EXACT DERIVATION OF THE VAN DER WAALS EQUATION*+

by
J.L. Lebowitz
Belfer Graduate School of Science
Yeshiva University
New York, New York 10033

It seems appropriate to speak here about a rigorous first principles derivation of the van der Waals equation despite the fact that my talk is based on work, done jointly by Oliver Penrose and myself, which has already been published [1,2].

As was made clear this morning, the original van der Waals equation is not exact for any real system, so I shall certainly not give a rigorous derivation of it. Rather, I shall show that in a certain well defined limit, corresponding to the attractive part of the potential between the molecules becoming infinitely long range and infinitely weak, a generalized form of the van der Waals-Maxwell equation of state is exact and indeed can be proven to be so.

I shall devote most of my talk not to the most general situation where this result is valid but to a more restricted class of potentials for which the proofs are sufficiently simple to be presented in some detail in a lecture like this. A general report on the current status of work in this field is now in preparation [3].

It was a great achievement of van der Waals to realize that the universal existence of a gas-liquid phase transition at low temperatures, terminating at some critical temperature T_c, can be understood qualitatively as arising in a simple way from some general features of the interaction between molecules. To this end van der Waals, in 1873, visualized the interaction potential between a pair of molecules a distance r apart, $u(r)$, as consisting of two separate additive parts, a short-range repulsive part $q(r)$ and a 'long-range' attractive part $w(r)$,

$$u(r) = q(r) + w(r). \tag{1}$$

The short-range part keeps the particles apart and is responsible for detailed correlations. The long-range part on the other hand sees only the gross macroscopic density profile of the fluid and is responsible for the condensation from the gas into the liquid below the critical temperatute T_c.

van der Waals' considerations led him to his famous equation of state. For $T > T_c$ the van der Waals equation of state gives a good qualitative representation of the isotherms of a real fluid: for $T < T_c$, however, each isotherm includes a section where the compressibility is negative, in violation of the thermodynamic stability principle. The primary reason for this failure is that the argument about the effect of the long-range potential $w(r)$ being independent of the correlations is only valid when the range of the correlations is small compared to the range of $w(r)$. This necessarily requires the system to be in a single phase; it does not allow for the possibility of coexisting liquid and vapour phases at different densities when there are correlations of an infinite range present in the system.

Maxwell showed that the coexistence region could be included in the theory by using the van der Waals equation of state for both liquid and vapour phases and using the thermodynamic equilibrium conditions that the two phases must have equal pressures and chemical potentials.

A very interesting derivation of van der Waals' equation of state with Maxwell's rule was given by van Kampen in 1964 [4]. In this derivation the volume occupied by the system is divided into a large number of cells, each small compared with the range of the long-range attractive

force, but large enough to contain many particles. Avoiding the pitfall of assuming a uniform distribution of particles over cells van Kampen obtained the distribution over cells by minimizing the free energy. His method leads to the Maxwell modified van der Waals equation of state, which implies a first-order phase transition.

Van Kampen's treatment, while containing the physics of the problem, was however not mathematically rigorous. In particular the conditions to be satisfied by the interactions were not specified and various limiting processes were only hinted at but not carried out explicitly.

An entirely different appraoch to the van der Waals's equation of state was taken by Kac, Uhlenbeck and Hemmer [5]. Their work concerned a one-dimensional system for which the short-range repulsive potential $q(r)$ is infinite for $r < r_o$ and vanishes for $r > r_o$ (hard rods of diameter r_o) and an attractive interaction $w(r)$ which contained an inverse range parameter γ

$$w(r,\gamma) = -a\gamma e^{-\gamma r} \qquad (2)$$

Using a formalism based on Wiener processes especially adapted to this problem which was first considered by Kac in 1959, Kac, Uhlenbeck and Hemmer were able to prove rigorously the validity of the van der Waals equation of state, together with the Maxwell rule, in the van der Waals limit $\gamma \to 0$.

The limit process $\gamma \to 0$, provides a clear distinction between the short range of $q(r)$ and the long (infinite as $\gamma \to 0$) range of $w(r,\gamma)$. This limit was first used by Baker in 1961 [6] in his study of spin systems.

Lebowitz and Penrose [1] combined the ideas of van Kampen with the use of the 'van der Waals limit' $\gamma \to 0$. They considered systems with interparticle potentials of the form (1) with $w(r)$ a <u>Kac potential</u> of the form

$$w(r,\gamma) = \gamma^\nu \, \varphi(\gamma r) \, , \quad \gamma > 0 \tag{3}$$

where ν is the dimensionality of the space considered. This reduces to (2) when $\nu = 1$ and $\varphi(x) = -ae^{-x}$. By imposing certain conditions on $q(r)$ (which have since been relaxed), and $\varphi(r)$, Lebowitz and Penrose showed that in the limit $\gamma \to 0$, taken in such a way that the range γ^{-1} remains small compared to the size of the system,

$$\lim_{\gamma \to 0} p(\rho,T,\gamma) \equiv p(\rho,T,0+) = p^o(\rho,T) - a\,\rho^2 + \text{Maxwell's rule}$$

$$\equiv MC \left\{ p^o(\rho,T) - a\rho^2 \right\} \, . \tag{4}$$

Here $p^o(\rho,T)$ is the pressure of the reference system, one for which $w(r,\gamma) = 0$, which may itself be having a transition of its own, and

$$a = -\tfrac{1}{2} \int w(r,\gamma)d\underline{r} = -\tfrac{1}{2} \int \varphi(x)d\underline{x}. \tag{5}$$

The right side of (5) is a Riemann integral over all of ν-dimensional space, whose existence is one of the conditions $\varphi(x)$ has to satisfy. The extension of (4) to quantum systems weas done by Lieb [7].

The results of Lebowitz and Penrose were generalized further by Gates and Penrose [8] , who showed in particular that there are some Kac potentials $\varphi(x)$ for which (4) does not hold. These potentials are of an oscillatory type and apparently produce in the system, for

some values of ρ and T, an oscillatory local density rather than a separation into only two phases. It is the latter situation which leads to (4).

Now to the proof. We are dealing here with a system whose total potential energy is a sum of pair interactions,

$$U(\underline{r}_1, \ldots, \underline{r}_N) = \Sigma \, q(r_{ij}) + \Sigma \, \gamma^\nu \, \varphi(\gamma r) = Q + W, \tag{6}$$

and we want to find the pressure of this system when $\gamma \to 0$. Now the first thing to notice is that if we consider a finite system of N particles in a region Ω in ν dimensional space then the free energy of the system, which we shall assume classical for the moment, is given by

$$F(N, \Omega, \gamma) = -kT \ln Z(N, \Omega, \gamma) \tag{7}$$

$$Z(N, \Omega, \gamma) = \lambda^{-\nu N} \frac{1}{N!} \int \ldots \int d\underline{r}_1 - d\underline{r}_N \, e^{-\beta U}, \quad \beta = (kT)^{-1} \tag{8}$$

where λ is the de Broigle wave length and I have omitted writing out the explicit dependence of F and Z on T.

Clearly if we let $\gamma \to 0$ with φ bounded we have

$$F(N, \Omega, 0) = F^o(N, \Omega)$$

where $F^o(N, \Omega)$ is just the free energy of the reference system, that is the system with interparticle potential $q(r)$. The same will be true also for the pressure to the extent to which the pressure can be defined precisely by equilibrium considerations alone for a finite system.

My qualitication about the definition of the pressure comes from the fact that there is a problem in the usual way of defining the pressure for a finite system as the negative of the derivative of the free energy with respect to the volume, i.e.

$$p(N, \Omega) = -\frac{\partial F(N, \Omega)}{\partial \Omega} \quad . \tag{9}$$

Since one does not specify the way the volume is changed in (9) the derivative will generally be different if Ω is a cube or a sphere which is being changed uniformly and who knows what it would be or what it would even mean if Ω is some complicated shape. Ah, you say, but if the volume is sufficiently large the shape will not matter $F(N, \Omega) \sim |\Omega| \, f(N/|\Omega|)$ where $|\Omega|$ is the volume of Ω. This is indeed true for a large class of potentials, including all those one is commonly interested in, as has been proven by people like van Hove, van Kampen, Ruelle, Fisher, Griffiths, Lieb, etc.[9]. To be precise what is proven is that if one takes a sequence of domains Ω_j of 'reasonable' shapes with volumes $|\Omega_j|$ and particle number N_j such that

$$N_j \to \infty \, , \, \Omega_j \to \infty \quad \frac{N_j}{|\Omega_j|} = \rho_j \to \rho$$

(this is called going to the thermodynamic limit) then

$$\frac{F(N_j, \Omega_j)}{|\Omega_j|} \equiv f(\rho_j \, ; \Omega) \to f(\rho) \, , \, \text{as } j \to \infty \, , \tag{10}$$

independent of the shapes of Ω_j (as long as the shapes are reasonable). As Bob Griffiths says there is nothing fancy about the thermodynamic limit. It is what an experimentalist does every time he reports a

specific heat measurement per gram of sample without specifying the shape of the sample.

Furthermore, as a by product of these proofs of the existence of the thermodynamic limit, one obtains that the thermodynamic free energy density $f(\rho)$ is convex, that is

$$f(\alpha \rho' + (1-\alpha) \rho'') \leq \alpha f(\rho') + (1-\alpha) f(\rho''), \quad 0 \leq \alpha \leq 1 \tag{11}$$

The convesity of the thermodynamic free energy density guarantees that $f(\rho)$ is a continuous function of the density and implies directly that the thermodynamic pressure, which is <u>defined</u> unambiguously in the thermodynamic limit by

$$p(\rho) = \rho^2 \frac{\partial}{\partial \rho} \left[f(\rho)/\rho \right] = \rho \frac{\partial f(\rho)}{\partial \rho} - f(\rho) , \tag{12}$$

is a monotone non-decreasing function of the density. Convexity also guarantees that this derivative of the free energy which defines the pressure will be a continuous function of the density almost everywhere. For the potentials one is usually interested in, say Lennard-Jones, . Dobrushin and others have actually proven (at least for classical systems) that the derivative of $f(\rho)$ exists everywhere which means that the pressure is a continuous function of the density. Hence for classical and presumably also for quantum systems, with Lennard-Jones type potentials, there are no 'anti-phase' transitions in which the pressure changes discontinuously as a function of the density. This discussion of thermodynamic limits and convexity tells us that if we want to go to the van der Waals limit $\gamma \to 0$ we had better do it after we have taken the thermodynamic limit $\Omega \to \infty$.

Actually one can do the two limits simultaneously if one is carefull always to keep the range of the Kac potential γ^{-1} small compared to the size of the system,

$$\gamma^{-1} \ll |\Omega|^{1/\nu} \quad .$$

To come to the main point then I will consider a pair potential of the form,

$$u(r) = q(r) + \gamma^{\nu} \varphi(\gamma r) \quad ,$$

with the following properties:

$$q(r) \geq A \frac{1}{r^{\nu+\epsilon}} \quad \text{for } r \leq r_o \tag{13}$$

$$q(r) = 0 \quad \text{for } r \geq t \ , \ q(r) \geq 0 \ , \quad \text{for all } r$$

$$\varphi(x) \leq 0, \quad |\varphi(x)| < B \quad \text{for all } x \ , \tag{14}$$

$$|\varphi(x)| \leq C \frac{1}{x^{\nu+\epsilon}}$$

with r_o, A, t, B, C, ϵ, and a defined in (5), positive constants. These conditions are sufficient to guarantee that the thermodynamic limit $f(\rho,\gamma)$ and $p(\rho,\gamma)$ exist and are continuous.

It turns out using properties of convex functions that proving (4) is equivalent to proving that

$$\lim_{\gamma \to 0} f(\rho,\gamma) \equiv f(\rho,0) = CE \{f^o(\rho) - a\rho^2\} \tag{15}$$

Here $f^o(\rho)$ is the thermodynamic free energy of the reference system and CE - convex envelope - means essentially that the Gibbs double tangent construction is applied to the function $f^o(\rho) - a\rho^2$ (see Fig. 1). More precisely the CE $\{\bar{\Psi}(\rho)\}$, of any function $\Psi(\rho)$, is the largest convex function which is less or equal to $\Psi(\rho)$.

We shall prove (15) by obtaining upper and lower bounds on the free energy which will coincide with the right side of (15) in the van der Waals limit $\gamma \to 0$. To obtain these bounds we assume that Ω is a cube.* Ω is then divided into M congruent smaller cubes (cells) $\omega_1 \ldots \omega_M$ and estimates are used for the interactions across cell boundaries to relate the free energy of Ω to the sum of the free energies of the cubes $\omega_1 \ldots \omega_M$.

From these upper and lower bounds, the free energy in the van der Waals limit is calculated by means of a succession of limit operations. First the thermodynamic free energy is calculated from $F(N, \Omega, \gamma)$ by taking the thermodynamic limit. The simplest way of taking this limit is to double the side of cube Ω repeatedly, adjusting N at each step to the value $\rho |\Omega|$, ($F(N, \Omega, \gamma)$ may be defined for

---

*Since in the thermodynamic limit the free energy density is independent of the shape of Ω this assumption is no restriction.

nonintegral N by linear interpolation). The next

operation is to take the van der Waals limit $\gamma \to 0$.

The upper and lower bounds on $F(N, \Omega, \gamma)$ which we shall derive

lead to upper and lower bounds on $f(\rho, 0)$. These bounds depend on

the volume ω of the cells ω_i used. The bounds can be simplified by

a third limit process $\omega \to \infty$, leading to (15). This triple limit

process corresponds to the following relationships:

$$r_o \ll \omega^{1/\nu} \ll \gamma^{-1} \ll |\Omega|^{1/\nu}$$

among the four characteristic lengths of our calculation: the range

of the short-range potential, the size of the cells, the range of the

Kac potential, and the size of the container.

Upper Bound on the Free Energy

Finding an upper bound on the free energy is equivalent to

finding a lower bound on the partition function. To obtain a lower

bound, we divide the cube Ω into M smaller cubical regions (cells)

$\omega_1 \ldots \omega_M$, each of side $(s + t)$ where t is defined in (13) and s is a

positive length such that $(s + t)$ is a submultiple of the side of

Ω (see Fig. 2). Since the small cubes completely fill Ω its volume

is given by

$$|\Omega| = M(s + t)^\nu .$$

For each ω_i, let ω_i' be the cube of side s consisting of all points

within ω_i whose distance from the boundary of ω_i is at least $\frac{1}{2}t$. A

lower bound on $Z(N, \Omega, \gamma)$ can be obtained by selecting any set of

integers $N_1, N_2, \ldots N_M$ which add up to N, and considering only the

contribution to the integral in (8) from configurations where there

are N_1 particles in the cell ω'_1, N_2 in ω'_2, and so on. The quantity so calculated we denote by $\tilde{Z}(N_1, N_2,...,N_M)$. There are $N!/(N_1!N_2! ...N_M!)$ ways of choosing the N_1 particles to go in the cell ω'_1, N_2 to go in ω'_2, etc., and since the particles are identical all these different ways give the same contribution to the integral. Multiplying a typical contribution by $N!/\prod_i N_i!$ we thus obtain

$$Z \geq \tilde{Z}(N_1, N_2 \ldots N_M)$$

$$= \prod_i \left[(1/N_i!) \, \lambda^{-\nu N_i/2} \right]$$

$$\times \int_{\omega'_1} \ldots \int_{\omega'_M} e^{-\beta U} \, d\underline{x}_1 \ldots d\underline{x}_N \, , \tag{16}$$

where the first N_1 of the N ν-fold integrations are taken over the region ω'_1, the next N_2 over ω'_2, and so on.

To obtain a lower bound on the integral in (16), we write

$$U = U' + U'' \tag{17}$$

where U' is the contribution to the total potential energy from pairs of particles that are both in the same cell, and U'' is the contribution from pairs that are in different cells. If U''_{max} is some upper bound on U'', then (16) implies

$$\tilde{Z} \geq \prod_i \left[(1/N_i!) \, \lambda^{-\nu N_i/2} \right]$$

$$\times \int_{\omega'_1} \ldots \int_{\omega'_M} \exp \left[-\beta U' - \beta U''_{max} \right] d\underline{x}_1 \ldots d\underline{x}_N$$

$$= \left[\prod_i Z(N_i, \omega', \gamma) \right] e^{-\beta U''_{max}} \, , \tag{18}$$

where $Z(N_i, \omega', \gamma)$ is the partition function for N_i particles in a cube ω', of side s.

To obtain a lower bound on the exponential factor in (18) we note that because of our assumption that $q(r) = 0$ for $r > t$ there will be no short range interaction between cells N_i' so that

$$U'' = W'' . \qquad (19)$$

A convenient upper bound for W'', the long-range contribution to U'', is

$$W'' \leq \sum_{i < j} N_i N_j w_{max}(\underline{k}_{ij}) \qquad (20)$$

where

$$w_{max}(\underline{k}_{ij}) \equiv \underset{\substack{x \in \omega_i \\ y \in \omega_j}}{Max} \; w(\underline{x} - \underline{y}, \gamma) = \underset{r \in \omega_o}{Max} \; w(\underline{k}_{ij} + 2\underline{r}, \gamma) \qquad (21)$$

where ω_0 is a cube of side (s + t) centered at the origin, and $\underline{k}_{ij}$ is the vector from the center of ω_i to that of ω_j. Using this bound we obtain,

$$F(N, \Omega, \gamma) \leq \sum_i F(N_i, \omega', \gamma)$$

$$+ \sum_{i < j}^{M} N_i N_j \; w_{max}(\underline{k}_{ij}) \qquad , \qquad (22)$$

This upper bound holds for any choice of $N_1, N_2, \ldots N_M$ whose sum is N.

To obtain an upper bound on the thermodynamic free energy in the van der Waals limit it is simplest to consider the case where all of N_1, N_2 ... N_M are equal:

$$N_1 = N_2 = \ldots = N_M = \rho(s + t)^{\nu} . \tag{23}$$

(22) then yields, after taking the limit $\Omega \rightarrow \infty$ and showing that the second term on the right side of (22) goes over into a sum over an infinite cubical lattice ,

$$f(\rho,\gamma) \leq F(\rho(s + t)^{\nu}, \omega', \gamma)/(s + t)^{\nu}$$

$$+ \tfrac{1}{2}\rho^2(s + t)^{\nu} \sum_{\underline{k}}{}' w_{max}(\underline{k}) . \tag{24}$$

where $\sum_{\underline{k}}'$ is an infinite sum over the complete infinite lattice of possible vectors $\underline{k}_{ij}$ except $\underline{k} = 0$.

Taking now the limit $\gamma \rightarrow 0$ we obtain, making use of the fact that on the scale γ^{-1} the size of the cells ω becomes infinitesimally small,

$$f(\rho,0) \leq F^{o}(\rho(s + t)^{\nu}, \omega')/(s + t)^{\nu} - a\rho^2 \tag{25}$$

The final limiting process is to make the cell size infinite by making $s \rightarrow \infty$. Applying this limit on both sides of (25) and using the continuity of $f^{o}(\rho)$, we obtain, since the volume of the cell ω' is s^{ν},

$$f(\rho,0) \leq f^{o}(\rho) - a\rho^2 . \tag{26}$$

The inequality (26) can be strengthened by using the fact, Eq. (11), that $f(\rho, \gamma)$ is a convex function of ρ; so that $f(\rho, 0)$, being the limit of a sequence of convex functions is also convex. Because of this (26) implies

$$f(\rho, 0) \leq CE\{f^o(\rho) - a\rho^2\} \tag{27}$$

Lower Bound on the Free Energy

We again divide the cube Ω into cubical cells $\omega_1 \ldots \omega_M$, each of side $s + t$. The partition function may be written

$$Z(N, \Omega, \gamma) = \sum_{N_1 \ldots N_M} Z(N_1, \ldots, N_M) \tag{28}$$

where the sum is over all sets of M nonnegative integers adding up to N and $Z(N_1, \ldots, N_M)$ means the contribution to Z from configurations with exactly N_i particles in cell ω_i (i = 1,2...M). Since there are $(N + M - 1)!/N!(M-1)!$ terms in the sum it has the upper bound

$$Z(N, \Omega, \gamma) \leq [(N + M - 1)!/N!(M - 1)!]$$
$$\times \underset{N_1 \ldots N_M}{\text{Max}} Z(N_1, \ldots, N_M), \tag{29}$$

the maximum being taken over all sets of nonnegative integers $N_1 \ldots N_M$ which add up to N. The combinatorial argument which led to (16) gives, when applied to $Z(N_1, \ldots, N_M)$, the formula

$$Z(N_1,\ldots,N_M) = \prod_i \left[(1/N_i!) \quad \lambda^{-\nu N_i/2} \right]$$

$$\times \int_{\omega_1} \cdots \int_{\omega_M} e^{-\beta U} \, d\underline{x}_1 \ldots d\underline{x}_N \tag{30}$$

where the first N_1 of the N integrations are over the cell ω_1, [not ω'_i as in (16)] the next N_2 over ω_2, and so on.

To obtain an upper bound on the integral in (30) we separate the potential energy U into three parts:

$$U = Q' + \widetilde{Q} + W \tag{31}$$

where Q' is the contribution to U from short-range interactions between particles that are in the same cell, $\widetilde{Q}$ the contribution from short-range interactions between particles that are in different cells and W is the total contribution from long-range interactions. If $\widetilde{Q}_{min}$ and W_{min} are lower bounds on $\widetilde{Q}$ and W, then (30) and (31) lead to the inequality, analogous to (18),

$$Z(N_1,\ldots,N_M) \leq \left\{ \prod_i Z^o(N_i, \omega) \right\} \exp\left[-(\widetilde{Q}_{min} + W_{min})/kT \right] \tag{32}$$

where $Z^o(N_i, \omega)$ is the partition function for N_i particles of the reference system in a cube of side s + t.

We now make use of the condition that $q(r) \geq 0$. This implies $\widetilde{Q} \geq 0$ or $\widetilde{Q}_{min} = 0$. A simple lower bound on W, the total long-range interaction, is given by

$$W \geq \tfrac{1}{2} \sum_i \sum_j N_i N_j \, w_{min}(\underline{k}_{ij}) \tag{33}$$

where

$$w_{min}(\underline{k}_{ij}) \equiv \text{Min } w(\underline{x} - \underline{y}, \gamma) = \underset{r \in \omega_o}{\text{Min }} w(\underline{k}_{ij} + 2\underline{r}, \gamma). \tag{34}$$

These formulas are analogous to (20) and (21) but the $i = j$ terms are now included. Since $N_i N_j \leq \frac{1}{2}N_i^2 + \frac{1}{2}N_j^2$, and $w_{min}(\underline{k}_{ij}) \leq 0$, we may deduce from (33) the inequality

$$W \geq \frac{1}{2} \sum_i \sum_j (\frac{1}{2}N_i^2 + \frac{1}{2}N_j^2) w_{min}(\underline{k}_{ij}) = \frac{1}{2} \sum_i N_i^2 \sum_j w_{min}(\underline{k}_{ij}). \tag{35}$$

The sum over j may be extended, without destroying the validity of (35), to include the infinite network of cells continuing outside Ω the pattern established inside it by the cells $\omega_1 \ldots \omega_M$.

Substituting this into (32) and using (29) we obtain

$$F(N, \Omega, \gamma) \geq kT \log [N! (M - 1)!/(N + M - 1)!]$$

$$+ \underset{N_1 \ldots N_M}{\text{Min}} \sum_{i=1}^{M} \{F^\circ(N_i, \omega) + \frac{1}{2}N_i^2 \sum_k w_{min}(\underline{k})\} \tag{36}$$

The second term on the right can be simplified by means of a simple property of the convex envelope of a function. This property is: for any f,

$$M^{-1} \sum_{i=1}^{M} f(N_i) \geq M^{-1} \sum_i CE\{f(N_i)\} \geq CE\{f(M^{-1} \sum_i N_i)\} \tag{37}$$

where the first inequality follows from the fact that $CE\{f\}$ is a lower bound on f, and the second from the fact that $CE\{f\}$ is convex. Thus we find

$$F(N, \Omega, \gamma) \geq kT \log [N! \, (M-1)!/(N+M-1)!]$$

$$+ M \, CE \left\{ F^o(N/M, \, \omega) + \tfrac{1}{2}(N/M)^2 \sum_k w_{min}(\underline{k}) \right\} \qquad (38)$$

where CE $\{ \quad \}$ is the convex envelope of the quantity in braces regarded as a function of N/M.

To apply the triple limiting process described earlier we first divide (38) by $|\Omega|$ and take the thermodynamic limit $\Omega \to \infty$, using Stirling's formula for the factorials which becomes exact in the limits we take. The result is

$$f(\rho, \gamma) \geq -kT \left\{ (s+t)^{-\nu} \log [1 + \rho(s+t)^\nu] \right.$$

$$+ \rho \log [1 + \rho^{-1}(s+t)^{-\nu}] \}$$

$$+ CE \left\{ (s+t)^{-\nu} F^o(\rho(s+t)^\nu, \, \omega) \right.$$

$$+ \tfrac{1}{2}\rho^2(s+t)^\nu \sum w_{min}(\underline{k}) \} \qquad . \qquad (39)$$

since

$$N/M = \rho \, |\Omega|/M = \rho(s+t)^\nu.$$

Taking the limit $\gamma \to 0$ we obtain

$$f(\rho,0) \geq -kT \left\{ (s + t)^{-\nu} \log \left[1 + \rho(s + t)^{\nu} \right] \right.$$

$$+ \rho \log \left[1 + \rho^{-1}(s + t)^{-\nu} \right] \right\}$$

$$+ CE \left\{ (s + t)^{-\nu} F^{o} (\rho(s + t)^{\nu} , \omega) - a\rho^2 \right\} , \qquad (40)$$

where use has been made of the fact, proven in [1], that for a sequence of functions $f_n(\xi)$ converging uniformly on an interval to $f(\xi)$, $\lim\limits_{n \to \infty} CE \{f_n(\xi)\} = CE \{f(\xi)\}$. Finally, taking the limit $s \to \infty$, where the convergence is again uniform , and using (4), we obtain

$$f(\rho,0) \geq CE \left\{ f^{o}(\rho) - a\rho^2 \right\} . \qquad (41)$$

Combined with (27) this gives (15) for the cases considered here.

Metastability

We have seen that a rigorous analysis of the equilibrium properties of a system with 'long range' attractive potentials yields below T_c a range of densities in which the free energy is given by the Gibbs double tangent construction, see Fig. 1. No meaning at all is given to the dotted line in Fig. 1 which was interpreted by Maxwell (an interpretation apparently supported by experiment; we shall hear more about this from Jim Langer) to represent metastable states.(The dashed line represents unstable states.) Indeed the whole problem of metastable states represents somewhat of an embarrassment to rigorous statistical mechanics at the present time. For while the van der Waals-Maxwell theory .

suggests that these states are the 'analytic continuations' of
the equilibrium state there are many who argue, Langer and Fisher
among them, that this is one of the qualitative features of the
infinite range potential limit which does not persist for finite
range potentials. It is argued that in first order phase transi-
tions in real systems there is an essential singularity blocking
analytic continuation. Even if this argument should turn out to
be incorrect the question still remains of how to define (with
or without analytic continuation) metastable states precisely,
with some justification from first principles.

A theory of metastability should describe the familiar
experimental facts about the large variety of metastable states
occurring in nature. These include supercooled vapors and liquids,
supersaturated solutions, and ferromagnets in the part of the
hysteresis loop where the magnetization and the applied magnetic
field are in opposite directions. They generally arise when some
thermodynamic parameter of the system, such as the temperature or
magnetic field, is changed from a value for which the stable
equilibrium state has a single thermodynamic phase, to one for
which it has at least part of the system in some new thermodynamic
phase. Instead of making the appropriate phase transition, however,
the system may go over continuously into a one-phase state, called
a metastable state, which appears, while it lasts, to be stationary
in time in the same manner as a stable equilibrium state. The
properties of the metastable state are found to be reproducible; that
is, they 'appear' to be completely determined by the values of the

thermodynamic parameters, in just the same way as those of a

stable equilibrium state. The distinguishing feature of a meta-

stable state is that, eventually, either through some external

disturbance or a spontaneous fluctuation which nucleates the

missing phase in some small part of the system, the system begins

an irreversible process which leads it inexorably to the correspond-

ing stable equilibrium state. Thermodynamically, the irreversibility

of this transition corresponds to a decrease in free energy or an

increase in entropy.

This indicates that we may characterize metastable thermodynamic

states by the following properties:

Only one thermodynamic phase is present   (a)

A system that starts in this state is likely to
take a long time to get out   (b)

Once the system has gotten out, it is unlikely to
return   (c)

One might add the statement that thermodynamics applies to

the metastable state - for example, the usual theory would apply if

a substance in such a state were taken around a Carnot cycle.

A complete theory of metastability must then describe both the

static properties of these states as well as the dynamics of their

persistence and decay. Some of the basic ideas underlying this

dynamics are already contained in Maxwell's discussion of metastable

states: Maxwell recognized the importance of nucleation; he saw that

to set up the metastable state we must be sure that none of the new

phase is present. (You will hear much more about this from Langer and

other speakers at this conference.)

The relationship between the static equilibrium like view of the metastable state that comes from the van der Waals-Maxwell like theories and the kinetic time-dependent view coming from nucleation and similar theories is very central to an understanding of the metastable state. It might indeed turn out that it is not possible in general to give a mathematically precise characterization of metastable states in a purely static formalism. I hope though that this will not be the case and what Penrose and I have done [2] is show how one can define, in the limit $\gamma \to 0$, through an extension of the formalism of equilibrium statistical mechanics, restricted equilibrium states which satisfy the above criteria of metastable thermodynamic states.

We begin by making precise the notion, inherent in the previous discussion, of imposing a restriction on the system which keeps its density roughly uniform. In general, such a restriction may be represented by confining the configuration of the system to a suitable region R in configuration space. In order for this region to correspond to a metastable state, the restrictions defining it should correspond to the imporision of a roughly uniform density, in accordance with the criterion (a), and it should also have properties corresponding to the conditions (b) and (c) mentioned earlier: If the dynamical state is initially in R, it is unlikely to escape quickly, and once it has escaped, it is unlikely to return.

To compute the conditional probabilities implicit in (b) we shall, as is usually done in statistical mechanics, use the Gibbs ensemble made up by taking an equilibrium ensemble and selecting from it at some initial time, $t = 0$, all those systems whose configurations are in R.

We call this ensemble a <u>restricted equilibrium ensemble</u>. The
conditional probability p(t) of the configuration being outside
R at time t is then equal to the fraction of the members of this
subensemble that are no longer in R at time t.

To avoid the complications arising from the walls of the con-
tainer, where nucleation of the new phase may proceed at a different
rate from that in the bulk liquid, we did our calculations for a
system with periodic boundary conditions. The escape rate we
estimate will therefore be the homogeneous nucleation rate, pro-
portional to the volume of the system. The effect of walls, which
is a special case of the nucleation of a new phase at the surface
of a foreign substance, is also discussed. We find that perfectly
elastic walls do not facilitate the formation of liquid droplets
in a supercooled vapor but may play an important role in serving
as loci for the formation of vapor cavities in superheated liquids.
A similar effect is observed for real walls.

This treatment of metastable states hinges on finding a suitable
region R in configuration space. The ideal choice would, perhaps, be
the one minimizing the escape rate - that is, the probability per unit
time for the configuration of the system to move out of R. We did not
attempt the difficult task of optimizing the region R. Instead, we
made our choice on physical grounds and showed that this choice leads
to a very small escape rate (so that the minimum escape rate must be
at least as small). The basic idea of our method is the following:
We take the cubical region Ω of volume $|\Omega|$, in which our system con-
taining N particles is confined and divide it up as before into M
cubical cells ω_i; i = 1,....M, $|\omega_i| = \omega = |\Omega|/M$.

Let n_i be the number of particles in ω_i, $\sum\limits_{i=1}^{M} n_i = N$, $\rho_i = n_i/\omega$, $\frac{1}{M} \sum \rho_i = \rho$. We may consider now the space $\{\rho_k\}$. A point in this space corresponds to a specification of the average density of each cell. We define the region R in the configuration space by restricting the ρ_i to a certain region in the $\{\rho_k\}$-space. We choose in particular the constraints

$$\rho^- < \rho_i < \rho^+ , \qquad \text{where} \quad \rho^- < \rho < \rho^+ \tag{42}$$

At $t = 0$ we assume that our system is represented by a canonical ensemble restricted to R, i.e.

$$\mu(x, t = 0) = \begin{cases} (h^{3N} N! \, Z_R)^{-1} e^{-\beta H(x)} , & x \in R \\ \\ 0, & \text{otherwise} \end{cases} \tag{43}$$

$$Z_R = \frac{1}{N! h^{3N}} \int_R e^{-\beta H(x)} \, dx$$

We let $p(t) = 1 - \int_R \mu(x,t) \, dx$; $\dfrac{dp(t)}{dt}$ is the escape rate and it is readily shown that for all t, $\dfrac{dp(t)}{dt} \leq \dfrac{dp(t)}{dt}\bigg/_{t=0} \equiv \lambda$.

We want to show that we can choose R in such a way that λ can be made arbitrary small in a certain limit even though the probability that the equilibrium system will be in R, Z_R/Z is vanishingly small. It turns out that this can be accomplished if we can show that the free energy for a given set densities ρ_i, $F(\{\rho_i\}, \omega, M)$ has the property that

$$[F(\{p'_k\}, \omega, M) - F(\{p\}, \omega, M)] \geq |\omega| \, [C + o(1)], \qquad (44)$$

where $\{p'_k\}$ is a point on the <u>boundary</u> of $\{p_k\}$-space, as specified in (42), and C is some positive constant. Here o(1) means a quantity which goes to zero as $|\omega| \to \infty$ in the limit considered. When (44) holds, we show that,

$$\lambda \leq K \, \frac{|\Omega|}{\omega^{1/\nu}} \, e^{-\beta |\omega| \, [C + o(1)]} \qquad (45)$$

where K is a constant which remains finite in the limit and ν is the dimensionality of the space considered. One has therefore to arrange things in such a way that as $|\Omega|$ and $|\omega| \to \infty$, the ratio $\omega^{-1} \ln |\Omega| \to 0$ and the inequality about the free energies, Eq. (44), remains valid.

To accomplish this we had to consider instead of the previous triple limit

$$r_o \leq \omega^{1/\nu} \ll \gamma^{-1} \ll |\Omega|^{1/\nu}$$

where each of the last three quantities went to infinity in succession a limit in which

$$r_o \ln |\Omega| \ll |\omega| \ll \gamma^{-\nu} \ll |\Omega|,$$

with all quantities going to infinity.

The limits are therefore not taken anymore independently. With this way of taking the limit everything for equilibrium systems goes as before and in addition we prove that (44) and hence (45) holds.

Thus if the system starts in R it remains there with probability 1 (in the limit). We actually showed this for only part of the mestastable region but Millard and Lund [10] extended our result to the whole region.

References

1. J.L. Lebowitz and O. Penrose, J. Math. Phys., $\underline{7}$, 98 (1966).

2. O. Penrose and J.L. Lebowitz, J. Stat. Phys. $\underline{3}$, 211 (1971).

3. P.C. Hemmer and J.L. Lebowitz,"Systems With Weak Long-Range Potentials", to appear <u>Critical Phenomena and Phase Transitions</u> C.Domb and M.Green, editors, Academic Press, 1974.

4. N.G. van Kampen, Phys. Rev. <u>135A</u>, 362 (1964); Physica <u>48</u>, 313 (1970).

5. M. Kac, G.E. Uhlenbeck and P.C. Hemmer, J. Math. Phys. $\underline{4}$, 216 (1963); $\underline{4}$, 229 (1963); $\underline{5}$, 60 (1964).

6. G.A. Baker, Phys. Rev., <u>122</u>, 1477 (1961).

7. E. Lieb, J. Math. Phys. $\underline{7}$; 1016 (1966).

8. D.J. Gates and O. Penrose, Comm. Math. Phys. <u>15</u>, 255; <u>16</u>, 231; <u>17</u>, 194 (1970).

9. D. Ruelle, <u>Statistical Mechanics</u>, Benjamin, 1969.

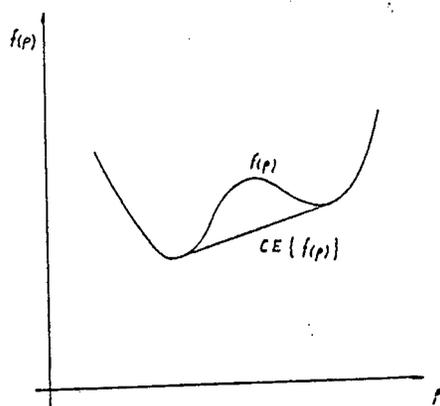10. K. Millard and L. Lund, J. Stat. Phys. $\underline{8}$, 225 (1973).
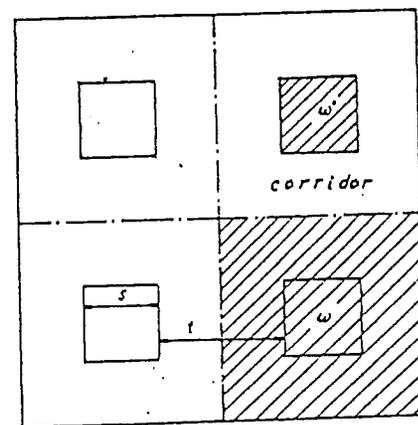
Fig. 1. A function $f(\rho)$ and its convex envelope.

Fig. 2. Division of Ω into cells.