Some Properties of the Informational Model of the Liquid State

Witold Brostow[1] and Yvon Sicotte[1]

Equilibrium properties of one-component liquids are obtainable—as suggested by Collins—from the coding procedure in terms of distances between neighboring molecules. The monatomic case is dealt with first, and consequences of some simplifying assumptions are explored. The connection between the probability $\psi(R)$ of an intermolecular distance R and the usual pair distribution function is considered. The treatment is then generalized to the case of heterogeneous multiatomic molecules.

1. INTRODUCTION

Among fundamental approaches to the liquid state, there are two relatively much less exploited than the rest: physical models and the informational method. These two approaches are also the most recent ones, which partly explains their current situation.

The method of physical models is frequently associated with the name of Bernal; binary radial distribution function $g(R)$ curves, obtained by Bernal and his colleagues counting distances R between balls, turned out to be comparable[1] to the results obtained for real liquids such as argon using X-ray and neutron diffraction techniques. Related studies have been made by Scott.[2] As discussed by Pryde,[3] an interesting result of Scott is reproduction by a system of balls of the volume change on melting of monatomic liquids.

---

[1] Département de Chimie, Université de Montréal, Montréal, P.Q., Canada.

Using ellipsoids instead of balls, Godbout and Sicotte[4] obtained similar results; their discussion is in terms of a quantity they call the coefficient of space occupancy ϕ_0 :

$$\phi_0 = N_A v_0 / V \tag{1}$$

where V is the molar volume, N_A is the Avogadro number, and v_0 is the proper volume of a molecule. For a substance which at the temperature of $0°K$ crystallizes in a hexagonal or a face-centered cubic lattice the latter parameter is given by[4]

$$v_0 = 0.7405 V(0°K)/N_A \tag{2}$$

The same parameter ϕ_0 turns out to be useful for considering dielectric polarization in liquids.[5] In general, the studies of models seem to suggest that a liquid may be represented by a compact assembly of irregular poly-hedra, with molecules at tops of each polyhedron. An alternative, namely placing molecules each in the middle of a polyhedron, has been discussed by Finney.[6] In either case spatial distributions of molecules are essential for understanding the behavior and properties of liquid phases.

The suggestion that information theory be applied to the liquid state is due to Collins.[7] He also proposed[7] a coding procedure in terms of distances between neighboring molecules, the links forming impenetrable barriers for other molecules. This reproduces features inferred from the models: A liquid is now represented by an assembly of irregular Delaunay polyhedra, with atoms as vertices of each polyhedron.

The space allotted to the motions of each molecule by such a procedure is called a Voronoi polyhedron (or a Dirichlet region, or a Wigner–Seitz cell); details of the respective geometric constructions are given by Finney[6] and Kiang.[8] The dual relationship between Delaunay and Voronoi polyhedra is discussed by Kiang. Apparently the mathematical formalism is the same whether dealing with the fragmentation of celestial bodies or with the attri-bution of microscopic spaces to molecules; in astrophysics, however, the Voronoi figures are more important than those of Delaunay,[8] while for our approach the reverse is true.

The object of the present work is to explore further the informational route to understanding liquids. In this context it might be worthwhile to recall the superior position of information theory as compared to statistical mechanics; we mean by this that the relations of statistical mechanics and thermodynamics may be obtained from the theory of information.[9] The approach might therefore turn out to be profitable in the long run even if it is not the easiest to handle. An indirect continuation of an earlier study[4] has also been an incentive for undertaking this work.

2. CODING FOR MONATOMIC LIQUIDS

The case of a two-dimensional liquid with summation over coordination numbers has been treated by Collins.[10] Information concerning the spatial distribution of monatomic molecules in a three-dimensional liquid may be transmitted over a communication channel in terms of interatomic distances (links) R, angles between links, or in a mixed way involving both distances and angles. The appearance of a redundant link (the length and location of which may be inferred from the data transmitted before) coincides with "closing a shell" around a given molecule. Denote by $\psi_z(R)\,dR$ the probability that link length is between R and $R + dR$ for a pair of neighboring molecules at least one of which has the coordination number z. Clearly, we have a certain distribution of probabilities $\psi_z(R)$. We also have a certain distribution of probabilities ω_z of values of z among atoms (the topological contribution). Further, we have the usual momentum or kinetic contribution, which can be expressed in terms of probabilities $\omega(p)$ of momenta p. For the general or informational entropy, assuming the coordination numbers for the individual atoms to be statistically independent, we write therefore

$$
\begin{aligned}
S = S^M + S^C + S^Z = \ & -3Nk \int_{-\infty}^{\infty} \omega(p) \ln \omega(p)\, dp \\
& - Nk \sum_z \omega_z \int_0^{\infty} \psi_z(R) \ln[\xi\psi_z(R)/R^2]\, dR \\
& - Nk \sum_z \omega_z \ln \omega_z
\end{aligned}
\tag{3}
$$

where N is the number of molecules (taken equal to the number of effectively uncorrelated links in the middle term), k is the Boltzmann constant (intro-duced since we intend to arrive eventually at the thermodynamic entropy); ξ is a measure constant, the same as used in Ref. 10 (for the use of measures in general, see, e.g., Ref. 11). Superscripts M, C, and Z refer to momentum, configurational, and topological terms, respectively.

Consistently with coding in terms of pairs, we represent the configura-tional energy U^C as a sum of pair interactions; we write for the total energy

$$
U = U^M + U^C = 3N \int_{-\infty}^{\infty} (p^2/2m)\, \omega(p)\, dp + \tfrac{1}{2}N \sum_z \omega_z z \int_0^{\infty} \psi_z(R)\, u(R)\, dR
\tag{4}
$$

where m is the molecular mass and $u(R)$ is the pair interaction potential.

For volume we have, correspondingly,

$$
V = N \sum_z \omega_z \mu_z \int_0^{\infty} R^3 \psi_z(R)\, dR
\tag{5}
$$

where each μ_z is a geometric factor which takes care of the spatial configuration (skewness) and also of the fact that a given atom participates in a number of polyhedra. With the volume of a regular tetrahedron of side a given by $a^3/(2 \cdot 6^{1/2})$, one would expect μ_z to be approximately proportional to $(2 \cdot 6^{1/2})^{-1}$.

The thermodynamic entropy S_T may be obtained by maximizing expression (3) subject to fluctuations of energy and volume. The following three normalizations have also to be taken into account:

$$\int_{-\infty}^{\infty} \omega(p)\, dp = \int_0^{\infty} \psi_z(R)\, dR = \sum_z \omega_z = 1 \tag{6}$$

Three of the five undetermined multipliers featuring in extremization are eliminated in the usual way $(\partial S^{\text{max}}/\partial \psi_z = \partial S^{\text{max}}/\partial \omega = \partial S^{\text{max}}/\partial \omega_z = 0)$. Two are evaluated from the thermodynamic identities $(\partial S/\partial V)_U = P/T$ and $(\partial S/\partial U)_V = T^{-1}$, where P denotes pressure. The final formula is

$$S_T = Nk \ln(Q_z/\xi) + \tfrac{3}{2} Nk \ln(2\pi mkT) + (H/T) \tag{7}$$

The enthalpy H appears here explicitly, and we can immediately rewrite Eq. (7) in terms of the Gibbs function G. The function Q_z is given by

$$Q_z = \sum_z \exp \left\{ -\int_0^{\infty} [\ln(\psi_z/R^2) + \Omega_z]\, \psi_z\, dR \right\} \tag{8}$$

$$\Omega_z = [zu(R)/2kT] + (\mu_z PR^3/kT) \tag{9}$$

$$\psi_z = R^2 e^{-\Omega_z} \Big/ \int_0^{\infty} R^2 e^{-\Omega_z}\, dR \tag{10}$$

Further,

$$\omega(p) = (2\pi mkT)^{-1/2} \exp(-p^2/2mkT) \tag{11}$$

One obtains, as one should, $U^M = 3NkT/2$. Specification of the parameters in (4), (5), and (7) is completed with

$$\omega_z = Q_z^{-1} \exp \left\{ -\int_0^{\infty} [\ln(\psi_z/R^2) + \Omega_z]\, \psi_z\, dR \right\} \tag{12}$$

As mentioned above, the case of two dimensions has been considered by Collins.[10] The relations of the present section, however, do not represent a direct generalization of his treatment to three dimensions. In his approach one more constraint, related to the coordination numbers, is imposed upon the system [Eq. (42) in Ref. 10]. As compared to the treatment of our Eq. (3), this introduces one more Lagrangian parameter in maximizing the entropy formula.

The length distribution functions $\psi_z(R)$ have been introduced first to perform coding; from the definition, however, it is fairly evident that there should be a connection to the binary radial distribution function $g(R)$ as used in diffractometry. Consider a shell of thickness dR around a given molecule. According to the definition of $g(R)$, the number of molecules in the shell is $4\pi R^2 N g(R) \, dR/V$; this, of course, is an average, independent of the value of z for a given atom. On the other hand, for an atom of coordination z we may apparently represent the same number by $z\psi_z(R)$; this should be valid as long as we are inside the range of Delaunay polyhedra of a given molecule. Within this range, then,

$$\sum_z \omega_z z \psi_z(R) = 4\pi R^2 (N/V) \, g(R) \tag{13}$$

As for relations valid without limitation on the R range, Collins[12] obtained the exact formula for $\psi(R)$ in terms of number density N/V for the two-dimensional perfect gas. Obtaining such a formula for a three-dimensional liquid appears prohibitively difficult.

The unresolved problem in the use of equations such as (4) is, as in the two-dimensional case,[10] that of the highest value of z that ought to be included in the summations. Our coordination number resulting from the coding procedure in general is not identical with the diffractometric value. Even if the identity would be assured, there are—as discussed by Mikolaj and Pings[13]—no less than four distinct methods of evaluating z from X-ray and slow neutron scattering. Unless recourse to topology turns out to be profitable, it will be necessary to seek this information from equilibrium properties as known from experiments.

3. EQUILIBRIUM PROPERTIES FOR SOME SIMPLE MONATOMIC CASES

Collins[7] considered a simplified case of a three-dimensional liquid, with the topological contribution S^z neglected; he argued that the contribution of this term is small, except in the liquid–solid transition region. Accordingly, he assumed, instead of (3), a simpler formula

$$S = -3Nk \int_{-\infty}^{\infty} \omega \ln \omega \, dp - Nk \int_0^{\infty} \psi \ln(\xi \psi/R^2) \, dR \tag{14}$$

along with

$$U = 3N \int_{-\infty}^{\infty} (p^2/2m) \, \omega \, dp + \tfrac{1}{2} N\bar{z} \int_0^{\infty} \psi u \, dR \tag{15}$$

$$V = N\mu \int_0^{\infty} R^3 \psi \, dR \tag{16}$$

In the above equations the average coordination number $\bar{z}$ has been introduced. The $\omega(p)$ distribution is the same as in the preceding section [Eq. (11)]. The other distribution is given by

$$\psi(R) = \frac{R^2}{Q} \exp\left(-\frac{\bar{z}u}{2kT} - \frac{\mu P R^3}{kT}\right) \tag{17}$$

with

$$Q = \int_0^\infty R^2 \exp\left(-\frac{\bar{z}u}{2kT} - \frac{\mu P R^3}{kT}\right) dR \tag{18}$$

The thermodynamic entropy formula follows either from the formalism of this section or from averaging over coordination numbers in Eq. (7). The result is

$$S_T = Nk \ln(Q/\xi) + \tfrac{3}{2}Nk \ln(2\pi mkT) + (H/T) \tag{19}$$

The question arises, under what circumstances may the simplified relations of this section be valid? A good reference case is a solid with a periodic lattice, where the average $\bar{z}$ is equal to the local coordination number of each of the atoms. Decreasing the density should introduce atoms of low coordination numbers; at the same time differences between the values of z of neighboring atoms should increase. One concludes that the approximation may be useful for dense fluids, but worse at, say, the vapor–liquid critical region.

In the present approximation we can also try to relate $\psi(R)$ to $g(R)$ within the range of Delaunay polyhedra. By reasoning such as led to (13) we obtain

$$\psi(R) = (4\pi R^2 N/V\bar{z}) \, g(R) \tag{20}$$

Substituting the last result into the second term of (15), one obtains the well-known "standard configurational energy formula":

$$U^C = (N^2/2V) \int_0^\infty 4\pi R^2 ug \, dR \tag{21}$$

One notices also that since $u(R) \to 0$ rapidly when $R \to \infty$, one may substitute the definite integral in (21) and in the second term of (15) by the indefinite one (or vice versa, cf., e.g., Ref. 14), and compare directly the respective integrands.

Further, we can introduce ψ as given by (17) into (16) and, using μR^3 as the independent variable, integrate by parts. The result is

$$PV = NkT - \tfrac{1}{6}Nz \int_0^\infty R(du/dR) \, \psi \, dR \tag{22}$$

Substituting (20) into the last result leads to the so-called standard pressure equation[3,14]:

$$PV = NkT - (N^2/6V) \int_0^\infty 4\pi R^3 (du/dR) \, g \, dR \tag{23}$$

Following our informational route—we have recovered Eqs. (21) and (23), known from the statistical mechanics of fluids. The consistency between the two approaches has thus been demonstrated.

The virial of intermolecular forces $W = -\sum\sum_{i \neq j} R_{ij} \, du(R_{ij})/dR$ can now be written in terms of parameters of the present theory. From the pressure equation (22) this is

$$W = \tfrac{1}{2} N \bar{z} \int_0^\infty R(du/dR) \, \psi \, dR \tag{24}$$

The behavior of parameters such as μ and $\bar{z}$, and in particular their dependence on volume and temperature, is of interest. The parameter μ may be related to the coefficient of space occupancy as defined by (1) through

$$\mu = v_0 / \phi_0 \overline{R^3} \tag{25}$$

where v_0 is given by (2), and the bar denotes the usual system average; it has to be remembered that the $\mu(T)$ dependence is hidden in (25). As for the parameter Q, its volume dependence may be obtained from (19) as

$$\left(\frac{\partial \ln Q}{\partial T} \right)_V = \frac{H}{NkT^2} - \frac{3}{2T} - \left(\frac{\partial P}{\partial T} \right)_V \frac{V}{NkT} \tag{26}$$

Consider now in more detail the integration range for equations such as (16), (18), or (22). The upper limit of infinity is not to be taken literally—see, e.g., the discussion in Ref. 15 and also comments following Eq. (21) here. Evaluating the distance l within which a molecule contributes to thermodynamic functions is in general difficult, though an approximate proposal of Debye for calculating l^2 should be recalled.[16] In the present treatment formation of Delaunay figures assures a cutoff, but the range l clearly varies from one molecule to another. Kiang[8] performed Monte Carlo computations for a system of 80 particles on 6400 sites to establish the size distribution of random Voronoi polyhedra. As one might have expected, he finds that the curve falls fairly rapidly beyond a certain value of R (roughly equal to l, in view of the dual character of Voronoi and Delaunay figures). The curve remains continuous, though; a tail is present, which should be a consequence of random irregularity of the system.

Returning now to Eq. (22), a special case may be obtained by substituting (17) and (18) into (22) and comparing directly the integrands in (18) and

in the second term of (22). Dropping integration limits is appropriate for (22), but because of the situation just discussed, such a step clearly lacks rigor for (18). The consequences are of some interest, however; one obtains

$$du/dR = 6(NkT - PV)/N\bar{z}R \tag{27}$$

Thus, with increasing R, du/dR is predicted to decrease—as it does beyond the potential well. One could attempt to integrate Eq. (27). This would give $u(R)$ as an inherent part of the model, instead of invoking it from the outside. On the other hand, the results would be, at best, useful between the potential energy minimum and the "border" of Delaunay figures.

4. MULTIATOMIC CASE

We turn now to multiatomic molecules, which for generality will be assumed heterogeneous. Thus a molecule contains r segments with

$$r = \sum_y r_y \tag{28}$$

where the summation refers to kinds of segments present. A segment might be a single atom or a group of atoms (e.g., the methyl group or a polymeric segment).

We find it very convenient for handling the multiatomic case to introduce the internal coordination number, call it ζ, which is the number of nearest-neighbor segments belonging to the same molecule. Thus, local ζ may vary from one segment to another, while the average value $\bar{\zeta}$ is fixed for each molecule.

The essential step again is formulating a prescription for coding. We propose coding in terms of: (1) distances; (2) angles (for using distances and angles see the beginning of Section 2); (3) number given to each molecule; (4) internal coordination number; (5) indication of whether or not a given segment is connected to the one encoded immediately before; (6) for heterogeneous molecules the kind of segment. We note that in the general case indicating connectedness is essential. Consider, say, a two-dimensional two-ring molecule with information as in points 1–4 specified; we can still have at least two possibilities,
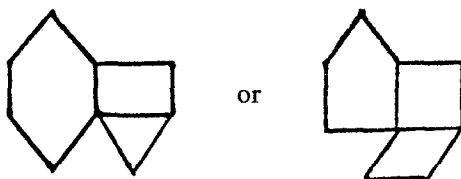
and of course we do not propose that the decoder consult independent sources (books of chemistry or physics, information which has not been transmitted over the channel).

We are now able to write for the entropy

$$S = S^D + S^M + S^C + S^Z + S^\zeta$$

$$= S^D - 3Nk \int_{-\infty}^{\infty} \omega(p) \ln \omega(p) \, dp - Nk \sum_y r_y$$

$$\times \sum_{y'} \int_0^{\infty} \psi_{yy'}(R) \ln \frac{\xi \psi_{yy'}(R)}{R^2} \, dR - Nk \sum_z \omega_z \ln \omega_z - Nk \sum_{z-\zeta} \omega_\zeta \ln \omega_\zeta$$

$$(29)$$

where S^D denotes internal (e.g., vibrational) contributions, which can be calculated by the usual methods; the momentum term is per whole molecule; the ω_ζ distribution refers to $z - \zeta$ external contacts of each segment, and is assumed independent of kinds of segments y; similarly, we have introduced ω_z instead of a series of ω_{zy}.

We now perform summations over all possible kinds of pairs. Denoting a pair of interacting segments by indices y and y', we have for the configurational energy

$$U^C = \frac{N}{2 \sum_y r_y \overline{(z - \zeta_y)}} \sum_y \sum_{y'} r_y r_{y'} \overline{(z - \zeta_y)} \overline{(z - \zeta_{y'})} \int_0^{\infty} u_{yy'}(R) \, \psi_{yy'}(R) \, dR$$

$$(30)$$

Consistent with the above, the volume is given by

$$V = \frac{N\mu}{2 \sum_y r_y \overline{(z - \zeta_y)}} \sum_y \sum_{y'} r_y r_{y'} \overline{(z - \zeta_y)} \overline{(z - \zeta_{y'})} \int_0^{\infty} R^3 \psi_{yy'}(R) \, dR \quad (31)$$

where $\psi_{yy'}(R)$ is given by the generalization of (20):

$$\psi_{yy'}(R) = [4\pi R^2 N r_y / V \overline{(z - \zeta_y)}] \, g_{yy'}(R) \tag{32}$$

The respective generalization of the pressure equation (22) is

$$PV = NkT - \frac{N}{6 \sum_y r_y \overline{(z - \zeta_y)}}$$

$$\times \sum_y \sum_{y'} r_y r_{y'} \overline{(z - \zeta_y)} \overline{(z - \zeta_{y'})} \int_0^{\infty} R \frac{du_{yy'}(R)}{dR} \, \psi_{yy'}(R) \, dR \tag{33}$$

5. CONCLUDING REMARKS

The informational approach seems to have at least one property in common with the—unattainable—exact theory of the liquid state: the

capability to describe all three basic states of matter. While the main framework of the theory is constructed, clearly much remains to be done. Information on the behavior of parameters of the theory such as μ's and z's and in particular on their temperature dependences is essential; this suggests recourse to the data of experimental thermodynamics.

We have recently embarked on the respective calculational project. With a judicious choice of $\bar{z}$ and μ, calculations of U^c from Eq. (15), V from (16), and then approximate prediction of the $g(R)$ curve within the Delaunay polyhedron from (20) is feasible for argon; the temperature and pressure are assumed, and the pair potential $u(R)$ is taken as known. To arrive at well-founded conclusions, valid not only for argon, it seems that fairly extensive computations might be necessary. While the relations in Section 4 are somehow anticipative of the treatment of multicomponent systems, mixtures should wait for the results of calculations for pure components. The treatment of mixtures should include that severe test of any approach to the liquid state: prediction of functions of mixing in terms of properties of pure components.

ACKNOWLEDGMENTS

REFERENCES

1. J. D. Bernal, *Proc. Royal Soc.* **A280**:299 (1964).
2. G. D. Scott, *Nature* **188**:908 (1960); **194**:956 (1962); G. D. Scott, A. M. Charlesworth, and M. K. Mak, *J. Chem. Phys.* **40**:611 (1964).
3. J. A. Pryde, *The Liquid State*, Hutchinson, London (1966), Chapter 8.
4. G. Godbout and Y. Sicotte, *Can. J. Chem.* **46**:967 (1968).
5. G. Godbout and Y. Sicotte, *J. Chim. Phys.* **65**:1944 (1968).
6. J. L. Finney, *Proc. Royal Soc. A* **319**:479, 495 (1970).
7. R. Collins, *Proc. Phys. Soc.* **86**:199 (1965).
8. T. Kiang, *Z. Astrophys.* **64**:433 (1966).
9. E. T. Jaynes, *Phys. Rev.* **106**:620 (1957); M. Tribus, *J. Appl. Mech.* **28**:1 (1961); A. Kossakowski, *Bull. Acad. Polon. Sci. Sér. Math. Astr. Phys.* **17**:263 (1969); J. Rothstein, in *Quantum Theory and Beyond*, T. Bastin, ed., Cambridge University Press, London (1971); C. Aslangul, *Compt. Rend.* **272**:1 (1972); M. Tribus and H. Motroni, *J. Stat. Phys.* **4**:227 (1972).
10. R. Collins, in *Proc. Battelle Colloquium on Phase Changes in Metals, Geneva-Villars*, McGraw-Hill, New York (1966).

11. J. F. C. Kingman and S. J. Taylor, *Introduction to Measure and Probability*, Cambridge University Press (1966).
12. R. Collins, *J. Phys. C* **1**:1461 (1968).
13. P. G. Mikolaj and C. J. Pings, *Phys. Chem. Liq.* **1**:93 (1968).
14. T. L. Hill, *Statistical Mechanics*, McGraw-Hill, New York (1956), Chapter 6.
15. W. Brostow, *Phys. Chem. Liq.* **3**:91 (1972).
16. P. Debye, *J. Chem. Phys.* **31**:680 (1959).