Report no. 00/14

An abstract formulation of the concept of entropy

J D P Donnelly

Entropy is presented as a concave function relating two sets of quantities called densities and fields. It allows a simple classification of the standard relations of classical thermodynamics and yields a simple derivation of the conditions for concavity of the entropy function. It also allows a formal derivation of the equations of fluid motion. Dissipation, mixtures, and phase changes may also be included in the theory in a natural manner.

Subject classifications: AMS(MOS): 76N15, 80A05

Key words and phrases: entropy, thermodynamics, fluid flow, concave functions

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD

August, 2000

1 Introduction

The equivalence between the existence of an entropy and symmetrizability for a system of conservation laws has been explained by Harten [1] bringing together results of Godunov [2] and Mock [3]. Harten takes as illustration the Euler equations for a polytropic gas. His work has subsequently been used by a number of authors some of whom have also looked at the theoretical aspects [4],[5],[6].

Their results, however, are often obtained by unduly complicated methods. It is the intention of this paper to set out a number of simple properties of physical entropy and show how they are related. In achieving this goal we shall make use of convex analysis as described by Rockafellar in his book [7]. A brief account of the most important results is given in [8], Appendix 2. The notation for physical quantities is mostly standard but for the convenience of readers we list at the end of the paper the symbols used in the main body of the paper.

The general plan of the paper is as follows. Section two discusses the general concept of entropy, section three develops the classical equilibrium thermodynamics of a stationary system, section four considers some general properties of convex functions and applies them to entropy, section five uses the perfect gas as an example, section six extends the theory to a convected fluid, section seven deduces the Eulerian equations of fluid motion from the first principles along the lines of Harten's paper [1], section eight introduces dissipative effects, section nine treats mixtures in the case where the individual variables, e.g. density, are confounded, section ten shows how changes of phase or state may be dealt with. Appendix A justifies the transformations which allow one to pass among the various variables while preserving their basic concavity or convexity.

Some of the basic ideas for the stationary fluid appeared some years ago [9], but the later sections have been developed since.

2 General concept of entropy

We give a brief account of the foundations of thermodynamics. For a fuller account see [10], [11].

The subject of thermodynamics has its origin in the observation that not all systems are conservative; situations are observed in which mechanical energy is not conserved and work done is not recoverable. It is also observed that such systems become warm and it is postulated that the missing energy has been converted into a form of energy whose presence is detected by a quality of 'hotness'.

To abstract we consider a *thermodynamic system* K defined by a set of macroscopic quantities at least one of which is not a deformation and is referred to as the *non-deformation* co-ordinate. The system is in *thermodynamic equilibrium* if all the defining quantities, including the non-deformation co-ordinate, are in equilibrium. It is an *adiabatic* enclosure if its equilibrium can be disturbed only by mechanical means.

Any change is called *quasi-static* if the work done on K is done precisely by the forces holding K in equilibrium and each state of K during the change is an equilibrium state. A change is called *reversible* if it is quasi-static and each infinitesimal change involved can be reversed i.e. if S_1 and S_2 are neighbouring states in the transition then it is possible to pass from S_1 to S_2 and from S_2 to S_1. Let us call the co-ordinates $x_1, \ldots, x_n$, where x_n is the non-deformation co-ordinate. The work done in any change is $W = \int \sum_{k=1}^{n-1} P_k dx_k$.

The so-called *zeroth law* states that any two systems which are in equilibrium with a third are in equilibrium with each other. Suppose that the equilibrium of the first pair is given by the functional relation $f(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0$ and that of the second pair by $g(y_1, \ldots, y_m, z_1, \ldots, z_l) = 0$. These will imply the equality of the third pair $h(x_1, \ldots, x_n, z_1, \ldots, z_l) = 0$ only if eliminating y_1 between them results in the simultaneous elimination of $y_2, \ldots, y_m$, which can only be if the dependence is through a single variable $\eta(y_1, \ldots, y_m)$. Similarly we deduce the existence of $\xi(x_1, \ldots, x_n)$ and $\zeta(z_1, \ldots, z_l)$ such that the functional relations become $f^*(\xi, \eta) = 0$, $g^*(\eta, \zeta) = 0$, and $h^*(\xi, \zeta) = 0$. Assuming that a form of the implicit function theorem holds these can be written as $\xi(x_1, \ldots, x_n) = \eta(y_1, \ldots, y_m) = \zeta(z_1, \ldots, z_l)$. Thus $t = \xi(x_1, \ldots, x_n)$ may be regarded as defining an empirical quantity which we may call the *empirical temperature* to distinguish it from the physical temperature which will be defined later.

The *first law* states that work done on a system in an adiabatic transition depends only on its terminal states so that if the system passes from S' through S'' to S''' then the work done may be written $F(x', x'') + F(x'', x''') = F(x', x''')$ which implies that we can write $F(x, x'') = U(x') - U(x'')$ for some U and that for an adiabatic change $W_0 + \Delta U = 0$, while for a nonadiabatic change we have a nonzero $Q = \Delta U + W$, referred to as the *heat* absorbed by the system.

We now come to the *second law*. Consider the integral form $\omega = \sum X_i dx_i$. This is integrable if there exists an integrating factor q such that $X_i = q\partial R/\partial x_i$ for some $R(x_1, \ldots, x_n)$. The necessary and sufficient condition for this is that the exterior product $\omega \wedge d\omega = 0$. This can be written $\sum (-1)^\nu X_i \partial X_j/\partial x_k = 0$, where $\nu = 1$ for an odd permutation and $\nu = 0$ for an even permutation of any fixed triplet of subscripts. We have the theorem of Carathéodory: *If every neighbourhood of a point contains points inaccessible from it along solution curves of* $\sum X_i dx_i = 0$ *then the equation is integrable.* This leads to Carathéodory's form of the second law: *For any state there are states which cannot be reached from it by an adiabatic change.* Combined with the theorem this implies the existence of an integrating factor for the form for an adiabatic change, namely the existence of λ and ϕ such that $\lambda d\phi = dU - \sum_i P_i dx_i$.

We consider isolated systems S_A and S_B in equilibrium i.e. with $dx_i = 0$ and in thermal equilibrium with each other i.e. at the same empirical temperature so that $\lambda_A d\phi_A = dU_A$, where $\lambda_A = \lambda_A(t, U_A)$ etc. They are brought into contact to form a combined system S_C. We have $U_C = U_A + U_B$. We want to have $\phi_C = \phi_C(\phi_A, \phi_B)$ i.e. $d\phi_C = (\partial\phi_C/\partial\phi_A)d\phi_A + (\partial\phi_C/\partial\phi_B)d\phi_B$. Comparing this with $\lambda_C d\phi_C = \lambda_A d\phi_A + \lambda_B d\phi_B$ we see that this is only possible if λ_A/λ_C and λ_B/λ_C are independent of t, i.e. $\lambda_A = T(t)\lambda'_A(\phi_A)$ etc. for some function $T(t)$. Now we can define $S(\phi)$ by $dS = \lambda'_A d\phi_A$ and write $TdS = dU$. We call S the *entropy* and T the *temperature*. We have a choice of signs and choose to make S and T positive.

We now need to appeal to physical observation namely that heat energy flows from a hotter body to the colder. Mathematically $T_1 < T_2$ corresponds to $U_1 < U_2$ or $dT/dU > 0$. This means that $d^2U/dS^2 = TdT/dU > 0$ so that U is a *convex* function of S and $d^2S/dU^2 = -T^{-2}dT/dU < 0$ so that S is a *concave* function of U, i.e. $S(\mu U_1 + (1-\mu)U_2) \geq \mu S(U_1) + (1-\mu)S(U_2)$ for $0 < \mu < 1$. This means that if we mix two systems with heats per unit mass U_1 and U_2 in the proportions $\mu : 1-\mu$ the entropy of the combined system is not less than the combination in the same proportions of the original entropies per unit mass.

We shall pursue this property of concavity in the next section for S as a function of U and volume or, as we prefer to consider it, the inverse of density.

3 Thermodynamics of a stationary system

From now on we shall use lower case for all physical variables except temperature which we shall continue to denote by T, i.e. we shall write s and u for entropy and heat. We introduce the following quantities: density ρ and its inverse (sometimes called *specific volume*) τ, pressure p, and $\theta = T^{-1}$. We shall also use the Gibbs free energy defined as $g = -sT + \tau p + u$. We shall not use the enthalpy $h = u + p\tau$ and the free energy $f = u - Ts$.

Of these s, u, ρ, and τ are quantities that inhere in the substance, are measured in units per unit mass, and are called *extensive* quantities; while T, θ, p, and g are defined globally and are called *intensive* quantities or fields.

We start by observing that g is the Legendre transform of $-u$ (see [7] §26 or [12] ch 2), so that as $du = Tds - pd\tau$ we have $dg = -sdT + \tau dp$ and if $u(s,\tau)$ is a convex function $g(T,p)$ is a concave function and vice versa. This allows us to define g by

$$g = \inf_{(s,\tau)}(-sT + \tau p + u)$$

and u by

$$u = \sup_{(T,p)}(sT - \tau p + g)$$

which will be valid even when u is not differentiable as happens at a phase change (see section 10).

We have seen that u is a convex function of s. It is also a convex function of s and τ jointly. It should be stressed that this a stronger result than convexity in s and τ separately which is all that is usually proved and is sometimes referred to as stability.

Theorem 1 *Necessary and sufficient conditions for $g(T,p)$ to be a concave function and hence for $u(s,\tau)$ to be a convex function are $c_p = T(\partial S/\partial T)_p > 0$ and $\kappa_s = -\tau^{-1}(\partial\tau/\partial p)_s > 0$ or equivalently $c_v = T(\partial S/\partial T)_\tau > 0$ and $\kappa_T = -\tau^{-1}(\partial\tau/\partial p)_T > 0$.*

Remark The quantities c_v and c_p are called the specific heats at constant volume and constant pressure respectively and the quantities κ_T and κ_s are called the isothermic and isentropic compressibilities. They satisfy the relation $c_v\kappa_T = c_p\kappa_s$.

We give the proof at the end of this section.

We can now appeal to Proposition 4 in Appendix A and deduce that $s(u,\tau)$ is concave in both its arguments jointly. Together with Proposition 5 this justifies the formal division by T that takes $du = Tds - pd\tau$ to $ds = \theta du + (\theta p)d\tau$. The conjugate function of s is seen to be θg. A further application of Propositions 4 and 5 allows us to multiply through by ρ in the expressions for dg and $d(\theta g)$ and then to take conjugates. The definition of g can thus be written in the following four ways

$$\begin{aligned} \theta p &= -\theta(\rho u) + (\theta g)\rho + \rho s, && (3.1a)\\ \theta g &= \theta u + (\theta p)\tau - s, && (3.1b)\\ g &= -sT + \tau p + u, && (3.1c)\\ p &= T(\rho s) + \rho g - \rho u, && (3.1d)\end{aligned}$$

with the dual pairs $(\rho s,\ \theta p)$, $(s,\ \theta g)$, $(u,\ g)$, and $(\rho u,\ p)$, where the first of each pair is an extensive and the second an intensive variable. Writing each as a function of the appropriate variables $(\rho s)(\rho u, \rho)$, $s(u,\tau)$, $g(T,p)$, and $(\theta g)(\theta, \theta p)$ are concave, while $(\theta p)(\theta, \theta g)$, $p(T,g)$, $u(s,\tau)$, and $(\rho u)(\rho s, \rho)$ are convex. We also have the differentials

$$\begin{aligned} d(\rho s) &= \theta d(\rho u) - (\theta g)d\rho & d(\theta p) &= -(\rho u)d\theta + \rho d(\theta g) && (3.2a)\\ ds &= \theta du + (\theta p)d\tau & d(\theta g) &= ud\theta + \tau d(\theta p) && (3.2b)\\ du &= Tds - pd\tau & dg &= -sdT + \tau dp && (3.2c)\\ d(\rho u) &= Td(\rho s) + gd\rho & dp &= (\rho s)dT + \rho dg && (3.2d)\end{aligned}$$

which may also be interpreted as subdifferentials.

We now prove the theorem, namely that $g(T,p)$ is concave which as we saw will imply that $u(s,\tau)$ and $(\rho u)(\rho s,\rho)$ are convex and $s(u,\tau)$ and $(\rho s)(\rho u,\rho)$ are concave.

Proof Since g concave is equivalent to $-g$ convex we may start from

$$-dg = sdT - \tau dp$$

write

$$\begin{aligned} ds &= \alpha dT + \beta dp\\ -d\tau &= \beta dT + \gamma dp\end{aligned}$$

and apply the conditions for the matrix

$$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$$

to be positive definite, which are $\alpha > 0$ and $\gamma - \beta^2/\alpha > 0$ or else $\gamma > 0$ and $\alpha - \beta^2/\gamma > 0$. But these are $(\partial s/\partial T)_p > 0$, $-(\partial\tau/\partial p)_s > 0$ or $-(\partial\tau/\partial p)_T > 0$, $(\partial s/\partial T)_\tau > 0$. Applying this test to any other of the relations leads to one or other of these criteria. ■

4 Entropy as a concave function

As we have noted f convex is the same as $-f$ concave and vice versa. We shall state everything in terms of convex functions as this is the usual convention. If we wish to state the corresponding result for concave functions then it is necessary to replace sup by inf.

Let $x = (\xi_1, \ldots, \xi_n) \in R^n$ and $f : R^n \to R^n$ be a convex function, that is to say

$$f(\mu x_1 + (1-\mu)x_2) \leq \mu f(x_1) + (1-\mu) f(x_2) \text{ for } 0 < \mu < 1.$$

If f is differentiable we may write the vector of partial derivatives $x^* = \partial f / \partial x$. If we let $\mu \to 1$ in the definition we find

$$f(z) - f(x) \geq \langle z - x, x^* \rangle$$

for any $z \in R^n$. If f is not differentiable this relation may still hold and any x^* satisfying it is called a *subgradient* of f. The set of all such x^* is called the *subdifferential* of f and is denoted $\partial f(x)$. In general it is a closed convex set in R^n which reduces to a point when f is differentiable. In the case where $n = 1$ this means that, where f has a 'corner', $\partial f(x)$ is a closed interval.

We define the *conjugate* f^* of f by

$$f^*(x^*) = \sup_x (\langle x, x^* \rangle - f(x)).$$

Since the subgradient inequality may be written

$$\langle x, x^* \rangle - f(x) \geq \langle z, x^* \rangle - f(z), \qquad \forall z,$$

we have that the following are equivalent

$$\text{(a)} \quad x^* \in \partial f(x), \tag{4.1a}$$

$$\text{(b)} \quad \langle z, x^* \rangle - f(z) \text{ achieves its maximum at } z = x, \tag{4.1b}$$

$$\text{(c)} \quad f(x) + f(x^*) \leq \langle x, x^* \rangle, \tag{4.1c}$$

$$\text{(d)} \quad f(x) + f(x^*) = \langle x, x^* \rangle. \tag{4.1d}$$

In general $\partial f(x)$ is a multivalued mapping from R^n to R^n. Substituting x_0 and x_1 for x and z in the definition and adding we find $\langle x_1 - x_0, x_1^* - x_0^* \rangle \geq 0$, i.e. ∂f is a *monotone* mapping from R^n to R^n. This concept can be extended to include an indefinite number of points. A multivalued mapping ρ from R^n to R^n is said to be *cyclically monotone* if

$$\langle x_1 - x_0, x_0^* \rangle + \langle x_2 - x_1, x_1^* \rangle + \cdots + \langle x_0 - x_m, x_m^* \rangle \leq 0$$

for any set of pairs (x_i, x_i^*) such that $x_i^* \in \rho(x_i)$ for $i = 1, \ldots, m$, where m is arbitrary. This set forms the *graph* of ρ. A *maximal* cyclically monotone mapping is one whose graph is not properly contained in any other cyclically monotone mapping. Using the subgradient inequality it is easy to show that if f is a convex function ∂f is cyclically monotone. In fact the following can be proved.

Theorem 2 *Let ρ be a multivalued mapping from R^n to R^n. In order for there to exist a closed proper convex function f on R^n such that $\rho(x) \subset \partial f(x)$ for every x it is necessary and sufficient that ρ be cyclically monotone.*

Proof The condition is clearly necessary since ∂f is cyclically monotone. Now suppose that ρ is cyclically monotone. Take any pair (x_0, x_0^*) in the graph of ρ (if ρ is empty the conclusion is trivial) and define f on R^n by

$$f(x) = \sup\{\langle x - x_m, x_m^*\rangle + \cdots + \langle x_1 - x_0, x_0^*\rangle\},$$

where the supremum is taken over all finite pairs (x_i, x_i^*), $i = 1, \ldots, m$, in the graph of ρ. Since f is the supremum of a finite collection of affine functions it is a closed convex function. By the cyclical monotonicity of ρ, $f(x_0) = 0$ and hence f is proper (not identically $-\infty$). Let x and x^* satisfy $x^* \in \rho(x)$. We shall show that $x^* \in \partial f(x)$. It will suffice to show that for any $\alpha < f(x)$ and any $y \in R^n$

$$f(y) > \alpha + \langle y - x, x^*\rangle.$$

Given $\alpha < f(x)$ there exist pairs $(x_i, x_i^*), i = 1, \ldots, m$ such that $x_i^* \in \rho(x_i)$ and

$$\alpha < \langle x - x_m, x_m^*\rangle + \cdots + \langle x_1 - x_0, x_0^*\rangle.$$

Setting $x_{m+1} = x$ and $x_m^* = x^*$ we have by the definition of f

$$f(y) \geq \langle y - x_{m+1}, x_{m+1}^*\rangle + \langle x_{m+1} - x_m, x_m^*\rangle + \cdots + \langle x_1 - x_0, x_0^*\rangle > \langle y - x, x^*\rangle + \alpha$$

as required. This proves that $\rho(x) \subset \partial f(x)$. ■

We conclude by quoting without proof one further theorem.

Theorem 3 *The subdifferential mappings of the closed proper convex functions on R^n are the maximally cyclically monotone mappings from R^n to R^n. The function is uniquely determined by its subdifferential mapping up to an additive constant.*

We can see that this provides an alternative method of defining s as

$$s(u) = \sup\{\frac{u - u_m}{T_m} \cdots + \frac{u_1 - u_0}{T_0}\},$$

where the supremum is taken over all finite pairs (u_i, T_i), $i = 1, \ldots, m$ in a sequence of reversible changes. This generalizes the classical definition

$$s = \int \frac{du}{T}.$$

There is one further interesting observation concerning properties of convex functions. Let us consider a *concave* function $s(u)$ where $u = (v_1, \ldots, v_n)$ may be considered as a general entropy function. Let A be an $n \times m$ matrix of rank m, where $n > m$. Let $\tilde{u} = (\tilde{v}_1, \ldots, \tilde{v}_m)$ and define $\tilde{s}(\tilde{u}) = \sup\{s(u) : Au = \tilde{u}\}$. Then the conjugate function $\tilde{g}$ of $\tilde{s}$ is given by

$$\tilde{g}(\tilde{\Theta}) = \inf_{\tilde{u}}\{\langle \tilde{u}, \tilde{\Theta}\rangle - \tilde{s}(\tilde{u})\} = \inf_{\tilde{u}} \inf_{Au=\tilde{u}}\{\langle \tilde{u}, \tilde{\Theta}\rangle - s(u)\} = \inf_{u}\{\langle u, A^*\tilde{\Theta}\rangle - s(u)\} = g(A^*\tilde{\Theta}),$$

where g is the conjugate function of s and since

$$\tilde{s}(\tilde{u}) + \tilde{g}(\tilde{\Theta}) = \langle \tilde{u}, \tilde{\Theta} \rangle = \langle Au, \tilde{\Theta} \rangle = \langle u, A^*\tilde{\Theta} \rangle = s(u) + g(A^T\Theta),$$

we see that $\tilde{\Theta} \in \partial\tilde{s}(\tilde{u})$ if and only if $A^T\tilde{\Theta} \in \partial s(u)$ with $Au = \tilde{u}$.

As an application let us take $A = (I, 0)$ so that $\tilde{u} = (v_1, \ldots, v_m)$. Then we find that if $\Theta = (\theta_1, \ldots, \theta_n)$ then $\tilde{\Theta} = (\theta_1, \ldots, \theta_m, 0, \ldots, 0)$ and that $\tilde{s}$ is found by using $\theta_{m+1} = \cdots = \theta_n = 0$ to solve for $v_{m+1}, \ldots, v_n$ in terms of $v_1, \ldots, v_m$, i.e. $(\theta_1, \ldots, \theta_m) \in \partial\tilde{s}(v_1, \ldots, v_m)$ if and only if $(\theta_1, \ldots, \theta_m, 0, \ldots, 0) \in \partial s(v_1, \ldots, v_n)$, so that $\theta_1, \ldots, \theta_m$ bear the same relation to $v_1, \ldots, v_m$ in both s and $\tilde{s}$.

We shall return to this theme in section 9.

5 Perfect gases

For a perfect or ideal gas we have

$$p\tau = RT,$$

where R is a constant called the universal gas constant. This relation is equivalent to $\tau(\theta p) = R$, so that from the expression (3.2b) for $\mathrm{d}(\theta g)$ we have

$$\left(\frac{\partial u}{\partial(\theta p)}\right)_\theta = \left(\frac{\partial \tau}{\partial \theta}\right)_{(\theta p)} = 0$$

and hence $u = u(\theta)$ i.e. u depends only on temperature.
We may use $\mathrm{d}u = T\mathrm{d}s - p\mathrm{d}\tau$ to obtain

$$\left(\frac{\partial}{\partial \tau}\right)_s (p\tau - RT) = \left(\frac{\partial p}{\partial \tau}\right)_s \tau + p + R\frac{\mathrm{d}T}{\mathrm{d}u}p = 0,$$

which may also be written

$$\left(\frac{\partial p}{\partial \rho}\right)_s = c^2(T) = RT\left(1 + R\frac{\mathrm{d}T}{\mathrm{d}u}\right) = RT\gamma(T),$$

where $c(T)$ is the isentropic sound speed, $\gamma(T)$ is a dimensionless quantity satisfying

$$R = (\gamma(T) - 1)\frac{\mathrm{d}u}{\mathrm{d}T},$$

and $\frac{\mathrm{d}u}{\mathrm{d}T}$ is the specific heat at constant volume. We note that (3.2b) may also be written

$$\mathrm{d}(\theta g) = (u + \tau p)\mathrm{d}\theta + \tau\theta\mathrm{d}p.$$

The coefficient of $\mathrm{d}\theta$ is of course the enthalpy h defined in Section 3. Its derivative with respect to T is called the specific heat at constant pressure and has the value

$$\frac{\mathrm{d}u}{\mathrm{d}T} + R = \gamma(T)\frac{\mathrm{d}u}{\mathrm{d}T}$$

so that $\gamma(T)$ may be identified as the ratio of the two specific heats.

A gas for which $\frac{\mathrm{d}u}{\mathrm{d}T}$ and hence $\gamma(T)$ is a constant is referred to as a *polytropic* gas. Using $u = c_v T$ i.e. $\theta = c_v/u$ in (3.2b) we have the perfect differential of

$$s - s_0 = c_v \ln\, u + R\, \ln\, \tau,$$

which may also be written

$$u = \tau^{1-\gamma} \exp((s - s_0)/c_v).$$

In terms of quantities per unit volume we have

$$\rho s = \rho(s_0 + c_v \ln(\rho u) - c_p \ln \rho), \tag{5.1}$$

where $c_p = c_v + R$, the specific heat at constant pressure.

6 Thermodynamics of a fluid in motion

For a moving fluid the energy density per unit mass is given by

$$e = u + \tfrac{1}{2} v^i v^i,$$

where to the internal energy u we must add the kinetic energy $\frac{1}{2} v^i v^i$. (The repeated index implies summation).

We see immediately that since $u(s, \tau)$ is convex so is $e(s, \tau, v^i)$ and

$$\mathrm{d}e = T\mathrm{d}s - p\mathrm{d}\tau + v^i \mathrm{d}v^i.$$

Let us denote the left hand side of (3.1c) by g_0 and redefine the Gibbs free energy as

$$g = g_0 - \tfrac{1}{2} v^i v^i = -sT + \tau p - v^i v^i + e,$$

so that $-g$ is the convex dual of e, i.e. $g(T, p, v^i)$ is concave and

$$\mathrm{d}g = -\ s\mathrm{d}T + \tau \mathrm{d}p - v^i \mathrm{d}v^i.$$

The complete set of dual equations (3.1) becomes

$$\theta p = -\theta(\rho e) + (\theta g)\rho + \theta \rho v^i v^i + \rho s \tag{6.1a}$$
$$\theta g = \theta e + (\theta p)\tau - \theta v^i v^i - s \tag{6.1b}$$
$$g = -sT + \tau p - v^i v^i + e \tag{6.1c}$$
$$p = T(\rho s) + \rho g + \rho v^i v^i - \rho e. \tag{6.1d}$$

We have the concave functions $\rho s(\rho e, \rho, \rho v^i)$, $s(e, \tau, v^i)$, $g(T, p, v^i)$, $(\theta g)(\theta, \theta p, \theta v^i)$ and the convex functions $(\theta p)(\theta, \theta g, \theta v^i)$, $p(T, g, v^i)$, $e(s, \tau, v^i)$, $(\rho e)(\rho s, \rho, \rho v^i)$.

The differentials are given by

$$\begin{aligned}
\mathrm{d}(\rho s) &= \theta \mathrm{d}(\rho e) - (\theta g)\mathrm{d}\rho - \theta v^i \mathrm{d}(\rho v^i) && (6.2a)\\
\mathrm{d}(\theta p) &= -(\rho e)\mathrm{d}\theta + \rho \mathrm{d}(\theta g) + (\rho v^i)\mathrm{d}(\theta v^i) && (6.2b)\\
\mathrm{d}s &= \theta \mathrm{d}e + (\theta p)\mathrm{d}\tau - (\theta v^i)\mathrm{d}v^i && (6.2c)\\
\mathrm{d}(\theta g) &= e\mathrm{d}\theta + \tau \mathrm{d}(\theta p) - v^i \mathrm{d}(\theta v^i) && (6.2d)\\
\mathrm{d}e &= T\mathrm{d}s - p\mathrm{d}\tau + v^i \mathrm{d}v^i && (6.2e)\\
\mathrm{d}g &= -s\mathrm{d}T + \tau \mathrm{d}p - v^i \mathrm{d}v^i && (6.2f)\\
\mathrm{d}(\rho e) &= T\mathrm{d}(\rho s) + g\mathrm{d}\rho + v^i \mathrm{d}(\rho v^i) && (6.2g)\\
\mathrm{d}p &= (\rho s)\mathrm{d}T + \rho \mathrm{d}g + (\rho v^i)\mathrm{d}v^i. && (6.2h)
\end{aligned}$$

7 Fluid flow I

We consider the motion of an inviscid, perfectly conducting fluid. Since we can neglect viscous stresses and heat conduction, we may write the equations of motion in conservation form

$$u_t + f^i_{,i} = 0, \tag{7.1}$$

where u is a vector of unknowns, $f^1, \cdots, f^{\mathrm{d}}$ are the flux vectors of the same dimension, and d is the space dimension.

Suppose that there exists a new vector of unknowns v for which (7.1) can be written in symmetric form

$$u_v v_t = f^i_v v_{,i} = 0,$$

where $v_{,i}$ stands for the derivative with respect to x_i and we sum over the repeated index. By symmetric form we mean that u_v, f^i_v are symmetric matrices and u_v is positive definite. The symmetry implies that u_v, f^i_v are gradients with respect to v i.e. there exist scalar functions $q(v), r^i(v)$ such that

$$q_v = u^T, \;\; r^i_v = (f^i)^T, \tag{7.2}$$

the superscript T denoting transpose. The positive definiteness of u_v is equivalent to the convexity of q.

Denote by $U(u)$ the convex dual of q

$$U(u) = \langle u, v\rangle - q(v), \tag{7.3}$$

which, for sufficient smoothness, is the Legendre transform, and denote by F^i the Legendre transform of r^i

$$F^i(f^i) = \langle f^i, v\rangle - r^i(v). \tag{7.4}$$

We have

$$\mathrm{d}U = \langle v, \mathrm{d}u\rangle, \;\; \mathrm{d}F^i = \langle v, \mathrm{d}f^i\rangle,$$

so that v^T transforms (7.1) into

$$U_t + F^i_{,i} = 0. \tag{7.5}$$

This is the result of Godunov [2], [1] that, if (7.1) can be symmetrized, there exists an entropy $U(u)$ and entropy fluxes $F^i(f^i)$ such that (7.5) holds.

The converse, due to Mock [3], [1], is that, if U, F^i exist such that (7.5) holds, then (7.1) is symmetrizable. To see this we use (7.3), (7.4) to define the duals q and r^i such that (7.2) holds, which again implies that u_v, f^i_v are symmetric and u_v is positive definite.

Let us now deduce the equations of fluid motion from the hypothesis that the entropy flux equation (7.5) takes the form

$$(\rho s)_t + (\rho s v^i)_{,i} = 0. \tag{7.6}$$

Now the dual of ρs is θp. Let us guess the dual of the flux $\rho s v^i$ to be $\theta p v^i$. Using (6.2b) we find.

$$\begin{aligned} \mathrm{d}(\theta p v^i) &= v^i \mathrm{d}(\theta p) + p\mathrm{d}(\theta v^i) - p v^i \mathrm{d}\theta \\ &= -(\rho e + p)v^i \mathrm{d}\theta + \rho v^i \mathrm{d}(\theta g) + (\rho v^i v^j + p\delta^{ij})\mathrm{d}(\theta v^j), \end{aligned} \tag{7.7}$$

where δ^{ij} is the Kronecker delta. We can now see that the entropy flux satisfies

$$\mathrm{d}(\rho s v^i) = \theta \mathrm{d}((\rho e + p)v^i) - \theta g \mathrm{d}(\rho v^i) - \theta v^j \mathrm{d}(\rho v^i v^j + p\delta^{ij}). \tag{7.8}$$

We may note that (6.2g), and (6.2h) also yield a dual pair deducible from (7.8), viz. $(\rho e + p)v^i$ and pv^i. The other two pairs based on densities per unit mass do not yield conservation forms of the equations of fluid motion.

Let us now write down the equations yielded by (7.8). They are

$$\begin{aligned} (\rho e)_t + ((\rho e + p)v^i)_{,i} &= 0 \\ \rho_t + (\rho v^i)_{,i} &= 0 \\ (\rho v^j)_t + (\rho v^i v^j)_{,i} + p_{,j} &= 0, \end{aligned} \tag{7.9}$$

where we have unparcelled the final p term.

We have further confirmation that the dual variable p is the pressure as commonly understood.

It is interesting to note that if we start with (6.2e) then the equations (not in flux form) of fluid flow motion may be written

$$\begin{aligned} Ds &= 0 \\ D\tau &= \tau\delta^i_j v^i_{,j} \\ Dv^j &= -\tau\delta^i_j p_{,i}, \end{aligned}$$

where $D = \frac{\partial}{\partial t} + v^i\frac{\partial}{\partial x^i}$ is the convected time derivative. Since p and v^i are the dual variables of T and v^i the right hand side is a symmetric rank-two matrix of the form $ab^T + ba^T$ multiplying the x^i derivative of the intensive variable (T, τ, v^i), where a and b are vectors, in this case unit vectors. It follows from the contravariant relations analyzed in Appendix A that this result holds for the other three sets of equations (6.2a), (6.2c), (6.2g). The inverse of a finite rank modification of a matrix is of course a modification of the inverse of the same rank. It is possible that this result may have some numerical application.

8 Fluid flow II

In order to include dissipative effects such as viscosity and heat conduction in the theory of fluid flow we shall modify (7.1) by incorporating an additional divergence term so that it becomes

$$u_t + f^i_{,i} + Q^i_{,i} = 0. \tag{8.1}$$

Taking the inner product with v we obtain the modified versions of (7.5)

$$U_t + F^i_{,i} + \langle v, Q^i_{,i} \rangle = 0, \tag{8.2}$$

which, on integration over a finite region Ω, becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega U + \int_{\partial\Omega} F^i n_i = \int_\Omega \langle \mathrm{grad} v, Q \rangle - \int_{\partial\Omega} \langle v, Q^i \rangle n^i,$$

where (n_i) is the normal derivative on $\partial\Omega$. The first integrand on the right hand side is the local rate of generation of entropy in the fluid, which is assumed, by traditional hypothesis, to be positive. For small gradients the dependence of Q on grad v will be approximately linear, $Q = L$ grad v, and the requirement of positivity demands that L be positive definite matrix. For larger gradients some nonlinear dependence may be assumed e.g. Q is the subgradient of a convex function of grad v that vanishes at the origin. The simple linear form of dependence allows an existence proof in the case of a stationary fluid based on the results of Brezis [13] (see [9]). The phenomenological theory of Onsager derives the result from more fundamental principles. For a simple exposition consult [14] Chapter 7.

Let us return to (7.10) and include a number of simple dissipative terms. Consider the forces acting on the fluid. These are of two types, a body force and a pressure tensor. We denote the first by a vector (f_j) and the second by (p^{ji}), where p^{ji} is the force in direction j on a surface whose normal is in direction i. The force on a volume element due to pressure is

$$-\int_{\partial\Omega} p^{ji} n_i = -\int_\Omega p^{ji}_{,i}$$

where n_i is the *outward* normal and the third (momentum) equation of (7.10) becomes

$$(\rho v^j)_t = (\rho v^i v^j)_{,i} + p^{ji}_{,i} = f^j,$$

where the scalar pressure had been included in the pressure tensor.

The kinetic energy will be increased by the work done on the fluid by these forces, while the internal energy will be increased by internal generation of heat and heat fluxes in the fluid. Denoting the heat flux by q^i, the total heat flux into Ω is

$$-\int_{\partial\Omega} q^i n_i = -\int_\Omega q^i_{,i},$$

so that the first (energy) equation of (7.10) is now

$$(\rho e)_t + (\rho e v^i)_{,i} + (v^j p^{ji})_{,i} + q^i_{,i} = f^j v^j + \phi,$$

where ϕ is the internal heating. We leave the second (continuity) equation unaltered not wishing to admit the creation of matter.

We now rewrite (7.6) using the dual variables to multiply the modified equations. We find

$$(\rho s)_t + (\rho s v^i)_{,i} = \theta p v^i_{,i} - \theta p^{ji} v^j_{,i} - \theta q^i_{,i} + \theta\phi.$$

To make sense of this we need to use some general results concerning the possible linear dependence of vectors and tensors on other vectors and tensors. We shall make use of [14] Chapter 8.

The fourth, scalar, term is straightforward, $\theta\phi > 0$ implies that θ and ϕ increase or decrease together. Since $\theta = T^{-1}$ and ϕ is an internal or spontaneous heat source this means that raising T must inhibit ϕ.

The third, vector, term should be considered as acting on a body of fluid

$$-\int_\Omega \theta q^i_{,i} = -\int_{\partial\Omega} \theta q^i n_i + \int_\Omega \theta_{,i} q^i,$$

of which the first term on the right hand side is minus the heat flux out of the fluid and the second is the net entropy increase which is positive provided

$$q^i = \kappa\theta_{,i},$$

where κ is a positive constant. As $\theta = T^{-1}$ this says that heat flows down a temperature gradient.

The first and second terms need to be taken together. The pressure tensor may be decomposed into three parts, the trace, an anti-symmetric part and the remaining symmetric part. The anti-symmetric part is usually neglected as it is the dissipation due to internal rotation of the fluid and only contributes in the case of "rough" molecules. We are left with a symmetric tensor for which the most general isotropic linear dependence may be written $p^{ij} = p\delta^{ij} + \pi^{ij}, p$ being the pressure, and π^{ij} is given by

$$\pi^{ij} = a_{ijkl} v^k_{,l}$$

with

$$a_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}),$$

where μ is the coefficient of shear viscosity, familiar from elementary accounts of the subject, and λ is called the second coefficient of viscosity. The quantity $\lambda + \frac{2}{3}\mu$ is called the bulk (or volume) viscosity. It is the constant of proportionality between the traces π^{ii} and $v^k_{,k}$. The contribution to the entropy is positive proved $\lambda + \frac{2}{3}\mu$ and μ are both positive.

9 Fluid mixtures

We turn now to a subject that can be made to appear unnecessarily complicated. It is therefore desirable to state at the beginning the basic assumptions. They are simply that the mixture is treated as having a single density, energy density, and momentum density, and consequently single Gibbs energy, temperature, and velocity. The motion of the different components within the fluid is assumed to be wholly diffusive and subject to the principle of entropy increase given in the previous paragraph. Thus ρs is to be a function of ρe, ρ, ρv^j and the vector (n^i) of the relative concentrations of the components.

Let us revert to the notation of section 4. We start with a first attempt to write down S and G for a mixture

$$S = \textstyle\sum n_i S_i - k \sum n_i \log n_i \quad \text{with } \sum n_i = 1.$$

The second term is called the entropy of mixing and the expression we have written down is the simple form to be found in any text book on thermodynamics. It is in fact the function f of the example following [7] Thm 16.4. Assume that $U = \sum n_i U_i$ and calculate the conjugate of S. As before the variable conjugate to U will be denoted by θ and the variable conjugate to n_i, usually referred to as the chemical potential, will be denoted by μ_i. We have

$$\begin{aligned} G(\theta, (\mu_i)) &= \inf_{(U_i, n_i)} \{\textstyle\sum n_i \mu_i + \langle \theta, \sum n_i U_i \rangle - \sum n_i S_i + k f((n_i))\} \\ &= \inf_{(n_i)} \{\textstyle\sum n_i (G_i(\theta) + \mu_i) + k f((n_i))\} \\ &= -k \sup_{(n_i)} \{\textstyle\sum n_i k^{-1}(-G_i(\theta) - \mu_i) - f(n_i)\} \\ &= -k f^*(-(\mu_i + G_i(\theta))/k) \\ &= -k \log \textstyle\sum_i \exp(-(\mu_i + G_i(\theta))/k), \end{aligned}$$

using the expression for f^*, the *convex* conjugate of f found in the example. The assumption $U = \sum n_i U_i$ is a simplification. Let us calculate the conjugate of G.

$$\begin{aligned} S(U, (n_i)) &= \inf_{\theta, (\mu_i)} \{\langle \theta, U \rangle + \textstyle\sum n_i \mu_i + k f^*(-(\mu_i + G_i)/k)\} \\ &= \inf_{\theta} \{\langle \theta, U \rangle - k \sup_{(\mu_i)} \{-n_i \mu_i / k - f^*(-(\mu_i + G_i)/k)\}\} \\ &= \inf_{\theta} \{\langle \theta, U \rangle - \textstyle\sum n_i G_i(\theta)\} - k f((n_i)). \end{aligned}$$

We may now use [7] Thm 16.4. For a concave function g and $\lambda > 0$, the conjugate of λg is simply

$$\begin{aligned} (\lambda g)^*(x^*) &= \inf_x \{\langle x, x^* \rangle - \lambda g(x)\} \\ &= \lambda \inf_x \{\langle x, \lambda^{-1} x^* \rangle - g(x)\} \\ &= \lambda g^*(\lambda^{-1} x). \end{aligned}$$

So $(n_i G_i)^*(V_i) = n_i G^*(n_i^{-1} V_i)$, and the conjugate of $\sum n_i G_i$ is given by

$$\begin{aligned} (\textstyle\sum n_i G_i)^*(U) &= \sup\{\textstyle\sum (n_i G_i)^*(V_i) \mid \sum V_i = U\} \\ &= \sup\{\textstyle\sum n_i G_i^*(n^{-1} V_i) \mid \sum V_i = U\} \\ &= \sup\{\textstyle\sum n_i G_i^*(U_i) \mid \sum n_i U_i = U\}, \end{aligned}$$

where we have set $U_i = n_i^{-1}V_i$. Thus $S(U,(n_i)) = S_0(U,(n_i)) - kf((n_i))$, where

$$S_0(U,(n_i)) = \sup\{\textstyle\sum n_i S_i(U_i) | \sum n_i U_i = U\}.$$

The entropy is thus made up of two parts: the first part is the sum of the entropies of the individual components, the second is an entropy of mixing. The individual entropies are calculated using the individual values of U_i given by the infimum (which is attained). Because the U_i are chosen in this way the corresponding $\theta \in \partial S_i(U_i)$ are all the same. If we chose to measure the individual values of U_i then we should have a more elaborate model — with more variables — but it would no longer be a mixture in the above sense and, because the θ_i for each component would not necessarily be the same, it would no longer be in equilibrium.

10 Change of phase

It may be that a substance can exist in two phases, liquid and solid, gas and liquid, or different alloys. A simple model for a change of phase is to set $G = \min(G_1, G_2)$ where G_1 and G_2 are the Gibbs functions associated with the two phases. Then [7] Thm 16.5 tells us that

$$S(U) = \inf\{(1-\lambda)S_1(U_1) + \lambda S_2(U_2) \mid (1-\lambda)U_1 + \lambda U_2 = U\}.$$

As a simple example in a single variable consider

$$f_1 = \tfrac{1}{2}x^2 \quad , \quad f_2 = \tfrac{1}{2}x^2 + bx$$

with $b > 0$.
We have (as an easy exercise)

$$f_1^*(x^*) = \tfrac{1}{2}(x^*)^2, \quad f_2(x^*) = \tfrac{1}{2}(x^* - b)^2.$$

Let $f = \sup(f_1, f_2)$. Then

$$\begin{aligned} f(x) = f_1(x) &\qquad x \leq 0 \\ f_2(x) &\qquad x \leq 0; \end{aligned} \tag{10.1}$$

and $\partial f(x)$ consists of a single value for $x \neq 0$.
Now

$$f^*(x^*) = \inf\{(1-\lambda)f_1^*(x_1^*) + \lambda f_2^*(x_2^*) \mid x^* = (1-\lambda)x_1^* + \lambda x_2^*\},$$

and we find

$$\begin{aligned} f^*(x^*) = f_1^*(x^*) &\qquad x^* \leq 0, \\ f_2^*(x^*) &\qquad x^* \geq b, \\ 0 &\qquad 0 < x^* < b. \end{aligned}$$

If we calculate $\partial f(0)$ we may use the subgradiant inequality. If $x^* \in \partial f(0)$ then $\frac{1}{2}z^2 \geq zx^*$ for $z < 0$ implying $x^* \geq 0$ and $\frac{1}{2}z^2 + bz \geq zx^*$ for $z > 0$ implying $x^* \leq b$. Hence $\partial f(0)$ is closed interval $[0, b]$. Geometrically any line through the origin whose gradient lies in this interval supports the graph of f at $x = 0$.

More generally we see that for a single pair of dual variables U, θ we may expect a value of θ_0 for which $G_1(\theta_0) = G_2(\theta_0)$, where the value of U jumps by a discrete amount from $[G_1'(\theta_0), G_2'(\theta)]$. At the same time $S(U)$ also jumps by a discrete amount. For any U there is a value of λ for which the above infimum is attained and this λ may be interpreted as the proportion of phase two present. For vectors U, θ the set of θ for which $G_1(\theta) = G_2(\theta)$ will be a hypersurface and, for these values of $\theta, \partial G(\theta)$ will be a set. Again the value of λ for which the infimum is attained may be interpreted as the proportion of phase two.

References

[1] A. Harten. High resolution schemes for conservation laws. *Journal of Computational Physics*, 49:357–393, 1983.

[2] S.K. Godunov. An interesting class of quasilinear systems. *Soviet Mathematics*, 2(2):947–949, 1961.

[3] M.S. Mock. Systems of conservation laws of mixed type. *Journal of Differential Equations*, 37:70–88,1980.

[4] F. Bourdel, J.P. Croisille, P. Delorme and P.A. Mazet. On the approximation of K-diagonizable hyperbolic systems by finite-elements—applications to the Euler equations and to gasaeous-mixtures. *Recherche Aerospatiale*, 5:15–34, 1989.

[5] J.P. Croisille and P. Delorme. Kinetic symmetrizations and pressure laws for the Euler equations. *Physics D*, 57(3-4):395–416, 1992.

[6] F. Dubois. Concavity of the thermostatic entropy and convexity of the Lax mathematical entropy. *Recherche Aerospatiale*, 3:77–80, 1990.

[7] R.T. Rockafellar. *Convex Analysis.* Princeton, 1970.

[8] G. A. Maugin. The Thermomechanics of Plasticity and Fracture. Cambridge University Press, 1992

[9] J.D.P. Donnelly. A model for non-equilibrium thermodynamic processes involving phase changes. *Journal of the Institute of Mathematics and Its Applications*, 24:425–38, 1979

[10] C. J. Adkins. *Equilibrium Thermodynamics.* Cambridge University Press, 1983.

[11] H. A. Buchdahl. *The Concepts of Classical Thermodynamics.* Cambridge University Press, 1966

[12] M. J. Sewell. *Maximum and Minimum Principles : a Unified Approach, with Applications.* Cambridge University Press, 1992

[13] H. Brezis. Problèmes unilatéraux. *Journal de Mathématiques pure et appliqués* 51:1–168, 1972.

[14] L.C. Woods. *The Theormodynamics of Fluid Systems.* Clarendon Press, Oxford, 1975.

Appendix A Convex inverse functions

In section 3 we wrote down four pairs of conjugate functions. Transition among these pairs was effected by inversion of some of the functions with respect to one argument. In this appendix we show that these inversions carry concave functions to convex ones or vice versa. We shall also show how subgradients transform.

Proposition 4 *Let f be jointly convex (resp. concave) in a scalar ξ and a vector x and strictly increasing in ξ. Then $f(\xi, x) = \zeta$ defines a unique inverse function g by $g(\zeta, x) = \xi$ jointly concave (resp. convex) in ζ and x.*

Proof Since f is strictly increasing in ξ, a unique inverse function exists. We show that it is concave.
Let $\zeta_i = f(\xi_i, x_i)$ $i = 1, 2$. By the convexity of f we have for λ satisfying $0 < \lambda < 1$

$$f((1-\lambda)\xi_1 + \lambda\xi_2, (1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)\zeta_1 + \lambda\zeta_2,$$

and by monotonicity in the first argument

$$(1-\lambda)\xi_1 + \lambda\xi_2 \leq g((1-\lambda)\zeta_1 + \lambda\zeta_2, (1-\lambda)x_1 + \lambda x_1).$$

As $\xi_i = g(\xi_i, x_i)$ this shows that g is concave. Since the correspondence between f and g is one-to-one the transition from a concave to a convex function is also valid. ∎

Corollary 4A *If $f = \max(f_1, f_2)$, where f_1, f_2, have the properties f in Proposition 4, then $g = \min(g_1, g_2)$, where g, g_1, g_2 are the inverse functions of f, f_1, f_2.*

Remark In the application to section 3 monotonicity is a consequence of density and temperature being positive.

Proposition 5 *Let f and g be as in Proposition 4. Then $(\mu, y) \in \partial f(\xi, x)$ is equivalent to $(\eta^{-1}, -\eta^{-1}y) \in \partial g(\zeta, x)$.*

Proof Set $\zeta' = f(\xi', x')$ for abitrary ξ', x'. Then the subgradient inequality for f may be written

$$\zeta' - \zeta \geq (\xi' - \xi)\eta + (x' - x, y)$$

or

$$\xi' - \xi \leq (\zeta' - \zeta)\eta^{-1} + (x' - x, -\eta^{-1}y),$$

which is the subgradient inequality for g. ∎
Rockafellar eschews the use of the more correct "supergradient" in this context.

Corollary 5A *The conjugate functions* f^* *and* g^* *are related by*

$$\eta^{-1}f^*(\eta, y) = g^*(\eta^{-1}, -\eta^{-1}y).$$

Proof From (4.1d) $u^* \in \partial\phi(u)$ if and only if $\phi(u) + \phi^*(u) = \langle u, u^*\rangle$. ■

Let us assume that f and g are twice differentiable and let us set $\omega = \eta^{-1}, w = -\eta^{-1}y$. We may then write

$$\begin{aligned} \mathrm{d}\zeta &= \eta \mathrm{d}\xi + y^T \mathrm{d}x, \\ \mathrm{d}\xi &= \omega \mathrm{d}\zeta + w^T \mathrm{d}x, \end{aligned}$$

which are obviously equivalent and have already in effect been used to define ω and w. Proceeding formally we have

$$\begin{aligned} \mathrm{d}^2\zeta &= \eta \mathrm{d}^2\xi + \mathrm{d}\eta \mathrm{d}\xi + \langle \mathrm{d}y, \mathrm{d}x\rangle \\ \mathrm{d}^2\xi &= \omega \mathrm{d}^2\zeta + \mathrm{d}\omega \mathrm{d}\zeta + \langle \mathrm{d}w, \mathrm{d}x\rangle. \end{aligned}$$

Multiplying the first equation by η^{-1} and adding we obtain

$$0 = \eta^{-1}\mathrm{d}\eta \mathrm{d}\xi + \langle \eta^{-1}\mathrm{d}y, \mathrm{d}x\rangle + \mathrm{d}\omega \mathrm{d}\zeta + \langle \mathrm{d}w, \mathrm{d}x\rangle.$$

Let J be the Jacobian matrix $\frac{\partial(\zeta,x)}{\partial(\xi,x)}$. Then this last equation tells us that the Jacobian matrix $\frac{\partial(\eta,y)}{\partial(\omega,w)}$ is $-\eta J^T$, a result that may also be obtained by direct computation.

As in section 7 let us write D for the total convected derivative and $(\quad)_{,i}$ for the derivative in the ith spatial direction.

Proposition 6 *Suppose that we have an equation*

$$D\begin{pmatrix}\xi \\ x\end{pmatrix} = A^i\begin{pmatrix}\eta \\ y\end{pmatrix}_{,i},$$

then in the transformed variable this becomes

$$D\begin{pmatrix}\zeta \\ x\end{pmatrix} = -\eta J A^i J^T \begin{pmatrix}\omega \\ w\end{pmatrix}_{,i},$$

so that if one system is symmetric so is the other; if one has a rank two matrix so does the other.

Proof The symmetry is obvious. If $A^i = a_i b_i^T + b_i a_i^T$ then $JA^iJ^T = (Ja_i)(Jb_i)^T + (Jb_i)(Ja_i)^T$. ■

In section 8 we mentioned models for diffusive effects.

Proposition 7 *Suppose we have an equation*

$$\begin{pmatrix}\xi \\ x\end{pmatrix}_t = \mathrm{div} L \mathrm{grad}\begin{pmatrix}\eta \\ y\end{pmatrix},$$

where L *is a symmetric matrix. Then modulo terms involving* $|\ \mathrm{grad}\ |^2$ *we have*

$$\begin{pmatrix}\zeta \\ x\end{pmatrix}_t = -\mathrm{div}(\eta J L J^T)\mathrm{grad}\begin{pmatrix}\omega \\ w\end{pmatrix}.$$

Proof What we actually obtain on transforming is

$$\begin{pmatrix} \zeta \\ x \end{pmatrix}_t = -\eta J \operatorname{div} L J^T \operatorname{grad} \begin{pmatrix} \omega \\ w \end{pmatrix}$$

and the extra term which we have neglected is

$$-\langle \operatorname{grad}(\eta J), L \operatorname{grad} \begin{pmatrix} \omega \\ w \end{pmatrix} \rangle.$$

■

Remark Notice that the sign of the right changes. This is because the original equation is entropy-increasing, while the second corresponds to a convex quality which will be decreasing.

Appendix B Homogeneous convex functions

The following example is given as an application of [7] Thm 16.4. Consider the function

$$\begin{aligned} f(\xi_1, \ldots, \xi_n) = \xi_1 \log \xi_1 + \cdots + \xi_n \log \xi_n \text{ if } \quad & \xi_j \geq 0, j = 1, \ldots, n \\ & \text{and } \xi_1 + \cdots + \xi_n = 1 \\ +\infty \text{ otherwise.} & \end{aligned}$$

We may write

$$f(x) = g(x) + \delta(x|D),$$

where

$$D = \{x = (\xi_1, ..., \xi_n) | \xi_1 + ... + \xi_n = 1|\},$$

and

$$g(x) = k(\xi_1) + ... + k(\xi_n),$$

with

$$\begin{aligned} k(\xi) = \xi \log \xi \qquad & \xi > 0, \\ 0 \qquad & \xi = 0, \\ +\infty \qquad & \xi < 0. \end{aligned}$$

Although int D is actually empty ri D, the interior relative to the hyperplane containing D, is the whole of D and the theorem still applies i.e.

$$f^*(x^*) = \inf_{y*} \{g^*(x^* - y^*) + \delta^*(y^*|D)\},$$

where the infimum is attained. Now

$$g^*(x^*) = k^*(\xi_1^*) + \cdots + k^*(\xi_n^*)$$

and by modifying the calculation of the dual of e^x we find

$$k^*(\xi^*) = e^{\xi^*-1}.$$

Also

$$\begin{array}{rl}\delta^*(x^*) = \lambda & \text{for } x^* = (\lambda, \dots, \lambda), \quad \forall \lambda \in R, \\ +\infty & \text{otherwise;}\end{array}$$

so that

$$\begin{aligned} f^*(x^*) &= \min_{\lambda}\{\lambda + \textstyle\sum_{j=1}^{n} e^{\xi^*-\lambda-1}\} \\ &= \log(e^{\xi_1^*} + \cdots + e^{\xi_n^*}) \end{aligned}$$

by setting to zero the derivative with respect to λ.

Appendix C Notation

e	energy per unit mass - sum of internal energy and kinetic energy
f	free energy per unit mass
g	Gibbs free energy per unit mass
h	enthalpy per unit mass
p	pressure
R	gas constant
s	entropy per unit mass
T	temperature measured from absolute zero
u	internal energy
v	velocity
Θ	T^{-1}
μ	chemical potential
ρ	density
τ	ρ^{-1}