HERRAMIENTAS MATEMÁTICAS DE LA TERMODINÁMICA

TOMÁS S. GRIGERA

Instituto de Física de Líquidos y Sistemas Biológicos (IFLYSIB), CONICET y Universidad Nacional de La Plata, Calle 59 no. 789, B1900BTE La Plata

Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Calle 49 y 115, c.c. 67, B1900 La Plata

Resumen. Se resumen conceptos matemáticos utilizados en el desarrollo teórico de la Termodinámica, cuyo dominio es necesario para poder seguir un curso de Termodinámica más allá del nivel introductorio. Luego de una revisión de resultados básicos de la teoría de funciones de muchas variables, se trata sobre el diferencial, las funciones implícitas, funciones homogéneas y transformada de Legendre.

Las herramientas matemáticas necesarias para desarrollar la teoría termodinámica son esencialmente la teoría de funciones reales de muchas variables y las transformadas de Legendre. Expondremos aquí un resumen de los resultados más importantes desde el punto de vista a de su aplicación a la termodinámica, en la mayor parte de los casos sin demostración. Para más detalles sobre la teoría se puede consultar algún texto universitario de análisis matemático, como Piskunov (1977, Cap. VIII) o Rey Pastor, Pi Calleja y Trejo (1963).

1. Resultados básicos

Consideramos funciones reales de más de una variable $f(\mathbf{x}) : D_f \longrightarrow \mathbb{R}$, donde el dominio D_f es algún subconjunto de $\mathbb{R}^n$. Para denotar las componentes de $\mathbf{x}$ escriberemos $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, o bien, cuando sea conveniente por brevedad y claridad, x, y, z en lugar de x_1, x_2, x_3.

1.1. Continuidad. Se dice que $f(\mathbf{x})$ es *continua* en $\mathbf{x}_0$ si y sólo si

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0). \tag{1}$$

Date: Agosto de 2024.

En virtud de la definición de límite, esto equivale a decir que $f(\mathbf{x})$ es continua en $\mathbf{x}_0$ si y sólo sí para cualquier $\epsilon > 0$ existe un $\delta > 0$ tal que $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon$ cuando $|\mathbf{x} - \mathbf{x}_0| < \delta$ y $\mathbf{x} \in D_f$[1].

1.2. Derivada direccional y derivadas parciales. Dado un punto $\mathbf{x} \in D_f$ y una dirección dada por el vector unitario $\breve{\mathbf{n}}$, la *derivada direccional* en $\mathbf{x}$ es

$$\left.\frac{\partial f}{\partial \breve{\mathbf{n}}}\right|_{\mathbf{x}} = \lim_{h\to 0} \frac{f(\mathbf{x} + h\breve{\mathbf{n}}) - f(\mathbf{x})}{h}. \tag{2}$$

Dicho de otro modo, $\partial f/\partial \breve{\mathbf{n}}$ es la derivada de la función $k(h) = f(\mathbf{x}+h\breve{\mathbf{n}})$ evaluada en $h = 0$. La existencia de la derivada direccional *no implica* continuidad de la función (aunque exista en todas direcciones).

En general son de especial interés las derivadas en las direcciones de los ejes coordenados ($\breve{\mathbf{i}}_1 = (1, 0, \ldots, 0)$, $\breve{\mathbf{i}}_2 = (0, 1, 0, \ldots, 0)$, etc.), que se denominan *derivadas parciales*,

$$\left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}} \equiv \left.\frac{\partial f}{\partial \breve{\mathbf{i}}_i}\right|_{\mathbf{x}} = \lim_{h\to 0} \frac{f(x_1, x_2, \ldots, x_i + h, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{h}. \tag{3}$$

Notemos que en el numerador del cociente de Newton la función está evaluada en dos puntos que sólo difieren en la coordenada i-ésima. La derivada es “parcial” porque si se mantienen constantes todas las variables menos una, entonces se trata simplemente de una derivada de una función de una variable (la variable elegida), actuando las otras como constantes. Por lo tanto, dada por ejemplo $f(x, y, z)$, al escribir la derivada parcial $\partial f/\partial z$ *no es necesario* aclarar que x e y son constantes, puesto que *por definición* efectivamente se mantienen constantes al armar el cociente de Newton:

$$\left.\frac{\partial f}{\partial z}\right|_{(x_0, y_0, z_0)} = \lim_{h\to 0} \frac{f(x_0, y_0, z_0 + h) - f(x_0, y_0, z_0)}{h}. \tag{4}$$

La notación

$$\left(\frac{\partial f}{\partial z}\right)_{x,y}, \tag{5}$$

que se lee “derivada parcial de f con respecto a z, con x e y constantes”, es redundante si del contexto está claro que f depende de las variables x, y y z. En termodinámica, sin embargo, se la suele emplear, porque muchas veces no es obvio cuál es el conjunto de variables independientes. Por ejemplo, el estado de un sistema simple puede especificarse, como veremos, dando las variables (U, V, N), o bien las variables (T, V, N). Otro observable, por ejemplo la presión, será entonces función de las variables que definen el estado, pero según sea la elección será $P = f(U, V, N)$ o $P = g(T, V, N)$. Si queremos expresar la razón de cambio de la presión respecto del volumen usaremos naturalmente una derivada parcial, pero según el conjunto

[1]Recordemos que las barras verticales indican el *módulo* o longitud de un vector, $|x| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$ y que el módulo de la diferencia $|\mathbf{x} - \mathbf{y}|$ es la *distancia euclídea* entre los puntos $\mathbf{x}$ e $\mathbf{y}$.

de variables, estaremos derivando dos funciones distintas:

$$\left(\frac{\partial P}{\partial V}\right)_{U,N} = \frac{\partial f}{\partial V}, \tag{6}$$

$$\left(\frac{\partial P}{\partial V}\right)_{T,N} = \frac{\partial g}{\partial V}. \tag{7}$$

El sentido de la notación (5) es entonces especificar claramente de qué función se está tomando la derivada sin necesidad de introducir los nombres f y g, y no recalcar que x e y son constantes, cosa que sabemos en virtud de que se trata de una derivada parcial.

1.3. Derivación y regla de la cadena. De lo dicho se sigue que las reglas de derivación estudiadas para el caso de una variable se aplican directamente a las derivadas parciales, tratando al resto de las variables como constantes. Sólo es necesario generalizar la "regla de la cadena" para derivar funciones compuestas: dada $f(x, y, z)$, si se definen tres funciones $x(u, v, w)$, $y(u, v, w)$ y $z(u, v, w)$ entonces puede definirse la función compuesta

$$g(u, v, w) = f\big(x(u, v, w), y(u, v, w), z(u, v, w)\big).$$

Las derivadas parciales se calculan mediante

$$\frac{\partial g}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial u}, \tag{8}$$

es decir se aplica la regla de la cadena reiteradamente, llegando a la variable u a través de todas las variables originales que se han hecho depender de ésta.

Retomando el ejemplo de la presión expresada en función de (U, V, N), o de (T, V, N), sabemos que esto sucede porque se ha definido $T(U, V, N)$ y luego se ha invertido la función para expresar U en función de (T, V, N), así $g(T, V, N) = f(U(T, V, N), V, N)$. La regla de la cadena nos permite entonces relacionar las derivadas (6) y (7):

$$\frac{\partial g}{\partial V} = \frac{\partial f}{\partial U}\frac{\partial U}{\partial V} + \frac{\partial f}{\partial V}, \tag{9}$$

que en termodinámica se suele expresar con la notación

$$\left(\frac{\partial P}{\partial V}\right)_{T,N} = \left(\frac{\partial P}{\partial U}\right)_{V,N}\left(\frac{\partial U}{\partial V}\right)_{T,N} + \left(\frac{\partial P}{\partial V}\right)_{U,N}. \tag{10}$$

Una forma práctica de aplicar la regla de la cadena de forma sistemática es mediante el método de los jacobianos (sec. 3.1).

1.4. Derivadas sucesivas. Puesto que la derivada parcial es a su vez una función de varias variables, se definen las derivadas de mayor orden por

$$\frac{\partial^n f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_n}} = \frac{\partial}{\partial x_{i_1}}\left(\frac{\partial}{\partial x_{i_2}}\left(\dots\frac{\partial f}{\partial x_{i_n}}\right)\right). \tag{11}$$

En general el orden en que se toman las derivadas es importante; la igualdad

$$\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{\mathbf{x}_0} = \left.\frac{\partial^2 f}{\partial y \partial x}\right|_{\mathbf{x}_0} \qquad (?), \tag{12}$$

no puede asegurarse siempre, pero si ambas existen en un entorno de $\mathbf{x}_0$ y son continuas en $\mathbf{x}_0$, entonces se puede afirmar (12). Los potenciales termodinámicos son analíticos excepto en las transiciones de fase, de modo que se en la mayor parte del espacio de variables termodinámicas se puede utilizar (12). En una transición de fase la igualdad puede no ser válida (por ejemplo en una transición de fase de primer orden las derivadas parciales son discontinuas, de modo que sólo existen las derivadas segundas laterales).

1.5. Diferenciabilidad. Se dice que $f(\mathbf{x})$ es *diferenciable* en $\mathbf{x}_0$ si existe una función lineal[2] $\lambda(\mathbf{h})$ tal que

$$\lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0)-\lambda(\mathbf{h})}{|\mathbf{h}|}=0. \tag{13}$$

Si $\lambda(\mathbf{h})$ existe, entonces es única. La existencia de derivadas parciales (aunque sea en todas direcciones), no implica diferenciabilidad. Una condición suficiente para que una función sea diferenciable en un punto es que en ese punto existan *y sean continuas* todas las derivadas parciales. A diferencia de la derivabilidad, el hecho de que una función sea diferenciable en un punto, implica que es continua en ese punto.

La diferenciabilidad asegura la existencia de una aproximación lineal a la función en los puntos donde es diferenciable, puesto que la ec. (13) implica que

$$f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0)=\lambda(\mathbf{h})+r(\mathbf{h}), \tag{14}$$

con un resto $r(\mathbf{h})$ que se anula más rápido que linealmente al acercarse a $\mathbf{x}_0$, es decir que $\frac{r(\mathbf{h})}{|\mathbf{h}|}\xrightarrow[\mathbf{h}\to 0]{} 0$. En otras palabras, la diferenciabilidad asegura la existencia del polinomio de Taylor de primer orden (como veremos, $\lambda(\mathbf{h})$ no es otra cosa que el plano tangente). La noción de diferenciabilidad no es necesaria en la teoría de funciones de una variable, pues en ese caso (13) se sigue directamente de la existencia de la derivada.

2. El diferencial

Cuando existe la aproximación lineal a una función (la función $\lambda(\mathbf{h})$ de (13) o (14)), ésta función se llama *diferencial* de $f(\mathbf{x})$ en el punto $\mathbf{x}_0$, y se nota $df_{\mathbf{x}_0}(\mathbf{h})$. Insistimos en que el diferencial es una *función lineal* (o forma lineal) de su argumento $\mathbf{h}$; el subíndice $\mathbf{x}_0$ nos recuerda que puede definirse una tal función lineal en cada punto en donde $f(\mathbf{x})$ es diferenciable.

Poniendo sucesivamente $\mathbf{h}=h\breve{\mathbf{i}}_1$, $\mathbf{h}=h\breve{\mathbf{i}}_2$, etc. en (14), dividiendo por h y tomando $h\to 0$ se encuentra que los coeficientes de $\lambda(\mathbf{h})$, coinciden con las derivadas parciales evaluadas en $\mathbf{x}_0$:

$$df_{\mathbf{x}_0}(\mathbf{h})=\sum_i \left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}_0} h_i. \tag{15}$$

Frecuentemente se escribe $\Delta\mathbf{x}$ en lugar de $\mathbf{h}$, puesto que $\mathbf{h}$ indica un desplazamiento respecto de $\mathbf{x}_0$. Podemos expresar el hecho de que el diferencial es la aproximación lineal a nuestra función escribiendo

$$\Delta f \equiv f(\mathbf{x}+\boldsymbol{\Delta}\mathbf{x})-f(\mathbf{x})\approx df_{\mathbf{x}}(\Delta x)=\sum_i \left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}} \Delta x_i. \tag{16}$$

[2]Una función $g(\mathbf{x})$ es lineal en $\mathbf{x}$ si y sólo si $g(a\mathbf{x}+b\mathbf{y})=ag(\mathbf{x})+bg(\mathbf{y})$. Tiene entonces que tener la forma $g(\mathbf{x})=\mathbf{a}\cdot\mathbf{x}=\sum_i a_i x_i$.

En el caso de una variable, $df_x(\Delta x) = f'(x)\Delta x$. En particular, para la función identidad $f(x) = x$, es $df(\Delta x) \equiv dx(\Delta x) = \Delta x$ para todo x. Por esto se suele escribir

$$df = \sum_i \frac{\partial f}{\partial x_i}\, dx_i, \tag{17}$$

omitiendo el punto $\mathbf{x}_0$ y el argumento del diferencial ($\Delta\mathbf{x}$, que es igual a $(\Delta x_1, \Delta x_2, \ldots) = (dx_1, dx_2, \ldots)$). Esta notación, sin embargo, no debe hacernos perder de vista que el diferencial no es otra cosa que una función lineal de sus argumentos $\Delta x_i = dx_i$, así como función (en general no lineal) del punto de diferenciación $\mathbf{x}$.

Finalmente, notemos que también se puede escribir el diferencial usando el gradiente de f, $\boldsymbol{\nabla} f \equiv (\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_n)$:

$$df_{\mathbf{x}}(d\mathbf{x}) = \boldsymbol{\nabla} f \cdot d\mathbf{x}. \tag{18}$$

2.1. El diferencial y las reglas de derivación. Las reglas de derivación se aplican también al diferencial, en virtud de ser éste una combinación lineal de las derivadas parciales. Si tenemos $f(\mathbf{x})$ y $g(\mathbf{x})$, es evidente que

$$d_{\mathbf{x}}(f+g) = d_{\mathbf{x}}f + d_{\mathbf{x}}g, \tag{19}$$

así como

$$d_{\mathbf{x}}(fg) = \sum_i \left[\frac{\partial f}{\partial x_i} g(\mathbf{x}) + f(\mathbf{x})\frac{\partial g}{\partial x_i}\right] dx_i = g(\mathbf{x}) d_{\mathbf{x}} f + f(\mathbf{x}) d_{\mathbf{x}} g, \tag{20}$$

recordando que estamos “diferenciando respecto de $\mathbf{x}$”, esto es que las derivadas (parciales) que aparecen son respecto al punto de diferenciación, y no respecto a los argumentos del diferencial.

Si definimos ahora un conjunto de funciones tal que $\mathbf{x} = \mathbf{x}(\mathbf{y})$[3], podemos definir la función compuesta $g(\mathbf{y}) = f(\mathbf{x}(\mathbf{y}))$. El diferencial de esta última, en virtud de la regla de la cadena, es

$$dg = \sum_j^m \sum_i^n \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial y_j} dy_j. \tag{21}$$

Esta fórmula se puede obtener también de la notación (17), con la siguiente ventaja mnemotécnica. Si se ha escrito df utilizando las x_i como variables, puede obtenerse (21) considerando que al haber introducido las variables y_i, los dx_i ya no son simplemente iguales a los Δx_i, sino que son a su vez aproximaciones lineales de las nuevas funciones,

$$dx_i = \sum_j^m \frac{\partial x_i}{\partial y_j} dy_j, \tag{22}$$

y (21) sigue sustituyendo (22) en (17).

[3]Es decir $x_1 = x_1(\mathbf{y})$, $x_2 = x_2(\mathbf{y})$, etc. No es necesario que la cantidad de variables y_i sea igual al número de variables x_i.

2.2. Diferenciales superiores. Si la función $f(\mathbf{x})$ y sus derivadas parciales hasta orden $n+1$ son continuas en un entorno de $\mathbf{x}_0$, entonces se puede desarrollar $f(\mathbf{x})$ en un polinomio de orden n (polinomio de Taylor),

$$\begin{aligned} f(\mathbf{x}_0 + \Delta\mathbf{x}) = f(\mathbf{x}_0) + \sum_i \left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}_0} \Delta x_i + \frac{1}{2}\sum_{ij} \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathbf{x}_0} \Delta x_i \Delta x_j + \dots \\ + \frac{1}{n!}\sum_{i_1\dots i_n} \left.\frac{\partial^n f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_n}}\right|_{\mathbf{x}_0} \Delta x_{i_1}\Delta x_{i_2}\dots\Delta x_{i_n} + R_n, \end{aligned} \tag{23}$$

donde R_n es una función que se anula más rápido que $|\mathbf{x}-\mathbf{x}_0|^n$ cuando $\mathbf{x}\to\mathbf{x}_0$. El término lineal en los Δx_i es como hemos visto el diferencial $df_{\mathbf{x}_0}(\Delta\mathbf{x})$; los términos sucesivos (a menos del factor $1/n!$) definen los diferenciales segundo, tercero, etc. Formalmente se pueden obtener diferenciando el diferencial de orden inmediato inferior. Por ejemplo

$$\begin{aligned} d^2 f = d(df) &= d\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) \\ &= d\left(\frac{\partial f}{\partial x}\right)dx + \frac{\partial f}{\partial x}d(dx) + d\left(\frac{\partial f}{\partial y}\right)dy + \frac{\partial f}{\partial y}d(dy). \end{aligned} \tag{24}$$

Puesto que la diferenciación opera sobre $\mathbf{x}$ y no sobre los dx_i (argumentos del diferencial), decimos que $d(dx) = d^2x = 0$ y diferenciamos las derivadas parciales,

$$d\left(\frac{\partial f}{\partial x}\right)dx = \frac{\partial^2 f}{\partial^2 x}(dx)^2 + \frac{\partial^2 f}{\partial y \partial x}dydx,$$

y análogamente para y, de modo que

$$d^2_{\mathbf{x}} f = \frac{\partial^2 f}{\partial^2 x}(dx)^2 + 2\frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial^2 y}(dy)^2. \tag{25}$$

El diferencial segundo es entonces una *forma cuadrática* en los dx_i.

2.3. Diferencial “inexacto”. Dada una función vectorial $\mathbf{F}(\mathbf{x})$, puede construirse la función

$$f_{\mathbf{x}}(\mathbf{h}) = \mathbf{F}(\mathbf{x})\cdot\mathbf{h}, \tag{26}$$

a la que, poniendo $\mathbf{h} = \Delta\mathbf{x}$, puede dársele la forma de un diferencial, o sea (con dos variables)

$$f_{\mathbf{x}}(\Delta\mathbf{x}) = F_x(\mathbf{x})\Delta x + F_y(\mathbf{x})\Delta y = F_x dx + F_y dy \equiv \text{“}d\phi\text{”}. \tag{27}$$

Podemos preguntarnos si $f_{\mathbf{x}}(\mathbf{h})$ es realmente un diferencial, es decir si es en cada $\mathbf{x}$ la aproximación lineal de alguna función $\phi(\mathbf{x})$. Si las F_α tienen derivadas continuas, la condición para que se trate de un diferencial es

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}, \tag{28}$$

Si esto se cumple, existe una $\phi(\mathbf{x})$ tal que $\mathbf{F}(\mathbf{x}) = \boldsymbol{\nabla}\phi(\mathbf{x})$ y efectivamente

$$d\phi = \boldsymbol{\nabla}\phi\cdot d\mathbf{x} = f_{\mathbf{x}}(d\mathbf{x}). \tag{29}$$

Algunos textos llaman “diferencial exacto” a $d\phi$, en este caso, mientras que de no cumplirse (28) dicen que (26) es un “diferencial inexacto”. Nosotros preferimos decir que $d\phi$ es un diferencial en el primer caso, mientras que en el segundo no lo es.

2.3.1. Factor integrante. Ante una expresión como (26) o (29), donde $\mathbf{F}(\mathbf{x})$ *no* es un gradiente ("diferencial inexacto"), cabe preguntarse si existe una función escalar $\mu(\mathbf{x})$ tal que

$$\mu(\mathbf{x})\mathbf{F}(\mathbf{x}) = \boldsymbol{\nabla}\phi, \tag{30}$$

o bien, que transforme el "diferencial inexacto" en "diferencial exacto". Si $\mu(\mathbf{x})$ existe, se denomina *factor integrante.*

2.4. Diferencial e integrales. Dada $\mathbf{F}(\mathbf{x})$ y una curva C parametrizada por una función $\mathbf{r}(t)$, $t \in [a, b]$ contenida en su dominio, su *integral de línea* es

$$\int_C \mathbf{F}(\mathbf{x}) \cdot d\mathbf{l} \equiv \int_C \mathbf{F}(\mathbf{x}) \cdot \breve{\mathbf{t}}(\mathbf{x}) ds \equiv \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \frac{d\mathbf{r}}{dt}\, dt, \tag{31}$$

donde $\breve{\mathbf{t}}(\mathbf{x})$ es el vector tangente unitario, $\breve{\mathbf{t}}(\mathbf{x}) = \mathbf{r}'(t)/|\mathbf{r}'(t)|$, y la integral de arco es

$$\int_C f(\mathbf{x})\, ds \equiv \int_a^b f(\mathbf{r}(t)) \left|\frac{d\mathbf{r}}{dt}\right| dt. \tag{32}$$

Las definiciones pueden recordarse más fácilmente utilizando la notación formal $d\mathbf{l} = d\mathbf{r} = \frac{d\mathbf{r}}{dt}dt$, y $ds = |d\mathbf{l}| = |\frac{d\mathbf{r}}{dt}||dt|$ (la integral $\int_C ds$ da la longitud de la curva C).

El integrando de la integral de línea (31) tiene la forma de un diferencial (26), aunque no necesariamente lo es (puede ser "inexacto"). Esta equivalencia formal es más evidente si se escribe (limitándonos por claridad a dos variables) $\mathbf{r}(t) = (x(t), y(t))$, con lo que $d\mathbf{l} = (x'(t), y'(t))dt = (dx, dy)$, y

$$\int \mathbf{F}(\mathbf{x}) \cdot d\mathbf{l} = \int F_x(x, y)\, dx + F_y(x, y)\, dy \equiv \int d\phi. \tag{33}$$

La notación es consistente si se recuerda que x e y no son variables independientes sino que están ligadas por pertenecer a la curva $\mathbf{r}(t)$, pero el segundo miembro de (33) es un tanto desafortunado porque sugiere que la integral original se puede descomponer en dos integrales, una que involucra sólo la variable x y una que sólo involucra y. Esto no es cierto en general, puesto que esta interpretación implica cambiar la curva C, transformándola en una quebrada que une los puntos $\mathbf{r}(a)$ y $\mathbf{r}(b)$ con un segmento de recta paralelo al eje X y otro paralelo al eje Y. Por otro lado, nos resistimos a llamar "diferencial" a $d\phi$, puesto que no se trata ya de una función lineal (los dx no son argumentos de una función, sino que indican las variables de integración). Sin embargo, en el caso en que $\mathbf{F}(\mathbf{x})$ se pueda escribir como el gradiente de una función $\phi(\mathbf{x})$ (es decir, se cumple la condición (28) para que $d\phi$ sea "diferencial exacto"), la integral depende sólo de los extremos de integración y no de la curva C, con lo cual efectivamente se puede calcular integrando en dos tramos rectos como sugiere (33), y

$$\int_C \mathbf{F}(\mathbf{x}) \cdot d\mathbf{l} = \int_C \boldsymbol{\nabla}\phi(\mathbf{x}) \cdot d\mathbf{l} = \phi\big(\mathbf{r}(b)\big) - \phi\big(\mathbf{r}(a)\big) = \Delta\phi, \tag{34}$$

como sugiere el último miembro de (33). También tiene sentido una expresión como

$$\int PdV + VdP = \int d(PV) = \Delta(PV), \tag{35}$$

ya sea que se trate de una función de una variable $P(V)$ (o $V(P)$), o que P y V sean funciones de una o más variables, ya que si fuera por ejemplo $P(x, y)$ y $V(x, y)$

la primera integral está indicando la integral de línea de $\mathbf{F}(x,y) = \boldsymbol{\nabla}(PV)$, como se puede verificar desarrollando dV y dP.

El uso de la palabra diferencial para designar al integrando de una integral de línea es un tanto desafortunado, pero es un abuso de lenguaje fácil de comprender. Por ejemplo en termodinámica es frecuente querer calcular el trabajo involucrado en determinado proceso. Generalmente uno puede escribir el trabajo para un desplazamiento infinitesimal, digamos $P(V,T)dV$ y lo llama "diferencial de trabajo" dW. La cantidad requerida es la integral de esa expresión, que se escribe $\Delta W = \int_C dW = \int_c P\big(V(t), T(t)\big) dV/dt\, dt$ (t es la variable que parametriza a la curva). Esto es otro abuso de notación, puesto que resulta que en general el trabajo infinitesimal *no* es expresable como un gradiente, por lo que la integral dependerá de la curva C elegida, y ΔW es un número que depende de C (una funcional de la curva) y no una la diferencia entre valores de alguna función W evaluada en los extremos de la misma.

2.5. Diferencial y extremos. La condición necesaria para que un punto en donde la función $f(\mathbf{x})$ tiene derivadas parciales sea un extremo local es que se anulen en ese punto todas las derivadas parciales, lo que se puede expresar en forma compacta como

$$\boldsymbol{\nabla} f|_{\mathbf{x}} = 0, \tag{36a}$$

o bien

$$df_{\mathbf{x}} = 0. \tag{36b}$$

La primera de las ecuaciones implica que todas las derivadas parciales son nulas porque se trata de una igualdad vectorial. La segunda ecuación debe leerse como una identidad entre funciones: estamos pidiendo que el diferencial sea una función nula (es decir, que $df(d\mathbf{x}) = 0\ \forall d\mathbf{x}$) lo que implica que deben ser nulos todos sus coeficientes. Frecuentemente en termodinámica resulta cómodo expresar la condición en la forma diferencial (36b) porque se evita de esta manera explicitar las variables independientes. Así, diremos por ejemplo que el equilibrio termodinámico requiere de un máximo de la entropía, de modo que exigiremos

$$dS = 0.$$

Esta condición es efectivamente válida en general, pero el lector deberá en cada caso particular saber reconocer las variables independientes para saber qué derivadas parciales habrá que tomar (y pedir que se anulen) para construir el diferencial.

La condición $df = 0$ no asegura por sí sola que el punto sea un extremo. Si la corrección cuadrática (tercer término de (23)) puede tomar ambos signos según los valores del incremento $d\mathbf{x}$, el punto no es un extremo sino un *punto de ensilladura.* Si en cambio $d^2 f > 0$ ($d^2 f < 0$) para todo $d\mathbf{x}$, el punto es un mínimo (máximo) local. Resumiendo, $\mathbf{x}$ es máximo local si cumple

$$df = 0, \qquad d^2 f < 0, \tag{37}$$

donde tanto igualdad como desigualdad deben leerse en sentido funcional, en particular la desigualdad impone que la forma cuadrática sea *definida negativa,* es decir que $d^2 f(d\mathbf{x}) < 0\ \forall d\mathbf{x}$. Esto se traduce en condiciones sobre las derivadas segundas que resulta más conveniente expresar diciendo que todos los autovalores de la forma cuadrática deben ser negativos. En el caso de dos variables, esto se traduce en las condiciones

$$\frac{\partial^2 f}{\partial x^2} < 0, \qquad \frac{\partial^2 f}{\partial y^2} < 0, \qquad \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} \geq \frac{\partial^2 f}{\partial x \partial y} \qquad \text{(máximo)}. \tag{38}$$

En el caso de un mínimo local, el diferencial segundo debe ser *definido positivo,* que en dos variables equivale a pedir

$$\frac{\partial^2 f}{\partial x^2} > 0, \qquad \frac{\partial^2 f}{\partial y^2} > 0, \qquad \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} \geq \frac{\partial^2 f}{\partial x \partial y} \qquad \text{(mínimo)}. \tag{39}$$

3. Función implícita

Los resultados de esta sección se exponen con más detalle por ejemplo en Piskunov (1977, § VIII.11) o Margenau y Murphy (1956, §1.5).

Consideremos una función de tres variables (la generalización a más variables es directa) $F(x, y, z)$, diferenciable y monótona en un entorno del punto $\mathbf{x}_0$. La condición $F(x, y, z) = 0$ define una superficie en el entorno en donde F es monótona, ligando las variables de modo que sólo dos son independientes. Se dice entonces que x es una *función implícita* de y, z (o bien y de x, z, etc.). Encontrar una expresión explícita para tal función puede ser muy complicado, pero pueden expresarse sus derivadas parciales. Para esto diferenciamos la condición, obteniendo

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial z}dz = 0, \tag{40}$$

igualdad que es válida como identidad funcional (el diferencial se anula en tanto función lineal, no sólo evaluado en algún punto) pero solamente sobre la superficie (es decir cuando $x = x(y, z)$). Una vez elegida la variable que se quiere hacer dependiente, se desarrolla el respectivo diferencial, por ejemplo

$$\begin{aligned} \frac{\partial F}{\partial x}\left(\frac{\partial x}{\partial y}dy + \frac{\partial x}{\partial z}dz\right) + \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial z}dz &= \\ \left(\frac{\partial F}{\partial x}\frac{\partial x}{\partial y} + \frac{\partial F}{\partial y}\right)dy + \left(\frac{\partial F}{\partial x}\frac{\partial x}{\partial z} + \frac{\partial F}{\partial z}\right)dz &= 0. \end{aligned} \tag{41}$$

Este último diferencial debe ser idénticamente nulo, pues corresponde al diferencial de $G(y, z) = f(x(y, z), y, z)$, que es idénticamente nula (el desarrollar el diferencial dx tiene el efecto de aplicar automáticamente la regla de la cadena). Los coeficientes de dy y dz deben por lo tanto anularse, de modo que podemos deducir

$$\left(\frac{\partial x}{\partial y}\right)_z = -\frac{\partial F/\partial y}{\partial F/\partial x}, \qquad \left(\frac{\partial x}{\partial z}\right)_y = -\frac{\partial F/\partial z}{\partial F/\partial x}. \tag{42}$$

Tomando como variables independientes (x, y) o (x, z) se deduce también

$$\left(\frac{\partial z}{\partial x}\right)_y = -\frac{\partial F/\partial x}{\partial F/\partial z}, \qquad \left(\frac{\partial z}{\partial y}\right)_x = -\frac{\partial F/\partial y}{\partial F/\partial z}, \tag{43}$$

$$\left(\frac{\partial y}{\partial x}\right)_z = -\frac{\partial F/\partial x}{\partial F/\partial y}, \qquad \left(\frac{\partial y}{\partial z}\right)_x = -\frac{\partial F/\partial z}{\partial F/\partial y}. \tag{44}$$

Con las expresiones anteriores pueden verificarse fácilmente las identidades

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z = 1, \tag{45}$$

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1. \tag{46}$$

Aunque la segunda puede resultar algo sorprendente desde el punto de vista de la simplificación de diferenciales (regla mnemotécnica que "suele" funcionar), su sentido y justificación deberían resultar claros.

Si agregamos una segunda condición $G(x, y, z) = 0$, quedarán determinadas dos variables en función de la tercera (digamos z). Los diferenciales dx y dy son entonces funciones lineales de dz, de modo que

$$dF = \left[\frac{\partial F}{\partial x}\frac{dx}{dz} + \frac{\partial F}{\partial y}\frac{dy}{dz} + \frac{\partial F}{\partial z}\right] dz = 0, \tag{47a}$$

$$dG = \left[\frac{\partial G}{\partial x}\frac{dx}{dz} + \frac{\partial G}{\partial y}\frac{dy}{dz} + \frac{\partial G}{\partial z}\right] dz = 0. \tag{47b}$$

Las derivadas dx/dz y dy/dz pueden hallarse en función de las derivadas parciales de F y G resolviendo el sistema lineal dado por el requisito de anulación de los coeficientes entre corchetes en (47).

3.1. Método de los jacobianos. En termodinámica es frecuente la necesidad de realizar cambios de variables y consecuentemente expresar ciertas derivadas parciales en términos de derivadas calculadas respecto a las variables independientes antes del cambio. Tales relaciones se obtienen por medio de la regla de la cadena, pero tratándose de decenas de cantidades que dependen de un conjunto de tres o más variables independientes la cantidad de relaciones posibles es muy grande y resulta conveniente disponer de un método que sistemática y sencillamente logre expresar las nuevas derivadas en función de las conocidas en las variables anteriores al cambio, sin hacer aparecer otras que requieran a su vez ser reescritas mediante la regla de la cadena. Un método que cumple este requisito es el de los jacobianos (Margenau y Murphy, 1956, §1.12, Landau y Lifshitz, 1959, §16).

El jacobiano es el determinante de una matriz de derivadas (la matriz jacobiana):

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}. \tag{48}$$

Una derivada única puede escribirse

$$\left(\frac{\partial u}{\partial x}\right)_y = \frac{\partial(u, y)}{\partial(x, y)}. \tag{49}$$

Es directo mostrar las siguientes propiedades: por las propiedades del determinante,

$$\frac{\partial(u, v)}{\partial(x, y)} = -\frac{\partial(v, u)}{\partial(x, y)}, \tag{50}$$

por la regla de la derivada del producto (desarrollando el jacobiano),

$$\frac{d}{dt}\left(\frac{\partial(u, v)}{\partial(x, y)}\right) = \frac{\partial(du/dt, v)}{\partial(x, y)} + \frac{\partial(u, dv/dt)}{\partial(x, y)}, \tag{51}$$

y finalmente mediante la regla de la cadena,

$$\begin{aligned}\frac{\partial(u,v)}{\partial(t,s)}\frac{\partial(t,s)}{\partial(x,y)} &= \begin{vmatrix} \left(\frac{\partial u}{\partial t}\right)_s\left(\frac{\partial t}{\partial x}\right)_y + \left(\frac{\partial u}{\partial s}\right)_t\left(\frac{\partial s}{\partial x}\right)_y & \left(\frac{\partial u}{\partial t}\right)_s\left(\frac{\partial t}{\partial y}\right)_x + \left(\frac{\partial u}{\partial s}\right)_t\left(\frac{\partial s}{\partial y}\right)_x \\ \left(\frac{\partial v}{\partial t}\right)_s\left(\frac{\partial t}{\partial x}\right)_y + \left(\frac{\partial v}{\partial s}\right)_t\left(\frac{\partial s}{\partial x}\right)_y & \left(\frac{\partial v}{\partial t}\right)_s\left(\frac{\partial t}{\partial y}\right)_x + \left(\frac{\partial v}{\partial s}\right)_t\left(\frac{\partial s}{\partial y}\right)_x \end{vmatrix} \\ &= \begin{vmatrix} \left(\frac{\partial u}{\partial x}\right)_y & \left(\frac{\partial u}{\partial y}\right)_x \\ \left(\frac{\partial v}{\partial x}\right)_y & \left(\frac{\partial v}{\partial y}\right)_x \end{vmatrix} = \frac{\partial(u,v)}{\partial(x,y)}.\end{aligned} \tag{52}$$

Esta última es la identidad más útil para efectuar los cambios de variables buscados, junto con las ya conocidas

$$\left(\frac{\partial u}{\partial x}\right)_y\left(\frac{\partial x}{\partial u}\right)_y = 1, \tag{53}$$

$$\left(\frac{\partial u}{\partial x}\right)_y\left(\frac{\partial x}{\partial y}\right)_u\left(\frac{\partial y}{\partial u}\right)_x = -1. \tag{54}$$

El ejemplo del cambio de variables (U,V,N) a (T,V,N) a través de una función $U(T,V,N)$ (pág. 3) se haría mediante el jacobiano como sigue:

$$\begin{aligned}\left(\frac{\partial P}{\partial V}\right)_{T,N} &= \frac{\partial(P,T)}{\partial(V,T)} = \frac{\frac{\partial(P,T)}{\partial(V,U)}}{\frac{\partial(V,T)}{\partial(V,U)}} \\ &= \left[\left(\frac{\partial P}{\partial V}\right)_U\left(\frac{\partial T}{\partial U}\right)_V - \left(\frac{\partial P}{\partial U}\right)_V\left(\frac{\partial T}{\partial V}\right)_U\right]\left[\left(\frac{\partial T}{\partial U}\right)_V\right]^{-1} \\ &= \left(\frac{\partial P}{\partial V}\right)_U - \left(\frac{\partial P}{\partial U}\right)_V\frac{\left(\frac{\partial T}{\partial V}\right)_U}{\left(\frac{\partial T}{\partial U}\right)_V} \\ &= \left(\frac{\partial P}{\partial V}\right)_U + \left(\frac{\partial P}{\partial U}\right)_V\left(\frac{\partial U}{\partial V}\right)_T.\end{aligned} \tag{55}$$

En este caso simple (cambia una sola variable) resulta más sencillo aplicar la regla de la cadena *ad hoc* como hicimos antes. Lo sistemático del método de los jacobianos resulta más evidente ante un ejemplo más complicado. Supongamos ahora que se desea expresar la derivada $\left(\frac{\partial U}{\partial S}\right)_V$ en términos de las derivadas de $U(P,T)$ (y de las funciones del cambio de variables, $S(P,T)$, $V(P,T)$), entonces

$$\left(\frac{\partial U}{\partial S}\right)_V = \frac{\partial(U,V)}{\partial(S,V)} = \frac{\frac{\partial(U,V)}{\partial(P,T)}}{\frac{\partial(S,V)}{\partial(P,T)}} = \frac{\left(\frac{\partial U}{\partial P}\right)_T\left(\frac{\partial V}{\partial T}\right)_P - \left(\frac{\partial V}{\partial P}\right)_T\left(\frac{\partial U}{\partial T}\right)_P}{\left(\frac{\partial S}{\partial P}\right)_T\left(\frac{\partial V}{\partial T}\right)_P - \left(\frac{\partial S}{\partial T}\right)_P\left(\frac{\partial V}{\partial P}\right)_T}. \tag{56}$$

El cociente de jacobianos asegura que todas las cantidades están expresadas en términos de funciones de las variables deseadas, P y T. Si se desea se puede multiplicar y dividir por $\left(\frac{\partial T}{\partial V}\right)_P$ para obtener

$$\left(\frac{\partial U}{\partial S}\right)_V = \frac{\left(\frac{\partial U}{\partial P}\right)_T - \left(\frac{\partial V}{\partial P}\right)_T\left(\frac{\partial U}{\partial V}\right)_P}{\left(\frac{\partial S}{\partial P}\right)_T - \left(\frac{\partial V}{\partial P}\right)_T\left(\frac{\partial S}{\partial V}\right)_P}, \tag{57}$$

pero esta forma más compacta puede no ser tan conveniente como la anterior, ya que involucra derivadas de las funciones $U(V,P)$ y $S(P,V)$, que no necesariamente se conocen.

4. Funciones homogéneas

Se dice que una función es *homogénea de grado k* si cumple que

$$f(\lambda \mathbf{x}) = \lambda^k f(\mathbf{x}), \qquad \forall \lambda \in \mathbb{R}. \tag{58}$$

La condición de homogeneidad impone fuertes restricciones a la función y sus derivadas. En primer lugar, se cumple el *teorema de Euler,* que afirma que $f(\mathbf{x})$ es homogénea de grado k sí y sólo si

$$k f(\mathbf{x}) = \mathbf{x} \cdot \nabla f|_{\mathbf{x}} = \sum_i x_i \left. \frac{\partial f}{\partial x_i} \right|_{\mathbf{x}}, \qquad \text{(ecuación de Euler)}. \tag{59}$$

Para ver que la homogeneidad implica (59) basta derivar (58) respecto de λ y poner $\lambda = 1$. Para mostrar que la ecuación de Euler implica homogeneidad, definir $\Phi(\lambda) = f(\lambda \mathbf{x})$ y utilizar la ecuación de Euler para encontrar que $d\Phi/d\lambda = k\Phi(\lambda)/\lambda$, con lo que $\Phi(\lambda) = \Phi(1)\lambda^k$, es decir $f(\lambda \mathbf{x}) = f(\mathbf{x})\lambda^k$.

El teorema de Euler dice que *las derivadas parciales determinan completamente a la función* (a menos que $k = 0$). Como sabemos, siempre se puede escribir la función en términos de sus derivadas (como podríamos inferir también del hecho de que se trata de $n+1$ funciones de n variables, de modo que, al menos localmente, debería ser posible invertir y expresar $\mathbf{x}$ en función de las $\frac{\partial f}{\partial x_i}$). En el caso de las funciones homogéneas, el teorema de Euler nos da la forma explícita de esa relación y muestra además que no existen constantes aditivas arbitrarias, como en el caso general (es decir, un conjunto de derivadas corresponde a sólo una función homogénea).

La homogeneidad puede hacerse manifiesta escribiendo la función en una forma alternativa. Por ejemplo, para dos variables, eligiendo $\lambda = 1/y$ en (58) encontramos

$$f(x, y) = y^k f(x/y) \equiv y^k g(x/y), \tag{60}$$

donde $g(x)$ no tiene por qué ser homogénea. Otro ejemplo importante en termodinámica: la entropía de un sistema simple es función de la energía, el volumen y el número de partículas, con la propiedad de ser una función homogénea de grado 1 en las tres variables,

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N). \tag{61}$$

Eligiendo $\lambda = 1/N$, se concluye que esta función puede escribirse como el producto de N por una función de la energía por partícula y la inversa de la densidad:

$$S(E, V, N) = N S(E/N, V/N, 1) \equiv N s(e, v), \tag{62}$$

con $e = E/N$, $v = V/N$.

Es fácil ver que las derivadas parciales de una función homogénea son también funciones homogéneas pero de grado $k - 1$. En efecto, derivando (58) respecto de $\mathbf{x}_i$:

$$\lambda \frac{\partial f}{\partial x_i} = \lambda^k \frac{\partial f}{\partial x_i}. \tag{63}$$

En el caso particular $k = 1$, que es el que ocurre en termodinámica, esto implica que las derivadas son en realidad función de $n - 1$ variables (considerar $k = 0$ en (60)), con lo que debe poder escribirse una de las derivadas en función de las demás.

En dos variables y llamando $P(x,y)$ y $Q(x,y)$ a las derivadas, tenemos

$$\begin{aligned} f(x,y) &= xP(x,y) + yQ(x,y), \\ P(x,y) &= g(x/y), \\ Q(x,y) &= h(x/y), \end{aligned}$$

de modo que en principio es posible escribir $Q(x,y) = h\Big(g^{-1}\big(P(x,y)\big)\Big)$. No es posible dar la relación explícitamente en general, pero sí en forma de ecuación diferencial. Derivando la primera de las ecuaciones anteriores sucesivamente respecto de x e y se encuentra que

$$x\frac{\partial P}{\partial x} + y\frac{\partial Q}{\partial x} = 0, \tag{64a}$$

$$x\frac{\partial P}{\partial y} + y\frac{\partial Q}{\partial y} = 0, \tag{64b}$$

que, suponiendo conocida $P(x,y)$ constituyen un par de ecuaciones diferenciales que permiten obtener $Q(x,y)$ a menos de una constante de integración.

En termodinámica, la función homogénea (de grado 1) $f(\mathbf{x})$ será la llamada *relación fundamental* (la entropía o la energía expresada en función de las variables que determinan el estado de equilibrio). Sus derivadas $P_i \equiv \partial f/\partial x_i$ son cantidades accesibles experimentalmente y las relaciones $P_i(\mathbf{x})$ se denominan ecuaciones de estado. En virtud de lo anterior, resulta entonces que conociendo las n ecuaciones de estado (en principio posibles de determinar mediante experimentos o teorías microscópicas), queda unívocamente determinada la relación fundamental. Si se conocen sólo $n-1$ ecuaciones de estado, se puede obtener la restante por integración de ecuaciones diferenciales tipo las (64) y así obtener la relación fundamental a menos de una constante aditiva. En el caso de n variables es conveniente escribir estas ecuaciones en la forma de un diferencial. Diferenciando la ecuación de Euler,

$$f = \sum_i P_i x_i, \tag{65}$$

$$df = \sum_i P_i dx_i + x_i dP_i. \tag{66}$$

pero puesto que $df = \sum_i P_i dx_i$, resulta que

$$\sum_i x_i dP_i = 0, \qquad \text{(Gibbs-Duhem).} \tag{67}$$

Tomando las x_i como variables independientes, como hasta ahora, la ecuación de Gibbs-Duhem contiene todas las relaciones tipo (64), puesto que en ese caso es necesario desarrollar los dP_i, escribiendo $dP_i = \sum_j (\partial P_i/\partial x_j)dx_j$. En tanto igualdad funcional, Gibbs-Duhem requiere que se anulen todos los coeficientes de los dx_j, es decir $\sum_i x_i \partial P_i/\partial x_j = 0\ \forall j$.

Pero también, si se conocen las primeras $n-1$ de las P_i, puede optarse por considerar éstas como variables y considerar $P_n \equiv Q$ una función de las P_i, $i = 1, \ldots, n-1$. Invirtiendo las relaciones $P_i = P_i(x_1/x_n, \ldots, x_{n-1}/x_n)$ tendremos $x_i/x_n = g_i(P_i)$, y Gibbs-Duhem se leerá

$$\sum_{i=1}^{n-1} g_i(P_1 \ldots P_{n-1}) dP_i + dQ = 0, \tag{68}$$

o sea

$$\sum_{i=1}^{n-1}\left[g_i(P_1\dots P_{n-1})+\frac{\partial Q}{\partial P_i}\right]dP_i=0. \tag{69}$$

La integración de las $n-1$ ecuaciones contenidas en (69) permitirá obtener $Q(P_1,\dots,P_{n-1})$ a menos de una constante de integración.

4.1. Funciones homogéneas generalizadas. Una función que cumple que

$$f(\lambda^a x,\lambda^b y)=\lambda f(x,y) \tag{70}$$

se dice homogénea generalizada. En termodinámica, éste tipo de funciones aparece al tratar fenómenos críticos: la llamada hipótesis de escala consiste en suponer que la energía libre intensiva se puede separar en dos contribuciones aditivas (una regular y otra singular), y que la parte singular de la energía libre intensiva es una función homogénea generalizada.

5. Transformada de Legendre

Como acabamos de ver, en la ecuación de Gibbs-Duhem puede ser conveniente considerar todas o algunas de las derivadas $P_i(\mathbf{x})$ como variables independientes. Podemos preguntarnos si es posible hacer este cambio de variables *sin perder información* de la constante de integración que requiere la resolución de las ecuaciones de Gibbs-Duhem. La herramienta que permite hacer esto es la *transformada de Legendre* (Binney et al. 1992, §1.5, Zia et al. 2009).

Supongamos que tenemos una función $f(x)$ y que deseamos realizar un cambio de variable $x\to p$, donde la nueva variable independiente es $p(x)=f'(x)$ (relación que naturalmente debe ser invertida para dar $x(p)$. En termodinámica por ejemplo f puede ser la entropía y x la energía. Esta última es una cantidad de gran importancia física, pero experimentalmente es en general más conveniente controlar y medir la temperatura (que es una derivada de la entropía). Si simplemente hacemos el cambio de variables $x\to p$,

$$\begin{aligned} p(x)=f'(x) \qquad &\longrightarrow \qquad x=x(p) \qquad\qquad (f\ \text{monótona})\\ y=f(x) \qquad &\longrightarrow \qquad y=g(p)=f(x(p)). \end{aligned}$$

Podríamos recuperar $f(x)$ a partir de $g(p)$ y $p(x)$, pero si sólo tenemos $g(p)$, no podemos recuperar completamente $f(x)$:

$$f(x)=g(p(x)), \qquad p(x)=df/dx, \qquad \longrightarrow \qquad f(x)=g\left(\frac{df}{dx}\right), \tag{71}$$

es decir, obtenemos una ecuación diferencial para $f(x)$. Además de la dificultad que presenta el problema inverso, habremos perdido información (representada en la constante arbitraria de integración[4]). Geométricamente, lo que sucede es que todas las curvas desplazadas horizontalmente darán la misma función transformada.

[4] Una constante puede parecer poca cosa, pero si la función es de varias variables y decidimos transformar una de ellas, esa constante respecto de la variable transformada puede contener una dependencia no trivial en las otras variables. Considérese como ejemplo $f(x,y)=(x-4y)^2$. Cambiando $x\to p=\partial f/\partial x$ resulta en $g(p,y)=p^2/4$. Al intentar invertir obtendremos $f(x)=(x+A(y))^2$.

Para evitar estos problemas, la solución es definir una representación alternativa a $f(x)$, denominada transformada de Legendre: dada una $f(x)$ convexa hacia arriba (o hacia abajo) en todo su dominio, definimos

$$p = \frac{df}{dx}, \qquad \longrightarrow \qquad x = x(p), \tag{72}$$

$$g(p) = x(p)p - f\big(x(p)\big). \tag{73}$$

Esta representación $g(p)$ contiene toda la información original, y además simplifica el problema de la inversión:

$$\frac{dg}{dp} = \frac{dx}{dp}p + x(p) - \frac{df}{dx}\frac{dx}{dp} = x(p), \tag{74}$$

pues $df/dx = p$, es decir que invirtiendo $x = dg/dp|_p$ recuperamos la función $p(x)$. La función original es entonces

$$f(x) = xp(x) - g(p(x)), \tag{75}$$

es decir que la transformada de Legendre así definida es su propia inversa.

En resumen,

$$f(x) + g(p) = px, \tag{76}$$

$$\frac{df}{dx} = p, \qquad \frac{dg}{dp} = x, \tag{77}$$

y derivando (77) podemos encontrar la relación entre las convexidades, pues $d^2f/dx^2 = dp/dx$ y $d^2g/dp^2 = dx/dp$, con lo que

$$\frac{d^2f}{dx^2}\frac{d^2g}{dp^2} = 1. \tag{78}$$

Por razones históricas, en termodinámica se usa la transformada de Legendre con el signo invertido, rompiendo así la simetría entre la transformada y antitransformada.

$$G(P) = F(X) - XP, \qquad F(X) = G(P) + XP, \tag{79a}$$

$$P = \frac{dF}{dX}, \qquad X = -\frac{dG}{dP}, \tag{79b}$$

$$\frac{d^2F}{d^2X}\frac{d^2G}{dP^2} = -1 \tag{79c}$$

En particular, definida de este modo la transformada tiene la convexidad contraria a la función.

Referencias

Binney, J. J., Dowrick, N. J., Fisher, A. J. y Newman, M. E. J. (1992), *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*, Clarendon Press.

Landau, L. D. y Lifshitz, E. M. (1959), *Statistical Physics (Part I)*, volume 5 of *Course of Theoretical Physics*, Pergamon Press, Oxford, 3 edición.

Margenau, H. y Murphy, G. M. (1956), *The Mathematics of Physics and Chemistry*, D. van Nostrand, Princeton, 2 edición.

Piskunov, N. (1977), *Cálculo Diferencial e Integral*, volume I, Mir, Moscú.

Rey Pastor, J., Pi Calleja, P. y Trejo, C. A. (1963), *Análisis Matemático*, volume II, Kapelusz, Buenos Aires, 7 edición.

Zia, R. K. P., Redish, E. F. y McKay, S. R. (2009), Making sense of the Legendre transform. *Am. J. Phys.* **77**, 614–622.