n), \quad (2.11)$$

(*) Esto es debido a que a tiempos largos la partícula "olvida" las condiciones iniciales de las que partió.

Con $t_1 < t_2 < \dots < t_{n-1} < t_n$ nos dice que la densidad de probabilidad condicional Π_n , de que $X_n(t_n)$ este en $(x_n, x_n + dx_n)$ es independiente de lo que sucedió antes de t_{n-1} , esto es, el proceso estocástico no tiene memoria. Para un proceso de Markov podemos considerar la siguiente expresión

$$\int \Pi_2(x_1, t_1; x, t) \Pi_2(x, t; x_2, t_2) dx.$$

la integral se toma sobre los valores de x que ocurren para t que satisface $t_1 < t < t_2$. Ya que para un proceso de Markov todas las densidades de probabilidad se reducen a Π_2 , la expresión anterior es igual a la probabilidad de encontrar $X(t_2)$ en $(x_2, x_2 + dx_2)$, si $X(t_1) = x_1$. Para un proceso de Markov, se puede deducir la ecuación de Chapman-Kolmogorov¹³⁾.

$$\Pi_2(x_1, t_1; x_2, t_2) = \int \Pi_2(x_1, t_1; x, t) \Pi_2(x, t; x_2, t_2) dx. \quad (2.12)$$

Esta ecuación nos da la probabilidad de encontrar $X(t_2)$ en $(x_2, x_2 + dx_2)$, si $X(t_1) = x_1$. El orden del tiempo es esencial: t se encuentra entre t_1 y t_2 .

Un ejemplo de un proceso de Markov ocurre en el movimiento Browniano. Una molécula en una solución líquida sufre aproximadamente 10^{21} colisiones/seg., con las partículas del medio. Si $t_i - t_{i-1}$ son pequeñas macroscópicamente, estos impactos destruyen la correlación entre t_i y lo que sucedió antes de t_{i-1} . Por tanto, el movimiento aleatorio de una partícula Browniana es un proceso de Markov. Cuando un proceso $X(t)$ es Markoviano la probabilidad $w_3(x_1, t_1; x_2, t_2; x_3, t_3) dx_1 dx_2 dx_3$ de encontrar

$X(t_1)$ en $(x_1, x_1 + dx_1)$, $X(t_2)$ en $(x_2, x_2 + dx_2)$ y $X(t_3)$ en $(x_3, x_3 + dx_3)$ se expresa como:

$$\omega_2(x_1, t_1; x_2, t_2) dx_1 dx_2 \Pi_2(x_2, t_2; x_3, t_3) dx_3,$$

y por tanto,

$$\omega_3(x_1, t_1; x_2, t_2; x_3, t_3) = \omega_2(x_1, t_1; x_2, t_2) \Pi_2(x_2, t_2; x_3, t_3), \quad (2.13)$$

la extensión de (2.7) se mantiene solo para un proceso de Markov. Continuando este algoritmo se encuentran sucesivamente todas las probabilidades ω_n . Esta propiedad hace los procesos de Markov manejables, esta es la razón de por que estos procesos son a menudo útiles en muchas aplicaciones.

5. PROCESOS DE WIENER^(3,4)

Un proceso de Wiener es un proceso de Markov, Gaussiano. Un proceso de Wiener es un conjunto de variables aleatorias Gaussianas $W(t)$ y tiene las siguientes propiedades:

- $\langle W(t) \rangle = 0$,
- $W(0) = 0$,
- los incrementos $W(t') - W(t'')$; para $t', t'' \geq 0$ son estacionarios e independientes: esto es, $W(t'+r) - W(t''+r) = W(t') - W(t'')$ para $t'+r, t''+r \geq 0$; $W(t'_j) - W(t'_j)$, $W(t'_k) - W(t'_k)$ son independientes para $t'_i > t'_j \geq t'_k > t'_l \geq 0$.

A continuación se analiza la variancia; ya que el valor medio de $W(t)$ es cero, su variancia es igual al segundo momento, y para cualesquiera $t_1, t_2 \geq 0$ se tiene⁽³⁾

$$\begin{aligned} V(W(t_1 + t_2)) &= \langle (W(t_1 + t_2))^2 \rangle \\ &= \langle (W(t_1 + t_2) - W(t_1) + W(t_1) - W(0))^2 \rangle \\ &= \langle (W(t_1 + t_2) - W(t_1))^2 \rangle + \langle (W(t_1) - W(0))^2 \rangle, \end{aligned}$$

como $W(t_1 + t_2) - W(t_1)$ y $W(t_1) - W(0)$ son independientes y tienen media cero, y de la propiedad estacionaria para los incrementos, se encuentra finalmente que

$$V(W(t_1 + t_2)) = V(W(t_1)) + V(W(t_2)). \quad (2.14)$$

Si se escribe

$$V(W(t)) = h(t), \quad (2.15)$$

donde $t > 0, h(t) \geq 0$, se puede expresar la ecuación (2.14) como

$$h(t_1 + t_2) = h(t_1) + h(t_2), \quad (2.16)$$

puesto que para un entero positivo $h(r) = r h(1)$, inmediatamente se sigue que ésto es cierto para r expresado como el cociente de dos enteros positivos. Esto vale también para un número positivo irracional, si se toma ésto como la intersección de dos límites subsecuentes de números racionales. Entonces la función $h(t)$ es

proporcional a t y la ecuación (2.15) para la variancia de $W(t)$ se expresa como⁽³⁾

$$V(W(t)) = \sigma^2 t,$$

donde σ^2 es una cantidad positiva. Más generalmente,

$$\begin{aligned} V(W(t) - W(s)) &= V(W(t - s) - W(0)) \\ &= V(W(t - s)), \end{aligned}$$

la cual puede escribirse como

$$V(W(t) - W(s)) = \sigma^2 |t - s|, \quad (2.17)$$

puesto que σ^2 es positiva, la variancia siempre es positiva. Para un proceso de Wiener, con $0 \leq s \leq t$ se tiene⁽³⁾

$$\begin{aligned} \langle W(s) W(t) \rangle &= \langle W(s)(W(s) + W(t) - W(s)) \rangle \\ &= \langle W(s) W(s) \rangle + \langle (W(s) - W(0))(W(t) - W(s)) \rangle \\ &= V(W(s)) = \sigma^2 s, \end{aligned}$$

Similarmente, cuando $0 \leq t \leq s$,

$$\langle W(s) W(t) \rangle = \langle W(t) W(s) \rangle = \sigma^2 t,$$

cubriendo ambos casos con la siguiente relación,

$$\langle W(s) W(s) \rangle = \sigma^2 \min(s, t), \quad (s, t \geq 0) \quad (2.18)$$

donde $\min(s, t)$ significa el valor mínimo de s y t . Este resultado será de importancia en el tratamiento de la ecuación de Langevin que se dará en el cap. III. Supongase ahora que $t_1 < t_2$ y $t'_1 < t'_2$, para el caso $t_1 < t'_1 < t_2 < t'_2$, se encuentra de (2.18), que

$$\begin{aligned} \langle (W(t_2) - W(t_1)) (W(t'_2) - W(t'_1)) \rangle &= \\ &= \langle W(t_2) W(t'_2) \rangle + \langle W(t_1) W(t'_1) \rangle - \langle W(t_2) W(t'_1) \rangle - \langle W(t_1) W(t'_2) \rangle \\ &= \sigma^2 (t_2 + t_1 - t'_1 - t_1) \\ &= \sigma^2 (t_2 - t'_1), \end{aligned}$$

donde $t_2 - t'_1$ es la intersección de los intervalos $t_2 - t_1$ y $t'_2 - t'_1$, ver la figura 2.

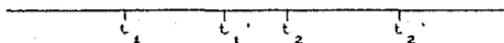


Figura 2. Los valores del tiempo en el argumento en la correlación del proceso de Wiener $W(t)$.

Seleccionando las diferentes posiciones de t'_1 y t'_2 relativo a t_1 y t_2 se puede verificar que esto es generalmente cierto, por

lo tanto se tiene^(a)

$$\langle (W(t_2) - W(t_1)) (W(t'_2) - W(t'_1)) \rangle = \sigma^2 [(t_2 - t_1) \cap (t'_2 - t'_1)] \quad (2.19)$$

Si se hace $t_1 = t$, $t_2 = t + dt$, $t'_1 = t'$, $t'_2 = t' + dt'$, entonces

$$W(t + dt) - W(t) = dW(t)$$

$$W(t' + dt') - W(t') = dW(t')$$

la relación c) de la definición del proceso de Wiener permite asegurar que éstas diferenciales son estacionarias, de (2.19) se deduce que

$$\langle dW(t) dW(t') \rangle = \sigma^2 (dt \cap dt') \quad (2.20)$$

en particular para $t' = t$, se encuentra

$$\langle (dW(t))^2 \rangle = \sigma^2 dt \quad (2.21)$$

este resultado será también de utilidad en el capítulo III en el contexto de la ecuación de Langevin. Sea $f(t, t')$ una función continua no-estocástica, y consideremos la integral^(b)

$$\begin{aligned} \int \int f(t, t') \langle \frac{dW(t)}{dt} \cdot \frac{dW(t')}{dt'} \rangle dt dt' &= \\ &= \langle \int \int f(t, t') \frac{dW(t)}{dt} \cdot \frac{dW(t')}{dt'} dt dt' \rangle \end{aligned}$$

$$= \int \int f(t, t') \langle dW(t) dW(t') \rangle$$

$$= \sigma^2 \int \int f(t, t') (dt \, n \, dt')$$

$$= \sigma^2 \int f(t, t) dt.$$

De aquí, se introduce la función delta de Dirac⁽³⁾

$$\int \int f(t, t') \left\langle \frac{dW(t)}{dt} \cdot \frac{dW(t')}{dt'} \right\rangle dt dt' = \int \int f(t, t') \sigma^2 \delta(t-t') dt dt',$$

donde se encuentra que

$$\left\langle \frac{dW(t)}{dt} \cdot \frac{dW(t')}{dt'} \right\rangle = \sigma^2 \delta(t - t'). \quad (2.22)$$

Considerese un conjunto $W_1(t), W_2(t), \dots, W_q(t)$, donde cada $W_i(t)$ es un proceso de Wiener. Supongamos que las W 's con diferentes subíndices no son correlacionadas, y como son Gaussianas éstas son independientes, esto es

$$\langle W_i(t) W_j(s) \rangle = \langle W_i(s) \rangle \langle W_j(t) \rangle = 0,$$

para $i \neq j$, además, por (2.18), se obtiene

$$\langle W_i(t) W_i(s) \rangle = \sigma^2 \min(s, t). \quad (2.23)$$

Sea $\vec{W}(t)$ un vector columna y $\vec{W}(t)^T$ un vector renglón, esto es su transpuesta, entonces

y por lo tanto,

$$\left\langle \frac{dW_i(t)}{dt} \cdot \frac{dW_j(t')}{dt'} \right\rangle = \delta_{ij} \sigma^2 \delta(t - t'), \quad (2.25)$$

donde δ_{ij} es la delta de Kronecker, que vale 1 si $i = j$ y cero si $i \neq j$. A manera de resumen, puede decirse que los procesos de Wiener $W(t)$ no son estacionarios. Sin embargo, ya que el incremento $W(t + dt) - W(t)$ es estacionario, $dW(t)$ y entonces $dW(t)/dt$ son un proceso estacionario. Ya que $\delta(\lambda)$ es una función de λ , se deduce, poniendo $t = 0$, $t' = t$, de la ecuación (2.22), que la función de autocorrelación de $dW(t)/dt$ es $\sigma^2 \delta(t)$.

6. FUNCIONES Y TIEMPOS DE CORRELACIÓN^(3,0)

Los resultados en muchos de los cálculos en procesos estocásticos son expresados en términos de las funciones de correlación. Las funciones de correlación-temporal tienen una amplia aplicación en el estudio del movimiento Browniano.

Considérense algunos sistemas físicos en los que se toman dos variables aleatorias continuas A y B del sistema, tal como la velocidad lineal del centro de masa de una molécula o su velocidad angular y un vector unitario a través del centro de masa de una molécula en rotación, los armónicos esféricos cuyos ángulos están relacionados con dicho vector. A y B en general dependen de las variables de posición y la velocidad ó de las variables de posición y el momento conjugado del sistema, que son denotadas por $\vec{r}(t)$.

Para un sistema de estado estacionario, la función de correlación temporal de A y B se define^(3,0) como el promedio

sobre el ensemble^(*) $\langle A^*(0) B(t) \rangle$ de $A^*(0)$ y $B(t)$

$$\langle A^*(0) B(t) \rangle = \int d\vec{v}(0) f(0) A^*(0) B(t), \quad (2.26)$$

$f(t)$ es la función de densidad de probabilidad del sistema al tiempo t en ausencia de campos externos. Por lo tanto, para cualquier tiempo s se tiene⁽³⁾

$$f(s) d\vec{v}(s) = f(0) d\vec{v}(0),$$

en un sistema de estado estacionario la probabilidad de encontrar la variable aleatoria \vec{v} entre $(\vec{v}, \vec{v} + d\vec{v})$ es independiente del tiempo. Por lo tanto

$$\begin{aligned} \langle A^*(s) B(t+s) \rangle &= \int d\vec{v}(0) f(0) A^*(0) B(t+s) \\ &= \int d\vec{v}(s) f(s) A^*(s) B(t+s), \end{aligned}$$

haciendo un cambio en la variable de integración de s a t , resulta que

$$\langle A^*(s) B(t+s) \rangle = \langle A^*(0) B(t) \rangle, \quad (2.27)$$

ambos promedios sobre el ensemble se toman sobre el tiempo cero, y en general se pueden tomar para cualquier tiempo. Cuando A y B son

(*) En este y en los posteriores capítulos se usará la palabra ensemble en el sentido de la mec. estadística de equilibrio.

idénticas, se obtiene la función de autocorrelación de A representada por $\langle A^*(0) A(t) \rangle$. Se deduce de (2.26) que

$$\langle A^*(0) A(-t) \rangle = \langle A^*(t) A(0) \rangle^*$$

si $A^*(t)$ conmuta con $A(0)$, entonces

$$\langle A^*(0) A(-t) \rangle = \langle A^*(0) A(t) \rangle^* \quad (2.28)$$

esto es, la función de autocorrelación al tiempo $-t$ es igual al complejo conjugado de la función de autocorrelación al tiempo t .

La función de autocorrelación normalizada de A se define como la cantidad adimensional

$$\langle A^*(0) A(t) \rangle / \langle A^*(0) A(0) \rangle$$

donde $\langle A^*(0) A(0) \rangle = |A(0)|^2$ es una cantidad real.

Las funciones de correlación tienen una conexión especial con las densidades espectrales. Tomemos para un proceso estocástico estacionario una función real del tiempo $X(t)$, y definamos $\tilde{X}(\omega, \tau)$

$$\tilde{X}(\omega, \tau) = \int_{-\tau/2}^{\tau/2} X(t) e^{-i\omega t} dt \quad (2.29)$$

entonces se define^{3,4} la densidad espectral o espectro de potencia $\tilde{X}(\omega)$ de $X(t)$ como

$$X(\omega) = \lim_{T \rightarrow \infty} \frac{|\tilde{X}(\omega, T)|^2}{T}$$

Por otra parte, del teorema ergódico^(3,4) se tiene que, el promedio temporal es igual al promedio sobre el ensemble en la ausencia de un campo externo.

$$\overline{(MC(0) MC(t))} = \langle (MC(0) MC(t)) \rangle,$$

a esta relación se le llama función de correlación de $MC(t)$ y se encuentra que,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} X(\omega) d\omega &= \overline{X(t') X(t'+t)} \\ &= \langle X(t') X(t'+t) \rangle, \end{aligned}$$

usando el resultado (2.27) en el lado derecho de la relación anterior se llega a la expresión

$$\langle X(0) X(t) \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} X(\omega) d\omega, \quad (2.30)$$

Si $\int_{-\infty}^{\infty} |X(\omega) X(t)| dt$ tiene un valor finito, se puede invertir el resultado anterior (2.30), se obtiene

$$X(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \langle X(0) X(t) \rangle dt, \quad (2.31)$$

ésto es, la densidad espectral y la función de autocorrelación son transformadas de Fourier una de otra. Esta relación es conocida como el teorema de Wiener-Khinchine⁽²⁾. El resultado es de gran importancia práctica, por ejemplo, en algunos experimentos de dispersión de luz se mide la densidad espectral del campo eléctrico de luz dispersada, y como consecuencia se determina el espectro de dispersión de luz por medio de las funciones de autocorrelación del campo eléctrico y del detector.

Se puede relacionar la función de autocorrelación con el tiempo de correlación. Se define^(2,6) el tiempo de correlación \mathcal{T}_A para la variable aleatoria $A(t)$ como la integral con respecto a t de 0 a ∞ de su función de autocorrelación normalizada como⁽²⁾

$$\mathcal{T}_A = \int_0^{\infty} \frac{\langle A^*(0) A(t) \rangle}{\langle A^*(0) A(0) \rangle} dt, \quad (2.31)$$

la integral claramente existe, si $\int_{-\infty}^{\infty} |\langle A^*(0) A(t) \rangle| dt$ tiene un valor finito. En los cálculos siguientes $\langle A^*(0) A(t) \rangle$ siempre será real, la integral existe y por tanto el tiempo de correlación será real. Podemos considerar a \mathcal{T}_A como el límite cuando $\omega \rightarrow 0$ de la transformación de Fourier de la función de autocorrelación normalizada tomando solo valores positivos de t , ésto es

$$J_A = \lim_{\omega \rightarrow 0} \int_0^{\infty} \frac{\langle A^*(0) A(t) \rangle}{\langle A^*(0) A(0) \rangle} e^{-\omega t} dt. \quad (2.32)$$

Si la función de autocorrelación se grafica contra t , entonces J_A es el área en el plano positivo entre la curva y el eje t .

7. DISCUSIÓN

El objetivo de este capítulo fué describir los procesos estocásticos. Primeramente, se dió una discusión de la necesidad de introducir dichos procesos desde el punto de vista mesoscópico, ésto es, de la mecánica estadística se tiene que un sistema se puede reemplazar con un ensemble de sistemas escogidos adecuadamente, los cuales nos describen la dinámica del sistema en términos de valores promedios.

Se introdujeron y desarrollaron los conceptos fundamentales para los procesos de Markov y de Wiener y finalmente se dió una descripción de las funciones de correlación. Además, se obtuvo el teorema de Wiener-Khinchine, que relaciona la densidad espectral con la transformada de Fourier de la función de correlación para procesos estacionarios. En la sección 4 se introdujo el concepto de un proceso de Markov ("sin memoria" o "ruido blanco"^(*)), utilizando el ejemplo conocido en física que es el movimiento Browniano.

Consideremos ahora la siguiente cuestión. En principio, se

(*) Ruido blanco se entiende como la no dependencia de la función de densidad espectral $X(\nu)$ en ν .

tiene que para cualquier sistema físico aislado y cerrado se puede describir un proceso de Markov introduciendo todas las variables microscópicas como componentes de X . La pregunta física es, se puede encontrar un conjunto de variables cuyo comportamiento en el tiempo sea descrito como un proceso de Markov?. Sabemos bien, de los hechos experimentales que esto está lejos de ocurrir para cualquier sistema en la naturaleza⁽⁹⁾. Aunque las hipótesis de proceso de Markov modelan relativamente bien ciertos fenómenos físicos como el movimiento Browniano.

En el siguiente capítulo veremos hasta que punto estas hipótesis pueden ser consistentes físicamente.

CAPITULO III

ECUACIONES DIFERENCIALES ESTOCASTICAS

1. INTRODUCCIÓN

En este capítulo se discuten las ecuaciones de difusión y de Langevin. Como se mencionó anteriormente, estas ecuaciones surgen en el contexto de la teoría del movimiento Browniano. No obstante, la aplicación de estas ecuaciones ha sido generalizada⁽²⁾ a situaciones donde la "partícula Browniana" no es una partícula real, sino en lugar de ello es alguna propiedad colectiva de un sistema macroscópico. Un ejemplo es el momento dipolar total eléctrico de una gran muestra de fluido. Aquí, la fluctuación irregular en el tiempo del momento eléctrico corresponde al movimiento irregular de las partículas del polen. Esta generalización es de suma importancia en el caso del movimiento Browniano rotacional.

En la siguiente sección, se discute la diferencia entre una ecuación determinista y una ecuación estocástica; se introduce la ecuación de difusión, que se utiliza, como referencia, en el tratamiento de la ecuación de Langevin.

El objetivo principal de este trabajo, es el de revisar la ecuación de Langevin en el contexto del movimiento Browniano y discutir una inconsistencia matemática surgida al considerar tiempos muy pequeños en las funciones de correlación (la que se hará ver en las secciones 3 y 4). En la sección 5 se da una discusión de los resultados obtenidos, los que sugieren un

tratamiento más cuidadoso del problema, mismo que se hará en el capítulo IV.

2. ECUACIONES DIFERENCIALES ESTOCÁSTICAS Y LA ECUACIÓN DE DIFUSIÓN^(9,10)

Una ecuación diferencial en la cual uno o más coeficientes son aleatorios se llama ecuación diferencial estocástica. La solución de tales ecuaciones nos da una función aleatoria.

A menudo las ecuaciones diferenciales se usan como modelos para describir el comportamiento de sistemas físicos. una distinción entre las ecuaciones diferenciales estocásticas y deterministas se verá en los siguientes párrafos.

a) Ecuación Diferencial Determinista⁽⁹⁾

A manera de ejemplo tomemos una partícula esférica de masa m inmersa en un fluido, la fuerza friccional que se opone al movimiento de la partícula está dada por la ley de Stokes⁽¹¹⁾

$$\vec{F}_c = -\gamma \vec{v}. \quad (3.1)$$

donde $\gamma = 6\pi\eta\alpha$ es el coeficiente de fricción, α es el radio de la partícula esférica y η es la viscosidad dinámica del fluido. La ecuación de movimiento para la partícula, esta dada por (en ausencia de gravedad):

$$m \dot{\vec{v}} + \gamma \vec{v} = 0 \quad \text{o} \quad \dot{\vec{v}} + \gamma/m \vec{v} = 0, \quad (3.2)$$

donde $\frac{1}{\tau} = 1/\tau$, τ es conocido como el tiempo de relajación. La solución de la ecuación (3.2), es dada como

$$\vec{v}(t) = \vec{v}(0)e^{-t/\tau} = \vec{v}(0)e^{-\gamma/m t} \quad (3.3)$$

donde $\vec{v}(0)$ es la velocidad inicial de la partícula inmersa en el fluido. La ecuación diferencial (3.2) es una ecuación determinista, es decir, la velocidad $\vec{v}(t)$ al tiempo t está completamente determinada por sus valores iniciales, dados por la solución (3.3). La relación (3.2) es válida solo si la masa m de la partícula es grande como para que las fluctuaciones en su velocidad debida a las fluctuaciones térmicas sean insignificantes.

b) Ecuación Diferencial Estocástica⁽⁶⁾

El mismo problema puede ser tratado de la siguiente manera. De la ley de la equipartición de la energía para una partícula con un grado de libertad, se tiene que

$$\langle E \rangle = \frac{1}{2} m \langle v^2 \rangle = \frac{1}{2} k T, \quad (3.4)$$

donde E es la energía cinética, k es la constante de Boltzmann, T es la temperatura del baño y m es la masa, que se supone muy pequeña. Entonces, para una partícula, la velocidad térmica

$$v_{term} = \langle v^2 \rangle^{1/2} = \left(\frac{k T}{m} \right)^{1/2}, \quad (3.5)$$

es apreciable, por lo tanto la velocidad de una partícula

"pequeña" no puede ser descrita exactamente por (3.2) con la solución (3.3). Si la masa de la partícula pequeña es aún grande comparada con las masas de las moléculas, se espera que (3.2) sea válida aproximadamente. La ecuación (3.2), sin embargo, debe modificarse para que nos conduzca a la energía térmica correcta (3.4). Esta modificación consiste en añadir una fuerza fluctuante $\vec{F}_f(t)$ en el lado derecho de (3.2), es decir, las fuerzas que actúan sobre la partícula pequeña se descomponen en una fuerza de amortiguación continua $\vec{F}_c(t)$, dada por (3.2) y una fuerza fluctuante $\vec{F}_f(t)$, es decir,

$$m \vec{v} = \vec{F}_c(t) + \vec{F}_f(t) = -\gamma \vec{v}(t) + \vec{F}_f(t). \quad (3.5b)$$

Esta fuerza $\vec{F}_f(t)$ es una fuerza aleatoria ó estocástica, sus propiedades se dan sólo en promedio.

Si se hubiera tratado el problema exactamente, se habrían resuelto las ecuaciones de las fuerzas acopladas al movimiento de todas las moléculas del fluido y de la partícula pequeña, y no aparecerían fuerzas estocásticas. Debido al gran número de moléculas del fluido (del orden de 10^{23}), no se pueden resolver estas ecuaciones acopladas. Además, como no se conocen los valores iniciales de todas las moléculas del fluido, no es posible calcular el movimiento exacto de la partícula pequeña inmersa en el fluido. Si se usara cualquier otro sistema (partícula y fluido) idéntico al primero excepto por los valores iniciales de el fluido, resultaría un movimiento diferente para la partícula pequeña.

Como se vió en la sección 2 del capítulo anterior, se considera el promedio sobre el ensemble de tales sistemas. La fuerza $\vec{F}_f(t)$

varía entonces de sistema a sistema y sólo se pueden considerar los promedios de esta fuerza para el ensemble. Por lo tanto, aplicando el resultado anterior (3.5b), en la ecuación (3.2), la ecuación de movimiento está dada

$$m \ddot{\vec{v}} + \gamma \dot{\vec{v}} = \vec{F}_r(t), \quad (3.6)$$

la ecuación (3.6) es una ecuación diferencial estocástica porque contiene la fuerza estocástica $\vec{F}_r(t)$, a esta ecuación se le llama ecuación de Langevin y será tratada en la sección 3.

La Ecuación de Difusión. ⁽¹⁰⁾

La ecuación de difusión puede ser vista como un tipo muy simple de ecuación diferencial estocástica; sus resultados y conclusiones se aplicarán a la ecuación de Langevin.

Primeramente, considerese la teoría clásica de transporte. Supongase que se tiene un recipiente que contiene un fluido en equilibrio térmico y que en algún lugar del recipiente se introduce una pequeña cantidad del fluido de un tipo diferente. El nuevo fluido comenzará a extenderse por todo el recipiente, pero lo hará con la presencia del fluido original. A este lento proceso de extensión se le llama difusión. La difusión se debe principalmente a los choques que las moléculas del fluido nuevo reciben de las moléculas de el fluido original. Después de un gran número de colisiones, las moléculas del fluido nuevo acaban extendidas más o menos uniformemente por todo el volumen. Esto es, la difusión determina únicamente la movilidad de las moléculas en el fluido debido a las fuerzas internas producidas por las

colisiones al azar (autodifusión).

De acuerdo a la mecánica estadística, las partículas suspendidas en un fluido ejercen una presión sobre el contenedor, dada por

$$p = n k T, \quad (3.7)$$

donde n es el número de partículas por unidad de volumen (concentración). El coeficiente de difusión D se define⁽¹⁰⁾ de manera fenomenológica por

$$\dot{i} = - D \text{grad } n, \quad (3.8)$$

donde \dot{i} es la densidad de corriente de partículas. En el caso de una dimensión, esta ecuación se puede escribir como:

$$\dot{i} = - D \frac{d n}{d x}. \quad (3.9)$$

De otra manera, si una fuerza externa F_{ext} y una resistencia γ actúan sobre una partícula, entonces, en el caso estacionario, se tiene que⁽¹⁰⁾

$$v = \frac{F_{ext}}{\gamma}, \quad (3.10)$$

de lo cual se obtiene la relación

$$\dot{i} = n v = n \frac{F_{ext}}{\gamma}. \quad (3.11)$$

En un estado estacionario donde hay un flujo de corriente como

resultado de un gradiente de presión, la fuerza por unidad de volumen se mantiene en equilibrio, dada por:

$$n F_{\text{ext}} = - \frac{d p}{d x} = - k T \frac{d n}{d x} . \quad (3.12)$$

por lo tanto,

$$F = - \frac{k T}{\gamma} \frac{d n}{d x} ,$$

y que comparando con (3.9),

$$D = \frac{k T}{\gamma} . \quad (3.13)$$

D es el coeficiente de difusión, esta relación es conocida como la fórmula de Einstein. Es conveniente considerar que la distribución espacial de las moléculas del fluido está descrita por una función continua de x , y , z que llamaremos n . Por $n(x, y, z)$ se entiende la densidad numérica de moléculas en un pequeño elemento de volumen centrado en (x, y, z) .

En el caso de una dimensión, se tiene⁽¹⁰⁾ que $\phi(s)$ es la densidad de probabilidad para una partícula que sufre un desplazamiento s a un tiempo τ . Por lo tanto, el número de partículas entre s y $s + ds$ está dado como⁽¹⁰⁾:

$$dn = n \phi(s) ds, \quad (3.14)$$

De la definición de la función $\phi(s)$, se puede obtener el número de partículas que están localizadas al tiempo $t + \tau$ entre

dos planos perpendiculares al eje x , cuya abscisa es x y $x + dx$, integrando (3.14) sobre todo el espacio, se obtiene

$$n(x, t + \tau) dx = dx \int_{-\infty}^{\infty} n(x + s, t) \phi(s) ds. \quad (3.15)$$

Puesto que τ es muy pequeño, se puede poner

$$n(x, t + \tau) = n(x, t) + \tau \frac{\partial n}{\partial t},$$

además, si se expande $n(x + s, t)$ en series de potencias de s :

$$n(x + s, t) = n(x, t) + s \frac{\partial n(x, t)}{\partial x} + \frac{s^2}{2!} \frac{\partial^2 n}{\partial x^2} + \dots$$

Por lo tanto, para τ pequeño, de la ecuación anterior (3.15) se tiene

$$\begin{aligned} n(x, t) + \frac{\partial n}{\partial t} \tau &= n(x, t) \int_{-\infty}^{\infty} \phi(s) ds + \frac{\partial n}{\partial x} \int_{-\infty}^{\infty} s \phi(s) ds \\ &+ \frac{1}{2} \frac{\partial^2 n}{\partial x^2} \int_{-\infty}^{\infty} s^2 \phi(s) ds + \dots \quad (3.16) \end{aligned}$$

Si se supone que los términos de orden mayor son muy pequeños para un orden mayor en τ , se tiene

$$\frac{\partial n}{\partial t} = D \frac{s^2}{\tau} \frac{\partial^2 n}{\partial x^2}. \quad (3.17)$$

dada la siguiente relación,

$$\frac{\partial n}{\partial t} + \text{div } v = 0, \quad (3.18)$$

se tiene que la ecuación de difusión es de la forma:

$$\frac{\partial n}{\partial t} = D \frac{\partial^2 n}{\partial x^2}. \quad (3.19)$$

Comparando ahora las dos ecuaciones (3.18) y (3.19) para $\frac{\partial n}{\partial t}$,

se obtiene

$$D = \frac{1}{2} \frac{\overline{s^2}}{\tau},$$

y

$$\overline{s^2} = \frac{2 k T}{\gamma} \tau, \quad (3.20)$$

las cuales relacionan el coeficiente de difusión D con el desplazamiento cuadrático promedio $\overline{s^2}$.

La ecuación de difusión (3.19) tiene como solución

$$n(x, t) = \frac{e^{-x^2/4Dt}}{(4\pi Dt)^{1/2}}, \quad (3.19')$$

donde $n(x, t)dx$ es la densidad de probabilidad de que la partícula este en x y $x + dx$ al tiempo t .

Por otro lado, la ecuación (3.10) puede ser tratada en términos estocásticos considerando simplemente que su parte derecha,

\vec{F}_{ext}/γ , es "una fuerza estocástica \vec{F}_f " generada por las moléculas del fluido y que es la responsable de la difusión.

En forma tridimensional, la ecuación diferencial estocástica, resulta ser

$$\dot{\vec{r}}(t) = \frac{d}{dt} \vec{r}(t) = \vec{F}_f(t) \equiv \vec{f}, \quad (3.21)$$

en donde $\vec{r}(t)$ es la posición al tiempo t y $\vec{f}(t)$ es la "fuerza fluctuante". Supongamos que \vec{f} es Gaussiana, estacionaria y satisface

$$\langle \vec{f}(t) \rangle = 0, \quad (3.22)$$

$$\langle \vec{f}_i(t) \vec{f}_j(s) \rangle = 2D \delta(t-s) \delta_{ij}, \quad (3.23)$$

con D el coeficiente de difusión, dado por la expresión (3.13). La función $\vec{f}(t)$ es un proceso de Markov, también llamado "ruido blanco" debido a la dependencia en la función delta de Dirac de (3.23). Las relaciones anteriores (3.21)-(3.23) permiten encontrar la distribución de probabilidad condicional

$$n(\vec{r}_1 t_1; \vec{r}_2 t_2) = [4\pi D(t_2-t_1)]^{-3/2} \exp\left[-\frac{|\vec{r}_1 - \vec{r}_2|^2}{4D(t_2-t_1)}\right] \quad (3.19'')$$

esto implica que si una partícula que se difunde está en \vec{r}_1 al tiempo t_1 , entonces la probabilidad al tiempo $t_2 > t_1$ de que esté entre \vec{r}_2 y $\vec{r}_2 + d\vec{r}_2$ es dada por $n(\vec{r}_1 t_1; \vec{r}_2 t_2) d\vec{r}_2$.

La expresión (3.19'') es la forma tridimensional de la (3.19'),

esto demuestra que las dos formulaciones anteriores del proceso de difusión son equivalentes.

La velocidad promedio del cambio de la componente r^2 se determina con (3.19"), y está dada por

$$\frac{\sqrt{\langle [r_i(s+t) - r_i(s)]^2 \rangle}}{t} \xrightarrow{t \rightarrow 0} \frac{\sqrt{2D}}{t^{1/2}} \quad (3.24)$$

donde se usa $\langle \dots \rangle$ para denotar el promedio sobre f^+ y $\langle \dots \rangle$ para denotar el promedio sobre la distribución de posiciones.

La expresión (3.24) presenta una divergencia para tiempos muy pequeños. Esto significaría que la partícula se mueve en un tiempo muy pequeño con velocidad infinita, lo cual es absurdo. La razón se debe a que se asumió que el desarrollo de los eventos al tiempo t son vistos como fenómenos independientes de los eventos en el tiempo anterior (procesos de Markov). Pero esto es muy difícil de justificar al escoger tiempos muy pequeños. Einstein^(13,16) refirió claramente que el problema está en suponer la función delta de Dirac en la correlación de (3.23), puesto que la ecuación diferencial (3.21) describe el proceso de difusión para tiempos grandes ($t \gg m\gamma$). La inconsistencia surgida en el límite de (3.24), sugiere considerar otras hipótesis para la función f^+ que tomen en cuenta la fuerte dependencia de la función de correlación en el tiempo. La versión de Stratonovich del cálculo de Ito resuelve esta inconsistencia surgida al tomar este límite, dicha versión corresponde a un proceso no-Markoviano o "ruido de color", mismo que será tratado en el siguiente capítulo.

En la introducción de este trabajo se ilustro graficamente la

dinámica de una partícula inmersa en un fluido, en ella se representaba el desplazamiento cuadrático promedio contra el tiempo (ver figura 1). Enseguida se muestra la gráfica del desplazamiento cuadrático promedio contra el tiempo. Es decir, la correlación de posiciones calculada con la densidad (3.24), y en la cual se usan los promedios (3.22)-(3.24).

$$\langle [r(s+t) - r(s)]^2 \rangle$$

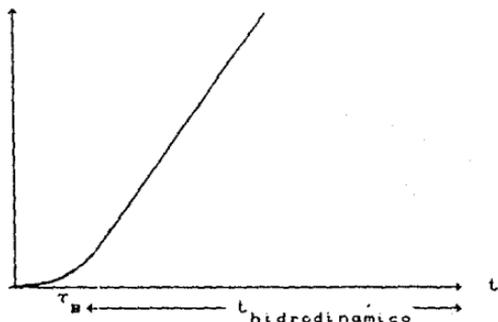


Figura 3. Gráfica de la velocidad promedio del cambio de la posición \vec{r} , dada por la expresión (3.24). ($\tau_B = 10^{-8}$ seg. Tomado de Berne & Pecora: Ref. 19).

La función de correlación para el desplazamiento relaciona las distancias recorridas por la partícula en una sucesión muy grande de experiencias. Como se puede observar de la figura 3, el intervalo de tiempo más corto, para el cual aún se tiene difusión (i.e. la línea recta), es del orden de 10^{-8} seg., esto es, cuando se toma el límite $t \rightarrow 0$ en la expresión (3.24), entendemos físicamente que el límite tiende a un tiempo tan pequeño como 10^{-13} seg. (tiempo que corresponde a la fluctuación en $f(t)$) y no al cero matemático. Una discusión más detallada de este punto se realizará en la sección 3.

3. LA ECUACIÓN DE LANGEVIN^(*)

Un tipo más general de ecuación diferencial estocástica que la ecuación de difusión es la ecuación de Langevin. En el resto del capítulo se hará una discusión de esta ecuación en términos de su conexión con el movimiento Browniano.

Se vió en la sección 2, que la ecuación de movimiento para una partícula Browniana, conocida como la ecuación de Langevin es

$$m \frac{d^2 \vec{r}}{dt^2} = -\gamma \frac{d\vec{r}}{dt} + \vec{F}(t), \quad (3.25)$$

donde $\gamma = 6\pi\eta a$, es el coeficiente de fricción para una partícula esférica de radio a (partícula Browniana)^(*), m es la masa y η es la viscosidad cortante del fluido, $\vec{F}(t)$ es la fuerza fluctuante, esto es, $\vec{F}(t)$ provee una representación fenomenológica de el efecto de miríadas de colisiones rápidas entre la partícula Browniana y las moléculas del fluido. Es por tanto, natural suponer que $\vec{F}(t)$ tiene componentes vectoriales independientes, Gaussianas y con primeros y segundos momentos de la forma^(3, p.13)

$$\langle \vec{F}(t) \rangle = 0, \quad (3.26)$$

$$\begin{aligned} \langle F_i(t) F_j(t') \rangle &= 2q \delta(t - t') \delta_{ij} \\ &= 2q \delta(t) \delta_{ij}. \end{aligned} \quad (3.27)$$

(*) Esta fórmula es válida solo para números de Reynolds muy pequeños y se conoce como fórmula de Stokes (ver ref:11).

donde $i, j = 1, 2, 3$ son las componentes cartesianas de \vec{F} , el valor de q es independiente de t , y posteriormente será calculado. La presencia de la función delta de Dirac en (3.27) manifiesta la conclusión fenomenológica de que la escala de tiempo para las fluctuaciones en la correlación ($\tau \approx 10^{-13}$) es despreciablemente corto relativo a la otra escala de tiempo en la descripción de Langevin, el tiempo de relajación $\tau_{\text{R}} = m/\gamma = 10^{-8}$ seg.

La fuerza estocástica con correlación $\delta(\tau)$, como en el caso de difusión, tiene ruido blanco ó proceso de Markov ("sin memoria"). Para verlo se analizará la función de densidad espectral. La densidad espectral $S_{ij}(\omega)$, de la fuerza estocástica $\vec{F}(t)$, dada por el teorema de Wiener-Khintchine (ecuación (2.31)), es la transformada de Fourier de la función de correlación de $\vec{F}(t)$,

$$S_{ij}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \langle F_i(t) F_j(t') \rangle dt,$$

utilizando la relación (3.27) para la correlación, se obtiene

$$S_{ij}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} q \delta(\tau) \delta_{ij} d\tau = q \delta_{ij}, \quad (3.28)$$

y es independiente de la frecuencia ω . Por lo tanto, la fuerza estocástica de (3.27) con un δ se llama ruido blanco, o un proceso de Markov (en la luz blanca el espectro de potencia no depende de ω).

La solución de la ecuación de Langevin (3.25), con condiciones iniciales $\vec{r}(0) = 0$ y $d\vec{r}/dt = \vec{v}(0)$ es

$$\vec{r}(t) = \frac{\gamma}{m} \left[1 - \exp\left[-\frac{\gamma}{m} t\right] \right] \vec{v}(0) + \frac{1}{m} \int_0^t dt' \int_0^{t'} dt'' \exp\left[-\frac{\gamma}{m} (t' - t'')\right] F(t''), \quad (3.29)$$

si se usa el hecho de que

$$\vec{v}(t) = \frac{d\vec{r}(t)}{dt}, \quad (3.30)$$

entonces

$$\vec{v}(t) = \vec{v}(0) e^{-\gamma/m t} + \frac{1}{m} \int_0^t e^{-\gamma/m(t-t')} \vec{F}(t') dt'. \quad (3.31)$$

Se supone que el valor inicial de la velocidad, $\vec{v}(0)$, está determinada por la distribución de Maxwell

$$W_1(\vec{v}(0)) = \left(2\pi \frac{kT}{m} \right)^{-3/2} \exp\left[-\frac{m \vec{v}(0) \cdot \vec{v}(0)}{2kT} \right],$$

donde k es la constante de Boltzmann y T es la temperatura del fluido. Esta distribución tiene la forma de una distribución Gaussiana.

Las relaciones (3.30) y (3.31) permiten realizar el cálculo de la energía cinética media. Para esto se utilizan los siguientes resultados

$$\langle \vec{v}(t) \rangle = \exp\left[-\frac{\gamma}{m} t\right] \vec{v}(0)$$

y

$$\langle \vec{v}(t) \rangle = 0.$$

entonces,

$$\begin{aligned} \frac{1}{2m} m \langle \vec{v}(t) \cdot \vec{v}(t) \rangle &= \frac{1}{2m} m \langle \vec{v}(0)^2 \rangle e^{-\gamma/m(2t)} + \\ &+ \frac{1}{2m} \int_0^t \int_0^t e^{-\gamma/m(t_1-t_2+t_1-t_2')} q \delta(t_2-t_2') dt_2 dt_2' \\ &= \frac{3}{2} kT \exp\left[-2\frac{\gamma}{m} t\right] + \frac{3}{2} \frac{q}{\gamma} \left[1 - \exp\left[-2\frac{\gamma}{m} t\right]\right]. \end{aligned} \quad (3.32)$$

Si $t \rightarrow \infty$, la partícula Browniana alcanza el equilibrio térmico con el fluido en el cual está inmerso, por ello el valor medio de la energía cinética debe ser $3/2 kT$. Esto solo se cumple si

$$q = k T \gamma. \quad (3.33)$$

éste es el teorema de fluctuación-disipación⁽⁵⁾, es decir, el coeficiente q relacionado con las fuerzas fluctuantes se relaciona con el coeficiente de fricción γ (disipación). El resultado (3.33) permite decir que

$$\frac{1}{2} m \{ \langle \dot{v}(t)^2 \rangle \} = \frac{3}{2} k T, \quad (3.34)$$

para todo t , es decir, el proceso es estacionario.

Calculemos ahora la correlación de posiciones. Es claro que

$$\{ \langle \vec{r}(t) \rangle \} = 0, \quad (3.35)$$

Además, para $t \geq t'$

$$\begin{aligned} \{ \langle r_i(t) r_j(t') \rangle \} &= \frac{mkT}{\gamma^2} \left\{ \exp\left[-\frac{\gamma}{m} t'\right] + \exp\left[-\frac{\gamma}{m} t\right] + 2 \frac{\gamma}{m} t' - \right. \\ &\quad \left. - 1 - \exp\left[-\frac{\gamma}{m} (t - t')\right] \right\} \delta_{ij} \quad (3.36) \end{aligned}$$

haciendo $t = t'$,

$$\{ \langle r_i(t) r_j(t) \rangle \} = \frac{2kT}{\gamma} \left\{ t - \frac{m}{\gamma} + \frac{m}{\gamma} \exp\left[-\frac{\gamma}{m} t\right] \right\} \delta_{ij} \quad (3.37)$$

Para t muy grandes (3.37) reproduce la parte de difusión de la figura 3

$$\frac{2kT}{\gamma} \left\{ t - \frac{m}{\gamma} + \frac{m}{\gamma} \exp\left[-\frac{\gamma}{m} t\right] \right\} \xrightarrow{t \rightarrow \infty} \frac{2kT}{\gamma} t, \quad (3.38)$$

No obstante, para tiempos muy pequeños la ec. (3.37) da

$$2 \frac{kT}{\gamma} \left\{ t - \frac{m}{\gamma} + \frac{m}{\gamma} \exp\left[-\frac{\gamma}{m} t\right] \right\} \xrightarrow{t \rightarrow 0} \frac{kT}{\gamma} t^2. \quad (3.39)$$

En consecuencia, la velocidad promedio del cambio de la componente de $\vec{r}(t)$ es obtenida de (3.29), y es dada por,

$$\sqrt{\frac{\langle [r_i(t'+t) - r_i(t')]^2 \rangle}{t}} \xrightarrow{t \rightarrow 0} \left[\frac{kT}{m} \right]^{1/2}. \quad (3.40)$$

el cual es consistente, esto es, $\vec{r}(t)$ es diferenciable.

Analogamente, utilizando el resultado (3.31) para la velocidad $\vec{v}(t)$, se calcula la correlación para las velocidades

$$\begin{aligned} \langle v_i(t) v_j(t') \rangle &= \exp\left[-\frac{\gamma}{m}(t-t')\right] \langle v_i(0) v_j(0) \rangle + \\ &+ \frac{2q}{m^2} \delta_{ij} \int_0^{t_2} ds \int_0^{t_1} ds' \exp\left[-\frac{\gamma}{m}(t_2-s+t'-s')\right] \delta(s-s') \\ &= \frac{kT}{m} \exp\left[-\frac{\gamma}{m}(t'-t)\right] \delta_{ij}, \quad \text{para } t' \geq t. \end{aligned} \quad (3.41)$$

De (3.41) se sigue que

$$\langle v_i(t) v_i(t') \rangle = \frac{kT}{m} \exp\left[-\frac{\gamma}{m}|t-t'|\right]. \quad (3.42)$$

Por lo tanto, se obtiene la aceleración promedio del cambio de la velocidad $\vec{v}(t)$ es,

$$\sqrt{\frac{\langle [v_x(t'+t) - v_x(t')]^2 \rangle}{t}} \xrightarrow{t \rightarrow 0} \left[\frac{kT\gamma}{m^2} \right]^{1/2} \frac{1}{t^{1/2}}, \quad (3.43)$$

la cual diverge. Es decir, la velocidad es consistente (en términos de (3.40)), mientras que la aceleración no existe. El límite $t \rightarrow 0$ se entiende nuevamente como $t \approx 10^{-13}$ seg, es decir, el tiempo mínimo entre fluctuaciones.

La inconsistencia de (3.43), al igual que en el caso de difusión, está en suponer que la función de correlación para $F(t)$ (ec. (3.27)) es una delta de Dirac, que nos da un proceso de Markov. Por tanto, hay que suponer un proceso en el cual la función de correlación de $F(t)$ sea descrita por un proceso no-Markoviano (cálculo de Ito-Stratonovich), para que la ecuación (3.25) tenga un sentido tanto físico como matemático.

3. PROCESOS DE ORNSTEIN-UHLENBECK⁽³⁾

Enseguida consideraremos la solución de la ecuación de Langevin (3.25) en una dimensión propuesta por Ornstein y Uhlenbeck⁽³⁾:

$$\frac{dv(t)}{dt} = -\gamma/m v(t) + F(t), \quad (3.44)$$

hagamos

$$dv(t) = -\gamma/m v(t) dt + dU(t), \quad (3.45)$$

Puede entenderse (3.45) de la siguiente manera. Sea $f(t)$

cualquier función continua no estocástica que depende de t , multiplicando (3.45) por $f(t)$ e integrando, se tiene

$$\int_{t=a}^b f(t) dv(t) = -\gamma/m \int_{t=a}^b f(t) v(t) dt + \int_{t=a}^b f(t) dU(t). \quad (3.46)$$

en particular, si $f(t) \equiv 1$ resulta que

$$v(b) - v(a) = -\gamma/m \int_a^b v(t) dt + (U(b) - U(a)), \quad (3.47)$$

para el movimiento Browniano puede entenderse como que el cambio en la velocidad en el tiempo $b - a$, es igual a $\int_a^b v(t) dt$ y es generado por el movimiento aleatorio, $U(b) - U(a)$ del baño.

Si se hace $a = 0$, $b = t$ en U , y se toma t muy grande, el proceso $U(t) - U(0)$ puede dividirse en intervalos,

$$U(t) - U(0) = \sum_{k=1}^n (U(t_k) - U(t_{k-1})), \quad (3.48)$$

donde $t_0 = 0 < t_1 < t_2 < \dots < t_n = t$. En el movimiento Browniano los incrementos dados por

$$U(t_1) - U(0), U(t_2) - U(t_1), \dots, U(t_n) - U(t_{n-1}), \quad (3.49)$$

son independientes y $U(t)$ es un proceso de Markov (como los tratados en la sección 3 del capítulo II). Si suponemos que el movimiento aleatorio del baño térmico está en un estado estacionario, entonces los incrementos de (3.49) son también estacionarios y distribuidos idénticamente. Debido a la

aleatoriedad del movimiento $\langle U(t) \rangle = 0$, y escogiendo el tiempo de origen en $t = 0$, entonces, $U(0) = 0$. De lo anterior, ya que cada incremento de (3.49) tiene promedio cero, y aplicando el teorema de límite central a (3.48) deducimos que $U(t) = [V(U(t))]$ ^{1/2} es una variable Gaussiana con media cero y variancia la unidad. Por tanto $U(t)$ es una variable Gaussiana con media cero.

En resumen, se tienen todos los requerimientos especificados en la sección 4 del capítulo II para un proceso de Wiener. Si además, no obstante la no existencia de $dv(t)/dt$, se quiere por conveniencia escribir la ecuación de Langevin en una dimensión para el movimiento de una partícula Browniana en ausencia de fuerzas externas, la ecuación (3.46) toma la forma

$$\frac{d v(t)}{dt} = -\gamma/m v(t) + \frac{dW(t)}{dt}, \quad (3.50)$$

donde $W(t)$ es un proceso de Wiener. Resolviendo (3.46) con la condición inicial $t=0$ y $v(0) = v(0) = \text{cte}$, haciendo $f(t) = e^{\gamma/m t}$, se encuentra

$$\int_{s=0}^t e^{\gamma/m s} dv(s) = -\gamma/m \int_0^t e^{\gamma/m s} v(s) ds + \int_{s=0}^t e^{\gamma/m s} dW(s), \quad (3.51)$$

integrando por partes el lado izquierdo y dividiendo por $e^{\gamma/m t}$ resulta que:

$$v(t) = v(0) e^{-\gamma/m t} + \int_{s=0}^t e^{-\gamma/m(t-s)} dW(s). \quad (3.52)$$

Expresando $dW(s)$ como $W(s+ds) - W(s)$ y viendo que $dW(s)$ es

Gaussiana y centrada, se deduce tomando el promedio de (3.52) que

$$\langle v(t) \rangle = v(0) e^{-\gamma/m t}, \quad (3.53)$$

Así que $v(t)$ es una variable aleatoria Gaussiana y con media $v(0) e^{-\gamma/m t}$. Para el cálculo de la variancia se usan los resultados (3.52) y (3.53) en la definición

$$\begin{aligned} & \langle (v(t) - \langle v(t) \rangle)^2 \rangle \\ &= \left\langle \int_0^t e^{-\gamma/m(t-t_1)} \frac{dW(t_1)}{dt_1} dt_1 \int_0^t e^{-\gamma/m(t-t_2)} \frac{dW(t_2)}{dt_2} dt_2 \right\rangle \end{aligned}$$

utilizando la ecuación (2.22) del capítulo II, se tiene

$$\begin{aligned} &= \sigma^2 e^{-2\gamma/m t} \int_0^t dt_1 \int_0^t dt_2 e^{-\gamma/m(t_1-t_2)} \delta(t_1 - t_2), \\ &= \sigma^2 e^{-2\gamma/m t} \int_0^t e^{\gamma/m t_1} dt_1, \end{aligned}$$

es decir,

$$\langle (v(t) - \langle v(t) \rangle)^2 \rangle = \frac{\sigma^2(1 - e^{-2\gamma/m t})}{2\gamma/m}. \quad (3.54)$$

Esto es, si conocemos σ , la media y la variancia de $v(t)$ puede ser calculada, su función de distribución de probabilidad se calcula como en (1.3) (ver cap. I) con las relaciones anteriores (3.53) y (3.54). El valor de σ^2 se obtiene de la condición de que

cuando t tiende a infinito el sistema se aproxima a un estado de equilibrio estadístico. Ya que la energía promedio está asociada con la velocidad $v(t)$, entonces

$$\begin{aligned} \langle E \rangle &= \frac{1}{2} k T \\ &= \frac{1}{2} m \lim_{t \rightarrow \infty} \langle (v(t))^2 \rangle \\ &= \frac{1}{2} m \frac{\sigma^2}{2\gamma/m} . \end{aligned}$$

Por tanto,

$$\sigma^2 = \frac{2\gamma/m k T}{m} , \quad (3.55)$$

de (3.54) y (3.55), se encuentra finalmente que

$$\{ \langle (v(t) - \langle v(t) \rangle)^2 \rangle \} = \frac{k T}{m} (1 - e^{-2\gamma/m t}) , \quad (3.56)$$

Considerando el mismo límite tomado al final de la sección 3 para la velocidad, (3.43), se encuentra

$$\sqrt{\frac{\langle (v(t'+t) - \langle v(t') \rangle)^2 \rangle}{t}} \xrightarrow{t \rightarrow 0} \left[\frac{k T \gamma}{m^2} \right]^{1/2} \frac{1}{t^{1/2}} , \quad (3.57)$$

este resultado es idéntico al encontrado en (3.43), lo que indica que la inconsistencia ocurre en ambos tratamientos (procesos de

Markov y Wiener). La hipótesis de Markovianidad para $F(t)$ y $U(t)$ no es consistente a tiempos muy pequeños.

5. DISCUSION

En este capítulo se introdujo el concepto de ecuación diferencial estocástica y se discutió su diferencia con una ecuación diferencial determinista. El tratamiento de la difusión se realizó con la teoría clásica de transporte y, de una forma alternativa, con los procesos estocásticos Markovianos, Gaussianos. La formulación estocástica permitió evidenciar la inconsistencia de la hipótesis de Markovianidad para tiempos muy pequeños. En el próximo capítulo se tratará la solución a esta inconsistencia.

En el caso de la ecuación de Langevin, que se resolvió bajo la condición de Markovianidad de la fuerza fluctuante (cálculo de Stratonovich), o alternativamente, con el proceso de Ornstein-Uhlenbeck (cálculo de Ito), una inconsistencia similar se presentó para la aceleración.

La Markovianidad es una idealización hecha por conveniencia matemática, pero en realidad los procesos físicos nunca son estrictamente Markovianos.

En el siguiente capítulo se dará una formulación generalizada de las ecuaciones de difusión y Langevin, lo cual tiene por objeto introducir un término de memoria (la versión de Stratonovich del cálculo de Ito o cálculo de Ito-Stratonovich) y permitir la "descripción matemática" correcta de la difusión y el movimiento Browniano.

CAPITULO IV

LA ECUACION DE LANGEVIN GENERALIZADA

1. INTRODUCCION

En el capítulo anterior se discutieron las ecuaciones de difusión y de Langevin con ruido blanco. Se mostró que al tomar ciertos límites, para tiempos pequeños ($t \approx 10^{-13}$ seg) ocurren divergencias. En un trabajo reciente, Fox⁽¹²⁾ discute esta inconsistencia, y propone las ecuaciones generalizadas de difusión y de Langevin que eliminan esta dificultad.

El propósito de este capítulo es fundamentar y analizar este tipo de ecuaciones generalizadas de difusión y de Langevin que conducen a un proceso no-Markoviano, Gaussiano.

En la siguiente sección se introduce la ecuación de difusión generalizada, y se discute su consistencia matemática en el caso de procesos no-Markovianos (cálculo de Ito-Stratonovich). En la sección 3 se da una discusión sobre la formulación microscópica de la ecuación de Langevin generalizada, esto es, dado el Hamiltoniano del sistema (partícula Browniana más baño térmico), se deduce la ecuación de Langevin generalizada a partir de un modelo de osciladores armónicos. Posteriormente, una vez obtenida la ecuación de Langevin generalizada se da una breve discusión de ésta, en la cual se demuestra que para procesos no-Markovianos la ecuación de Langevin generalizada (y sus resultados) tiene sentido físico y matemático (ver sección 4). Finalmente, en la sección 5 se da una discusión de los resultados obtenidos.

2. LA ECUACIÓN DE DIFUSIÓN GENERALIZADA⁽¹³⁾

La ecuación de difusión para una partícula en un fluido, tratada en el capítulo anterior, fue resuelta suponiendo Markovianidad. Esto, para tiempos infinitesimalmente pequeños es difícil de justificar. La razón está en que para estos tiempos la interacción de la partícula con el baño "recuerda" a cada instante t , lo que ocurrió a un instante t' anterior. La función de correlación en la fuerza estocástica deja de ser una delta de Dirac y en consecuencia pasa a ser una función más complicada de la diferencia de tiempos $t - t'$.

Las ecuaciones de difusión y Langevin generalizadas que obedecen procesos no-Markovianos se pueden obtener de la dinámica microscópica⁽¹⁴⁾. Más adelante veremos como se puede hacer esto en el caso de la ecuación de Langevin. Si el proceso es Gaussiano, pero no-Markoviano, las matemáticas aun pueden ser tratadas.

Enseguida, se muestra como se pueden evitar las inconsistencias en el caso de difusión. La ecuación básica es

$$\frac{d}{dt} \vec{r}(t) = \vec{f}(t), \quad (4.1)$$

Ya que $\vec{f}(t)$ es una función Gaussiana, se tienen el valor promedio y la función de correlación de $\vec{f}(t)$ dados por⁽¹⁵⁾

$$\begin{aligned} \langle \vec{f}(t) \rangle &= 0 \\ \langle f_i(t) f_j(s) \rangle &= D(t-s) \delta_{ij}, \end{aligned} \quad (4.2)$$

en analogía con las condiciones (3.22)-(3.23) del capítulo anterior, con la diferencia de que ahora la función de correlación $D(t-s)$ no es una delta, y en tal caso, se dice que este proceso es no-Markoviano con "memoria" o "ruido de color".

Sea $\vec{r}(0) = 0$ las condiciones iniciales, se tiene⁽¹³⁾ que para $t \geq s$ el valor promedio y la función de correlación de $\vec{r}(t)$ son:

$$\langle \vec{r}(t) \rangle = 0$$

$$\langle r_i(t) r_j(s) \rangle = \int_0^t dt' \int_0^{t'} ds' D(t'-s') \delta_{ij}, \quad (4.3)$$

para resolver la integral doble de la ecuación (4.3), es conveniente escoger en forma explícita $D(t-s)$, la cual tiene las siguientes características generales⁽¹³⁾

$$D(t-s) = \frac{D}{\tau} \exp\left[-\frac{|t-s|}{\tau}\right], \quad (4.4)$$

para $|t-s| \gg \tau$ la ecuación (4.4) se comporta como una función delta de Dirac y se recupera el caso Markoviano. Sustituyéndola en la ecuación (4.2) para $t \geq s$ se obtiene⁽¹³⁾

$$\begin{aligned} \langle r_i(t) r_j(s) \rangle &= \int_0^t dt' \int_0^{t'} ds' \frac{D}{\tau} \exp\left[-\frac{|t'-s'|}{\tau}\right] \delta_{ij} \\ &= [2Ds + D\tau (\exp[-t/\tau] + \exp[-s/\tau] - \exp[-(t-s)/\tau] - 1)] \delta_{ij}, \end{aligned} \quad (4.5)$$

cuando $\tau \rightarrow 0$, $(1/\tau) \exp(-|t|/\tau) \rightarrow 2\delta(t)$ y en el límite de $\tau \rightarrow 0$ la ecuación (4.5) es precisamente $2Ds \delta_{ij}$. Sin embargo, para $0 \leq s \leq t \leq \tau$, la ec. (4.5) se aproxima en el límite $t/\tau \rightarrow 0$ como:

$$\langle r_i(t) r_j(s) \rangle \xrightarrow{t/\tau \rightarrow 0} \frac{D}{\tau} t s. \quad (4.8)$$

En particular, se tiene⁽¹³⁾

$$\langle r_i(t) r_j(t) \rangle \xrightarrow{t/\tau \rightarrow 0} \frac{D}{\tau} t^2. \quad (4.6)$$

De manera análoga a (3.30) del capítulo anterior, la velocidad es⁽¹³⁾

$$\frac{\sqrt{\langle (Cr_i(s+t) - r_i(s))^2 \rangle}}{t} \xrightarrow{t \rightarrow 0} \left(\frac{D}{\tau}\right)^{1/2}. \quad (4.7)$$

la cual es correcta. El comportamiento particular de la función delta de Dirac está implícito dentro de la presencia de τ en el denominador de (4.7), es decir, para que tenga sentido el primer término de (4.7) como una velocidad, los intervalos de tiempo t deben tomarse suficientemente pequeños, lo cual está garantizado en el caso anterior por la función delta de Dirac (ver cap. III), en este caso la situación es salvada por el intervalo de tiempo τ .

Para tiempos grandes ($t \gg \gamma/m$) no se requieren cálculos estocásticos especiales, esto es, las ecuaciones de difusión y Langevin son válidas. Para tiempos pequeños ($t \ll \gamma/m$) se utiliza la versión de Stratonovich del cálculo de Ito (i.e. el cálculo de Ito-Stratonovich o ecuaciones con memoria). Los cálculos de

Ito-Stratonovich describen el análisis estocástico de los procesos no-Markovianos, el cual no presenta las dificultades diferenciales surgidas en los procesos Markovianos, correspondiendo más con la realidad física.

3. FORMULACIÓN MICROSCÓPICA DE LA ECUACIÓN DE LANGEVIN GENERALIZADA⁽⁸⁾

En esta sección se realiza la discusión de los modelos para la ecuación de Langevin generalizada, y algunas de sus propiedades generales. Enseguida se introduce un Hamiltoniano para un sistema acoplado no-lineal en un baño térmico, con el cual se deriva la ecuación de Langevin generalizada exacta para este modelo.

El modelo Hamiltoniano, describe las interacciones de un sistema en un baño térmico de osciladores armónicos, donde se supone que los momentos y coordenadas del sistema (partícula Browniana) son v y x . (por simplicidad todas las masas serán igual a la unidad). Si la energía potencial $U(x)$ del sistema es arbitraria, p_j y q_j son los momentos y las coordenadas del j -ésimo oscilador. La frecuencia del j -ésimo oscilador es ω_j , y γ_j es una constante de acoplamiento. Entonces se tiene, el Hamiltoniano del sistema dado como⁽⁹⁾

$$H = \frac{1}{2} v^2 + U(x) + \sum \frac{1}{2} p_j^2 + \sum \frac{1}{2} \omega_j (q_j - \frac{\gamma_j}{\omega_j^2} x)^2, \quad (4.8)$$

nótese que el acoplamiento del sistema y el baño de los osciladores es bilineal, esto hace posible la derivación de la ecuación de Langevin generalizada⁽⁸⁾. El propósito de esta

derivación, es encontrar las ecuaciones de movimiento para el sistema donde no se involucren explícitamente las variables del baño, los términos se identifican con la disipación y el ruido, es decir, con el proceso. La derivación comienza con las ecs. de Hamilton del movimiento⁽⁸⁾

$$\dot{x} = v,$$

$$\dot{v} = -U'(x) + \sum \gamma_j \left(q_j - \frac{\gamma_j}{\omega_j^2} x \right),$$

$$\dot{q}_j = p_j, \quad (4.9)$$

$$\dot{p}_j = -\omega_j^2 q_j + \gamma_j x.$$

Las ecuaciones de movimiento para el j-ésimo oscilador contienen un término de acoplamiento $\gamma_j x(t)$, que es el campo externo dependiente del tiempo. Estas ecuaciones son lineales, y se pueden resolver explícitamente, esto es

$$q_j(t) = q_j(0) \cos \omega_j t + p_j(0) \frac{\sin(\omega_j t)}{\omega_j} + \gamma_j \int_0^t ds v(s) \frac{\cos \omega_j(t-s)}{\omega_j}. \quad (4.10)$$

El primer y segundo término provienen de las condiciones iniciales, y el tercer término muestra el movimiento del oscilador influenciado por el sistema. La integración por partes nos conduce a una forma más útil, con el cual el resultado anterior queda

$$\begin{aligned}
 q_j(t) = & \left[q_j(0) - \frac{\gamma_j}{\omega_j^2} x(0) \right] \cos \omega_j t + p_j(0) \frac{\sin \omega_j t}{\omega_j} - \\
 & - \gamma_j \int_0^t ds \gamma(s) \frac{\cos \omega_j (t-s)}{\omega_j^2}. \quad (4.11)
 \end{aligned}$$

sustituyendo el resultado (4.11) en la ecuación (4.9) para dv/dt , y agrupando términos, queda:

$$\dot{x}(t) = v(t),$$

$$\dot{v}(t) = -U'(x(t)) - \int_0^t ds K_N(t-s) v(s) + F_N(t). \quad (4.12)$$

El subíndice N nos indica las "ecuaciones de Langevin no-lineales o generalizadas". El kernel $K_N(t)$, o función de memoria, está dado ⁽⁸⁾

$$K_N(t) = \sum \frac{\gamma_j^2}{\omega_j^2} \cos \omega_j t. \quad (4.13)$$

Note que éste es completamente independiente del estado del sistema y el baño, y está completamente determinado por las frecuencias del oscilador y las constantes de acoplamiento. La cantidad $F_N(t)$ representa el "ruido", y está dada como:

$$F_N(t) = \sum \gamma_j [q_j(0) - \frac{\gamma_j^2}{\omega_j^2} x(0)] \cos \omega_j t + \sum \gamma_j p_j(0) \frac{\sin \omega_j t}{\omega_j} \quad (4.14)$$

El término de memoria, ec. (4.13), es una fuerza friccional proporcional a la velocidad del sistema, es no-Markoviana, sin embargo, hay muchos osciladores en el baño⁽⁸⁾ y si sus frecuencias y constantes de acoplamiento se escogen apropiadamente, entonces la función de memoria se aproxima a una función delta y la fricción es aproximadamente Markoviana. En la siguiente sección 4 se usará la función no-Markoviana.

El término de "ruido" $F_N(t)$, ecuación (4.14), está determinado⁽⁸⁾ por los estados iniciales del sistema. Si estos estados iniciales son especificados, entonces $F_N(t)$ se conoce como una función del tiempo y por lo tanto no es verdaderamente aleatoria. Si en otro caso, el estado inicial no se especifica precisamente, pero si en un sentido estadístico, entonces se conocen las propiedades estadísticas de $F_N(t)$. En este caso es apropiado referirse a $F_N(t)$ como ruido aleatorio.

En cualquier experimento individual, $F_N(t)$ se desarrollará en un sentido particular, pero si el experimento se repite con diferentes condiciones iniciales, la nueva $F_N(t)$ será diferente. El comportamiento estadístico de $F_N(t)$ en muchas repeticiones de un experimento es determinado por el ensemble estadístico de las condiciones iniciales; el promedio sobre este ensemble se denotará por $\langle \dots \rangle^0$. Si el ensemble inicial es tal que el primer momento se anula, se tiene⁽⁸⁾

$$\langle p_j(t) \rangle^0 = 0,$$

$$\langle q_j(t) - \frac{\gamma_j}{\omega_j^2} x(t) \rangle^0 = 0. \quad (4.15)$$

Por lo tanto, de la ecuación (4.14) se tiene que el valor promedio del ruido se anula, esto es,

$$\langle F_N(t) \rangle^0 = 0, \quad (4.16)$$

como es de esperarse para el ruido convencional de Langevin.

Si el ensemble es tal que los segundos momentos tienen valores en el equilibrio, la función de correlación de $p(t)$ con las condiciones iniciales está dada como ⁽⁸⁾

$$\langle p_j(t) p_k(0) \rangle^0 = kT \delta_{jk},$$

$$\langle [q_j(t) - \frac{\gamma_j}{\omega_j^2} x(t)]^2 \rangle^0 = \frac{kT}{\omega_j^2} \delta_{jk}, \quad (4.17)$$

donde δ_{jk} es la delta de Kronecker. Entonces, utilizando estos resultados en la ecuación de ruido $F_N(t)$ ec. (4.14), se obtiene que la función de correlación del ruido es proporcional a la función de memoria,

$$\langle F_N(t) F_N(t') \rangle^0 = kT K_N(t - t'), \quad (4.18)$$

a este resultado se le llama ruido de color (con memoria), también se le llama el teorema de fluctuación-disipación generalizado.

4. LA ECUACIÓN DE LANGEVIN GENERALIZADA^(8,12)

En esta sección se resuelve la ecuación de Langevin generalizada para el caso de un proceso no-Markoviano, es decir, con "memoria". En el capítulo III se mostraron las inconsistencias inherentes a la teoría Markoviana. Se mostrará que no se presenta alguna inconsistencia para el caso no-Markoviano, el cual surge de manera natural de la física microscópica.^(8,13)

La ecuación de Langevin generalizada es dada como

$$m \frac{d}{dt} \vec{v}(t) = - \int_0^t \gamma(t-s) \vec{v}(s) ds + \vec{F}(t), \quad (4.19)$$

donde $\gamma(t-s)$ es el kernel de memoria, $\vec{F}(t)$ es la fuerza fluctuante, y tiene las siguientes propiedades Gaussianas

$$\langle \vec{F}_i(t) \rangle = 0$$

$$\langle \vec{F}_i(t) \vec{F}_j(s) \rangle = kT \gamma(|t-s|) \delta_{ij}, \quad (4.20)$$

En el caso especial de que $\gamma(|t-s|) = \gamma \delta(t-s)$, la ecuación (4.19) se reduce a la ecuación (3.21) dada en el capítulo anterior, donde γ es el coeficiente de fricción de la ley de Stokes. Esto es, cuando $\gamma(t-s)$ tiene un ancho en el tiempo, la descripción es no-Markoviana o "ruido de color", como se mostrará

enseguida.

La propiedad Gaussiana implica que las matrices de correlación para dos tiempos determinan todas las funciones de distribución. En este caso, todas las funciones de distribución no pueden ser determinadas por las funciones de distribución de probabilidad $\vec{F}(t_1, t_2; \vec{r}_1, \vec{r}_2)$ o la probabilidad condicional P_2 . Ya que el proceso es no-Markoviano, la solución de (4.19) se escribe⁽¹³⁾ como

$$\vec{v}(t) = \chi(t) \vec{v}(0) + \frac{1}{m} \int_0^t \chi(t-s) \vec{F}(s) ds, \quad (4.21)$$

donde $\chi(t)$ es la matriz de autocorrelación de la velocidad, y se define⁽¹³⁾ como su transformada inversa de Laplace

$$\hat{\chi}(z) = (z + \hat{\gamma}(z) / m)^{-1}, \quad (4.22)$$

y $\hat{\gamma}(z)$ es la transformada de Laplace de $\gamma(t)$, dada por

$$\hat{\gamma}(z) = \int_0^{\infty} \exp[-zt] \gamma(t) dt. \quad (4.23)$$

Por otro lado, se tiene que el valor inicial de la velocidad se determina por medio de la distribución de Maxwell, esto es

$$\langle \vec{v}(0) \vec{v}(0) \rangle = (2\pi m T/m)^{3/2} \exp[-m \vec{v}(0) \cdot \vec{v}(0) / 2kT]. \quad (4.24)$$

De la ecuación (4.21), se sigue que el promedio de la velocidad es

$$\langle \vec{v}(t) \rangle^0 = \chi(t) \vec{v}(0)$$

y,

$$\langle \langle \vec{v}(0) \rangle \rangle^0 = 0, \quad (4.25)$$

donde $\langle \dots \rangle^0$ denota promedio sobre el ensemble de las condiciones iniciales. Y además, se obtiene⁽¹³⁾ la matriz de autocorrelación de velocidad para dos tiempos, $\chi_{ij}(t_2 - t_1)$, para $t_2 \geq t_1$, dada por

$$\begin{aligned} \chi_{ij}(t_2 - t_1) &= \\ &= \langle \langle v_i(t_2) v_j(t_1) \rangle \rangle^0 \\ &= \chi(t_2) \chi(t_1) \langle v_i(0) v_j(0) \rangle + \frac{1}{m^2} \int_0^{t_2} ds \int_0^{t_1} ds' \chi(t_2 - t_1) \langle \vec{F}_i(s) \vec{F}_j(s') \rangle \\ &= \frac{kT}{m} \chi(t_2) \chi(t_1) \delta_{ij} + \frac{kT}{m^2} \int_0^{t_2} ds \int_0^{t_1} ds' \chi(t_2 - s) \chi(t_1 - s') \langle |s - s'| \rangle \delta_{ij}. \end{aligned} \quad (4.26)$$

La integral doble en el lado derecho de (4.26) se calcula usando la transformada doble de Laplace⁽¹³⁾, esto es

$$\begin{aligned}
& \int_0^{\infty} dt_2 \exp[-zt_2] \int_0^{\infty} dt_1 \exp[-z't_1] \int_0^z \int_0^{t_1} ds \int_0^{s'} \chi(t_2-s) \chi(t_1-s') \gamma(|s-s'|) \\
&= \int_0^{\infty} dt_2 \int_0^{\infty} dt_2' \int_0^{\infty} ds_1 \int_0^{\infty} dt_1 \exp[-z(t_2-s_2)] \exp[-z'(t_1-s_1)] \\
&\quad \times \chi(t_2-s_2) \chi(t_1-s_1) \exp[-zs_2] \exp[-z's_1] \gamma(|s_2-s_1|) = \\
&= \int_0^{\infty} ds_2 \int_0^{\infty} dt_2 \int_0^{\infty} ds_1 \int_0^{\infty} dt_1 \exp[-z\tau_2] \chi(\tau_2) \exp[-z'\tau_1] \chi(\tau_1) \\
&\quad \times \exp[-zs_2] \exp[-z's_1] \gamma(|s_2-s_1|) \\
&= \hat{\chi}(z) \hat{\chi}(z') \int_0^{\infty} ds_2 \int_0^{\infty} ds_1 \exp[-z's_2] \exp[-z's_1] \gamma(|s_2-s_1|).
\end{aligned}$$

(4.27)

La integral anterior de (4.27) es otra vez una transformada doble de Laplace, y se obtiene para esta doble integral

$$\int_0^{\infty} ds_2 \int_0^{\infty} ds_1 \exp[-zs_2] \exp[-z's_1] \gamma(|s_2-s_1|) =$$

$$\begin{aligned}
&= \int_0^{\infty} ds_2 \int_0^{\infty} ds_1 \exp[-z(s_2 - s_1)] \gamma(|s_2 - s_1|) \exp[-(z + z')s_1] \\
&= \int_0^{\infty} ds_1 \int_{-s_1}^{\infty} d\sigma \exp[-z\sigma] \gamma(|\sigma|) \exp[-(z + z')s_1] \\
&= \int_0^{\infty} ds_1 \left[\hat{\alpha}(z) + \int_{-s_1}^0 d\sigma \exp[-z\sigma] \gamma(|\sigma|) \right] \exp[-z\sigma] \gamma(|\sigma|) \\
&= \frac{\hat{\gamma}(z)}{z + z'} - \int_0^{\infty} ds_1 \left(\frac{d}{dt} \exp[-(z + z')s_1] \right) \frac{1}{z + z'} \int_{-s_1}^0 d\sigma \exp[-z\sigma] \gamma(|\sigma|) \\
&= \frac{\hat{\gamma}(z)}{z + z'} + \int_0^{\infty} ds_1 \exp[-(z + z')s_1] \frac{1}{z + z'} \exp[zs_1] \gamma(s_1) \\
&= \frac{\hat{\gamma}(z) + \hat{\gamma}(z')}{z + z'}. \tag{4.28}
\end{aligned}$$

Sustituyendo este resultado en (4.27) se obtiene:

$$\begin{aligned}
&\int_0^{\infty} dt_2 \int_0^{\infty} dt_1 \exp[-zt_2] \exp[-z't_1] \\
&\quad \times \int_0^{t_2} ds_2 \int_0^{t_1} ds_1 \chi(t_2 - s_2) \chi(t_1 - s_1) \gamma(|s_2 - s_1|) =
\end{aligned}$$

$$= \hat{\chi}(z) \hat{\chi}(z') \frac{\hat{\gamma}(z) + \hat{\gamma}(z')}{z + z'} \quad (4.29)$$

De la ecuación (4.28) se obtiene que

$$\hat{\chi}(z) \hat{\gamma}(z) = m(1 - z \hat{\chi}(z))$$

y,

$$\hat{\chi}(z') \hat{\gamma}(z') = m(1 - z' \hat{\chi}(z')). \quad (4.30)$$

Usando estas identidades en el lado derecho de la ec. (4.29) nos da:

$$\hat{\chi}(z) \hat{\chi}(z') \frac{\hat{\gamma}(z) + \hat{\gamma}(z')}{z + z'} = m \left(\frac{\hat{\chi}(z) + \hat{\chi}(z')}{z + z'} - \hat{\chi}(z) \hat{\chi}(z') \right).$$

(4.31)

Comparando con el resultado (4.28) se muestra que el lado derecho satisface

$$m \left(\frac{\hat{\chi}(z) + \hat{\chi}(z')}{z + z'} - \hat{\chi}(z) \hat{\chi}(z') \right) =$$

$$= m \int_0^{\infty} dt_2 \int_0^{\infty} dt_1 \exp[-zt_2] \exp[-z't_1] (\chi(|t_2 - t_1|) - \chi(t_2) \chi(t_1)).$$

(4.32)

Comparando este resultado con la relación (4.27) nos da:

$$\begin{aligned}
 & \int_0^{t_2} ds \int_0^{t_1} ds' \chi(t_2 - s) \chi(t_1 - s') \mathcal{K}(|s - s'|) = \\
 & = m (\chi(t_2 - t_1) - \chi(t_2) \chi(t_1)). \quad (4.33)
 \end{aligned}$$

Por lo tanto, la matriz de autocorrelación de velocidad dada por la ec. (4.26) se escribe en la forma:

$$\chi_{i,j}(t_2 - t_1) = \frac{kT}{m} \chi(t_2 - t_1) \delta_{i,j}, \quad (4.34)$$

el cual muestra la estacionaridad del proceso, es decir, la matriz de autocorrelación se determina por la matriz de correlación para dos tiempos⁽¹³⁾ dada por:

$$\begin{aligned}
 & \left[\begin{array}{cc} \langle \langle \dot{v}_i(t_1) \dot{v}_j(t_1) \rangle \rangle & \langle \langle \dot{v}_i(t_1) \dot{v}_j(t_2) \rangle \rangle \\ \langle \langle \dot{v}_i(t_2) \dot{v}_j(t_2) \rangle \rangle & \langle \langle \dot{v}_i(t_2) \dot{v}_j(t_1) \rangle \rangle \end{array} \right] = \\
 & = \frac{kT}{m} \begin{bmatrix} \delta_{i,j} & \chi(t_2 - t_1) \delta_{i,j} \\ \chi(t_2 - t_1) \delta_{i,j} & \delta_{i,j} \end{bmatrix}. \quad (4.35)
 \end{aligned}$$

La matriz inversa está dada como

$$\frac{m}{kT} \frac{1}{1 - \chi^2(t_2 - t_1)} \begin{bmatrix} \delta_{i,j} & -\chi(t_2 - t_1) \delta_{i,j} \\ -\chi(t_2 - t_1) \delta_{i,j} & \delta_{i,j} \end{bmatrix}. \quad (4.36)$$

en la cual, el determinante está dado por

$$\left(\frac{m}{kT}\right)^6 (1 - \chi^2 |t_2 - t_1|)^{-3}. \quad (4.37)$$

Además, la función de distribución para dos tiempos es

$$P_2(\vec{v}_1, t_1; \vec{v}_2, t_2) = \left[(2\pi)^6 \left(\frac{kT}{m}\right)^6 (1 - \chi^2 |t_2 - t_1|)^3 \right]^{-1/2} \\ \times \exp \left[-\frac{1}{2} \frac{m}{kT} \frac{(\vec{v}_1 \cdot \vec{v}_1 + \vec{v}_2 \cdot \vec{v}_2 - 2\vec{v}_1 \cdot \vec{v}_2 \chi |t_2 - t_1|)}{1 - \chi^2 |t_2 - t_1|} \right]. \quad (4.38)$$

La distribución condicional para dos tiempos está dada por⁽¹⁹⁾

$$P_2(\vec{v}_1, t_1; \vec{v}_2, t_2) \equiv \frac{P_2(\vec{v}_1, t_1; \vec{v}_2, t_2)}{P(\vec{v}_1, t_1)}. \quad (4.39)$$

De las ecuaciones (4.25), y (4.34), haciendo $t_2 = t_1$, se obtiene⁽¹⁹⁾

$$P_1(\vec{v}_1, t_1) = \left[2\pi \frac{kT}{m} \right]^{-3/2} \exp \left[-\frac{P}{2kT} \vec{v}_1 \cdot \vec{v}_1 \right]. \quad (4.40)$$

Además, la probabilidad condicional P_2 está dada por:

$$\begin{aligned} \Pi_2(\vec{v}_1, t_1; \vec{v}_2, t_2) &= \left[2\pi \frac{kT}{M} (1 - \chi^2 |t_2 - t_1|) \right]^{-3/2} \\ &\times \exp \left[-\frac{1}{2} \frac{m}{kT} \frac{(\vec{v}_2 - \chi |t_2 - t_1| \vec{v}_1) \cdot (\vec{v}_2 - \chi |t_2 - t_1| \vec{v}_1)}{1 - \chi^2 |t_2 - t_1|} \right]. \end{aligned} \quad (4.41)$$

Si el proceso es Markoviano, entonces la ecuación de Chapman-Kolgomorov (ver cap. III), es

$$\Pi_2(\vec{v}_1, t_1; \vec{v}_3, t_3) = \int_{-\infty}^{\infty} \Pi_2(\vec{v}_1, t_1; \vec{v}, s) \Pi_2(\vec{v}, s; \vec{v}_3, t_3) d^3 \vec{v}. \quad (4.42)$$

Usando la relación (4.30), se requiere que se de la condición

$$\chi |t_3 - s| \chi |s - t_1| = \chi |t_3 - t_1|, \quad (4.43)$$

para todo $t \geq s \geq t$. Esto es, esta ecuación solamente se satisface⁽⁴³⁾ para

$$\chi |t - t'| = \exp(-|t - t'|D), \quad (4.44)$$

para una constante D. En conjunto con (4.22), se tiene el siguiente resultado

$$\hat{\gamma}(z) = mD, \quad (4.45)$$

el cual es equivalente con $\gamma(|t-s|) = mD \delta(t-s)$ en el límite

Markoviano. Sin embargo, (4.41) no se puede satisfacer, puesto que la relación (4.40) no describe un proceso de Markov.

De las ecuaciones (4.26) y (4.34) se obtiene la correlación de la velocidad $\vec{v}(t)$ dada como

$$\{ \langle \vec{v}_i(t) \vec{v}_j(s) \rangle \} = (kT/m) \chi(|t-s|) \delta_{ij}, \quad (4.46)$$

donde $\chi(t)$ se define a través de la transformada de Laplace $\hat{\chi}(z)$, en términos de la transformada de Laplace de $\gamma(t)$, $\hat{\gamma}(z)$

$$\hat{\chi}(z) \equiv (z + \hat{\gamma}(z)/m)^{-1}. \quad (4.47)$$

Haciendo $i=j$ y $t=s$, la relación (4.36) nos da

$$\{ \langle \vec{v}_i(t) \vec{v}_i(t) \rangle \} = kT/m, \quad (4.48)$$

por lo tanto, un cálculo similar a (3.57), da para la aceleración:

$$\frac{\sqrt{\langle \langle [\vec{v}_i(s+t) - \vec{v}_i(s)]^2 \rangle \rangle}}{t} \xrightarrow{t \rightarrow 0} \left[\left(\frac{2kT}{m} \right)^{1/2} \frac{\sqrt{1 - \chi(t)}}{t} \right]_{t \rightarrow 0} \quad (4.49)$$

Para ver que este resultado no diverge, se supone que la función $\gamma(|t-s|)$ esta dada por

$$\gamma(|t-s|) = \frac{\gamma}{\tau} \exp\left[-\frac{|t-s|}{\tau} \right]. \quad (4.50)$$

Entonces la transformada de Laplace de esta ecuación es

$$\hat{\chi}(z) = \frac{\gamma}{\tau} \left(\frac{1}{z + \tau^{-1}} \right). \quad (4.51)$$

Por lo tanto, se obtiene para $\chi(z)$ la siguiente expresión,

$$\hat{\chi}(z) = \frac{z + \tau^{-1}}{z^2 + z\tau^{-1} + (\gamma/m)\tau^{-1}}. \quad (4.52)$$

la cual es la transformada de Laplace de

$$\frac{1}{a-b} (a \exp[at] - b \exp[bt]) + \frac{1}{\tau} \frac{1}{a+b} (\exp[at] - \exp[bt]), \quad (4.53)$$

donde a y b están definidas por

$$a = -\frac{1}{2\tau} + \frac{1}{2\tau} \left(1 - \frac{4\gamma\tau}{m} \right)^{1/2}$$

y,

$$b = -\frac{1}{2\tau} - \frac{1}{2\tau} \left(1 - \frac{4\gamma\tau}{m} \right)^{1/2}. \quad (4.54)$$

en general es físicamente razonable suponer que $4\gamma\tau/m \ll 1$, lo cual es obvio cuando $t \rightarrow 0$. Utilizando (4.53) en la relación $\chi(t)$, de la ecuación (4.49) se tiene⁽¹³⁾

$$\frac{\sqrt{1 - \chi(t)}}{t} \xrightarrow{t \rightarrow 0} \left(\frac{\gamma}{2m\tau} \right)^{1/2}. \quad (4.55)$$

Por tanto, el resultado específico de $\chi(|t - s|)$ en la ecuación

(4.49) es

$$\lim_{t \rightarrow 0} \frac{\langle [v_1(s+t) - v_1(s)]^2 \rangle}{t} \longrightarrow \left(\frac{k T \gamma}{m^2 r} \right)^{1/2} \quad (4.56)$$

la cual claramente no diverge. Como en (3.57) la dependencia en r muestra el origen del comportamiento particular en el caso de la delta de Dirac. Podemos ver que en los procesos no-Markovianos no aparecen las dificultades matemáticas que surgen en los procesos Markovianos. Por lo tanto, la ecuación (4.29) de Langevin generalizada es consistente matemáticamente.

5. DISCUSIÓN

En este capítulo se discutió que para un proceso no-Markoviano las ecuaciones de difusión y de Langevin generalizadas son consistentes, a tiempos pequeños, en el sentido físico y matemático. Esta consistencia ha sido mostrada al suponerse, en ambos casos, que la función de correlación para la fuerza estocástica $F(t)$ no contiene una función delta, esto es, se tiene un proceso con "memoria" ó "ruido de color". En consecuencia, se deduce que la velocidad (para difusión) y la aceleración (para Langevin) tienen un valor finito.

El hecho esencial de que se considere un proceso no-Markoviano permite que los resultados obtenidos a partir de la ecuación de Langevin generalizada describan el comportamiento del sistema (partícula Browniana más baño térmico), el cual da una descripción más cercana a la realidad.

Esta discusión ha sido útil para distinguir la diferencia entre un proceso de Markov (cálculo de Stratonovich) y un proceso no-Markoviano (cálculo de Ito-Stratonovich). De igual manera, se ha recurrido a la formulación microscópica de la ecuación diferencial estocástica de Langevin. Este último aspecto reviste especial interés porque la formulación precisa de modelos microscópicos para deducir estas ecuaciones ha sido objeto de amplios estudios^(7,8,14,17). Aquí, se trató un modelo físico simple que permite, no obstante, formular conclusiones muy generales.

CONCLUSIONES GENERALES

Las ecuaciones diferenciales estocásticas han irrumpido recientemente en una gran cantidad de áreas de la ciencia^(1,2,9).

La investigación de éstas, en el aspecto formal y en las aplicaciones ha generado importantes resultados, que de una u otra forma, arrojan nueva luz para el entendimiento de los sistemas que se comportan de manera estocástica.

En este trabajo se discutió, de manera básica, la formulación fenomenológica de las ecuaciones de difusión y de Langevin en el contexto del movimiento Browniano. Las hipótesis de solución, conocidas como ruido blanco, resultaron ser inconsistentes cuando se toman en cuenta tiempos muy pequeños (10^{-13} seg $\approx t \ll m/\gamma$). Esto, motivó la introducción de ecuaciones y de hipótesis de solución más complejas; las ecuaciones de difusión y de Langevin generalizadas y los procesos Gaussianos no-Markovianos. Con esto, las inconsistencias, a tiempos cortos, fueron removidas. Para tiempos largos ($t \gg m/\gamma$) es posible recuperar la Markovianidad, y con ellas, las ecuaciones clásicas de difusión y Langevin.

La formulación del problema en términos de la ecuación de Langevin es conocida como el cálculo de Stratonovich; la formulación en términos de los procesos de Wiener es llamado el cálculo estocástico de Ito. En el caso lineal la total equivalencia entre estas formulaciones fue mostrada en el capítulo III. Para el caso no lineal (generalizado) esta equivalencia no ocurre pero la relación entre uno y otro resulta trivial⁽¹⁰⁾.

El tratamiento no-Markoviano a menudo es pasado por alto en la

mayoría de la literatura relacionada con el Movimiento browniano. Esto es debido por un lado a la enorme dificultad matemática del tratamiento no-Markoviano (cuyos aspectos esenciales fueron mostrados en el capítulo IV), y por otro lado a que el cálculo de ciertos promedios (por ejemplo correlaciones en la posición) dan resultados finitos y describen adecuadamente las respectivas situaciones físicas. A pesar de ello, vimos en el capítulo III que otros promedios (p. e., la correlación en la velocidad para la ecuación de Langevin) llevan a resultados en total contradicción con los resultados experimentales. El remedio a esta situación es considerar situaciones ligeramente más complicadas, es decir, procesos no-Markovianos.

Como punto final mencionamos que el tipo de procesos no-Markovianos, adecuados en el tratamiento de la ecuación de Langevin es motivo de investigación actual. Concluimos que éste es un tema en el cual estamos interesados en un futuro inmediato.

REFERENCIAS

1. E. Braun, Un Movimiento en Zigzag. Serie: La ciencia desde México Vol.13 SEP (1986).
2. B.J.L. Arauz Lara, Propiedades dinámicas de suspensiones de partículas Brownianas interactuantes. Tesis doctoral Cinvestav (1985).
3. J. McConnell, Rotational Brownian Motion and Dielectric Theory. Academic Press (1995).
4. W. Feller, Int. a la Teoría de Probabilidades. Limusa Vol.1 y Vol.2 (1985).
5. H. Haken, Synergetics (an Introduction). Springer-Verlag-Berlin (1983).
6. H. Goldstein, Mecánica Clásica. Aguilar (1977).
7. N.G. Van Kampen, Stochastic Processes in Physics and Chemistry North-Holland (1981).
8. R. Zwanzig, Problems in Nonlinear Transport Theory. Springer-Verlag (1980).
9. H. Risken, The Fokker-Planck Equation. Springer-Verlag Berlin (1984).
10. W. Pauli, Statistical Mechanics, Pauli Lectures on Physics vol 4, (1978)
11. L.D. Landau, Mecánica de Fluidos. Reverte (1976).
12. R.F. Fox, J. of Stat. Phys. Vol. 46. nos.56 (1987).
13. R.F. Fox, Phys. Rep. , 48:179 (1978).
14. H. Mori, Prog. Theor. Phys., 53, 1617 (1975).
15. H.B. Callen, Thermodynamics. John Wiley (1960).

16. A. Einstein, Investigation on the Theory of the Brownian Movement. Dover (1956).
17. G.W. Ford, M.Kac, P. Mazur, J. of Math. Phys., Vol 6. no.4 (1965).
18. S. Pick, Physica 103A (1980) 630-632.
19. P. Pusey and R. Tough, Dynamic Light Scattering and Velocimetry: Applications of Photon Correlation Spectroscopy. R. Pecora, (1987).