

UNIVERSIDAD NACIONAL DE CÓRDOBA  
FACULTAD DE MATEMÁTICA, ASTRONOMÍA, FÍSICA Y  
COMPUTACIÓN

---

SERIES "C"

**TRABAJOS DE FÍSICA**

N.º 17/2022

**Notas complementarias de Termodinámica**

Guido A. Raggio



Editor: Miguel A. Chesta

CIUDAD UNIVERSITARIA 5000, CÓRDOBA  
REPÚBLICA ARGENTINA

# Notas complementarias de Termodinámica

G. A. Raggio

Marzo 2022

## 0. Advertencia

Estas notas se distribuyeron fragmentariamente a los alumnos del curso *Termodinámica y Mecánica Estadística I* de la Licenciatura en Física de FAMAFA en 2019. No constituyen una introducción a la termodinámica clásica. Sólo pretenden ilustrar, comentar o complementar ciertos aspectos o detalles de la presentación de el curso mencionado que se basó en el libro de H.B. Callen (*Thermodynamics (an introduction to the physical theories of equilibrium thermostatics and irreversible thermodynamics)*. J. Wiley & Sons, New York 1960.<sup>1</sup>). La palabras “concave” y “convex” no figuran en el índice del excelente libro de Callen. Sin embargo: ¿Quién puede dudar que la convexidad/concavidad de las funciones termodinámicas en sus variables naturales constituye la propiedad matemática fundamental<sup>2</sup>?

Aquí deducimos la concavidad de la entropía en sus variables extensivas naturales a partir de postulados generales (homogeneidad y aditividad). Las desigualdades que involucran a las magnitudes como compresibilidades, coeficientes de expansión, capacidades caloríficas, etc. surgen como consecuencias de la concavidad y no al revés.

## 1. Homogeneidad de grado 1 y concavidad de la entropía

En el libro de Callen (o en cualquier otro que presenta las leyes de la termodinámica por medio de un razonamiento deductivo a partir de postulados y no inductivamente usando evidencia empírica) la saga de la termodinámica comienza postulando la existencia de una función de estado –la entropía  $S$ – que posee ciertas propiedades específicas derivadas de (o motivadas por) leyes fenomenológicas. Se postula también que el sistema está caracterizado en cuanto a sus propiedades macroscópicas por un conjunto de variables extensivas que incluyen la energía interna  $U$ , el volumen  $V$ , así como variables adicionales que juntamos a una  $k$ -upla

---

<sup>1</sup>Hay segunda edición con cambios importantes.

<sup>2</sup>A aquellos interesados en presentaciones axiomáticas de las leyes de la termodinámica les recomiendo el libro de R. Giles (*Mathematical Foundations of Thermodynamics*. Pergamon Press, Oxford 1964) o el artículo de E.H. Lieb, J. Yngvason (*The Physics and Mathematics of the Second Law of Thermodynamics*. *arXiv:cond-mat/9708200*).

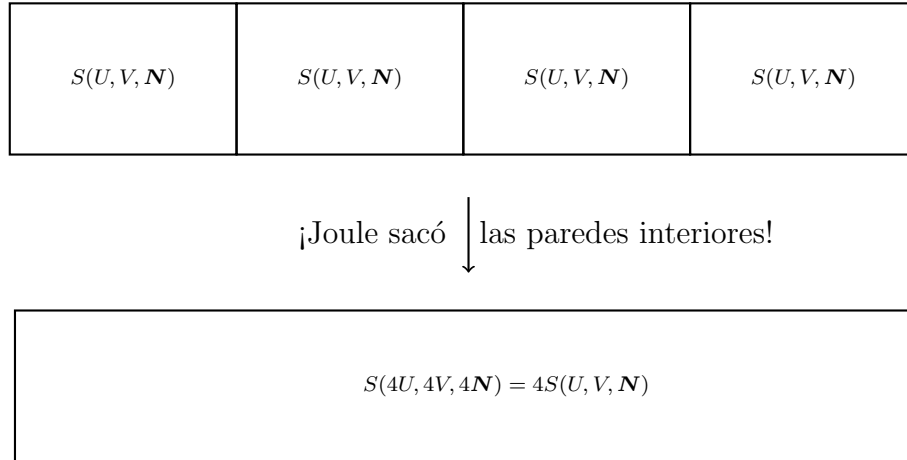


Figura 1: 4-aditividad

$\mathbf{X} = (X_1, X_2, \dots, X_k)$ . Dependiendo del sistema estas variables extensivas adicionales serán: números molares  $N_j$  que caracterizan la composición (o cualquier otra variable indicativa de la masa de la  $j$ -ésima componente con identidad química); momentos eléctricos o magnéticos (en presencia de campos electromagnéticos) o más precisamente la magnitud de la proyección de estos momentos sobre el campo correspondiente; componentes de tensiones o módulos elásticos en sistemas sólidos. Como función de estado,  $S$  depende de las variables  $(U, V, \mathbf{X})$  que caracterizan a este.

¿¿Adonde varia  $(U, V, \mathbf{X})$ ? Claramente,  $V > 0$  y también  $N_j > 0$ . En los casos de sistemas electromagnéticos o en el caso de sólidos, las variables correspondientes no tienen signo, i.e. varían en  $\mathbb{R}$ . Para la energía interna  $U$  la cosa no es tan simple –se miden solamente diferencias de energía– pero se introduce un estado de referencia (fiduciario o imaginario) para un sistema específico al que si se le puede asignar la energía cero, pero esto se complica cuando se tiene que ver con sistemas distintos. En general se supone, y esto es lo que hacemos aquí, que  $(U, V, \mathbf{X})$  varia en un subconjunto  $\mathcal{E}$  de  $\mathbb{R}^{2+k}$  que es un cono convexo. Vale decir: si  $(U, V, \mathbf{X}) \in \mathcal{E}$  y  $a > 0$  entonces también  $(aU, aV, a\mathbf{X}) \in \mathcal{E}$  y, además, si  $(U_j, V_j, \mathbf{X}^{(j)}) \in \mathcal{E}$  para  $j = 1, 2$  y  $0 < \lambda < 1$  entonces  $(\lambda U_1 + (1 - \lambda)U_2, \lambda V_1 + (1 - \lambda)V_2, \lambda \mathbf{X}^{(1)} + (1 - \lambda)\mathbf{X}^{(2)}) \in \mathcal{E}$ . Por ejemplo, típicamente, para un fluido puro (una sola componente)  $(U, V, N) \in \mathbb{R}_+^3 \equiv (0, \infty)^{\times 3}$ .

En cuanto a las propiedades de la función entropía  $(U, V, \mathbf{X}) \mapsto S(U, V, \mathbf{X})$  consideraremos solamente dos de las muchas variantes que se postulan en base a evidencia empírica.

La  $n$ -aditividad postula que si se consideran  $n$  copias idénticas de un sistema mutuamente aisladas entre si de entropía  $S(U, V, \mathbf{X})$  cada una, el sistema de  $n$  copias aisladas tiene entropía  $nS(U, V, \mathbf{X})$  y, si se levantan los aislamientos y estas copias se unen en un sólo sistema sin alterar el resto del universo entonces la entropía del sistema compuesto es  $S(nU, nV, n\mathbf{X})$  e igual a  $nS(U, V, \mathbf{X})$  (véase la figura 1).

Decimos que se tiene  $p$ -aditividad si considerando un sistema de entropía  $S(U, V, \mathbf{X})$  que se

---

<sup>3</sup>Se puede demostrar que  $\mathcal{K} \subset \mathbb{R}^d$  es un cono convexo si y sólo si  $\mathbf{x}, \mathbf{y} \in \mathcal{K}$  implica que  $\mathbf{x} + \mathbf{y} \in \mathcal{K}$ .

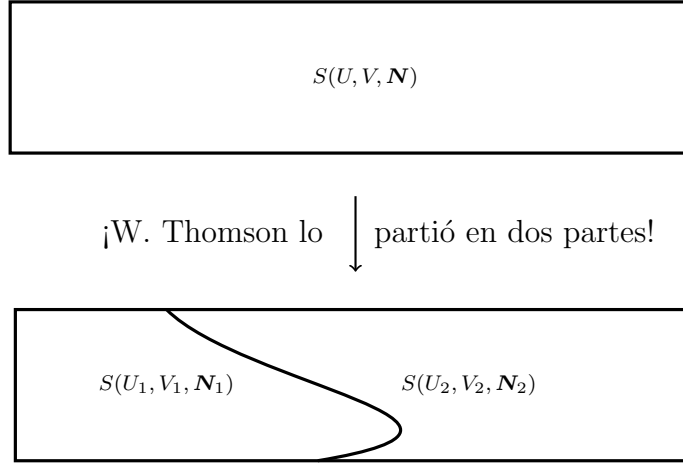


Figura 2: p-aditividad.  $\frac{U_1}{U_2} = \frac{V_1}{V_2} = \frac{N_{1j}}{N_{2j}}, j = 1, 2, \dots, r$

secciona sin alterar el resto del universo en dos subsistemas que no interactúan de entropías  $S(U_1, V_1, \mathbf{X}^{(1)})$  y  $S(U_2, V_2, \mathbf{X}^{(2)})$  respectivamente de tal modo que  $U = U_1 + U_2$ ,  $V = V_1 + V_2$ ,  $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$  con  $U_1/U_2 = V_1/V_2 = X_k^{(1)}/X_k^{(2)}$  (donde  $k$  recorre los índices de las variables extensivas adicionales) entonces se tiene  $S(U_1, V_1, \mathbf{X}^{(1)}) + S(U_2, V_2, \mathbf{X}^{(2)}) = S(U, V, \mathbf{X})$  (vease la figura 2).

Reemplazo los postulados II y III de Callen por:

**Postulado II<sub>o</sub>:** Para todo sistema compuesto existe una función de estado (llamada entropía)  $S$  de las variables extensivas completas que tiene las siguientes dos propiedades:

- $S$  es o bien  $n$ -aditiva para todo  $n \in \mathbb{N}$  o bien  $S$  es  $p$ -aditiva.
- En todo proceso natural en el cual un sistema completamente aislado pasa de un estado de equilibrio a otro estado de equilibrio, la entropía no disminuye.

Además,  $S$  es no-negativa, continuamente diferenciable y  $(\frac{\partial S}{\partial U})_{\mathbf{X}} > 0$ .

Con este postulado demostramos homogeneidad y concavidad de la entropía. La no-negatividad postulada no se usa.

### De la aditividad a la homogeneidad

Haciendo referencia a la figura 1, la entropía total de  $n$  copias idénticas del mismo sistema en el mismo estado cuando estas copias están mutuamente aisladas entre si, es  $nS(U, V, \mathbf{N})$ . La  $n$ -aditividad postula que al remover las divisiones aislantes entre las copias la entropía no

cambia<sup>4</sup>:

$$S(nU, nV, n\mathbf{M}) = nS(U, V, \mathbf{N}) , \quad n = 2, 3, \dots .$$

Esto es la homogeneidad de grado 1 para enteros positivos. Con el Lema 2 (Apéndice 6) obtenemos la homogeneidad de grado 1 de  $S$ .

La p-aditividad en cambio, postula (figura 2) que al partir al sistema en dos partes aisladas entre si de modo que  $U = U_1 + U_2$ ,  $V = V_1 + V_2$ , y  $\mathbf{N} = \mathbf{N}_1 + \mathbf{N}_2$  y las razones de las variables extensivas correspondientes sean constantes  $U_1/U_2 = V_1/V_2 = N_{11}/N_{21} = N_{12}/N_{22} = \dots$  se tiene

$$S(U_1, V_1, \mathbf{N}_1) + S(U_2, V_2, \mathbf{N}_2) = S(U_1 + U_2, V_1 + V_2, \mathbf{N}_1 + \mathbf{N}_2) .$$

La n-aditividad se desprende de la p-aditividad inmediatamente. Y esto conduce a la homogeneidad de grado 1. Hemos obtenido el

*Primer resultado:  $S$  es una función homogénea de grado 1 en las variables extensivas.:*

$$S(aU, aV, a\mathbf{N}) = a S(U, V, \mathbf{N}) , \quad a > 0 .$$

#### Del no-crecimiento y la homogeneidad a la concavidad

Considero –refiriendo a la figura 3– dos sistemas del mismo fluido uno en el estado de equilibrio  $(U_1, V_1, \mathbf{N}_1)$  y el otro en el estado de equilibrio  $(U_2, V_2, \mathbf{N}_2)$  separados por una compuerta que aísla a una parte de la otra. La entropía del sistema compuesto es entonces  $S(U_1, V_1, \mathbf{N}_1) + S(U_2, V_2, \mathbf{N}_2)$  donde  $S$  es la función de estado correspondiente al fluido usado. Al sacar la compuerta, se inicia un proceso en general irreversible! Al alcanzar el equilibrio la energía interna es  $U_1 + U_2$ , el volumen es  $V_1 + V_2$ , y la composición es  $\mathbf{N}_1 + \mathbf{N}_2$  ya que todo está aislado del resto del universo.

Ahora, por la propiedad de no-decrecimiento,

$$S(U_1 + U_2, V_1 + V_2, \mathbf{N}_1 + \mathbf{N}_2) \geq S(U_1, V_1, \mathbf{N}_1) + S(U_2, V_2, \mathbf{N}_2) ;$$

pero, con la homogeneidad de grado 1,

$$\begin{aligned} S(U_1 + U_2, V_1 + V_2, \mathbf{N}_1 + \mathbf{N}_2) &= S\left(2 \frac{U_1 + U_2}{2}, 2 \frac{V_1 + V_2}{2}, 2 \frac{\mathbf{N}_1 + \mathbf{N}_2}{2}\right) \\ &= 2 S\left(\frac{U_1 + U_2}{2}, \frac{V_1 + V_2}{2}, \frac{\mathbf{N}_1 + \mathbf{N}_2}{2}\right) , \end{aligned}$$

y por ende

$$S\left(\frac{U_1 + U_2}{2}, \frac{V_1 + V_2}{2}, \frac{\mathbf{N}_1 + \mathbf{N}_2}{2}\right) \geq \frac{1}{2} S(U_1, V_1, \mathbf{N}_1) + \frac{1}{2} S(U_2, V_2, \mathbf{N}_2) ;$$

---

<sup>4</sup>Esto puede verse también como una propiedad de los estados de equilibrio. Un estado de equilibrio no se altera si uno piensa un dado fluido en equilibrio como constituido por dos partes divididas por una pared aislante imaginaria.

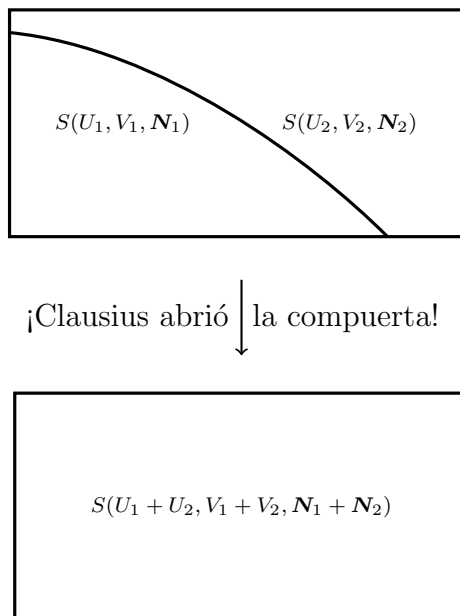


Figura 3: ¡Una especie inversión de la p-aditividad!

que es la concavidad en el punto medio. Con el Lema 4 (Apéndice 7) obtenemos el

*Segundo resultado:  $S$  es cóncava como función de las variables extensivas:*

$$\begin{aligned}
 & S(\lambda U_1 + (1 - \lambda)U_2, \lambda V_1 + (1 - \lambda)V_2, \lambda \mathbf{N}_1 + (1 - \lambda)\mathbf{N}_2) \\
 & \geq \lambda S(U_1, V_1, \mathbf{N}_1) + (1 - \lambda)S(U_2, V_2, \mathbf{N}_2), \quad 0 \leq \lambda \leq 1.
 \end{aligned}$$

Quiero recalcar que tanto en la n-aditividad como en la p-aditividad se parte de un sistema en equilibrio y se lo secciona (por ejemplo imaginariamente) en subsistemas *de modo de no lacerar el equilibrio*. En cambio en la situación de la figura 3, se parte de dos sistemas separados (del mismo material) cada uno en equilibrio por sí sólo y se postula que pasa cuando se levanta la separación. Los procesos de separación o unión son totalmente idealizados ya que no pueden aportar energía alguna. De esto se encargan los Sres. Joule, W. Thomson y Clausius que tienen atributos divinos.

Consideramos estados de equilibrio dados por puntos  $(U, \mathbf{X})$  en un subconjunto convexo  $\mathcal{K}$  de  $\mathbb{R} \times \mathbb{R}^n$  donde  $U$  –que es la energía interna– y las variables  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  –que son extensivas– alcanzan para la caracterización termodinámica del sistema. Vimos que la función entropía  $(U, \mathbf{X}) \mapsto S(U, \mathbf{X})$  es homogénea de grado 1 y cóncava. Además,

$$\frac{1}{T} := \left( \frac{\partial S}{\partial U} \right)_{\mathbf{X}} > 0,$$

de modo que la entropía es estrictamente creciente en la energía interna al dejar las demás variables extensivas fijas. La positividad de  $T$  es también todo lo que necesitamos para aplicar

el Teorema de la Función Inversa a la función  $F : \mathcal{K} \rightarrow \mathbb{R} \times \mathbb{R}^n$  definida por  $F(U, \mathbf{X}) = (S(U, \mathbf{X}), \mathbf{X})$  cuyo jacobiano es ( $\nabla$  es el gradiente respecto de las variables  $\mathbf{X}$  manteniendo a  $U$  constante)

$$\left( \begin{array}{c|ccc} 1/T = (\partial S/\partial U)_{\mathbf{X}} & \nabla S & & \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \cdots & 1 \end{array} \right) ;$$

cuya determinante es  $(\partial S/\partial U)_{\mathbf{X}} > 0$ . Con este teorema hay una función inversa  $G$  que aplica la imagen de  $\mathcal{K}$  bajo  $F$  en  $\mathcal{K}$  tal que su jacobiano es

$$\left( \begin{array}{c|ccc} T = 1/(\partial S/\partial U)_{\mathbf{X}} & -T(\nabla S) & & \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \cdots & 1 \end{array} \right) .$$

$G$  define la función  $U : (S, \mathbf{X}) \mapsto G_1(S, \mathbf{X})$  tal que:

$$(1-1) \quad S(G_1(Z, \mathbf{X}), \mathbf{X}) = Z \quad , \quad U(S(Y, \mathbf{X}), \mathbf{X}) = Y .$$

Se tiene cuando  $a > 0$ , que

$$U(aS(Y, \mathbf{X}), a\mathbf{X}) = U(S(aY, a\mathbf{X}), a\mathbf{X}) = aY = aU(S(Y, \mathbf{X}), \mathbf{X}) ,$$

debido a la homogeneidad de grado 1 de la función  $S$ . Esto verifica que  $(S, \mathbf{X}) \mapsto U(S, \mathbf{X})$  es homogénea de grado 1.

Para demostrar la convexidad, considero dos valores  $S_1$  y  $S_2$  de la entropía, dos vectores fijos  $\mathbf{X}^{(1)}$  y  $\mathbf{X}^{(2)}$ , y sus correspondientes valores de energía interna  $U_1$  y  $U_2$  tales que

$$(1-2) \quad S_1 = S(U_1, \mathbf{X}^{(1)}) , \quad S_2 = S(U_2, \mathbf{X}^{(2)}) , \quad U_1 = U(S_1, \mathbf{X}^{(1)}) , \quad U_2 = U(S_2, \mathbf{X}^{(2)}) .$$

Ya que la entropía es cóncava tenemos para  $0 < \lambda < 1$  que

$$\lambda S_1 + (1 - \lambda)S_2 = \lambda S(U_1, \mathbf{X}^{(1)}) + (1 - \lambda)S(U_2, \mathbf{X}^{(2)}) \leq S(\lambda U_1 + (1 - \lambda)U_2, \lambda \mathbf{X}^{(1)} + (1 - \lambda)\mathbf{X}^{(2)}) .$$

Y, como a  $\mathbf{X}$  fijo el mapa  $S \mapsto U(S, \mathbf{X})$  es creciente pues  $T = (\partial U/\partial S)_{\mathbf{X}} > 0$ , tengo que

$$\begin{aligned} & U(\lambda S_1 + (1 - \lambda)S_2, \lambda \mathbf{X}^{(1)} + (1 - \lambda)\mathbf{X}^{(2)}) \\ & \leq U(\underbrace{S(\lambda U_1 + (1 - \lambda)U_2)}_{=Y}, \underbrace{\lambda \mathbf{X}^{(1)} + (1 - \lambda)\mathbf{X}^{(2)}}_{=\mathbf{X}}), \underbrace{\lambda \mathbf{X}^{(1)} + (1 - \lambda)\mathbf{X}^{(2)}}_{=\mathbf{X}} ; \end{aligned}$$

pero el miembro derecho de esta desigualdad es  $U(S(Y, \mathbf{X}), \mathbf{X})$  que por (1-1) es igual a  $Y = \lambda U_1 + (1 - \lambda)U_2$ . Por lo tanto, usando (1-2),

$$U(\lambda S_1 + (1 - \lambda)S_2, \lambda \mathbf{X}^{(1)} + (1 - \lambda)\mathbf{X}^{(2)})$$

$$\leq \lambda U_1 + (1 - \lambda)U_2 = \lambda U(S_1, \mathbf{X}^{(1)}) + (1 - \lambda)U(S_2, \mathbf{X}^{(2)}) .$$

Esto demuestra que

$U$  es convexa y homogénea de grado 1 en sus variables naturales  $(S, \mathbf{X})$  .

De la convexidad y homogeneidad se obtiene la sub-aditividad:

$$U(S_1 + S_2, \mathbf{X}_1 + \mathbf{X}_2) \leq U(S_1, \mathbf{X}_1) + U(S_2, \mathbf{X}_2) .$$



## 2. Potenciales termodinámicos

Generalmente los potenciales termodinámicos se introducen por medio de la transformación de Legendre de funciones diferenciables. Las hipótesis necesarias para ello (diferenciabilidad y inyectividad de las derivadas) no se explicitan usualmente quedando tacitas o implícitas.

Aquí introducimos los potenciales termodinámicos como transformadas de Legendre-Fenchel de  $U(S, V, \mathbf{X})$  respecto de las variables básicas  $(S, V)$ . Los potenciales surgen de un problema variacional que no necesita de hipótesis alguna sobre la función  $U$  y que permiten demostrar inmediatamente las propiedades de convexidad/concavidad y de homogeneidad pertinentes. Comentamos sobre la equivalencia de ambas maneras de introducir los potenciales en el apéndice 7.

1. Dada una función  $f$  a valores reales definida en un conjunto  $K \subset \mathbb{R}^n$ , su **transformada de Legendre-Fenchel** es la función

$$(2-1) \quad f^\#(\mathbf{y}) := \sup\{\mathbf{x} \cdot \mathbf{y} - f(\mathbf{x}) : \mathbf{x} \in K\}, \quad \mathbf{y} \in \mathbb{R}^n.$$

$f^\#$  puede tomar el valor  $+\infty$  (e.g., la transformada de Legendre-Fenchel de la función  $\mathbf{x} \mapsto \alpha \mathbf{x}$ ) pero no toma el valor  $-\infty$  ni es idénticamente  $+\infty$ . Obsérvese que no se hacen hipótesis adicionales sobre  $K$  y tampoco sobre  $f$ .

**Lema 1**  $f^\#$  es convexa.

Demostración: Si  $0 \leq \lambda \leq 1$ ,

$$\begin{aligned} f^\#(\lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2) &= \sup_{\mathbf{x} \in K} \{\lambda [\mathbf{y}_1 \cdot \mathbf{x} - f(\mathbf{x})] + (1 - \lambda) [\mathbf{y}_2 \cdot \mathbf{x} - f(\mathbf{x})]\} \\ &\leq \sup_{\mathbf{x} \in K} \{\lambda [\mathbf{y}_1 \cdot \mathbf{x} - f(\mathbf{x})]\} + \sup_{\mathbf{x} \in K} \{(1 - \lambda) [\mathbf{y}_2 \cdot \mathbf{x} - f(\mathbf{x})]\} = \lambda f^\#(\mathbf{y}_1) + (1 - \lambda) f^\#(\mathbf{y}_2). \end{aligned}$$

La desigualdad es consecuencia de que el supremo de una suma de funciones es menor o igual a la suma de los supremos de cada función. Esto prueba la afirmación cuando  $f^\#(\mathbf{y}_1)$  y  $f^\#(\mathbf{y}_2)$  son ambos finitos y también cuando  $f^\#(\mathbf{y}_1)$  o  $f^\#(\mathbf{y}_2)$  son  $+\infty$ .

Hemos definido a  $f^\#$  en  $\mathbb{R}^n$ . En termodinámica interesa en general la función  $f^\#$  restringida a un conjunto convexo de  $\mathbb{R}^n$ ; más precisamente a un cono convexo como  $\mathbb{R}_+^n = (0, \infty)^{n \times n}$ . Más sobre la transformada de Legendre-Fenchel (L-F) en el apéndice sobre funciones convexas (apéndice 7).

2. Definimos la **energía libre de Helmholtz**  $F(T, V, \mathbf{X})$  como menos la transformada de L-F de  $S \mapsto U(S, V, \mathbf{X})$ :

$$F(T, V, \mathbf{X}) = - \sup_{S > 0} \{TS - U(S, V, \mathbf{X})\} = \inf_{S > 0} \{U(S, V, \mathbf{X}) - TS\}, \quad T > 0.$$

En virtud de su definición como menos la transformada de L-F de  $U$  con respecto a  $S$ ,  $F$  es cóncava en  $T$ .

Para  $a > 0$  tenemos –en virtud de la homogeneidad de  $U$ – que

$$\begin{aligned} F(T, aV, a\mathbf{X}) &= \inf_{S>0} \{U(S, aV, a\mathbf{X}) - TS\} = \inf_{S>0} \{aU(a^{-1}S, V, \mathbf{X}) - TS\} \\ &= a \inf_{S>0} \{U(a^{-1}S, V, \mathbf{x}) - T(a^{-1}S)\} = a \inf_{Z>0} \{U(Z, V, \mathbf{X}) - TZ\} = aF(T, V, \mathbf{x}) \end{aligned}$$

lo que demuestra que  $(V, \mathbf{X}) \mapsto F(T, V, \mathbf{X})$  es homogénea de grado 1.

Veamos que  $F$  es convexa en sus variables extensivas. Si  $0 < \lambda < 1$ , tenemos

$$\begin{aligned} F(T, \lambda V_1 + (1 - \lambda)V_2, \lambda\mathbf{X}_1 + (1 - \lambda)\mathbf{X}_2) &= \inf_S \{U(S, \lambda V_1 + (1 - \lambda)V_2, \lambda\mathbf{X}_1 + (1 - \lambda)\mathbf{X}_2) - TS\} \\ &= \inf_{S_1, S_2} \{U(\lambda S_1 + (1 - \lambda)S_2, \lambda V_1 + (1 - \lambda)V_2, \lambda\mathbf{X}_1 + (1 - \lambda)\mathbf{X}_2) - T[\lambda S_1 + (1 - \lambda)S_2]\} \\ &\leq \inf_{S_1, S_2} \{\lambda[U(S_1, V_1, \mathbf{X}_1) - TS_1] + (1 - \lambda)[U(S_2, V_2, \mathbf{X}_2) - TS_2]\} \\ &= \inf_{S_1} \{\lambda[U(S_1, V_1, \mathbf{X}_1) - TS_1]\} + \inf_{S_2} \{(1 - \lambda)[U(S_2, V_2, \mathbf{X}_2) - TS_2]\} \\ &= \lambda F(T, V_1, \mathbf{X}_1) + (1 - \lambda)F(T, V_2, \mathbf{X}_2), \end{aligned}$$

en virtud de la convexidad de  $U$  en sus variables naturales.

$$\boxed{T \mapsto F(T, V, \mathbf{X}) \text{ es cóncava a } (V, \mathbf{X}) \text{ fijo}}$$

$$\boxed{(V, \mathbf{X}) \mapsto F(T, V, \mathbf{X}) \text{ es convexa y homogénea de grado 1 a } T \text{ fijo}}$$

Observamos que el ínfimo en la definición de  $F$  se asume cuando

$$(2-2) \quad T = \left( \frac{\partial U}{\partial S} \right)_{V, \mathbf{X}}$$

de modo que la  $T$  en el argumento de  $F$  es la temperatura termodinámica y podemos entonces escribir

$$F(T, V, \mathbf{X}) = U(S(V, T, \mathbf{X}), V, \mathbf{X}) - TS(V, T, \mathbf{X}),$$

donde  $S(V, T, \mathbf{X})$  es la solución de la ecuación diferencial (2-2).

Suponiendo ahora (y no antes) que  $F$  es diferenciable, de la concavidad de  $F$  como función de  $T$  y de  $(\partial F / \partial T)_{V, \mathbf{X}} = -S$  obtenemos  $C_v \geq 0$ , ya que

$$0 \geq \left( \frac{\partial^2 F}{\partial T^2} \right)_{V, \mathbf{X}} = - \left( \frac{\partial S}{\partial T} \right)_{V, \mathbf{X}} = - \frac{C_v}{T}.$$

De la convexidad de  $F$  como función de  $V$  y de  $(\partial F / \partial V)_{T, \mathbf{X}} = -p$  obtenemos positividad de la compresibilidad isotérmica

$$\kappa_T := - \frac{1}{V} \left( \frac{\partial V}{\partial p} \right)_{T, \mathbf{X}} \geq 0,$$

ya que

$$0 \leq \left( \frac{\partial^2 F}{\partial V^2} \right)_{T, \mathbf{X}} = - \left( \frac{\partial p}{\partial V} \right)_{T, \mathbf{X}} = \frac{1}{V \kappa_T} .$$

3. La **entalpía**  $H(p, S, \mathbf{X})$  es menos la transformada de L-F de  $U$  respecto de  $V$ :

$$H(S, p, \mathbf{X}) = - \sup_V \{-pV - U(S, V, \mathbf{X})\} = \inf_S \{U(S, V, \mathbf{X}) + pV\} , \quad p > 0 .$$

En virtud de su definición  $H$  es cóncava en  $p$ . Y exactamente igual a como lo hicimos con  $F$ , se demuestra que  $H$  es convexa y homogénea de grado 1 en sus variables extensivas  $(S, \mathbf{X})$ .

$$\boxed{p \mapsto H(S, p, \mathbf{X}) \text{ es cóncava a } (S, \mathbf{X}) \text{ fijo}}$$

$$\boxed{(S, \mathbf{X}) \mapsto H(S, p, \mathbf{X}) \text{ es convexa y homogénea de grado 1 a } p \text{ fijo}}$$

Tenemos

$$H(S, p, \mathbf{X}) = U(S, V(S, p, \mathbf{X}), \mathbf{X}) + pV(S, p, \mathbf{X}) , \quad \left( \frac{\partial U}{\partial V} \right)_{S, \mathbf{X}} = -p .$$

Del mismo modo que en el caso de  $F$ , obtenemos positividad de la compresibilidad adiabática<sup>5</sup>

$$\kappa_S := - \frac{1}{V} \left( \frac{\partial V}{\partial p} \right)_{S, \mathbf{X}} \geq 0 ,$$

de la concavidad en  $p$  de  $H$ ; y la positividad de la capacidad calorífica a presión constante (o del calor específico a presión constante) a partir de la convexidad de  $H$  en  $S$ .

4. La **energía libre de Gibbs** es menos la transformada de L-F de  $U$  con respecto a las variables  $(S, V)$ :

$$G(T, p, \mathbf{X}) = - \sup_{V, S} \{-pV + TS - U(S, V, \mathbf{X})\} = \inf_{V, S} \{U(S, V, \mathbf{X}) + pV - TS\} , \quad T, p > 0 .$$

En virtud de su definición  $G$  es cóncava en  $(T, p)$ . Y como lo hicimos con  $F$  se demuestra que  $G$  es convexa y homogénea de grado 1 en sus variables extensivas  $\mathbf{X}$ .

$$\boxed{(T, p) \mapsto G(T, p, \mathbf{X}) \text{ es cóncava a } \mathbf{X} \text{ fijo}}$$

$$\boxed{X \mapsto G(T, p, X) \text{ es convexa y homogénea de grado 1 a } (T, p) \text{ fijo}}$$

Tenemos

$$G(T, p, \mathbf{X}) = U(S(T, p, \mathbf{X}), V(T, p, \mathbf{X}), \mathbf{X}) + pV(T, p, \mathbf{X}) - TS(T, p, \mathbf{X}) .$$

---

<sup>5</sup> $\kappa_S = c_v \kappa_T / c_p$ .

Nuevamente la concavidad de  $G$  en  $p$  y  $T$  por separado implica la positividad de  $\kappa_T$  y de  $C_p$ . Pero también, por la concavidad en  $(p, T)$ , la forma cuadrática asociada a

$$\begin{pmatrix} \left( \frac{\partial^2 G}{\partial p^2} \right)_{T, \mathbf{X}} & , & \left( \frac{\partial}{\partial p} \left( \frac{\partial G}{\partial T} \right)_{p, \mathbf{X}} \right)_{T, \mathbf{X}} \\ \left( \frac{\partial}{\partial T} \left( \frac{\partial G}{\partial p} \right)_{T, \mathbf{X}} \right)_{p, \mathbf{X}} & , & \left( \frac{\partial^2 G}{\partial T^2} \right)_{p, \mathbf{X}} \end{pmatrix}$$

es negativa semi-definida con lo cual expresando las derivadas en términos de  $C_p$ ,  $\kappa_T$  y el coeficiente de expansión térmica

$$\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_{p, \mathbf{X}} ,$$

se obtiene

$$\frac{C_p \kappa_T}{VT} \geq \alpha^2 .$$

Esta última desigualdad se obtiene también de la convexidad de  $(S, V) \mapsto U(S, V, \mathbf{X})$ .

5. De la convexidad y la homogeneidad de grado 1 en sus variables extensivas, se obtiene la subaditividad de todos los potenciales  $F$ ,  $H$  y  $G$  en sus variables extensivas:

$$F(T, V_1, \mathbf{X}_1) + F(T, V_2, \mathbf{X}_2) \geq F(T, V_1 + V_2, \mathbf{X}_1 + \mathbf{X}_2) ;$$

$$H(S_1, p, \mathbf{X}_1) + H(S_2, p, \mathbf{X}_2) \geq H(S_1 + S_2, p, \mathbf{X}_1 + \mathbf{X}_2) ;$$

$$G(T, p, \mathbf{X}_1) + G(T, p, \mathbf{X}_2) \geq G(T, p, \mathbf{X}_1 + \mathbf{X}_2) .$$

### 3. Principios variacionales para los potenciales termodinámicos

Recordemos el Principio de Entropía maximal: *A energía interna fija, los valores de las variables extensivas no fijadas son tales que maximizan la entropía para esa energía interna y los valores fijados de las variables extensivas.*

Si  $(U_o, \mathbf{X}^{(o)})$  es un punto estacionario de la entropía con  $U_o$  y ciertas  $X_j^{(o)}$  ( $j \in J \subset \{1, 2, \dots, n\}$ ) fijas – vale decir  $(\partial S / \partial X_k)(U_o, \mathbf{X}^{(o)}) = 0$  para  $k \in \{1, 2, \dots, n\} \setminus J$ – entonces  $(U_o, \mathbf{X}^{(o)})$  es un máximo global de la entropía  $S(U, \mathbf{X})$  a  $U = U_o$  fijo y  $X_j = X_j^{(o)}$  ( $j \in J$ ) fijas. Mostramos que con  $S_o = S(U_o, \mathbf{X}^{(o)})$  (que es el valor máximo sujeto a los vínculos explicitados), tenemos que  $U(S_o, \mathbf{X}^{(o)})$  es un mínimo global de  $U(S, \mathbf{X})$  sujeta a los vínculos  $S = S_o$  y  $X_j = X_j^{(o)}$  para  $j \in J$ . Tenemos, para  $k \notin J$

$$\left( \frac{\partial U}{\partial X_k} \right)_{U=U_o, X_j=X_j^{(o)}, etc} = - \frac{\left( \frac{\partial S}{\partial X_k} \right)_{U=U_o, X_j=X_j^{(o)}, etc}}{\left( \frac{\partial S}{\partial U} \right)_{X_j=X_j^{(o)}, etc}} = -T \left( \frac{\partial S}{\partial X_k} \right)_{U=U_o, X_j=X_j^{(o)}, etc} ,$$

lo que se anula evaluando en el punto estacionario ( $U_o, \mathbf{X}^{(o)}$ ). Pero por la convexidad de  $U$  este punto estacionario corresponde necesariamente a un mínimo global. También podemos argumentar a la inversa partiendo de un punto estacionario de  $U$  para obtener uno de  $S$ . Esto prueba la equivalencia del Principio de Entropía maximal y el Principio de Energía Interna Mínima: *A entropía fija, los valores de las variables extensivas no fijadas son tales que minimizan la energía interna para esa entropía y los valores fijados de las variables extensivas.*

De esto se desprenden inmediatamente los principios extremales de los potenciales termodinámicos  $F(T, V, \mathbf{X})$ ,  $H(S, p, \mathbf{X})$  y  $G(T, p, \mathbf{X})$  a saber: los valores de las variables extensivas no fijadas son tales que minimizan el valor del potencial termodinámico cuando se fijan las variables intensivas correspondientes (i.e.,  $T$  para  $F$ ,  $p$  para  $H$  y  $(T, p)$  para  $G$ ) y, posiblemente, alguna variable extensiva.

En todas las instancias de aplicación de uno de los principios variacionales, los valores de las magnitudes no constreñidas son únicos si la propiedad de convexidad/concavidad pertinente es estricta. El estado de equilibrio predicho es único. Sino, los valores que garantizan la ‘extremalidad’ conforman un conjunto convexo y se está en presencia de un fenómeno de transición de fases o análogo. Esto se discute con muy buen anclaje físico en S. Cannas: *Notas de Mecánica Estadística*. Editorial de la UNC, Córdoba 2018.

### 3.1. Trabajo máximo: algunas acotaciones

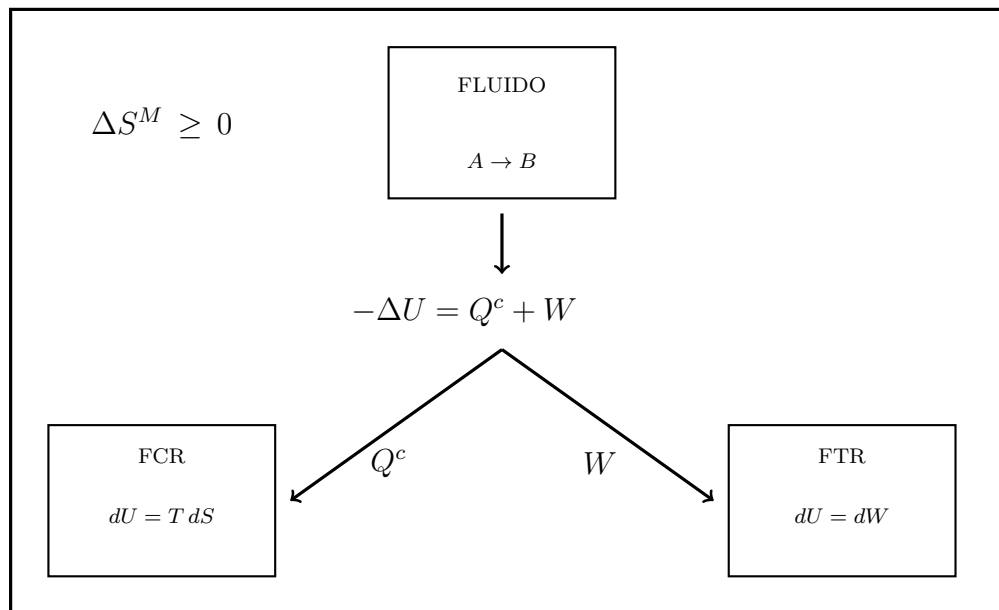


Figura 4: La maquina

Anoto aquí algunas cuestiones aclaratorias al tratamiento del llamado *Teorema del Trabajo Máximo* como se presenta en el libro de H.B. Callen y en el libro de S. Cannas mencionado.

Las maquinas  $M$  que considero son sistemas (termodinámicos) compuestos por tres (sub-) sistemas: un sistema termodinámico (específicamente un fluido simple) que realiza un proceso de un estado de equilibrio  $A$  a un estado de equilibrio  $B$ ; una fuente de calor reversible (FCR), y una fuente de trabajo reversible (FTR). Sobre las fuentes no tengo nada que agregar a lo presentado en las notas de Cannas. Es crucial que para la FCR, en tanto y en cuanto está mecánicamente aislada y sólo puede transferir calor, se tiene  $dU = TdS$ . Mientras que para la FTR se tiene  $dS = 0$ .

Supongo que  $\Delta U = U_B - U_A$  es dado. La argumentación que sigue es independiente del signo de  $\Delta U$ . La energía libre disponible para las fuentes es entonces  $-\Delta U$  que se reparte en una transferencia de calor  $Q^c$  del fluido a la FCR y un trabajo realizado  $W$  sobre la FTR de modo que (por la conservación de la energía interna válida para  $M$ ):

$$-\Delta U = Q^c + W .$$

El supra-índice  $c$  refiere a la FCR. Téngase en cuenta que desde el punto de vista del fluido, este realiza trabajo  $-W$  y transfiere calor  $-Q$ . Con nuestra convención sobre los signos,  $W > 0$  significa que el fluido realiza trabajo sobre la FTR y  $Q^c > 0$  significa que el fluido pierde energía interna en forma de calor que es transferida a la FCR.

El balance entrópico para  $M$  es entonces el siguiente.

$$0 \leq \Delta S^M = \Delta S^c + \Delta S ,$$

con igualdad si el proceso  $A \rightarrow B$  es reversible ya que las fuentes son ambas reversibles. Se tiene  $\Delta S = S_B - S_A$  y

$$(3-1) \quad \Delta S^c = \Delta S^M - \Delta S \geq -\Delta S ,$$

con igualdad si el proceso de base es reversible. El balance energético es:

$$\Delta U^M = 0 , \quad Q^c = \Delta U^c = \int_{S_0^c}^{S_0^c + \Delta S^c} T^c(\zeta, V^c, \mathbf{X}) d\zeta , \quad W = -Q^c - \Delta U = -Q^c - (U_B - U_A) .$$

Aquí,  $T^c$  es la función de la entropía, el volumen de la FCR, y otras posibles variables extensivas que da la temperatura de la FCR<sup>6</sup>.  $S_0^c$  es la entropía del la FCR para la situación inicial donde  $A$  es el estado del fluido y no hay intercambio alguno con las fuentes. Claramente  $W$  será maximal si  $Q^c$  es minimal. Ahora, usando (3-1),

$$Q^c = \Delta U^c = \int_{S_0^c}^{S_0^c + \Delta S^c} T^c(\zeta, V^c, \mathbf{X}) d\zeta$$

---

<sup>6</sup>Estamos suponiendo que el proceso de transferencia de calor en la FCR no altera las variables extensivas  $(V, \mathbf{X})$  que es parte de la idealización que conlleva la caracterización de una FCR.

$$= \int_{S_o^c}^{S_o^c - \Delta S + \Delta S^M} T^c(\zeta, V^c, \mathbf{X}) d\zeta = \int_{S_o^c}^{S_o^c - \Delta S} T^c(\zeta, V^c, \mathbf{X}) d\zeta + \int_{S_o^c - \Delta S}^{S_o^c - \Delta S + \Delta S^M} T^c(\zeta, V^c, \mathbf{X}) d\zeta .$$

Pero la primera integral (de  $S_o^c$  hasta  $S_o^c - \Delta S$ ) coincide con el calor  $Q^c$  en el caso reversible ( $\Delta S^c = -\Delta S$ ):

$$(3-2) \quad Q_{rev}^c = \int_{S_o^c}^{S_o^c - \Delta S} T^c(\zeta, V^c, \mathbf{X}) d\zeta ;$$

y para la segunda integral (de  $S_o^c - \Delta S$  hasta  $S_o^c - \Delta S + \Delta S^M$ ) el Teorema del Valor Medio de la Integración nos dice que

$$\int_{S_o^c - \Delta S}^{S_o^c - \Delta S + \Delta S^M} T^c(\zeta, V^c, \mathbf{X}) d\zeta = T^c(\zeta_o, V^c, \mathbf{X}) \Delta S^M ,$$

con  $\zeta_o$  un valor de la entropía de la FCR intermedio entre los límites de integración. Ya que la temperatura (termodinámica) es positiva, esta última integral es, por (3-1), no-negativa y nula si y sólo si  $\Delta S^M = 0$ , o sea en el caso reversible<sup>7</sup>. Hay un incremento no-negativo de  $Q^c$  respecto de  $Q_{rev}^c$ :

$$(3-3) \quad Q^c - Q_{rev}^c \geq 0 ,$$

con igualdad para el caso reversible. Se desprende de esto que  $Q_{rev}^c$  es entonces minimal y, en el caso reversible  $W$  es maximal e igual a

$$(3-4) \quad W_{max} := -Q_{rev}^c + U_A - U_B .$$

$$(3-5) \quad W \leq W_{max} ,$$

donde hay igualdad si y sólo si el proceso es reversible. Las ecuaciones (3-5), (3-3) complementadas con (3-4) y (3-2) constituyen el llamado *Teorema del Trabajo Máximo*.

Observe ahora que el signo de  $Q_{rev}^c$  está fijado por el signo de  $\Delta S$ , ya que el integrando en (3-2) es positivo:

$$\Delta S > 0 \iff Q_{rev}^c < 0 ; \quad \Delta S < 0 \iff Q_{rev}^c > 0 .$$

Con esto obtenemos que:

$$W_{max} > -\Delta U \text{ si } \Delta S > 0 ; \quad W_{max} < -\Delta U \text{ si } \Delta S < 0 .$$

Si  $\Delta U < 0$  y  $\Delta S > 0$  entonces  $W_{max} > 0$  o sea que si el proceso  $A \rightarrow B$  es reversible el fluido realiza trabajo vía la FTR y “absorbe” calor de la FCR ( transformación de calor en trabajo).

<sup>7</sup>Podemos argumentar directamente: el integrando es positivo y el límite superior de integración es mayor o igual al límite inferior de integración. Luego la integral es no-negativa anulándose si y sólo si  $\Delta S^M = 0$ .

Si, en cambio  $\Delta U > 0$  y  $\Delta S < 0$  entonces en el caso reversible el fluido entrega calor a la FCR y la FTR realiza trabajo sobre el fluido (transformación de trabajo en calor). En los otros dos casos, i.e.  $\Delta U < 0$  con  $\Delta S < 0$  resp.  $\Delta U > 0$  con  $\Delta S > 0$ , habrá transferencia de calor del fluido a la FCR (resp. al revés) pero no podemos inferir nada a priori sobre el signo de  $W_{max}$  que dependerá de la magnitud relativa de  $Q_{rev}^c$  y  $\Delta U$  que tienen distinto signo.

Por el Teorema del Valor Medio de la integral,

$$Q^c = T^c(\zeta', V, \mathbf{X}) \Delta S^c$$

donde  $\zeta'$  es un valor de la entropía de la FCR intermedio entre  $S_o^c$  y  $S_o^c + \Delta S^c$ .

## 4. El gas perfecto puro

Para un gas puro (una sola componente química) que satisface la leyes de Boyle y de Joule:

- (Ley de Boyle) A temperatura constante, el producto  $pV$  es constante;
- (Ley de Joule) La energía interna depende solamente de la temperatura.

En ciertos experimentos de Joule, él estudió gases que realizan una expansión libre en la cual no se realiza trabajo desde el medio en condiciones adiabáticas. No se observa cambio de temperatura, o sea

$$\left(\frac{\partial T}{\partial V}\right)_U = 0.$$

Pero

$$\left(\frac{\partial T}{\partial V}\right)_U = \frac{-1}{\left(\frac{\partial U}{\partial T}\right)_T \left(\frac{\partial U}{\partial V}\right)_T} = \frac{-1}{C_V} \left(\frac{\partial U}{\partial V}\right)_T = 0,$$

implica que

$$(4-1) \quad \left(\frac{\partial U}{\partial V}\right)_T = 0 \quad (\text{Ley de Joule}) .$$

Ahora

$$(4-2) \quad 0 = \left(\frac{\partial U}{\partial V}\right)_T = \left(\frac{\partial U}{\partial V}\right)_S + \left(\frac{\partial U}{\partial S}\right)_V \left(\frac{\partial S}{\partial V}\right)_T = -p + T \left(\frac{\partial S}{\partial V}\right)_T .$$

Considero la función de estado  $F=U-TS$  cuyo diferencial es

$$dF = dU - TdS - SdT = -SdT - pdV .$$

Con esto obtengo

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial}{\partial V} \left(\frac{\partial(-F)}{\partial T}\right)_V\right)_T = \left(\frac{\partial}{\partial T} \left(\frac{\partial(-F)}{\partial V}\right)_T\right)_V = \left(\frac{\partial p}{\partial T}\right)_V \quad \left(= \frac{\alpha}{\kappa_T}\right) ,$$



y con (4-2)

$$\left(\frac{\partial p}{\partial T}\right)_V = p/T .$$

La integración de esta ecuación diferencial produce  $\ln(p) = \ln(T) + f(V)$  o sea

$$(4-3) \quad p\Phi(V) = T .$$

La ley de Boyle la escribo como

$$(4-4) \quad pV = \Psi(T) ;$$

y esto con (4-3) nos da

$$p = \frac{\Psi(T)}{V} = \frac{T}{\Phi(V)} ,$$

de modo que

$$\Phi(V) = \frac{T}{\Psi(T)} V$$

que debe ser independiente de  $T$ . Esto indica que  $\Psi(T) = cT$  y que  $\Phi(V) = c^{-1}V$ , además  $pV = cT$ . Ya que  $V$  es extensiva,

$$(4-5) \quad \boxed{pV = NRT} ,$$

donde  $R$  es la llamada constante molar de los gases ideales de dimensión  $JK^{-1}mol^{-1}$ . Es notable que la ecuación de estado (4-5) es consecuencia (¡termodinámica!) de la Ley de Joule (4-1) y de la Ley de Boyle (4-4). No hay hipótesis específicas sobre como depende  $U$  de  $T$ . Observamos que  $\alpha = 1/T$  y  $\kappa_T = 1/p$ .

Suponemos ahora que  $C_v = Nc_v = (\partial U/\partial T)_V$  es constante (no depende de  $T$ ; ya asumimos que no hay dependencia<sup>8</sup> de  $V$ ); entonces

$$(4-6) \quad \boxed{U = Nc_v T} ,$$

que es la más simple de las formas que puede adoptar la ley de Joule. En lo que sigue será útil introducir el parámetro sin dimensión (que es constante)

$$\zeta_V := c_v/R ,$$

con el cual  $U = NR\zeta_V T$ .

Queremos obtener la entropía. Con las magnitudes molares

$$u := U/N , \quad v := V/N , \quad s := S/N . ,$$

sabemos que

$$S(U, V, N) = Ns(u, v) .$$

---

<sup>8</sup> $(\partial C_v/\partial V)_T = ((\partial/\partial V)(\partial U/\partial T)_V)_T = ((\partial/\partial T)(\partial U/\partial V)_T)_V = 0$ .

Tenemos

$$ds = \frac{1}{T} du + \frac{p}{T} dv ,$$

que con  $T = u/c_v$  y  $p = RT/v$ , nos da

$$ds = \frac{R\zeta_v}{u} du + \frac{R}{v} dv .$$

Integramos esto a lo largo del camino  $(u_o, v_o) \rightarrow (u, v_o)$  con  $v = v_o$  constante y luego  $(u, v_o) \rightarrow (u, v)$  con  $u$  constante obteniendo

$$s(u, v) = s(u_o, v_o) + \zeta_v R \ln(u/u_o) + R \ln(v/v_o) .$$

Entonces, con  $s_o = s(u_o, v_o) = S(U_o, V_o, N_o)/N_o$ ,

$$(4-7) \quad \boxed{S(U, V, N) = Ns(u, v) = Ns_o + NR\zeta_v \ln\left(\frac{UN_o}{U_oN}\right) + RN \ln\left(\frac{VN_o}{V_oN}\right)}$$

Usando la ecuación de estado (4-5) y la ecuación (4-6) también tenemos

$$(4-8) \quad \boxed{S(T, V, N) = Ns_o + \zeta_v RN \ln(T/T_o) + RN \ln\left(\frac{VN_o}{V_oN}\right)}$$

$$(4-9) \quad \boxed{S(T, p, N) = Ns_o + NR(1 + \zeta_v) \ln(T/T_o) - RN \ln(p/p_o)}$$

donde la temperatura de referencia  $T_o$  y la presión de referencia  $p_o$  están dadas por

$$T_o = \frac{U_o}{N_o R \zeta_v} , \quad p_o = \frac{N_o R T_o}{V_o} = \frac{U_o}{V_o \zeta_v} .$$

Dado que  $(U_o, V_o)$  es arbitrario también  $(T_o, V_o)$  o  $(T_o, p_o)$  lo son. Esta es la ecuación fundamental para un gas ideal de una sola componente química definido como un fluido que tiene a (4-5) como ecuación de estado y cuya energía interna está dada por (4-6). Alternativamente, un gas ideal es un fluido que cumple con la Ley de Joule (4-1), con la Ley de Boyle (4-4) y para el cual  $c_V$  es constante.

¿Cual es la ecuación de una adiabática, i.e. la relación entre  $p$  y  $V$  a  $S$  constante? Hemos visto que

$$\left(\frac{\partial V}{\partial p}\right)_S = -V \kappa_T \frac{C_p}{C_v} .$$

Para un gas ideal ya observamos que  $\kappa_T = 1/p$ . Por lo tanto

$$\left(\frac{\partial V}{\partial p}\right)_S = -\frac{V C_v}{p C_p} .$$

También hemos visto que

$$C_p = C_v + \frac{VT\alpha^2}{\kappa_T} ;$$

entonces, ya que  $\alpha = 1/T$  para un gas ideal,

$$C_p = C_v + \frac{pV}{T} = C_v + NR ,$$

o bien

$$c_p = c_v + R .$$

Entonces,

$$\gamma := c_p/c_v = 1 + \frac{1}{\zeta_v} = \frac{1 + \zeta_v}{\zeta_v}$$

es constante. Con esto podemos integrar

$$\left( \frac{\partial V}{\partial p} \right)_S = -\gamma^{-1} \frac{V}{p} ,$$

a

$$-\gamma \ln(V) = \ln(p) + a(S) ,$$

o sea:

$$pV^\gamma = A(S) .$$

Observamos –por último– que el factor multiplicativo  $NR(1+\zeta_v)$  en el sumando con  $\ln(T/T_0)$  de la expresión para  $S(T, p, N)$  es igual a  $Nc_p = C_p$ .

## 4.1. Mezcla ideal de gases ideales

Consideramos  $r$  ( $r \geq 2$ ) gases ideales en una mezcla con composición dada por  $\mathbf{N} = (N_1, N_2, \dots, N_r)$ . Introducimos  $N = N_1 + N_2 + \dots + N_r$  y las fracciones molares  $x_j = N_j/N$  ( $j = 1, 2, \dots, r$ ). Se tiene  $\sum_{j=1}^r x_j = 1$ . Suponemos “idealidad” de la mezcla en el sentido que la ec. de estado es

$$(4-10) \quad pV = NRT .$$

Podemos introducir las “presiones parciales” formales  $p_j = N_jRT$  y entonces  $p = \sum_{j=1}^r p_j$ .

En general

$$(4-11) \quad dS = \frac{1}{T}dU + \frac{p}{T}dV - \sum_{j=1}^r \frac{\mu_j}{T}dN_j .$$

#### 4.1.1. Primera variante (aditividad de la energía interna):

Para cada  $j$ ,

$$U_j = N_j c_v^{(j)} T = N_j R \zeta_v^{(j)} T .$$

Y

$$(4-12) \quad U = \sum_{j=1}^r U_j = T \sum_{j=1}^r N_j c_v^{(j)} = NR \zeta_v T ,$$

donde

$$(4-13) \quad \zeta_v := \sum_{j=1}^r x_j \zeta_v^{(j)} ,$$

es constante.

Para integrar (4-11), tomamos un estado de referencia  $(U_o, V_o, \mathbf{N})$  cuya composición  $\mathbf{N}$  es la de mezcla. Entonces, ya que

$$T = \frac{U}{NR \zeta_v} , \quad \frac{p}{T} = NR/V ,$$

tenemos

$$dS = \frac{NR \zeta_v}{U} dU + \frac{NR}{V} dV ,$$

en cualquier camino en el espacio  $(U, V, \mathbf{N})$  para el cual  $dN_1 = dN_2 = \dots = dN_r = 0$ . Ya que  $\zeta_v$  es constante, integrando desde  $(U_o, V_o, \mathbf{N})$  a  $(U, V_o, \mathbf{N})$  con  $V = V_o$  y sin cambios en  $\mathbf{N}$  y luego desde  $(U, V_o, \mathbf{N})$  a  $(U, V, \mathbf{N})$  sin variación en  $U$  ni en  $\mathbf{N}$ , obtenemos

$$(4-14) \quad S(U, V, \mathbf{N}) = S(U_o, V_o, \mathbf{N}) + NR \zeta_v \ln(U/U_o) + NR \ln(V/V_o) .$$

La dependencia de  $S$  en  $\mathbf{N}$  no es explícita y no hemos integrado completamente al diferencial  $dS$ . Sin embargo, expresando  $U$  y  $U_o$  con la relación  $U = NR \zeta_v T$  teniendo en cuenta que tanto  $N$  como  $\zeta_v$  son constantes,

$$S(T, V, \mathbf{N}) = S(T_o, V_o, \mathbf{N}) + RN \zeta_v \ln(T/T_o) + RN \ln(V/V_o) .$$

#### 4.1.2. Segunda variante (aditividad de la entropía)

Tomo un punto de referencia arbitrario pero fijo  $(T_o, V_o, N_o) \in \mathbb{R}_+^3$  que uso para obtener la entropía de cada componente de acuerdo a (4-8):

$$S^{(j)}(T, V, N_j) = N_j s_o^{(j)} + RN_j \zeta_v^{(j)} \ln(T/T_o) + RN_j \ln\left(\frac{VN_o}{V_o N_j}\right) .$$

Aquí,

$$s_o^{(j)} = S^{(j)}(U_o^{(j)}, V_o, N_o)/N_o = S^{(j)}(T_o, V_o, N_o)/N_o , \quad U_o^{(j)} = N_o R \zeta_v^{(j)} T_o , \quad \zeta_v^{(j)} = c_v^{(j)}/R .$$

Apelando a que la entropía de la mezcla es la suma de las entropías de cada componente

$$\begin{aligned}
S(T, V, \mathbf{N}) &= \sum_{j=1}^r S^{(j)}(T, V, N_j) \\
&= \sum_{j=1}^r N_j s_o^{(j)} + R \sum_{j=1}^r N_j \zeta_V^{(j)} \ln(T/T_o) + R \sum_{j=1}^r N_j \ln(V/V_o) - R \sum_{j=1}^r N_j \ln(N_j/N_o) \\
&= \sum_{j=1}^r N_j s_o^{(j)} + RN \zeta_V \ln(T/T_o) + RN \ln(V/V_o) - RN \sum_{j=1}^r x_j \ln(x_j N/N_o) \\
&= \sum_{j=1}^r N_j s_o^{(j)} + RN \zeta_V \ln(T/T_o) + RN \ln(V/V_o) - RN \sum_{j=1}^r x_j \ln(x_j) - RN \ln(N/N_o) .
\end{aligned}$$

Definimos la **entropía de mezcla**

$$(4-15) \quad \boxed{S_{mezcla} = -RN \sum_{j=1}^r x_j \ln(x_j)} ;$$

y observamos que  $S_{mezcla}$  es:

- solamente una función de la composición  $\mathbf{N}$  y es homogénea de grado 1;
- estrictamente positiva<sup>9</sup>;
- maximal para  $N_j = N/r$  con valor  $NR \ln(r)$ .

Con esto

$$(4-16) \quad \boxed{S(T, V, \mathbf{N}) = \sum_{j=1}^r N_j s_o^{(j)} + RN \zeta_V \ln(T/T_o) + RN \ln(VN_o/V_oN) + S_{mezcla}}$$

Esta es una integración completa en la medida que  $N$ ,  $\zeta_v$  y las fracciones molares  $x_j$  son todas funciones explícitas de  $\mathbf{N}$ . Pero está no es la ecuación fundamental para lo cual debemos expresar a  $S$  como función de  $(U, V, \mathbf{N})$ . Esta tarea es inmediata; a partir de que

$$U = NR \zeta_v T$$

obtenemos

$$(4-17) \quad \boxed{S(U, V, \mathbf{N}) = \sum_{j=1}^r N_j s_o^{(j)} + RN \zeta_v \ln \left( \frac{U}{NR \zeta_v T_o} \right) + RN \ln(VN_o/V_oN) + S_{mezcla}}$$

como ecuación fundamental  $S(U, V, \mathbf{N})$ . Aquí todos los argumentos de los logaritmos naturales son adimensionales.

<sup>9</sup>ya que  $t \ln(t) < 0$  para todo  $0 < t < 1$ .

--

*Nota y advertencia:* Introduciendo

$$U_o = NR\zeta_v T_o = \sum_{j=1}^r N_j \frac{U_o^{(j)}}{N_o} = \sum_{j=1}^r N_j u_o^{(j)}, \quad u_o^{(j)} = \frac{U_o^{(j)}}{N_o},$$

notamos que  $U_o$  –que es la energía interna de la mezcla de la misma composición pero a temperatura  $T_o$ – es función de la composición (homogénea de grado 1). Podemos reescribir (4-17) como

$$S(U, V, \mathbf{N}) = \sum_{j=1}^r N_j s_o^{(j)} + RN\zeta_v \ln(U/U_o) + RN \ln(VN_o/V_oN) + S_{mezcla};$$

lo que es correcto pero hay que tener cuidado. En (4-17) se ve manifiestamente que  $S$  es función homogénea de grado 1 en sus variables naturales  $(U, V, \mathbf{N})$  pero esto no es el caso en la ec. que acabamos de escribir. Allí, la homogeneidad se manifiesta cuando no olvidamos escalar a  $U_o$  ya que  $\mathbf{N} \rightarrow \lambda\mathbf{N}$  implica que  $U_o \rightarrow \lambda U_o$ .

Hay otra posibilidad. Definiendo

$$U'_o = N_o R\zeta_v T_o = \frac{N_o}{N} U_o = \sum_{j=1}^r x_j U_o^{(j)},$$

obtengo

$$S(U, V, \mathbf{N}) = \sum_{j=1}^r N_j s_o^{(j)} + RN\zeta_v \ln(UN_o/U'_oN) + RN \ln(VN_o/V_oN) + S_{mezcla}.$$

Ahora  $U'_o$  depende de la composición solamente via las fracciones molares que son invariantes ante  $\mathbf{N} \rightarrow \lambda\mathbf{N}$  (con lo cual  $U'_o$  es homogénea de grado cero en  $\mathbf{N}$ ). En esta última versión de  $S(U, V, \mathbf{N})$  la homogeneidad de grado 1 es manifiesta.

--

Eliminemos  $V$  en términos de  $(p, T)$ . Con la ec. (4-10),  $V = NRT/p$  y definiendo

$$p_o = N_o RT_o / V_o,$$

$$(4-18) \quad \boxed{S(T, p, \mathbf{N}) = \sum_{j=1}^r N_j s_o^{(j)} + RN(\zeta_v + 1) \ln(T/T_o) - RN \ln(p/p_o) + S_{mezcla}}.$$

La comparación con  $\sum_{j=1}^r S^{(j)}(p, T, N_j)$  nos muestra que

$$S(T, p, \mathbf{N}) = \sum_{j=1}^r S^{(j)}(T, p, N_j) + S_{mezcla};$$

la entropía de la mezcla es igual a la suma de las entropías de las componentes a la misma temperatura y presión que tiene la mezcla aumentada con la entropía de mezcla.

### 4.1.3. Funciones de respuesta. ‘Adiabáticas’

De la ec. de estado (4-10) obtenemos que para la mezcla;

$$\alpha = 1/T, \quad \kappa_T = 1/p.$$

Y de (4-12),

$$C_V = \sum_{j=1}^r N_j c_v^{(j)} = NR\zeta_v, \quad c_v = C_V/N = \sum_{j=1}^r x_j c_v^{(j)} = R\zeta_v,$$

y análogamente usando  $S(T, p, \mathbf{N})$  que

$$C_p = \sum_{j=1}^r N_j c_p^{(j)}, \quad c_p = \sum_{j=1}^r x_j c_p^{(j)}.$$

Con esto se obtiene de  $c_p^{(j)} = c_v^{(j)} + R$  que

$$c_p = c_v + R, \quad \gamma = c_p/c_v = 1 + \frac{1}{\zeta_v}.$$

Todo como en el gas ideal de una sola componente pero con  $\zeta_v$  dado por (4-13) que depende de la composición.

De la relación general para  $(\partial V/\partial p)_S$  obtenemos la misma ec. para las ‘adiabáticas’:

$$pV^\gamma = A(S).$$

### 4.1.4. Paradoja de Gibbs

Considere dos gases ideales en sendos compartimentos de un cilindro completamente cerrado (rígido, impermeable, adiabático). Los compartimentos tienen volúmenes  $V_1$  y  $V_2$  respectivamente y están separados por una pared impermeable, rígida, pero diatérmica. Cuando hay equilibrio en este sistema las temperaturas  $T_1$  y  $T_2$  son iguales a  $T$ , y la entropía viene dada por

$$S_i = S^{(1)}(T, V_1, N_1) + S^{(2)}(T, V_2, N_2).$$

Se saca ahora la pared que separa de modo que los gases disponen del volumen total  $V = V_1 + V_2$ . Previsiblemente los gases difunden y, cuando se llega al equilibrio, tenemos una mezcla homogénea de ambos gases a la misma temperatura  $T$  con composición  $\mathbf{N} = (N_1, N_2)$ . En este estado final (el proceso es irreversible), tenemos

$$S_f = S(T, V_1 + V_2, N_1, N_2)$$

dada por la expresión anterior. El cambio de la entropía es

$$\Delta S = S_f - S_i.$$

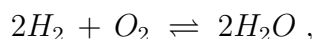
Tomando las expresiones desarrolladas para  $S_f$ ,  $S^{(1)}$  y  $S^{(2)}$ , obtenemos

$$\Delta S = RN_1 \ln \left( \frac{V_1 + V_2}{V_1} \right) + RN_2 \ln \left( \frac{V_1 + V_2}{V_2} \right) .$$

Observamos que ya que los argumentos de los logaritmos naturales son mayores a 1,  $\Delta S > 0$  (como debe ser). Esta fórmula es válida incluso cuando los gases son idénticos y aquí está la paradoja de Gibbs: Si los gases son idénticos al abrir la compuerta no se mezclan y no hay cambio alguno de la entropía total. El problema subyacente es: ¿que quiere decir idénticos? Y la paradoja se levantará cuando se aclare este concepto (en mecánica estadística). Las expresiones obtenidas para  $S(T, V, \mathbf{N})$  son incorrectas cuando dos o más de los gases ideales son iguales.

## 5. Equilibrio Químico

Considero sistemas termodinámicos con múltiples especies químicas que participan en una reacción química. Por ejemplo un gas mezcla de agua, oxígeno (molecular) e hidrógeno (molecular)



como es costumbre en Química. Los coeficientes estequiométricos  $\nu_{H_2O} = 2$ ,  $\nu_{H_2} = 2$  y  $\nu_{O_2} = 1$  nos indican las proporciones de los reactante  $H_2O$ ,  $H_2$  y  $O_2$ . La doble flecha  $\rightleftharpoons$  indica que la reacción puede suceder en ambos sentidos: formación de agua a partir de oxígeno e hidrógeno o bien descomposición de agua en sus constituyentes elementales. Claramente podemos reescribir esta reacción como

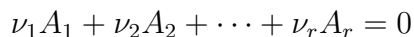


Para lo que sigue es conveniente reescribir la reacción simbólicamente como

$$-2H_2 - O_2 + 2H_2O = 0 , \quad \text{o bien como } 3H_2 + \frac{3}{2}O_2 - 3H_2O = 0 .$$

Los coeficientes estequiométricos están determinados hasta multiplicación por un factor real no nulo arbitrario.

Consideramos entonces una reacción química que involucra a  $r$  especies químicas  $A_j$ ,  $j = 1, 2, \dots, r$ , escrita simbólicamente como



Los coeficientes estequiométricos  $\nu_j$  (no todos nulos) no tienen todos el mismo signo. Si multiplicamos cada uno de los coeficientes estequiométricos por el mismo número real no-nulo obtenemos la misma reacción química. Es conveniente convenir que las especies cuyo coeficiente estequiométrico es positivo serán los *productos* y que los *eductos* serán aquellas especies cuyos coeficientes son negativos. Esto (vía multiplicación de todos los coeficientes con  $-1$ ) es puramente convencional.

El número molar de la  $j$ -ésima especie será  $N_j$ ,  $j = 1, 2, \dots, r$ . Supongamos que hay un cambio



de los números molares  $\Delta N_j$  de los reactantes producido por la reacción. Entonces en virtud de la conservación de masa y la identidad misma (supuesta estable) de las especies químicas involucradas, estos incrementos se encuentran en la misma relación de proporción que la relación de proporción entre los coeficientes estequiométricos correspondientes

$$\Delta N_1 : \Delta N_2 : \cdots : \Delta N_r = \nu_1 : \nu_2 : \cdots : \nu_r .$$

En otras palabras

$$(5-1) \quad \frac{\Delta N_1}{\nu_1} = \frac{\Delta N_2}{\nu_2} = \cdots = \frac{\Delta N_r}{\nu_r} .$$

En el caso de la reacción  $-2H_2O + O_2 + 2H_2 = 0$ , si desaparecen  $5 \text{ mol}$  ( $\Delta N_{H_2O} = -5 \text{ mol}$ ) de agua deben aparecer  $5 \text{ mol}$  de hidrógeno molecular ( $\Delta N_{H_2} = 5 \text{ mol}$ ) y  $2,5 \text{ mol}$  de oxígeno molecular ( $\Delta N_{O_2} = 2,5 \text{ mol}$ ):

$$\frac{\Delta N_{H_2O}}{-2} = \frac{\Delta N_{H_2}}{2} = \frac{\Delta N_{O_2}}{1} = \frac{5}{2} .$$

La proporción (5-1) nos permite introducir una variable real  $\xi$  tal que

$$\Delta \xi := \frac{\Delta N_1}{\nu_1} = \frac{\Delta N_2}{\nu_2} = \cdots = \frac{\Delta N_r}{\nu_r} .$$

Entonces

$$(5-2) \quad N_j = N_j^{(o)} + \nu_j \xi , \quad j = 1, 2, \dots, r , \quad \mathbf{N} = \mathbf{N}^{(o)} + \xi \boldsymbol{\nu} .$$

donde  $\mathbf{N}^{(o)} = (N_1^{(o)}, N_2^{(o)}, \dots, N_r^{(o)})$  es una composición de referencia (arbitraria) que corresponde a  $\xi = 0$ . Ahora, ya que  $N_j(\xi) = N_j^{(o)} + \nu_j \xi$  no puede ser negativo, obtenemos cotas de variabilidad para  $\xi$  en función de la composición de referencia y los coeficientes estequiométricos. En efecto, de  $\nu_j \xi \geq -N_j^{(o)}$  ( $\leq 0$ ) obtenemos:

$$\xi \geq -\frac{N_j^{(o)}}{\nu_j} \text{ para todo } j \text{ con } \nu_j > 0 ;$$

$$\xi \leq -\frac{N_j^{(o)}}{\nu_j} \text{ para todo } j \text{ con } \nu_j < 0 .$$

A fortiori,

$$\xi_{min} \leq \xi \leq \xi_{max} ,$$

donde

$$\xi_{max} := \min \left\{ -\frac{N_j^{(o)}}{\nu_j} : j , \nu_j < 0 \right\} = \min \left\{ \frac{N_j^{(o)}}{|\nu_j|} : j , \nu_j < 0 \right\} ,$$

$$\xi_{min} := \max \left\{ -\frac{N_j^{(o)}}{\nu_j} : j , \nu_j > 0 \right\} = -\min \left\{ \frac{N_j^{(o)}}{\nu_j} : j , \nu_j > 0 \right\} .$$

Observese que  $\xi_{min} \leq 0$  y que  $\xi_{max} \geq 0$ . Además,  $\xi_{max} = 0$  si  $N_j^{(o)} = 0$  para todos los eductos (la reacción no puede proceder de eductos a productos) mientras que  $\xi_{min} = 0$  si no hay productos (la reacción no puede proceder de productos a eductos).

Si  $p$  y  $T$  son constantes la composición de equilibrio de la mezcla reactiva está determinada por la minimización de  $G(T, p, \mathbf{N})$  a  $p, T$  constantes. Fijando  $p, T$  y definiendo

$$\Phi(\xi) := G(T, p, \mathbf{N}^{(o)} + \xi \boldsymbol{\nu}), \quad \boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_r),$$

se trata de minimizar una función de una variable real en el intervalo<sup>10</sup>  $[\xi_{min}, \xi_{max}]$  para obtener la composición vía (5-2). Este problema es enormemente simplificado por el hecho de que  $\Phi$  es una función convexa. Esto es consecuencia de la energía libre de Gibbs como función de sus variables extensivas: convexidad de  $\mathbf{N} \mapsto G(T, p, \mathbf{N})$ . Entonces, dados  $\xi_1$  y  $\xi_2$  en el intervalo de variabilidad, tenemos –para  $0 < \lambda < 1$ –

$$\begin{aligned} \Phi(\lambda \xi_1 + (1 - \lambda) \xi_2) &= G(T, p, \mathbf{N}^{(o)} + (\lambda \xi_1 + (1 - \lambda) \xi_2) \boldsymbol{\nu}) \\ &= G(T, p, \lambda [\mathbf{N}^{(o)} + \xi_1 \boldsymbol{\nu}] + (1 - \lambda) [\mathbf{N}^{(o)} + \xi_2 \boldsymbol{\nu}]) \\ &\leq \lambda G(T, p, \mathbf{N}^{(o)} + \xi_1 \boldsymbol{\nu}) + (1 - \lambda) G(T, p, \mathbf{N}^{(o)} + \xi_2 \boldsymbol{\nu}) = \lambda \Phi(\xi_1) + (1 - \lambda) \Phi(\xi_2). \end{aligned}$$

En virtud de la convexidad, hay dos posibilidades: o bien hay un punto crítico de  $\Phi$  en el intervalo abierto en cuyo caso ese punto corresponde automáticamente a un mínimo global, o bien el mínimo se asume en el borde. Si la función  $\Phi$  es convexa pero no estrictamente convexa no podemos garantizar de que en el caso de que el punto crítico existe el mínimo sea único<sup>11</sup>.

¿Cual es la condición de estacionaridad para un punto  $\xi$ ? A partir de  $dG = \sum_{j=1}^r \mu_j dN_j = 0$  obtenemos

$$\sum_{j=1}^r \nu_j \mu_j(T, p, \mathbf{N}(\xi)) = 0;$$

equivalentemente

$$0 = \frac{d\Phi}{d\xi}(\xi) = \sum_{j=1}^r \left( \frac{\partial G}{\partial N_j} \right)_{T, p, N_k (k \neq j)} \nu_j = \sum_{j=1}^r \nu_j \mu_j(T, p, \mathbf{N}(\xi)).$$

Claramente –como debe ser– esta condición es invariante ante multiplicación por cualquier real no-nulo de cada uno de los coeficientes estequiométricos.

Si la temperatura es constante, ya que  $H = G + TS$ , se tiene

$$\left( \frac{\partial H}{\partial \xi} \right)_{T, p} = \left( \frac{\partial G}{\partial \xi} \right)_{T, p} + T \left( \frac{\partial S}{\partial \xi} \right)_{T, p} = \sum_{j=1}^r \nu_j \mu_j + T \left( \frac{\partial S}{\partial \xi} \right)_{T, p};$$

<sup>10</sup>Este intervalo no se reduce a un único punto  $\xi = 0$  cuando la composición de referencia no es nula.

<sup>11</sup>Considere una función constante que es automáticamente convexa (y cóncava).

el segundo sumando representa el *calor de reacción* (a  $T, p$  constantes). Con  $S = -(\partial G/\partial T)_{p, \mathbf{N}}$ , obtenemos

$$T \left( \frac{\partial S}{\partial \xi} \right)_{T, p} = -T \sum_{j=1}^r \nu_j \left( \frac{\partial \mu_j}{\partial T} \right)_{p, \mathbf{N}} .$$

En equilibrio

$$\left( \frac{\partial H}{\partial \xi} \right)_{T, p} (T, p, \mathbf{N}^{eq}) = - \sum_{j=1}^r \nu_j \left( \frac{\partial \mu_j}{\partial T} \right)_{p, \mathbf{N}} (T, p, \mathbf{N}^{eq}) ;$$

la reacción se dice *endotérmica* si esta magnitud es positiva y *exotérmica* en el caso contrario. Un cambio (infinitesimal) positivo en  $T$  provoca un cambio (infinitesimal) en  $\xi$  que sera positivo si la reacción es endotérmica y negativo en el caso exotérmico.

## El caso de una mezcla ideal de gases ideales - Ley de acción de masas

El potencial químico de la  $j$ -ésima especie  $A_j$  para una mezcla ideal de gases ideales es

$$\mu_j(T, p, \mathbf{N}) = RT[(\phi_j(T) + \ln(px_j))] ,$$

donde  $\phi_j$  es función únicamente de la temperatura una vez elegido un estado de referencia común a todos las especies<sup>12</sup>. Entonces la condición de régimen estacionario se escribe

$$K := \prod_{j=1}^r (x_j)^{\nu_j} = (p)^{-\sum_{j=1}^r \nu_j} \exp\left\{-\sum_{j=1}^r \nu_j \phi_j(T)\right\} ,$$

conocida como Ley de acción de masas<sup>13</sup>. Esta relación entre las fracciones molares, la presión y temperatura de la mezcla determina los valores de las fracciones molares de equilibrio.

## Muchas reacciones

Si consideramos  $L$  reacciones

$$\sum_{j=1}^r \nu_j^{(k)} A_j = 0 , \quad k = 1, 2, \dots , L ,$$

entre  $r$  especies (no todas las especies deben necesariamente participar en cada una de las reacciones), entonces la ley de proporciones (5-1) es válida para cada reacción y se introducen  $L$  “avances de reacción”  $\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(L)}$  de modo que, con los  $L$  vectores  $\boldsymbol{\nu}^{(k)} =$

<sup>12</sup> $\phi_j(T) = \text{const.} + c_p[\ln(T/T_o) - (T - T_o)/T]$ .

<sup>13</sup>Nuevamente, la identidad es invariante ante multiplicación de los coeficientes estequiométricos con un factor común.

$(\nu_1^{(k)}, \nu_2^{(k)}, \dots, \nu_r^{(k)})$ ,

$$N_j = N_j^{(o)} + \sum_{k=1}^L \nu_j^{(k)} \xi^{(k)}, \quad \mathbf{N} = \mathbf{N}^{(o)} + \sum_{k=1}^L \xi^{(k)} \boldsymbol{\nu}^{(k)} = \mathbf{N}^{(o)} + \mathbf{S}\boldsymbol{\xi},$$

donde se introduce el vector  $L$ -dimensional  $\boldsymbol{\xi}$  de componentes  $\xi_k = \xi^{(k)}$  ( $k = 1, 2, \dots, L$ ) y la matriz ( $r \times L$ ) estequiométrica  $S$  de componentes  $S_{jk} = \nu_j^{(k)}$  ( $j = 1, 2, \dots, r; k = 1, 2, \dots, L$ )<sup>14</sup>. Nuevamente, con un cálculo análogo (idéntico) al hecho para una sola reacción, la función

$$\Phi(\boldsymbol{\xi}) = G(T, p, \mathbf{N}^{(o)} + \mathbf{S}\boldsymbol{\xi})$$

resulta convexa en el vector  $\boldsymbol{\xi}$  en el paralelepípedo<sup>15</sup>

$$[\xi_{min}^{(1)}, \xi_{max}^{(1)}] \times [\xi_{min}^{(2)}, \xi_{max}^{(2)}] \times \dots \times [\xi_{min}^{(L)}, \xi_{max}^{(L)}] \subset \mathbb{R}^L$$

donde los valores extremales  $\xi_{min}^{(k)} < \xi_{max}^{(k)}$  se definen reacción por reacción con los mismos valores de  $\mathbf{N}^{(o)}$ . Se quiere entonces minimizar a  $\Phi$  en este paralelepípedo; la convexidad implica que si el paralelepípedo admite un punto estacionario en su interior, este corresponde a un mínimo global, y en caso contrario el mínimo se asume en el borde. La condición para un punto estacionario (anulación del gradiente de  $\Phi$ ) es equivalente a

$$\sum_{j=1}^r \nu_j^{(k)} \mu_j = 0, \quad \text{para } k = 1, 2, \dots, L.$$

En equilibrio hay entonces equilibrio en cada una de las reacciones y viceversa.

## 6. Apéndice: Funciones homogéneas

Consideramos funciones a valores reales definidas en un cono convexo  $\mathcal{C} \subset \mathbb{R}^d$ . Una de estas funciones  $f$  se dice **homogénea de grado**  $\alpha$  si para algún real  $\alpha$  (fijo):

$$(6-1) \quad f(\lambda \mathbf{x}) = \lambda^\alpha f(\mathbf{x}), \quad \text{para todo } \lambda > 0 \text{ y todo } \mathbf{x} \in \mathcal{C}$$

1. Si  $f$  homogénea de grado  $\alpha$  es diferenciable con respecto a la  $j$ -ésima variable  $x_j$  entonces la derivada parcial con respecto a  $x_j$  es una función homogénea de grado  $\alpha - 1$ . En efecto, para todo  $\lambda > 0$ , y cualquier  $h \in \mathbb{R}$  tal que  $x_j + h \geq 0$

$$\frac{f(\lambda x_1, \lambda x_2, \dots, \lambda x_j + h, \dots, \lambda x_n) - f(\lambda \mathbf{x})}{h}$$

<sup>14</sup>Observe que  $r \geq L$ . En la práctica conviene trabajar con una matriz estequiométrica de rango maximal  $L$ , o sea con reacciones linealmente independientes. Esto, sin embargo, no cambia ni simplifica lo que sigue.

<sup>15</sup>Que es un conjunto convexo en  $\mathbb{R}^L$ ; i.e., contiene el segmento de recta que une a cualquier par de puntos de él.

$$\begin{aligned}
&= \frac{\lambda^\alpha f(x_1, x_2, \dots, x_j + (h/\lambda^\alpha), \dots, x_n) - \lambda^\alpha f(\mathbf{x})}{h} \\
&= \frac{f(x_1, x_2, \dots, x_j + (h/\lambda^\alpha), \dots, x_n) - f(\mathbf{x})}{(h/\lambda^\alpha)}.
\end{aligned}$$

La afirmación se obtiene tomando el límite  $h \rightarrow 0$ .

2. (Relación de Euler). Si  $f$  admite derivadas parciales respecto de cada una de sus variables en un abierto  $A \subset \mathcal{C}$  entonces

$$(6-2) \quad (\nabla f)(\mathbf{x}) \cdot \mathbf{x} = \alpha f(\mathbf{x}), \quad \mathbf{x} \in A.$$

E inversamente, si se cumple (6-2) en un abierto  $A \subset \mathcal{C}$  entonces se cumple (6-1) para todo  $\lambda > 0$  suficientemente cercano a 1.

3. **Lema 2** *La función continua  $f$  definida en  $\mathcal{C}$  es homogénea de grado  $\alpha$  si y sólo si (6-1) se cumple para todo  $\lambda \in \mathbb{N}$ .*

Demostración: La necesidad de la condición es evidente.

Si para todo  $p \in \mathbb{N}$  tenemos  $f(p\mathbf{x}) = p^\alpha f(\mathbf{x})$  cualquiera sea  $\mathbf{x} \in \mathcal{C}$ , entonces cualquiera sea  $\mathbf{y} \in \mathcal{C}$  tenemos  $\mathbf{z} = (1/p)\mathbf{y} \in \mathcal{C}$  y por ende

$$f\left(\frac{1}{p}\mathbf{y}\right) = f(\mathbf{z}) = p^{-\alpha} p^\alpha f(\mathbf{z}) = p^{-\alpha} f(p\mathbf{z}) = \left(\frac{1}{p}\right)^\alpha f(\mathbf{y})$$

Entonces si  $p, q \in \mathbb{N}$  y  $\mathbf{x} \in \mathcal{C}$  tenemos

$$\left(\frac{p}{q}\right)^\alpha f(\mathbf{x}) = \frac{1}{q^\alpha} f(p\mathbf{x}) = f\left(\frac{p}{q}\mathbf{x}\right),$$

que demuestra (6-1) para  $\lambda$  racional y positivo. Dado  $\lambda > 0$  real, hay una sucesión de racionales que convergen a  $\lambda$  y la continuidad de  $f$  nos produce (6-1).

## 7. Apéndice: Funciones convexas.

Para  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $\mathbf{x}\mathbf{y}$  denota el producto escalar usual y  $\|\mathbf{x}\| := \sqrt{\mathbf{x}\mathbf{x}}$  la norma (o distancia a  $\mathbf{0}$ ).

Considero un conjunto convexo  $\mathcal{K}$  de  $\mathbb{R}^n$ , o sea que  $\mathcal{K}$  es tal que para cada par  $(\mathbf{x}, \mathbf{y})$  de puntos  $\mathbf{x}$  e  $\mathbf{y}$  de  $\mathcal{K}$ , el segmento (de recta)

$$\{\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} : 0 \leq \lambda \leq 1\},$$

está en  $\mathcal{K}$ . Los ejemplos más relevantes para nosotros son  $\mathcal{K} = \mathbb{R}_+^n$ ; o  $\mathcal{K} = \mathbb{R}^n$ .

Una función a valores reales definida en  $\mathcal{K}$  se dice **convexa** si

$$(7-1) \quad f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y})$$

se cumple para  $\mathbf{x}$  y  $\mathbf{y}$  puntos arbitrarios de  $\mathcal{K}$  y todo  $0 < \lambda < 1$ . En caso de que esto valga reemplazando  $\leq$  por  $\geq$ , la función se dice **cóncava**. Alternativamente,  $f$  es cóncava si  $-f$  es convexa. Si las desigualdades son estrictas hablamos de **estrictamente** convexa o cóncava.

Si  $\mathcal{K}$  es tal que con  $\mathbf{x} \in \mathcal{K}$  también  $a\mathbf{x} \in \mathcal{K}$  para todo  $a > 0$  entonces  $\mathcal{K}$  es un cono. Si  $\mathcal{K}$  es un cono convexo entonces  $x + y \in \mathcal{K}$  cuando  $\mathbf{x}, \mathbf{y}$  están en  $\mathcal{K}$  ya que  $x + y = \lambda(1/\lambda)\mathbf{x} + (1 - \lambda)(1/(1 - \lambda))\mathbf{y}$  para cualquier  $0 < \lambda < 1$ . Si  $f$  está definida en un cono convexo  $\mathcal{K}$ , toma valores reales, es convexa y homogénea de grado 1, entonces es sub-aditiva:  $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$ <sup>16</sup>. Análogamente, si  $f$  es sub-aditiva y homogénea de grado 1 entonces es convexa.

**Lema 3** (*J.L.V.W. Jensen*)  $f$  definida en  $\mathcal{K}$  es convexa si y sólo si para cualquier conjunto finito  $\{\mathbf{x}_j : j = 1, 2, \dots, N\}$  en  $\mathcal{K}$  y cualquier conjunto  $\{\lambda_j : \lambda_j \geq 0, j = 1, 2, \dots, N\}$  tal que  $\sum_{j=1}^N \lambda_j = 1$ , se tiene

$$(7-2) \quad f\left(\sum_{j=1}^N \lambda_j \mathbf{x}_j\right) \leq \sum_{j=1}^N \lambda_j f(\mathbf{x}_j).$$

Demostración: La suficiencia de (7-2) es inmediata. Para ver la necesidad, considere  $N > 2$  y suponga que  $\lambda_j > 0$  para todo  $j = 1, 2, \dots, N$ . Sea  $\mu_1 = \sum_{j=1}^{N-1} \lambda_j$ . Entonces, con

$$\mathbf{y}_1 = \sum_{j=1}^{N-1} \frac{\lambda_j}{\mu_1} \mathbf{x}_j,$$

tenemos

$$\sum_{j=1}^N \lambda_j \mathbf{x}_j = \mu_1 \mathbf{y}_1 + (1 - \mu_1)\mathbf{x}_N$$

de modo que —ya que  $1 - \mu_1 = \lambda_N$ —

$$f\left(\sum_{j=1}^N \lambda_j \mathbf{x}_j\right) = f(\mu_1 \mathbf{y}_1 + (1 - \mu_1)\mathbf{x}_N) \leq \mu_1 f(\mathbf{y}_1) + (1 - \mu_1)f(\mathbf{x}_N) = \mu_1 f(\mathbf{y}_1) + \lambda_N f(\mathbf{x}_N).$$

Si  $N - 1 > 2$ , sea  $\mu_2 = \sum_{j=1}^{N-2} (\lambda_j / \mu_1)$ . Con

$$\mathbf{y}_2 = \sum_{j=1}^{N-2} \frac{\lambda_j}{\mu_1 \mu_2} \mathbf{x}_j,$$

---

<sup>16</sup>Use  $x + y = \lambda(1/\lambda)\mathbf{x} + (1 - \lambda)(1/(1 - \lambda))\mathbf{y}$ .

obtenemos  $\mathbf{y}_1 = \mu_2 \mathbf{y}_2 + (1 - \mu_2) \mathbf{x}_{N-1}$  y entonces

$$f(\mathbf{y}_1) = f(\mu_2 \mathbf{y}_2 + (1 - \mu_2) \mathbf{x}_{N-1}) \leq \mu_2 f(\mathbf{y}_2) + (1 - \mu_2) f(\mathbf{x}_{N-1})$$

de modo que –ya que  $\mu_1(1 - \mu_2) = \lambda_{N-1}$ –

$$\begin{aligned} f\left(\sum_{j=1}^N \lambda_j \mathbf{x}_j\right) &\leq \mu_1 f(\mathbf{y}_1) + (1 - \mu_1) f(\mathbf{x}_N) \leq \mu_1 \mu_2 f(\mathbf{y}_2) + \mu_1(1 - \mu_2) f(\mathbf{x}_{N-1}) + \lambda_N f(\mathbf{x}_N) \\ &= \mu_1 \mu_2 f(\mathbf{y}_2) + \lambda_{N-1} f(\mathbf{x}_{N-1}) + \lambda_N f(\mathbf{x}_N). \end{aligned}$$

Podemos iterar esto descomponiendo en el  $k$ -ésimo paso  $\mathbf{y}_{k-1}$  mientras sea posible (i.e.,  $N - (k - 1) > 2$ ), y esto completa la demostración.

1. (Convexidad en el punto medio). Una función sobre  $\mathcal{K}$  es convexa en el punto medio si

$$f\left(\frac{\mathbf{x} + \mathbf{y}}{2}\right) \leq \frac{f(\mathbf{x}) + f(\mathbf{y})}{2},$$

para puntos arbitrarios  $\mathbf{x}$  y  $\mathbf{y}$  de  $\mathcal{K}$ .

**Lema 4** (*J.L.V.W. Jensen, 1905*) Una función continua sobre  $\mathcal{K}$  es convexa si y sólo si es convexa en el punto medio.

Demostración: La necesidad es evidente. Para demostrar la suficiencia, probamos primero que si  $\mathbf{x}_j$ ,  $j = 1, 2, \dots, n$ , son  $n$  ( $n \in \mathbb{N}$ ) puntos de  $\mathcal{K}$ , entonces

$$(7-3) \quad f\left(\frac{1}{n} \sum_{j=1}^n \mathbf{x}_j\right) \leq \frac{1}{n} \sum_{j=1}^n f(\mathbf{x}_j).$$

Esto procede en tres pasos.

*Primer paso:* Demostramos (7-3) para  $n = 2^k$  con  $k \in \mathbb{N}$  por inducción. La convexidad en el punto medio nos dice que (7-3) es cierto para  $k = 1$ . Si vale para algún  $k \in \mathbb{N}$  entonces

$$\begin{aligned} f\left(\frac{1}{2^{k+1}} \sum_{j=1}^{2^{k+1}} \mathbf{x}_j\right) &= f\left(\frac{1}{2^k} \sum_{j=1}^{2^k} \frac{\mathbf{x}_{2j-1} + \mathbf{x}_{2j}}{2}\right) \leq \frac{1}{2^k} \sum_{j=1}^{2^k} f\left(\frac{\mathbf{x}_{2j-1} + \mathbf{x}_{2j}}{2}\right) \\ &\leq \frac{1}{2^k} \sum_{j=1}^{2^k} \left(\frac{f(\mathbf{x}_{2j-1}) + f(\mathbf{x}_{2j})}{2}\right) = \frac{1}{2^{k+1}} \sum_{j=1}^{2^{k+1}} f(\mathbf{x}_j). \end{aligned}$$

*Segundo paso:* Demostramos que si (7-3) es cierto para algún  $n_o + 1$  entonces también vale para  $n_o$ . Supongamos entonces la validez de (7-3) para  $n_o + 1 \in \mathbb{N}$  con  $n_o > 1$ . Entonces cualesquiera sean  $\{\mathbf{x}_j : j = 1, 2, \dots, n_o\} \subset \mathcal{K}$ , definimos

$$\mathbf{x}_{n_o+1} := \frac{1}{n_o} \sum_{j=1}^{n_o} \mathbf{x}_j ;$$

entonces

$$\sum_{j=1}^{n_o+1} \mathbf{x}_j = \left( \sum_{j=1}^{n_o} \mathbf{x}_j \right) + \frac{1}{n_o} \sum_{j=1}^{n_o} \mathbf{x}_j = \left( 1 + \frac{1}{n_o} \right) \sum_{j=1}^{n_o} \mathbf{x}_j = \frac{n_o + 1}{n_o} \sum_{j=1}^{n_o} \mathbf{x}_j = (n_o + 1) \mathbf{x}_{n_o+1} ,$$

lo que demuestra que

$$\mathbf{x}_{n_o+1} = \frac{1}{n_o + 1} \sum_{j=1}^{n_o+1} \mathbf{x}_j .$$

En tal caso, por la hipótesis sobre  $n_o$ ,

$$\begin{aligned} f(\mathbf{x}_{n_o+1}) &= f\left(\frac{1}{n_o + 1} \sum_{j=1}^{n_o+1} \mathbf{x}_j\right) \leq \frac{1}{n_o + 1} \sum_{j=1}^{n_o+1} f(\mathbf{x}_j) \\ &= \frac{1}{n_o + 1} \sum_{j=1}^{n_o} f(\mathbf{x}_j) + \frac{1}{n_o + 1} f(\mathbf{x}_{n_o+1}) ; \end{aligned}$$

o sea que

$$\frac{n_o}{n_o + 1} f(\mathbf{x}_{n_o+1}) \leq \frac{1}{n_o + 1} \sum_{j=1}^{n_o} f(\mathbf{x}_j) ,$$

vale decir

$$f\left(\frac{1}{n_o} \sum_{j=1}^{n_o} \mathbf{x}_j\right) = f(\mathbf{x}_{n_o+1}) \leq \frac{1}{n_o} \sum_{j=1}^{n_o} f(\mathbf{x}_j) ,$$

que es la afirmación de este segundo paso.

*Tercer paso:* Dado  $n \in \mathbb{N}$  arbitrario, basta encontrar un  $p$  tal que  $n \leq 2^p$  y a partir de  $2^p$  donde vale (7-3) por el primer paso, bajar en  $n - 2^p$  pasos para obtener (7-3) para  $n$  (y para todo natural menor o igual que  $2^p$ ) usando el segundo paso. Esto completa la demostración de (7-3).

La segunda y última parte es la siguiente. Si ahora  $k, n \in \mathbb{N}$  con  $k < n$ , y  $\mathbf{x}, \mathbf{y}$  son puntos arbitrarios de  $\mathcal{K}$  tenemos –tomando  $\mathbf{x}_j = \mathbf{x}$  para  $j = 1, 2, \dots, k$  y  $\mathbf{x}_j = \mathbf{y}$  para  $j = k + 1, \dots, n$ –

$$f\left(\frac{k}{n} \mathbf{x} + \left(1 - \frac{k}{n}\right) \mathbf{y}\right) = f\left(\frac{k\mathbf{x} + (k - n)\mathbf{y}}{n}\right)$$



$$\leq \frac{1}{n} \left( \sum_{j=1}^k f(\mathbf{x}) + \sum_{j=k+1}^n f(\mathbf{y}) \right) = \frac{k}{n} f(\mathbf{x}) + \left( 1 - \frac{k}{n} \right) f(\mathbf{y}),$$

lo que demuestra que (7-1) se cumple para todo  $\lambda \in (0, 1)$  que sea racional. De la continuidad de  $f$  deducimos (7-1).

2. (*Convexidad y diferenciabilidad*). La siguiente caracterización de la convexidad por medio de la diferenciabilidad debería ser familiar: Si  $f$  admite derivadas parciales de segundo orden que son continuas entonces es convexa si y sólo si el Hessiano  $\mathbf{H}(\mathbf{x})$  –cuyos elementos de matriz son

$$H_{jk}(\mathbf{x}) = \left( \frac{\partial^2 f}{\partial x_j \partial x_k} \right) (\mathbf{x}), \quad j, k = 1, 2, \dots, n -$$

es positivo semi-definido<sup>17</sup> para todo  $\mathbf{x} \in \mathcal{K}$ ; esto es el caso si los autovalores de la matriz  $\mathbf{H}(\mathbf{x})$  son no-negativos para todo  $\mathbf{x} \in \mathcal{K}$ .

Si  $f$  es diferenciable<sup>18</sup> entonces es convexa si y sólo si

$$(7-4) \quad f(\mathbf{y}) \geq f(\mathbf{x}) + (\nabla f(\mathbf{x})) \cdot (\mathbf{y} - \mathbf{x}) \text{ para todo } \mathbf{x}, \mathbf{y} \in \mathcal{K}.$$

Alternativamente, si y sólo si

$$(7-5) \quad (\mathbf{x} - \mathbf{y}) \cdot (\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})) \geq 0 \text{ para todo } \mathbf{x}, \mathbf{y} \in \mathcal{K}.$$

En ambos casos la convexidad es estricta si y sólo si hay desigualdad estricta para  $\mathbf{x} \neq \mathbf{y}$  en (7-4) y (7-5).

3. Si  $f$  es convexa, todo mínimo local es un mínimo global.  
Si  $f$  es convexa entonces es continua en el interior de  $\mathcal{K}$ <sup>19</sup>.

---

<sup>17</sup>Esto significa que la forma cuadrática (dependiente de  $\mathbf{x}$ )

$$\mathbb{R}^n \ni \mathbf{r} \mapsto \langle \mathbf{r}, \mathbf{H}(\mathbf{x})\mathbf{r} \rangle = \sum_{j,k=1}^n r_j H_{jk}(\mathbf{x}) r_k$$

toma valores no negativos para todo  $\mathbf{r} \in \mathbb{R}^n$ .

<sup>18</sup> $f : O \rightarrow \mathbb{R}$  donde  $O \subset \mathbb{R}^n$  es abierto, es diferenciable en  $\mathbf{x} \in O$  si existe una transformación lineal  $L : \mathbb{R}^n \rightarrow \mathbb{R}$  tal que

$$\lim_{\mathbf{h} \rightarrow 0} \frac{|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - L\mathbf{h}|}{\|\mathbf{h}\|} = 0.$$

En tal caso las derivadas parciales  $f_{x_j}$  de  $f$  existen en  $\mathbf{x}$  y se tiene

$$L\mathbf{y} = \sum_{j=1}^n y_j f_{x_j}(\mathbf{x}) = \mathbf{y} \cdot \nabla f(\mathbf{x}).$$

<sup>19</sup>Más es cierto:  $f$  es continua en el interior relativo de  $\mathcal{K}$  (el interior puede ser vacío, e.g.  $\mathcal{K}$  un segmento de recta en  $\mathbb{R}^2$ ).

4. Si  $f$  es convexa y congelamos  $m$  variables (donde  $0 \leq m < n$ ) entonces  $f$  es convexa en las  $n - m$  variables no congeladas. En efecto, congelando las primeras  $m$  variables,

$$\begin{aligned} & f(x_1, x_2, \dots, x_m, \lambda z_{m+1} + (1 - \lambda)y_{m+1}, \dots, \lambda z_n + (1 - \lambda)y_n) \\ &= f(\lambda(x_1, x_2, \dots, x_m, z_{m+1}, z_{m+2}, \dots, z_n) + (1 - \lambda)(x_1, x_2, \dots, x_m, y_{m+1}, y_{m+2}, \dots, y_n)) \\ &\leq \lambda f(x_1, x_2, \dots, x_m, z_{m+1}, z_{m+2}, \dots, z_n) + (1 - \lambda)f(x_1, x_2, \dots, x_m, y_{m+1}, y_{m+2}, \dots, y_n). \end{aligned}$$

En particular  $f$  es convexa en cada una de sus variables separadamente (dejando las otras fijas). Pero la inversa no es cierta: Considere

$$f(x, y) = x^2 e^{-y},$$

en  $\mathcal{C}_2$  o en  $\mathbb{R}^2$ . Calculamos  $f_x(x, y) = 2xe^{-y}$ ,  $f_{xx}(x, y) = 2e^{-y} > 0$  con lo cual  $x \mapsto f(x, y_0)$  es convexa para todo  $y_0$  fijo. También  $f_y(x, y) = -x^2 e^{-y}$ ,  $f_{yy}(x, y) = x^2 e^{-y} \geq 0$  de modo que a  $x_0$  fijo  $y \mapsto f(x_0, y)$  es convexa. Sin embargo, con  $f_{xy}(x, y) = -2xe^{-y}$ ,

$$\det(\mathbf{H}(x, y)) = -2x^2 e^{-2y} \leq 0$$

así que  $\mathbf{H}(x, y)$  no es positiva semi-definida y por ende  $f$  no es convexa.

5. Ahí van ejemplos de funciones convexas que no son homogéneas:  $f(x, y) = x^2 + y$ , y  $g(x, y) = e^{x-y}$ . En cambio  $f(x, y) = x \ln(x/(x+y))$  en  $\mathbb{R}_+^2$  es homogénea de grado 1 pero no es ni convexa ni cóncava (en la variable  $x$ ) aunque es convexa en la variable  $y$ .
6. La discusión de la transformada de L-F necesita de la introducción de la llamada semi-continuidad inferior que es una variante debilitada de la continuidad. Referimos al libro de R.T. Rockafellar (*Convex Analysis. Princeton University Press, Princeton 1970*) o al de J. van Tiel (*Convex Analysis. J. Wiley & Sons, Chichester 1984*). Sigue un listado de resultados indicando solamente las demostraciones que son simples.

- a)  $f^\#$  está definida en  $\mathbb{R}^n$  cuando  $f : S \rightarrow \mathbb{R}$  siendo  $S$  un subconjunto cualquiera de  $\mathbb{R}^n$ . Vimos (Lema 1) que  $f^\#$  es convexa.

Ya que

$$(7-6) \quad \mathbf{x}\mathbf{y} - f(\mathbf{x}) \leq f^\#(\mathbf{y}), \quad \text{para todo } \mathbf{x} \in S \text{ y todo } \mathbf{y} \in \mathbb{R}^n$$

y el miembro izquierdo de la desigualdad es finito  $f^\#(\mathbf{y})$  no puede ser  $-\infty$ . Sin embargo  $f^\#$  puede tomar el valor  $+\infty$ ; incluso se puede tener  $f^\# \equiv +\infty$  (vea el ejemplo en nota al pie del siguiente item).

- b) Habiendo definido  $f^\#$  en  $\mathbb{R}^n$  podemos definir su transformada de Legendre-Fenchel

$$(f^\#)^\#(\mathbf{x}) := \sup_{\mathbf{y} \in \mathbb{R}^n} \{\mathbf{x}\mathbf{y} - f^\#(\mathbf{y})\},$$

que es convexa y para la cual vale –use (7-6)–  $(f^\#)^\#(\mathbf{x}) \leq f(\mathbf{x})$  para todo  $\mathbf{x} \in S$ . Se obtiene entonces una minorante (función de menor valor en cada punto) de la  $f$

original que es convexa en  $\mathbb{R}^n$  (y por ende convexa en  $S$  cuando  $S$  es un convexo)<sup>20</sup>. El teorema es: **Si  $f$  es convexa y continua en el convexo  $S \subset \mathbb{R}^n$  entonces  $f = (f^\#)^\#$  en  $S$ .**

c) Resumimos aquí el contenido de la sección 26 (The Legendre Transformation) del libro de Rockafellar mencionado.

Si  $f$  a valores reales es diferenciable en el convexo abierto  $\mathcal{K}$ , entonces hay igualdad en (7-6) si y sólo si  $\mathbf{y} = (\nabla f)(\mathbf{x})$ : Dado cualquier  $\mathbf{x} \in \mathcal{K}$ , para todo  $\mathbf{y} \in \mathbb{R}^n$

$$(7-7) \quad \mathbf{y}\mathbf{x} \leq f(\mathbf{x}) + f^\#(\mathbf{y}), \text{ con igualdad si y sólo si } \mathbf{y} = \nabla f(\mathbf{x}).$$

Esto permite definir la transformada de Legendre de  $f$  como sigue. Sea  $I_{\mathbf{y}} := \{\mathbf{x} \in \mathcal{K} : \nabla f(\mathbf{x}) = \mathbf{y}\}$ . Si  $I_{\mathbf{y}}$  contiene un único punto  $\mathbf{x}_o$  entonces definimos

$$g(\mathbf{y}) := \mathbf{x}_o\mathbf{y} - f(\mathbf{x}_o).$$

Si  $I_{\mathbf{y}}$  contiene más de un punto de  $\mathcal{K}$ , y  $\mathbf{x}_1, \mathbf{x}_2 \in I_{\mathbf{y}}$ , entonces por (7-7)  $\mathbf{x}_1\mathbf{y} - f(\mathbf{x}_1) = \mathbf{x}_2\mathbf{y} - f(\mathbf{x}_2)$ .  $I_{\mathbf{y}}$  es convexo: sea  $\mathbf{x} = \lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$  y  $\mathbf{y}_o := \nabla f(\mathbf{x}_1) = \nabla f(\mathbf{x}_2)$ . Por (7-7)  $\mathbf{y}_o\mathbf{x} \leq f(\mathbf{x}) + f^\#(\mathbf{y}_o)$ . Por otro lado

$$\begin{aligned} \mathbf{x}\mathbf{y}_o - f(\mathbf{x}) &= \lambda\mathbf{x}_1\mathbf{y}_o + (1 - \lambda)\mathbf{x}_2\mathbf{y}_o - f(\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \\ &\geq \lambda[\mathbf{x}_1\mathbf{y}_o - f(\mathbf{x}_1)] + (1 - \lambda)[\mathbf{x}_2\mathbf{y}_o - f(\mathbf{x}_2)] = \lambda f^\#(\mathbf{y}_o) + (1 - \lambda)f^\#(\mathbf{y}_o) = f^\#(\mathbf{y}_o), \end{aligned}$$

usando (7-7) en la penúltima igualdad, de modo que  $\mathbf{x}\mathbf{y}_o = f(\mathbf{x}) + f^\#(\mathbf{y}_o)$  con lo que (7-7) implica que  $\nabla f(\mathbf{x}) = \mathbf{y}_o$ . Se define a la transformada de Legendre  $g$  de  $f$  sin ambigüedad por

$$g(\mathbf{y}) := \mathbf{x}\mathbf{y} - f(\mathbf{x}), \quad \mathbf{x} \in I_{\mathbf{y}};$$

quedando así definida en toda la imagen  $D$  de  $\mathcal{K}$  bajo el gradiente  $\nabla$ . El resultado básico y general es que  $g$  coincide con  $f^\#$  (la transformada de Legendre-Fenchel) en  $D$ . Sin embargo, en general,  $D$  no es convexo.

Cuando  $f$  es estrictamente convexa y cumple la condición: para toda sucesión  $\{\mathbf{x}_k : k = 1, 2, \dots\}$  en  $\mathcal{K}$  que converge a un punto en el borde de  $\mathcal{K}$

$$(7-8) \quad \lim_{k \rightarrow \infty} \|\nabla f(\mathbf{x}_k)\| = \infty ;$$

entonces se tiene  $g = f^\#$  en el convexo donde  $f^\#$  es finita. Además,  $f^\#$  es estrictamente convexa; diferenciable cumpliendo la condición (7-8); su transformada de Legendre es  $f$  y se tiene  $\nabla f^\# = (\nabla f)^{-1}$ .

---

<sup>20</sup>Con  $S = (0, \infty)$  y  $f(x) = -1/x$  (que es cóncava) se calcula  $f^\# \equiv +\infty$  y luego  $(f^\#)^\# \equiv -\infty$  que es la única minorante convexa de  $f$ .