

## Capítulo 5

# REPRESENTACIONES GRÁFICAS

**Autores:** José María García Palanco  
Marta Sánchez-Cabezudo Tirado

## 5 REPRESENTACIONES GRÁFICAS

Cualquier experimento tiene por finalidad comprobar la validez de un modelo teórico, contrastando los valores experimentales con los predichos por el modelo, o bien, estudiar un fenómeno y, de la información obtenida experimentalmente, elaborar un modelo que describa ese fenómeno.

Los datos obtenidos a partir de las medidas en un laboratorio deben presentarse de manera que los demás obtengan la mayor cantidad y calidad de información posible. Para lograr esto recurrimos a las *tablas* y a las *representaciones gráficas*. Las tablas nos permiten ver el conjunto de los datos obtenidos sin tener que irlos persiguiendo a lo largo del informe. Con las gráficas no sólo conseguimos una información cuantitativa de la magnitud medida sino también su relación con los parámetros del experimento.

### 5.1 Tablas

Cuando sea posible se registrarán las medidas en forma de tablas, puesto que se trata de la forma más compacta y sencilla de presentar los resultados, de acuerdo con las siguientes recomendaciones:

- En las tablas se representarán tanto los datos directos de las medidas del laboratorio como los pasos intermedios relevantes y los resultados buscados.
- Las medidas de una misma magnitud se escribirán preferiblemente sobre una misma columna vertical, ya que es más sencillo comparar a simple vista un conjunto de datos verticales.
- En la cabecera de cada columna se indicará el nombre de la magnitud y/o el símbolo seguido por las unidades. Al indicar en la cabecera las unidades ya no es necesario repetirlas después de cada medida; con esto se ahorra tiempo, energía y se hace más claro el informe.
- Es conveniente elegir las unidades (o las potencias de 10 adecuadas) para que los números queden expresados en el rango entre 0.1 y 1000. En el caso de que, como consecuencia de lo indicado, todos los valores de una columna de datos incluyan un mismo factor potencia de 10, se podrá prescindir del mismo, indicando en la cabecera de la columna la magnitud correspondiente multiplicada por el factor inverso del que se ha prescindido en la columna de valores experimentales.
- Los errores en la estimación de cada magnitud se pueden poner en la cabecera de la columna correspondiente, si son comunes a todas las medidas; si no, pueden ponerse detrás de cada medida usando comillas para no tener que repetir su escritura innecesariamente.

## Ejemplo

A modo de ejemplo se puede considerar la siguiente tabla:

**Tabla 5.1** Tabla en la que se han seguido las recomendaciones para la presentación de datos

Determinación experimental del campo magnético de un solenoide				
Frec. ( $\pm 10$ ) (Hz)	$I_0$ ( $\pm 0,01$ ) (mA)	$V_{pp}$ ( $\pm 0,1$ div) (medida)	$V_{pp}$ (V)	$B_{exp}$ (mT)
400	9,62	5,0 div $\times$ 0,02 V/div	0,100 $\pm$ 0,002	0,159
600	9,58	3,2 div $\times$ 0,05 V/div	0,160 $\pm$ 0,005	0,160
800	9,52	4,0 div $\times$ 0,05 V/div	0,200 $\pm$ 0,005	0,159
1000	9,46	4,8 div $\times$ 0,05 V/div	0,240 $\pm$ 0,005	0,158
Media: 0,159 $\pm$ 0,001				

- El título que aparece justo encima de la tabla nos permite identificar el tipo de experimento realizado, o la fase en la que nos encontramos del mismo.
- En la tabla observamos que las tres primeras columnas corresponden a los datos obtenidos directamente de los aparatos, la cuarta columna a un cálculo intermedio y la última al resultado final buscado. La tercera y cuarta columnas representan la misma magnitud; al tomar los datos en el laboratorio siempre es preferible hacerlo tal y como aparecen en la tercera columna, ya que será mucho más fácil detectar cualquier posible error cometido, pero en el informe final podemos prescindir de esta columna.
- La estimación del error que aparece en las cabeceras de las tres primeras columnas nos indica que el error absoluto es el mismo para todos los datos de la columna. El error correspondiente a la magnitud final, quinta columna, se ha obtenido sin embargo por medio de la estimación estadística y aparece junto a la media de la magnitud en la última fila. Esto puede hacerse así porque sabemos que los datos de la última columna deben representar un mismo valor obtenido en diferentes condiciones.
- El criterio seguido tanto para la estimación del error absoluto de cada valor de la cuarta columna como para el cálculo por estimación estadística de la magnitud media de la quinta columna, debe quedar claro en el informe al que pertenece la tabla.
- Las unidades de cada magnitud han sido elegidas para que los números que figuran en cada columna sean los más sencillos posibles.
- La columna que recoge el resultado final  $B_{exp}$ (mT) podría haberse encabezado como  $B_{exp} \times 10^6$ (T) para utilizar unidades del sistema internacional, pero los datos recogidos en la columna serían: 159, 160, 159 y 158 en cada uno de los experimentos, indicando que los resultados reales son:  $156 \times 10^{-6}$  T,  $160 \times 10^{-6}$  T,  $159 \times 10^{-6}$  T y  $158 \times 10^{-6}$  T respectivamente.
  - NOTA: En algunas publicaciones se puede encontrar otro criterio en esta cuestión, según el cual el encabezamiento  $B_{exp} \times 10^{-6}$ (T) significa que cada dato de la columna encabezada está multiplicado por el factor  $10^{-6}$ , igual que el criterio anterior. Sólo una observación del orden de magnitud medida nos permitirá discernir cual de los dos se ha utilizado en cada caso.

## 5.2 Representaciones Gráficas

El punto de partida para la elaboración de un gráfico, es el análisis de la tabla de valores obtenidos del experimento. Este análisis nos permite tomar decisiones y proceder sobre los siguientes aspectos:

### 1. Escoger variables

En un experimento se suele variar una magnitud (*variable independiente*) con el fin de observar el efecto que se produce sobre otra (*variable dependiente*). Por convenio se representa la variable independiente en *abscisas* (eje horizontal) y la variable dependiente en *ordenadas* (eje vertical).

### 2. Ejes y escalas

- Como ya hemos dicho, los ejes tienen, por convenio, una función predeterminada: sobre el eje horizontal se representa la variable independiente (la que nosotros variamos) y sobre el vertical la variable dependiente (la que nos muestra el efecto). Los ejes deben llevar claramente indicada la magnitud que representan, el intervalo de medida y las unidades en que se expresan los datos.
- La elección de los intervalos que determinan la escala no es arbitraria: el intervalo representado en el eje debe concordar con el rango de la magnitud representada, de manera que todos los datos figuren dentro de la gráfica y ocupen la mayor parte del área de ésta.
- Los ejes deben llevar indicaciones del valor de la magnitud a intervalos regulares, que no tiene por qué coincidir con los valores de los puntos experimentales; es decir, una misma longitud de eje no puede corresponder a dos intervalos distintos de valores de la magnitud.
- No es necesario marcar el valor de todos y cada uno de los intervalos y sí deben ser elegidos de forma que el valor de la magnitud se lea con comodidad (por ejemplo, múltiplos o submúltiplos de 10).
- No es necesario que el origen, el punto de coordenadas (0,0), esté incluido en la gráfica, incluso puede llegar a ser contraproducente.

### 3. Representación de los datos medidos

Las representaciones gráficas deben llevar claramente indicados los datos obtenidos experimentalmente; para ello es necesario señalarlos sobre la gráfica con un símbolo cuyo tamaño y forma permita apreciarlos (y distinguir unos de otros cuando correspondan a diferentes conjuntos de medidas, que se suelen denominar *series*).

Puesto que todos los puntos experimentales van acompañados de un cierto grado de error, es recomendable, siempre que sea posible, hacer la representación de dichos puntos acotando una región dentro de la cual cualquier punto puede representar el valor real de la observación, de acuerdo con la escala empleada. La no acotación de un trazo, se interpreta como la imposibilidad de representar el error (por ser muy pequeño) en estos casos es conveniente representar los puntos mediante un pequeño círculo.



Fig. 5.1 Formas de presentar la existencia o no del error en una gráfica.

En el ejemplo siguiente se representa  $x = f(t)$  de acuerdo con los resultados de la Tabla 5.2.

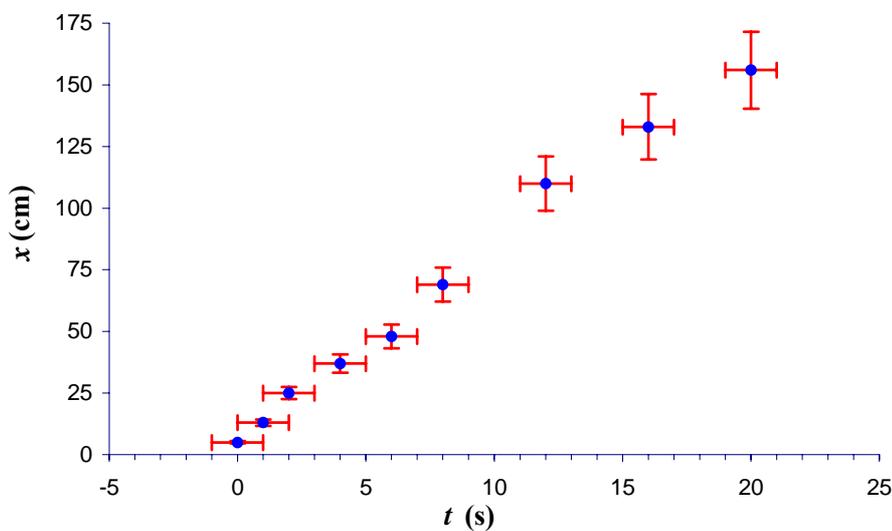


Fig. 5.2 Representación de la posición,  $x$ , con el tiempo,  $t$

Tabla 5.2 Posición,  $x$ , en función del tiempo,  $t$ , para un móvil

$t (\pm 1)$ (s)	0	1	2	4	6	8	12	16	20
$x (\pm 10\%)$ (cm)	5,0	13,0	25,0	37,0	48,0	69,0	110,0	133,0	156,0

Para el tiempo, los trazos son del mismo tamaño para todos los puntos, ya que representan 1 segundo por encima del valor y 1 segundo por debajo del valor. Para las posiciones, la longitud de los trazos varía debido a que el error de cada medida (10 %) es diferente.

#### 4. Representación del comportamiento de las magnitudes

Los puntos obtenidos experimentalmente no son la única información que puede deducirse de un experimento, el “espacio” entre los puntos debe ser completado. Para ello en las representaciones gráficas hay que añadir una línea que indique la tendencia (la ley física) que rige el comportamiento de una magnitud frente a la otra. Esta línea debe tener en cuenta que los datos experimentales están afectados por errores, de los que ya hemos hablado, y que por tanto las líneas no tienen por que pasar por todos y cada uno de esos puntos experimentales. Como criterio general (el sentido común es aplicable siempre) debemos tener en cuenta que:

- Si se representan variables continuas en cada uno de los ejes, la curva debe ser continua. Las variables no sufren casi nunca cambios bruscos, por lo que las líneas no están compuestas nunca por segmentos rectos; las inflexiones son siempre suaves, y por tanto deben trazarse líneas curvas (o una única recta) que represente el comportamiento de las magnitudes involucradas en el experimento.
- Las líneas deben pasar por la mayoría de las zonas de observación de los puntos experimentales; si uno de ellos se aleja demasiado del comportamiento general del resto de los puntos, debe ser estudiado y, en consecuencia, corregido o desechado.
- Los puntos experimentales deben estar distribuidos, en lo posible, uniformemente a cada lado de la curva.

Como ejemplo, podríamos trazar la recta correspondiente al ejemplo anterior, lo cual se observa en la Fig. 5.3.

Al aplicar este método geométrico, siempre existe la posibilidad de trazar más de una recta, aunque se apliquen todos los criterios anteriormente expuestos. La forma adecuada de ajustar una expresión analítica a los puntos experimentales se basa en los métodos de regresión que se estudian en el apartado 5.3.

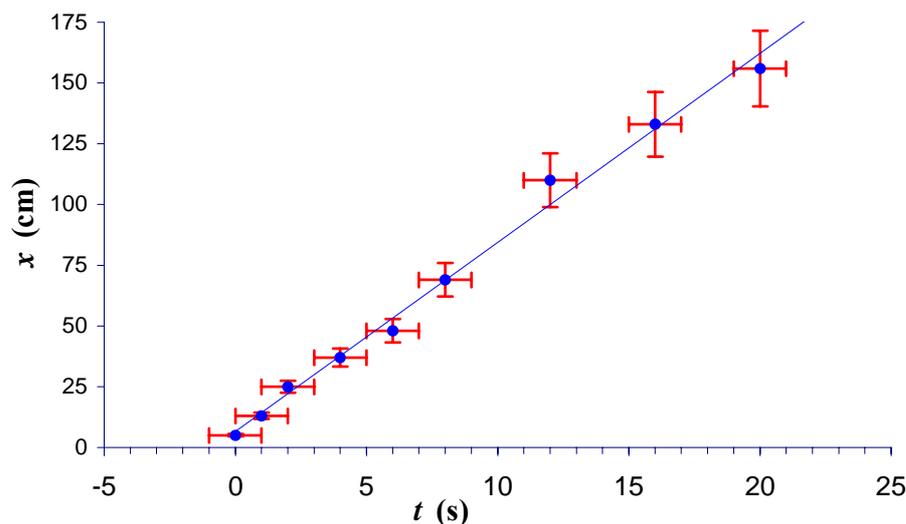


Fig. 5.3 Representación de la variación del espacio,  $x$ , con el tiempo,  $t$

Cuando sea posible realizar un ajuste por regresión (como por ejemplo el de mínimos cuadrados) la línea del ajuste debe representarse sobre los datos experimentales.

## 5. Identificación y título del gráfico

Hay que tener en cuenta que toda representación debe ir acompañada de un título que indique claramente el estudio que se ha realizado y el sistema físico que es objeto de estudio. Este rótulo suele situarse en la parte superior o inferior de la gráfica y no debe contener información ya recogida en los ejes (por ejemplo, las unidades de la magnitud representada).

## 6. Papeles especiales

En algunos casos la representación de la gráfica puede exigir la utilización de un papel especial por diversos motivos:

- La magnitud (o magnitudes) representada cubre un intervalo muy grande o afecta a distintas escalas (por ejemplo al medir el comportamiento de un sistema respecto de la frecuencia en distintos rangos de frecuencias desde los Hz hasta los MHz).
- Se quiere “linealizar” una curva (por ejemplo una variación exponencial).
- Se quiere representar una magnitud en función de la dirección en la que se mide.
- Se quiere representar la dependencia entre tres variables en lugar de dos, como es habitual.

En los dos primeros casos se utilizará un papel semilogarítmico o logarítmico. En el papel semilogarítmico el espaciado entre divisiones en uno de los ejes es proporcional al logaritmo decimal de la magnitud, así el espaciado entre 1 y 2 (o entre 1 y 10) es el mismo que entre 10 y 20 (o entre 10 y 100). En el papel logarítmico esto ocurre para los dos ejes.

Para el tercer caso se suele utilizar el “papel polar” en el que las divisiones corresponden a radios y circunferencias en lugar de líneas horizontales y verticales.

En el cuarto caso se utiliza un papel con representación triangular.

## 7. Ejemplos y errores frecuentes

Utilizaremos un ejemplo para mostrar los aspectos mencionados en los apartados anteriores.

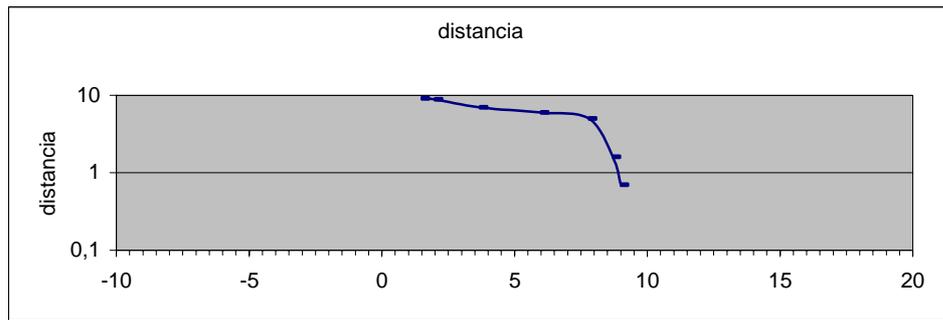
Intentaremos representar la curva de variación del campo magnético en el interior de un solenoide. Para ello medimos la fuerza electromotriz,  $V$ , inducida en un carrete en función de su posición en el solenoide,  $X$ .

Así, obtenemos los datos de la tabla que figura a continuación:

**Tabla 5.3 Fuerza electromotriz inducida en un carrete,  $V$ , y posición de éste,  $X$ , dentro de un solenoide**

$X$ (cm)	$V$ (mV)
0	9,41
1	9,41
2	9,24
3	8,74
4	7,39
5	5,38
6	3,36
7	2,02
8	1,34

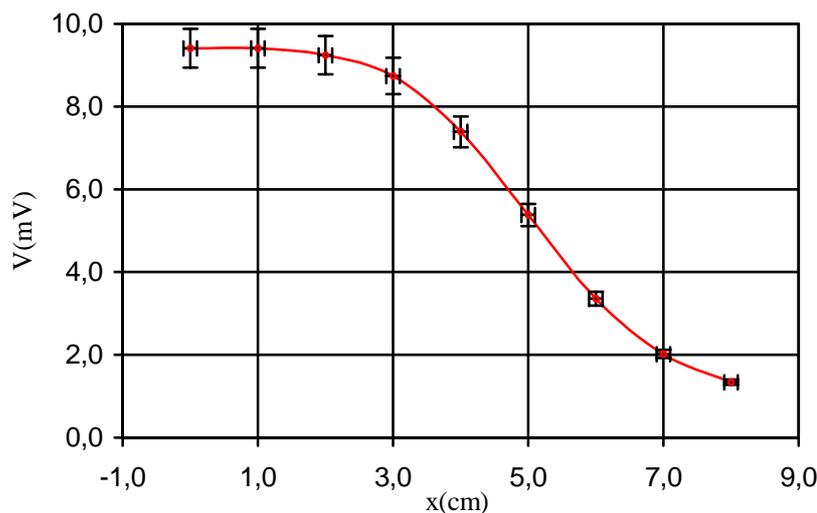
Puede tomarse esta gráfica realizada a partir de los datos de la tabla anterior como modelo de las cosas que **NO** deben hacerse:



**Fig. 5.4** Gráfica que muestra los errores más frecuentes en las representaciones gráficas

- No todos los puntos experimentales están representados en la figura.
- Los puntos experimentales no figuran o no se distinguen.
- Los puntos están unidos por líneas (segmentos escalonados) que no corresponden a una variación continua.
- El espacio ocupado por los puntos experimentales dentro de la gráfica representan una parte mínima del espacio total debido a una mala elección del intervalo de representación.
- El color añadido al fondo estorba al restar claridad a la representación.
- Los ejes X e Y tienen longitudes muy dispares. El tamaño del eje vertical es demasiado pequeño.
- El título de la gráfica no da ninguna información sobre lo que se pretende representar, y es redundante ya que “dice” lo mismo que el rótulo del eje Y.
- Falta el rótulo del eje X. Faltan las unidades del eje Y.
- Los intervalos entre marcas en el eje Y no están equiespaciados ni representan intervalos iguales entre datos.
- La abundancia de marcas en el eje X, y la falta de ellas en el eje Y hacen que sea imposible estimar sobre la gráfica los valores de los puntos experimentales.
- Puede ser útil incluir un *pie de figura* que aclare el significado de la representación.

Una gráfica adecuada podría ser la siguiente:



**Fig. 5.5** Variación de la f.e.m. inducida en el interior de un solenoide,  $V$ , en función de la posición,  $x$

### 5.2.1 Escalas Lineales

Ya se ha visto que cualquier dependencia, funcional o no, entre dos variables, puede ser representada gráficamente en un plano y se han dado las normas básicas que habitualmente se siguen para la confección y presentación de tales gráficos.

La ventaja principal que se obtiene al representar datos numéricos de forma gráfica es que se hace visible la relación que estos datos guardan entre sí.

Generalmente, las escalas que se utilizan en cada uno de los ejes de coordenadas son escalas **lineales**, es decir, la unidad de la magnitud que se representa se corresponde con un segmento de longitud constante a lo largo de todo el eje o, dicho de otro modo, a diferencias numéricas iguales corresponden segmentos de igual longitud. En la Fig 5.6 se indican dos gráficas preparadas ambas con escalas lineales.

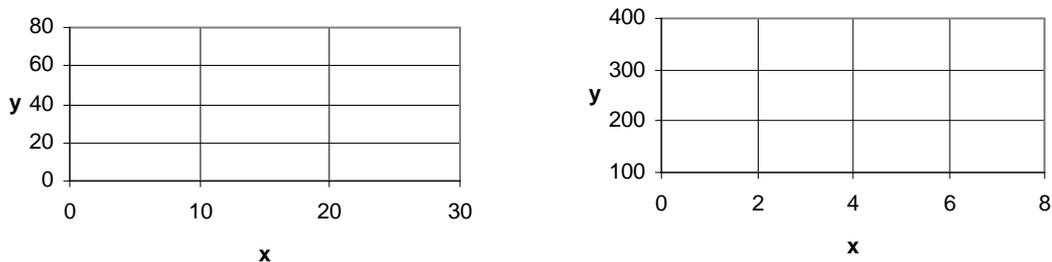


Fig. 5.6 Gráficas preparadas con escalas lineales

Cuando en **ambos** ejes, la unidad de la magnitud que se representa se corresponde con un segmento de igual longitud, se dice que las escalas son isótropas. Cuando se representa una recta en escalas isótropas, entonces (y solamente en este caso) el valor numérico de la pendiente física de la recta es igual al valor de su pendiente geométrica. La pendiente física tiene unidades y la pendiente geométrica no las tiene. En la Fig 5.7 se indica una función lineal con este tipo de escalas.

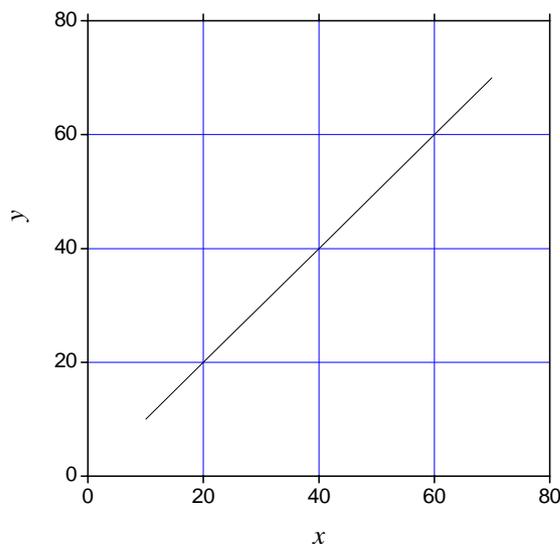


Fig. 5.7 Función lineal representada en un gráfico de escalas isótropas

## 5.2.2 Escalas Logarítmicas

Muchas veces, al representar gráficamente los datos obtenidos en un experimento no resultan ajustables a una recta. En el caso más general, el cálculo numérico nos proporciona herramientas para ajustar tales datos experimentales, por ejemplo, a un polinomio de un cierto grado y obtener después los coeficientes del polinomio.

Una forma intermedia de actuar, que en general conduce a resultados aceptables, consiste en representar los datos que resultan no lineales, en papel logarítmico para ver si de este modo se pueden ajustar a una recta. Básicamente, esto ocurrirá cuando los datos obtenidos en el experimento respondan a una relación potencial o exponencial.

### 1. Función potencial

Una función potencial tiene la forma:

$$[5.1] \quad y = Ax^n$$

donde  $x$  e  $y$  son las variables cuyos valores obtenemos experimentalmente y  $A$  y  $n$  son las constantes físicas a determinar.

Un ejemplo sencillo de este tipo de función puede ser el movimiento rectilíneo uniformemente acelerado:  $y$  sería el espacio y  $x$  el tiempo.

Si tomamos logaritmos decimales en la expresión [5.1] obtenemos:

$$[5.2] \quad \log y = \log A + n \log x$$

y haciendo el cambio de variables:

$$[5.3] \quad \begin{aligned} y' &= \log y \\ x' &= \log x \end{aligned}$$

queda:

$$[5.4] \quad y' = \log A + nx'$$

que es la ecuación de una recta de pendiente  $n$  y ordenada en el origen  $\log A$ .

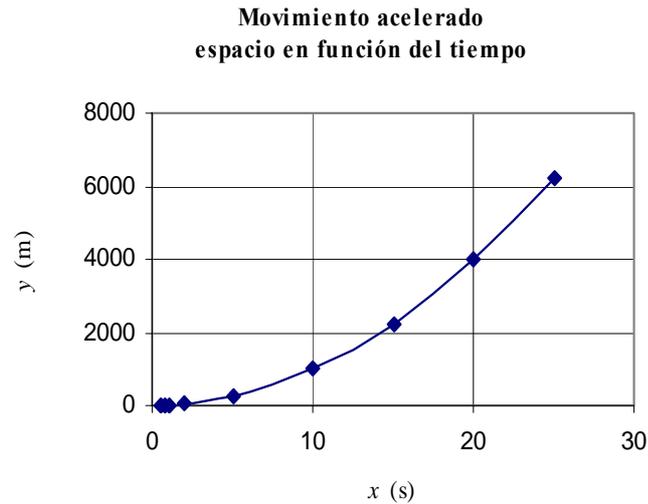
Veamos con un ejemplo la forma de actuar.

Disponemos de los datos experimentales de la Tabla 5.4.

**Tabla 5.4 Datos ajustables a una función potencial**

$x$	0,5	0,75	1	2	5	10	15	20	25
$y$	2,5	5,5	9,8	39,7	252	995	2251	4010	6245

Si los representamos en papel milimetrado, Fig. 5.8, vemos claramente que no son ajustables a una recta, aunque tienen un aspecto de parábola.



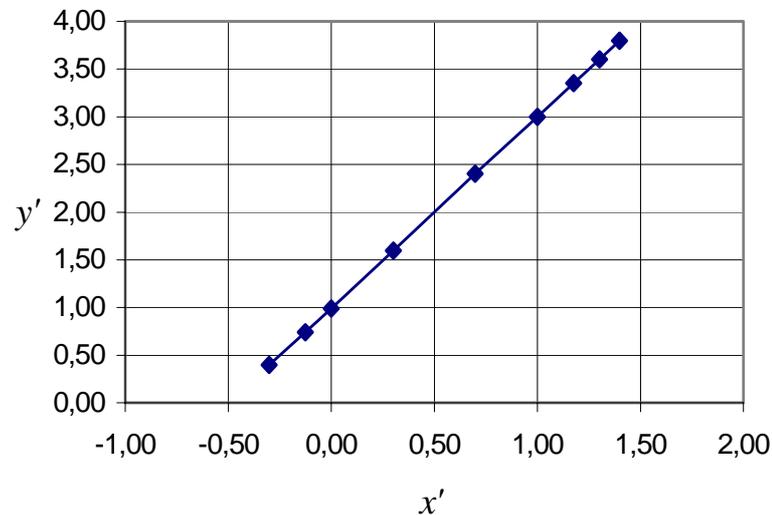
**Fig. 5.8** Gráfica de una función potencial de tipo parabólico

En la Tabla 5.5 aparecen los valores de  $\log x$  y  $\log y$  correspondientes a los datos de la Tabla 5.4.

**Tabla 5.5** Logaritmos de los datos ajustables a una función potencial

$x' = \log x$	-0,30	-0,12	0,00	0,30	0,70	1,00	1,18	1,30	1,40
$y' = \log y$	0,40	0,74	0,99	1,60	2,40	3,00	3,35	3,60	3,80

Y en la Fig. 5.9 se muestra su representación gráfica, donde vemos que ahora sí pueden ser ajustados a una recta.



**Fig. 5.9** Representación de los logaritmos de los datos ajustables a una función potencial

Obsérvese que se han representado  $y' \equiv \log y$  en función de  $x' \equiv \log x$  y en las escalas aparecen valores suficientes para los rangos de  $x'$  y de  $y'$ .

Utilizando la recta de la Fig. 5.9 se podrían obtener las constantes  $n$  y  $A$ . En efecto, la pendiente de la recta representada viene dada por:

$$[5.5] \quad n = \frac{y'_2 - y'_1}{x'_2 - x'_1}$$

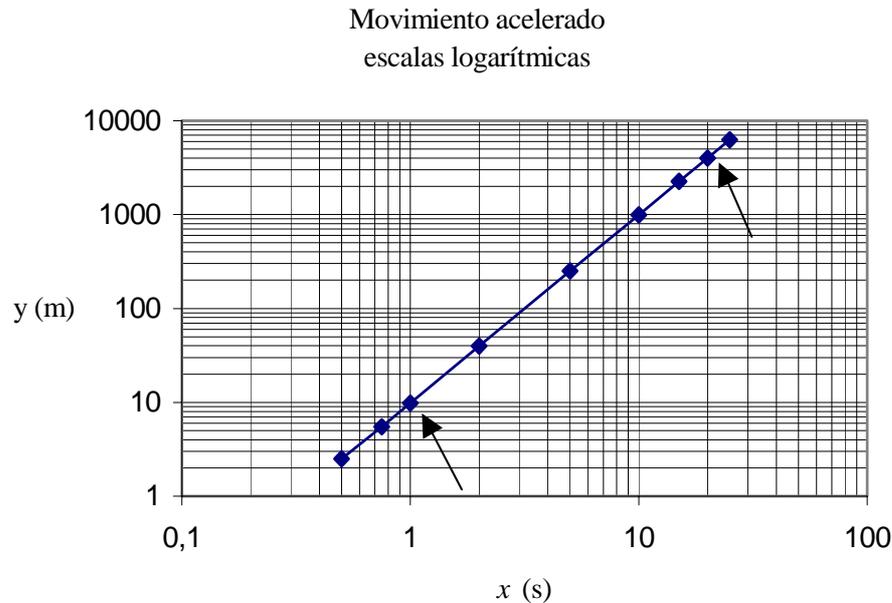
y tomando, por ejemplo, los puntos (0, 1) y (1, 3) se obtiene  $n = 2$ .

Para  $x' = 0$ , se lee  $y' = 1$  con lo que  $\log A = 1$  de donde  $A = 10$ , y la relación buscada queda de la forma:

$$[5.6] \quad y = 10x^2$$

Sin embargo, la forma habitual de procedimiento no es la que se ha descrito. Lo que realmente se hace, para evitar tener que tomar logaritmos en toda la tabla de datos, es representar los datos  $x$  e  $y$  en ESCALAS LOGARÍTMICAS, que son las mismas que las anteriores (Fig. 5.9) pero en lugar de aparecer rotulados los rangos de valores de  $x'$  e  $y'$  aparecen rotulados los rangos de valores de  $x$  e  $y$ .

En estas escalas logarítmicas (Fig. 5.10) se **LEE** el valor de  $x$  o de  $y$  pero **ESTÁ REPRESENTADO SU LOGARITMO**.



**Fig. 5.10 Representación en escalas logarítmicas de los datos ajustables a una función potencial**

Para hallar la pendiente de la recta en este gráfico, se hace de la siguiente forma. Como por la fórmula [5.5]:

$$[5.7] \quad n = \frac{y'_2 - y'_1}{x'_2 - x'_1}$$

restituyendo el cambio de variables:

$$[5.8] \quad n = \frac{\log y_2 - \log y_1}{\log x_2 - \log x_1} = \frac{\log\left(\frac{y_2}{y_1}\right)}{\log\left(\frac{x_2}{x_1}\right)}$$

y tomando, por ejemplo, los puntos (1, 10) y (20, 4000), tenemos:

$$[5.9] \quad n = \frac{\log\left(\frac{4000}{10}\right)}{\log\left(\frac{20}{1}\right)} = \frac{2,60}{1,30} = 2$$

y para  $x = 1$  (o lo que es lo mismo  $x' = 0$ ), se lee  $y = 10$ , de donde  $A = 10$ .

La ordenada en el origen se lee para  $x = 1$  porque  $\log 1 = 0$ .

El papel pautado como aparece en la Fig. 5.10 se llama de **escalas doblemente logarítmicas**.

## 2. Función exponencial

Una función exponencial es de la forma:

$$[ 5.10] \quad y = Ba^{nx}$$

y tomando logaritmos:

$$[ 5.11] \quad \log y = \log B + nx \log a$$

y llamando:

$$[ 5.12] \quad m = n \log a$$

queda:

$$[ 5.13] \quad \log y = mx + \log B$$

Evidentemente, las parejas de valores  $(x_i, y_i)$  que obtengamos a partir de algún experimento, representadas en papel milimetrado normal, no responderán a una gráfica lineal, pero si se representa ahora  $y' = \log y$  en función de  $x$ , se obtiene una recta cuya pendiente se calcula, como siempre, tomando dos puntos de la recta,  $(x_1, y'_1)$   $(x_2, y'_2)$  y la pendiente es:

$$[ 5.14] \quad m = \frac{y'_2 - y'_1}{x_2 - x_1} = \frac{\log\left(\frac{y_2}{y_1}\right)}{x_2 - x_1}$$

y la ordenada en el origen, para  $x = 0$  es  $y'(0) = \log B$ , de donde se obtiene el valor de  $B$ .

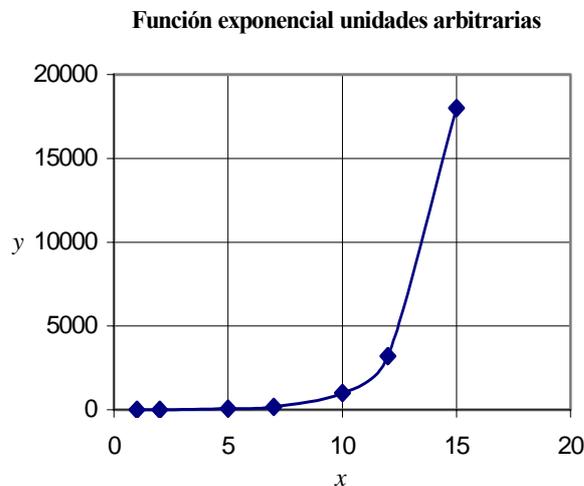
En este caso, vemos que solamente una de las variables se representa en escala logarítmica, lo que se consigue utilizando el **papel semilogarítmico**.

En la Tabla 5.6 se indica un conjunto de pares de valores que podrían responder a una función exponencial.

**Tabla 5.6 Datos ajustables a una función exponencial**

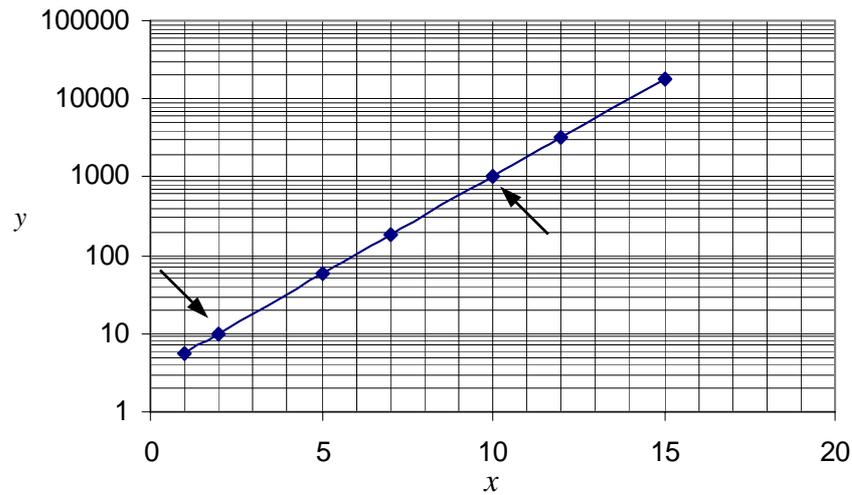
$x$	1	2	5	7	10	12	15
$y$	5,6	10,1	56,9	180	1012	3200	18000

En la Fig. 5.11 se representan tales datos en escalas lineales y se observa que no están sobre una recta.



**Fig. 5.11 Gráfica de una función exponencial**

En la Fig. 5.12 se representan los mismos datos en escala semilogarítmica (eje de abscisas lineal) y se ve que ya aparecen alineados, por lo que corresponden a una función exponencial.



**Fig. 5.12 Representación en escalas semilogarítmicas de los datos ajustables a una función exponencial**

Para obtener el valor  $m$  de la pendiente tomamos, por ejemplo, los puntos: (2, 10), (10, 1000) y operando:

$$[5.15] \quad m = \frac{\log\left(\frac{y_2}{y_1}\right)}{x_2 - x_1} = \frac{\log\left(\frac{1000}{10}\right)}{10 - 2} = \frac{2}{8} = 0,25$$

La ordenada en el origen se lee directamente, prolongando la recta hasta que corte al eje  $y$  en  $x = 0$ ,

$$[5.16] \quad y'(x=0) = \log B \Rightarrow B \approx 3,1$$

pero hay que advertir que, en muchas ocasiones, se puede cometer un error excesivo y habrá que recurrir a procedimientos de regresión lineal.

Conocida  $m$  se puede restituir la función [5.11] a su representación original indicada en [5.10], para una base cualquiera conocida.

Naturalmente, en todos los casos anteriores, se puede hacer un tratamiento de regresión lineal como se verá más adelante, en lugar de un simple ajuste "a ojo".

### 5.3 Regresiones y Ajustes

Con mucha frecuencia, en el desarrollo de proyectos técnico-científicos se plantea un mismo tipo de problema, aunque puede manifestarse a través de una doble vertiente:

- ¿Existe algún tipo de correlación entre dos o más magnitudes experimentales de un mismo sistema?. Es decir que un cambio en una de las variables suponga un cambio proporcional en las otras.
- Si la relación entre variables existe, ¿puede deducirse una ley matemática simple que la reproduzca analíticamente?

En realidad ambas cuestiones se encuentran completamente relacionadas, llegando a ser las dos caras de un mismo problema. Si somos capaces de encontrar una ley matemática que ligue ciertas variables de un sistema, es evidente que guardarán algún tipo de relación entre sí, y viceversa. De momento mantendremos separadas ambas cuestiones y profundizaremos en ellas mediante el análisis de dos supuestos concretos.

#### *Supuesto n° 1*

Imaginemos que tengamos una sospecha razonada de que un cierto factor cuantificable pueda influir en el comportamiento de una variable experimental. Es desde luego una situación cotidiana en cualquier ámbito tecnológico. Necesitaremos de algún procedimiento matemático para evaluar la correlación.

Por ejemplo, podríamos plantear un estudio para ver las posibles ventajas de incluir pequeñas concentraciones del ión  $\text{Cr}^{3+}$  en el crecimiento de ciertas plantas. Podríamos diseñar el experimento adecuado y obtener unos resultados que vamos a suponer semejantes a los que aparecen en la siguiente figura:

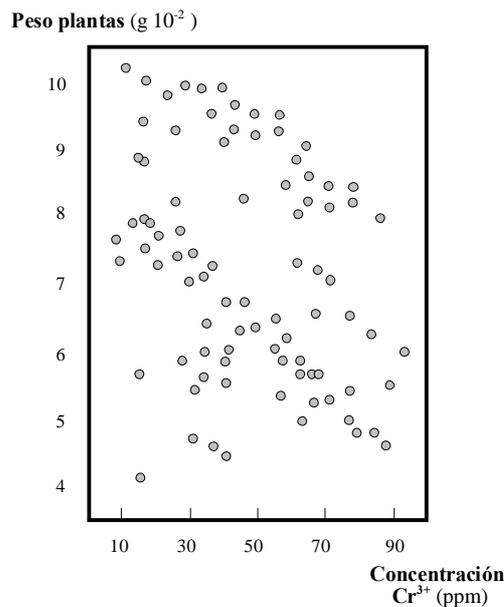


Fig. 5.13 Peso de las plantas en función de la concentración de  $\text{Cr}^{3+}$

A la vista de los resultados obtenidos, volvemos a plantearnos la cuestión que dio origen a este supuesto estudio. ¿Existe alguna influencia entre el ión  $\text{Cr}^{3+}$  y el crecimiento de las plantas?. Si los resultados presentados en la gráfica fueran reales, el lector podría ya obtener sus propias conclusiones al respecto. De lo que sí estamos seguros es de que este problema, analizado por un número suficiente de lectores repartiría las interpretaciones a favor y en contra de una forma bastante equitativa.

La razón en el fondo es bien simple. Se trata de un conjunto de resultados que están más allá de la capacidad humana para distinguir objetivamente un patrón geométrico regular. Necesitaríamos de algún procedimiento matemático que evalúe por nosotros, y de una forma más rigurosa, si estos datos están correlacionado y en definitiva, si debemos dar alguna validez a la hipótesis inicial.

A lo largo de este capítulo veremos que los llamados *métodos de regresión* son una de las herramientas disponibles para abordar este tipo de cuestiones.

### Supuesto n° 2

En otras ocasiones la correlación empírica entre dos o más variables de un mismo sistema, está ya plenamente verificada, pero nos falta aún saber si esta correlación se manifiesta en forma de una función matemática simple. Esto supondría grandes ventajas, ya que de este modo podríamos cuantificar el comportamiento del sistema.

Por ejemplo, se ha verificado adecuadamente que la resistencia a la tracción del acero (una aleación de hierro y carbono), es sensible a la proporción de otros metales minoritarios presentes en la misma. ¿Podríamos predecir la resistencia a la tracción de un cierto tipo de acero en función de su contenido en wolframio?. Para ello mediríamos experimentalmente esta propiedad en diferentes muestras de acero y posteriormente representaríamos los resultados obtenidos. Para continuar con el ejemplo, supongamos que los resultados obtenidos son los que aparecen en la siguiente gráfica:

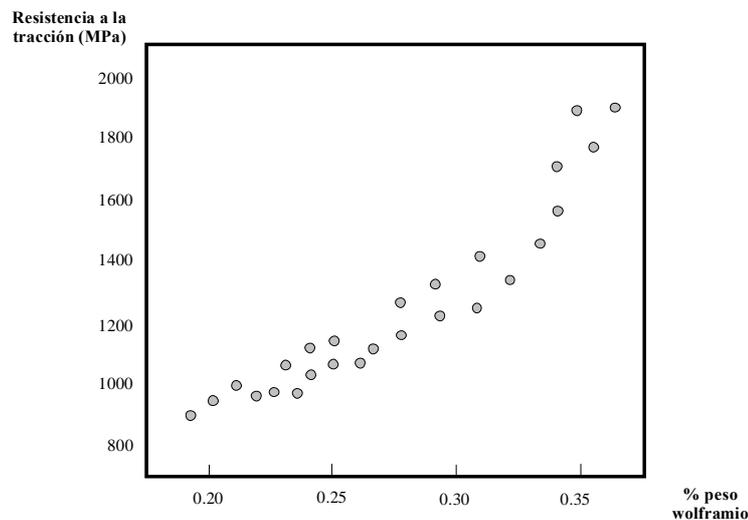


Fig. 5.14 Resistencia a la tracción de un acero en función de la cantidad de wolframio

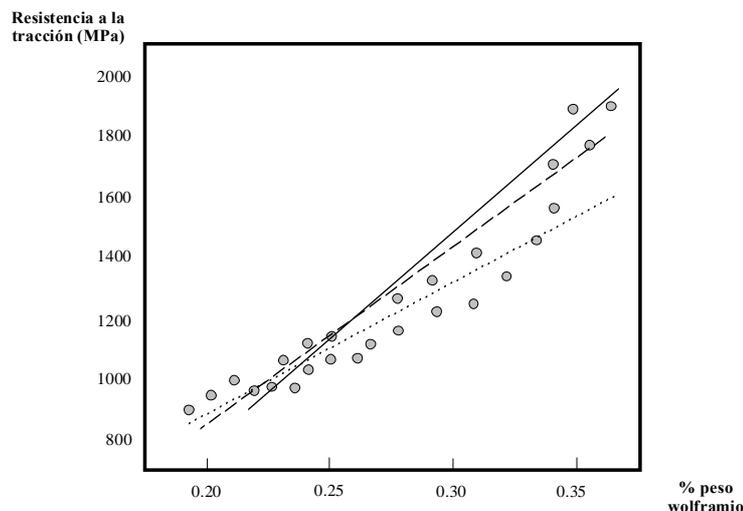
A simple vista ya es fácil observar que existe una notable correlación entre los factores estudiados (lo cual ya era un requisito previo del estudio), pero ¿existe alguna función matemática simple que reproduzca este comportamiento?. Este sería el segundo, y principal problema, al que pretendemos dar respuesta a lo largo del presente capítulo.

### 5.3.1 Correlación de Variables

Sigamos de nuevo trabajando con los datos del último supuesto, referido al acero. A la vista de los resultados gráficos, estaríamos dispuestos a concluir que si son presentados para su interpretación, simplemente visual, a un grupo de personas, la mayoría de ellas concluirían que hay una clara correlación de variables. La siguiente cuestión sería saber la función matemática que representa esta correlación. Si dejamos que este conjunto de observadores opine también a este respecto, seguramente muchos podrían pensar que una simple recta podría ser suficiente para reproducir estos datos. La verdadera dificultad se genera cuando les pidamos que tracen esta recta. Podríamos encontrarnos con una situación parecida a la que aparece en la Fig. 5.15.

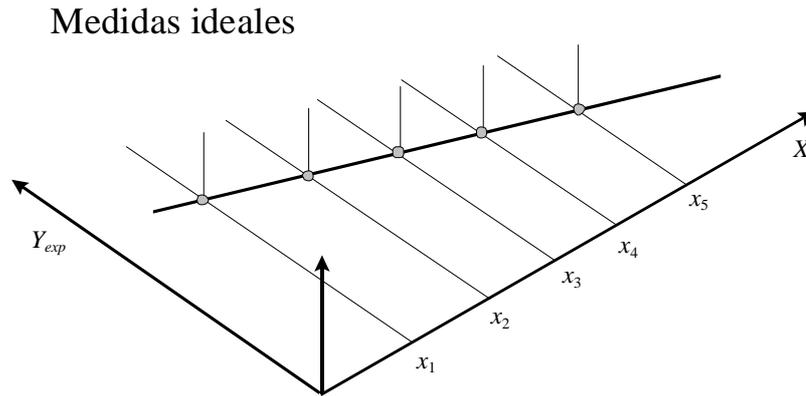
Es decir, diferentes personas trazarían diferentes rectas de ajuste. Todos, en principio, pueden llevar parte de razón, pero de todas las rectas posibles, habrá una que sea la mejor de todas: *aquella que se acerque a un mayor número de puntos experimentales.*

Este podría ser un buen planteamiento para encontrar la mejor recta pero no es muy eficaz desde el punto de vista matemático (los humanos podemos contar fácilmente, pero los métodos matemáticos no pueden hacerlo. Solo pueden establecer relaciones entre variables). Lo cambiaremos por otro planteamiento, más operativo, y que al final, veremos que engloba esta misma idea.



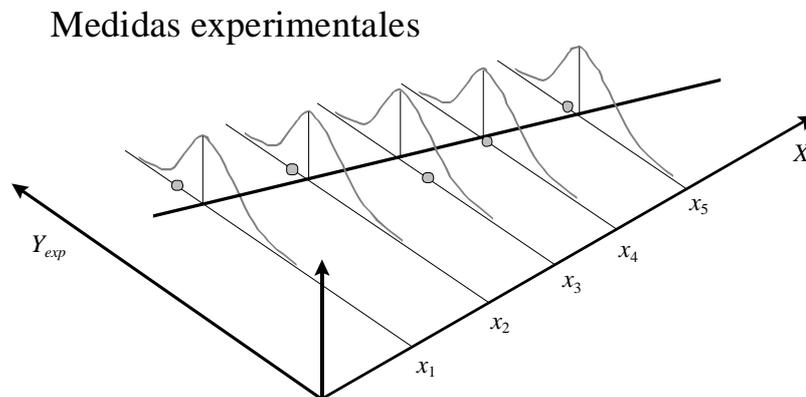
**Fig. 5.15** Posibles rectas que ajustan los datos de resistencia a la tracción de un acero

La cuestión de partida será ahora preguntarnos por qué los puntos experimentales no están todos correctamente alineados en un hábito concreto (por ejemplo una recta, parábola, exponencial, etc). Ya sabemos suficientemente que las medidas experimentales vienen afectadas inevitablemente por errores e imprecisiones, y que en el mejor de los casos este error se debe simplemente a la influencia errática del azar. De esta forma podríamos suponer como hipótesis de partida que la dispersión de puntos obtenida, se debe exclusivamente a la influencia del azar. Si esta influencia no estuviera presente, nuestros puntos estarían perfectamente alineados sobre una curva matemática tal y como se observa en la Fig. 5.16.



**Fig. 5.16 Puntos experimentales perfectamente alineados**

Observamos en la figura 5.17 que cada dato experimental ( $y_{exp}$ ) aparece disperso según una distribución, que previsiblemente será gaussiana en torno al valor central marcado por la función matemática en cada punto  $x_i$ .



**Fig. 5.17 Puntos experimentales dispersos de acuerdo a una distribución gaussiana**

Si realizamos correctamente una sola medida de la variable experimental, sabemos que el valor obtenido estará más o menos próximo al valor exacto (aunque desconocido) y que sucesivas medidas en las mismas condiciones, irían generando resultados distribuidos con una probabilidad gaussiana a lo largo de un cierto intervalo. Por eso, cabe esperar que cada una de nuestras medidas esté dispersa en torno al valor exacto, por el que debería pasar la función.

Si las medidas de todos los puntos experimentales han sido realizadas en similares condiciones, es de esperar que la varianza  $\sigma^2$  de la distribución sea la misma para cada punto. La probabilidad de obtener el valor experimental  $y_1$  al realizar una medida es:

$$[ 5.17] \quad P(y_1) = \frac{1}{\sigma\sqrt{2}} e^{-(y_1 - y_1^{exacto})^2 / 2\sigma^2}$$

Como estamos admitiendo que hay una cierta función matemática  $Y(x)$  que pasa por los valores exactos, tendremos:

$$[ 5.18] \quad P(y_1) = \frac{1}{\sigma\sqrt{2}} e^{-(y_1 - Y(x_1))^2 / 2\sigma^2}$$

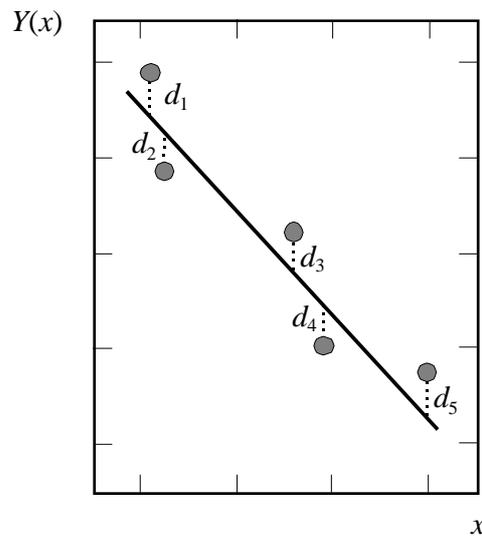
y en general, la probabilidad de obtener los  $n$  datos del experimento será:

$$[5.19] \quad P(y_1, y_2, \dots, y_n) = \frac{1}{\sigma\sqrt{2}} e^{-\sum_{i=1}^n [y_i - Y(x_i)]^2 / 2\sigma^2}$$

Esta probabilidad será máxima cuando al argumento de la exponencial sea un mínimo:

$$[5.20] \quad D = \sum_{i=1}^n d_i^2 = \sum_{i=1}^n [y_i - Y(x_i)]^2 = \min$$

lo cual se corresponde con una idea fácilmente intuitiva. En efecto, cada uno de los términos  $d_i$  representa la distancia entre el valor obtenido experimentalmente ( $y_i$ ) en una medida y el valor exacto de la misma ( $Y_i$ ). Por tanto la mejor función de ajuste será aquella que haga mínimas las distancias entre los puntos experimentales y la función de ajuste.



**Fig. 5.18** Distancia  $d_i$  entre el valor experimental y el exacto para encontrarse alineado

Al principio de este epígrafe, intentábamos encontrar la función que pasaba por un mayor número de puntos experimentales, argumento que no era directamente computable. Ahora lo hemos reemplazado por algo equivalente, pero que sí resulta abordable mediante simple cálculo.

Cualquier procedimiento encaminado a encontrar la mejor función de ajuste a un conjunto de puntos, se denomina *Método de Regresión*. Con frecuencia, en la bibliografía también se les denomina ajuste por *Mínimos Cuadrados*, aludiendo a la minimización de la suma de distancias al cuadrado, entre los puntos experimentales y la recta de ajuste. En cualquier caso, recuerde que no todos los métodos de regresión se basan en los mínimos cuadrados, existiendo otros métodos de ajuste a los que no haremos referencia.

### 5.3.2 Regresión Lineal

Llegados a este punto se hace imprescindible hacer alguna hipótesis sobre la función  $Y(x)$  antes de seguir avanzando. La situación más sencilla con la que nos podemos encontrar es que la función de ajuste sea una recta. Cualquier recta tiene por ecuación general:

$$[5.21] \quad Y(x_i) = a x_i + b$$

De esta forma, a través de los parámetros  $a$  y  $b$  podremos especificar cualquiera de las infinitas rectas posibles. De todas ellas, ¿cuál es la que nos interesa?. Por supuesto aquella que haga mínimo el sumatorio  $D$ :

$$[5.22] \quad D = \sum_{i=1}^n [y_i - (a x_i + b)]^2 = \min$$

A partir de aquí el desarrollo tan solo consiste en un proceso de cálculo elemental. La condición de mínimo se alcanza cuando la derivada primera se anula (también damos por sentado, aunque se podría demostrar, que la derivada segunda será un máximo en dichas circunstancias).

$$[5.23] \quad \frac{\partial D}{\partial a} = 2 \sum_{i=1}^n [y_i - (a x_i + b)] (-x_i) = 0$$

$$[5.24] \quad \frac{\partial D}{\partial b} = 2 \sum_{i=1}^n [y_i - (a x_i + b)] (-1) = 0$$

Estas dos condiciones simultáneas generan un sistema lineal de dos ecuaciones:

$$[5.25] \quad \left. \begin{array}{l} \sum_{i=1}^n [y_i - (a x_i + b)] x_i = 0 \\ \sum_{i=1}^n [y_i - (a x_i + b)] = 0 \end{array} \right\} \rightarrow \begin{array}{l} \sum_{i=1}^n y_i x_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0 \\ \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - n b = 0 \end{array}$$

que puede ser expresado de forma más compacta como:

$$[5.26] \quad \left. \begin{array}{l} S_1 - a S_2 - b S_3 = 0 \\ S_4 - a S_3 - n b = 0 \end{array} \right\}$$

donde hemos llamado:

$$[5.27] \quad S_1 = \sum_{i=1}^n x_i y_i, \quad S_2 = \sum_{i=1}^n x_i^2, \quad S_3 = \sum_{i=1}^n x_i, \quad S_4 = \sum_{i=1}^n y_i, \quad S_5 = \sum_{i=1}^n y_i^2$$

Finalmente las soluciones del sistema toman la forma:

$$[5.28] \quad Y(x) = a x + b, \quad a = \frac{n S_1 - S_3 S_4}{n S_2 - S_3^2}, \quad b = \frac{S_2 S_4 - S_1 S_3}{n S_2 - S_3^2}$$

## Ejemplo

Se ha estudiado en el laboratorio la absorbancia ( $A$ ) de un cierto catión  $B^{3+}$ , a una longitud de onda de 327 nm, y a diferentes concentraciones. Los resultados aparecen en la siguiente tabla:

**Tabla 5.7 Absorbancia de un catión a diferentes concentraciones**

Absorbancia	0,102	0,214	0,314	0,424	0,518
Concentración (mol·litro <sup>-1</sup> )	0,02	0,04	0,06	0,08	0,10

La absorbancia y la concentración de una sustancia absorbente se encuentran relacionadas mediante la ley de Beer-Lambert:

$$[5.29] \quad A = k [B^{3+}]$$

donde  $k$  es una constante característica de cada sistema, desconocida en este caso, pero que puede obtenerse mediante el análisis de regresión de estos datos experimentales.

Para realizar el estudio de regresión lineal, podemos comenzar por construir una tabla en la que calculemos todos los sumatorios necesarios.

**Tabla 5.8 Sumatorios para realizar la regresión lineal**

$i$	$A \equiv y_i$	$[B^{3+}] \equiv x_i$	$x_i y_i$	$x_i^2$	$x_i$	$y_i$
1	0,102	0,02	0,00204	0,0004	0,02	0,102
2	0,214	0,04	0,00856	0,0016	0,04	0,214
3	0,314	0,06	0,01884	0,0036	0,06	0,314
4	0,424	0,08	0,03392	0,0064	0,08	0,424
5	0,518	0,10	0,05180	0,0100	0,10	0,518
$n = 5$			$S_1 = 0,11516$	$S_2 = 0,0220$	$S_3 = 0,30$	$S_4 = 1,572$

La recta de ajuste vendrá dada por la ecuación:

$$[5.30] \quad Y(x) = ax + b$$

donde los coeficientes son:

$$[5.31] \quad a = \frac{n S_1 - S_3 S_4}{n S_2 - S_3^2} = \frac{5 \times 0,11516 - 0,30 \times 1,572}{5 \times 0,0220 - 0,30^2} = 5,21 \text{ M}^{-1}$$

$$[5.32] \quad b = \frac{S_2 S_4 - S_1 S_3}{n S_2 - S_3^2} = \frac{0,0220 \times 1,572 - 0,11516 \times 0,30}{5 \times 0,0220 - 0,30^2} = 0,0018$$

Cabe aún preguntarse por qué obtenemos un valor distinto de cero (aunque pequeño) para la ordenada en el origen, en contradicción con ley de Beer-Lambert. La explicación está en la presencia de errores aleatorios. Nuestra función ajusta *ciegamente* los datos, sin ceñirse a una justificación teórica.

$$[5.33] \quad A = 5,21 [B^{3+}] + 0,0018$$

### 5.3.3 Estimación del Error en un Ajuste

Pese a que hemos mencionado la idea de que, para una función dada, un ajuste por regresión proporciona los coeficientes que mejor reproducen los datos experimentales, estos no necesariamente son valores exactos.

Por simplicidad consideremos el caso de una recta de regresión en la que determinamos su pendiente ( $a$ ) y su ordenada en el origen ( $b$ ). Si repetimos las medidas experimentales para los mismos valores de las abscisas  $x_i$ , obtendremos nuevos valores de  $y_i$ , que en definitiva nos llevarán a coeficientes de ajuste algo diferente a los anteriores. Si repetimos este proceso muchas veces veremos como los distintos valores obtenidos para la pendiente  $a$  tendrían una distribución gaussiana en torno a un valor central, e igualmente obtendríamos para  $b$ .

¿Qué ventaja podríamos sacar a esta situación?. En realidad puede demostrarse que resulta posible calcular la varianza (es decir una medida del error o el margen de incertidumbre) del ajuste, entendida como la desviación de los puntos experimentales con respecto al correspondiente valor obtenido por regresión. La expresión para esta varianza  $s_r^2$  es:

$$[5.34] \quad s_r^2 = \frac{1}{(n-2)} \left\{ \sum_{i=1}^n y_i^2 - \frac{1}{n} \left( \sum_{i=1}^n y_i \right)^2 - a^2 \left[ \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \right] \right\}$$

$$[5.35] \quad s_r^2 = \frac{1}{(n-2)} \left[ S_5 - \frac{1}{n} S_4^2 - a^2 \left( S_2 - \frac{1}{n} S_3^2 \right) \right]$$

donde el significado de los términos  $S_j$  son los que ya definimos en el estudio de la regresión lineal. En su momento se definió también el término  $S_5$  [5.27], el cual no interviene en el cálculo de la ordenada en el origen ni en el de la pendiente. Podemos comprobar ahora que sí interviene en el cálculo de la incertidumbre total del ajuste.

Igualmente puede demostrarse que la varianza de los coeficientes de ajuste se calcula de:

$$[5.36] \quad s_a^2 = \frac{n s_r^2}{n S_2 - S_3^2}, \quad s_b^2 = \frac{s_r^2 S_2}{n S_2 - S_3^2}$$

Normalmente la ordenada en el origen y la pendiente de una recta de ajuste suelen tener algún significado físico concreto, y su determinación suele ser el objeto del estudio. A la vista de las expresiones anteriores observamos que mediante un procedimiento de regresión podemos también determinar el error y, por tanto, el límite de confianza de las magnitudes físicas que representan los coeficientes de la recta de ajuste. En realidad, la raíz cuadrada de estas varianzas puede ser identificada con el error absoluto de dichas magnitudes físicas.

### 5.3.4 Coeficiente de Correlación

Hasta el momento hemos dado por supuesto que podemos realizar el ajuste de un conjunto de datos experimentales a una determinada función. Sin embargo esto no significa que la función elegida represente suficientemente bien los datos de partida. Sería conveniente disponer de algún procedimiento para medir objetivamente si, dentro de la natural dispersión de los puntos, este conjunto puede ser descrito por una función de ajuste. Puede demostrarse que, para un ajuste lineal, el llamado *coeficiente de correlación* ( $r$ ), definido como:

$$[ 5.37] \quad r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

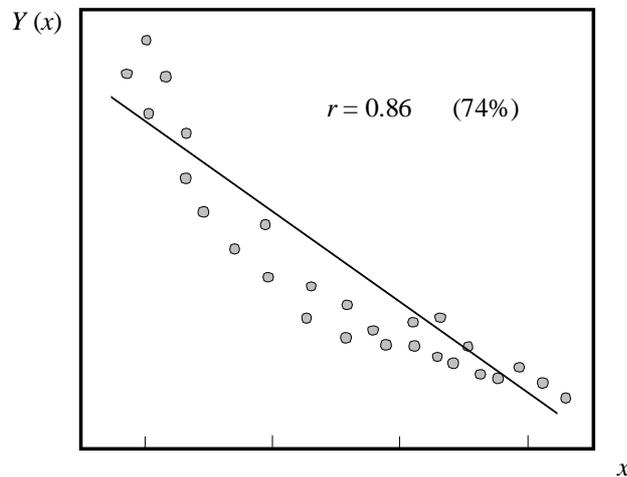
cumple esta importante misión.

El coeficiente solo puede variar entre los límites  $\pm 1$ , de forma tal que:

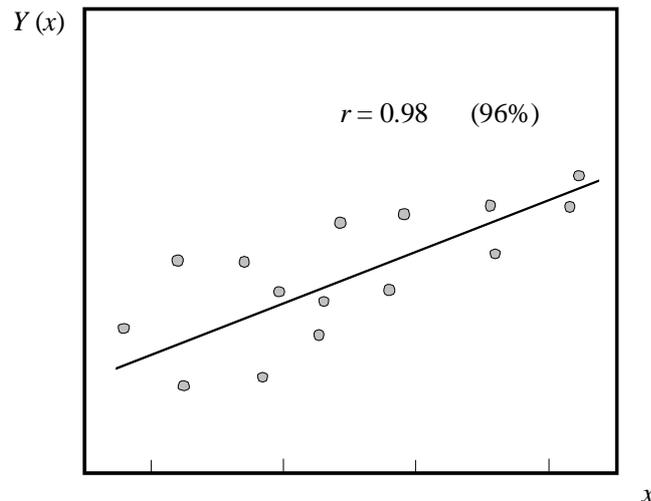
- Ajuste perfecto:  $|r| = 1$
- Ajuste nulo:  $r = 0$

Más aún, su cuadrado puede entenderse como el tanto por uno de efectividad del ajuste:

$$[ 5.38] \quad \text{Grado de ajuste (\%)} = 100 r^2$$



**Fig. 5.19 Ajuste con baja correlación e incluso con poca dispersión**



**Fig. 5.20 Ajuste con alta correlación e incluso con alta dispersión**

Podríamos pensar que un ajuste realizado con una efectividad del 74% (Fig. 5.19) puede ofrecernos grandes garantías de reproducir los datos experimentales, y sin embargo no es así. En efecto, el coeficiente de correlación solo da cuenta del grado en que los cambios de la variable  $x$  generan cambios proporcionales en la variable  $y$ , con independencia de la dispersión de puntos. Por tanto un grado de ajuste del 74%, supone también que hay un 26% de los cambios que no pueden ser explicados mediante regresión lineal. Tal vez en este caso, deberíamos elegir otro tipo de función para el ajuste. Si la recta fuera la elección correcta estaríamos en correlaciones muy próximas al 100%, con independencia de la mayor o menor dispersión de puntos (Fig. 5.20).

## 5.4 Interpolación

Para terminar este capítulo, haremos referencia a un problema que en cierto modo guarda relación con los procedimientos de correlación. Hasta ahora nos hemos centrado en el caso de disponer de un conjunto de datos experimentales, con la lógica dispersión inherente al proceso de medida, y que necesitamos ajustar a algún tipo de función matemática para alcanzar un doble objetivo: Por un lado facilitar el cálculo y la predicción de nuevos resultados, y por otro, obtener conclusiones acerca del error y el grado de confianza de los resultados.

En otras ocasiones nos enfrentamos a una circunstancia algo distinta: necesitamos conocer el valor de una función  $Y(x)$  en un punto  $x_0$  a partir de valores de la función en puntos próximos a éste, es decir, interpolar.

Pensemos por ejemplo que estamos estudiando un problema en el que necesitamos conocer la presión de vapor del agua. Podemos acudir a un socorrido *handbook* en el que sin duda encontraremos tabuladas las presiones de vapor del agua entre  $0^{\circ}\text{C}$  y  $100^{\circ}\text{C}$ . Tal vez encontremos que los incrementos de la tabla son de un grado centígrado, pero nuestro problema requiere que conozcamos la presión de vapor a  $86.4^{\circ}\text{C}$ . Son por tanto 100 puntos que deberíamos representar gráficamente, y tratar de ajustar a alguna función matemática. Sin duda un esfuerzo que pocas personas estarían dispuestas a asumir.

En otras ocasiones es una función matemática algo enrevesada la que nos plantea dificultades.

Por ejemplo tal vez estamos interesados en calcular la integral:

$$[ 5.39] \quad I(x) = \int_0^x e^{-z^2} dz$$

pero al no ser analítica, no tendremos resultados directos para cualquier valor de  $x$ . En la bibliografía encontraremos tablas de esta función, para una gran cantidad de valores de  $x$ . Por desgracia tal vez verifiquemos una vez más la indemostrable *Ley de Murphy*, y el valor que vamos buscando sea precisamente uno de los que faltan en la tabla.

Tanto en este caso, como en el anterior, la solución más sensata no será ajustar todo el conjunto de datos, ni resolver numéricamente la integral, sino tratar de *interpolar* el valor buscado a partir de datos conocidos, y que sean próximos a él.

El caso más simple de interpolación es aquel que utiliza solo dos puntos, uno de ellos situado algo por debajo del valor buscado y el otro algo por encima. Como es evidente, la función matemática de más bajo orden, que pasa por dos puntos, es la línea recta y si disponemos de una pareja de puntos  $(x_1, y_1)$  e  $(x_2, y_2)$ , tal vez estemos en condiciones de interpolar el valor  $Y(x_m)$ , donde  $x_m$  es una abscisa intermedia a  $x_1$  y  $x_2$ .

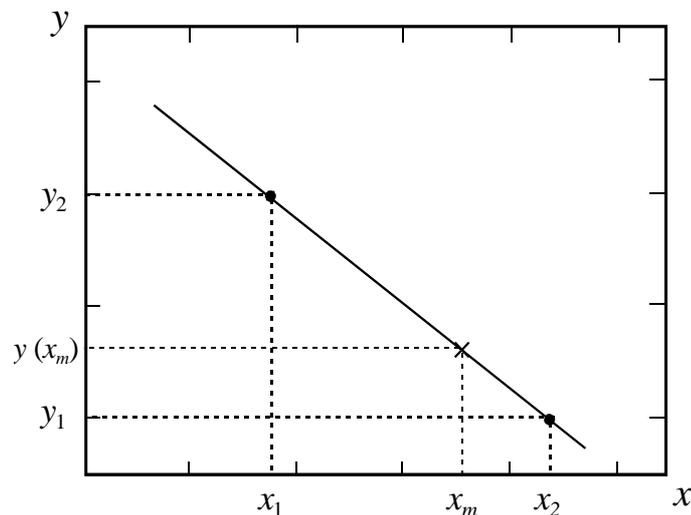


Fig. 5.21 Recta que interpola el valor entre dos puntos

Para ello tan solo tendremos que calcular la ecuación de dicha recta:

$$[ 5.40] \quad y(x) = \alpha + \beta x \quad \rightarrow \quad \left. \begin{array}{l} \alpha + \beta x_1 = y_1 \\ \alpha + \beta x_2 = y_2 \end{array} \right\} \rightarrow \quad \begin{array}{l} \alpha = \frac{(x_2 y_1 - x_1 y_2)}{(x_2 - y_1)} \\ \beta = \frac{(y_2 - y_1)}{x_2 - x_1} \end{array}$$

y aplicarla al punto de nuestro interés:

$$[ 5.41] \quad Y(x_m) \approx y(x_m) = \alpha + \beta x_m$$

El resultado obtenido seguramente se parecerá mucho al que obtendríamos si conociésemos el valor de la función en el punto  $x_m$ , pero no será idéntico. La razón es que estamos tratando de aproximar una función matemática, con una cierta curvatura, mediante una simple recta. Si el incremento de separación entre las ordenadas  $y_1$  e  $y_2$  no es muy grande, tal vez sea razonable aproximar este tramo mediante una recta. En caso contrario, el error cometido puede ser significativo.

La alternativa inmediata a los casos difíciles es emplear otro tipo de función interpoladora, tal vez un polinomio de orden 2. Sin embargo veremos que esta solución no es plenamente satisfactoria.

Para encontrar la ecuación de un polinomio de segundo orden, necesitamos conocer 3 puntos por los que pase dicha curva:

$$[5.42] \quad y(x_i) = \alpha + \beta x_i + \gamma x_i^2$$

y de esta forma plantear un sistema de ecuaciones cuya solución determine el valor de los coeficientes  $\alpha$ ,  $\beta$ ,  $\gamma$ .

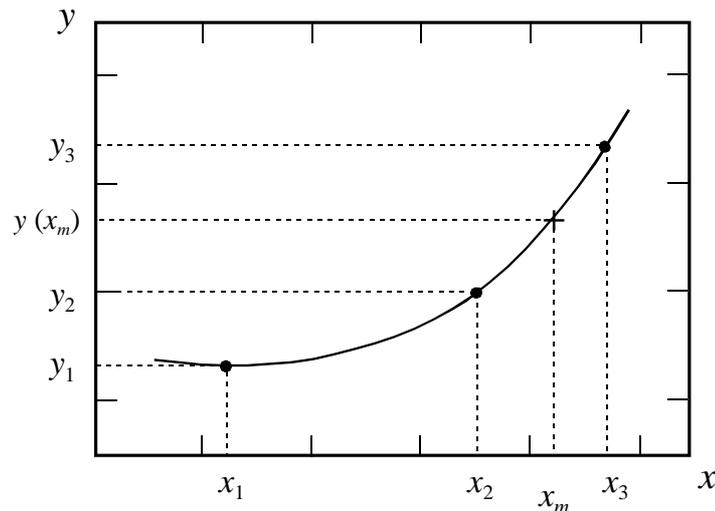


Fig. 5.22 Interpolación con un polinomio de segundo orden

Podemos comprobar en la figura que el nuevo valor que necesitamos interpolar  $x_m$ , estará situado necesariamente entre dos de los puntos, dejando el tercero en una situación más asimétrica. Si los puntos experimentales no pueden ser representados exactamente por un polinomio de segundo orden, esto se traducirá en una pérdida de la capacidad para interpolar correctamente el valor buscado. Consideremos ahora un polinomio de tercer grado:

$$[5.43] \quad y(x) = \alpha + \beta x + \gamma x^2 + \delta x^3$$

Para determinar sus coeficientes necesitamos un total de 4 puntos.

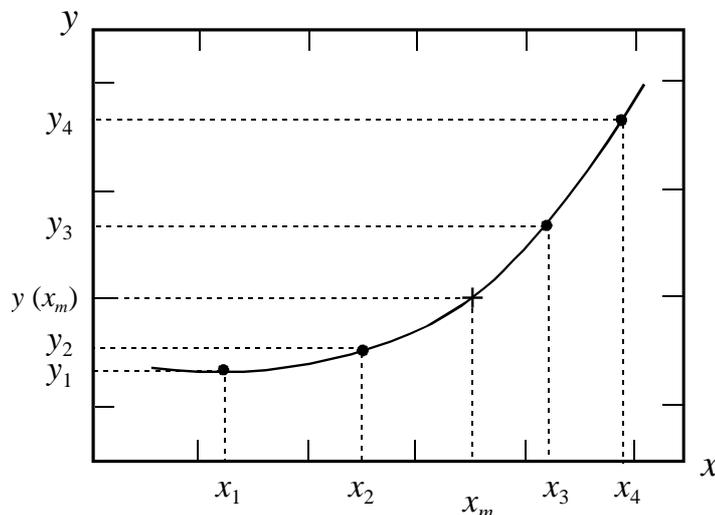


Fig. 5.23 Interpolación con un polinomio de tercer orden

De esta forma la coordenada sobre la que necesitamos interpolar, sí puede colocarse en una posición central con respecto a los *puntos base* de interpolación. Los resultados obtenidos serán así más próximos al valor exacto.

La conclusión final que debemos extraer es que las interpolaciones mediante polinomios deben llevarse a cabo, preferentemente, con polinomios de orden impar (1, 3, 5, etc), situando siempre el punto a interpolar en una posición central.

#### Ejemplo

Realizar una interpolación sobre la tabla de niveles de confianza de una distribución normal, y determinar cual es el factor z que corresponde a un nivel de confianza del 75%.

**Tabla 5.9 Niveles de confianza para una distribución normal**

Factor z	Nivel de confianza (%)	Factor z	Nivel de confianza (%)
0,67	50	1,96	95
1,00	68	2,00	96
1,29	80	2,58	99
1,64	90	3,00	99.7

Podemos comenzar por realizar una interpolación lineal. Para un nivel de confianza del 75%, los dos valores tabulados más próximos son: (68, 1,00) y (80, 1,29).

A partir de ellos podemos plantear el sistema:

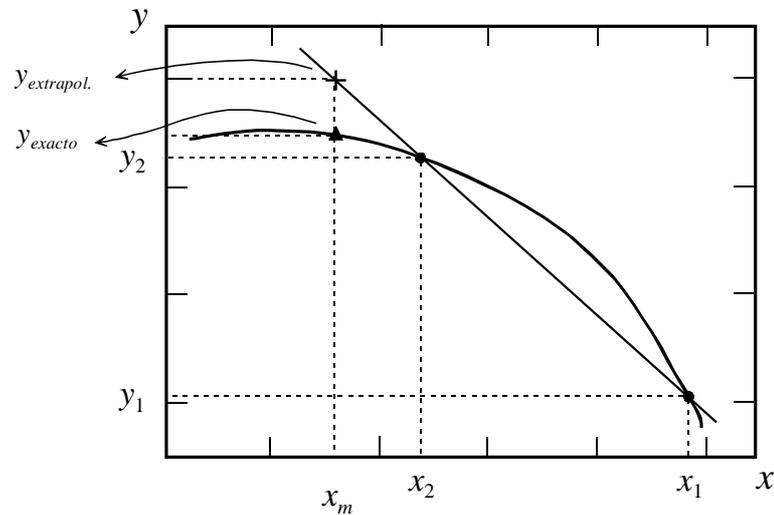
$$[5.44] \quad \left. \begin{array}{l} \alpha + \beta 68 = 1,00 \\ \alpha + \beta 80 = 1,29 \end{array} \right\} \rightarrow \begin{array}{l} \alpha = -0,6433 \\ \beta = 0,0242 \end{array} \rightarrow Y(75) = -0,6433 + 0,0242 \times 75 = 1,1717$$

Podemos también ensayar la interpolación mediante un polinomio de 3<sup>er</sup> orden. En este caso el sistema será:

$$[5.45] \quad \left. \begin{array}{l} \alpha + \beta 50 + \gamma 50^2 + \delta 50^3 = 0,67 \\ \alpha + \beta 68 + \gamma 68^2 + \delta 68^3 = 1,00 \\ \alpha + \beta 80 + \gamma 80^2 + \delta 80^3 = 1,29 \\ \alpha + \beta 90 + \gamma 90^2 + \delta 90^3 = 1,64 \end{array} \right\} \rightarrow Y(75) = 1,1558$$

La solución exacta a este problema es  $Y(x) = 1,15035$ , poniéndose de manifiesto el mejor grado de aproximación con el polinomio de tercer grado, aun cuando la interpolación lineal sea razonablemente próxima, y desde luego mucho más rápida de calcular.

Por último deberíamos mencionar un problema relacionado con los anteriores, la *extrapolación*. Ahora se trata de encontrar un valor que se encuentra próximo, pero fuera del intervalo de puntos conocidos. La siguiente figura puede ayudarnos a comprender mejor esta situación:



**Fig. 5.24** Gráfica explicativa del procedimiento de extrapolación

El procedimiento a seguir en una extrapolación es el mismo que en la interpolación, pero en general los resultados obtenidos serán menos precisos. La razón podemos observarla en la figura anterior. Al estar fuera del rango de los datos, desconocemos el hábito que describe la colección de puntos, y puede que la función de extrapolación nos sitúe en una posición alejada del verdadero valor.

En general, los resultados de una extrapolación deben ser manejados con extrema cautela, y solamente podremos considerarlos aceptables, si se conoce bien el hábito de la función, o si estamos en un punto muy próximo al extremo de los datos disponibles.

## 5.5 Cálculos Gráficos

Con mucha frecuencia el análisis de una colección de datos experimentales requiere la integración o derivación de los mismos.

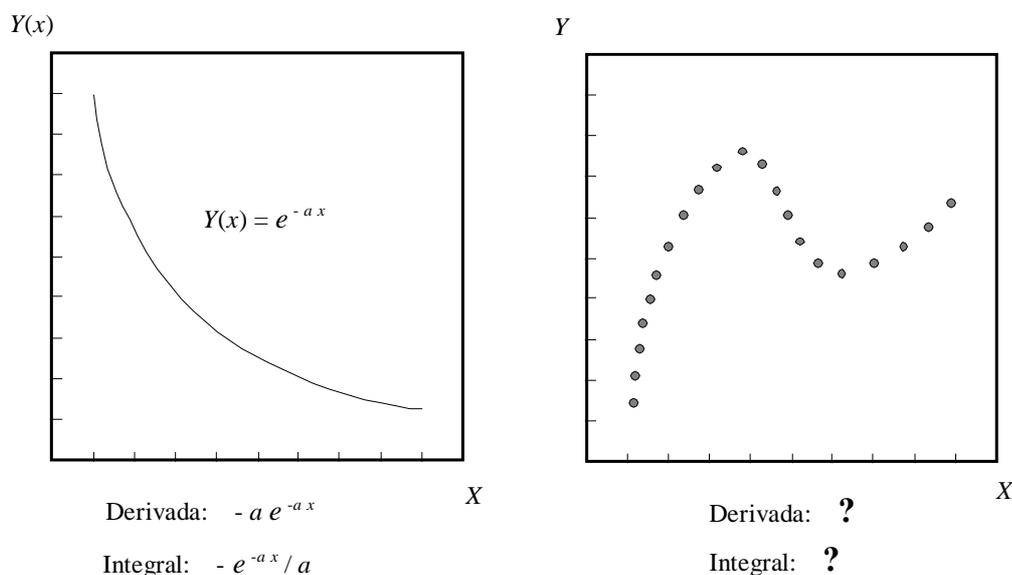
Por ejemplo, si medimos la masa  $m$  que hay depositada en distintos puntos  $x$  de un alambre, estaríamos en disposición de poder calcular el momento de inercia  $I$ , alrededor de su eje de giro. Para ello, al menos en principio, nos bastaría con calcular la integral:

$$[ 5.46] \quad I = \int_0^M x^2 dm$$

Parece también bastante natural que si disponemos de datos acerca de la concentración de un reactivo  $C_A$  a medida que progresa el tiempo de reacción  $t$ , podríamos calcular la velocidad de la reacción  $v_A$  mediante el simple cálculo de la derivada:

$$[ 5.47] \quad v_A = -\frac{dC_A}{dt}$$

Calcular la derivada o la integral de una sencilla función continua (Fig. 5.25a) puede resultar muy sencillo, pero es posible que se planteen mayores dificultades si lo que tiene es un conjunto de puntos aislados (Fig. 5.25b).

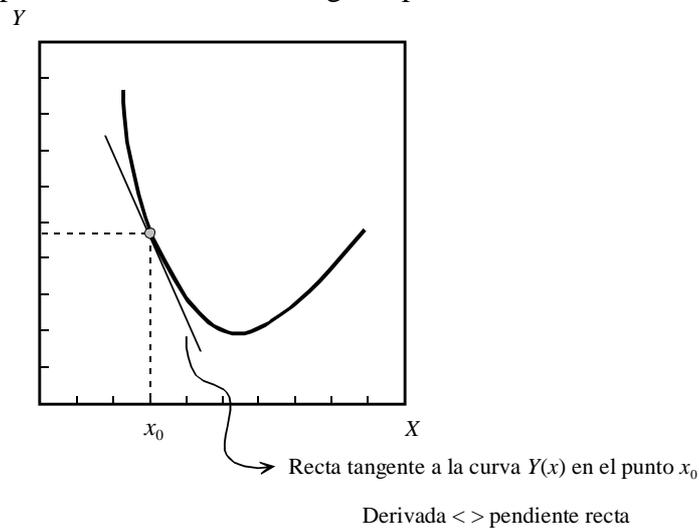


**Fig. 5.25 Derivada e integral obtenidas de una función analítica, planteamiento del problema en el caso de un conjunto de puntos aislados**

Puesto que en los estudios experimentales los resultados obtenidos suelen consistir en una colección de datos discretos, resultará de máximo interés para nosotros disponer de procedimientos que nos permitan calcular derivadas, integrales, e incluso otro tipo de operaciones sobre estos datos.

### 5.5.1 Interpretación Gráfica de la Derivada

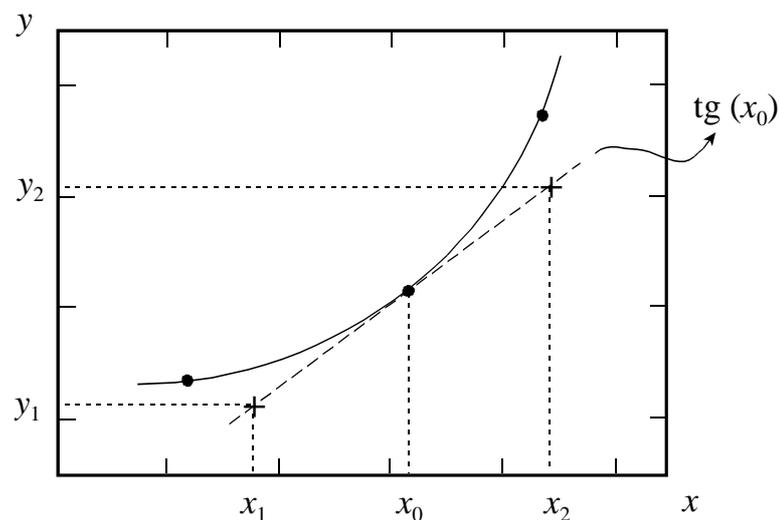
La derivada de una función en un punto, coincide con la pendiente de la recta tangente a la función en dicho punto. Parece un trabalenguas, pero en realidad es un concepto muy simple.



**Fig. 5.26 Significado geométrico de la derivada**

Si disponemos de una colección de datos experimentales tales que al ser representados proporcionen un hábito coherente, podemos calcular de forma gráfica, y algo aproximada, el valor de la derivada en uno de los puntos. Para ello basta recordar el significado que hemos dado al concepto de la derivada.

Situados en la representación gráfica, sobre el punto en que queremos calcular la derivada (llamémosle  $x_0$ ), trazaremos una recta tangente a la curva que pasa por los puntos experimentales. La recta tangente debe pasar necesariamente por el punto  $x_0$ .



**Fig. 5.27 Recta tangente a la curva que pasa por los puntos experimentales en el punto  $x_0$**

Naturalmente este procedimiento carece de una gran exactitud, y solo debemos entenderlo como una estimación aproximada. Basta darse cuenta que dos personas distintas probablemente trazarán rectas algo diferentes, pero que cada uno de ellos juzgará como tangente a la curva en dicho punto.

En cualquier caso, una vez que disponemos de la recta tangente, elegiremos una pareja de puntos situados sobre la misma recta y leeremos sus coordenadas. En nuestra figura se corresponden con los pares  $(x_1, y_1)$  e  $(x_2, y_2)$ .

Podemos ahora calcular la ecuación de la recta tangente:

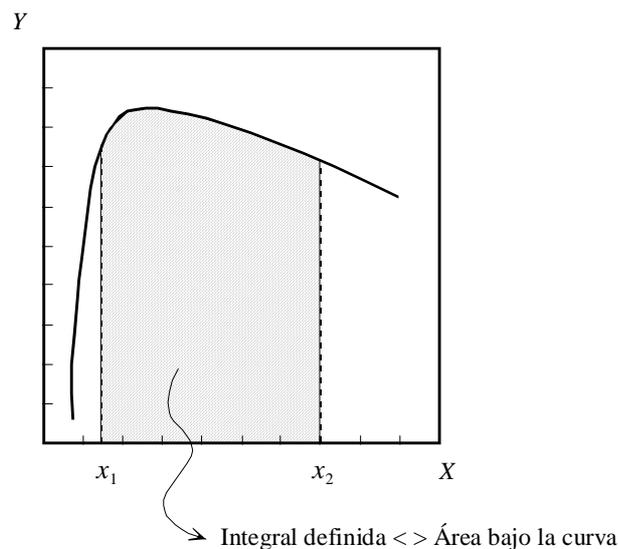
$$[ 5.48] \quad y = mx + c$$

y más concretamente, su pendiente  $m$ , que como ya hemos mencionado, coincide con el valor de la derivada de la función  $y(x)$  en el punto  $x_0$ :

$$[ 5.49] \quad m = \frac{(y_2 - y_1)}{(x_2 - x_1)} \equiv \left. \frac{dy(x)}{dx} \right]_{x=x_0}$$

### 5.5.2 Interpretación Gráfica de la Integral

La integral de una función  $y = F(x)$  a lo largo de un cierto intervalo cerrado, coincide con el área bajo dicha función y el eje de abscisas colocado éste en la posición 0 del eje de ordenadas.



**Fig. 5.28 Significado geométrico de la integral**

Calcular algunas propiedades mecánicas de un sistema móvil, como por ejemplo el trabajo, puede resultar bastante fácil. Podría bastar con hacer uso de la definición:

$$[ 5.50] \quad W = \int_a^b F(x)dx$$

El problema surge si la fuerza aplicada sobre el móvil, va cambiando de una forma complicada, o no bien conocida, a medida que este se desplaza. Podríamos solucionar este problema si realizamos medidas experimentales en las que determinemos la fuerza aplicada sobre el móvil en determinados puntos del recorrido. Tendríamos así un conjunto de  $n$  datos numéricos que podrían ser representados en una gráfica como la que aparece a continuación:

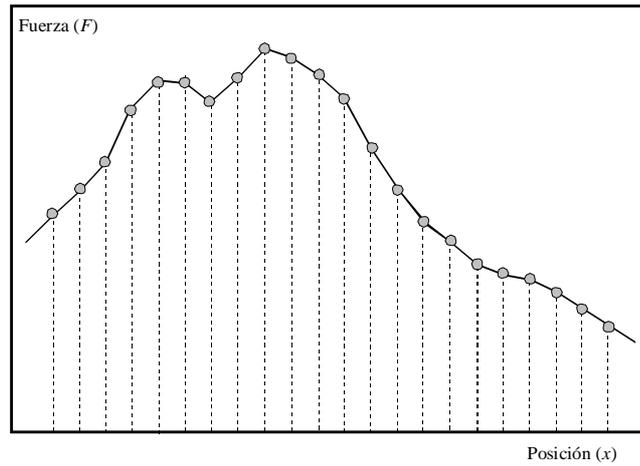


Fig. 5.29 Fuerza aplicada sobre un móvil en función de la posición

La solución está ahora más cercana: puesto que el trabajo es la integral de la curva  $F(x)$  a lo largo del recorrido, bastará con calcular el área bajo la curva, ¿pero cómo es posible este cálculo?.

Antes de la llegada de las calculadoras electrónicas, el cálculo del área bajo la curva se convertía en una cuestión que requería buenas dosis de paciencia y habilidad. Básicamente había dos procedimientos:

- Tomar una regla y un lápiz, para descomponer pacientemente el área engendrada por la curva en varias figuras regulares, cuyo área es conocida, y posteriormente sumar todas las áreas parciales.

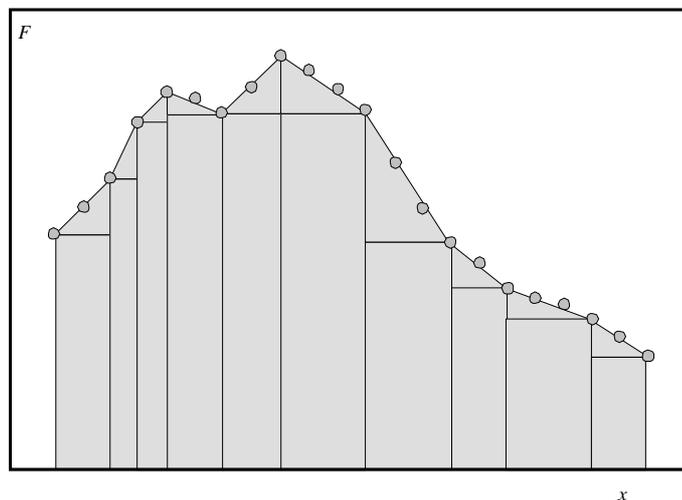
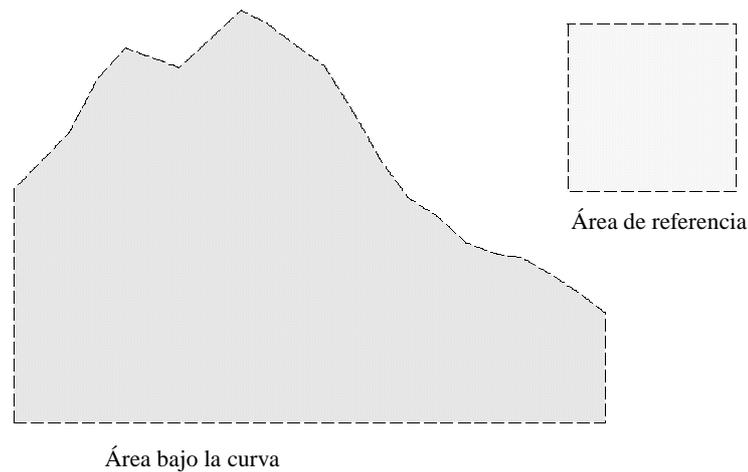


Fig. 5.30 Descomposición de una función  $F$  en figuras regulares de área conocida

- El segundo procedimiento es aún más sorprendente. Recortar cuidadosamente el trozo de papel que corresponde al área bajo la curva y pesarlo en una balanza de precisión. Posteriormente recortamos, sobre la misma hoja de papel que hemos utilizado en la representación, otro área bien conocida (pongamos por caso un cuadrado), y pesamos también este segundo trozo de papel. Como la densidad y espesor del papel es aproximadamente igual en los dos casos, las masas ( $M$ ) y las áreas ( $S$ ) de ambas figuras serán proporcionales.



**Fig. 5.31 Trozos de papel de forma igual al área bajo la curva y a un área de referencia**

$$[ 5.51] \quad \frac{M_{exp}}{S_{exp}} = \frac{M_{ref}}{S_{ref}}$$

$$[ 5.52] \quad S_{exp} = S_{ref} \frac{M_{exp}}{M_{ref}}$$

## 5.6 Métodos Numéricos

Métodos ingeniosos, ¿verdad?, pero poco prácticos y limitados. Los cálculos basados en gráficos no pueden ofrecer una gran confianza en la elaboración de resultados. Más aún, si tenemos que calcular, por ejemplo la derivada, en cada uno de una larga lista de puntos experimentales, ni siquiera será un proceso rápido.

Más aún, volvamos al ejemplo del móvil y consideremos otra situación bastante habitual. Tal vez sea posible conocer de forma explícita la fuerza que actúa sobre el móvil por lo que podría parecer que el cálculo del trabajo resulta inmediato, pero supongamos que el móvil se mueve gracias a la acción de una fuerza oscilatoria un tanto peculiar:

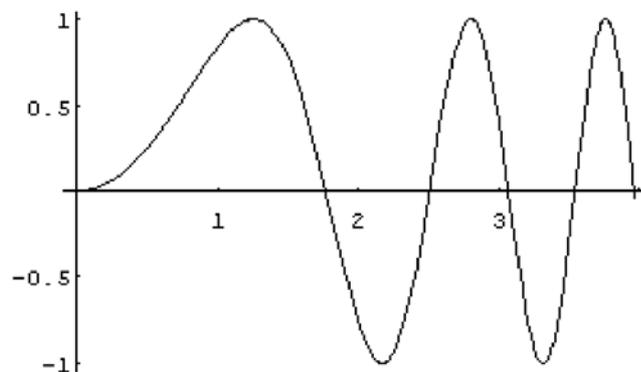


Fig. 5.32 Fuerza oscilatoria en función de la posición

[ 5.53] 
$$F(x) = \text{sen}(x^2) \rightarrow W = \int_a^b \text{sen}(x^2) dx$$

Esta integral tiene un aspecto muy inocente, pero en realidad no es posible obtener su primitiva de forma analítica (no se confunda la función  $\text{sen } x^2$ , sin primitiva, con  $\text{sen}^2 x$ , que sí la tiene).

¿Qué podemos hacer en este caso?, o incluso en cualquier otro problema de cálculo en el que tengamos que resolver ecuaciones complicadas, derivar, integrar o resolver ecuaciones diferenciales. Sencillamente recurrir a los potentes procedimientos que nos ofrece el cálculo numérico o incluso a programas de ordenador donde este tipo de algoritmos ya están automatizados.

Hoy en día disponemos de herramientas numéricas muy eficaces y fácilmente programables en un ordenador. Son los llamados *métodos numéricos*, *cálculo con variables discretas* o *métodos de cálculo con elementos finitos*, de los cuales la integración y la derivación son solo dos ejemplos concretos. Los métodos numéricos se han convertido en una herramienta fundamental para la ciencia e ingeniería, donde el estudio de modelos a escala y la realización de simulaciones por ordenador, se han hecho imprescindibles.

A continuación presentaremos solo algunas ideas referentes a la integración y derivación de conjuntos de datos numéricos, lo cual nos permitirá resolver muchos problemas frecuentes en un laboratorio.

### 5.6.1 Cálculo Numérico de la Derivada

Sin necesidad de entrar en grandes desarrollos matemáticos, que nos alejarían de nuestro objetivo, bastaría con presentar algunos procedimientos para calcular la derivada numérica de un conjunto de puntos discretos (o de una función continua de la que tomemos solo valores concretos).

Si representamos las coordenadas de cada uno de los puntos experimentales por  $(x_i, f_i)$ , podremos calcular la derivada en un punto concreto, llamémosle  $(x_0, f_0)$ , si conocemos las coordenadas de un punto adyacente.

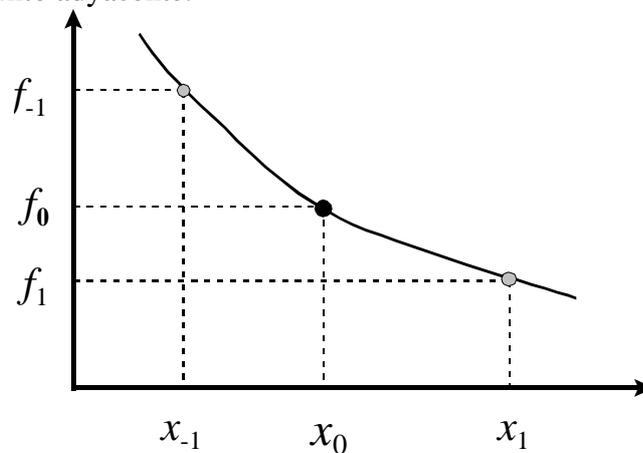


Fig. 5.33 Puntos de una curva que nos servirán para calcular la derivada en  $(x_0, f_0)$

En efecto, la derivada primera se puede calcular como:

$$[5.54] \quad \frac{df(x_0)}{dx} \approx \frac{\Delta f(x_0)}{\Delta x} = \frac{f(x_1) - f(x_0)}{\Delta x} = \frac{f(x_0 + h) - f(x_0)}{h}$$

siendo  $h$  la distancia de separación entre las abscisas de ambos puntos, a la que frecuentemente se la denomina *paso de derivación*.

$$[5.55] \quad h = x_1 - x_0$$

Igualmente podríamos haber calculado la derivada haciendo uso del punto inmediatamente anterior al requerido. En este caso la derivada será

$$[5.56] \quad \frac{df(x_0)}{dx} \approx \frac{\Delta f(x_0)}{\Delta x} = \frac{f(x_{-1}) - f(x_0)}{\Delta x} = \frac{f(x_0 - h) - f(x_0)}{(-h)}$$

Como es fácil comprobar, el origen de este procedimiento numérico no esconde grandes misterios. En realidad bastaría con calcular el límite de cocientes incrementales para alcanzar la verdadera definición de derivada

$$[5.57] \quad \frac{df(x_0)}{dx} \approx \lim_{h \rightarrow 0} \frac{\Delta f(x_0)}{\Delta x}$$

De esta forma, justificamos también de una manera natural, el requerimiento básico para poder calcular una derivada numérica con suficiente exactitud, y es que la distancia de

separación entre las abscisas de dos medidas experimentales adyacentes, **debe ser muy pequeña**. En caso contrario los errores cometidos en este cálculo, tienden a ser muy elevados.

El algoritmo que acabamos de presentar se denomina genéricamente de *un solo punto*, pero es posible desarrollar otros procedimientos más elaborados, y mucho más exactos para calcular las derivadas. Sin entrar en más profundidades, un algoritmo simétrico de 4 puntos, nos permitiría calcular la derivada primera según:

$$[5.58] \quad \frac{df(x_0)}{dx} \approx \frac{2f(x_0+2h) + f(x_0+h) - f(x_0-h) - 2f(x_0-2h)}{10h}$$

Como puede observarse, para calcular la derivada en el punto  $(x_0, f_0)$ , se requieren las coordenadas de dos puntos por encima y dos puntos por debajo del indicado. El esfuerzo de cálculo es algo mayor, pero la calidad de los resultados mejora notablemente.

En muchas ocasiones, la obtención de resultados a partir de datos experimentales puede requerir el cálculo de derivadas de orden superior. Como una simple extensión de los procedimientos ya descritos, podremos terminar el epígrafe señalando que la derivada segunda puede calcularse según:

$$[5.59] \quad \frac{d^2f(x_0)}{dx^2} \approx \frac{f(x_0+2h) - 2f(x_0+h) + f(x_0)}{h^2}$$

### 5.6.2 Métodos Numéricos de Integración

Como ya hemos mencionado anteriormente, si nos enfrentamos a la necesidad de calcular una integral definida:

$$[5.60] \quad I = \int_a^b f(x)dx$$

en el que la función  $f(x)$  es complicada, o lo que es más frecuente en un laboratorio, solo conocemos el valor de la función en ciertos puntos concretos, la integración solo podrá realizarse mediante un *método numérico*.

Son varios los procedimientos disponibles para realizar esta integración (Simpson, interpolación polinómica, cuadratura, Euler, Newton-Cotes, Monte Carlo, etc), cada uno de ellos con sus ventajas y dificultades, pero todos bien descritos en los libros especializados. A continuación describiremos solamente la forma de trabajar con alguno de ellos.

## 1. Método de Euler

Consideremos un conjunto de puntos  $(x_i, f_i)$  que representan respectivamente los valores de abscisas y ordenadas de la función a integrar. Una forma muy simple de estimar el valor del área encerrada por la curva continua que pasaría por esos puntos, sería aproximarla por una suave curva escalonada, tal y como vemos en la figura:

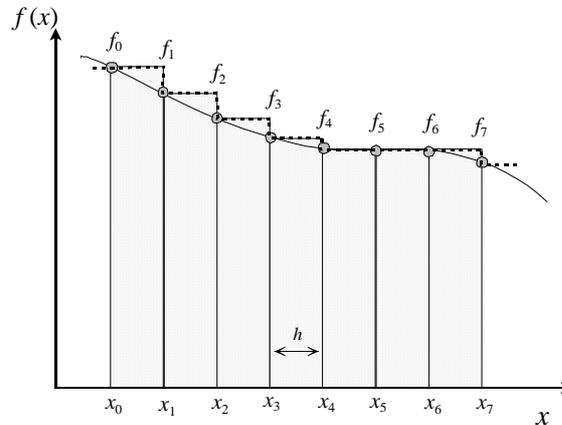


Fig. 5.34 Aproximación de una función a una curva escalonada

Integrar esta función escalonada ya resulta muy sencillo, puesto que será la suma de las áreas de cada uno de los rectángulos:

$$[5.61] \quad I = \int_a^b f(x) dx \approx (x_1 - x_0)f_0 + (x_2 - x_1)f_1 + \dots + (x_7 - x_6)f_6$$

o bien si todas las abscisas se encuentran igualmente espaciadas por una distancia  $h$ , podremos expresar la integral en forma de sumatorio compacto

$$[5.62] \quad I = \int_{x_0}^{x_n} f(x) dx \approx h \sum_{i=0}^{n-1} f_i$$

Observamos pues la sencillez de cálculo del valor de una integral definida a través del método de Euler, aunque debería ser también patente la poca exactitud que supone aproximar la función mediante una curva escalonada. Veamos un ejemplo a este respecto.

Ejemplo1

La integral:

$$[5.63] \quad I = \int_0^1 e^{-x} dx = 0,6321$$

es sin duda inmediata de calcular y nos servirá para ilustrar la bondad del método de Euler. Podemos elegir un paso de integración  $h = 0,2$ , es decir trabajar con 5 puntos, y construir la siguiente tabla:

Tabla 5.10 Valores de la función  $e^{-x}$

$f(x_i)$	1,0	0,8187	0,6703	0,5488	0,4493
$x_i$	0,0	0,2	0,4	0,6	0,8

de forma tal que el valor estimado para la integral será:

$$[5.64] \quad I \approx h \sum_{i=0}^5 f_i = 0,2(1 + 0,8187 + 0,6703 + \dots + 0,4493) = 0,6974$$

Observamos así que el resultado está sobrestimado con respecto al valor correcto de la integral, algo que no debería sorprendernos a la vista del tipo de aproximación que plantea el método de Euler. Podemos en cualquier caso tratar de mejorar los resultados utilizando un paso de integración más pequeño (y por tanto un escalonamiento más suave). Probemos ahora con  $h = 0,05$ , lo que genera 20 puntos de integración.

$$[ 5.65] \quad I \approx h \sum_{i=0}^{19} f_i = 0,6481$$

resultado que sigue siendo casi un 3% superior al valor correcto, y todo ello pese al esfuerzo de calcular y sumar 20 términos.

Las conclusiones que podemos extraer con respecto al método de integración de Euler es que se trata en un algoritmo muy simple, y que los resultados pueden no ser demasiado satisfactorios. De hecho hemos probado el algoritmo con una exponencial, que es una función muy *suave*. Si la función tiene cambios de pendiente más pronunciados, los errores se incrementarían. Veamos el siguiente ejemplo:

### Ejemplo 2

Las tres integrales siguientes han sido calculadas mediante un algoritmo de Euler que utiliza un paso de integración  $h = 0,5$ , comparándose el resultado obtenido con su valor analítico.

$$[ 5.66] \quad \int_2^4 \frac{1}{(x+3)^{3/2}} dx = 0,1385, \quad I_{Euler} = 0,1477, \quad Error = 6,6\%$$

$$[ 5.67] \quad \int_0^6 e^{-x} \text{sen}(x) dx = 0,4992, \quad I_{Euler} = 0,4784, \quad Error = 4,1\%$$

$$[ 5.68] \quad \int_0^3 \text{sen}(x) \cos(2x) dx = -0,6765, \quad I_{Euler} = -0,7516, \quad Error = 11,1\%$$

## 2. Método del Trapecio

A la vista de que el anterior procedimiento de integración, puede llevarnos a cometer importantes errores, cabría reflexionar acerca de los puntos débiles del procedimiento empleado.

Sin duda, la aproximación de la curva real por una curva escalonada, es la principal causa del comportamiento erróneo, máxime cuando consideramos pasos de integración grandes (con relación al intervalo total de integración) o cuando la curva no tiene un hábito muy suave (fuertes cambios de pendiente).

Podríamos así tratar de aproximar la curva real por elementos geométricos más próximos a ella, y que al mismo tiempo nos permitan calcular fácilmente su área. En este sentido, el siguiente grado de aproximación consistiría en utilizar trapecios en lugar de rectángulos.

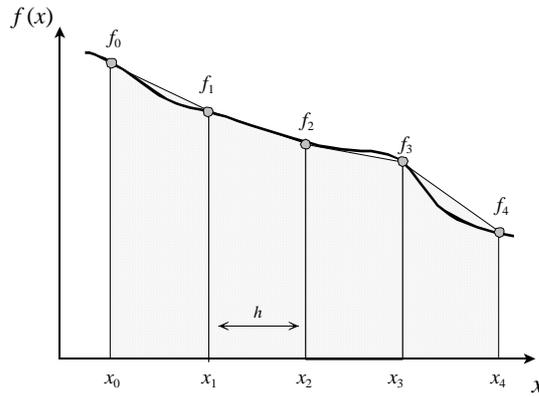


Fig. 5.35 Aproximación de una función a curvas en forma de trapecio

El área de cada trapecio es, como se observa en la Fig. 5.36:

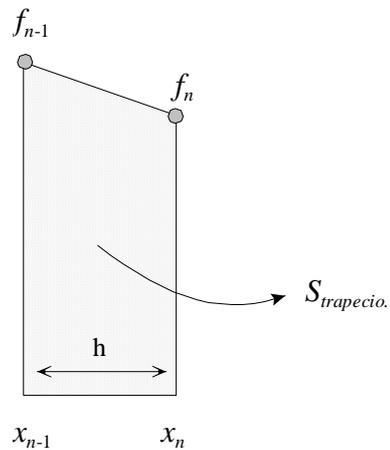


Fig. 5.36 Dimensiones de un trapecio entre dos ordenadas consecutivas

$$[ 5.69] \quad S_{trapecio} = \frac{h}{2}(f_{n-1} + f_n)$$

y finalmente la integral puede ser calculada mediante:

$$[ 5.70] \quad I = \int_a^b f(x) dx \approx \frac{h}{2}(f_0 + f_1) + \frac{h}{2}(f_1 + f_2) + \dots + \frac{h}{2}(f_3 + f_4)$$

o de forma más general:

$$[ 5.71] \quad I = \int_{x_0}^{x_n} f(x) dx \approx \frac{h}{2} \sum_{i=0}^{n-1} (f_i + f_{i+1})$$

## Ejemplo 3

En el ejemplo anterior se había estimado el valor de tres integrales de prueba. Los resultados obtenidos con la *regla del trapecio*, utilizando el mismo paso de integración  $h = 0,5$ , son:

$$[ 5.72] \quad \int_2^4 \frac{1}{(x+3)^{3/2}} dx = 0,1385, \quad I_{Trap} = 0,1388, \quad Error = 0,2\%$$

$$[ 5.73] \quad \int_0^6 e^{-x} \text{sen}(x) dx = 0,4992, \quad I_{Trap} = 0,4784, \quad Error = 4,1\%$$

$$[ 5.74] \quad \int_0^3 \text{sen}(x) \cos(2x) dx = -0,6765, \quad I_{Trap} = -0,7177, \quad Error = 6,1\%$$

### 3. Método de Simpson

Si aún quisiéramos mejorar los resultados proporcionados por los anteriores métodos de integración, tendríamos que trabajar con una aproximación de mayor orden sobre la curva experimental. El método de Euler consistía en una aproximación puntual, el método del trapecio en una aproximación lineal, por tanto el siguiente paso lógico consiste en probar una aproximación polinómica de orden 2, es decir una parábola.

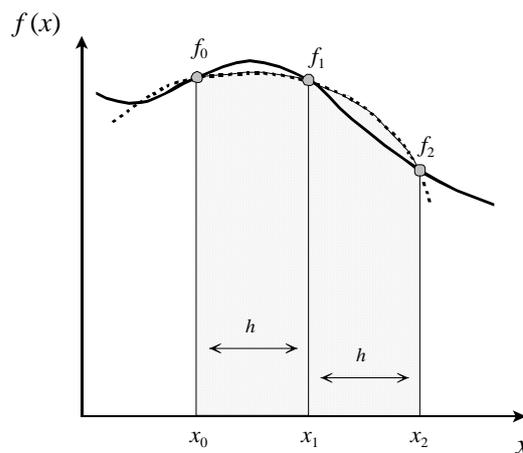


Fig. 5.37 Aproximación de una función a curvas en forma de parábola

En este caso, el conjunto de figuras en que descomponemos el área total bajo la curva, ya no son tan regulares como en los casos anteriores, puesto que su límite superior es un segmento de parábola en lugar de una línea recta. Aún así puede encontrarse una forma analítica de calcular el área de estos elementos.

Obviaremos en este escrito el desarrollo matemático formal y remitimos al lector a la bibliografía más especializada, pero no resulta difícil demostrar que el área encerrada bajo una curva  $f(x)$ , en un intervalo de valor  $2h$ , siendo de nuevo  $h$  el paso de integración, es igual a:

$$[ 5.75] \quad I = \int_{x_0}^{x_2} f(x) dx \approx \frac{h}{3} (f_0 + 4f_1 + f_2)$$

donde la única restricción importante que debemos imponer es que el paso de integración elegido, sea tal que podamos dividir el intervalo de integración en un número par de subintervalos.

$$[5.76] \quad (x_n - x_0) = nh, \quad n = \frac{(x_n - x_0)}{h} \quad (\text{par})$$

La fórmula de integración de Simpson, para un intervalo completo puede escribirse cómo:

$$[5.77] \quad I = \int_{x_0}^{x_n} f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + f_2) + \frac{h}{3}(f_2 + 4f_3 + f_4) + \dots + \frac{h}{3}(f_{n-2} + 4f_{n-1} + f_n)$$

#### Ejemplo 4

Vamos a obtener la aproximación de Simpson para la integral:

$$[5.78] \quad I = \int_0^6 \frac{x^2}{x^3 + 8} dx = 1,1107$$

utilizando un paso de integración  $h=1$ . Puesto que la longitud del intervalo de integración es igual a 6, con este valor de  $h$  nos estamos asegurando un número par de subintervalos, tal y como requiere el método de Simpson.

Para ello comenzamos por construir la tabla de valores necesarios para la integración:

**Tabla 5.11 Tabla de valores de la función que se integra**

$f(x_i)$	0,0	0,1111	0,2500	0,2571	0,2222	0,1880	0,1607
$x_i$	0,0	1,0	2,0	3,0	4,0	5,0	6,0

Pudiendo ahora calcular la integral según:

$$[5.79] \quad I \approx \frac{h}{3}(f_0 + 4f_1 + f_2) + \frac{h}{3}(f_2 + 4f_3 + f_4) + \frac{h}{3}(f_4 + 4f_5 + f_6) =$$

$$= \frac{h}{3}(f_0 + f_6) + \frac{h}{3}(4f_1 + 2f_2 + 4f_3 + 2f_4 + 4f_5) = 1,1100$$

## 4. Consideraciones Finales

Disponemos ya de los elementos necesarios para poder calcular de forma numérica la integral definida de funciones analíticas, o lo que es más interesante para nuestros fines, la integral sobre una colección de datos discretos, que probablemente provengan de un estudio experimental.

De los diferentes algoritmos de integración ilustrados, ¿cual es el más conveniente?. Como puede intuir el lector la respuesta no puede ser única, ya que cada situación concreta aconsejará la utilización de uno de los métodos.

A la hora de alcanzar una buena operatividad de trabajo, debemos tratar de alcanzar un equilibrio entre rapidez de cálculo y exactitud de los resultados.

Se podría demostrar que los errores inherentes a los distintos procedimientos descritos son proporcionales al valor de ciertas derivadas de la función, calculadas en un punto indeterminado del intervalo. Más en concreto:

Método de Euler:  $Error \approx h^2 f'(z)$

Método del trapecio:  $Error \approx h^2 f''(z)$

Método de Simpsom:  $Error \approx h^2 f'''(z)$

siendo  $z$ , un valor indeterminado, aunque comprendido dentro el intervalo  $(x_0, x_n)$ . De esta forma vemos que el método de Euler solo puede proporcionar resultados exactos cuando la función a integrar es una constante, que el método del Trapecio solo es exacto cuando la función es lineal y que el método de Simpson proporciona resultados exactos, incluso para polinomios de orden 3. Sin embargo esto difícilmente nos ayudará si disponemos de un conjunto de datos experimentales. Más aún, no debemos confundir el error del método de integración con el error en la determinación experimental de una magnitud.

Volvamos de nuevo al ejemplo inicial de este epígrafe. Si estamos interesados en calcular experimentalmente el trabajo realizado por un móvil, el error en la determinación final será la suma de las imprecisiones debidas a cada dato experimental más el error correspondiente al método de integración

### Ejemplo

Si una magnitud experimental  $F$ , ha sido calculada mediante integración con el método del trapecio, de unos datos experimentales  $(x_i, f_i)$  que a su vez están sujetos a error, la estimación del error total debería hacerse de la siguiente manera:

$$[5.80] \quad F = \int_{x_0}^{x_4} f(x) \approx \frac{h}{2} (f_0 + f_1 + f_1 + f_2 + f_2 + f_3 + f_3 + f_4) = \frac{h}{2} (f_0 + 2f_1 + 2f_2 + 2f_3 + f_4) = \frac{h}{2} S$$

El error relativo para  $F$  será:

$$[5.81] \quad \frac{\Delta F}{F} = \frac{\Delta k}{k} + \frac{\Delta S}{S}$$

donde:

$$[5.82] \quad \frac{\Delta S}{S} = \frac{\Delta f_0 + 2\Delta f_1 + 2\Delta f_2 + 2\Delta f_3 + \Delta f_4}{f_0 + 2f_1 + 2f_2 + 2f_3 + f_4}$$

si todas las medidas  $f_i$  han sido realizadas en iguales circunstancias, pueden tener el mismo error absoluto, con lo cual:

$$[5.83] \quad \frac{\Delta S}{S} = \frac{8\Delta f_0}{f_0 + 2f_1 + 2f_2 + 2f_3 + f_4}$$

y finalmente, el error absoluto para  $F$ , será:

$$[5.84] \quad \Delta F = F \left( \frac{\Delta h}{h} + \frac{\Delta S}{S} \right) + E_{Trap}$$

siendo  $E_{Trap}$  el error cometido en el procedimiento de integración:

$$[5.85] \quad E_{Trap} \approx h^3 f''(z), \quad z \in (0, 4)$$

No resulta sencillo calcular exactamente este último término, ya que habría que calcular la derivada segunda en cada punto del intervalo de integración y elegir el mayor de estos valores (sería la estimación del máximo error posible).

Como procedimiento aproximado podemos utilizar:

$$[5.86] \quad f''(z) \approx \frac{f(z)}{h^2}, \quad E_{\text{trap}} \approx h f(z) = h \bar{f}(x) = \frac{h}{5} \sum_{i=0}^4 f_i$$

La consecuencia práctica que debemos extraer es que todos los procedimientos experimentales basados en la integración de datos, e igualmente podríamos decir para la derivación, son inherentemente más imprecisos que aquellos procedimientos en los que hay una medida directa de la magnitud. Aún así, en muchas ocasiones resulta imposible prescindir de este tipo de procedimientos y hay que tratar de minimizar el error.

Por tanto, salvo que estemos estimando solamente el orden de magnitud de una medida experimental, tal vez deberíamos tratar de trabajar siempre con el algoritmo de mayor orden que nos permitan nuestros datos experimentales.

En este sentido aún nos faltaría por decir que los procedimientos de integración de datos discretos no se acaban en la regla de Simpson. Son numerosos los métodos de integración de orden superior que podemos encontrar descritos en la bibliografía, a la que remitimos al lector con mayor inquietud en este campo. Sin duda en los métodos numéricos encontrará un extenso campo donde invertir su tiempo.