

102

PLANTEAMIENTO DE PROBLEMAS  
EN CIENCIAS EXACTAS E INGENIERIA

por **102**

DIEGO BRICIO HERNANDEZ

1989

CIMAT

Apartado Postal 402  
36000 Guanajuato, Gto.  
M é x i c o

Tel: (473) 2 02 58 y 2 58 05  
Telefax: (473) 2 57 49  
E-Mail: CIMAT@UNAMVM1.BITNET



## CONTENIDO

	Pag.
Prefacio.	
Conferencia 1.	
Un Ejemplo: La Mecánica de Partículas.	1
1.1 La representación del tiempo y del espacio.	1
1.2 Partículas en un campo de fuerza.	6
Conferencia 2.	
Los Elementos del Modelado Matemático.	11
2.1 La cuerda vibrante y los componentes de un modelo.	11
2.2 Un ejemplo: la conducción de calor en sólidos.	19
Conferencia 3.	
El Carácter Relativo del Determinismo.	27
3.1 Determinismo e indeterminismo.	27
3.2 Un ejemplo: la desintegración radiactiva.	31
Conferencia 4.	
Lo Microscópico y lo Macroscópico.	40
4.1 La difusión molecular.	40
4.2 Reacción y difusión en dimensión uno.	47
Conferencia 5.	
Algunas Consideraciones Metodológicas.	56
5.1 Reglas para el planteamiento de problemas.	56
5.2 Validación de los modelos matemáticos.	65
Apéndice.	
El Modelo Clásico de la Gravitación Universal.	71
A.1 La fuerza de atracción newtoniana.	71
A.2 Las leyes de Kepler.	77
Bibliografía.	85

## PREFACIO.

Esta monografía contiene cinco conferencias dictadas por el autor ante estudiantes y profesores de la Universidad de Sonora (Hermosillo), durante la última semana de marzo de 1989. A dichas conferencias asistió una audiencia muy variada, con miembros del Departamento de Matemáticas y de la Escuela de Ciencias Químicas, así como estudiantes de las licenciaturas en Matemáticas, Geología, Ingeniería Industrial y de Sistemas e Ingeniería Química.

Figura en ella un apéndice dedicado a la ley de la Gravitación Universal, en su versión newtoniana, incluyendo su aplicación a la descripción de las órbitas planetarias. Este último material fué expuesto ante estudiantes de la Maestría en Educación Matemática de la Universidad de Guanajuato, en abril de 1989. Por otro lado, la parte referente a los modelos determinísticos a base de ecuaciones diferenciales ordinarias fué impartido como cursillo para estudiantes de licenciatura durante el XVIII Congreso de la Sociedad Matemática Mexicana, celebrado en Mérida, Yucatán en Noviembre de 1984.

Para ser justos, debemos decir que estas notas constituyen una parte del material preparado durante la impartición del curso Modelos Matemáticos, en 1983 y 1985, como parte de la Licenciatura en Matemáticas que ofrece la Universidad Autónoma Metropolitana (Iztapalapa), en la Cd. de México. Faltaron de incluir aquí un buen número de ejemplos, así como algunas técnicas de aproximación para la solución de ecuaciones algebraicas y diferenciales, y técnicas de simulación de Monte Carlo. Aparte de los ejemplos, el material faltante ha sido ya cubierto por el autor al menos en parte, en las monografías "Análisis Numérico" y "El método de monte Carlo" que se citan en la Bibliografía.

La selección de material obedece a una vieja preocupación del autor por entender la relación de las Matemáticas con el mundo real. En particular, entender cómo se originan los conceptos matemáticos a partir de situaciones concretas, mediante lo que comunmente se denomina "planteamiento de problemas". Para ello, se presenta un buen

número de ejemplos, tomados de áreas de la Física Clásica como la Mecánica de Partículas y del Medio Continuo, la Conducción de Calor en Sólidos y la teoría elemental de los Circuitos Eléctricos; se dan ejemplos también de otras áreas, mejor identificadas con la Ingeniería Química, como son los fenómenos de Difusión y Reacción. Sin embargo, ésta no es una colección de ejemplos (clásicos y modernos), así como tampoco es una colección de técnicas de solución de ecuaciones. Su objetivo es más bien metodológico, y los ejemplos sirven para motivar e ilustrar los conceptos importantes.

Uno de los conceptos que más se ha querido enfatizar es el de DETERMINISMO, que resulta fundamental para el planteamiento de problemas. En dos conferencias (la 3 y la 4) se ha buscado ilustrar el carácter relativo de este aspecto, que más bien es una propiedad de los modelos matemáticos de la realidad y no de ésta. Previamente, en la segunda conferencia se dan los elementos que constituyen un modelo determinístico, digamos uno a base de ecuaciones diferenciales. Dos elementos muy importantes de todo modelo son las representaciones matemáticas del espacio y del tiempo, y en la primera conferencia nos ocupamos de ellos. Como subproducto de todo lo anterior, en la primera parte de la última conferencia se formulan algunas "reglas" para el planteamiento de problemas, usando la metodología de "elegir primero una ley de conservación y luego una forma de calcular los distintos términos de la misma". La última conferencia presenta además algunas consideraciones metodológicas sobre la validación de los modelos matemáticos; las técnicas de la inferencia estadística resultan fundamentales para ello. Finalmente, el apéndice ofrece un ejemplo muy completo de un modelo matemático determinístico y algunas de sus consecuencias. Véase el inicio de cada una de las seis partes; ahí se da una descripción específica del tema tratado.

Algunos libros de ejemplos de modelado matemático que vale la pena recomendar son [Noble, 1967], [Haberman, 1977], [Andrews-McClone, 1980], [Boyce, 1981], [Burghes-Borrie, 1981], [Harte, 1985] y [Beltrami, 1987], entre otros. Un libro excelente orientado hacia las técnicas de solución de las ecuaciones es [Strang, 1986]; dos libros de "técnicas de simulación" que vale la pena mencionar son

[Ferziger,1981] y [Yakowitz,1977], para no hablar del monumental [Press et al.,1988]. Libros donde se desarrollan con gran amplitud los temas matemáticos aquí tratados (modelos determinísticos), en especial la parte de reacción y difusión, son [Aris,1975], [Fife,1979] y, con alcance enciclopédico y alto nivel matemático, [Smoller,1983]; para la parte estocástica, véanse [Nelson,1967] y [Schuss,1980]. Finalmente, dos libros sobre modelado matemático que sobresalen del resto, pues contienen una multitud de ejemplos interesantes que combinan tanto técnicas de solución como aspectos metodológicos, son [Aris,1978] y [Lin-Segel,1974]. Por otro lado, todo buen libro de Física o Ingeniería sirve para ilustrar este material.

En todo el trabajo se ha buscado ilustrar y motivar las ideas sin complicar excesivamente la presentación desde el punto de vista matemático. El nivel matemático presupone un dominio de las Ecuaciones Diferenciales al nivel de [Boyce-DiPrima,1977], [Braun,1984] o [Simmons,1972], de la Probabilidad al nivel de [Feller,1968], y del Cálculo Avanzado al nivel de [Marsden-Tromba, 1987]. Fuera de las Matemáticas, conviene contar con un conocimiento de la Mecánica y Electricidad al nivel de [Resnick-Halliday,1980]. Finalmente, los conceptos de Termodinámica y Fenómenos de Transporte que aquí se manejan pueden consultarse en [Callen,1960] y [Bird et al.,1961], respectivamente. Se recomienda la referencia [Cárdenas et al.,1978], donde se desarrollan algunos temas de Cálculo Avanzado muy en relación con su aplicación termodinámica.

Las figuras, fórmulas, etc. se numeran internamente a cada sección. En cuanto a las referencias internas, la fórmula  $k$  de la sección  $n$  de la conferencia  $m$  se llama  $(k)$  en la misma sección,  $(n.k)$  en otra sección de la misma conferencia y  $(m.n.k)$  cuando se le menciona desde otra conferencia. Finalmente, se agradece al Dr. Alfinio Flores por sus comentarios y sugerencias sobre algunos puntos del Apéndice, y a la Srta. Elizabeth Esparza por su eficiente labor de mecanografiado.

D.B.H., Valenciana, Julio de 1989.

## CONFERENCIA 1.

### UN EJEMPLO: LA MECANICA DE PARTICULAS.

En esta conferencia se examina el modelo matemático por excelencia, la *Mécanica Clásica*, en lo que se refiere al movimiento unidimensional de partículas puntuales. En la primera sección se discuten las diferentes posibilidades para representar el espacio y el tiempo en términos matemáticos. Se concluye que, a pesar de lo limitado de la experiencia, que sólo permite efectuar un número finito de mediciones, conviene representar tanto el espacio como el tiempo mediante los puntos de la recta real (o de un espacio euclidiano de dimensión adecuada). En la segunda sección se dan algunos ejemplos de partículas que se mueven en un campo de fuerza, con énfasis en los aspectos relacionados con la conservación de la energía mecánica.

#### 1.1 La representación del tiempo y del espacio.

Problema: "Una piedra cae y se desea predecir su trayectoria".

Comencemos por precisar lo que vamos a entender por *trayectoria*: se trata de especificar, en cada instante de tiempo, la posición de la piedra. En particular, nos damos cuenta de que para hablar de trayectoria hay que hacerlo en términos de *tiempo* y *espacio*, por lo que convendrá precisar primero estos conceptos. Lo haremos desde una perspectiva empirista, tomando en cuenta los instrumentos de medición a nuestro alcance, que suponemos funcionan correctamente.

En la práctica, medimos el tiempo mediante un reloj, que nos da un cierto grado de precisión. Si ésta es de más de 1 seg., lo más natural es tener como representante del tiempo al conjunto  $\mathbb{Z}$  de los números enteros, y decir que  $\{0\}$  es el *presente*,  $\{\dots, -2, -1\}$  el *pasado* y  $\{1, 2, \dots\}$  el *futuro*. Además de representar las restricciones que nos imponen los instrumentos de medición, esta elección de representante matemático para el tiempo refleja el orden lineal entre instantes, mediante la relación  $\leq$  en  $\mathbb{Z}$ : dados dos instantes  $s, t \in \mathbb{Z}$ , decimos que  $s$  es *anterior* a  $t$  si  $s \leq t$ . Esta elección de representante matemático para

el tiempo nos lleva a construir un modelo a *tiempo discreto*.

Pero supongamos que hemos elegido un reloj más fino, que nos da una precisión mayor, digamos de fracciones de segundo. Más aún, tomamos en cuenta la posibilidad de que refinamientos tecnológicos sucesivos nos den relojes capaces de distinguir entre instantes cada vez más próximos. Aún cuando es natural que tales refinamientos tengan un límite, al menos en principio podemos continuar el proceso indefinidamente. Es claro que si la descripción del tiempo es discreta no podemos tomar en cuenta esa posibilidad; necesitamos un representante matemático del tiempo en el cual sea cierto que "entre dos instantes cualesquiera hay otro". Es decir, se trata de conservar la posibilidad de subdividir el tiempo en lapsos arbitrariamente cortos, sin perder la relación de precedencia temporal.

Lo anterior se puede lograr tomando como representante del tiempo al conjunto  $\mathbb{Q}$  de los números racionales con su relación  $\leq$ , en lugar de  $(\mathbb{Z}, \leq)$ . En efecto, dados dos instantes  $s, t \in \mathbb{Q}$ ,  $s \leq t$  se interpreta diciendo que  $s$  es anterior a  $t$ , y si  $s < t$  entonces  $\frac{1}{2}(s+t) \in (s, t)$ . Así pues, el tiempo resulta infinitamente divisible y puede darse  $\{0\}$  como el presente, con el pasado (resp. futuro) representado mediante el conjunto de los racionales negativos (resp. positivos).

Consideraciones análogas se aplican a la descripción del espacio: conviene dividir el espacio en fragmentos cada vez más pequeños, según la precisión que nos permitan nuestros instrumentos de medición, para lo cual conviene tomar como su representante también al conjunto  $\mathbb{Q}$  de los números racionales, a sabiendas de que será posible dar una descripción unidimensional del fenómeno. En este caso la relación  $\leq$  en  $\mathbb{Q}$  nos permite ordenar las posiciones recorridas por la piedra a lo largo de su caída.

Conviene señalar que la visión que lleva a describir el movimiento de la piedra en forma unidimensional supone que basta especificar la posición de una de las partes de la piedra para tener una descripción completa del movimiento. Para simplificar las cosas, representamos a la piedra como un punto material, sin dimensiones

físicas: la posición de la piedra es la de dicho punto material. Esto, desde luego, es una idealización. Más aún, despreciamos también cualquier efecto que puedan tener otros cuerpos sobre la piedra, incluso el viento, la rotación de la Tierra, etc. Tampoco interesa la forma de la piedra, ni su material.

Las simplificaciones anteriores son idealizaciones de un observador a escala humana, parado sobre la superficie de la Tierra y provisto de sus instrumentos de medición (reloj y cinta métrica). O quizá, además del reloj nuestro observador dispone de una cámara fotográfica que produce imágenes intermitentes pero próximas en el tiempo. Según sus observaciones, la piedra cae a lo largo de una recta que pasa por el centro de la Tierra, por lo que efectivamente bastará con especificar en cada instante la altura a la que se encuentra la piedra en relación con el piso para describir su caída.

Resumiendo: convenimos por ahora en representar tanto el tiempo como el espacio mediante números racionales; en cada instante, la piedra ocupa un solo punto del espacio y su único atributo de interés es la altura a la que se halla en relación al piso.

Para un experimento dado, obtenemos una tabla de valores del tipo de

Tiempo:	$t_1$	$t_2$	...	$t_m$
Altura:	$x_1$	$x_2$	...	$x_m$

donde

$$0 =: t_0 < t_1 < \dots < t_m =: T,$$

en tanto que

$$x_0 > x_1 > \dots > x_m$$

si  $x_0$  es la altura inicial.

Podemos refinar nuestras mediciones y entonces la tabla de resultados será

Tiempo:	$t'_0$	$t'_1$	$t'_2$	...	$t'_n$
Espacio:	$x'_0$	$x'_1$	$x'_2$	...	$x'_n$

donde  $t'_0 = t_0$ ,  $t'_n = T$ , con  $n > m$ , por lo que los lapsos entre mediciones son ahora más cortos. Vemos, además, la caída de la piedra y tenemos la convicción de que ésta "pasa por todos los puntos de la recta a lo largo de la cual cae". Todo esto nos lleva a postular la existencia de una función  $x: \mathbb{Q} \rightarrow \mathbb{Q}$ , de la cual sólo conocemos los valores registrados en el experimento, a los cuales interpola, de tal manera que

$x(t) :=$  altura de la piedra en el instante  $t$ .

Interesa además hablar de cantidades dinámicas que relacionan espacio y tiempo. Por ejemplo, la *velocidad media* entre dos instantes  $t$  y  $t+h$  es

$$v_t(h) := \frac{x(t+h) - x(t)}{h};$$

la *aceleración media* entre los instantes  $t$ ,  $t+h_1$  y  $t+h_1+h_2$  será

$$a_t(h_1, h_2) := \frac{v_{t+h_1}(h_2) - v_t(h_1)}{h_2}$$

Conceptualmente al menos, puede pensarse en refinar las mediciones, tomando instantes de tiempo cada vez más próximos (es decir,  $h$  cada vez más pequeño) y considerar el comportamiento de estas cantidades dinámicas en función de  $h$ . Por ejemplo, tomemos la velocidad media  $v_t(h)$ , para  $t$  fijo, que empíricamente se puede determinar para  $h > \varepsilon$ , donde  $\varepsilon$  es la precisión de nuestro reloj. Podemos graficar  $v_t(h)$  con respecto a  $h$  y estudiar que sucede al "extrapolar  $h$  a cero", según se indica en la figura 1.

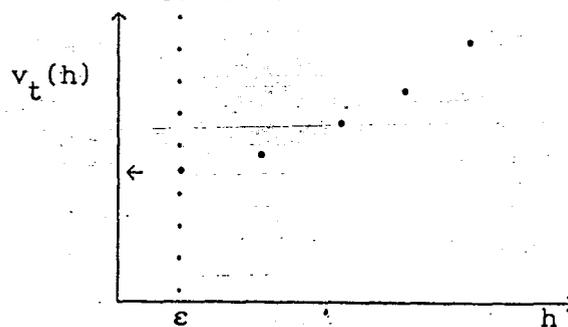


Fig. 1: Comportamiento de la velocidad media.

La contrapartida matemática de esta operación de laboratorio es la de "tomar el límite cuando  $h$  tiende a cero", es decir el límite

$$\lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h};$$

si existe, dicho límite se llamará la *velocidad instantánea* en el instante  $t$ , y será denotado como  $v(t)$ . Análogamente, la *aceleración instantánea*  $a(t)$  se define mediante el límite

$$\lim_{h \rightarrow 0} \frac{v(t+h) - v(t)}{h}$$

Así pues, las primeras dos derivadas de la trayectoria  $x$  nos dan la velocidad y la aceleración de la partícula.

Sin embargo, puede suceder que para alguna sucesión  $\{h_n\}$  en  $\mathbb{Q}$ , con  $h_n \downarrow 0$ , la sucesión  $\{v_n\}$  de números racionales dada por

$$v_n := \frac{x(t+h_n) - x(t)}{h_n}$$

arroje valores cada vez más parecidos entre sí y que, sin embargo, no converja a ningún racional. Para quitar estos problemas, conviene tomar como representante del tiempo (y también del espacio) al conjunto  $\mathbb{R}$  de los números reales. Este es suficientemente rico como para no dar lugar a anomalías como la ya señalada: basta con que

$$\lim_{m, n \rightarrow \infty} |v_m - v_n| = 0$$

para asegurar que el límite  $\lim_{n \rightarrow \infty} v_n$  existe (como número real). Esta propiedad es la *completitud* o *completez* de la recta real [Rudin, 1976].

Así pues, tomaremos como representante del tiempo al conjunto  $\mathbb{R}$  de los números reales, con su relación  $\leq$  y la operación de  $+$ ; si el presente es  $\{0\}$ , el pasado es  $(-\infty, 0)$  y el futuro es  $(0, +\infty)$ .

La trayectoria de la partícula está dada por una función  $x: \mathbb{R} \rightarrow \mathbb{R}$ .

## 1.2 Partículas en un campo de fuerza.

Dado que empíricamente podemos determinar velocidades y aceleraciones instantáneas, la trayectoria  $x: \mathbb{R} \rightarrow \mathbb{R}$  debe ser al menos dos veces continuamente diferenciable. Además, debe cumplir las *condiciones iniciales*

$$x(0) = x_0, \quad \dot{x}(0) = 0, \quad (1)$$

puesto que la piedra se soltó desde una altura  $x_0$  respecto al piso.

Bajo condiciones normales, la piedra se suelta y queda bajo la acción de la fuerza de gravedad, por lo cual la aceleración será constante e igual a la de la gravedad; es decir  $x$  es solución de la ecuación diferencial

$$\ddot{x} = -g, \quad (2)$$

donde  $g = 9.81 \text{ cm/seg}^2$ . De ser así, la trayectoria es

$$x(t) = x_0 - \frac{1}{2} g t^2 \quad (3)$$

### Ejemplo 1.

Si, en lugar de soltar la piedra a partir del reposo, se le da una velocidad inicial  $v_0$ , entonces las condiciones iniciales son

$$x(0) = x_0, \quad \dot{x}(0) = v_0, \quad (4a)$$

por lo que la trayectoria queda determinada como

$$x(t) = x_0 + v_0 t - \frac{1}{2} g t^2 \quad (4b)$$

### Ejemplo 2.

Supongamos que la partícula está cargada eléctricamente, con carga  $q$  y masa  $m$ , y que es atraída por otra carga puntual  $Q$ , fija a

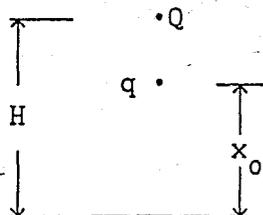


Fig. 2: La partícula cargada que cae en un campo eléctrico.

una altura  $H$  sobre el nivel del piso a lo largo de la vertical. De acuerdo a la ley de Coulomb, el efecto electrostático se traduce en una aceleración adicional de

$$\frac{k}{m} \cdot \frac{q Q}{[H-x(t)]^2} \frac{\text{cm}}{\text{seg}^2}$$

en el instante  $t$ , donde  $k$  es una constante de proporcionalidad dada por las propiedades dieléctricas del medio. Entonces, la ecuación de movimiento de la partícula es

$$\ddot{x} = -g + \frac{k}{m} \cdot \frac{qQ}{(H-x)^2} \quad (5)$$

La trayectoria queda de nuevo determinada mediante las condiciones iniciales (4a), si  $x_0 \neq H$ .

### Ejemplo 3.

Generalizando los ejemplos anteriores, supongamos que la partícula se mueve bajo la acción de un campo de fuerza dado por un potencial total  $U$ ; entonces, la fuerza que actúa sobre la partícula es  $-U'(x)$ , por lo que la ecuación de movimiento es

$$m\ddot{x} = -U'(x).$$

La función energía total, también llamada función *hamiltoniana* es

$$H(x, v) := \frac{1}{2} mv^2 + U(x)$$

y, a lo largo del movimiento, sus valores son

$$E(t) := H(x(t), \dot{x}(t)).$$

Así pues, la tasa de variación de la energía total es

$$\frac{dE}{dt} = m\dot{x} \left[ -\frac{U'(x)}{m} \right] + U'(x)\dot{x} = 0.$$

En otras palabras, la energía total se conserva a lo largo del movimiento: *en el plano de fase* (la fase es la pareja posición - velocidad), *las trayectorias de la partícula están contenidas en las curvas de nivel de la función hamiltoniana.*

En particular, para la piedra en caída libre  $U(x)=mgx$  y la ley de conservación anterior dice que

$$m \frac{v(t)^2}{2} + mgx(t) = mgx_0, \quad t \geq 0.$$

Al momento de tocar el suelo ( $x=0, t=t_0$ ) la velocidad de la piedra es entonces  $v(t_0) = \sqrt{2gx_0}$ . Para saber cuando ocurrirá eso, de (3) obtenemos la respuesta  $t_0 = \sqrt{2x_0/g}$ .

Para el ejemplo 2,

$$U(x) = mgx + \frac{kqQ}{H-x};$$

en el plano  $x-v$ , el movimiento de la partícula se da a lo largo de las curvas de nivel de la función hamiltoniana

$$\frac{1}{2}mv^2 + mgx + \frac{kqQ}{H-x}.$$

Para las condiciones iniciales (4a), posición y velocidad están ligadas mediante la ecuación

$$\frac{1}{2}m(v^2 - v_0^2) + mg(x - x_0) + kqQ \left( \frac{1}{H-x} - \frac{1}{H-x_0} \right) = 0. \blacksquare$$

#### Ejemplo 4.

En el ejemplo 3, supongamos que la partícula se encuentra sumergida en un fluido (agua, aire, aceite, etc.). La acción del fluido se manifiesta mediante una fuerza que se opone al movimiento de la partícula, y que sólo actúa si ésta se mueve.

Sea  $f: \mathbb{R} \rightarrow \mathbb{R}$  tal fuerza viscosa con  $f(v) :=$  valor de dicha fuerza cuando la partícula se mueve con velocidad  $v$ ; entonces, debe cumplirse que  $f(0) = 0$ . Además, la gráfica de  $f$  debe encontrarse en los cuadrantes segundo y cuarto, es decir  $vf(v) \leq 0$  [véase la figura 3], pues la fuerza viscosa se opone al movimiento de la partícula.

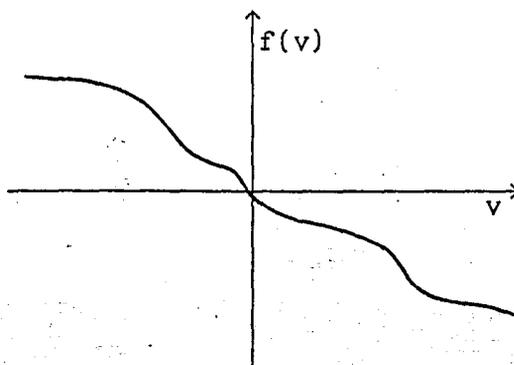


Fig. 3: La forma típica de una fuerza viscosa.

Si, además de esta fuerza, actúa sobre la partícula un campo de fuerza dado mediante un potencial  $U$ , la ecuación de movimiento es

$$m\ddot{x} = -U'(x) + f(\dot{x})$$

Veamos qué sucede a la energía total a lo largo de las trayectorias de

este sistema dinámico:

$$\frac{dE}{dt} = m\dot{x} \left[ \frac{-U'(x) + f(\dot{x})}{m} \right] + U'(x)\dot{x}$$

e.d.

$$\frac{dE}{dt} = f(\dot{x}(t))\dot{x}(t) \leq 0.$$

Así pues, la energía no puede aumentar y de hecho disminuye si la partícula se mueve. Es en este sentido que la fuerza viscosa es *dissipativa*. ■

#### Ejemplo 5.

Supongamos ahora que la fuerza viscosa es lineal, es decir,  $f(v) = -\beta v$ , con  $\beta > 0$ . Supongamos además que sobre la partícula actúa una fuerza variable  $F$ , que no depende de la posición ni de la velocidad de la misma, sólo del tiempo; supongamos que  $F: \mathbb{R} \rightarrow \mathbb{R}$  es continua. Entonces, la ecuación de movimiento es

$$m\ddot{x} = -U'(x) - \beta\dot{x} + F(t)$$

En el espacio fase, la ecuación de movimiento es

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= -\frac{1}{m} U'(x) - \frac{\beta}{m} v + \frac{1}{m} F(t) \end{aligned}$$

y un cálculo sencillo muestra que

$$\frac{dE}{dt} = -\beta v^2 + F(t)v$$

Integrando entre  $t_1$  y  $t_2$ , la relación anterior se transforma en

$$\Delta E = Q + W,$$

donde

$$Q := -\int_{t_1}^{t_2} \beta v(t)^2 dt,$$

$$W := \int_{t_1}^{t_2} F(t)v(t) dt.$$

Obsérvese que  $W$  es el trabajo realizado sobre la partícula por la fuerza externa  $F$ , en tanto que  $Q$  es el calor absorbido por la partícula, todo durante el intervalo  $[t_1, t_2]$ . Claramente  $Q < 0$ , lo cual indica que dicho calor realmente se pierde, es decir, es energía que cede la partícula al medio circundante.

La relación

$$\Delta E = Q + W$$

expresa la ley de la Conservación de la Energía en forma cercana a la Primera Ley de la Termodinámica.

## CONFERENCIA 2.

### LOS ELEMENTOS DEL MODELADO MATEMATICO.

El objetivo de esta segunda conferencia es presentar los elementos que constituyen un modelo matemático construido bajo una óptica determinística. Se comienza por poner en evidencia dichos elementos en la sección 1, mediante el análisis de una cuerda vibrante. Se da una presentación crítica del tema, enfatizando claramente las diferentes hipótesis y cómo se traducen éstas en diferentes modelos para la cuerda. En la segunda sección se da un modelo matemático para la conducción de calor en una barra metálica. Al hacerlo se busca ilustrar tanto el uso de la ley de conservación de la energía como el papel que juegan las hipótesis y los diferentes elementos del modelado matemático.

#### 2.1 La cuerda vibrante y los componentes de un modelo.

En la sección anterior hemos construido modelos matemáticos de situaciones sencillas, que se refieren al movimiento de una partícula bajo la acción de diversos tipos de fuerza externa. En cada caso, el principal atributo de interés es la posición de la partícula, pero suponemos que también se puede medir su velocidad; así pues, el atributo combinado de interés es la fase (posición-velocidad). En cada caso, se observa un fenómeno porque el atributo de interés cambia con el tiempo.

Igual que en la primera sección se argumentó hasta usar el conjunto  $\mathbb{R}$  de números reales para medir el tiempo y el espacio, podemos defender la elección de  $\mathbb{R}^2$  como conjunto de valores de los atributos de interés. En general, si se desea estudiar  $n$  atributos, se representarán sus valores mediante los elementos de un conjunto  $A \subset \mathbb{R}^n$ .

En cada uno de los casos se estudia el movimiento de una partícula que se supone es una masa puntual. De manera análoga podría estudiarse el movimiento de un sistema de  $N$  partículas, cada una un punto material; cada partícula tiene una posición y una velocidad y

durante el desarrollo del fenómeno los atributos bajo estudio (posición y velocidad) cambian no sólo en el tiempo, sino que cambian de partícula a partícula en cada instante. Así pues, los atributos se distribuyen en el tiempo y en el espacio durante el desarrollo del fenómeno. El espacio está representado en este caso mediante un conjunto finito con  $N$  elementos, digamos  $\{1, \dots, N\}$ ; se dice entonces que se trata de un modelo a *espacio discreto*.

Sin embargo, hay otras situaciones en las que el espacio se considera *continuo*. Por ejemplo, si cae un objeto de dimensiones finitas y queremos tener cuenta del movimiento de cada una de sus partes, habrá que representarlo mediante algún subconjunto  $S$  de  $\mathbb{R}^3$ . Si el cuerpo es suficientemente rígido esto no es necesario, pues dada la fase de uno cualquiera de sus puntos la de los demás queda fijada automáticamente. Si, por el contrario, el cuerpo en cuestión no es rígido (como por ejemplo la membrana de un tambor) entonces se debe por fuerza dar una representación continua del espacio y debe estudiarse el fenómeno en tanto que redistribución espacio-temporal de la fase.

Algunas veces, mediante consideraciones de simetría se podrá simplificar la representación del espacio, y entonces  $S$  será un subconjunto de  $\mathbb{R}^1$  o  $\mathbb{R}^2$ . En general, en un modelo a espacio continuo éste se representará mediante un subconjunto  $S \subset \mathbb{R}^d$ , con  $1 \leq d \leq 3$ ; típicamente se querrá calcular derivadas de funciones definidas en  $S$ , por lo que convendrá que  $S$  sea abierto.

#### Ejemplo.

"Una cuerda uniforme y ligera se mantiene razonablemente tensa sujetándola por los extremos. En un momento dado se pulsa la cuerda y comienza a vibrar, sin que se suelte. Se busca describir las vibraciones resultantes".

Comencemos por precisar las condiciones bajo las que se desarrolla el fenómeno, pues hay demasiados aspectos que deben especificarse todavía. En efecto, el fenómeno puede ser sumamente complejo si admitimos la posibilidad de movimientos verdaderamente

tridimensionales, es decir, que saquen a la cuerda del plano vertical que la contiene inicialmente (véase la figura 1).

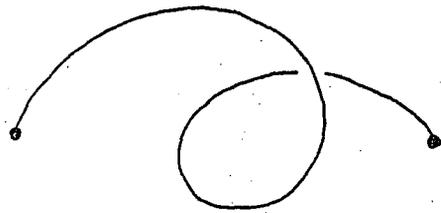


Fig. 1: Una forma posible para una cuerda que oscile en tres dimensiones.

Supondremos entonces que las oscilaciones de la cuerda están confinadas a dicho plano. Más aún, la cuerda no se dobla, como en la figura 2; es decir, supondremos que cada punto de la cuerda oscila a lo largo de la vertical que lo contiene al inicio. Lo anterior equivale a suponer que la cuerda se mantiene tensa en todo momento.



Fig. 2: Una forma posible para una cuerda que oscile en un plano vertical pero que se dobla.

Finalmente, consideremos a la cuerda como un objeto unidimensional, ignorando la diferencia en posición y velocidad (fase) de los puntos en una sección transversal dada de la misma. Véase la figura 3.

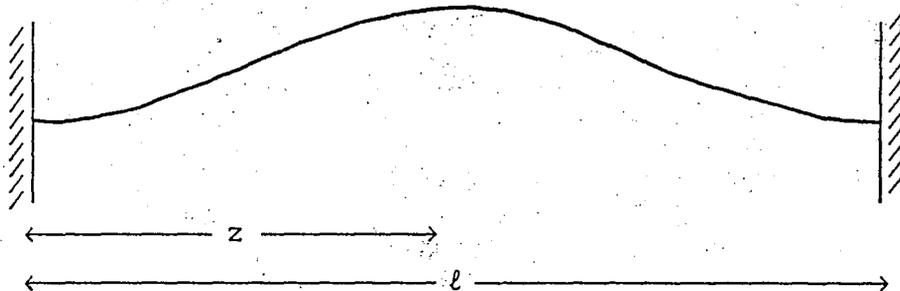


Fig. 3: Forma de una cuerda tensa que oscila en un plano vertical y cuyos extremos se mantienen fijos.

De acuerdo con lo anterior, podemos representar el espacio mediante  $S := (0, \ell)$ , donde  $\ell$  es la longitud de la cuerda tensa en reposo; de nuevo tomaremos como representante del tiempo a la recta

real. Dado que interesa llevar un registro de la posición y velocidad de cada uno de los puntos de la cuerda, midiendo la posición respecto al estado de equilibrio, resulta que conviene tomar como espacio de atributos a  $\mathbb{R}^2$ .

Las hipótesis que hemos introducido implican la existencia en cada instante de una función  $x:S \rightarrow A$  ("cada punto de la cuerda tiene una fase bien definida"). Sea  $x(z) = (u(z), v(z))$ , con  $u$  el perfil de posición y  $v$  el de velocidad. A menos que la cuerda se rompa, el perfil de posición debe ser una función continua y lo mismo el de velocidad. Además, debe cumplirse que

$$u(0)=0, u(\ell)=0 \quad (1)$$

para cada perfil  $x$ , puesto que los extremos de la cuerda permanecen fijos.

Con objeto de especificar plenamente el conjunto  $E$  de los perfiles admisibles debemos detallar aún más la descripción del sistema en cuestión. Comencemos por suponer que la tensión se reparte uniformemente a lo largo de la cuerda, con una magnitud  $T$  unidades de fuerza (digamos  $T$  dinas, en el sistema cgs). Para cada  $z \in S$ , sea  $\theta(z)$  el ángulo que forma la tensión con respecto a la horizontal; sean  $z, z+h \in S$ , con  $h > 0$ , como se muestran en la figura 4.

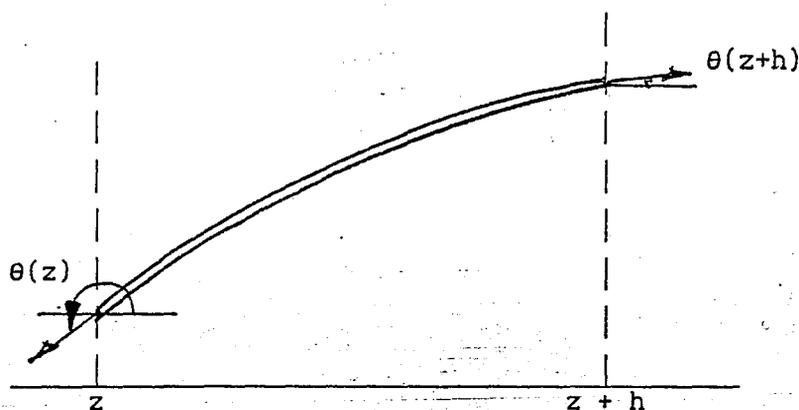


Fig. 4: Análisis de un segmento de la cuerda vibrante.

La componente vertical neta de la fuerza de tensión sobre el tramo de la curva proyectada sobre  $[z, z+h]$  es

$$T \operatorname{sen}\theta(z+h) + T \operatorname{sen}\theta(z),$$

es decir

$$T \frac{\operatorname{tg}\theta(z+h)}{\sqrt{1+\operatorname{tg}^2\theta(z+h)}} - T \frac{\operatorname{tg}\theta(z)}{\sqrt{1+\operatorname{tg}^2\theta(z)}},$$

pues  $\operatorname{sen}\theta(z) < 0$ ,  $\operatorname{tg}\theta(z) > 0$ , según la figura 4.

Supongamos ahora (por mera conveniencia) que  $u$  es diferenciable, de tal manera que

$$u' = \operatorname{tg}\theta$$

y entonces la tensión resultante es

$$T \left\{ \frac{u'(z+h)}{\sqrt{1+u'(z+h)^2}} - \frac{u'(z)}{\sqrt{1+u'(z)^2}} \right\}.$$

Si suponemos, además, que  $u$  es dos veces continuamente diferenciable, el Teorema del Valor Medio nos autoriza a expresar dicha fuerza como

$$T \frac{d}{dz} \left[ \frac{u'}{\sqrt{1+u'^2}} \right] (z+\alpha h)h, \quad (2)$$

para algún  $\alpha \in (0,1)$ . Conviene entonces tomar como perfiles de posición al conjunto de todas las funciones  $u: [0,\ell] \rightarrow \mathbb{R}$ , continuas en  $[0,\ell]$ , dos veces continuamente diferenciables en  $(0,\ell)$  y que satisfacen (1). Para tales perfiles, la expresión (2) nos permite calcular la tensión que actúa sobre el tramo de cuerda que en reposo ocupa la porción de  $z$  a  $z+h$ . En cuanto a los perfiles de velocidad aceptables, serán aquellas funciones continuas  $v: [0,\ell] \rightarrow \mathbb{R}$ , que satisfacen además  $v(0)=0$ ,  $v(\ell)=0$  ("los extremos no se mueven").

Ahora bien, a cada instante  $t \in \mathbb{R}$  le debe corresponder un perfil de atributos, por lo que debe existir una función  $\omega: \mathbb{R} \rightarrow E$ , llamada la *trayectoria* de la cuerda en el espacio de perfiles. Introduzcamos una notación más convencional, definiendo

$$x(t,z) := \omega(t)(z),$$

$$(u(t,z), v(t,z)) := x(t,z).$$

Nótese que  $v = \frac{\partial u}{\partial t}$ . Además, debe poderse hablar de aceleración para cada punto de la cuerda. Conviene entonces restringir a trayectoria  $t \mapsto (u(t, \cdot), v(t, \cdot))$  para los cuales  $\frac{\partial u}{\partial t}$  y  $\frac{\partial^2 u}{\partial t^2}$  a) estén definidas para cada  $(t,z)$ , con  $t \in \mathbb{R}$ ,  $z \in S$  y b) sean continuas ahí.

Si las velocidades que se alcanzan no son demasiado altas, podemos calcular la fuerza disipativa que ejerce el medio sobre nuestra porción de cuerda tal y como se hizo en el ejemplo 1.2.5; es decir, dicha fuerza está dada por

$$-\int_z^{z+h} \mu \frac{\partial u}{\partial t}(t, \zeta) d\zeta, \quad (3)$$

donde  $\mu > 0$ .

Supondremos conocida también la densidad lineal de la cuerda,  $\rho$  (constante, medida en g/cm), de tal manera que

$$\int_z^{z+h} \rho \frac{\partial^2 u}{\partial t^2}(t, \zeta) d\zeta$$

es el término "masa  $\times$  aceleración" que aparece en la Segunda Ley de Newton.

Despreciando la acción de la gravedad sobre la cuerda (se trata de un material ligero), las únicas fuerzas externas son (2) y (3), por lo que la Segunda Ley de Newton afirma que

$$\int_z^{z+h} \rho \frac{\partial^2 u}{\partial t^2}(t, \zeta) d\zeta = -\int_z^{z+h} \mu \frac{\partial u}{\partial t}(t, \zeta) d\zeta + T \frac{\partial}{\partial z} \left\{ \frac{\partial u / \partial z}{\sqrt{1 + (\partial u / \partial z)^2}} \right\} (t, z+\alpha h) h$$

Basta dividir entre  $h$  y tomar el límite cuando  $h \rightarrow 0$  para obtener que

$$\rho \frac{\partial^2 u}{\partial t^2} + \mu \frac{\partial u}{\partial t} = T \frac{\partial}{\partial z} \left\{ \frac{\frac{\partial u}{\partial z}}{\sqrt{1 + \left(\frac{\partial u}{\partial z}\right)^2}} \right\}. \quad (4)$$

Por lo común se supone que las vibraciones son de pequeña amplitud. Ello implica que  $\theta(z)$  toma valores pequeños, por lo que  $\text{sen}\theta(z) \approx \text{tg}\theta(z) = \frac{\partial u}{\partial z}$  y entonces (4) da lugar a

$$\rho \frac{\partial^2 u}{\partial t^2} + \mu \frac{\partial u}{\partial t} = T \frac{\partial^2 u}{\partial z^2} \quad (5)$$

llamada la *ecuación del telégrafo*. Si no hay fricción ( $\mu=0$ ), la ecuación del telégrafo se reduce a

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial z^2}, \quad (6)$$

donde  $c^2 := T/\rho$ . Esta ecuación se llama normalmente *de la cuerda vibrante*; en realidad, sabemos que se obtiene a partir de (4), bajo la suposición de que no hay fricción y de que las oscilaciones son pequeñas.

Conviene observar que las unidades de  $c$  en el sistema cgs son cm/seg, es decir,  $c$  es una *velocidad*: resulta ser la velocidad de propagación de las ondas generadas por la cuerda vibrante.

Así pues, a manera de conclusión: las trayectorias en el espacio de los perfiles son las funciones  $t \mapsto (u(t, \cdot), u(t, \cdot))$ , soluciones del sistema de ecuaciones en derivadas parciales

$$\frac{\partial u}{\partial t} = v$$

$$\frac{\partial u}{\partial t} = -\frac{\mu}{\rho} v + c^2 \frac{\partial}{\partial z} \left\{ \frac{\frac{\partial u}{\partial t}}{\sqrt{1 + \left[ \frac{\partial u}{\partial z} \right]^2}} \right\};$$

o cualquiera de sus versiones simplificadas (5) y (6).

Del modelo anterior y de los datos en las secciones precedentes podemos extraer algunas conclusiones generales acerca de los modelos que hemos venido construyendo.

Se parte dando una representación para el tiempo y también para el espacio. Esta última (S) es un conjunto finito (o a lo más

numerable), cuando se está modelando un sistema de partículas; si se trata de modelar un medio continuo,  $S$  es un subconjunto abierto de  $\mathbb{R}^d$ , para  $d \leq 3$ . El tiempo se representa mediante un subconjunto de  $\mathbb{Z}$  (modelos a tiempo discreto) o un intervalo en  $\mathbb{R}$  (modelos a tiempo continuo).

En cada caso, se trata de estudiar un cierto número  $n$  de atributos de interés, cuyos valores se representarán mediante los elementos de un conjunto  $A \subset \mathbb{R}^n$ . En cada instante se tendrá una distribución especial de atributos, representada mediante una función  $x: S \rightarrow A$ , llamada perfil de atributos. Los perfiles aceptables constituyen un conjunto  $E$ , y quedarán fijados mediante consideraciones de tipo físico o matemático dictadas por la misma situación bajo análisis.

El objetivo del modelado es representar el fenómeno, para lo cual se debe especificar la repartición de los atributos en el espacio y en el tiempo. Esto se puede lograr dando la trayectoria temporal de los perfiles. En otras palabras, interesan las funciones  $\omega: T \rightarrow E$ , cada una de las cuales constituye una posible trayectoria; de nuevo, cada problema limita la clase de las trayectorias admisibles, las cuales constituyen un conjunto  $\Omega$ .

Queda así establecida la forma general que toma un modelo matemático, a saber

$$\langle T, S, A, E, \Omega \rangle,$$

donde  $T$ ,  $S$  y  $A$  son conjuntos,

$T$ : el tiempo,  $T \subset \mathbb{R}$

$S$ : el espacio,  $S$  finito o  $S$  abierto en  $\mathbb{R}^d$ .

$A$ : los atributos,  $A \subset \mathbb{R}^n$ .

Además

$E$ : los perfiles admisibles, un conjunto de funciones de  $S$  en  $A$ .

$\Omega$ : las trayectorias admisibles, un conjunto de funciones de  $T$  en  $E$ .

## 2.2 Un ejemplo: la conducción de calor en sólidos.

Problema: "El extremo izquierdo de una barra metálica está empotrado en un material que se mantiene a temperatura constante, en tanto que el derecho está libre (véase la Fig.1). La atmósfera que rodea la barra es muy extensa y recibe calor de la barra a todo lo largo. Se desea modelar la variación espacio-temporal de la temperatura en la barra".

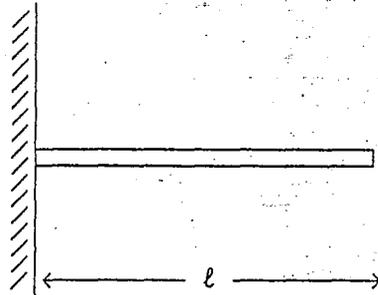


Fig. 1: Una barra empotrada en una pared vertical

La situación anterior puede presentarse en un análisis simplificado de los fenómenos que suceden en una aleta de enfriamiento: se busca disipar calor de un cuerpo caliente dotándolo de aletas a lo largo de él, para así aumentar la superficie de transferencia de energía al exterior. Veamos cómo esta situación nos permite ilustrar los elementos de un modelo matemático tal y como se describieron en la sección anterior, a partir del análisis de la cuerda vibrante.

Suponiendo que la barra sea esbelta, la temperatura en cada una de sus secciones transversales será aproximadamente constante. Para simplificar la descripción de este fenómeno, supongamos que la temperatura es de hecho constante sobre cada sección transversal de la barra, es decir, que varía únicamente en dirección axial, y esto en cada instante. Esta consideración nos sugiere tomar  $S := (0, l)$  como representante del espacio, donde  $l$  es la longitud de la barra. Como representante del tiempo tomaremos  $T := \mathbb{R}$ . Dado que el único atributo de interés es la temperatura (medida en  $^{\circ}\text{K}$ ), que no puede ser negativa, tomaremos  $A := [0, \infty)$  como espacio de atributos.

Sea  $T_0$  la temperatura de la pared,  $T_a$  la del medio ambiente, ambas medidas en  $^{\circ}\text{K}$ . Supongamos que la barra cede calor al ambiente en proporción directa a la diferencia de temperatura y al área expuesta (ley de Newton sobre el enfriamiento). Supongamos además que el extremo izquierdo de la barra está en todo momento a la temperatura de la pared.

Los perfiles de temperatura serán funciones  $x: (0, \ell) \rightarrow [0, \infty)$ , que supondremos continuas, es decir, no hay saltos en los valores de la temperatura a lo largo de la barra. Desde luego, debe cumplirse que  $x(0^+) = T_0$ , según lo que hemos supuesto. En general,  $x(z)$  es la temperatura en aquella sección de la barra que se encuentra a  $z$  cm de la pared ( $0 \leq z \leq \ell$ ).

En el extremo de la derecha, se pierde calor al exterior. Si la barra es cilíndrica y tiene radio  $R$ , su sección transversal tiene área  $\pi R^2$ , por lo que las pérdidas de calor al exterior son

$$\gamma(\pi R^2)[x(\ell) - T_a] \frac{\text{cal}}{\text{seg}},$$

según la ley de Newton del enfriamiento. Aquí  $\gamma$  es una constante positiva, que supongamos conocida; nótese que sus unidades son  $\text{cal} \cdot \text{seg}^{-1} \cdot \text{cm}^{-2} \cdot ^{\circ}\text{K}^{-1}$ .

Por otro lado, dentro de la barra se tienen secciones a distintas temperaturas. Es un hecho empírico que dicha circunstancia genera una transferencia neta de energía, de las partes más calientes a las más frías. Este fenómeno se puede estudiar midiendo la temperatura a lo largo del tubo en un instante dado, mediante termopares colocados en las posiciones  $z_1, \dots, z_p$ . Dado este arreglo experimental, dichas mediciones se efectúan frecuentemente, digamos en los instantes  $t_1, \dots, t_p$ , de todo lo cual resulta una tabla de valores

$$t_i, T_{z_1}^{(i)}, \dots, T_{z_p}^{(i)}, i=1, \dots, r,$$

donde  $T_j^{(i)}$  denota el valor de la temperatura medida en el instante  $t_i$  por el termopar colocado en el punto  $z_j$ . De dicha tabla se puede

estimar los gradientes de temperatura, mediante los cocientes

$$g_{ij} := \frac{T_{j+1}^{(i)} - T_j^{(i)}}{z_{j+1} - z_j} \quad j=1, \dots, p-1; \quad i=1, \dots, r$$

A su vez, se determina el flujo de calor  $Q_{ij}$  a través de cada una de las  $p$  secciones transversales estudiadas, y esto también en cada uno de los  $r$  instantes considerados. Aquí  $Q_{ij}$  está medido en  $\frac{\text{cal}}{\text{seg} \cdot \text{cm}^2}$ . Es

un hecho empírico, verificado en un gran número de casos, que  $Q_{ij}/g_{ij}$  es constante (al menos dentro del error experimental considerado permisible) y toma valores negativos, que dependen solo del material de que esté constituida la barra. En particular, el calor se transmite en la dirección en que disminuye la temperatura, como ya hemos señalado.

Idealizando la situación anterior, supongamos que  $x: (0, \ell) \rightarrow [0, \infty)$  es diferenciable. Entonces, las observaciones anteriores se traducen en la llamada *ley de Fourier*, según la cual el flujo de calor por unidad de área a lo largo de la barra es  $-kx'(z)$   $\frac{\text{cal}}{\text{seg} \cdot \text{cm}^2}$  (medida en la sección  $z$ ). Aquí  $k$  es un número positivo, que supondremos conocido; sus unidades son  $\frac{\text{cal}}{\text{seg} \cdot \text{cm} \cdot ^\circ\text{K}}$  y se llama la *conductividad térmica* del material de que está hecha la barra. En particular, el flujo de calor en el interior de la barra, medido en su extremo libre es

$$-k(\pi R^2)x'(\ell^-) \frac{\text{cal}}{\text{seg}},$$

cuando el perfil de temperatura es  $x$ .

En vista de lo anterior, y puesto que no se acumula energía en el extremo derecho, cada perfil diferenciable por la izquierda en  $z = \ell$  debe cumplir la condición

$$-kx'(\ell) = \gamma[x(\ell) - T_a].$$

De hecho convendrá tomar perfiles que son dos veces continuamente diferenciables en  $(0, \ell)$ , según se verá más adelante. Así pues,

$$E := \left\{ x \in C^1[0, \ell] \cap C^2(0, \ell) : x(0) = T_0, kx'(\ell) + \gamma x(\ell) = \gamma T_a \right\}$$

Sea  $\omega: T \rightarrow E$  una trayectoria en el espacio de los perfiles admisibles; de nuevo, utilizaremos la notación

$$u(t, z) := \omega(t)(z).$$

Es decir,  $u(t, \cdot) \in E$  es el perfil de temperatura en la barra, medido en el instante  $t$ . Supongamos que  $\frac{\partial u}{\partial t}$  es continuamente diferenciable.

Igual que hicimos para la cuerda vibrante en la sección anterior, examinemos los eventos que suceden en la sección de la barra comprendida entre  $z$  y  $z+h$ , con  $0 < z < z+h < \ell$  (ver figura 2), en un instante fijo  $t$ . Sea ahora  $\rho$  la densidad del material de la barra, medida en  $\text{g/cm}^3$ ,  $C_p$  su capacidad calorífica, medida en  $\frac{\text{cal}}{\text{g} \cdot ^\circ\text{K}}$ .

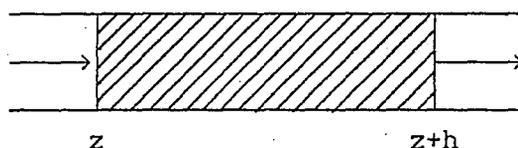


Fig. 2: Sección de la barra metálica de la figura 1.

El contenido total de energía en esta sección es  $(\pi R^2) \int_z^{z+h} \rho C_p u(t, \zeta) d\zeta$ , por lo que su tasa de acumulación es

$$\frac{d}{dt} (\pi R^2) \int_z^{z+h} \rho C_p u(t, \zeta) d\zeta = (\pi R^2) \int_z^{z+h} \rho C_p \frac{\partial u}{\partial t}(t, \zeta) d\zeta \quad \frac{\text{cal}}{\text{seg}}. \quad (1)$$

Por otro lado, a lo largo de la barra se conduce el calor; según la ley de Fourier en  $z$  entra calor con una tasa

$$-k(\pi R^2) \frac{\partial u}{\partial z}(t, z)$$

Así pues,

$$+k(\pi R^2) \left[ \frac{\partial u}{\partial z}(t, z+h) - \frac{\partial u}{\partial z}(t, z) \right]$$

da el calor neto transferido por conducción a través de la sección bajo estudio. Dado que  $u(t, \cdot) \in C^2(0, \ell)$ , el Teorema del Valor Medio nos

permite expresar dicha cantidad en la forma de

$$k(\pi R^2)h \frac{\partial^2 u}{\partial z^2}(t, z+\alpha h) \frac{\text{cal}}{\text{seg}}, \quad (2)$$

donde  $\alpha \in (0, 1)$ .

Finalmente, se transmite calor al medio circundante: según la ley de Newton del enfriamiento, la tasa de pérdidas de calor por dicho concepto es

$$\gamma(2\pi R) \int_z^{z+h} [u(t, \zeta) - T_a] d\zeta \frac{\text{cal}}{\text{seg}}. \quad (3)$$

Puesto que se conserva la energía, la tasa neta de acumulación (2) - (3) debe coincidir con (1), es decir

$$\int_z^{z+h} \rho(p) \frac{\partial u}{\partial t}(t, \zeta) d\zeta = kh \frac{\partial^2 u}{\partial z^2}(t, z+\alpha h) - \frac{2\gamma}{R} \int_z^{z+h} [u(t, \zeta) - T_a] d\zeta,$$

para cada  $t$ , y donde  $\alpha \in (0, 1)$ . Basta dividir entre  $h$  y tomar el límite cuando  $h \rightarrow 0$  para obtener que

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial z^2} - \beta(u - T_a), \quad (4a)$$

donde

$$\alpha = \frac{k}{\rho C_p}, \quad \beta = \frac{2\gamma}{\rho R C_p}. \quad (4b)$$

Así pues, las trayectorias admisibles constituyen un conjunto  $\Omega$ , formado por aquellas funciones  $u: \mathbb{R} \times (0, l) \rightarrow [0, \infty)$  tales que

- i)  $u(t, \cdot) \in C^1[0, l] \cap C^2(0, l)$ ,  $t \in \mathbb{R}$   
 $\frac{\partial u}{\partial t} \in C(\mathbb{R} \times (0, l))$
- ii)  $u(t, 0) = T_0$ ,  $k \frac{\partial u}{\partial z}(t, l) + \gamma u(t, l) = \gamma T_a$ ,  $t \in \mathbb{R}$
- iii)  $u$  satisface la ecuación diferencial (4).

Esto concluye la especificación de un modelo matemático para la situación planteada. ■

Igual que para las partículas en un campo de fuerza estudiadas en la sección 1.2, también para la barra metálica en un baño térmico que

acabamos de estudiar podemos investigar el comportamiento de la energía. En este caso la energía mecánica del sistema no cambia, pues no se registra ningún desplazamiento, pero sí cambia su energía térmica. Análogamente a la función hamiltoniana que nos da la energía mecánica total en la sección 1.2, podemos dar el funcional entalpia, que para cada perfil  $x$  de temperatura da la correspondiente energía térmica en la barra:

$$H(x) := (\pi R^2) \int_0^\ell \rho C_p x(z) dz.$$

A lo largo de una trayectoria  $t \mapsto u(t, \cdot)$  en el espacio  $E$  de los perfiles de temperatura, la energía térmica está dada en función del tiempo mediante

$$E(t) := H(u(t, \cdot))$$

es decir,

$$E(t) = (\pi R^2) \int_0^\ell \rho C_p u(t, z) dz.$$

Un cálculo simple basado en (4) lleva a que

$$\frac{dE}{dt} = -(\pi R^2) k \frac{\partial u}{\partial z}(t, 0) - \gamma (\pi R^2) [u(t, \ell) - T_a] - (2\pi R) \gamma \int_0^\ell [u(t, z) - T_a] dz, \quad (5)$$

que no es otra cosa que un balance global de energía térmica: la tasa neta de acumulación en la barra es la diferencia entre la tasa con que la energía entra a la misma por el extremo izquierdo empotrado en la pared a alta temperatura y la tasa con que la barra cede calor al exterior a través de su paredes.

Por otro lado, hay circunstancias en las que no se observan cambios temporales en la temperatura de los diferentes puntos de la barra, es decir, en sus perfiles de temperatura. De un perfil de temperatura que, una vez alcanzado, ya no cambia, se dice que es *estacionario*.

Sea  $\bar{u} \in E$  un tal perfil. Nótese que, a lo largo de la trayectoria estacionaria correspondiente, la energía  $E(t)$  es constante, por lo que el balance global de energía (5) se reduce a

$$(\pi R^2) k \bar{u}'(0) = \gamma (\pi R^2) [\bar{u}(\ell) - T_a] + (2\pi R) \gamma [\langle \bar{u} \rangle - T_a],$$

donde

$$\langle \bar{u} \rangle := \frac{1}{l} \int_0^l \bar{u}(z) dz$$

es la temperatura media a lo largo de la barra.

Ahora bien, la trayectoria estacionaria satisface (4), por lo que el perfil estacionario  $\bar{u}$  es solución del problema de valores a la frontera

$$\alpha \frac{d^2 \bar{u}}{dz^2} - \beta (\bar{u} - T_a) = 0 \quad (6a)$$

$$\bar{u}(0) = T_0, \quad k u'(l) + \gamma \bar{u}(l) = \gamma T_a \quad (6b)$$

Un cálculo simple, aunque algo tedioso, muestra que

$$\bar{u}(z) = T_a + (T_0 - T_a) \varphi(z), \quad (7a)$$

donde

$$\varphi(z) := \frac{(kc - \gamma) \cosh c(l+z) + \gamma \cosh c(l-z) + (kc + \gamma) \sinh c(l+z) - kc \cdot \sinh c(l-z)}{kc \cosh cl + \gamma \sinh cl} \quad (7b)$$

con  $c := \sqrt{\beta/\alpha}$ .

Un perfil arbitrario  $f \neq \bar{u}$  no será estacionario, por lo que si dicho perfil de temperatura se alcanza en un momento dado (digamos en  $t=0$ ) a partir de ahí se desatará un proceso termodinámico irreversible, en el que la temperatura en cada punto de la barra se verá cambiar continuamente en el tiempo. Sin embargo, a la larga se restablecerá el estado estacionario  $\bar{u}$ . En efecto, sea  $u \in \Omega$  la trayectoria correspondiente, que puede obtenerse mediante el método de separación de variables aplicado al problema de valores iniciales correspondiente a (4) con la condición inicial  $u(0, \cdot) = f$ . Sea  $\delta := f - \bar{u}$ ,  $v := u - \bar{u}$ . Entonces,  $v$  satisface el problema de valores iniciales y a la frontera.

$$\frac{\partial v}{\partial t} = \alpha \frac{\partial^2 v}{\partial z^2} - \beta v, \quad (8a)$$

$$v(t, 0) = 0 \quad k \frac{\partial v}{\partial z}(t, l) + \gamma v(t, l) = 0 \quad (8b)$$

$$v(0, \cdot) = \delta. \quad (8c)$$

El método de separación de variables conduce al problema de valores propios

$$\alpha y'' - \beta y = \lambda y \quad \text{en } (0, \ell) \quad (9a)$$

$$y(0)=0, \quad ky'(\ell) + \gamma y(\ell)=0, \quad (9b)$$

cuya solución es

$$\lambda_k = -\beta - \alpha \omega_k^2, \quad \psi_k(z) = \text{sen} \omega_k z, \quad k=1, 2, \dots \quad (9c)$$

con  $\{\omega_k, k=1, 2, \dots\}$  son las soluciones positivas de la ecuación trascendente

$$k\omega + \gamma \text{tg} \omega \ell = 0.$$

Entonces, la solución del problema (7) queda expresada en la forma

$$v(t, z) = \sum_{k=1}^{\infty} c_k e^{\lambda_k t} \psi_k(z), \quad (10)$$

donde  $\{c_k, k=1, 2, \dots\}$  son los coeficientes de Fourier de la función  $\delta$  respecto al sistema ortogonal  $\{\psi_k\}$ , es decir

$$\delta(z) = \sum_{k=1}^{\infty} c_k \psi_k(z).$$

De (10) es claro que  $v(t, z) \rightarrow 0$  cuando  $t \rightarrow +\infty$ , por lo que  $u(t, z) \rightarrow \bar{u}(z)$  y con ello queda establecida la última afirmación en cursivas: el perfil (7) es estable (se puede mantener) pues, si se le perturba, el sistema reacciona en forma tal que a la larga se restituyen las condiciones deseadas.

En la segunda parte de la conferencia 4 se construyen modelos matemáticos muy cercanos al que nos ha ocupado en esta sección. Ahí se examinan con mayor atención varios aspectos ya encontrados en el fenómeno de la conducción de calor, pero con referencia a la difusión molecular. Por ejemplo, ahí se estudian con detenimiento las diferentes condiciones a la frontera que puedan presentarse, así como la caracterización matemática de los flujos (de materia o energía, según se trate). En cierto sentido, dicha sección puede considerarse como una continuación de ésta.

## CONFERENCIA 3.

### EL CARACTER RELATIVO DEL DETERMINISMO.

Los modelos considerados hasta ahora han sido construidos desde una óptica determinística: existe un *estado*, cuyo conocimiento en un instante dado determina el futuro, siempre que se conozcan los factores externos que actúan sobre el sistema en cuestión. En esta conferencia ilustramos el tipo de consideraciones que se hacen cuando no se puede asignar un estado. En la sección 1 se aborda de nuevo el fenómeno de la partícula que cae, pero desde un punto de vista distinto; se descubre que el mismo fenómeno puede aparecer determinístico o no, según la óptica desde la cual se estudie. En la sección 2 se ilustra lo anterior mediante la desintegración radiactiva: para tiempos largos resulta determinística, en tanto que para tiempos cortos no. Se da el espacio probabilístico como modelo no determinístico por excelencia.

#### 3.1 Determinismo e indeterminismo.

Regresemos a la situación estudiada en la primera conferencia, acerca del objeto que cae. Hemos visto cómo es posible construir un modelo matemático para tal situación bajo diversas suposiciones acerca de la forma, tamaño, etc., del objeto, así como de las características del medio que lo rodea. En particular, hemos construido un modelo según el cual el tiempo que tarda en caer una moneda que se suelta desde una altura  $x_0$  es de  $\sqrt{2x_0/g}$ , suponiendo que la aceleración de la gravedad  $g$  es constante a lo largo de la trayectoria.

Fijémonos ahora en otro atributo de la moneda, como es la cara que queda hacia arriba una vez que toca el piso. En efecto, la experiencia nos indica que en la inmensa mayoría de los casos la moneda reposará en el suelo con águila (A) o sol (S) hacia arriba, y que sólo en raras ocasiones descansará sobre su canto, o rodará y se perderá de vista. Además, nunca jamás se ha visto que la moneda desaparezca a medio camino hacia el suelo, sino que ciertamente cae. Así pues, sin demasiados escrúpulos podemos afirmar que

- i) la moneda cae,
- ii) presenta A o S,

y nada más.

También la experiencia en este tipo de lanzamientos nos recuerda lo difícil que resulta predecir qué cara va a salir al lanzar la moneda: por más que repetimos el lanzamiento, no sabemos qué factores fijar para que nuestra predicción concuerde con lo observado. Una ronda típica de volados tiene como resultados

A, S, S, A, S, A, A, A. (1)

Sin embargo, algo que sí podemos hacer es repetir el experimento un cierto número  $N$  de veces, contando el número  $N_A$  de ellos en que sale águila. La frecuencia relativa correspondiente es

$$f_A^{(N)} := \frac{N_A}{N},$$

Para los resultados (1), las frecuencias relativas son

N	1	2	3	4	5	6	7	8
$f_A^{(n)}$	1.000	0.500	0.333	0.500	0.400	0.500	0.571	0.625

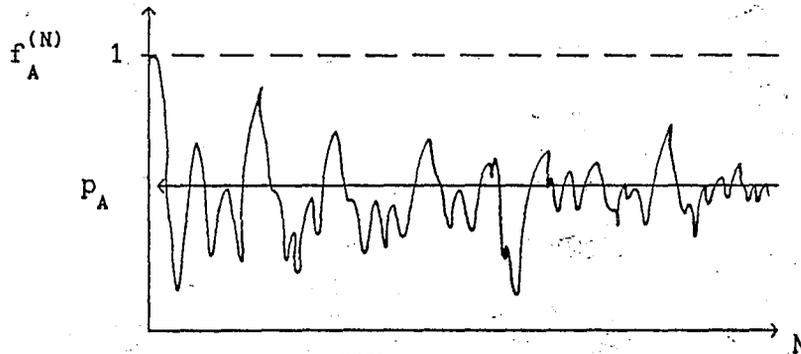


Fig. 1: Comportamiento de la frecuencia relativa

A la larga se observan comportamientos del tipo ilustrado en la figura 1: los valores  $f_A^{(N)}$  oscilan en forma errática, pero para  $N$  grande, las oscilaciones son cada vez menores (con algunas desviaciones), en torno a un valor  $p_A \in [0, 1]$ . Se repite este experimento un número  $r$  de veces, obteniendo así números

$$p_A^{(1)}, \dots, p_A^{(r)}$$

que, en términos generales, difieren "poco". Luego entonces, algo que sí puede predecirse es este valor  $p_A$ , al menos dentro de un aceptable margen de error: el valor  $p_A \in [0,1]$  así determinado se llama la probabilidad de que salga águila al lanzar la moneda en cuestión.

Esta construcción se puede realizar para un sinfín de situaciones en las que en principio no se detecta ninguna reproducibilidad: se elige un evento observable, se repite el experimento un "gran" número de veces y se determina el "límite" de la sucesión de frecuencias relativas cuando el número de repeticiones crece indefinidamente. El resultado es la *probabilidad* del evento en cuestión, o mejor dicho, una estimación de dicha probabilidad. Para afinar la estimación de la probabilidad de un evento, conviene determinarla un número "grande" de veces y tomar a su promedio como estimador.

El enfoque anterior para la determinación de probabilidades se conoce como *frecuentista*, y corresponde a una visión empirista del origen del conocimiento. Su base fundamental, acerca del comportamiento asintótico de las frecuencias relativas de un evento al repetirse un experimento, se llama *Principio de la Regularidad Estadística* y es una generalización de la experiencia. Este enfoque tiene sus limitaciones y hay enfoques alternativos; también hay maneras de darle la vuelta a las críticas que se le hacen. Basta por ahora decir que no hay ninguna solución enteramente satisfactoria al problema de aclarar el origen del concepto de probabilidad. El enfoque frecuentista es suficientemente bueno, y especialmente atractivo para la aplicación de la probabilidad a las ciencias experimentales, pues según él la probabilidad es una propiedad física que puede medirse experimentalmente.

Todo esto es bien conocido, y el resultado final es un modelo probabilístico para este aspecto del fenómeno "lanzamiento de una moneda". Veamos las características de dicho modelo.

Primeramente, tenemos los *resultados* posibles del experimento en cuestión, que no son sino los elementos del conjunto

$$\Omega = \{A, S\},$$

según hemos convenido. En seguida, reconocemos los *eventos* observables asociados a este experimento:

- que salga águila
- que salga sol
- que salga algo ("águila o sol")
- el resto,

donde "el resto" incluye todos esos eventos concebibles que sin embargo hemos clasificado como imposibles. Representamos todos los eventos anteriores mediante subconjuntos de  $\Omega$ , a saber  $\{A\}$ ,  $\{S\}$ ,  $\Omega$  y  $\phi$ , respectivamente. Tenemos pues la familia de eventos

$$\mathcal{A} = \{\phi, \{A\}, \{S\}, \Omega\}$$

como segundo ingrediente de este modelo.

Después, siguiendo el enfoque frecuentista determinamos  $p_A$  (la probabilidad de que salga águila) y entonces  $p_S = 1 - p_A$  (la probabilidad de que salga sol). Claramente la probabilidad de que salga algo es uno ( $p_\Omega = 1$ ) y la probabilidad de que salga algo imposible es  $p_\phi = 0$  (nunca se observa). En otras palabras, tenemos una función  $P: \mathcal{A} \rightarrow [0, 1]$  que asigna a cada evento su probabilidad:

$$P(A) = p_A \qquad P(S) = 1 - p_A$$

$$P(\Omega) = 1 \qquad P(\phi) = 0$$

Dicha función se denomina la *probabilidad* y completa la terna ordenada

$$(\Omega, \mathcal{A}, P), \qquad (1)$$

llamada *espacio probabilístico*: este es el modelo matemático utilizado con mayor frecuencia para describir situaciones en las que interviene el azar. Pero hay alternativas, sólo que aquí no hablaremos de ellas.

Del análisis anterior parece desprenderse la conclusión

siguiente: el mismo fenómeno (la caída libre de una moneda) presenta dos aspectos, uno de los cuales puede tratarse en forma determinística en tanto que el otro parece requerir un tratamiento no determinístico. Sin embargo, a continuación veremos que las cosas no son así de simples, y que hay otras variables que influyen también. Por ejemplo la escala de observación.

Supongamos que el aspecto determinístico del fenómeno (el tiempo que la moneda tarda en llegar al suelo cuando se le suelta desde cierta altura) se estudia con una moneda pequeña, armados de un reloj de baja precisión. Si se suelta la moneda de alturas menores de 1 m, el tiempo de caída será menor de 0.5 seg. Dada la precisión del reloj, es concebible que los resultados de medir dicho tiempo ocho veces sean

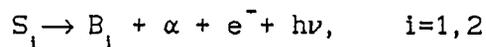
$$1, 0, 0, 1, 0, 1, 1, 1 \quad (3)$$

compárese con (1). Más aún, igual que cuando se estudia el otro aspecto de los volados, al repetir el experimento que dio lugar a (3) no se obtiene ninguna reproducibilidad. Bajo estas condiciones, un observador concluiría que este aspecto del fenómeno no puede tratarse determinísticamente, y optaría por una descripción probabilística afín a (2). En dicho modelo  $\Omega$  los resultados posibles serían 0 y 1, de tal manera que  $\Omega = \{0, 1\}$ ,  $\mathcal{A} = \{\emptyset, \{0\}, \{1\}, \Omega\}$  y P se determinaría según un enfoque frecuentista.

¿Qué pasó? El ejemplo anterior nos enseña que una misma situación puede aparecer determinista desde una escala de observación y no determinista desde otra. Por lo tanto, el determinismo resulta más bien ser una óptica para el análisis de la realidad física, antes que una propiedad intrínseca de la misma. En la siguiente sección profundizaremos en el estudio de estos aspectos.

### 3.2 Un ejemplo: La desintegración radiactiva.

Problema. "Una sustancia radiactiva S presenta dos isótopos,  $S_1$  y  $S_2$ . Ambos son radiactivos, y se descomponen emitiendo partículas  $\alpha$ ,  $\beta$  y  $\gamma$ , con lo que dan lugar a sendos isótopos estables  $B_1$  y  $B_2$ , según las reacciones



$S_1$  se descompone muy lentamente,  $S_2$  muy rápidamente. Por lo tanto, el estudio experimental de estas descomposiciones requiere instrumentación distinta: en el primer caso se registra la masa del radioisótopo  $S_1$  mediante una balanza analítica; en el segundo, se cuenta el número de partículas alfa desprendidas usando un contador Geiger. En ambos casos se mide el tiempo mediante un reloj bastante preciso. Se desea describir el fenómeno en términos matemáticos".

En ambos casos se puede trabajar a tiempo continuo, es decir, tomando  $T := \mathbb{R}$ . Además, el espacio aquí no juega ningún papel, puesto que sólo se registra la radiactividad en forma global, mediante el instrumento del caso. Por lo tanto, en ambos casos representaremos el espacio mediante un solo punto.

En cuanto a los atributos, aquí sí hay diferencia: en el primer caso se registra una variable continua (la masa), en tanto que en el segundo se registra una variable entera (el número de partículas alfa); en ambos casos se trata de variables no negativas, por lo que

$$A_1 := [0, \infty), \quad A_2 := \mathbb{N} \cup \{0\},$$

donde  $\mathbb{N}$  denota el conjunto de los números naturales. Dada la elección de representante espacial ( $S_1 = S_2 = \{1\}$ ), de manera natural se identifican los perfiles con los atributos, igual que en los ejemplos de mecánica de partículas vistos en el capítulo anterior. Así pues,

$$E_1 := [0, \infty), \quad E_2 := \mathbb{N} \cup \{0\}.$$

Por lo que se refiere a las trayectorias, podemos hacer las siguientes consideraciones. Evidentemente, en el primer caso se tendrán funciones reales no negativas, definidas en toda la recta real. Tomando en cuenta que puede medirse empíricamente la tasa instantánea de desintegración, mediante el procedimiento ya descrito en la sección 1.1, es natural pedir que  $\Omega_1$  conste de funciones reales continuamente diferenciables,  $m: \mathbb{R} \rightarrow [0, \infty)$ , con

información presente. Como quiera, se trata de una idealización cuyas predicciones deben tomarse con cautela, sobre todo para tiempos largos: cuando  $t \rightarrow \infty$ ,  $e^{-\lambda t} m_0$  tomará valores por debajo de la precisión de nuestra balanza analítica, así que en la práctica se concluye que el radioisótopo se agota. De manera análoga, en el pasado remoto la masa del radioisótopo tiene un valor grande, pero seguramente no tiende a  $+\infty$  cuando  $t \rightarrow -\infty$  como predice el modelo.

Con todo y sus limitaciones, este modelo es sumamente útil para identificar un radioisótopo, mediante la determinación de la constante de desintegración radiactiva  $\lambda$  que le corresponde. Para ello, basta medir el tiempo necesario para que una masa dada de  $S_1$  se reduzca a la mitad (el tiempo de vida media,  $t_{1/2}$ ). En efecto, sabiendo que

$$\frac{m_0}{2} = m_0 e^{-\lambda t_{1/2}},$$

se obtiene de ahí que

$$\lambda = \frac{\log 2}{t_{1/2}}. \quad (2)$$

Resíprocamente, si se conoce la constante de desintegración radiactiva  $\lambda$  de un cierto radioisótopo  $S_1$ , este modelo puede utilizarse para determinar la edad de objetos que lo contengan: basta observar que

$$t_1 = \frac{\log \left( \frac{m_1}{m_0} \right)}{\lambda};$$

la Geocronología a desarrollado métodos para calcular la edad de las rocas mediante esta fórmula: se mide el valor actual  $m_1$  y se estima  $m_0$ , con lo cual se obtiene la edad  $t_1$ .

Estas técnicas, por supuesto, suponen posible realizar las mediciones con precisión absoluta. En la práctica esto no sucede, por lo que conviene proceder de manera diferente; ilustremos lo anterior mediante una técnica para estimar la constante de desintegración radiactiva  $\lambda$ . Se miden las masas remanentes  $m_1, \dots, m_N$  en instantes  $t_1, \dots, t_N$ ; por otro lado, se sabe que al graficar los puntos

$$\left( t_i, \frac{m_0}{m_i} \right), \quad i=1, \dots, N$$

en escala semilogaritmica se obtienen puntos "más o menos" alineados a lo largo de una recta de pendiente positiva a través del origen de coordenadas.

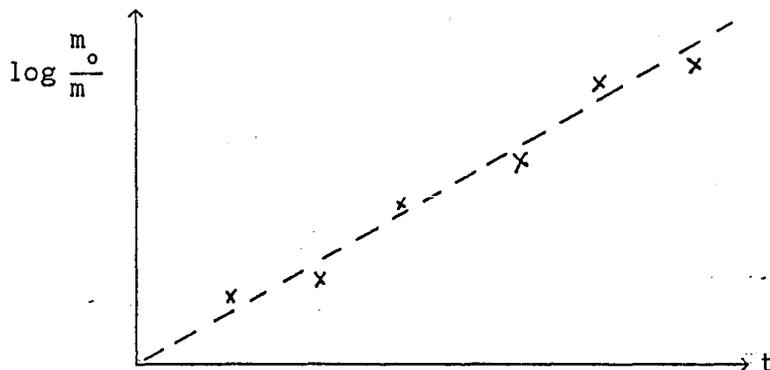


Fig.1: Datos experimentales sobre el comportamiento temporal de la masa residual de radioisótopo.

Se tiene entonces que

$$\log \frac{m_0}{m_i} = \lambda t_i + \varepsilon_i \quad i=1, \dots, N$$

y se determina  $\lambda$  por el método de los mínimos cuadrados, buscando minimizar la suma de residuos cuadráticos  $\sum_{i=1}^N \varepsilon_i^2$ . Se obtiene entonces la fórmula

$$\lambda = \frac{\sum_{i=1}^n t_i \log \frac{m_0}{m_i}}{\sum_{i=1}^n t_i^2} \quad (3)$$

en lugar de (2).

En cuanto a la determinación de  $\Omega_2$ , es claro que también consta de funciones no negativas sobre la recta, pero cuyos valores son enteros. De hecho, las trayectorias son funciones escalonadas no decrecientes  $N: \mathbb{R} \rightarrow \mathbb{N} \cup \{0\}$ , donde

$N(t)$  := número de partículas alfa emitidas hasta el instante  $t$ .

Para cada trayectoria puede distinguirse una sucesión de instantes

$$0 = \tau_0 < \tau_1 < \tau_2 < \dots, \quad (4)$$

donde

$\tau_1$  = instante en que se desprende la  $i$  ésima partícula  $\alpha$ .

Dado que cualquier muestra de radioisótopo  $S_2$  no contiene más de un número finito de núcleos y cada núcleo puede dar lugar a unas cuantas partículas alfa, en la práctica la sucesión (4) debe terminar. Sin embargo, el número finito de núcleos es muy grande, del orden del número de Avogadro  $6.023 \times 10^{23}$ , por lo que parece legítimo suponer que (4) es infinita: el fenómeno puede proceder en forma indefinida en el tiempo.

Así pues,  $\Omega_2$  consta de funciones escalonadas, con valores no negativos y cuyos saltos son todos positivos, de magnitud 1. Por conveniencia, conviene considerar que son funciones continuas por la derecha, de tal modo que

$$N(\tau_k^+) = N(\tau_k^-) = N(\tau_k^-) + 1, \quad k = 1, 2, \dots \quad (5)$$

En otras palabras:

$$\Omega_2 = \left\{ N: \mathbb{R} \rightarrow \mathbb{N} \cup \{0\} \mid \text{vale (5) para alguna sucesión (4)} \right\};$$

una trayectoria típica tiene la forma dada en la figura 2

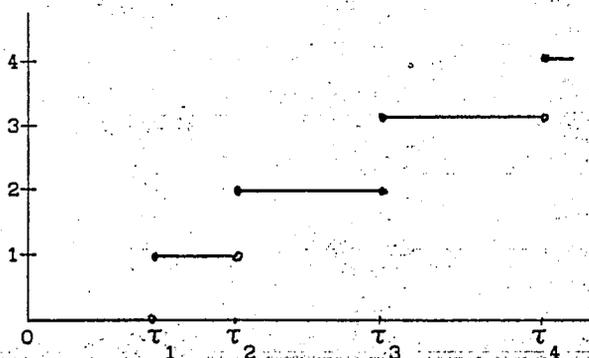


Fig. 2: Registro temporal del número de partículas alfa desprendidas del radioisótopo  $S_2$  hasta el instante actual

Si se repite el experimento un cierto número de veces, se obtendrán otras tantas trayectorias, todas distintas. Sin embargo, es posible detectar una regularidad de tipo estadístico, eligiendo eventos observables del tipo de

"se desprenden  $n$  partículas alfa en  $t$  segundos".

Es decir se conduce un gran número  $k$  de experimentos de duración  $t$  segundos, y se anota el número  $k_n$  de aquellos en los que se desprendieron exactamente  $n$  partículas. Se forma la frecuencia relativa  $k_n/k$  y se determina su comportamiento asintótico, observando que se asienta en torno a un valor, al cual llamamos  $p_n(t)$ : la probabilidad de que se desprendan exactamente  $n$  partículas alfa en  $t$  segundos.

Siguiendo a [Feller, 1968], podemos construir un modelo probabilístico para este fenómeno. Para ello conviene basarse en hipótesis convenientes, y las siguientes parecen razonables:

i) Si  $0 < s < t$ , el número de partículas alfa emitidas durante  $[s, t]$  no depende del número que fue emitido durante  $[0, s]$ .

ii) La probabilidad de que se emita exactamente una partícula alfa durante  $[t, t+h]$  es  $\lambda h + o(h)$ , donde  $\lambda$  es una constante positiva.

iii) La probabilidad de que se emita más de una partícula alfa durante  $[t, t+h]$  es  $o(h)$ .

iv) Para cada  $n \geq 0$ ,  $p_n(\cdot)$  es continuamente diferenciable.

v) Es seguro que en un intervalo de longitud cero no se desprenda ninguna partícula alfa.

iv) El comportamiento del fenómeno no depende del instante en que se comienza a observarlo, sino sólo de la duración de la observación.

Sea  $A_n(s, t)$  el evento consistente en que se desprendan exactamente  $n$  partículas durante el intervalo  $[s, t]$  con  $n \geq 0$ ,  $0 \leq s \leq t$ . Claramente este evento se observa siempre que se registre una trayectoria con exactamente  $n$  saltos entre los instantes  $s$  y  $t$ , por lo que  $A_n(s, t)$  puede identificarse con el conjunto de dichas trayectorias. Además, cualquier asignación de probabilidad a los eventos en  $\Omega_2$  debe ser tal que

$$P(A_n(s,t)) = p_n(t-s),$$

en virtud de vi).

Claramente

$$A_0(0,t+h) = A_0(0,t) \cap A_0(t,t+h),$$

por lo que, de i), ii), iii)

$$\begin{aligned} p_0(t+h) &= p_0(t)p_0(h) \\ &= p_0(t)(1-\lambda h - o(h)), \end{aligned}$$

es decir

$$\frac{p_0(t+h) - p_0(t)}{h} = -\lambda p_0(t) + \frac{o(h)}{h}$$

En virtud de iv) y v), se tiene entonces que

$$\dot{p}_0 = -\lambda p_0 \quad p_0(0)=1. \quad (5a)$$

Analogamente, para  $n \geq 1$

$$\begin{aligned} A_n(0,t+h) &= A_n(0,t) \cap A_0(t,t+h) \\ &+ A_{n-1}(0,t) \cap A_1(t,t+h) + \sum_{k=2}^n A_{n-k}(0,t) \cap A_k(t,t+h), \end{aligned}$$

donde  $+$  y  $\sum$  denotan unión ajena. Evidentemente, en virtud de i) tenemos que

$$p_n(t+h) = p_n(t)p_0(h) + p_{n-1}(t)p_1(h) + R_n,$$

donde  $R_n$  es la probabilidad del tercer término, que se refiere al desprendimiento de dos o más partículas alfa en el intervalo  $[t,t+h]$ .

Dado que

$$\sum_{k=2}^n A_{n-k}(0,t) \cap A_k(t,t+h) \subset \sum_{k=2}^{\infty} A_k(t,t+h),$$

de iii) resulta que

$$0 \leq R_n \leq P \left[ \sum_{k=2}^{\infty} A_k(t,t+h) \right] = o(h).$$

Tomado en cuenta lo anterior, además de i) y ii), obtenemos que

$$\frac{p_n(t+h) - p_n(t)}{h} = -\lambda p_n(t) + \lambda p_{n-1}(t) + o(h).$$

Finalmente, de iv) y v) se obtiene

$$\dot{p}_n = -\lambda p_n + \lambda p_{n-1}, \quad p_n(0) = 0, n \geq 1 \quad (5b)$$

al tomar límite cuando  $h \rightarrow 0$ .

De lo anterior resulta que la sucesión de probabilidades  $\{p_n, n = 0, 1, 2, \dots\}$  satisface el sistema de ecuaciones diferenciales ordinarias (5). Este es un problema de valores iniciales que se resuelve en forma recursiva sin mayor dificultad, llegando a que

$$p_n(t) = \frac{e^{-\lambda t} (\lambda t)^n}{n!} \quad n = 0, 1, 2, \dots \quad (6)$$

En el lenguaje usual de la Probabilidad, el resultado anterior se expresa diciendo que "el número de partículas alfa desprendidas en un intervalo de longitud  $t$  es Poisson con parámetro  $\lambda t$ ".

De esta manera se han probabilizado los eventos de la forma  $A_n(s, s+t)$ , para  $n \geq 0, s \geq 0, t \geq 0$ . Hay muchos otros eventos de interés aparte de estos, por lo que resulta de gran importancia completar la construcción de un espacio probabilístico

$$(\Omega_2, \mathcal{A}, P),$$

donde  $\mathcal{A}$  es una familia suficientemente grande de eventos, con

$$A_n(s, s+t) \in \mathcal{A}, \quad n \geq 0, s \geq 0, t \geq 0$$

y  $P: \mathcal{A} \rightarrow [0, 1]$  es una probabilidad en el sentido técnico del término. El espacio probabilístico resultante se denomina *proceso estocástico de Poisson*. Para los detalles de su construcción referimos al lector a [Doob, 1953]. Véase también el capítulo XVII de [Feller, 1968] para otros ejemplos de procesos estocásticos afines al de Poisson.

Como quiera, el ejemplo anterior ilustra de nuevo la importancia de la escala de observación sobre la percepción de la realidad; en este caso se trata de una escala temporal. En efecto, el mismo fenómeno aparece como determinístico, o no, según se le examine para tiempos largos (primer caso) o para tiempos cortos (segundo caso).

## CONFERENCIA 4.

### LO MICROSCOPICO Y LO MACROSCOPICO.

Esta conferencia busca ilustrar el efecto del punto de vista desde el que se estudia un fenómeno sobre el modelo que de él se hace. La tesis fundamental es que la óptica con la que se examina el mundo real determina nuestra percepción del mismo y, por ende, la descripción que de él damos. Aquí se profundiza lo ya enunciado en la conferencia anterior, acerca de la relatividad del determinismo. Se considera el fenómeno de la difusión molecular, el cual se estudia primero desde un punto de vista microscópico y luego desde un punto de vista macroscópico. En la sección 1 se da un modelo estocástico, el de Ornstein-Uhlenbeck para el movimiento de una partícula en un baño térmico, y se indican algunos casos límites del mismo. En la segunda sección se presenta el modelo clásico fenomenológico de la difusión molecular, tomando en cuenta la posibilidad de producción y aniquilamiento de materia, y enfatizando el planteamiento de las condiciones a la frontera.

#### 4.1 La difusión molecular.

Problema: "Se deja caer una gota de tinta negra en un recipiente con agua. Sin agitar, se deja formar la mezcla hasta que se llega a una solución homogénea de color gris oscuro (ver figura 1). Describese el fenómeno en términos matemáticos".

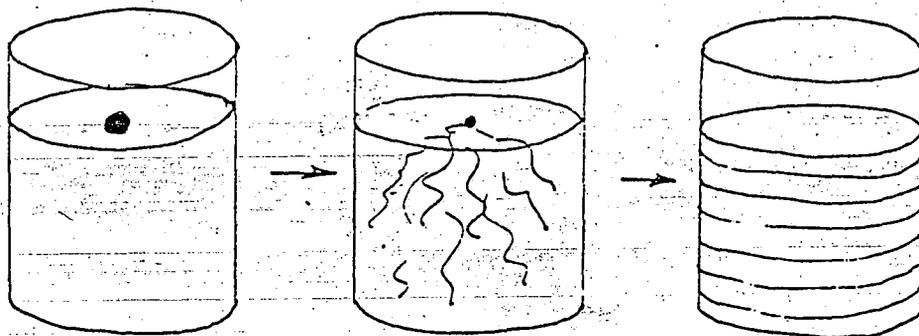


Fig. 1: El fenómeno de la difusión molecular

Evidentemente, las partículas de la tinta se mueven en el interior del fluido circundante (el agua), pues de otro modo no se

entiende cómo se da la homogeneización. Por cierto, las partículas que constituyen la tinta son gránulos de tamaño mucho mayor que el de las moléculas del agua, por lo que hay una clara distinción entre unas y otras. Por otro lado, las moléculas del agua también están en incesante movimiento, por lo que debe darse un sin número de choques entre ellas y cada uno de dichos gránulos. Estos choques, junto con la influencia que sobre la partícula de tinta puede ejercer un campo de fuerza externa, da como lugar el que la partícula se mueva en el interior de la solución; en esto consiste el fenómeno de la *difusión molecular*.

Por observación directa se sabe que las trayectorias de las partículas de tinta son bastante irregulares, y que resulta difícil predecirlas. Por un lado, se puede saber con cierta precisión la posición del punto de partida, pero quizá no sea lícito decir que en el interior de la gota de tinta las partículas se encuentran inicialmente en reposo. Por otro, aún cuando se conozca la fase inicial para cada partícula, el grandísimo número de moléculas de agua que la bombardean a alta velocidad, desde todas las direcciones, provoca efectos muy complicados sobre aquella. Además, es en extremo difícil caracterizar esta fuerza adicional, debido precisamente a su naturaleza fluctuante, y a que los cambios suceden con extrema rapidez. Así pues, conviene aislar esta parte del fenómeno y tratarla por métodos *ad hoc* que especificaremos a continuación. Baste por ahora decir que supondremos representables estos efectos mediante alguna fuerza externa variable con el tiempo  $F(t)$  y así aparecerá en las ecuaciones del movimiento de la partícula.

Por lo que respecta a la parte lenta del fenómeno (el movimiento de la partícula en sí), corresponde a la situación descrita en el Ejemplo 1.2.6. Para facilitar las cosas supongamos que el movimiento se da en una sola dimensión; sea  $m > 0$  la masa de la partícula que se difunde, y supongamos que el campo de fuerza dentro del cual se mueve está dado por un potencial  $U \in C^1(\mathbb{R})$ ; supongamos además que el medio ejerce una fuerza viscosa lineal en la velocidad con constante de proporcionalidad  $\beta > 0$ . Del ejemplo antes citado, las ecuaciones de movimiento son

$$\dot{x} = v \quad (1a)$$

$$\dot{v} = -\frac{1}{m} U'(x) - \frac{\beta}{m} v + \frac{1}{m} F(t), \quad (1b)$$

que deben complementarse con las condiciones iniciales

$$x(0) = x_0, \quad v(0) = v_0, \quad (1c)$$

si se desea elegir una trayectoria de entre todas las posibles.

En efecto, supongamos por ahora que la fuerza externa  $F$  es continua y que el único campo de fuerza que actúa es el de la gravedad [ $U(x) = mgx$ ], por lo que el problema a resolver es

$$\dot{x} = v \quad x(0) = x_0 \quad (2a)$$

$$\dot{v} = -\frac{\beta}{m} v - g + \frac{1}{m} F(t), \quad v(0) = v_0 \quad (2b)$$

La segunda de estas ecuaciones se llama de *Langevin*. La solución de este problema es

$$x(t) = x_0 - \frac{mg}{\beta} t + \frac{m}{\beta} \left[ v_0 + \frac{mg}{\beta} \right] \left[ 1 - e^{-\frac{\beta t}{m}} \right] + \frac{1}{\beta} \int_0^t \left[ 1 - e^{-\frac{\beta}{m}(t-s)} \right] F(s) ds \quad (3a)$$

$$v(t) = -\frac{mg}{\beta} + \left[ v_0 + \frac{mg}{\beta} \right] e^{-\frac{\beta t}{m}} + \frac{1}{m} \int_0^t e^{-\frac{\beta}{m}(t-s)} F(s) ds \quad (3b)$$

Inmediatamente nos damos cuenta de algunas cosas:

a) Tanto la velocidad como la posición pueden descomponerse en un transitorio, una parte permanente y otra fluctuante.

b) Para la velocidad, el régimen permanente da una velocidad constante de asentamiento, cuya magnitud es  $mg/\beta$ . El transitorio desaparece exponencialmente.

c) Para la posición de la partícula, el régimen estacionario puede a su vez descomponerse en dos componentes: una de movimiento rectilíneo uniforme a la velocidad límite  $(x_0 - \frac{mg}{\beta} t)$  y una derivada constante  $\frac{mv_0}{\beta} + g \left( \frac{m}{\beta} \right)^2$ , la primera de cuyas partes se debe a la velocidad inicial, la segunda al tirón de la gravedad. Aquí también el transitorio desaparece con rapidez exponencial.

d) Por último, la fuerza fluctuante da lugar a fluctuaciones en posición y velocidad dadas por integrales del tipo

$$\int_0^t \varphi(s)F(s)ds, \quad (4)$$

donde  $\varphi$  es una función continua.

Así pues, sin la presencia de la fuerza fluctuante la partícula se desplazaría en movimiento rectilíneo uniforme. El efecto de los choques intermoleculares hace que en cambio se de el movimiento errático observado. En otras palabras, la fuerza fluctuante  $F$  no puede ser una función continua: de hecho, debe ser tal que las integrales como (4) sean erráticas, para cada instante  $t$ .

En [Hernández, 1988] hemos dado en forma elemental la construcción de un objeto matemático llamado *ruido blanco*, que es un proceso estocástico generalizado, para el cual tienen sentido las integrales del tipo

$$\rho(t) := \int_0^t \varphi(s)\xi(s)ds, \quad (5a)$$

llamadas *de Wiener*. Una tal integral define una variable aleatoria gaussiana para cada  $t$ , de tal manera que  $\{\rho(t), t \in \mathbb{R}\}$  constituye un proceso estocástico gaussiano. Además,

$$E \rho(t) = 0, \quad E \rho(t)^2 = \int_0^t |\varphi(s)|^2 ds; \quad (5b)$$

en general,

$$E \left[ \int_0^t \varphi(s)\xi(s)ds \int_0^t \psi(s)\xi(s)ds \right] = \int_0^t \varphi(s)\psi(s)ds \quad (5c)$$

Así pues, podemos representar la fuerza fluctuante  $F$  como un múltiplo escalar de ruido blanco, digamos  $F = \sigma\xi$ , con  $\sigma > 0$ ; las fluctuaciones en la posición y velocidad son los procesos estocásticos gaussianos

$$\left\{ \frac{\sigma}{\beta} \int_0^t \left[ 1 - e^{-\frac{\beta}{m}(t-s)} \right] \xi(s) ds, t \in \mathbb{R} \right\} \quad (6a)$$

y

$$\left\{ \frac{\sigma}{m} \int_0^t e^{-\frac{\beta}{m}(t-s)} \xi(s) ds, t \in \mathbb{R} \right\}, \quad (6b)$$

respectivamente. Un cálculo elemental basado en (5) muestra que la

media del proceso gaussiano bidimensional (6) es cero, en tanto que su matriz de covarianza es

$$C(t) := \sigma^2 \begin{pmatrix} \frac{1}{\beta^2} \left[ e^{-\frac{\beta t}{m}} - 1 + \frac{2\beta t}{m} \right] & \frac{1}{m\beta} \left[ 1 - e^{-\frac{\beta t}{m}} \right] \\ \frac{1}{\beta} \left[ 1 - e^{-\frac{\beta t}{m}} \right] & \frac{1}{m^2} \end{pmatrix} \quad (7a)$$

Esta es también la matriz de covarianza del proceso estocástico combinado  $\{(x(t), v(t)), t \in \mathbb{R}\}$ , cuya media es

$$\mu(t) = \begin{pmatrix} x_0 + \frac{m}{\beta} \left[ v_0 + \frac{mg}{\beta} - gt \right] \\ -\frac{mg}{\beta} \end{pmatrix} + \begin{pmatrix} \frac{m}{\beta} \\ 1 \end{pmatrix} e^{-\frac{\beta t}{m}} \quad (7b)$$

Asintóticamente, el proceso combinado  $\{(x(t), v(t)), t \in \mathbb{R}\}$  se comportará como un proceso gaussiano, cuya media y covarianza se obtiene fácilmente de (7).

Para la difusión de la tinta en el agua no se observa ningún asentamiento, lo que indica que la fuerza fluctuante producto de los choques intermoleculares es suficiente para neutralizar el tirón de la gravedad. Esto se debe ciertamente a que la masa de las partículas de tinta es pequeña, y sería muy distinto si se tratara de partículas mayores. Es de interés para este caso desprestigiar el efecto de la gravedad, tomando  $g=0$  en (2); supondremos además que la partícula parte del reposo.

De todo ello se obtiene que, para tiempos grandes

$$E|x(t) - x_0|^2 \approx \frac{\sigma^2}{m\beta} t.$$

Por otro lado, del análisis de [Einstein, 1905] se desprende que, si se alcanza el equilibrio a temperatura  $T$ , necesariamente

$$\sigma^2 = 2\beta kT, \quad (8)$$

donde  $k$  es la constante de Boltzmann. Así pues, se llega a que

$$E|x(t) - x_0|^2 \approx \left(\frac{2kT}{m}\right)t, \quad (9)$$

en el equilibrio (es decir, para tiempos grandes) a temperatura  $T$ . Esta última relación es muy importante, pues permite validar este modelo: en efecto, por medición directa se obtiene una estimación del desplazamiento medio cuadrático de la partícula a distintos instantes de tiempo; si no varía linealmente con  $t$ , el modelo se rechaza.

El análisis anterior se puede complicar cuando se tiene difusión molecular bajo la influencia de campos externos. Un caso en el que todavía se pueden obtener conclusiones en forma explícita corresponde a la difusión en un medio elástico, que ejerce una fuerza restitutiva dada por la ley de Hooke, es decir,  $U(x) = \frac{1}{2}\alpha x^2$ , con  $\alpha > 0$ . Si la fuerza fluctuante es de nuevo  $F = \sigma\xi$ , con  $\xi$  ruido blanco gaussiano estándar, las ecuaciones a resolver son

$$\begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{\alpha}{m} & -\frac{\beta}{m} \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\sigma}{m} \end{pmatrix} \xi(t); \quad \begin{pmatrix} x(0) \\ v(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix};$$

como son lineales, se pueden resolver en forma explícita usando la integral de Wiener. El análisis original de este modelo se debe a Ornstein y Uhlenbeck; data de 1930 y no usa la integral de Wiener. Véase los artículos correspondientes reimpresos en [Wax, 1954]. En general, el término  $U'(x)$  no será lineal, y cuando  $F = \sigma\xi$  (con  $\xi$  ruido blanco) resolver las ecuaciones diferenciales

$$\dot{x} = v, \quad x(0) = x_0 \quad (10a)$$

$$\dot{v} = -\frac{1}{m} U'(x) - \frac{\beta}{m} v + \frac{\sigma}{m} \xi, \quad v(0) = v_0 \quad (10b)$$

requiere de la integral de Ito [Schuss, 1980].

Por otro lado, se puede simplificar el modelo (10) procediendo como sigue. Suméjase éste en una familia infinita de modelos

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= -\frac{1}{m} U'(x) - \frac{\beta}{\epsilon m} v + \frac{\sigma}{m} \xi, \end{aligned}$$

uno para cada  $\epsilon \in (0, 1]$ ; aquí,  $\sigma_1 = \sigma$  y, en general, si queremos que valga la relación de Einstein para cada uno de estos modelos, debe tenerse

que

$$\sigma_\varepsilon^2 = 2\beta kT/\varepsilon, = \frac{\sigma^2}{\varepsilon^2}$$

por lo que los modelos a considerar son

$$\dot{x} = v \quad (11a)$$

$$\dot{v} = -\frac{1}{m} U'(x) - \frac{\beta}{\varepsilon m} v + \frac{\sigma}{m\sqrt{\varepsilon}} \xi, \quad (11b)$$

uno para cada  $\varepsilon \in (0, 1]$ .

Cambiamos de escala temporal, con objeto de estudiar este sistema para tiempos largos; definamos la nueva escala temporal  $s := \varepsilon t$ , las nuevas variables

$$x^\varepsilon(s) := x(s/\varepsilon), \quad v^\varepsilon(s) := \varepsilon v(s/\varepsilon), \quad \xi^\varepsilon(s) := \sqrt{\varepsilon} \xi(s/\varepsilon).$$

Hecha la substitución, la familia de modelos (11) da lugar a

$$\frac{dx^\varepsilon}{ds} = v^\varepsilon$$

$$\varepsilon \frac{dv^\varepsilon}{ds} = -\frac{1}{m} U'(x^\varepsilon) - \frac{\beta}{m} v^\varepsilon + \frac{\sigma}{m} \xi^\varepsilon,$$

es decir

$$\frac{\beta}{m} \frac{dx^\varepsilon}{ds} + \varepsilon \frac{dv^\varepsilon}{ds} = \frac{1}{m} U'(x^\varepsilon) + \frac{\sigma}{m} \xi^\varepsilon.$$

Formalmente, al hacer que  $\varepsilon \rightarrow 0$  desaparece el segundo término a la izquierda de la última igualdad. Si, además  $x^\varepsilon \rightarrow y$ ,  $v^\varepsilon \rightarrow u$ ,  $\xi^\varepsilon \rightarrow \eta$ , con  $\eta$  ruido blanco estándar, entonces se tendrá, en el límite,

$$\beta \dot{y} = -U'(y) + \sigma \eta. \quad (12)$$

Este nuevo modelo es precisamente el dado en 1916 por Smoluchowski; para el movimiento browniano, véase [Smoluchowski, 1916]. Para la justificación rigurosa del paso de (11) a (12), véase [Nelson, 1967]. En particular, si se trata de una partícula que se difunde libremente, el modelo de Smoluchowski se reduce al de Einstein. En general, es verdad que el comportamiento predicho para tiempos largos por el modelo de Smoluchowski (12) concuerda con las que arroja el de Ornstein-Uhlenbeck (10), por lo que no se puede distinguir uno del otro al menos desde este punto de vista. Véase el libro de Nelson arriba citado para una discusión al respecto.

## 4.2 Reacción y difusión en dimensión uno.

Consideremos la siguiente variante del problema estudiado en la sección anterior.

Problema: "Un tubo cilíndrico muy esbelto se llena de agua. En un instante dado se le inyecta tinta negra, que se deja difundir hasta que se uniformiza la solución en un color gris oscuro. Se desea estudiar la variación temporal de la distribución espacial de la tinta, bajo la hipótesis de que el tubo se encuentre sellado en los extremos".

Sea  $\ell$  la longitud del tubo. Dado que se trata de un tubo esbelto, ignoraremos las variaciones de concentración y cualquier otro atributo de interés en cada sección transversal, por lo que podemos tomar  $S = (0, \ell)$  como representante del espacio físico. El tiempo será representado por los elementos de  $T := \mathbb{R}$ . Hay un sólo atributo de interés, la concentración de tinta, y se mide mediante números no negativos; por lo tanto,  $A = [0, \infty)$ . El fenómeno se estudia tomando muestras a lo largo del tubo, en un número finito de puntos y analizando por fotometría el fluido colectado, para determinar la concentración de tinta. Dichas determinaciones se efectúan con frecuencia, y el resultado es una tabla

$$t_i, C_1^{(i)}, \dots, C_p^{(i)} \quad i=1, \dots, r \quad (1)$$

donde  $r$  es el número de repeticiones y  $p$  es el número de puntos muestreados, a saber  $z_1, \dots, z_p \in (0, \ell)$ .

Los perfiles de concentración de tinta serán funciones  $x: [0, \ell] \rightarrow A$ ; puesto que no se esperan variaciones bruscas de composición a lo largo del tubo, los perfiles se supondrán continuos. De hecho, por conveniencia matemática supondremos (véase la sección 2.2) que los perfiles son continuamente diferenciables en  $[0, \ell]$ , con segunda derivada continua en  $(0, \ell)$ . Esta suposición resultará útil, según veremos a continuación.

De la tabla (1) se obtiene la tabla derivada

$$\frac{C_{j+1}^{(1)} - C_j^{(1)}}{z_{j+1} - z_j} \quad j=1, \dots, p-1; \quad i=1, \dots, r$$

con lo que se estiman los gradientes de concentración a lo largo del tubo. A la vez, se determinan los cocientes

$$\frac{1}{\pi R^2} \frac{C_j^{(i+1)} - C_j^{(i)}}{t_{i+1} - t_i} \quad i = 1, \dots, r-1; \quad j=1, \dots, p$$

donde  $R$  es el radio interno del tubo. Así pues, se determina experimentalmente el transporte del soluto por unidad de área a través de cada una de las secciones transversales del tubo determinadas por los puntos de observación.

Observaciones experimentales muestran que el transporte de materia se da en la dirección en que disminuye la concentración, y que la tasa de transporte por unidad de tiempo y por unidad de área es proporcional al gradiente local de concentraciones. Esta afirmación empírica constituye la llamada *Ley de Fick* (recuérdese la ley de Fourier de la sección 2.2) y se puede traducir en términos matemáticos como sigue: si  $x[0, \ell] \rightarrow A$  es un perfil de concentraciones, y si  $z \in [0, \ell]$ , entonces el flujo por unidad de trabajo a través de la sección transversal determinada por  $z$  es

$$-D \frac{dx}{dz}, \quad (2)$$

donde  $D$  es un coeficiente positivo, llamado la *difusividad* de la tinta con el agua a la temperatura del experimento; en el sistema cgs, sus unidades son  $\text{cm}^2/\text{seg}$ . Esta ley, junto con la ley de la conservación de la materia, constituye la herramienta principal para el análisis de estas situaciones.

Por otro lado los perfiles aceptables deben reflejar el hecho de que no hay difusión en los extremos de la barra. Según la ley de Fick, ello requiere que

$$Dx'(0)=0, \quad Dx'(\ell)=0,$$

por lo que el conjunto de los perfiles admisibles es

$$E := \{x \in C^1[0, \ell] : x \in C^2(0, \ell), \quad Dx'(0)=0, \quad Dx'(\ell)=0\}.$$

A su vez, las trayectorias admisibles serán algunas funciones

$\omega: T \rightarrow E$ ; veamos cuales.

Para comenzar, introduzcamos la notación

$$C(t, z) := \omega(t)(z),$$

por lo que debe cumplirse que

$$D \frac{\partial C}{\partial z}(t, 0) = 0, \quad D \frac{\partial C}{\partial z}(t, \ell) = 0 \quad (3)$$

Sea  $h > 0$ , y sean  $z, z+h \in (0, \ell)$ ,  $t \in \mathbb{R}$ . Igual que hicimos en la sección 2.2 para estudiar la conducción de calor en una barra esbelta, analicemos los eventos que tienen lugar en el instante  $t$  en la sección del tubo determinado por dos planos perpendiculares a su eje, colocados en cada uno de estos dos puntos (véase la figura 1).

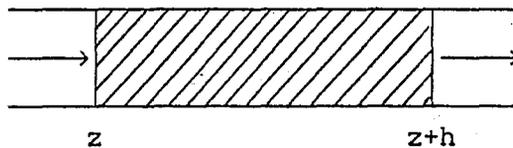


Fig. 1: Sección del tubo en el que se estudia la difusión

(compare con la figura 2.2.2). Por la sección de la izquierda entra un flujo de

$$-D(\pi R^2) \frac{\partial C}{\partial z}(t, z+h) \frac{g}{\text{seg}}$$

de tinta, en tanto que por la derecha sale un flujo de

$$\pi R^2 \left[ D \frac{\partial C}{\partial z}(t, z+h) - D \frac{\partial C}{\partial z}(t, z) \right].$$

Por otro lado, la cantidad de tinta contenida en dicha región es

$$(\pi R^2) \int_z^{z+h} C(t, \zeta) d\zeta.$$

Si  $\frac{\partial C}{\partial t}$  es continuo, entonces la tasa de variación del contenido de tinta en esa región es

$$(\pi R^2) \frac{d}{dt} \int_z^{z+h} C(t, \zeta) d\zeta = (\pi R^2) \int_z^{z+h} \frac{\partial C}{\partial t}(t, \zeta) d\zeta$$

Por lo tanto

$$\int_z^{z+h} \frac{\partial C}{\partial t}(t, \zeta) d\zeta = D \frac{\partial C}{\partial z}(t, z+h) - D \frac{\partial C}{\partial z}(t, z).$$

Basta dividir ambos miembros por  $h$  y tomar el límite cuando  $h \rightarrow 0$  para obtener que

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial z} \left( D \frac{\partial C}{\partial z} \right), \quad t \in \mathbb{R}, z \in (0, \ell) \quad (4)$$

Así pues, el conjunto de las trayectorias aceptables es

$$\Omega = \{C: \mathbb{R} \times (0, \ell) \rightarrow [0, \infty) \mid \frac{\partial C}{\partial t}, \frac{\partial^2 C}{\partial z^2} \in C[\mathbb{R} \times (0, \ell)], (2) \& (3)\}.$$

Aquí hemos supuesto que  $D$  es una función continuamente diferenciable de la posición. De hecho, a menudo se toma  $D$  constante, por lo que (4) se transforma en

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial z^2}, \quad t \in \mathbb{R}, 0 < z < \ell \quad (5)$$

En general, el coeficiente de difusión será una función continuamente diferenciable de  $C$ , por lo que (4) es una ecuación diferencial parcial no lineal. En el caso más sencillo en que  $D$  es constante, la ecuación diferencial que resulta (5) se llama la *ecuación de difusión*. En lo que sigue nos concentraremos en este caso particular, por lo que

$$\Omega = \{C: \mathbb{R} \times (0, \ell) \rightarrow [0, \infty) \mid \frac{\partial C}{\partial t}, \frac{\partial^2 C}{\partial z^2} \in C[\mathbb{R} \times (0, \ell)], (2) \& (4)\}.$$

Una variante de gran interés se obtiene cuando el fluido no permanece estático, sino que se mueve con velocidad conocida. De hecho, sea  $v: [0, \ell] \rightarrow \mathbb{R}$  el perfil de velocidades en el fluido, que supondremos se mantiene durante todo el desarrollo del fenómeno. Supongamos además que  $v$  es continua.

Bajo estas circunstancias, el análisis anterior permanece, siempre que se efectúe sobre un elemento que se mueve con el fluido. En otras palabras, los puntos del tubo se mueven a lo largo de las soluciones de la ecuación diferencial

$$\dot{z} = v(z),$$

y la función

$$t \mapsto C(t, z(t))$$

satisface la ecuación de difusión, es decir

$$\frac{d}{dt} C(t, z(t)) = D \frac{\partial^2 C}{\partial z^2} (t, z(t))$$

Aplicando la Regla de Cadena obtenemos sin dificultad que

$$\frac{\partial C}{\partial t} + v(z) \frac{\partial C}{\partial z} = D \frac{\partial^2 C}{\partial z^2} \quad (6)$$

Una extensión con mayor interés aún se obtiene si, en lugar de tinta, se difunde con el agua una sustancia que toma parte en una reacción química que se desarrolla en el seno del fluido.

Sea  $\varphi: [0, \infty) \rightarrow \mathbb{R}$  una función continuamente diferenciable tal que

$$\varphi(C) = \text{tasa de desaparición de la sustancia que se difunde, por unidad de volumen} \left[ \frac{\text{g}}{\text{cm}^3 \cdot \text{seg}} \right].$$

Entonces, en el interior del elemento bajo estudio se consumen

$$(\pi R^2) \int_z^{z+h} \varphi(C(t, \zeta)) d\zeta \quad \frac{\text{g}}{\text{seg}}$$

por lo que la tasa neta con que se acumula el soluto contenido en el elemento bajo estudio es

$$\pi R^2 \left\{ - \int_z^{z+h} \varphi(C(t, \zeta)) d\zeta + D \frac{\partial C}{\partial z}(t, z+h) - D \frac{\partial C}{\partial z}(t, z) \right\}.$$

En consecuencia, podemos afirmar que

$$\int_z^{z+h} \frac{\partial C}{\partial z}(t, \zeta) d\zeta = D \frac{\partial C}{\partial z}(t, z+h) - D \frac{\partial C}{\partial z}(t, z) - \int_z^{z+h} \varphi(C(t, \zeta)) d\zeta.$$

De nuevo, basta dividir entre  $h$  y tomar el límite cuando  $h \rightarrow 0$  para obtener

$$\frac{\partial C}{\partial t} + v(z) \frac{\partial C}{\partial z} = D \frac{\partial^2 C}{\partial z^2} - \varphi(C). \quad (7)$$

Un caso particular especialmente simple resulta cuando la reacción química en cuestión es de primer orden, es decir

$$\varphi(C) = kC,$$

pues entonces la ecuación (7) da lugar a

$$\frac{\partial C}{\partial t} + v(z) \frac{\partial C}{\partial z} = D \frac{\partial^2 C}{\partial z^2} - kC. \quad (8)$$

Esta ecuación diferencial parcial es lineal y, junto con las condiciones a la frontera (3), basta especificar una condición inicial  $C(0, \cdot)$  para determinar la solución  $C$ , usando la técnica de separación de variables que ya se utilizó en la sección 2.2.

Salvo en este caso particular, la ecuación diferencial de reacción con difusión (7) será no lineal, precisamente a causa del término de reacción  $\varphi(C)$ . En cambio, si sólo hay difusión y no hay reacción, la ecuación diferencial resultante será siempre lineal, salvo cuando el coeficiente de difusión sea función de la concentración de la sustancia que se difunde.

El método de separación de variables permite resolver los problemas de valores a la frontera siempre y cuando no presenten ninguna no linealidad. En general, para resolver los problemas de valores iniciales y a la frontera resultantes habrá que recurrir a los métodos numéricos u otras técnicas de aproximación.

Pero regresemos a la ley de Fick, según la cual el flujo por unidad de área perpendicular al eje del cilindro está dado por (1), cuando el perfil de concentración es  $x$ . Esta ley es una aproximación. En general dicho fluido en el punto  $z \in [0, \ell]$  será función de  $x'(z)$ , no necesariamente proporcional a este valor. Es también claro que el flujo va en la dirección en que disminuye la concentración, por lo que debe decrecer con  $x'(z)$ . Además, en muchas situaciones de interés el medio no será uniforme, por lo que la dependencia antes descrita puede cambiar con el punto  $z$ . Así pues, podemos postular que existe una función  $j: [0, \ell] \times [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ , tal que

$j(z, x(z), x'(z)) :=$  flujo por unidad de área perpendicular al eje del cilindro, a través de la sección transversal en  $z$ , cuando el perfil de concentración es  $x$ .

Conviene suponer además que  $j$  es continuamente diferenciable, y de la

discusión anterior se desprende que, si denotamos las coordenadas de  $[0, \ell] \times [0, \infty) \times \mathbb{R}$  como  $(z, u, w)$  entonces  $\frac{\partial j}{\partial w} < 0$ .

Bajo estas condiciones, la tasa neta de transporte a través del volumen achurado en la figura 1 es

$$\pi R^2 \left[ -j(z+h, x(z+h), x'(z+h)) + j(z, x(z), x'(z)) \right]$$

y el balance de materia ahí cuando además hay reacción dice que

$$\begin{aligned} & \int_z^{z+h} \frac{\partial C}{\partial t}(t, \zeta) d\zeta + j(z+h, x(z+h), x'(z+h)) - j(z, x(z), x'(z)) \\ &= - \int_z^{z+h} \varphi(C(t, \zeta)) d\zeta. \end{aligned}$$

De nuevo, basta dividir entre  $h$  y tomar el límite cuando  $h \rightarrow 0$  para obtener la relación

$$\frac{\partial C}{\partial t} + \frac{dj}{dz} \left( z, C, \frac{\partial C}{\partial z}, \frac{\partial^2 C}{\partial z^2} \right) = -\varphi(C), \quad (9)$$

donde

$$\frac{\partial j}{\partial z}(z, u, w, y) = \frac{\partial j}{\partial z}(z, u, w) + \frac{\partial j}{\partial u}(z, u, w)w + \frac{\partial j}{\partial w}(z, u, w)y;$$

(9) se llama *ecuación de continuidad*. Si además el fluido está en movimiento, obtenemos la *ecuación de continuidad para problemas con flujo*, a saber

$$\frac{\partial C}{\partial t} + v(z) \frac{\partial C}{\partial z} + \frac{dj}{dz} \left( z, C, \frac{\partial C}{\partial z}, \frac{\partial^2 C}{\partial z^2} \right) = -\varphi(C) \quad (10)$$

En cuanto a las condiciones a la frontera, en lugar de (3) se tendrá

$$j(0, C(t, 0), \frac{\partial C}{\partial z}(t, 0)) = 0, \quad j(\ell, C(t, \ell), \frac{\partial C}{\partial z}(t, \ell)) = 0$$

cómo expresión matemática del hecho que el tubo se encuentra aislado en los extremos. Con mayor generalidad, podemos pensar en fijar el flujo en esos puntos (no necesariamente anularlo), y entonces la condición a la frontera sería

$$j(0, C(t, 0), \frac{\partial C}{\partial z}(t, 0)) = \psi(0) \quad (11a)$$

$$j(l, C(t, l), \frac{\partial C}{\partial z}(t, l)) = \psi(l) \quad (11b)$$

donde  $\psi$  es una función conocida, definida en  $\{0, l\}$  y cuyos valores son los flujos en cuestión. Si de hecho se cumple la Ley de Fick, (11) se transforma en

$$-D \frac{\partial C}{\partial z}(t, 0) = \psi(0), \quad -D \frac{\partial C}{\partial z}(t, l) = \psi(l) \quad (12)$$

llamadas *condiciones de Neumann*. Entonces, las relaciones (11) son unas condiciones de Neumann generalizadas.

Si el tubo no está cerrado en el extremo inicial, sino que la corriente fluye a través de él, y si  $C_0$  es la concentración a la entrada, entonces se da un flujo de materia adicional en ese punto. Como primera aproximación podemos considerar que dicho flujo es proporcional a la diferencia de concentraciones  $C(t, 0) - C_0$ , con constante de proporcionalidad  $\gamma_0$ . El flujo hacia el interior del tubo es entonces  $\psi(0) = -\gamma_0 [C(t, 0) - C_0]$ . Análogamente, si el extremo de la derecha está abierto y hay una concentración  $C_1$  fuera del tubo, se genera un flujo de materia dado por  $\psi(l) = \gamma_1 [C(t, l) - C_1]$ , donde  $\gamma_1$  es una constante positiva conocida. Tanto  $\gamma_0$  como  $\gamma_1$  tienen unidades de velocidad.

Como en ninguna de las entradas del tubo se puede acumular el material, debemos tener que

$$j(0, C(t, 0), \frac{\partial C}{\partial z}(t, 0)) = -\gamma_0 [C(t, 0) - C_0] \quad (13a)$$

$$j(l, C(t, l), \frac{\partial C}{\partial z}(t, l)) = +\gamma_1 [C(t, l) - C_1] \quad (13b)$$

Si vale la ley de Fick, la condición (13) se transforma en

$$-D \frac{\partial C}{\partial z} + \gamma_0 C \Big|_{z=0} = \gamma_0 C_0 \quad (14a)$$

$$D \frac{\partial C}{\partial z} + \gamma_1 C \Big|_{z=l} = \gamma_1 C_1 \quad (14b)$$

Estas se llaman *condiciones de Robin*, por lo que las condiciones (13) serían de Robin generalizadas.

Nótese que también (12) y (14) pueden ser no lineales, cuando el coeficiente de difusión depende de la concentración. Supongamos que  $D$  es constante en (14) y dividamos (14a) por  $\gamma_0 C_0$ , (14b) por  $\gamma_1 C_1$ , para obtener

$$-\frac{1}{v_0} \frac{\partial u}{\partial \sigma} + u \Big|_{\sigma=0} = 1, \quad \frac{1}{v_1} \frac{\partial v}{\partial \sigma} + v \Big|_{\sigma=1} = 1,$$

donde hemos definido las variables adimensionales

$$\sigma := \frac{z}{l}, \quad u := \frac{C}{C_0}, \quad v := \frac{C}{C_1},$$

además de los grupos adimensionales

$$v_0 := \frac{\gamma_0 l}{D}, \quad v_1 := \frac{\gamma_1 l}{D}$$

Si estos grupos adimensionales toman valores muy grandes (la transferencia hacia el exterior le gana a la difusión interna), podemos hacer que  $v_0 \rightarrow \infty$ ,  $v_1 \rightarrow \infty$ . En este caso se tendrá  $u=1$  en  $\sigma=0$ ,  $v=1$  en  $\sigma=1$ , es decir

$$C(t,0) = C_0, \quad C(t,l) = C_1. \quad (15)$$

Estas últimas condiciones a la frontera se llaman *de Dirichlet*; como hemos visto, se aplican cuando vale la ley de Fick en el interior, pero la difusión hacia el exterior es más rápida que la difusión molecular en el interior del tubo: como consecuencia de ello, se establece rápidamente el equilibrio con las corrientes externas.

Desde luego, consideraciones análogas se aplican al modelo matemático para la conducción de calor en una barra construido en la sección 2.2.

## CONFERENCIA 5.

### ALGUNAS CONSIDERACIONES METODOLOGICAS.

El objetivo de esta última conferencia es el de aislar algunos principios generales que gobiernan tanto el planteamiento de problemas como la validación de los modelos resultantes. Se enfatiza la metodología consistente en "primero plantear un principio de conservación y luego elegir un modo de calcular los diferentes términos de la ecuación", para construir un modelo matemático. Dicha metodología ya había sido utilizada sistemáticamente en las cuatro conferencias precedentes, y aquí se pone una vez más de manifiesto mediante la consideración de circuitos eléctricos sencillos. En cuanto a la validación de los modelos, se propone en la sección 2 a la inferencia estadística (estimación y prueba de hipótesis) como marco conceptual, ayudada por la optimización. La perspectiva global es la de buscar un adecuamiento entre las predicciones del modelo y las observaciones practicadas sobre el sistema real.

#### 5.1 Reglas para el planteamiento de problemas.

Regresemos por un momento al ejemplo 2.6 de la conferencia 1, referente a una partícula que se mueve en un medio viscoso bajo la acción de un potencial y una fuerza externa variable. La energía total de dicho sistema está dada por la función hamiltoniana

$$H(x, v) = \frac{1}{2} m v^2 + U(x),$$

y las ecuaciones de movimiento son

$$\dot{x} = v$$

$$\dot{v} = -\frac{1}{m} U'(x) - \frac{\beta}{m} v + \frac{1}{m} F(t).$$

A lo largo de una trayectoria de este sistema en el espacio fase, la energía tome los valores

$$E(t) = H(x(t), v(t)),$$

por lo que

$$\frac{dE}{dt} = v(t)F(t) - \beta v(t)^2. \quad (1)$$

Basta que elijamos la fuerza externa en función del estado de movimiento como

$$F(t) = \beta v(t) \quad (2)$$

para que la energía total se conserve. Es decir, el control externo (2) alimenta energía al sistema con la tasa justa para contrarrestar la tasa de disipación de energía producto de la fuerza viscosa. En general, al sistema se alimenta energía con una tasa (potencia) instantánea de

$$F(t)v(t) \frac{\text{erg}}{\text{seg}},$$

y se extrae energía con una potencia instantánea de

$$\beta v(t)^2 \frac{\text{erg}}{\text{seg}}.$$

El resultado neto es una tasa de acumulación de energía dada por (1). Observamos que

$$\text{Tasa de acumulación} = \text{Tasa de entrada} - \text{tasa de salida}, \quad (3)$$

La relación (3) fue aplicada en este caso a la energía mecánica. Sin embargo, en la sección 2 de la conferencia anterior se aplicó a la masa, al plantear las ecuaciones que gobiernan la difusión molecular cuando se le estudia de un punto de vista microscópico; y en la sección 2.2 se aplicó a la energía térmica, en forma análoga. Las mismas ecuaciones de movimiento que dan lugar a (1) son un caso particular de (3), pero aplicada a la conservación del impulso lineal  $p := mv$ . En efecto, la segunda de dichas ecuaciones es

$$\frac{dp}{dt} = -U'(x) - \beta v + F(t)$$

es decir

$$\frac{dp}{dt} = \frac{d}{dt} \int_0^t [F(s) - U'(x(s))] ds - \frac{d}{dt} \int_0^t \beta U(s) ds.$$

Vemos que, en un sistema mecánico, las fuerzas externas generan un flujo de impulso lineal hacia el sistema, cuyo valor en el instante  $t$  es

$$\int_0^t [F(s) - U'(x(s))] ds.$$

Analogamente, el impulso lineal que fluye hacia el exterior del sistema en el instante  $t$  es

$$\int_0^t \beta v(s) ds.$$

Vemos pues que la Segunda Ley de Newton no es sino una expresión del Principio de Conservación del Impulso Lineal. En general, toda Ley de Conservación tendrá la apariencia de (3). El siguiente ejemplo servirá para ilustrar estos conceptos.

#### Ejemplo 1.

"Se alimenta una corriente de salmuera concentrada a un tanque de difusión. Ahí se mezcla con una corriente de agua pura y se homogeneiza la mezcla mediante un poderoso agitador. Se mide la concentración de las corrientes de salmuera a la entrada y a la salida del tanque. Se opera el tanque de tal manera que ni haya derrames ni se vacíe".

Un punto clave en la construcción de un modelo para esta situación es la especificación del espacio. Suponiendo que se trate de un tanque cilíndrico, sería natural tomar  $S$  como un cilindro en  $\mathbb{R}^3$ , con lo que los perfiles de concentración serían funciones no negativas de tres variables,  $C: S \rightarrow [0, \infty)$ . Sin embargo, se nos informa que el agitador es "poderoso", de lo cual surge la posibilidad de suponer instantánea la homogenización de la mezcla en el tanque. Así pues, suponiendo mezclado perfecto, la salmuera en el tanque tiene la misma concentración de la salmuera en la corriente de salida. Por lo tanto  $S$  se puede reducir a un punto y los perfiles de concentración se pueden identificar con los valores de la misma. Desde luego, el tiempo será continuo ( $T = \mathbb{R}$ ).

Sin embargo, la concentración no es el único atributo de interés para plantear la conservación de la materia (3): debemos recordar que la sal presente en el tanque en un instante dado se obtiene

multiplicando la concentración de sal por el volumen de la salmuera. Así pues,  $A = [0, \infty)^2$ : si  $(x, y) \in A$ ,  $x$  representa la concentración e  $y$  el volumen de la salmuera en el tanque.

De acuerdo con lo anterior, las trayectorias serán parejas de funciones no negativas, de variable real; las supondremos además continuamente diferenciables para poder hablar de tasa de variación y aplicar (3). Sea  $(C, v): \mathbb{R} \rightarrow A$  una trayectoria tal, de manera que

$C(t)$  = concentración de la salmuera en el tanque,

$v(t)$  = volumen de la salmuera en el tanque,

ambos en el instante  $t$ .

Sea  $C_0$  la concentración de la salmuera concentrada que se alimenta al tanque, con una rapidez (variable) de  $q_0(t)$   $\text{cm}^3/\text{seg}$  en el instante  $t$ . A su vez, el agua se alimenta a razón constante de  $q$   $\text{cm}^3/\text{seg}$  (véase la figura 1). Por otro lado, la salmuera se

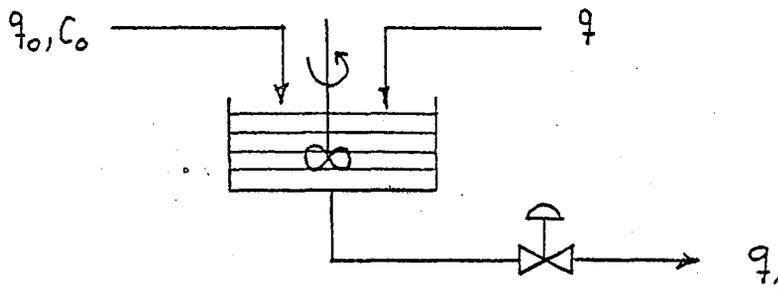


Fig. 1: Tanque agitado para la dilución de una salmuera.

extrae del tanque regulando el flujo de salida de manera que la tasa de extracción en el instante  $t$  sea de  $q_1(t)$   $\text{cm}^3/\text{seg}$ . Ello se logra mediante un controlador de nivel, que acciona la válvula para evitar tanto derrames como vaciados del tanque.

Una sencilla aplicación de (3) al flujo de sal a través del tanque muestra que

$$\frac{d}{dt} (CV) = q_0 C_0 - q_1(t)C;$$

además, despreciando cualquier efecto de cambio de volumen que pudiera suceder durante el mezclado, de la misma (3) se obtiene

$$\frac{dV}{dt} = q_0 + q - q_1,$$

con lo que

$$\frac{dC}{dt} = \frac{q_0 C_0 - (q_0 + q)C}{V}.$$

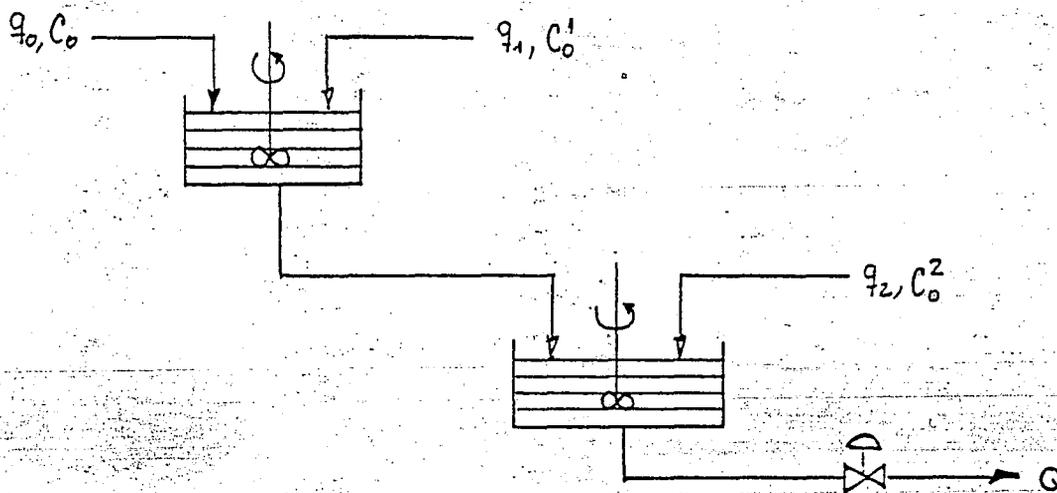
Las trayectorias quedan especificadas mediante este par de ecuaciones diferenciables ordinarias. Basta fijar una condición inicial

$$V(0) = V_0, \quad C(0) = C_0$$

para que quede determinada una trayectoria, siempre que  $V_0 > 0$ . ■

### Ejercicio 1.

Dé un modelo matemático que describa la variación temporal de la concentración de salmuera en cada uno de dos tanques agitados colocados en serie.



Las tres corrientes de alimentación contienen salmuera, de distintas concentraciones. Suponga que la agitación es perfecta en ambos tanques, además de que los volúmenes de solución se mantienen constantes.

### Ejercicio 2.

En el planteamiento del ejemplo anterior se ignoró la posibilidad de un derrame. Esto es razonable gracias a la presencia de un controlador de volumen. Sin embargo, pudiera darse que el controlador no cumpliera bien su cometido. ¿Cómo se modifica el planteamiento en ese caso?

### Ejercicio 3.

Volviendo al tanque agitado de la figura 1 suponga ahora que se alimentan dos corrientes (1 y 2). La corriente 1 contiene una sustancia A en solución, en tanto que la corriente 2 aporta una sustancia pura B. Ahora bien A y B reaccionan según



y la reacción es de primer orden con respecto a cada reactivo A: la tasa de reacción a la temperatura de trabajo es  $kC_A C_B$  g/cm<sup>3</sup>·seg, cuando la concentración de A es  $C_A$  y la de B es  $C_B$ ; k es la constante de velocidad de reacción correspondiente. Se mantienen las condiciones de operación descritas en el ejemplo 1. Dé un modelo matemático de esta situación.

Ya hemos visto que el principio de conservación (3) permite construir modelos matemáticos en diversas áreas, aplicándolo a la masa, la energía o el impulso lineal. El ejemplo siguiente muestra su aplicación a la conservación de la carga y la energía eléctrica. La clave para ello son las dos leyes de Kirchoff, respectivamente, aplicadas a un nodo y un lazo en un circuito eléctrico (figura 2).

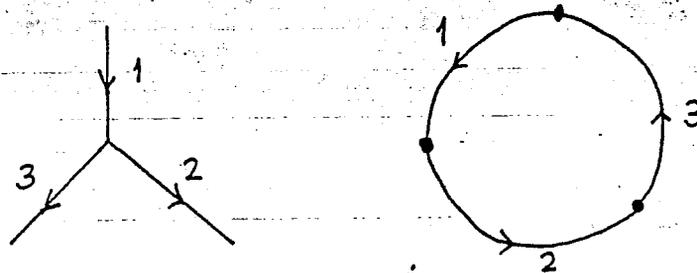


Fig. 2: Los elementos básicos en la topología de un circuito eléctrico.

Primera Ley de Kirchoff.

En un nodo al que concurren varias ramas de un circuito, la carga neta acumulada es nula.

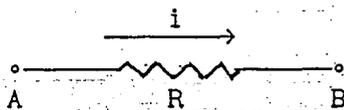
Segunda Ley de Kirchoff.

Al recorrer un lazo cerrado de un circuito, no hay ganancia neta de potencial eléctrico.

Resulta evidente que estas dos leyes no son sino casos particulares de (3): aseguran que no hay acumulación ni de carga en un nodo ni de energía en un lazo. Para aplicarlas, basta saber calcular la caída de potencial a través de cada una de las ramas que constituyen un circuito, relacionándola a la intensidad de corriente que circula por dicha rama.

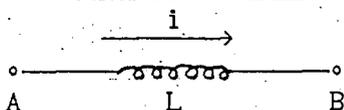
Ahora bien, entre las propiedades más importantes de los diversos elementos de un circuito eléctrico convencional figuran la *resistencia* (R, en ohms), la *inductancia* (L, henrys) y la *capacitancia* (C, en faradays). Disponemos de las siguientes leyes empíricas que describen el comportamiento de estos elementos, donde  $V_A - V_B$  es la caída de voltaje,  $i$  es la intensidad de corriente,  $q$  la carga eléctrica acumulada y  $t$  es el tiempo:

a) *Ley de Ohm*



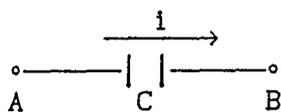
$$V_A - V_B = Ri$$

b) *Ley de Henry*



$$V_A - V_B = -L \frac{di}{dt}$$

c) Ley de Faraday

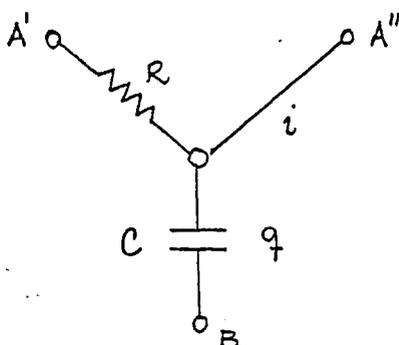


$$q = C(V_A - V_B)$$

$$= \int_0^t i(s) ds$$

Ejemplo 2.

Una conexión eléctrica se puede esquematizar como un nodo 0 al que convergen tres ramas, A'O, A''O y BO.



Las dos terminales A', A'' se mantienen al mismo voltaje  $V_{AB}$  con respecto a B. Se mide además la intensidad de corriente en la rama OA''. Se desea conocer la variación temporal de la carga acumulada en la rama OB''.

Sean  $q, i: \mathbb{R} \rightarrow \mathbb{R}$ , con  $i$  continua y  $q$  continuamente diferenciable, tal que

$i(t)$  = corriente medida

$q(t)$  = carga acumulada en el capacitor,

ambos en el instante  $t$ .

Por la primera Ley de Kirchoff, la corriente fluye a lo largo de A'O con una intensidad igual a  $\dot{q} - i$ . Por las leyes de Ohm y de Faraday, la caída de voltaje en las ramas A'O y BO son respectivamente,  $R(\dot{q}(t) - i)$  y  $q(t)/C$ , en cada instante  $t$ . En virtud de la segunda Ley de Kirchoff, se cumple que

$$R(\dot{q}(t) - i(t)) + \frac{1}{C} q(t) = V_{AB}$$

En otras palabras,  $q$  satisface la ecuación diferencial

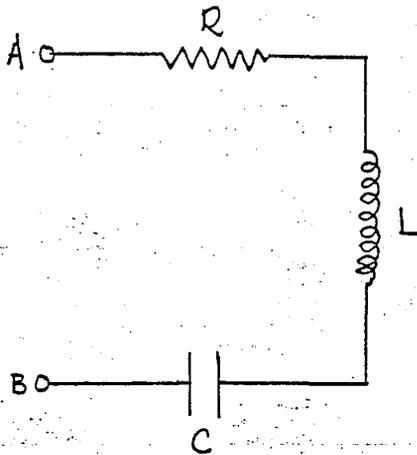
$$\dot{q} + \frac{1}{RC} q = i(t) + \frac{V_{AB}}{R}$$

#### Ejercicio 4.

Haga explícitos los cinco elementos  $T$ ,  $S$ ,  $A$ ,  $E$  y  $\Omega$  que constituyen el modelo matemático determinístico ya construido en el ejemplo anterior.

#### Ejercicio 5.

Una conexión eléctrica se puede esquematizar como una resistencia, una capacitancia y una inductancia pero conectados en serie, según se ilustra en la figura. Se mide la caída global de voltaje entre  $A$  y  $B$ , que resulta variable en el tiempo; lo mismo se puede decir de la intensidad de corriente en dicha rama. Dé un modelo matemático para la variación temporal de la intensidad de corriente medida y de la carga acumulada en el condensador.



#### Ejercicio 6.

En relación con el ejemplo de una partícula que da lugar a (1), escriba las ecuaciones de movimiento para el caso particular  $U(x) = \frac{1}{2}\alpha x^2$  (el oscilador armónico). Adimensionalice las ecuaciones resultantes, y también las ecuaciones de movimiento en el espacio carga-intensidad obtenidas en el ejercicio precedente. ¿Qué analogías percibe entre ambos sistemas? Diga cómo simularía el comportamiento del sistema mecánico mediante el sistema eléctrico.

#### Ejercicio 7.

Dé la región del espacio  $L$ ,  $R$ ,  $C$  en la cual se obtienen oscilaciones eléctricas. ¿Cuándo se mantienen dichas oscilaciones? ¿Qué parámetros deben ajustarse para mantener una frecuencia deseada?

De los ejemplos dados hasta ahora podemos inferir algunos principios generales para la construcción de modelos matemáticos en las ciencias y las ingenierías. Por conveniencia, los haremos explícitos:

I. *Elíjase una propiedad física P que se conserve, es decir, a la cual se pueda aplicar (3).*

II. *Exprésese P en términos de atributos de interés.*

III. *Identifíquense los componentes del sistema en los que se produce o se consume P.*

IV. *Para cada una de dichos componentes, relaciónense las tasas de producción o consumo de P a los valores de los atributos de interés.*

V. *Aplíquese el principio de conservación (3).*

En algunos campos, como la teoría de los Fenómenos de Transporte, se acostumbra llamar *ecuaciones constitutivas* a las que relacionan la tasa de producción o consumo de una propiedad con las variables o atributos de interés. Las Leyes de Henry, Faraday y Ohm son ejemplos de ecuaciones constitutivas; también la ley de Fick en la transferencia de masa y la ley de Fourier en la de calor, así como las expresiones que dan la velocidad de reacción en la cinética química. En la mecánica de partículas, la ley de Hooke y la que dice que la fuerza de fricción es proporcional a la velocidad son otros ejemplos de ecuaciones constitutivas, así como la segunda ley de Newton es un ejemplo de la ley de conservación. En general, las ecuaciones constitutivas se obtienen experimentalmente y lo mismo se puede decir de las leyes de conservación.

## 5.2 Validación de los modelos matemáticos.

La prueba última de validez de un modelo matemático la ofrece la confrontación entre las predicciones que de él se obtienen y las observaciones realizadas sobre el sistema real. Así, por ejemplo, el modelo de Ornstein-Uhlenbeck para la difusión molecular predice que el desplazamiento cuadrático medio de la partícula varía linealmente con el tiempo bajo condiciones de equilibrio; si el comportamiento

observado no corresponde a esto, dicho modelo no es aplicable a la situación observable y debe ser modificado. También, el modelo para la caída libre de una partícula dado en la segunda parte de la primera conferencia predice un tiempo de caída proporcional a la raíz cuadrada de la altura de la cual se deja caer la piedra; si en la práctica no es así, deben revisarse los postulados que subyacen a dicho modelo: o el medio ejerce un efecto apreciable, o la aceleración de la gravedad no es constante o hay otro campo de fuerza externo que no habíamos tomado en cuenta. Para la cuerda vibrante de la conferencia 2 puede ser importante tomar en cuenta el rozamiento, o no considerar que las vibraciones son pequeñas, según se observe. Para la descripción de la difusión desde un punto de vista macroscópico, si las mediciones de los perfiles de concentración no concuerdan con la hipótesis de que se trata sólo de difusión, habrá que incorporar también la reacción química para explicar las discrepancias observadas.

Ahora bien, la obtención de las predicciones cuantitativas de un modelo matemático requiere esencialmente resolver las ecuaciones que lo constituyen. Para obtener la predicción que da el modelo acerca del valor de los atributos en cierto punto del espacio y en cierto instante, debe ser posible calcular los valores numéricos correspondientes, pues son valores numéricos también los resultados de las mediciones experimentales. Si observamos con atención, todos los modelos que hemos utilizado en las conferencias anteriores contienen parámetros: valores de la masa, de la inductancia, resistencia y capacitancia, difusividades, constantes cinéticas, coeficientes de proporcionalidad de varios tipos, etc. Claramente, los valores de dichos parámetros deben conocerse ya al momento de calcular las predicciones del modelo.

Algunos parámetros, como la masa, la aceleración de la gravedad, etc., se pueden medir independientemente. Por lo general, el experimentador diseñará una situación sencilla en la que también aparezcan dichos parámetros, pero en la cual todas las variables restantes se puedan medir fácilmente, y de ahí se obtienen los valores de los parámetros. Luego se utilizan en las ecuaciones que gobiernan el fenómeno de interés, bajo la hipótesis de que efectivamente sean

los mismos parámetros. Esta, desde luego, es una hipótesis adicional, que tiene que ver con los fundamentos mismos de la disciplina de que se trate. Por ejemplo, consideremos la masa. Normalmente se determina pesando el objeto en cuestión, y al resultado le llamamos la *masa gravitacional*; luego se utiliza más bien como *masa inercial*, en las ecuaciones de movimiento de una partícula, como coeficiente de la aceleración. La equivalencia entre masa inercial y masa gravitacional es objeto de consideraciones no triviales en los tratados de Mecánica Clásica.

En principio, podríamos pensar en determinar así todos los parámetros de un modelo, pero en muchos casos resulta difícil pensar en situaciones simples que permitan su determinación directa en forma independiente. En tales casos debe por fuerza recurrirse a técnicas de estimación de parámetros, según veremos a continuación.

Supongamos que un modelo  $M$  contiene  $p$  parámetros desconocidos, digamos  $\theta_1, \dots, \theta_p$  y sea  $\Theta \subset \mathbb{R}^p$  el conjunto donde puede tomar valores el vector  $\theta := (\theta_1, \dots, \theta_p)$ . En realidad,  $M$  no es un modelo, sino un *metamodelo* con tantas instancias como valores tome el parámetro  $\theta$ . Se hablará entonces de una familia

$$M := \left\{ M_\theta, \theta \in \Theta \right\} \quad (1)$$

de modelos para la situación bajo análisis. A la vez, se dispone de observaciones, típicamente bajo la forma de mediciones de los atributos en un número finito de puntos  $p_1, \dots, p_n \in S$  y en un número finito de instantes de tiempo  $t_1, \dots, t_m \in T$ . Sea  $y_{ij}$  el valor medido en el instante  $t_i$ , en el punto  $p_j$ ,  $i=1, \dots, m$ ,  $j=1, \dots, n$ . El modelo  $M_\theta$  predice el valor  $a(t, p | \theta)$  para el atributo en el punto  $p$  medido en el instante  $t$ ; en particular, predice el valor  $a(t_i, p_j | \theta)$  para el valor que medido resulta  $y_{ij}$ . Se puede tomar como medida de la *discrepancia* entre la teoría y la realidad a la función  $\delta: \Theta \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$  dada por

$$\delta(\theta | Y) := \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \| a(t_i, p_j | \theta) - y_{ij} \|^2, \quad (2)$$

donde  $Y$  es la matriz  $m \times n$  de los datos. Esta definición de

discrepancia es a todas luces arbitraria y podría darse de muchas otras formas. Pero para fijar las ideas supondremos que ha sido definida de la manera arriba descrita.

Lo que se desea es una regla de decisión que permita elegir el valor más adecuado  $\theta \in \Theta$  compatible con las observaciones colectadas. Esto puede hacerse de una infinidad de maneras, basta especificar una función  $e: \mathbb{R}^{m \times n} \rightarrow \Theta$ ; una tal función se llamará un *estimador*.

Hay un mínimo de propiedades que convendría asegurar en un estimador para que sea admisible. Por ejemplo, es claro que los datos observados están sujetos al error experimental y sería sumamente inconveniente que pequeños errores en los datos dieran como resultado grandes errores en la determinación del parámetro. Así pues, es importante que los estimadores sean funciones continuas; de un estimador tal se dirá que es *robusto*. Esta y otras propiedades pueden servir para definir la clase  $\mathcal{E}$  de todos los estimadores admisibles.

Dada una medida de discrepancia  $\delta$  y una clase  $\mathcal{E}$  de estimadores admisibles, es natural decir que el mejor estimador en esa clase será aquel que dé la menor discrepancia para cualquier observación. En otras palabras, tiene sentido proponer el siguiente

#### Problema de Estimación.

Encontrar  $e^* \in \mathcal{E}$  tal que

$$\delta(e^*(Y)|Y) = \min_{\theta \in \Theta} \delta(\theta|Y) \quad (3)$$

cualquiera que sea  $Y \in \mathbb{R}^{m \times n}$ . Si existe,  $e^*$  es llamado un *estimador óptimo*.

Claramente, el carácter óptimo de un estimador es relativo a la medida de discrepancia  $\delta$  y a la clase  $\mathcal{E}$  que se considere admisible. Si  $\delta$  se define como en (2), el estimador se llama de *mínimos cuadrados*. Se observa que, para obtenerlo, se debe minimizar una función escalar de  $\theta \in \Theta$  para cada posible valor  $Y$  de las observaciones. En general,

dicha función escalar será una suma de cuadrados de otras funciones, es decir, de la forma

$$\sum_{k=1}^n f_k(\theta)^2, \quad (4)$$

y hay métodos eficientes que toman ventaja de esto, para hacer los cálculos más eficientes; por ejemplo el método iterativo llamado de Gauss-Newton, cuando  $\theta = \mathbb{R}^p$ . Este método, y el Maquardt (que también es muy recomendable para problemas de estimación por mínimos cuadrados) son variantes del método de Newton-Raphson, modificado para extraer ventaja de la forma especial de (4).

Se observa que para calcular los valores  $a(t_i, p_j)$  hay que resolver las ecuaciones de movimiento, por lo que el cálculo de las  $f_k$ 's en (4) puede resultar complejo. Métodos como el de Gauss-Newton requieren además el cálculo del gradiente de las  $f_k$ 's, lo cual añade al costo computacional. Hay otros, como Nelder-Mead o método del Simplejo, que se basan exclusivamente en los valores de la función a minimizar, por lo cual pueden resultar preferibles. Sin embargo, no es fácil dar indicaciones generales acerca del método más conveniente para realizar la minimización (3). En general, la forma de las ecuaciones del modelo y la elección de medida de discrepancia tendrán una influencia decisiva en la elección de método de optimización a utilizar.

Por otro lado, obsérvese que en la discusión anterior no se ha hecho ninguna referencia a formulaciones de tipo estadístico o probabilístico. En particular, subyace a todo lo anterior la idea de que se deben "resolver las ecuaciones del modelo", lo cual supone que se trata de modelos determinísticos. Para los modelos probabilísticos se pueden formular conceptos análogos. En efecto, para ello es frecuente que ni el conjunto  $\Omega$  de los posibles resultados ni la familia  $A$  de eventos a probabilizar dependan de ningún parámetro: es al especificar la probabilidad que se introducen éstos. Por ejemplo, considérese el modelo de la desintegración radiactiva para tiempos cortos. Aquí, los resultados posibles son todas las funciones escalonadas con valores no negativos, continuas por la derecha; el

conjunto de todas ellas es  $\Omega$ . En cuanto a la familia  $\mathcal{A}$ , son algunos de los subconjuntos de  $\Omega$ , y se puede decir cuales sin referencia a ningún parámetro. Sin embargo, las probabilidades  $p_n(t)$  dependen de un parámetro  $\lambda > 0$ . Así pues, en realidad se tiene una familia  $\{(\Omega, \mathcal{A}, P_\lambda), \lambda > 0\}$  de espacios probabilísticos como metamodelo aplicable a esta situación. Habrá que estimar el mejor valor del parámetro mediante algún estimador adecuado.

En general, cuando no hay determinismo se tendrá una familia (1) de modelos matemáticos, con

$$M_\theta := (\Omega, \mathcal{A}, P_\theta). \quad (5)$$

Para estimar el parámetro  $\theta$ , se podrá utilizar una técnica tipo mínimos cuadrados también, si se desea. O se podrá seguir un enfoque más probabilístico, como es el de máxima verosimilitud. En este caso, las observaciones se consideran muestra de una matriz aleatoria  $Y$ , cuya distribución  $\mu_\theta$  bajo  $P_\theta$  se conoce; su medida de verosimilitud no es otra cosa que  $\mu_\theta$ . En particular, si  $\mu_\theta$  tiene una densidad  $f(\cdot|\theta)$  y si  $Y_0$  es la muestra observada, entonces se dice que  $f(Y_0|\theta)$  es la verosimilitud de la muestra. Un *estimador de máxima verosimilitud* es uno que maximiza esta función, cualesquiera que sean las observaciones:

$$f(Y_0|e^*(Y_0)) = \max_{\theta \in \Theta} f(Y_0|\theta). \quad (6)$$

Se ve que (6) es un caso particular de (3), que se obtiene al tomar

$$\delta(\theta|Y_0) = -f(Y_0|\theta).$$

Para modelos probabilísticos como (5) tenemos además la posibilidad de discriminar entre dos modelos dados ( $\theta = \theta_1$  vs.  $\theta = \theta_2$ ) o, más en general, entre dos conjuntos de modelos ( $\theta \in \Theta_0$  vs.  $\theta \in \Theta - \Theta_0$ ), mediante la técnica inferencial conocida como *prueba de hipótesis estadísticas*. Mediante una prueba de este tipo, podemos rechazar la hipótesis de que  $\theta \in \Theta_0$  a la luz de la evidencia empírica (o no rechazarla, si es lo justo), cuidando mantener la probabilidad de rechazar una hipótesis verdadera por debajo de una cierta tolerancia  $\alpha$ .

## APENDICE: LA GRAVITACION UNIVERSAL

En este apéndice se presenta la dinámica newtoniana, incluyendo su génesis en las leyes de Kepler, así como en los trabajos experimentales y filosóficos que las precedieron. Se incluye este material porque la teoría newtoniana constituye el modelo matemático determinístico por excelencia, al cual se refieren todos los demás. La presentación que aquí se da expone los orígenes empíricos de la teoría, a la vez que presenta algunas de sus consecuencias. Por lo tanto constituye material ilustrativo de la metodología expuesta en las cinco conferencias que lo preceden.

Se recomienda a los lectores la parte histórica de la tesis [Flores Peñafiel, 1976], así como [Koestler, 1982], para obtener una visión del lado humano de estas ideas. Además, la parte técnica de la tesis de Flores Peñafiel contiene desarrollos teóricos importantes de la teoría newtoniana de la gravitación.

### 1. La síntesis newtoniana.

Dos fenómenos han intrigado al hombre desde los albores de la historia:

- i) la caída de los cuerpos en la superficie de la Tierra
- ii) el movimiento de los cuerpos celestes (el Sol, la Luna, los cometas, etc.).

El hombre moderno conoce las leyes que gobiernan dichos fenómenos. Más importante aún, sabe que esencialmente se trata del mismo fenómeno, conocido como *gravitación*. Este conocimiento, que ahora damos por descontado, representa una síntesis formidable; es resultado de un trabajo de miles de años acerca del Universo, que culmina en el Siglo XVII con la formulación por parte de Isaac Newton (Inglaterra, 1642-1727) de dos leyes que llevan su nombre y que son la base de a) la Dinámica, b) la Gravitación, en su versión clásica. Ellas son

• Dinámica (2ª Ley de Newton)

Bajo la acción de una fuerza, una partícula experimenta una aceleración igual a un múltiplo escalar positivo de dicha fuerza.

El inverso de dicha constante de proporcionalidad se conoce como masa de la partícula. Así pues, si se tiene una partícula de masa  $m$  sobre la que actúa una fuerza externa  $F$ , la partícula experimenta una aceleración  $a$  dada por

$$a = \frac{1}{m} F$$

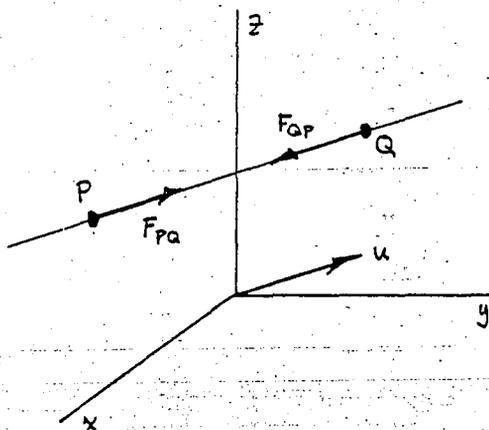
es decir,

$$F = ma. \quad (1)$$

Nótese que en esta ley figura una abstracción muy importante, que es el concepto de *partícula*: ente sin dimensiones físicas (es decir, un punto) al que se asocia una masa.

• Gravitación Universal (Ley de la ...)

— Dos partículas cualesquiera  $P$ ,  $Q$  se atraen con una fuerza proporcional al producto de sus masas e inversamente proporcional al cuadrado de la distancia entre ellas. Dicha fuerza de atracción actúa a lo largo de  $\overline{PQ}$  y es mutua.



Si  $m(P)$  y  $m(Q)$  denotan las masas respectivas,  $P$  es atraída por  $Q$  con una fuerza igual a

$$F_{PQ} = G \frac{m(P) \cdot m(Q)}{|P - Q|^2} u, \quad (2a)$$

donde  $u \in \mathbb{R}^3$  es un vector unitario paralelo a  $\overline{PQ}$  y que apunta en dirección de  $P$  a  $Q$ . A su vez,  $Q$  es atraído por  $P$ , y

$$F_{QP} = - F_{PQ} .$$

(2b)

La síntesis a que hacemos mención líneas arriba resulta de un gran número de contribuciones previas al gran salto conceptual dado por Newton. Brevemente, podemos mencionar las siguientes:

1. La formulación, por parte de Nicolás Copérnico (Torún, en la hoy Polonia, 1473-1543) de la teoría heliocéntrica para el sistema solar, en rompimiento con la visión geocéntrica del Universo, entonces en voga.

2. La idea renacentista, formulada por Galileo Galilei (Florencia, en la hoy Italia, 1564-1642) y otros, de recurrir al experimento para descubrir la verdad, y no sólo basarse en la especulación acerca del mundo.

3. La realización de mediciones precisas del curso de los planetas por parte de Tyko Brahe (Dinamarca, 1546-1601), que acumuló un cuantioso acervo de datos astronómicos.

4. El análisis que hizo Johannes Kepler (Bohemia, en la hoy Checoslovaquia, 1570- 1630) de los datos de Brahe, de quien había sido asistente.

Dicho análisis condujo a la formulación de los siguientes enunciados, en los que se sistematiza la gran masa de conocimiento empírico legado por Brahe, y que se conocen como *Leyes de Kepler*:

1ª Ley

*Los planetas se mueven alrededor del Sol en órbitas elípticas, con el Sol en uno de los focos.*

2ª Ley

*El movimiento de un planeta es tal que el radio vector que va del Sol al planeta barre áreas iguales en tiempos iguales.*

### 3ª Ley

Para dos planetas cualesquiera del sistema solar, los cuadrados de los periodos de recorrido de sus órbitas se encuentran en la misma proporción que los cubos de los semiejes mayores respectivos.

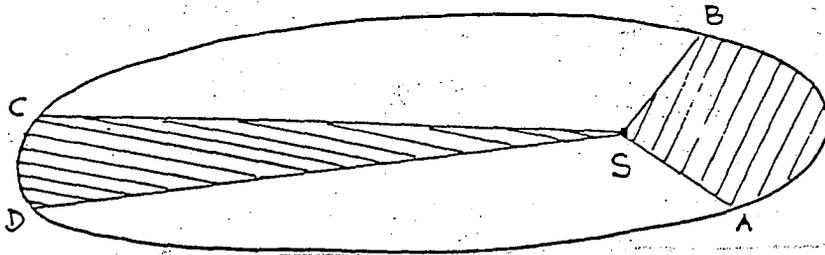


Figura 1: Órbita elíptica de un planeta que se mueve alrededor del sol.

Finalmente la coronación de los cuatro desarrollos anteriores vino con

5. La formulación de las dos leyes de Newton antes citadas, en los años 1665 a 1667, durante un periodo de vacaciones forzadas de la Universidad de Cambridge, cerrada a causa de un brote de peste.

En términos de nuestro hemisferio, este arco de tiempo cubre los siguientes eventos:

- a) El apogeo de los mexicas, pueblo guerrero y última de las culturas del altiplano mexicano (S.XIII a S.XV).
- b) El descubrimiento de América, como inicio de la expansión europea hacia estas tierras (1492).
- c) La conquista española de Tenochtitlán (1521).
- d) El establecimiento de la primera imprenta de América, en la Ciudad de México (1550).
- e) La fundación de la Real y Pontificia Universidad de México (1570)

- f) Florecimiento de la Nueva España, en el Siglo XVII, que produjo figuras como Sor Juana Inés de la Cruz (1648-1695) y, en general
- g) La colonización europea de toda la América (española, portuguesa, francesa y anglosajona).

La sociedad novohispana era más dada al cultivo de las Letras que al de las Ciencias, además de que la Iglesia Católica tenía una gran influencia, y por mucho tiempo desaprobó esa desacralización de los fenómenos celestes que resultaba de la síntesis newtoniana (recuérdese el juicio a Galileo). Sin embargo, sí hubo algunos intelectuales novohispanos que, a través de sus lecturas o de su correspondencia, pudieron acceder a la nueva visión del mundo propuesta por Newton. Al respecto, véanse [De Gortari, 1979], [Trabulsee, 1974, 1982].

Desde un punto de vista epistemológico, la teoría heliocéntrica abrió el camino para la síntesis newtoniana, de los fenómenos terrestres con los celestes. Además, el método experimental llevó a la recolección de la información pertinente. Pero el trabajo de Kepler fue doblemente importante, pues

- a) proporcionó un sólido sustento a las ideas copernicanas, que así se vió permiten explicar los fenómenos celestes de manera más simple, y
- b) permitió a Newton dar la dependencia correcta de la fuerza de atracción con la separación entre las partículas.

Finalmente, la idea newtoniana (por demás irreverente a los ojos de la Iglesia Católica) de aplicar las mismas leyes a los fenómenos terrestres y celestes llevó a demostrar la validez de las leyes de Kepler a partir de los principios de la dinámica newtoniana. Dicha demostración se hizo usando métodos matemáticos desarrollados por Newton mismo: el Cálculo Diferencial e Integral. Veamos a continuación como se llega a tales resultados.

Primeramente, conviene observar que ni los cuerpos terrestres ni los celestes son partículas puntuales; de hecho, la Mecánica Celeste se ocupa del movimiento de cuerpos de gran tamaño. Sin embargo, dichos cuerpos son esféricos, al menos en una primera aproximación, y el Cálculo Integral permite demostrar la siguiente consecuencia de (2a):

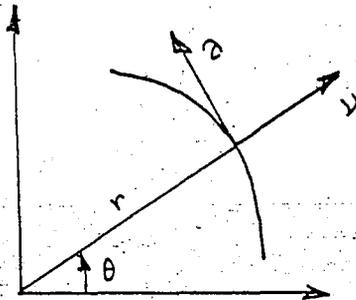
Lema 1.

Sea una distribución esférica de materia, de masa total  $M$  y centro en  $O$ . Supóngase que su densidad, no necesariamente uniforme, es esféricamente simétrica. Entonces, la fuerza que esta distribución de materia ejerce sobre una partícula externa  $P$  coincide con la que ejercería sobre  $P$  una partícula de masa  $M$  colocada en  $O$ .

Véase la sección 5.2 de [Collinson, 1980] (también la sección 16.6 de [Resnick-Halliday, 1980]) para una demostración de este hecho.

Análogamente, el Cálculo Diferencial permite demostrar el siguiente resultado, que establece las reglas que gobiernan la cinemática plana en coordenadas polares. Para enunciarlo, sea una curva  $C$  en el plano, descrita mediante el par de ecuaciones paramétricas

$$r = r(t), \quad \theta = \theta(t), \quad t \in \mathbb{R},$$



con  $r, \theta \in C^2$ . Sea  $P \in C$  y sean  $\tau = \tau(P), \nu = \nu(P) \in \mathbb{R}^2$  vectores unitarios con  $\tau$  tangente a  $C$  y  $\nu$  normal a  $C$  en el punto  $P$ . Desde luego, las funciones  $t \mapsto r(r(t), \theta(t)), t \mapsto \nu(r(t), \theta(t))$  son continuamente diferenciables.

Lema 2.

Sea  $C$  una curva en el plano, dos veces diferenciable, y sea una partícula que se desplaza a lo largo de  $C$ . Entonces, su velocidad y aceleración son

$$v = r\dot{\theta}\tau + \dot{r}\nu \quad (3a)$$

$$a = (r\ddot{\theta} + 2\dot{r}\dot{\theta})\tau + (\ddot{r} - r\dot{\theta}^2)\nu, \quad (3b)$$

respectivamente.

De acuerdo con (1), si la partícula en cuestión tiene masa  $m$ , sobre ella actúa en cada momento una fuerza  $F(t)$ , dada por

$$F = m(r\ddot{\theta} + 2\dot{r}\dot{\theta})\tau + m(\ddot{r} - r\dot{\theta}^2)\nu. \quad (4)$$

Se refiere a los lectores a la sección 2.2 de [Collinson, op.cit] (o al Tópico Suplementario I de [Resnick-Halliday, op.cit]) para la demostración del Lema 2.

## 2. Génesis de la ley de la gravitación universal.

Por el Lema 1.1, sabremos que en la Dinámica Newtoniana se puede considerar que los cuerpos celestes son partículas puntuales. Veamos a continuación que el movimiento de los planetas se da en un plano que contiene al Sol, como consecuencia del carácter central del campo gravitacional.

### Definición.

Un campo de fuerza es central, con centro en  $O$ , si la fuerza en cada punto  $P \in \mathbb{R}^3$  depende sólo de la longitud de  $OP$ .

Así pues, un campo de fuerza central está dado por  $f(|r|)\frac{r}{|r|}$  y, en el plano, del Lema 1.2 se sigue que el sistema de ecuaciones diferenciales

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0 \quad (1a)$$

$$\ddot{r} - r\dot{\theta}^2 = f(r) \quad (1b)$$

describe la trayectoria de una partícula que se mueve en dicho campo central.

### Proposición 1.

El movimiento de una partícula en un campo de fuerza central se da en un plano que contiene al centro.

Demostración.

Las ecuaciones de movimiento son

$$\dot{r} = v, \quad \dot{v} = f(r) \cdot \frac{r}{|r|},$$

por lo que, a lo largo de un movimiento cualquiera

$$\frac{d}{dt}[r(t) \times v(t)] = v(t) \times v(t) + \frac{f(r)}{|r|} r(t) \times r(t) = 0$$

y, si  $L := r(0) \times v(0)$ , entonces

$$r(t) \times v(t) = L \quad \forall t \in \mathbb{R}$$

(relación llamada la *Ley de Conservación del impulso angular*). Por consiguiente

$$L \cdot r(t) = r(t) \times v(t) \cdot r(t) = 0$$

y el movimiento se da en el plano perpendicular a  $L$  que pasa por el centro  $O$ . ■

Como consecuencia del resultado precedente, el estudio del movimiento de una partícula en un campo de fuerza central se reduce a la investigación de las soluciones del sistema de ecuaciones diferenciales (1).

Teorema 1.

Para una partícula  $P$  que se mueve en un campo de fuerza central, con centro en  $O$ , el radio vector  $OP$  describe áreas iguales en tiempos iguales.

Demostración.

El área descrita por el segmento que va de  $O$  a  $(r(s), \theta(s))$ , para  $0 \leq s \leq t$ , está dada por

$$A(t) = \frac{1}{2} \int_0^t r^2(s) \dot{\theta}(s) ds$$

Ahora bien, la trayectoria  $t \mapsto (r(t), \theta(t))$  satisface (1a), por lo cual

$$\frac{d}{dt} [r(t)^2 \dot{\theta}(t)] = r(t)[r(t)\ddot{\theta}(t) + 2\dot{r}(t)\dot{\theta}(t)] = 0$$

y entonces

$$r^2(t)\dot{\theta}(t) = h \quad \forall t \in \mathbb{R} \quad (2)$$

si  $h := r^2(0)\dot{\theta}(0)$ . Así pues,

$$A(t) = \frac{1}{2} h t, \quad (3)$$

que implica la afirmación a probar. ■

Vemos pues que la 2ª Ley de Kepler es consecuencia del carácter central del campo de fuerza del Sol, y no requiere el tipo de dependencia que prescribe la ley de Newton (1,2). De hecho vale también para el movimiento rectilíneo uniforme.

#### Teorema 2.

Para una partícula P que se mueve en un campo de fuerza O, su trayectoria es una cónica con un foco en O si y sólo si la fuerza es de la forma

$$f(r) = -\frac{k}{r^2}, \quad (4)$$

para alguna constante  $k > 0$ .

#### Demostración.

En efecto, en coordenadas polares con centro en O, la ecuación de una cónica con un foco en O es

$$r = \frac{pe}{1 + e \cos \theta}, \quad (5)$$

donde  $e \geq 0$  es la excentricidad de la cónica y  $p > 0$  es un factor de escala. Si  $e < 1$  se trata de una elipse (una circunferencia si  $e=0$ ), de una hipérbola si  $e > 1$  y de una parábola si  $e=1$ .

Sea  $v := \frac{1}{r}$ ; entonces en coordenadas  $(v, \theta)$  la cónica está descrita por la ecuación

$$v = \frac{1}{p} \left[ \frac{1}{e} + \cos \theta \right].$$

Un cálculo simple muestra que v satisface la ecuación diferencial

$$\frac{d^2 v}{d\theta^2} + v = \frac{1}{pe},$$

que llamamos la ecuación diferencial de una cónica.

Por otro lado, la trayectoria  $t \mapsto (r(t), \theta(t))$  satisface (1b), que

bajo el cambio de variable  $v = 1/r$ , se transforma en

$$\frac{d^2v}{d\theta^2} + v = - \frac{f(1/v)}{h^2 v^2};$$

esta última ecuación diferencial coincide con la de la cónica sólo si

$$f(1/v) = - \frac{h^2 v^2}{pe},$$

de donde resulta (4), con

$$k = \frac{h^2}{pe}. \quad (6)$$

Para la afirmación recíproca basta invertir el orden de los pasos anteriores. ■

Queda claro del resultado anterior que la 1ª Ley de Kepler ya expresa la dependencia de la fuerza de atracción con el inverso del cuadrado de la separación. Veamos qué se puede decir acerca del carácter de la cónica recorrida, es decir, de qué depende que sea efectivamente una elipse o no. Para ello nos conviene estudiar la función *energía total*, dada por

$$H(r, \theta, v, \omega) = \frac{1}{2} m (v^2 + r^2 \omega^2) - \frac{k}{r}$$

tomando en cuenta (1.3a) y (4).

#### Proposición 2.

Para una partícula  $P$  que se mueve en un campo de fuerza central (4),  $H$  es constante a lo largo de cualquier movimiento, digamos con valor  $E_0$ . Además, la trayectoria es elíptica si  $E_0 < 0$ , parabólica si  $E_0 = 0$  e hiperbólica si  $E_0 > 0$ .

#### Demostración.

A lo largo de un movimiento cualquiera  $t \mapsto (r(t), \theta(t))$  la energía está dada por

$$E(t) = H(r(t), \theta(t), \dot{r}(t), \dot{\theta}(t))$$

y un cálculo fácil muestra que  $E$  es constante. Sea  $E_0 := E(0)$ .

De (2), resulta que

$$E_0 = \frac{1}{2} \left[ \dot{r}(t)^2 + \frac{h}{r(t)^2} \right] - \frac{k}{r(t)}, \quad t \in \mathbb{R}.$$

Por otro lado, (5) supone que el rayo  $\theta=0$  corta a la cónica en un punto donde  $r$  toma un valor extremo. Sea  $t_0 \in \mathbb{R}$  tal que  $\theta(t_0)=0$ , por lo que  $\dot{r}(t_0)=0$  y entonces

$$E_0 = \frac{m h^2}{2r_0^2} - \frac{k}{r_0},$$

donde  $r_0 := r(t_0)$ . Por otro lado, haciendo  $\theta=0$  en (5) obtenemos

$$r_0 = \frac{p e}{1 + e}.$$

Basta eliminar  $r_0$  entre estas últimas dos relaciones para obtener que

$$e = \sqrt{1 + E_0 \left[ \frac{2 h^2}{k^2} \right]},$$

de donde resulta la afirmación acerca del carácter de la trayectoria de P. ■

En otras palabras, tanto los cometas recurrentes como los planetas del sistema solar son cuerpos de baja energía, y a ellos se aplica la 1ª Ley de Kepler. Hay otros cuerpos celestes cuya energía es suficientemente alta para que no se vean sino una sola vez, pues sus trayectorias son curvas abiertas (parábolas o ramas de hipérbola).

Concentrémonos en el contenido del movimiento de los cuerpos celestes de baja energía, para estudiar la dependencia del periodo con el que recorren sus órbitas elípticas respecto a la magnitud de las mismas.

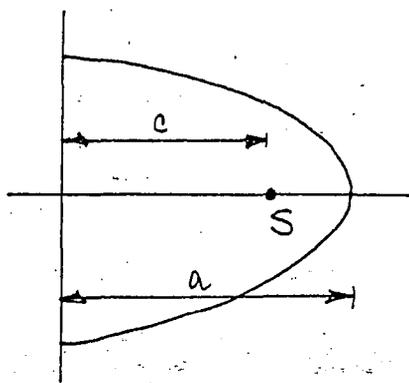
### Teorema 3.

Para una partícula P que se mueve en un campo de fuerza central (4), con energía  $E_0 < 0$ , sea T el periodo con que recorre su órbita y sea a el semieje mayor de la misma. Entonces

$$T^2 = \frac{4\pi^2}{k} a^3. \quad (7)$$

### Demostración.

Sea b la longitud del semieje menor de la órbita, de tal manera que



$$e = \frac{c}{a}, \quad c = \sqrt{a^2 - b^2}.$$

Entonces, de (3) tenemos que

$$\pi ab = \gamma \cdot \frac{hT}{2} \quad (8)$$

En  $\theta = 0$

$$\frac{p e}{1 + e} = a - c = a(1 - e) = \frac{h^2}{k(1 + e)},$$

en virtud de (6). Por lo tanto

$$\frac{h^2}{k} = a(1 - e^2) = \frac{b^2}{a},$$

es decir

$$b^2 = \frac{h^2}{k} a. \quad (9)$$

Basta eliminar  $b$  entre (8) y (9) para obtener (7). ■

Sean ahora dos planetas del sistema solar, digamos  $P_1$  y  $P_2$ , que recorren órbitas elípticas confocales en  $S$ , cuyos semiejes mayores son  $a_1$  y  $a_2$ , en periodos  $T_1$  y  $T_2$ , respectivamente. Del mismo modo, sean  $k_1$  y  $k_2$  las constantes de la fuerza (4) aplicables a los pares  $(S, P_1)$  y  $(S, P_2)$ , respectivamente. De (7) se tiene entonces que

$$\frac{T_1^2}{T_2^2} = \left[ \frac{k_2}{k_1} \right] \frac{a_1^3}{a_2^3},$$

que no necesariamente coincide con la 3ª Ley de Kepler. Si coincide si la constante de proporcionalidad  $k$  en (4) no depende sino de la partícula en  $S$ , es decir, del Sol, pues entonces  $k_1 = k_2$  y se habrá recuperado la 3ª Ley de Kepler.

Sea  $M$  la masa del Sol y definamos

$$G := \frac{k}{M},$$

de tal manera que (4) se transforma en

$$f(r) = - \frac{GM}{r^2}$$

y la fuerza con que el Sol atrae a un planeta de masa  $m$  es de magnitud

$$G \frac{Mm}{r^2}, \quad (10)$$

cuando están separados por una distribución  $r > 0$ . Aquí,  $G$  es una constante que depende en principio de las características del centro atractor, es decir, del Sol.

Newton fué más allá e introdujo el siguiente postulado:

Axioma.

*G es universal.*

Esto significa que la fuerza de atracción gravitatoria mutua entre dos masas puntuales cualesquiera  $M$  y  $m$  separadas por una distancia  $r$  está dada por (10), donde  $G$  es una constante con dimensiones adecuadas pero *independiente del par de masas de que se trate*. Basados en este axioma,  $G$  se puede determinar en el laboratorio, trabajando con masas y distancias conocidas: basta medir la fuerza de atracción correspondiente mediante una balanza o un dinamómetro. Dicha determinación se ha hecho en múltiples ocasiones y se ha obtenido siempre el mismo valor de  $G$ , dentro de los límites de error experimental. Su valor actualmente aceptado es

$$G = 6.672 \times 10^{-8} \frac{\text{dina} \cdot \text{cm}^2}{\text{g}^2}$$

Por otro lado, se ha utilizado este valor de  $G$  para cálculo de órbitas de satélites artificiales, de cuerpos celestes de otros sistemas planetarios, de otras galaxias, etc., obteniendo una buena concordancia con las mediciones. De hecho ha habido discrepancias entre la teoría newtoniana y el experimento, detectadas a principios de este siglo, lo cual indica que aquella tiene sus límites de validez. Dichas discrepancias han podido ser explicadas mediante una teoría más general, la de la *relatividad* de Einstein. Véase [Misner et

al., 1970]. Sin embargo, por casi tres siglos la síntesis newtoniana permaneció incontestada, y aún ahora sigue siendo utilizada cuando las velocidades son pequeñas comparadas con la velocidad de la luz y la densidad de la materia es relativamente baja; si eso no se cumple, las correcciones relativistas se vuelven importantes.

## BIBLIOGRAFIA

Adby, P.R. & M.A.H. Dempster, *Introduction to Optimization Methods*, Chapman & Hall, London, 1974.

Andrews, J.G. & R.R. McClone, eds., *Mathematical modelling*, Butterworths, London, 1980.

Aris, R., *The mathematical theory of diffusion and reaction in permeable catalysts* (2 vols.), Clarendon Press, Oxford, 1975.

Aris, R., *Mathematical modelling techniques*, Research Notes in Mathematics No. 24, Pitman, San Francisco, 1978.

Beltrami, E., *Mathematics for dynamic modelling*, Academic Press, Boston, 1987.

Bird, R.B., W.E. Stewart & E.N. Lightfoot, *Transport Phenomena*, John Wiley, New York, 1961.

Boyce, W.E. (ed.), *Case studies in mathematical modeling*, Pitman, Boston, 1981.

Boyce, W.E. & R. DiPrima, *Elementary differential equations and boundary value problems*, John Wiley & Sons, New York, 1977.

Braun, M., *Differential equations and their applications*, Springer Verlag, New York, 1984.

Burghes, D.N. & M.S. Borrie, *Modelling with differential equations*, Ellis Horwood, Chichester, 1981.

Callen, H.B., *Thermodynamics*, John Wiley & Sons, New York 1960.

Cárdenas, E., D.B. Hernández, C. Rincón & C. Velarde, *Metodos matematicos de la Termodinamica*, Editorial Universitaria (UNAM), Cd. de México, 1978.

Collinson, C.D., *Introductory Mechanics*, Edward Arnold, London, 1980.

De Gortari, E., *La ciencia en la historia de México*, Editorial Grijalbo, Cd. de México, 1979.

Doob, J.L., *Stochastic Processes*, J. Wiley & Sons, New York, 1953.

Einstein, A., *Ann. der Physik*, 17(1905), 549; English translation reprinted in R. Fürth, ed. *Investigations on the theory of the brownian motion*, Dover Publications Inc., New York, 1956.

Feller, W., *An Introduction to Probability Theory and its Applications*, John Wiley, New York, 1968.

Ferziger, J.H., *Numerical methods for engineering applications*, John Wiley & Sons, New York, 1981.

Fife, P., *Mathematical aspects of reacting and diffusing systems*, Springer Lecture Notes in Biomathematics 28, Springer Verlag, New York, 1979.

Flores Peñafiel, A., *El problema restringido de los tres cuerpos*, Tesis Profesional, Facultad de Ciencias, UNAM, Cd. de México, 1976.

Haberman, R., *Mathematical models: mechanical vibrations, population dynamics and traffic flow*, Prentice Hall Inc., Englewood Cliffs NJ, 1977.

Harte, J., *Consider a spherical cow: a course in environmental problem solving*, William Kaufmann Inc., Los Altos CA, 1985.

Hernández, D.B., *Análisis Numérico*, Texto del III Coloquio del Departamento de Matemáticas del CINVESTAV (La Trinidad, Tlax., Agosto de 1983). Cd. de México, 1983.

Hernández, D.B., "De los racionales al ruido blanco", *Memorias del XXI Congreso de la Soc. Mat. Mex. (Hermosillo, Son., Noviembre, 1988)*, Serie Comunicaciones, Aportaciones Matemáticas, Soc. Mat. Mex., Cd. de México, 1988.

Hernández, D.B. & L.J. Alvarez, *El método de Monte Carlo*. Informe Monográfico, Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa. Cd. de México, 1984.

Hurley, D., *Probabilidad y Estadística III: Inferencia Estadística*, Ed. Limusa, Cd. de México, 1984.

Koestler, A., *The sleepwalkers: a history of Man's changing views of the Universe*, Penguin Books, London, 1982.

Lin, C.C. & L.A. Segel, *Mathematics applied to deterministic problems in the natural sciences*, The Macmillan Publishing Co., New York, 1974 (reimpreso por SIAM, Philadelphia, 1988).

Marsden, J.E. & A.J. Tromba, *Cálculo Vectorial*, Addison Wesley Iberoamericana, Wilmington, Del., 1987.

Méndez, I., *Probabilidad y Estadística IV: Modelos Lineales*, Ed. Limusa, Cd. de México, 1984.

Misner, C.W., K.S. Thorne & J.A. Wheeler, *Gravitation*, W.H. Freeman & Co., San Francisco, 1970.

Nelson, E., *Dynamical theories of brownian motion*, Princeton University Press, Princeton NJ, 1967.

Noble, B., *Applications of undergraduate mathematics in engineering*, The Macmillan Co., New York 1967.

Press, W.H., B.P. Flannery, S.A. Teutolsky & W.T. Vetterling, *Numerical recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1988.

Resnick, R. & D. Halliday, *Física (Partes I y II)*, Cia. editora Continental S.A. de C.V., Cd. de México, 1980.

Rudin, W., *Principles of Mathematical Analysis*, McGraw Hill, New York, 1976.

Schuss, Z., *Theory and applications of stochastic differential equations*, John Wiley & Sons, New York, 1980.

Simmons, G.F., *Differential equations, with applications and historical notes*, Tata-McGraw Hill Publ. Co., New Delhi, 1972.

Smoller, J., *Shock waves and Reaction-Diffusion Equations*, Springer Verlag, New York, 1983.

Smoluchowski, M. "Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen", *Phys. Zeitschr.* 17(1916), 557-571, 585-599; English translation in R.S. Ingarden, ed. *Marian Smoluchowski: His Life and Scientific Work*, Polish Scientific Publishers, Warsaw, 1986.

Strang, G., *Introduction to Applied Mathematics*, Wellesley - Cambridge Press, Wellesley MA, 1986.

Trabulse, E., *Ciencia y religión en el Siglo XVII*, El Colegio de México, Cd. de México, 1974.

Trabulse, E., *El círculo roto*, Ed. SEP Ochenta, Cd. de México, 1982.

Wax, N. (ed.), *Selected papers on noise and stochastic processes*, Dover Publications Inc., New York, 1954.

Yakowitz, S.J., *Computational Probability and Simulation*, Addison Wesley, Reading MA, 1977.