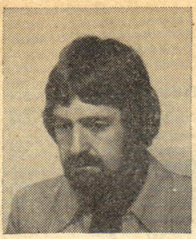

Thermodynamic Phase Changes and Catastrophe Theory

D. A. LAVIS and G. M. BELL, FIMA
Department of Mathematics, Chelsea College, London

1. Introduction

In 1873 van der Waals published his thesis[1] "On the continuity of the gaseous and the liquid state" in which he proposed his now famous equation of state as a basis for an understanding of the vapour–liquid phase transition. It was known that part of the loop of the continuous isotherm yielded by van der Waals' equation, for temperatures lower than the critical temperature, would correspond to metastable or unstable states of the system. A phase transition should be represented by drawing a tie-line parallel to the volume axis which would convert the isotherm into one of the general shape of those yielded by the experimental work of Andrews.[2] In 1875 Maxwell argued,[3] on the basis of the equivalence of heat and work, that the correct tie-line to choose would be that which formed with the isotherm two bounded areas of equal magnitude. It is now known that the equal areas rule is exactly equivalent to choosing the least minimum of an appropriate thermodynamic potential function. This latter principle we shall refer to as *Maxwell's rule.*

In the publications of Thom,[4–6] the principle whereby a first-order phase transition (a catastrophe) occurs when the potentials associated with two stable states (attractors) are equal is referred to as "Maxwell's convention." Although Thom makes it clear (reference 5, p. 56) that the physical validity of the convention is not his concern, he does regard it as "somewhat arbitrary" (reference 4, p. 105). He further remarks that the potential function "might well be replaced by any other function expressing roughly the stability of the attractor, . . . , and the structure of the catastrophe points would scarcely be altered" (reference 5, p. 56). It is beyond our competence to know whether, in the biological field, such modifications of the potential function have any significance, but it is certainly clear, from a study of the work of Fowler,[7] that an insufficiently careful choice of potential function can lead to serious errors in thermodynamics. To avoid confusion with the correctly formulated Maxwell's rule, we shall refer to the Maxwell convention of Thom by the name *Equal potentials convention.* We shall see in the next section that the distinction between the thermodynamically correct Maxwell's rule and the, not necessarily correct, Equal potentials convention rests on the choice of an appropriate and non-arbitrary potential function.

2. Thermodynamic potentials

For a system, considered in the context of catastrophe theory, the space $\mathbf{R}^r$ of all the variables $\{X_1, X_2, \ldots, X_r\}$, which can be independently manipulated by the experimenter, is called the *control space.*[8] The space $\mathbf{R}^n$ of the variables $\{Y_1, Y_2, \ldots, Y_n\}$, which are necessary, together with the control space variables, to specify a state of the system, is called the *state space.*[8] For *elementary catastrophes* there exists a smooth potential function $\varphi\colon \mathbf{R}^r \times \mathbf{R}^n \to \mathbf{R}$ such that the equilibrium states of the system are given by[9]

$$\frac{\partial \varphi}{\partial Y_j} = 0, \quad j = 1, 2, \ldots, n. \tag{2.1}$$

To a particular point in control space there may correspond a number of equilibrium states of the system. Which of these states is the stable equilibrium state may depend, as in the case of the Zeeman catastrophe machine,[8] on the means by which the state was attained. Part of the purpose of this paper is to discuss stability criteria for phase transitions in fluid systems.

We consider a fluid system of identical particles and define

thermal variables $\begin{cases} T\colon & \text{absolute temperature} \\ S\colon & \text{entropy} \end{cases}$

mechanical variables $\begin{cases} P\colon & \text{pressure} \\ V\colon & \text{volume} \end{cases}$

chemical variables $\begin{cases} \mu\colon & \text{chemical potential per particle} \\ N\colon & \text{number of molecules} \end{cases}$

U: thermodynamic energy.

These variables will be well defined for an equilibrium state of the system and, if the system undergoes an infinitesimal quasi-static change of equilibrium state

$$dU = T\,dS - P\,dV + \mu\,dN. \tag{2.2}^*$$

The control space for this system will be three dimensional and the variables of the space can be chosen to be either one of each of the bracketed pairs of variables given above. The remaining variables will then be given as functions with the

---

* The justification for this and subsequent results, quoted from thermodynamic theory, can be found in almost any text on the subject, e.g., Landsberg,[10] ter Haar and Wergeland.[11]

control space as domain. If we choose control space $\{S, V, N\}$ (a system in thermal, mechanical and chemical isolation) then we shall have, in this space, surfaces of constant energy of the form

$$U(S, V, N) = \text{constant}. \tag{2.3}$$

From (2.2) it is clear that the normal to such a surface at the point (S, V, N) will be in the direction $(T, -P, \mu)$. The thermodynamic energy is normally a homogeneous function of (S, V, N) and it follows from equation (2.2) and Euler's homogeneous function theorem that

$$U = TS - PV + \mu N. \tag{2.4}$$

Alternative choices of control space lead to surfaces specified by alternative functions.* We shall consider three.

(*i*) By defining the Helmholtz free energy

$$A = U - ST, \tag{2.5}$$

we can obtain from (2.2), by a Legendre transformation,

$$dA = -S\,dT - P\,dV + \mu\,dN. \tag{2.6}$$

The control space is now $\{T, V, N\}$ (a system in mechanical and chemical isolation and in thermal equilibrium with its environment). From (2.4)

$$A = -PV + \mu N. \tag{2.7}$$

The next two cases are symmetric in their use of the mechanical and chemical variables and we shall treat them in parallel.

(*ii*) By defining the Gibbs free energy

$$G = A + PV, \tag{2.8a}$$

we obtain from (2.6), by a Legendre transformation,

$$dG = -S\,dT + V\,dP + \mu\,dN. \tag{2.9a}$$

The control space is now $\{T, P, N\}$ (a system in thermal and mechanical equilibrium with its environment and in chemical isolation). From (2.7)

$$G = \mu N. \tag{2.10a}$$

(*iii*) By defining the function (unnamed in the literature)

$$D = A - \mu N, \tag{2.8b}$$

we obtain from (2.6), by a Legendre transformation,

$$dD = -S\,dT - P\,dV - N\,d\mu. \tag{2.9b}$$

The control space is now $\{T, V, \mu\}$ (a system in thermal and chemical equilibrium with its environment and in mechanical isolation). From (2.7)

$$D = -PV. \tag{2.10b}$$

Since classical thermodynamics is essentially concerned with equilibrium situations, a certain care is needed in defining these variables for non-equilibrium situations. Whenever this is possible the second law of thermodynamics allows the following analysis. Suppose that during some infinitesimal change of state of the system the entropy changes by dS, then

$$dS = d_e S + d_i S, \tag{2.11}$$

where $d_e S$ satisfies equation (2.2) and $d_i S$, the contribution to dS due to changes within the system satisfies the equation

$$d_i S \geqslant 0. \tag{2.12}$$

It follows from equations (2.2), (2.11) and (2.12) that

$$T\,dS - dU - P\,dV + \mu\,dN \geqslant 0, \tag{2.13}$$

where the equality sign corresponds to quasi-static equilibrium changes of state. By applying the Legendre transformations utilised above we conclude that, if the change of state occurs at a fixed point in the control space:

(*i*) $\{T, V, N\}$, then $dA \leqslant 0$,

(*ii*) $\{T, P, N\}$, then $dG \leqslant 0$,

(*iii*) $\{T, V, \mu\}$, then $dD \leqslant 0$.

We define non-equilibrium states of the system by introducing variables in addition to those of the control space. These new variables are the state space variables. For the control spaces in cases (*i*)–(*iii*) given above, the equilibrium states are then given by local minima of the potential functions A, G and D, respectively. The potential functions corresponding to alternative choices of control space can be derived by using further Legendre transformations in a way illustrated by the cases considered here. It is perhaps worth emphasising that, whatever permissible control space is chosen, there exists an appropriate potential function which can be derived using well established thermodynamic principles.

One way of obtaining well defined thermodynamic functions for non-equilibrium states is to include one or more of the fundamental set of six variables defined at the beginning of this section, beyond those necessary to specify a point in control space. We shall consider, in parallel, two cases corresponding to examples (*ii*) and (*iii*) given above.

Let the control space be $\{T, P, N\}$ and let the system be in an equilibrium state with respect to the variables T, V and N but not necessarily with respect to the applied

Let the control space be $\{T, V, \mu\}$ and let the system be in an equilibrium state with respect to the variables T, V and N but not necessarily with respect to the chemical

---

* On the basis of choice described above there are, of course, eight functions in all. Other functions can be generated by choosing U, A, G, D, etc., as control space variables.

pressure P. The Gibbs free energy is given from (2.8a) by

$$G(T,P,N;\,V) = A(T,V,N)+PV. \qquad (2.14a)$$

The state space is the one-dimensional space of V and the equilibrium states are given by

$$\frac{\partial G}{\partial V} = \left(\frac{\partial A}{\partial V}\right)_{T,N} + P = 0. \qquad (2.15a)$$

From (2.6), this equation defines precisely the equilibrium value of the pressure at a point in the control space $\{T, V, N\}$. A value of the volume is a stable equilibrium value if

$$\left(\frac{\partial^2 A}{\partial V^2}\right)_{T,N} > 0. \qquad (2.16a)$$

potential of the particles in the environment. The potential D is given from (2.8b) by

$$D(T,V,\mu;\,N) = A(T,V,N)-\mu N. \qquad (2.14b)$$

The state space is the one-dimensional space of N and the equilibrium states are given by

$$\frac{\partial D}{\partial N} = \left(\frac{\partial A}{\partial N}\right)_{T,V} - \mu = 0. \qquad (2.15b)$$

From (2.6), this equation defines precisely the equilibrium value of the chemical potential at a point in the control space $\{T, V, N\}$. A value for the number of particles is a stable equilibrium value if

$$\left(\frac{\partial^2 A}{\partial N^2}\right)_{T,V} > 0. \qquad (2.16b)$$

These are forms of the well known thermodynamic stability conditions.

It is clear from the above analysis that the thermodynamic equilibrium and stability conditions are strictly applicable only to states in a neighbourhood of an equilibrium state. They are local conditions. For a discussion of phase transitions they need to be extended to more general situations, which may be achieved by the following assertion.

The thermodynamic potential function, which is appropriate for the determination of local stability, is also that which determines the relative stability of all equilibrium states in state space which correspond to the same point in control space.

This assertion may be regarded as an additional axiom of thermodynamics and in that sense "conventional" but it is by no means arbitrary. It also has a long history of usage dating back to the work of Gibbs.[12]

Consider again our fluid system from the two parallel points of view given above and suppose that it can exist in vapour and liquid phases

with potential functions $G^{(1)}$ and $G^{(2)}$, respectively, for the control space $\{T, P, N\}$. If for a particular point in the control space there exist minima of the potential functions in each of the phases, with volumes V_1 and V_2 for the vapour and liquid, respectively, then according to our assertion the stable state is taken to be represented by the lower minimum. Thus we have the phase transition condition

$$G^{(1)}(T,P,N;\,V_1) = G^{(2)}(T,P,N;\,V_2), \qquad (2.17a)$$

which, from equation (2.10a), is equivalent to the condition that the particles in the two phases have equal chemical potentials.

For a classical equation of state of the form

$$P = f(T,V,N) \qquad (2.18a)$$

an isotherm (at constant N) in the P–V plane will exhibit the conditions necessary for a phase transition only if it has more than one solution for V at fixed (T, P, N). The simplest way this can occur is by means of a van der Waals type loop (see Fig. 1). In this case equation (2.18a) has, for fixed T (below the critical value) and over a certain range of P, three roots for V. We must distinguish between them according to their values of the potential function. This is given from equations (2.14a) and (2.18a) by

$$G(T,P,N;\,V) = PV - \int f(T,V,N)\,dV + \alpha(T,N), \qquad (2.19a)$$

where α is some unknown function of T and N. From equations (2.17a) and (2.19a) the Maxwell rule condition (2.17a) (with $G^{(1)} = G^{(2)}$) now takes the form

$$P(V_1-V_2) = \int_{V_2}^{V_1} f(T,V,N)\,dV. \qquad (2.20a)$$

with potential functions $D^{(1)}$ and $D^{(2)}$, respectively, for the control space $\{T, V, \mu\}$. If for a particular point in the control space there exist minima of the potential functions in each of the phases with numbers of particles N_1 and N_2 for the vapour and liquid, respectively, then according to our assertion the stable state is taken to be represented by the lower minimum. Thus we have the phase transition condition

$$D^{(1)}(T,V,\mu;\,N_1) = D^{(2)}(T,V,\mu;\,N_2), \qquad (2.17b)$$

which from equation (2.10b) is equivalent to the condition that the two phases are at equal pressures.

For a classical equation of the form

$$\mu = g(T,V,N) \qquad (2.18b)$$

an isotherm (at constant V) in the μ–N plane will exhibit the conditions necessary for a phase transition only if it has more than one solution for N at fixed (T, V, μ). The simplest way this can occur is by means of a van der Waals type loop. In this case equation (2.18b) has, for fixed T (below the critical value) and over a certain range of μ, three roots for N. We must distinguish between them according to their values of the potential function. This is given from equations (2.14b) and (2.18b) by

$$D(T,V,\mu;\,N) = -\mu N + \int g(T,V,N)\,dN + \beta(T,V), \qquad (2.19b)$$

where β is some unknown function of T and V. From equations (2.17b) and (2.19b) the Maxwell rule condition (2.17b) (with $D^{(1)} = D^{(2)}$) now takes the form

$$\mu(N_1-N_2) = \int_{N_2}^{N_1} g(T,V,N)\,dN. \qquad (2.20b)$$

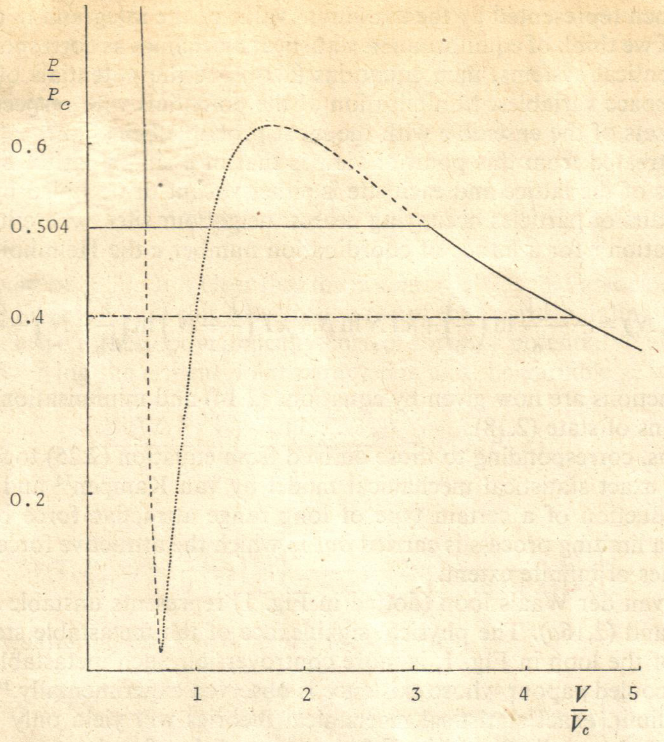

Fig. 1. The van der Waals isotherm for $(T/T_c) = 0{\cdot}85$. The broken part of the curve corresponds to metastable states and the dotted part to unstable states if the phase transition is calculated according to Maxwell's rule (or the equivalent equal areas rule). The pressure $(P/P_c) = 0{\cdot}4$ corresponds to the (incorrect) Riemann–Hugoniot catastrophe of Fowler[7]

Equation (2.20*a*) is an analytic expression of the usual equal areas rule. Equation (2.20*b*) is also an equal areas rule, but of a less familiar kind. It can be seen that neither of these is in any way an arbitrary convention, but both are consequences of well established thermodynamic practice.

As an example of this analysis we shall consider the van der Waals fluid. Here, as in most classical thermodynamic models, we are given only an equation of state of the form (2.18*a*):

$$P = \frac{NkT}{V-bN} - \frac{aN^2}{V^2}, \tag{2.21a}$$

where *a* and *b* are constants and *k* is Boltzmann's constant. It may be shown (reference 11, p. 35) that the heat capacity of this fluid, at constant volume, is everywhere outside the coexistence region equal to the perfect gas value $3Nk/2$ and in the low density limit the entropy gives asymptotically the Sackur–Tetrode equation. This enables us to evaluate the functions g, α and β to within an arbitrary constant which fixes the zero of the chemical potential. Neglecting this constant does not affect our calculations which give for (2.18b)

$$\mu = -kT\ln\left\{\left(\frac{2mkT\pi}{h^2}\right)^{\frac{3}{2}}\left(\frac{V}{N}-b\right)\right\} - \frac{2aN}{V} + kT\left(\frac{V}{bN}-1\right)^{-1} \tag{2.21b}$$

and potentials of the form

$$G(T,P,N;\,V) = PV - NkT\ln\left\{\left(\frac{2mkT\pi}{h^2}\right)^{\frac{3}{2}}\left(\frac{V}{N}-b\right)\right\} - \frac{aN^2}{V} - NkT \tag{2.22a}$$

and

$$D(T,V,\mu;\,N) = -\mu N - NkT\ln\left\{\left(\frac{2mkT\pi}{h^2}\right)^{\frac{3}{2}}\left(\frac{V}{N}-b\right)\right\} - \frac{aN^2}{V} - NkT, \tag{2.22b}$$

where *h* is Planck's constant and *m* is the mass per particle.

Let us now consider our two parallel cases from a statistical mechanical point of view where, at equilibrium, the potentials are given by[13]

$$G(T,P,N) = -kT\ln\Lambda(T,P,N), \tag{2.23a}$$

where Λ is the constant pressure partition function given in terms of the Helmholtz free energy by

$$\Lambda(T,P,N) = \int \exp -\frac{1}{kT}\{A(T,V,N)+PV\}\,dV. \tag{2.24a}$$

$$D(T,V,\mu) = -kT\ln\Xi(T,V,\mu), \tag{2.23b}$$

where Ξ is the grand canonical partition function given in terms of the Helmholtz free energy by

$$\Xi(T,V,\mu) = \sum_N \exp -\frac{1}{kT}\{A(T,V,N)-\mu N\}. \tag{2.24b}$$

It can be seen that the potentials of equations (2.14) are obtained from equations (2.24) by neglecting the integral or sum.

The equilibrium situation is then represented by the maximum value of the integrand in (2.24*a*) and the maximum term of the summation in (2.24*b*). If we think of equilibrium in statistical mechanics as corresponding to the average behaviour of an infinite ensemble of identical systems, then equations (2.14) are the potentials of subsets of the ensemble with particular values of the state space variables. Minimisation of the potentials with respect to the state space variables is then seen as selecting the subsets of the ensemble with the most probable behaviour.

An example which can be treated from this point of view is that of a simple lattice fluid. Here the volume $V = nv_0$, where n is the number of sites of the lattice and each site is either vacant or occupied by a particle. The particle interaction energy is confined to pairs of particles occupying nearest neighbour sites, where it has the value $-w$ $(w>0)$, and in the zeroth order approximation[13] for a lattice of coordination number c the Helmholtz free energy is given by

$$A(T, V, N) = -\frac{kTV}{v_0}\ln\left(\frac{V}{v_0}\right) + kTN\ln N + kT\left(\frac{V}{v_0} - N\right)\ln\left(\frac{V}{v_0} - N\right) - \frac{cwv_0N^2}{2V}. \tag{2.25}$$

The appropriate potential functions are now given by equations (2.14) and minimisation with respect to the state space variables leads to the equations of state (2.18).

Equilibrium $P-V$ isotherms, corresponding to those derived from equation (2.25) together with the equal areas rule, have been obtained from an exact statistical mechanical model by van Kampen[14] and Lebowitz and Penrose.[15] This derivation involves the introduction of a certain type of long range attractive force between pairs of particles. The isotherms are obtained when a limiting process is carried out in which the attractive force becomes infinitely weak at the same time as its range becomes of infinite extent.

The middle branch of the van der Waals loop (dotted in Fig. 1) represents unstable states of the system, as can be seen from equations (2.15*a*) and (2.16*a*). The physical significance of the metastable states, represented by the broken parts of the outer branches of the loop in Fig. 1, is more controversial. Such metastable states can be regarded as the super-heated liquid or super-cooled vapour whose existence is observed experimentally.[16] However, it is often supposed that, in the thermodynamic limit, exact statistical mechanical theories will yield only the stable states of the system (see, for example, reference 13, Appendix 9). Fisher[17] has suggested that for short range forces the transition point at the vapour–liquid boundary should correspond to an essential singularity of the Gibbs free energy. From this general point of view the continuity of the isotherms into metastable regions found in mean field theories, like that of van der Waals or the zeroth-order approximation, or in higher closed form approximations applied to lattice gases (see, for example, Burley[18] or Lavis and Bell[19]), is an artifact due to inexact treatment. However other authors take a different viewpoint. Gaunt and Baker[20] supposed that the metastable states can be obtained by analytic continuation and used series expansion techniques to locate the spinodal curve separating the metastable region from the unstable region. A resolution has been attempted by Domb[21] in terms of the critical rôle played by certain droplet configurations. Numerically these configurations make only a negligible contribution to the partition function and would therefore be undetected by the methods of Gaunt and Baker. Another approach, due to Lebowitz,[22] shows that metastable states can be obtained, in an exact theory, provided that a constraint is applied to prevent large density fluctuations. From this point of view metastable states presumably occur in approximate treatments because these implicitly impose homogeneity in each phase by using the density as an order variable.

3. Phase changes and catastrophes

In section 2 we considered a fluid system from two points of view, with control space $\{T, P, N\}$ and state space $\{V\}$ and with control space $\{T, V, \mu\}$ and state space $\{N\}$. In the first of these cases we can reduce the control space to two dimensions by fixing the number of particles at some convenient value (e.g., Avogadro's number, when we have one mole of fluid). In the second case we can similarly reduce the control space by fixing the volume. We are now, in each case, considering a system with control space $\{X_1, X_2\}$ and state space $\{Y\}$. The set $\mathscr{M}$ of attractors and repellers is the set of equilibrium states at which the derivative of φ with respect to Y is zero.

The main theorem of catastrophe theory tells us[9] that if we project $\mathscr{M}$ into the two dimensional space $\{X_1, X_2\}$ then the only singularities associated with this mapping are folds and cusps. Suppose that there is a cusp at the point (X_{1c}, X_{2c}) with the associated point of $\mathscr{M}$ given by Y_c. In the thermodynamic sense this is the *critical point* of a phase transition and the assertion of catastrophe theory is that in a neighbourhood of the critical point the phase transition can be represented by a *Riemann–Hugoniot catastrophe*. This means that there exists a function x of Y, such that $x(Y_c) = 0$, with respect to which the potential has, in the neighbourhood of the critical point, the asymptotic expansion

$$\varphi(X_1, X_2;\, x) \sim \zeta(X_1, X_2)\,x + \xi(X_1, X_2)x^2 + \sigma(X_1, X_2)\,x^3 + \chi(X_1, X_2)\,x^4, \tag{3.1}$$

where $\chi(X_{1c}, X_{2c}) > 0$, $\zeta(X_{1c}, X_{2c}) = \xi(X_{1c}, X_{2c}) = 0$ and σ is identically zero. There is, of course, a large number of functions which, given the set $\mathscr{M}$, can be constructed so that this set represents the vanishing of their first derivatives with respect to Y. It must however be emphasised that, if the correct phase transition curve is to be obtained using Maxwell's rule, then the potential must be constructed according to the principles of thermodynamics as outlined in section 2. This basic point seems to have been overlooked by Fowler[7] in his treatment of the van der Waals and Berthelot equations. In order to clarify this point we shall outline Fowler's treatment of van der Waals' equation. For this equation it is not difficult to show (see, for example, reference 11, p. 4) that the critical point is given by $T_c = 8a/(27bk)$, $P_c = a/(27b^2)$ and $V_c = 3Nb$. Fowler chooses the variable

$$x = (V_c/V - 1) \tag{3.2}$$

and expresses equation (2.21a) in the form

$$0=\frac{2}{3}\left(\frac{4T}{T_c}-\frac{P}{P_c}-3\right)+\frac{1}{3}\left(\frac{8T}{T_c}+\frac{P}{P_c}-9\right)x+x^3. \tag{3.3}$$

On the basis of this he produces the potential $\varphi^{(\mathrm{F})}$ given by

$$\varphi^{(\mathrm{F})}(T,P;\,x)=\frac{2}{3}\left(\frac{4T}{T_c}-\frac{P}{P_c}-3\right)x+\frac{1}{6}\left(\frac{8T}{T_c}+\frac{P}{P_c}-9\right)x^2+\frac{1}{4}x^4, \tag{3.4}$$

for which $\partial\varphi^{(\mathrm{F})}/\partial V=0$ gives equation (3.3). It is clear that the potential $\varphi^{(\mathrm{F})}$ is that associated with a Riemann–Hugoniot catastrophe. It is equally clear, however, that this catastrophe is not that corresponding to the phase transition in the van der Waals fluid. We now derive a thermodynamically correct form of potential Φ as a function of x. Translating equation (2.22a) into a form involving the critical point parameters and the variable x, we obtain

$$\frac{2\{G(T,P;\,V)-G(T,P;\,V_c)\}}{3P_cV_c}=\Phi(T,P;\,x), \tag{3.5}$$

where

$$\Phi(T,P;\,x)=\frac{16}{9}\left(\frac{T}{T_c}\right)\ln\left(\frac{2+2x}{2-x}\right)-2x\left\{1+\frac{1}{3(1+x)}\left(\frac{P}{P_c}\right)\right\}. \tag{3.6}$$

The Taylor expansion of this equation about $x=0$, to the fourth order, is

$$\Phi(T,P;\,x)\sim\frac{2}{3}\left\{\frac{4T}{T_c}-\frac{P}{P_c}-3\right\}x+\frac{2}{3}\left\{\frac{P}{P_c}-\frac{T}{T_c}\right\}x^2+\frac{2}{3}\left\{\frac{T}{T_c}-\frac{P}{P_c}\right\}x^3+\frac{1}{12}\left\{\frac{8P}{P_c}-\frac{5T}{T_c}\right\}x^4. \tag{3.7}$$

The identity between the linear terms in (3.4) and (3.7) is simply a matter of scaling, and the fact that equation (3.7), like (3.1) and (3.4), has coefficients which vanish up to the cubic term at the critical point is determined by the order of the phase transition (see, for example, Landau and Lifshitz[23], chapter 14). In equations (3.1) and (3.4) the cubic term is identically zero. That this is not the case for equation (3.7) illustrates an important point concerning the relationship between phase transitions and catastrophe theory. It is of course quite easy to eliminate this term in (3.7) by replacing the variable x by the variable x' defined by

$$x'=x+2\left\{\frac{T}{T_c}-\frac{P}{P_c}\right\}\left\{\frac{8P}{P_c}-\frac{5T}{T_c}\right\}^{-1}, \tag{3.8}$$

as is done by Schulman and Revzen.[24] It is tempting to suppose that we have, by this means, obtained a potential in the required form (3.1). This is however incorrect, since the Taylor expansion for the Gibbs free energy about $x'=0$ must be obtained from the exact expression (3.6) for Φ and not from (3.7). If we expand correctly with respect to x' we shall find a non-vanishing cubic term.

The absence of a cubic term in the expansion of the thermodynamic potential, with respect to some ordering variable, is associated with definite symmetry properties of the model, which have been reflected in the choice of expansion variable (see reference 23, chapter 14). It is presumably a consequence of catastrophe theory that, given a particular system and a particular two dimensional control space, then a parameter exists in terms of which the expansion of the correct thermodynamic potential has the Riemann–Hugoniot form. It is, however, by no means easy to determine this parameter in particular cases. It is certainly true that neither x nor x' is that parameter for a van der Waals fluid with control space $\{T,P\}$. On the other hand, for this fluid, with control space $\{T,\mu\}$ and state space $\{N\}$, we have been able to determine a parameter in terms of which the correct thermodynamic potential, expanded to the fourth order, has the Riemann–Hugoniot form. The rôles of N and V are now reversed and the number of particles has critical value $N_c=V/(3b)$. The critical value of the chemical potential is given by

$$\frac{\mu_c}{kT_c}=-\ln\left\{2b\left(\frac{2mkT_c\pi}{h^2}\right)^{\frac{3}{2}}\right\}-\frac{7}{4}. \tag{3.9}$$

Choosing the variable

$$y=(N/N_c-1) \tag{3.10}$$

and translating equation (2.22b) into a form involving the critical point parameters and the variable y we obtain

$$\frac{D(T,\mu;\,N)-D(T,\mu;\,N_c)}{kT_cN_c}=\Psi(T,\mu;\,y), \tag{3.11}$$

where

$$-\frac{3}{2}\ln\left(\frac{T}{T_c}\right)$$

$$\Psi(T,\mu;\,y)=\left(\frac{T}{T_c}\right)(1+y)\ln\left(\frac{2+2y}{2-y}\right)+y\left(\frac{T}{T_c}\right)\left\{\left(\frac{\mu_c}{kT_c}-\frac{\mu}{kT}\right)+\frac{3}{4}\right\}-\frac{9}{8}y(2+y). \tag{3.12}$$

$-\frac{3}{2}\ln\left(\frac{T}{T_c}\right)$

The Taylor expansion of this equation about $y = 0$, to the fourth order, is

$$\Psi(T,\mu;y)\sim\left\{\left(\frac{T}{T_c}\right)\left(\frac{\mu_c}{kT_c}-\frac{\mu}{kT}\right)+\frac{9}{4}\left(\frac{T}{T_c}-1\right)\right\}y+\frac{9}{8}\left(\frac{T}{T_c}-1\right)y^2+\frac{9}{64}\left(\frac{T}{T_c}\right)y^4. \tag{3.13}$$

It is also of interest to note the equations corresponding to (3.6) and (3.12) for the lattice gas with Helmholtz free energy (2.25). These are

$$\Phi_{\text{L.G.}}(T,P;x)=\frac{2}{3}\left(\frac{T}{T_c}\right)\frac{\{(1-x)\ln(1-x)+(1+x)\ln(1+x)\}}{(1+x)(2\ln 2-1)}+\frac{2x}{3(1+x)}\left\{\left(\frac{T}{T_c}\right)-\left(\frac{P}{P_c}\right)+\left(\frac{T}{T_c}-1-x\right)(2\ln 2-1)^{-1}\right\} \tag{3.14}$$

and

$$\Psi_{\text{L.G.}}(T,\mu;y)=\left(\frac{T}{T_c}\right)\{(1+y)\ln(1+y)+(1-y)\ln(1-y)\}-y\left(2+y-\frac{2\mu}{\mu_c}\right), \tag{3.15}$$

respectively, with Taylor expansions

$$\Phi_{\text{L.G.}}(T,P;x)\sim\frac{2}{3}\left\{\frac{T}{T_c}-\frac{P}{P_c}+\left(\frac{T}{T_c}-1\right)(2\ln 2-1)^{-1}\right\}x-\frac{2}{3}\left(\frac{T}{T_c}-\frac{P}{P_c}\right)x^2$$
$$+\frac{2}{3}\left(\frac{T}{T_c}-\frac{P}{P_c}\right)x^3+\left\{\frac{1}{9}\left(\frac{T}{T_c}\right)(2\ln 2-1)^{-1}-\frac{2}{3}\left(\frac{T}{T_c}-\frac{P}{P_c}\right)\right\}x^4 \tag{3.16}$$

and

$$\Psi_{\text{L.G.}}(T,\mu;y)\sim 2\left(\frac{\mu}{\mu_c}-1\right)y+\left(\frac{T}{T_c}-1\right)y^2+\frac{1}{6}\left(\frac{T}{T_c}\right)y^4. \tag{3.17}$$

Again we have a situation in which the control space $\{T,P\}$ with parameter x does not produce a Riemann–Hugoniot catastrophe, whereas the control space $\{T,\mu\}$ with parameter y does. This can be clearly understood in terms of the basic symmetry of the Ising model from which the simple lattice gas is derived. If we translate the Ising model into a lattice gas (reference 13, chapter 7) the constant field distribution, which carries the symmetry properties, must be translated into the grand canonical distribution. If we construct the constant pressure distribution the symmetry is destroyed.

Let us now return to the problem of determining the first order phase transition away from the critical point. Such a transition will occur on a line of points in the control space terminating at the critical point and lying in a region for which the equation of state has three roots. We have shown in section 2 that this phase transition curve can be obtained, for a fluid system, by applying Maxwell's rule to the correctly established thermodynamic potential. Since catastrophe theory is a local theory, valid only in some neighbourhood of the critical point, we can expect, in general, to obtain from it only the asymptotic properties of the phase transition curve. There is an exception to this when the potential is given by

$$\varphi(X_1,X_2;x)=\zeta(X_1,X_2)\,x+\varphi_1(X_1,X_2;x), \tag{3.18}$$

where φ_1 is an even function of x with second derivative equal to zero and fourth derivative strictly positive at the critical point. In this case the phase transition curve in the X_2–X_1 plane is given exactly by

$$\zeta(X_1,X_2)=0. \tag{3.19}$$

One example of this special case is that of the lattice gas with potential $\Psi_{\text{L.G.}}$ given by equation (3.15). Here the phase transition curve in the $\mu-T$ plane is

$$\mu=\mu_c. \tag{3.20}$$

Another example would be that for which the potential were given exactly by the Riemann–Hugoniot potential. Fowler's potential (3.4) for the van der Waals fluid is of this form and he derives the phase transition curve in the $P-T$ plane

$$\frac{P}{P_c}=\frac{4T}{T_c}-3. \tag{3.21}$$

Unfortunately, since he has used the wrong potential, it is clear that this equation does not correctly represent the phase transition curve. Even as a local asymptotic form of this curve it is not very satisfactory. Suppose we have, for our potential function, the asymptotic form (3.1) with σ zero at the critical point, but not necessarily identically zero. Then a short calculation, using Maxwell's rule, yields the equation

$$\sigma(X_1,X_2)\{4\xi(X_1,X_2)\,\chi(X_1,X_2)+\sigma^2(X_1,X_2)\}\sim 8\,\zeta(X_1,X_2)\,\chi^2(X_1,X_2) \tag{3.22}$$

for the asymptotic form for the phase transition curve. When σ is identically zero, as in equation (3.13), we have the asymptotic equivalent of equation (3.19). When σ is not identically zero we must use the complete form of (3.22). In the case of equation (3.7) this gives

$$0\sim\left\{\frac{4T}{T_c}-\frac{P}{P_c}-3\right\}\left\{\frac{8P}{P_c}-\frac{5T}{T_c}\right\}-4\left\{\frac{T}{T_c}-\frac{P}{P_c}\right\}^2\left\{\frac{10P}{P_c}-\frac{7T}{T_c}\right\}. \tag{3.23}$$

The curves in the P–T plane representing equations (3.21) and (3.23) agree in gradient but not in curvature at the critical point. In Figs. 2 and 3 we have represented examples of the behaviours of $\varphi^{(F)}$ and Φ, given respectively by equations (3.4) and (3.6), as functions of (V/V_c). From these the phase transition tie-lines have been obtained as represented on Fig. 1. That determined from Fowler's potential $\varphi^{(F)}$ is at a considerably lower pressure than the correct value obtained from Φ. Since $\varphi^{(F)}$ and Φ are smooth functions on the real line, without self-intersections, a diffeomorphism must exist between them which maps one on to the other. This, nevertheless, does not seem to be a sufficient condition to guarantee that $\varphi^{(F)}$ is an adequate function for representing the van der Waals phase transition.

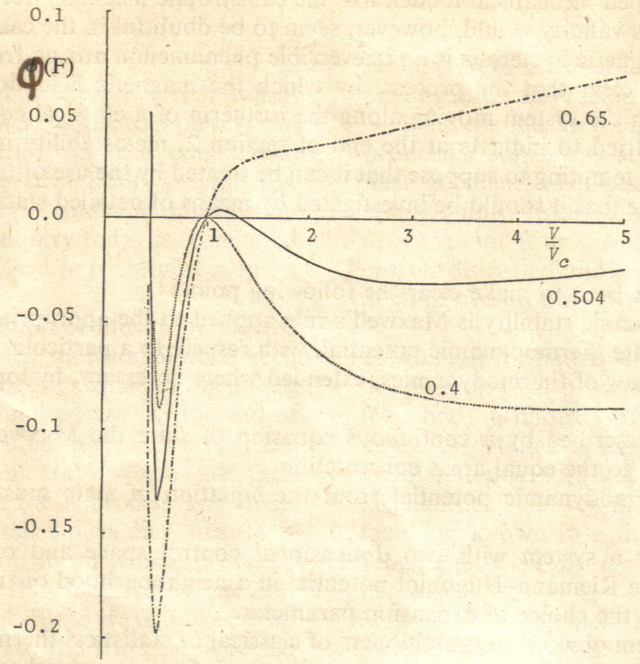

Fig. 2. Fowler's Riemann–Hugoniot potential corresponding to the van der Waals equation for $(T/T_c) = 0{\cdot}85$. The curves are parametrised by their values of (P/P_c) with $(P/P_c) = 0{\cdot}4$ corresponding to the catastrophe

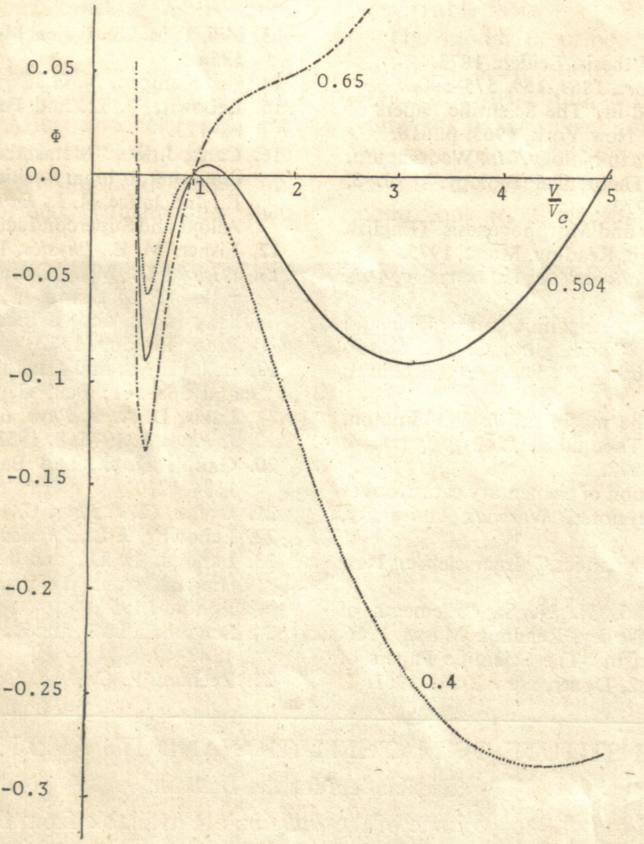

Fig. 3. The function Φ (which is to within a function of P and T the Gibbs free energy) for the van der Waals equation for $(T/T_c) = 0{\cdot}85$. The curves are parametrised by their values of (P/P_c) with $(P/P_c) = 0{\cdot}504$ corresponding to the phase transition calculated according to Maxwell's rule

There is another aspect of Fowler's paper[7] which gives rise to objections. Relying on the supposed conventionality of Maxwell's rule he proceeds to a discussion of alternative conventions. He gives two examples. The first he refers to as the *saturation convention* whereby "one minimum is preferred as soon as it comes into existence." For this he gives no justification and it need not detain us. The second, the so-called *delay convention* whereby "the system remains in a minimum until it disappears" has been discussed by Thom in relation to models in embryology (reference 5, p. 177) and by Zeeman in relation to the behaviour of his catastrophe machine[8] and to models for learning.[25] We are not in a position to embark upon a detailed discussion of the validity of this convention for biological and sociological models, and it is certainly clear that for a damped mechanical model, like the catastrophe machine, not subject to energy fluctuations, the delay convention is valid. Its validity would, however, seem to be doubtful in the case discussed by Fowler, namely that of magnetic hysteresis. Magnetic hysteresis is an irreversible phenomenon arising from frictional effects at domain boundaries. It is by no means clear that the process, by which the magnetic field does work in overcoming these frictional effects, corresponds to the system moving along the isotherm of a classical equation of state until it reaches the spinodal line. As we have tried to indicate at the end of section 2, metastability in thermodynamic systems is a complicated problem. It may be tempting to suppose that it can be treated by the use of the delay convention but we feel that it is much more appropriate that it should be investigated by means of detailed statistical mechanical models.

4. Conclusions

The purpose of this work has been to make clear the following points.

(*i*) The criterion for thermodynamic stability is Maxwell's rule applied to the appropriate thermodynamic potential.

(*ii*) The choice of the appropriate thermodynamic potential, with respect to a particular control space, is not arbitrary. It is determined by the second law of thermodynamics, extended where necessary, by long accepted practice, to include non-neighbouring states.

(*iii*) Whenever the system is described by a continuous equation of state the Maxwell rule, for a first-order phase transition, is exactly equivalent to the equal areas construction.

(*iv*) Determination of the thermodynamic potential from the equation of state must be done with due regard to thermodynamic laws.

(*v*) Whether the potential, for a system with two dimensional control space and one dimensional state space, is asymptotically equivalent to the Riemann–Hugoniot potential in a neighbourhood of the critical point depends on the choice of control space and on the choice of expansion parameter.

These points will perhaps seem obvious to practitioners of classical or statistical thermodynamics. They do, however, seem to be somewhat neglected by those engaged in the application of catastrophe theory.

We have been aided in this work by discussion with Dr. F. Jellett.

References

1. van der Waals, J. D., Doctoral thesis, Leiden, 1873.
2. Andrews, T., *Phil. Trans. R. Soc.*, 1869, **159**, 575–590.
3. Maxwell, J. C., 1875; reprinted in "The Scientific Papers of J. C. Maxwell," vol. 2, Dover, New York, 1965, p. 418.
4. Thom, R., "Topological models in biology," *in*, Waddington, C. H., *Editor*, "Towards a Theoretical Biology," vol. 3, Edinburgh University Press, 1970.
5. Thom, R., "Structural Stability and Morphogenesis" (English trans., Fowler, D. H.), Benjamin, Reading, Mass., 1975.
6. Thom, R., *Modèles mathématiques de la morphogenèse*, Coll. 10–18, Paris, 1974.
7. Fowler, D. H., "The Riemann–Hugoniot catastrophe and van der Waals equation," *in* Waddington, C. H., *Editor*, "Towards a Theoretical Biology," vol. 4, Edinburgh University Press, 1972.
8. Zeeman, E. C., "A catastrophe machine," *in* Waddington, C. H., *Editor*, "Towards a Theoretical Biology," vol. 4, Edinburgh University Press, 1972.
9. Zeeman, E. C., "The classification of elementary catastrophes of codimension$\leqslant 5$," Lecture notes, Warwick University, 1974.
10. Landsberg, P. T., "Thermodynamics," Interscience, New York, 1961.
11. ter Haar, D., and Wergeland, H. N. S., "Elements of Thermodynamics," Addison-Wesley, Reading, Mass., 1966.
12. Gibbs, J. W., 1873; reprinted in "The Scientific Papers of J. W. Gibbs," vol. 2, chapter 2, Dover, New York, 1961.
13. Hill, T. L., "Statistical Mechanics," McGraw-Hill, New York, 1956.
14. van Kampen, N. G., *Phys. Rev.*, 1964, **135A**, 362–369.
15. Lebowitz, J. L., and Penrose, O., *J. Math. Phys.*, 1966, **7**, 98–113.
16. Cahn, J. W., "Metastability, instability and the dynamics of unmixing in binary critical systems," *in* Mills, R. E., Ascher, E., and Jaffee, R. I., *Editors*, "Critical Phenomena in Metals, Alloys and Superconductors," McGraw-Hill, New York, 1971.
17. Fisher, M. E., *Physics*, 1967, **3**, 255–283.
18. Burley, D. M., "Closed form approximations for lattice systems," *in* Domb, C., and Green, M. S., *Editors*, "Phase Transitions and Critical Phenomena," vol. 2, Academic Press, London, 1972.
19. Bell, G. M., and Lavis, D. A., *J. Phys. A.*, 1970, **3**, 427–441 and 568–581; Bell, G. M., *J. Phys. C.*, 1972, **5**, 889–905; Lavis, D. A., *J. Phys. C.*, 1973, **6**, 1530–1545; Lavis, D. A., *J. Phys. A*, 1975, **8**, 1933–1951.
20. Gaunt, D. S., and Baker, G. A., *Phys. Rev.*, 1970, **1B**, 1184–1210.
21. Domb, C., *J. Phys. C*, 1973, **6**, 39–42.
22. Lebowitz, J. L., *Physica*, 1974, **73**, 48–60.
23. Landau, L. D., and Lifshitz, E. M., "Statistical Physics" (English trans. Peierls, E. and R. F.), Pergamon Press, London, 1959.
24. Schulman, L. S., and Revzen, M., *Collective Phenomena*, 1972, **1**, 43–49.
25. Zeeman, E. C., *Bull. IMA*, 1974, **10**, 272–281.