postulate about the localized description. Professor Musher called our attention to the connection between our closed shell discussion and the crystal field theory. Professors Charles A. Coulson, Benjamin M. Gimarc, and H. Bradford Thompson also have furnished helpful suggestions.

A New Model of Liquid Water

Oleh Weres[1] and Stuart A. Rice*

Contribution from the Department of Chemistry and The James Franck Institute, University of Chicago, Chicago, Illinois 60637. *Received April 5, 1972*

Abstract: A new model for liquid water is developed upon the basis of the intermolecular potential function proposed by Ben-Naim and Stillinger and some simple and clearly defined structural assumptions. It is primarily a model of the cell theory of fluids tradition, but it also includes a fairly well-developed treatment of the structural entropy contributions. It is demonstrated that the original Ben-Naim–Stillinger potential contains a rather serious error in its curvature in the rotational degrees of freedom and that the quantum effects in water are much too large to safely ignore. In the model proposed properties are expressed as functions of the concentrations of the various varieties of cells dealt with, and their equilibrium concentrations are obtained by minimizing the free energy. When proper account is taken of the potential function insufficiency, and of quantum effects, the calculated values of the thermodynamic properties and density are in fairly good agreement with experiment. Most of the discrepancy between calculation and experiment is easily explained, at least qualitatively, by the known defects of the cell model.

I. Introduction

The theoretical study of all but the simplest liquids has been and remains extremely difficult. The complexity of the intermolecular interactions has as yet in most cases precluded the development of realistic and tractable model intermolecular potential functions. Also, the existence of five or more degrees of freedom per molecule has made extensive numerical calculations nearly impossible. Water, because it is the simplest of the "nonsimple" liquids, and because of its importance, may be considered to be a test case for such studies.

Until recently theoretical structural and statistical mechanical studies of water have been limited to descriptions in terms of rather simply conceived structural entities whose interrelations and properties could be easily dealt with. All of the treatments of this school suffer from severe structural underdefinition, or excessive reliance on adjustable parameters, or both. Some of them have the additional quality of grossly disagreeing with available experimental data.

One group of simple models postulates the existence of microphases within the liquid. Usually the various microphases are supposed to have markedly different densities. Those models which require the existence of such regions of different densities are invalidated by the results of X-ray scattering studies by Narten, *et al.*,[2a] discussed by Narten and Levy,[2b] which deny the existence of density variations larger than those arising from thermal fluctuations. The most advanced models of this type are those proposed by Nemethy and Scheraga,[3] Vand and Senior,[4] and Jhon, *et al.*[5,6]

Another popular model of the structure of water is that of an ice-like hydrogen-bonded structure with non-hydrogen-bonded "interstitial" molecules in the lattice cavities. A typical example is the model proposed by Marchi and Eyring[7] and subsequently retracted by Eyring,[6] who cited Stevenson's[8] interpretation of the ultraviolet spectrum of water as strong evidence for the nonexistence of significant concentrations of non-hydrogen-bonded molecules at any temperature in the normal liquid range. The models proposed by Nemethy and Scheraga[3] and Vand and Senior[4] also fail on this score.

A "simple" model which we feel deserves more thorough discussion has been proposed by Angell.[9] He interprets the properties of water in terms of the concentrations of broken and unbroken hydrogen bonds without reference to other molecular properties. This scheme can be used to generate the classical two state thermodynamics often used to describe the properties of water. Since it is not at all necessary to associate the existence of microphases or non-hydrogen-bonded molecules with broken hydrogen bonds, this model does not conflict with the available spectroscopic data. Although it suffers from structural underdefinition, we find it rather elegant and note that it generates some impressive results. We consider Angell's model to be the best and least restrictive of the "simple" category. It may turn out to be a useful simplification for the classification of the molecular states derived from a more complete statistical mechanical model, but that is not now known. We do have one objection to Angell's model, however, in that it assumes that

---

[1] Fannie and John Hertz Foundation Fellow.

[2] (a) A. H. Narten, M. D. Danford, and H. A. Levy, *Discuss. Faraday Soc.*, **43**, 97 (1967); (b) A. H. Narten and H. A. Levy, *Science*, **165**, 447 (1969).

[3] G. Némethy and H. A. Scheraga, *J. Chem. Phys.*, **36**, 3382 (1962); **41**, 680 (1964).

[4] V. Vand and W. A. Senior, *ibid.*, **43**, 1878 (1965).

[5] M. S. Jhon, J. Grosh, T. Ree, and H. Eyring, *ibid.*, **44**, 1456 (1966).

[6] M. S. Jhon, E. R. Van Artsdalen, J. Grosh, and H. Eyring, *ibid.*, **47**, 2231 (1967).

[7] R. P. Marchi and H. Eyring, *J. Phys. Chem.*, **68**, 221 (1964).

[8] D. P. Stevenson, *ibid.*, **69**, 2145 (1965).

[9] C. A. Angell, *ibid.*, **75**, 3698 (1971).

the chemical potential of a hydrogen bond is temperature dependent only through the configurational entropy. This ignores the thermodynamic effects of the variation of intermolecular vibrational amplitudes with temperature.

The only "nonsimple" model (other than our own) is the one recently presented by Bell.[10] (Bell's paper appeared in print as we were making the final revisions, following the referee's comments, on this manuscript.) Bell's model is rather similar to our own though much simpler. It basically consists of a lattice gas on a body centered cubic lattice as does our model. Because of its similarity to our model we have chosen to discuss it in considerable detail. The length of this discussion has forced us to relegate it to an Appendix.[10a]

We feel that the most important single recent advance in the theoretical study of water has been the development of a new effective pairwise-additive-intermolecular potential energy function by Ben-Naim and Stillinger.[11] The form of this model potential is based on the assumption that the intermolecular potential may be divided into short range repulsive and electrostatic components. The short range (or dispersion) component is taken to be the Lennard-Jones potential suitable for the neon atom, which contains the same number of electrons as the water molecule, centered on the oxygen nucleus. The spherical symmetry of this component probably is not a bad approximation because the bond length of 0.96 Å is considerably smaller than the collision radius of neon, 1.41 Å. The electrostatic component is represented by a regular tetrahedron of charges of ±0.19 electron, 1 Å away from the oxygen nucleus and centered on it. Coulomb divergences are eliminated by multiplying the electrostatic term by a "switching function" which continuously reduces and finally eliminates the electrostatic term at oxygen-oxygen separations smaller than a certain value. The introduction of this "switching function" may, at first glance, seem unphysical, but actually is not, since it merely represents the short range repulsive effects of the very high electron densities in the immediate vicinities of the protons. The parameters of the electrostatic term are fitted to the second virial coefficient of water vapor and the experimentally determined first neighbor separation in ice.

Rahman and Stillinger[12] have carried out molecular dynamics calculations using the Ben-Naim-Stillinger potential (henceforth referred to as "the BNS potential") and obtained tolerably good results for the radial distribution function and certain other properties. They found that their results for the potential energy and self-diffusion coefficient could be improved considerably by simply increasing the strength of the potential by 6%. Ben-Naim has reported an attempt to solve the Percus-Yevick equation for water using the BNS potential.[13] It was only partially successful because of the strength of the intermolecular interactions involved.

Although the molecular dynamics and integral equation approaches might at first glance appear to offer possibilities for obtaining enticingly general descriptions of water, in reality they both have some major shortcomings. An as yet apparently insurmountable one is that both are incapable of dealing with quantum effects, which cannot be neglected due to the large spacing of the librational levels in water. The molecular dynamics technique does not, at present, lead to a direct evaluation of the entropy. Both approaches are just barely practical from the computational volume point of view; attempts to improve the results coming from them or to widen their scope could push them over the edge. We certainly feel that this would be the case if one were to attempt to treat aqueous solutions using the Percus-Yevick equation.

Before proceeding to a detailed description of our model in the next section, we wish to review the general considerations which led us to it. For the time being, we will restrict our attention to purely classical statistical mechanics.

Ideally, any problem in equilibrium classical statistical mechanics for which the appropriate intermolecular potential functions are known could be solved by exactly evaluating the partition function for the whole system. Since this is not a practical proposition, we seek a reasonable approximation to the partition function. In a dense substance such as liquid water, most of the phase space may immediately be dropped from consideration because it corresponds to very high energy repulsive configurations. In the case of water a similar consideration arises from the strength and strong angular dependence of the electrostatic interactions. If the intermolecular forces are strong enough and the system dense enough, the major contributions to the partition function may be expected to arise from a number of regions of phase space corresponding to maximally stable configurations with fairly "narrow" interconnections between them, as is the case in crystals. These phase space regions may be identified with certain "microstructures" in which the system is at a relatively low potential energy.

Indications for such a state of affairs include solid-like spectral properties, as is the case in water, and a sizable disparity between the periods of intermolecular vibration and the time required for molecular reorientation and translations. Eisenberg and Kauzmann estimate this ratio to be in the order of 50 in water on the basis of diffusion and dielectric relaxation data.[14] Further confirmation for the "amorphous solid" view is given by the molecular dynamics results of Rahman and Stillinger[12] who Fourier transformed various calculated momentum autocorrelation functions and found them to consist predominantly of high-frequency components. Also, the slow decay of their calculated dipole moment autocorrelation function argues for the absence of rapid molecular reorientations.

Once it has been ascertained that a "microstructure" description is appropriate and the main features of

---

(10) G. M. Bell, *Proc. Phys. Soc. London (Solid State Phys.)*, **5**, 889 (1972).

(10a) NOTE ADDED IN PROOF: After completion of this paper we were informed of the work of M. Weissmann and L. Blum, *Trans. Faraday Soc.*, 64, 2605 (1968), which we had accidentally missed. The cell model employed in that paper differs considerably from the one we use.

(11) A. Ben-Naim and F. H. Stillinger, Jr., in "Water and Aqueous Solutions, Structure, Thermodynamics, and Transport Processes," R. A. Horne, Ed., Wiley, New York, N. Y., 1972.

(12) A. Rahman and F. H. Stillinger, Jr., *J. Chem. Phys.*, **55**, 3336 (1971); F. H. Stillinger, Jr., and A. Rahman, submitted for publication.

(13) A. Ben-Naim, *J. Chem. Phys.*, **52**, 5531 (1970).

(14) D. Eisenberg and W. Kauzmann, "The Structure and Properties of Water," Oxford University Press, London, 1969, pp 152-153. The present authors strongly recommend this book as the most comprehensive review on the subject of water to date.

the microstructure have been identified, the problem quite naturally separates into configurational and vibrational components. In the case of water, this separation makes it possible to use the much simpler techniques appropriate for an amorphous solid. In the present work we begin by identifying the major local structural features and developing a description of the liquid in terms of them. The latter task includes deriving appropriate expressions for the configurational energy and entropy. Next we evaluate the strictly local properties by means of lattice dynamics and the cell theory of fluids. Then we find expressions for the various quantities which we wish to calculate in terms of the concentrations of the various local structures. Finally, we find the equilibrium concentrations of the local structures by minimizing the expression for the free energy in terms of them.

Of course, the exact form and degree of definition of the local structures chosen must be such that the course of calculation outlined above is tractable. An idealized assumed form for the local structures does not, however, inevitably lead to large errors in the calculated properties. If we accept the hydrogen bond as being a well-defined structural entity, it is possible to speak of the topology of the microstructure of water which is determined by the hydrogen-bonding and hard sphere-like packing properties of the molecules. As long as the topological characteristics of the "microstructure" are preserved in the idealized local structures, the calculated configurational entropy, which is probably the most structure-sensitive property of a substance, will not be greatly worsened by the idealization.

There is some dispute as to whether hydrogen bonds in liquid water may actually be considered to be discrete structural entities. The strongest evidence we know of against the discrete hydrogen bond view is the observation made by Rahman and Stillinger that the various typical molecular configurations which their molecular dynamics calculation generated do not appear to possess discrete hydrogen bonding.[12] We wish to point out, however, that the hydrogen bond concept that is used in our work includes the role of vibrational motion. We define it by stating: A hydrogen bond between a hydrogen atom of one molecule and a free electron pair (or negative point charge if the BNS potential is employed) of another molecule exists (or has existed) when the two have been in proximity for several periods of intermolecular vibration and the local minimum of their potential energy of interaction has been bracketed in all three relevant relative coordinates. This definition is restrictive enough to ensure the separability of the configurational and vibrational effects which our treatment is based on. The presence of such entities cannot be detected from the purely static pictures of molecular configurations upon which Rahman and Stillinger base their conclusion. We also wish to point out that their molecular dynamics results predict about one too many nearest neighbors, which probably tends to hide the presence of any discrete hydrogen bonds. Furthermore, the as yet incompletely known form of the potential function and the disregard of quantum effects in their calculation may well be sufficiently large to misrepresent the local structure to this degree.

In our work we use the BNS potential increased by 6% as suggested by Rahman and Stillinger.[12] (No further mention will be made of this modification.) The major difficulties we encountered were connected with the genesis of the BNS potential as an effective potential for classical calculations and the fact that the "BNS hydrogen bond's" curvature in the librational degrees of freedom leads to serious disagreement with spectroscopic data. When these deficiencies are properly accounted for, however, our model leads to some gratifyingly good results. The three major thermodynamic functions are in error by about 10% in the direction of predicted lower stability and the heat capacity is too low by about 25%. These errors are for the most part consistent with the known defects of the cell theory. Insofar as it has been possible to determine, our model does not conflict with the experimentally observed radial distribution function. The calculated values of the density are in excellent agreement with experiment, yielding a maximum error of only 1.4%. We have also generated an approximation to the translational frequency spectrum and obtained excellent agreement with the experimentally observed positions of the two major peaks in this region. Our calculations indicate that the number of hydrogen bonds per molecule is about 1.35 throughout the normal liquid range and that, surprisingly, there is probably very little structural change with temperature other than simple solid-like thermal expansion. We have also carried through all of the calculations without correcting the curvature error for comparison.

We present a general description of our model in section II and go into greater detail in sections III–VIII. In section IX we describe our method of approximating the translational frequency spectrum. Section X deals with the sources of some of the experimental data with which we compare our results. Section XI contains a detailed presentation and analysis of our results and in section XII we discuss some possibilities for future investigation.

Appendix I contains the detailed discussion of Bell's model alluded to previously. It should be read after section II. Appendix II contains a general discussion of lattice statistical methods which we decided to include for the benefit of those of our readers who are not chemical or solid state physicists. In it we attempt to describe a paradigm for use in treating problems of the configurational entropy sort. It should be read before section III. We discuss computational details in Appendix III.

Before proceeding further, the reader is forewarned that sections VIII through X, a large part of section V, and Appendix III deal largely with technical considerations and thus are peripheral to undertanding the major content and results of our work.

Throughout this paper we take the liberty of neglecting the distinctions between internal energy and enthalpy when dealing with condensed phases.

II. A General Description of the Model

We contend that much of the information contained in a full binary distribution function is redundant because most of the structure evident beyond the first minimum is probably due to indirect correlations. Since indirect correlations are ultimately composed of direct correlations, it stands to reason that the whole

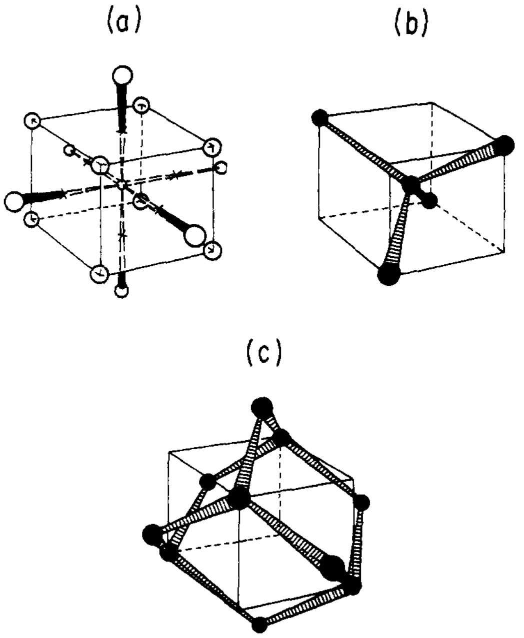

Figure 1. Geometric characteristics of the present model: (a) space lattice section employed; (b) first molecular framework; (c) second molecular framework. Heavy lines indicate hydrogen bonds.

radial distribution function may be expressed in terms of the direct correlation portion, *i.e.*, the first maximum region.

In determining the most likely arrangements of molecules within this region, we focus our attention upon the short range repulsive forces, the hydrogen-bonding geometry, the high density of liquid water, and its highly negative internal energy.

The heat of fusion of ice is only about one-eighth of the internal energy change accompanying sublimation. This indicates that many of the hydrogen bonds must remain intact in water at the melting point and that the potential energy increase associated with those that are broken is probably partially compensated for by strong electrostatic interactions between non-hydrogen-bonded near neighbors. This, coupled with the liquid-solid density difference, indicates that the structure of water has a large degree of hydrogen bonding along with many near non-hydrogen-bonded neighbor pairs. We feel that the most plausible structure reflecting all of these properties is one containing interpenetrating partially hydrogen-bonded networks which are locally independent in respect to hydrogen bonding. We include the assumption of local independence because this property most probably stabilizes the structure both energetically and entropically. The energetic stabilization comes about because local independence allows configurations with less distortion from ideal hydrogen-bonding angles than otherwise. The entropic stabilization is due to the possibility of more efficient packing and, hence, more possible configurations at a given density.

Some preliminary calculations led us to the conclusion that the most likely structure answering to the above description is a partially occupied ice VIII lattice[15] which consists of two interpenetrating diamond-like lattices.[16] The distribution of molecules in the fully occupied ice VIII lattice is of the body-centered cubic type.

The fundamental lattice section (henceforth referred to as the "basic cell") which we chose to employ is a regular octahedron in the body-centered cubic ice VIII lattice which contains a central site and its fourteen nearest and next-nearest neighbor sites. In full ice VIII occupancy it contains a central molecule and its four hydrogen-bonded nearest neighbors from one diamond-like sublattice and ten molecules from the other sublattice, four of which are first neighbors and six second neighbors of the central molecule (see Figure 1). We vary the dimensions of the basic cell linearly with temperature so that the nearest-neighbor distance changes from 2.82 Å at melting to 2.88 Å at boiling as indicated by the results of Narten, *et al.*[2] Such a semiempirical *ansatz* about the intermolecular distances seems necessary because our treatment is essentially of the quasiharmonic variety. It is, therefore, unsuitable for studying this type of thermal expansion which is a result of anharmonicity in the intermolecular potential. We feel that this choice for the fundamental region size is satisfactory because it includes all of the central molecule's neighbors within 4 Å, which is all that we believe necessary, and yet the number of configurations possible is small enough that the model is numerically tractable.

The local structures which our model deals with are the various possible distributions of the noncentral molecules over the sites of such basic cells, one of which is considered to be centered on each molecule in the liquid. (Henceforth we refer to these distributions as "basic cell configurations.") As a matter of convenience, we have restricted the range of basic cell configurations which we deal with to those which have two, three, or four hydrogen-bonded neighbors. The existence of non-hydrogen-bonded molecules is, of course, precluded by Stevenson's spectroscopic data cited earlier.[8] Stevenson also presents a convincing spectroscopic argument that there cannot be a significant concentration of singly hydrogen-bonded molecules which are hydrogen bonded through the oxygen. We contend that this proscription can be extended to singly hydrogen-bonded molecules which are bonded through a proton since there are no grounds for supposing that the latter species is thermodynamically any more stable than the former. It is quite possible that our restrictions upon the numbers of first and second neighbors are unrealistic. It will, however, later become apparent that the exact arrangement of non-hydrogen-bonded neighbor pairs does not directly enter into the calculation of any of the thermodynamic quantities. The only significant properties of such pairs are their total concentrations, which can surely be approximated by a moderately restricted assortment of basic cell configurations such as that employed in this work. We computer-generated all possible basic cell configurations satisfying these constraints and found that there are 374. The basic variables of our model are the concentrations (or "probabilities") of each of them.

In defining our basic cell configurations, we have not

---

(15) B. Kamb and B. L. Davis, *Proc. Nat. Acad. Sci. U. S.*, **52**, 1433 (1964).

(16) We also experimented with structures resembling the two close packed lattice types but found that their agreement with both structural and thermodynamic data was much worse.

distinguished between the hydrogen atoms and the free electron pairs on the molecules in them. To have done so would have led to an intractably large number of basic cell configurations. The effects of the possibility that various sets of molecular orientations are consistent with a given basic cell configuration are dealt with separately.

We feel that a major shortcoming of our model is that the results it generates are difficult to compare with the experimentally determined radial distribution function.[2a] For example, the second major peak which appears in the experimental radial distribution function is almost certainly due to hydrogen bonded second neighbors and, therefore, cannot be accurately studied on the basis of our model because of the small basic cell size employed. Also, the intermolecular distances in our basic cell configurations are precisely defined while the distances between the various types of neighbors in the real substance vary over a certain range. The degree of variation strongly influences the appearance of a given neighbor pair type's contribution to the radial distribution function. If the spread is small, reflecting strong intermolecular correlations, the contribution is a high, narrow peak. If it is large, reflecting weaker correlations, the contribution is a broad low peak, or possibly even a flat region. Since our idealized cell structure takes no account of such variation we can only make a crude comparison and come to conclusions the contents of which are no stronger than "does not disagree."

We will now attempt such a rough analysis and tentatively identify our first and second neighbor pairs with the main components of the short range portion of the experimental radial distribution function. The main feature in this region is the large first maximum. This we associate with the first neighbors, both hydrogen bonded and nonhydrogen bonded. The sharpness of this peak is to be expected because the short range forces, both electrostatic and dispersion, are very strong. (A quick glance at the radial distribution of liquid argon will quickly convince the reader that dispersion forces alone are quite strong enough to produce a sharp first maximum.)

Since these forces decrease rapidly with distance, it is reasonable to expect the non-hydrogen-bonded second neighbor contribution to be of the diffuse variety. Using this assumption, we identify the minimum and the minor peak between the first and second major peaks as the second neighbor contribution. In making this seemingly illogical identification of a minimum with a specific contribution, we use the observation that the bottom of this minimum is actually at about the bulk density. Therefore, we feel that it is valid if it is assumed that the correlations reflected by this region are weak. (There is some dispute as to whether the minor peak is real rather than merely a numerical artifact. The latter conclusion would not, however, invalidate our argument which is based only upon the average density in the region between the two major peaks.)

There are two other possible objections to this last identification which must be answered. First, the minor peak referred to is at about 3.6 Å rather than at about 3.3 Å as is to be expected from our assumed basic cell structure. To this we reply that the disagreement is only about 10%, certainly not overridingly significant at the level of approximation used throughout this work, especially if the possibility of an assymmetric distribution of second neighbor distances is granted. Second, the second hydrogen bonded neighbor maximum is rather pronounced even though it lies at an even greater distance. This, however, poses no contradiction since second hydrogen-bonded neighbors correlate through strong hydrogen-bonding interactions, even if indirectly.

An objection to structures of the type employed in this work which is related to the above discussion has been raised by Rahman and Stillinger on the basis of their molecular dynamics results.[12] They present a decomposition of their calculated radial distribution function in solid angles in the coordinate system of the reference molecule. Their results indicate a fairly low density of neighbors in the regions in which our structure's non-hydrogen-bonded nearest neighbors should lie. This objection is, we believe, invalid because their purely angular decomposition of space around the reference molecule is rather different from our decomposition into three dimensional neighbor positions (or volumes).

Our expression for the configurational entropy is based upon that for a lattice gas of molecules and hydrogen bonds (between nearest neighbor pairs of molecules) in a body-centered cubic lattice. The correlations of molecules and hydrogen bonds are dealt with by constraining their distribution to be consistent with the tetrahedral hydrogen bonding geometry and our previously introduced constraint that each molecule be hydrogen bonded to at least two others. This is discussed in section III.

There is also a "structural" entropy component which corresponds to the residual entropy of ice. We calculate this contribution using a simple extension of Pauling's calculation for ice.[17] (Henceforth we refer to this term as the "orientational entropy.") This is discussed in greater detail in section IV.

The above terms relate to the statistics of "linking up" the molecules to form a microstructure as discussed above. The remaining terms generate the thermodynamic properties of any such given microstructure.

The major remaining term is the "cell model term." The cells dealt with here are, of course, just our basic cells with the various arrangements of molecules in them. In order to reduce the volume of calculations to tractability, we treat the relatively small effects of non-hydrogen-bonded neighbors as perturbations using the appropriate classical expressions. We deal with effects of hydrogen-bonded neighbors directly by using the quantum mechanical harmonic oscillator expressions and frequencies calculated by means of the quasi-harmonic approximation. The translational and librational motions are dealt with separately. This formalism is necessary in order to be able to deal with quantum effects and to allow an insufficiency of the BNS potential which we discovered to be corrected. The effects of molecules not within a given molecule's basic cell are calculated using a dielectric continuum approximation. We account for the effects of temperature dependent variations of the intramolecular zero-

(17) Reference 14, pp 74–76.

point energies semiempirically. These matters are discussed in section V.

An apparent contradiction arises from including the distribution of non-hydrogen-bonded neighbors in the definition of the basic cell configurations but nearly ignoring them in the treatment of the configurational entropy. We include them explicitly in order to be able to treat their fairly important energetic effects more accurately by forcing their distribution in each cell to be consistent with our assumed structure. The same purpose is served by the various self-consistency constraints discussed in sections III and VII and Appendix II.

The basic defect of the cell model is its inability to deal with the effects of the correlation of motions of neighboring molecules. We initially attempted to treat these effects by the method with which we generated our approximation to the translational frequency spectrum. It turned out, however, that this approach was insufficiently accurate and too unwieldy to bother with in light of the fairly small magnitude of these effects indicated by our preliminary results. Therefore, we decided to use a simple approximation technique, which is described in section VI, to deal with them. A general discussion of correlation (or "communal") effects is included in sections VI and XI.

With the exception of the dielectric continuum and intramolecular zero-point energy contributions to the cell term, all of these quantities are functions of the basic cell probabilities. We find the equilibrium basic cell probabilities at each temperature by expressing the free energy in terms of the basic cell probabilities and minimizing it by the method of steepest descent. The minimization is performed under suitable constraints which ensure that the resulting calculated structure is self-consistent and obeys our various initial assumptions. This is discussed in section VII and Appendix III.

The heat capacity is determined by differentiating the calculated values of the enthalpy as discussed in section VIII, where we also describe our approximation to the Kirkwood correlation factor.[18]

We have not attempted to describe phase transitions using our model. We have, however, evaluated the thermodynamic properties of ice at the melting point, using techniques which are closely related to our treatment of water, in order to test our model by calculating the thermodynamic properties of this simpler related system. We have combined these results with our water results to evaluate the thermodynamic changes associated with melting. This is discussed in sections V and XI.

III. Configurational Entropy

Since no plausible general treatment of the configurational entropy of the noncrystalline substances problem has yet been proposed, we were forced to attempt to extend the lattice statistical methods discussed in Appendix II. The variant most suitable for our purposes is the theory of the dense lattice gas. We readily concede that the lattice gas approach has not been very good at all in dealing with the properties of simple liquids and gases. However, we are quite convinced that it is much more suitable in the case of water than in the case of simpler substances. The reason that the lattice gas approach fails in treating, say, liquid argon, is that it is based on the assumption that the system's phase may realistically be divided into distinct microstructures in the sense of the discussion in section I. These postulated microstructures are, of course, the various allowed configurations of the lattice gas. In that discussion, however, we concluded that just such a decomposition of phase space is a very reasonable assumption in the case of liquid water. We also concluded that, thermodynamically speaking, local structure is all. Therefore, we conclude that a lattice gas formalism is indeed appropriate.

Our original inclination was to treat the configurational entropy problem by means of the Ising model treatment developed by Hijmans and De Boer[19] using our basic cell as the basic figure. However, the full number of configurations possible in a basic figure of 15 sites, about 1000 in this case, made this technique too unwieldy. We believe that using the Hijmans–De Boer technique to calculate the configurational entropy could conceivably be more accurate than our treatment. Also, the lack of a liquid–solid transition in Bell's results[10] might be rectified by going over to a larger basic figure. As they stand, Bell's results indicate that a straight lattice statistical calculation of the configurational entropy using a fairly small basic figure might be sufficient for studies not involving phase transitions. This view is substantiated by the work of Kruseman-Aretz,[20] which indicates that calculations which employ fairly small basic figures (on the order of three or four sites) often generate results which are not significantly different from those obtained using larger basic figures. He also derives the Hijmans–De Boer technique and other related techniques (such as the Guggenheim–McGlashan[21] technique employed by Bell) as first order approximations to a formally exact theory.

The technique we use relies on the hypothesis of well-defined microstructure less than the various standard lattice statistical treatments. More specifically, all of the standard lattice statistical treatments of the lattice gas treat molecules and vacancies symmetrically; *i.e.*, vacancies are treated as a second component in a disordered alloy with a fixed size and undeformable shape. This implies that a vacancy in the structure causes correlations of molecules and other vacancies as strongly as does a molecule. This, of course, is unreasonable because a given volume of empty space is anything but undeformable. In our derivation, only the very first step employs a postulation of this symmetry. All other invocations of the microstructure hypothesis involve the structure surrounding molecules but not that surrounding vacancies.

The first factor in our estimate of the number of configurations is that corresponding to a random distribution of molecules over a lattice

$$\frac{(N/\rho)!}{N!\{N[(1/\rho)-1]\}!}$$

where N is the number of molecules.

---

[18] J. G. Kirkwood, *J. Chem. Phys.*, 7, 911 (1939).

[19] J. Hijmans and J. De Boer, *Physica (Utrecht)*, 21, 471, 485, 499 (1955).

[20] F. E. J. Kruseman-Aretz, Thesis, Universiteit van Amsterdam, 1962.

[21] E. A. Guggenheim and M. C. McGlashan, *Proc. Roy. Soc., Ser. A*, 206, 335 (1951).

In a strictly construed lattice gas the quantity ρ is simply the fraction of sites which are occupied. Since we have chosen to interpret the lattice gas rather liberally, we must find a somewhat different working definition. For the sake of simplicity, we have chosen to define it as $^1/_{14}$th of the average number of first and second neighbors. This seems to be the most reasonable choice when dealing with a structure defined in terms of basic cells of the type which we employ. The possibility of minor error in this interpretation does not concern us, because our preliminary calculations indicate that alterations of the magnitude possible and reasonable here do no more than add a relatively small, constant term to the various calculated quantities. In calculating the macroscopic density, we identify ρ with the fraction of sites filled in a perfect lattice. Again, the possibility of minor error is not important because the expression for macroscopic density is not used in any other calculation.

We do not assign any value to ρ because it is a function of the set of basic cell probabilities $\{P_i\}$, whose equilibrium values are determined by minimizing the free energy in terms of them at each temperature considered. The same applies to the other quantities which are introduced in this section.

In the minimization of the free energy we introduce the constraint that the ratio of the number of first neighbor pairs to the number of second neighbor pairs be 4/3, the same as the ratio of the number of first neighbor sites to second neighbor sites in our basic cell. This requirement is consistent with our procedure of starting off with a random distribution of molecules and hydrogen bonds, and is explicitly introduced to eliminate the possibility of obtaining an equilibrium solution contradicting this initial assumption. (See Appendix II in regard to the initial random distribution.) The introduction of the factors which compensate for the initial over-counting does not invalidate this requirement.

From this point on we do not use the microstructure hypothesis except when dealing with the immediate surroundings of molecules.

Using our assumption of body-centered cubic local structure and a direct corollary of the above discussion, we immediately conclude that the number of pairs of molecules in the random spatial distribution of molecules introduced above which may be hydrogen bonded, *i.e.*, first neighbor pairs, is $4N\rho$. Of course, not all first neighbor pairs are actually hydrogen bonded. Defining

$$P_h = \text{number of hydrogen bonds}/N$$

we may write the number of ways of distributing the NP_h hydrogen bonds over the $4N$ pairs which may be hydrogen bonded as

$$\frac{(4N\rho)!}{(NP_h)![N(4\rho - P_h)]!}$$

The product of these two factors is the initial estimate of the number of configurations which must now be modified to compensate for overcounting. The overcounting occurs because most of the initial random distributions contain molecules which have too few or too many hydrogen bonds or incorrect angles between the hydrogen bonds.

For the time being, we specify the number of hydrogen bonds each molecule engages in and which of the nearest neighbor sites in its basic cell are occupied by the hydrogen bonded neighbors and derive the factor corresponding to the $(^1/_{16})^N$ in the Pauling calculation.

Any given first neighbor position in any given basic cell has the probability $P_h/4$ of being occupied by a hydrogen-bonded neighbor in the above derived random ensemble of hydrogen bonds and molecules. Therefore, the probability of any molecule having i hydrogen-bonded neighbors in i specific first neighbor positions in its basic cell in such an ensemble is given by $(P_h/4)^i$. This factor occurs a total of $2NP_h$ times.

Another restriction factor is necessary to disallow the occurrence of non-hydrogen-bonded neighbors in the hydrogen-bonded neighbor positions of each cell. (This restriction is necessary to ensure consistency with our postulated structure.) The corresponding factor for each hydrogen-bonded neighbor position not occupied by a hydrogen-bonded neighbor is

$$(1 - \rho)/[1 - (P_h/4)]$$

where the denominator arises because we are dealing only with those first neighbor positions which have already been determined not to contain hydrogen-bonded neighbors. This factor occurs $N(4 - 2P_h)$ times.

Both of these considerations together introduce the factor

$$\left(\frac{P_h}{4}\right)^{2NP_h}\left(\frac{1 - \rho}{1 - (P_h/4)}\right)^{N(4-2P_h)}$$

At first glance it would seem that this is insufficient because it contains no factors of the type $[1 - (P_h/4)]$ to ensure that any given molecule is not hydrogen bonded in excess of the hydrogen-bonding arrangement assumed for it. This, however, is not a valid objection, since the number of factors $P_h/4$ appearing in the above expressions is $2NP_h$, *i.e.*, two per hydrogen bond. This means that the above factor specifies the position of each and every hydrogen bond by "fixing" both of its ends. Therefore, there are no excess hydrogen bonds left which may be incorrectly placed.

All of the above factors together yield the number of configurations in the random ensemble in which each molecule has a specific number and arrangement of hydrogen bonded neighbors. Of course, there are many such sets of numbers and arrangements of neighbors possible. It remains to derive an estimate of that number.

We define P_2, P_3, and P_4 as the fractions of molecules with two, three, and four hydrogen-bonded neighbors, respectively. The number of ways of assigning the N molecules to these categories is simply

$$N!/(NP_2)!(NP_3)!(NP_4)!$$

A molecule with a specified number of hydrogen-bonded neighbors may have them distributed over the nearest neighbor positions in its basic cell in one of several ways. The number of possible arrangements of two, three, or four hydrogen bonded neighbors is twelve, eight, or two, respectively. (The reader will readily convince himself of this if he draws a cube and marks off four alternate vertices to form a tetrahedron.)

The total contribution to the multiplicity of configurations from this source is

$$12^{NP_2}\,8^{NP_3}\,2^{NP_4}$$

Both of these factors together are the equivalent of the factor 6^N in the Pauling calculation.

Multiplying and rearranging, we write the total number of configurations as in eq 1.

$$\frac{(N/\rho)!}{N!\left[N\left(\frac{1}{\rho}-1\right)\right]!}\,\frac{(4N\rho)!}{(NP_{\mathrm{h}})![N(4\rho-P_{\mathrm{h}})]!}\times \frac{N!}{(NP_2)!(NP_3)!(NP_4)!}\times \left[12\left(\frac{P_{\mathrm{h}}}{4}\right)^2\right]^{NP_2}\left[8\left(\frac{P_{\mathrm{h}}}{4}\right)^3\right]^{NP_3}\times \left[2\left(\frac{P_{\mathrm{h}}}{4}\right)^4\right]^{NP_4}\left(\frac{1-\rho}{1-(P_{\mathrm{h}}/4)}\right)^{N(4-2P_{\mathrm{h}})} \quad (1)$$

IV. Orientational Entropy

Our expression for the number of orientational configurations available to an incompletely hydrogen-bonded set of water molecules is merely an extension of the Pauling calculation for ice.[17] The number of ways in which the NP_{h} protons involved in hydrogen bonding may be placed on molecules, disregarding for the moment the requirement of two protons per molecule, is $2^{NP_{\mathrm{h}}}$. The number of ways of distributing the unaccounted for $N(2-P_{\mathrm{h}})$ protons and $N(2-P_{\mathrm{h}})$ unbonded electron pairs over $2N(2-P_{\mathrm{h}})$ positions is

$$[2N(2-P_{\mathrm{h}})]!/\{[N(2-P_{\mathrm{h}})]!\}^2$$

Of the 16 possible occupancies of the four positions of an individual molecule with protons and free electron pairs, only six give the neutral molecule. Therefore, we introduce a restriction factor $^6/_{16}$ for each molecule and obtain

$$Q^{(0)} = 2^{NP_{\mathrm{h}}}\frac{2[N(2-P_{\mathrm{h}})]!}{\{[N(2-P_{\mathrm{h}})]\}^2}\left(\frac{6}{16}\right)^N$$

$$S^{(0)}/k = N(\ln 6 - P_{\mathrm{h}}\ln 2)$$

For $P_{\mathrm{h}} = 0$, *i.e.*, the dilute lattice gas with specified potential local hydrogen bonding geometry

$$S^{(0)}/k = N\ln 6$$

which is the rigorously correct result. For $P_{\mathrm{h}} = 2$, *i.e.*, ice

$$S^{(0)}/k = N\ln {}^3/_2$$

which is just the Pauling result.

V. Cell Model Contributions and the Insufficiencies of the BNS Potential

In the straightforward cell model treatment the expressions for the free energy, internal energy, and entropy are

$$A = N\sum_i P_i\left(-\frac{1}{\beta}\ln Q_i - E_{ti}\right)$$

$$E = N\sum_i P_i\left(-\frac{1}{Q_i}\frac{\partial Q_i}{\partial\beta} - E_{ti}\right)$$

$$S = \frac{N}{T}\sum_i P_i\left(-\frac{1}{Q_i}\frac{\partial Q_i}{\partial\beta} + \frac{1}{\beta}\ln Q_i\right)$$

where E_{ti} is one-half the potential energy of interaction of the central molecule in cell type i with its neighbors in the cell when all the molecules are fixed in their lattice positions and orientations, and Q_i is the phase integral of the central molecule moving about the cell with its neighbors fixed at their lattice positions.

Our initial inclination was simply to use the above expressions. However, we quickly concluded that they are neither practical nor appropriate for our purposes. First of all, treating each set of molecular orientations consistent with any given configuration separately would require an impossibly large number of six dimensional integrals to be evaluated numerically. This problem we quickly resolved by deciding to treat the effects of non-hydrogen-bonded neighbors as perturbations. Second, it appears that the interactions of molecules with other molecules too far away to be included in their basic cells are too large to neglect. We chose to approximate these effects by employing a dielectric continuum approximation. Third, the variation of the hydrogen bonding potential in the rotational degrees of freedom is so rapid that classical phase integrals are seriously in error in the temperature range studied. This we remedied by replacing the actual BNS potential with a quadratic approximation and employing the thermodynamic functions of the quantum mechanical harmonic oscillator. This also required some reinterpretation of the physical significance of the BNS potential which was, after all, constructed using purely classical criteria. Finally, we discovered that the BNS potential predicts librational frequencies much higher than those experimentally observed. Fortunately, our use of the quantum mechanical harmonic oscillator formalism allows this to be compensated for by means of a simple rescaling of the calculated librational frequencies. We now proceed to discuss these matters in greater detail.

In dealing with the librational frequencies we found it much easier to focus our attention on the spectral properties of ice rather than of liquid water, since the simpler and much better known structure of ice enormously simplifies interpretation and comparison.

The magnitude of the librational frequencies (which in the case of ice at the melting point appear to be distributed over the range 400–1000 cm^{-1} corresponding to about $2\text{-}5kT^{22-25}$) makes the use of classical phase integrals impermissible. As we mentioned before, we originally attempted to employ classical phase integrals alone in our treatment of the basic cell properties. The entropy of ice calculated in this way, which is presented in Table I, indicates just how inappropriate a procedure this was. When dealing with such strong intermolecular forces the classical phase integrals do not differ much in value from phase integrals over sets of classical harmonic oscillators. The magnitude of the error is due to the fact that the entropy of a quantum mechanical harmonic oscillator, *i.e.*, a "real" harmonic oscillator, goes to zero at high values of $h\nu/kT$ while that of a classical harmonic oscillator decreases without limit to spurious negative values.

---

[22] R. Zimmerman and G. C. Pimental in *Advan. Mol. Spectrosc., Proc. Int. Meet., 4th*, 762 (1962).

[23] A. J. Leadbetter, *Proc. Roy. Soc., Ser. A*, **287**, 403 (1965).

[24] H. Prask, H. Boutin, and S. Yip, *J. Chem. Phys.*, **48**, 3367 (1968).

[25] R. E. Shawyer and P. Dean, *Discuss. Faraday Soc.*, **No. 48**, 102 (1969).

Table I. Thermodynamic Properties of Ice

	G, cal	H, cal	S, eu	C_p, cal/(mol °K)
Experimental[a]	−12,738	−10,030	9.91	9.25
Classical BNS	−11,469[b]	−10,090[b]	4.65	12.08[b,c]
Quantum BNS	−10,975	−9,078	7.03	7.69[b]
Curvature rescaled quantum BNS	−12,454	−10,335	7.69	7.77[b]

[a] Reference 27. [b] Does not include intramolecular zero-point energy change contributions. [c] Calculated by using analytic differentiation of internal energy rather than numerical as in water calculations. Probably unreliable due to numerical error.

We separately calculate the translational and librational Einstein frequencies[26] for cells with two, three, and four hydrogen-bonded neighbors. We calculate these frequencies separately for each temperature studied since, as previously discussed, we vary the cell dimensions linearly with temperature. In deriving the quadratic approximation to the translational potential we average (with appropriate weighting) the BNS potential between hydrogen bonded neighbors over the two relative orientations consistent with hydrogen bonding. This approximation probably does not affect the results appreciably, since our preliminary results indicate that the hydrogen bonding potentials of the two configurations do not differ by more than about 10% throughout the important part of the potential well. In the case of the librational contributions, the derived thermodynamic quantities must be averaged over all six possible orientations of the central molecule, since each of the six corresponds to a different set of principal inertial axes. In calculating the inertia tensors we use the gas phase bond angle but increase the bond lengths from 0.96 to 1 Å, which we believe to be a reasonable estimate of the elongation due to hydrogen bonding in condensed phases. The same logic was used by Ben-Naim and Stillinger[11] in setting the charge-center distance in their potential at 1 Å. The librational potential is approximated in terms of the sum of squares of the angular deviations of each of the charges involved in hydrogen bonding from the ideal hydrogen-bonding directions. The hydrogen-bonded neighbor molecules are, of course, taken to be fixed in their ideal lattice positions and orientations. This form for the quadratic approximation of the librational potential contains only one parameter. This, along with the separation of translational and librational motions allows the correction of the BNS potential discussed below to be introduced through a simple rescaling of the librational frequencies.

The main difficulty involved in going over to a quantum mechanical formalism is that the BNS potential was devised by fitting it to experimental data using purely classical techniques. In other words, it is an effective pair potential for use in classical mechanical formulations. Its efficacy in this respect is evidenced by the excellent agreement with experiment of the classical calculation of the enthalpy of ice (see Table I)[27] and of the molecular dynamics calculation of the potential energy of water by Rahman and Stillinger.[12] This, of course, does not at all mean that it is an appropriate effective potential for quantum mechanical formulations. In addition to the above discussed entropy effect, the appearance of intermolecular vibrational zero-point energies in a quantum mechanical formulation makes the BNS potential in its original form even more unsuitable in this application. There is little choice but to modify it in such a way that it can serve as an effective potential in quantum mechanical calculations.

We have found that the simplest such modification which appears to be both effective and reasonable is to redefine it as the effective potential which encompasses all potential energy effects except those arising from the variation of intramolecular zero point energies. We have, therefore, added on these effects separately in our quantum mechanical formulation. As is evident from Table I, this redefinition just about cancels out the original BNS potential's intermolecular zero point energy related insufficiency. We also argue that it is probably unsound to attempt to include the effects of intramolecular zero point energy changes in a temperature independent effective potential such as the BNS. They are almost certainly strongly temperature dependent in that they are a function of the average instantaneous (on the time scale of molecular motions) molecular environments and, as such, can be expected to be extremely sensitive to the variations of cell structure and intermolecular vibrational amplitudes with temperature. This argument also makes our result of little structural variation (to be discussed in section XII) with temperature consistent with the experimentally observed considerable variation of the positions of intramolecular vibrational spectral peaks. Another way of describing this effect is to say that hydrogen bonds grow weaker with increasing temperature and that this weakening results in decreased hydrogen-bonding effects on the intramolecular vibrations. We decided to use a semiempirical treatment of these effects rather than attempt a more fundamental treatment because the conceptual machinery necessary for a fundamental treatment is not yet available.

Adhering to the above discussion, we add on a term representing the potential energy contribution of the intramolecular zero-point energy decrease on transition from vapor to condensed phases. In the case of ice at the melting point Whalley[28] has calculated this contribution to be −1.4 kcal/mol and, therefore, much too large to be ignored. It is, unfortunately, difficult to calculate for liquid water because the O–H stretching region of the ir spectrum is rather complicated and, furthermore, the available experimental data are inadequate.

We decided to employ the vapor-phase frequencies quoted by Eisenberg and Kauzmann[29] and the liquid phase frequencies given by Bayly, *et al.*,[30] considering the stretching mode distribution to be centered at the average of the frequencies of the main absorption peaks in the corresponding spectral region. To estimate the temperature variations of these frequencies we use the

---

(26) The analysis of the various molecular dynamics calculated momentum autocorrelation functions carried out by Rahman and Stillinger indicates little coupling between translational and librational motions.[12] This conclusion supports our separation of the two.

(27) Calculated from the liquid water data obtained as described in section X and the thermodynamic constants of melting quoted by D. Eisenberg and W. Kauzmann, ref 14, p 100.

(28) E. Whalley, *Trans. Faraday Soc.*, **53**, 1578 (1957).

(29) D. Eisenberg and W. Kauzmann, ref 14, p 8.

(30) J. G. Bayly, V. B. Kartha, and W. M. Stevens, *Infrared Phys.*, 3, 211 (1963).

results of Fox and Martin.[31] The contribution at the melting point calculated in this way is −900 cal/mol, in good agreement with the above-mentioned result for ice.

A rescaling of the librational curvature is necessary because the BNS potential predicts an average Einstein librational frequency of 990 cm^{-1} and a maximum librational Einstein frequency of 1270 cm^{-1} for ice at the melting point, both of which are in gross disagreement with available experimental data. The ir absorption maximum in this region is at 795 cm^{-1}.[22] Leadbetter's detailed analysis of the low temperature heat capacity of ice indicates a distribution centered at about 650 cm^{-1}.[23] Inelastic neutron scattering studies reveal a large peak at about 660 cm^{-1} and possibly another smaller one in the range 800–950 cm^{-1}.[24] The numerical evaluation of the frequency spectrum of ice carried out by Shawyer and Dean indicates a distribution of frequencies over the range 430–730 cm^{-1}.[25] There is no evidence whatsoever for any significant librational spectral density above 1000 cm^{-1}.[32]

We feel that it is difficult to reach any definite conclusions regarding the exact distribution of librational frequencies because of the apparent inconsistencies of the available data and the inevitable complexity of any theoretical analysis of coupled librational motions. However, we may at least attempt to adjust the librational curvature to minimize its conflict with the data presented above. We feel that this purpose is best served by decreasing the curvature so that the average librational Einstein frequency of ice falls in the range 650–700 cm^{-1}. This agrees well with Leadbetter's[23] thermodynamic analysis and moderately well with Shawyer and Dean's[25] calculated spectrum. The two lower Einstein frequencies fall in the range of about 500–650 cm^{-1} and are no more than about 120 cm^{-1} apart. Thus we may tentatively identify the major neutron-scattering peak at 660 cm^{-1} with modes composed mostly of the two lower frequency Einstein modes. The highest frequency Einstein mode falls in the range 850–910 cm^{-1} and thus may be identified with the ir maximum and the possible second peak in the neutron scattering spectrum. We feel that the unpredictable nature of the ir absorption selection rules which may be expected in the case of a disordered molecular crystal invalidates any objections to the relatively low choice for the average frequency on the grounds that the ir peak is at a somewhat higher frequency. We have chosen to rescale the librational curvature by a factor 0.458 in order to set the average frequency equal to 650 cm^{-1} in agreement with Leadbetter's estimate.

In calculating the perturbation terms representing the effects of non-hydrogen-bonded neighbors we use the unmodified BNS potential and classical phase integrals as a matter of convenience. We calculate the perturbation terms separately for each neighbor, for each orientation of the neighbor relative to the central molecule, for each number of hydrogen bonded neighbors and each temperature. To represent the potential between hydrogen bonded pairs we use the average of the BNS potential over the two relative orientations consistent with hydrogen bonding as discussed previously. Finally, we average these terms over all nine relative orientations between non-hydrogen-bonded neighbors possible in the lattice and over the three orientations of the central molecule relative to the given basic cell configuration.

To give an example, each perturbation term for the free energy corresponding to the non-hydrogen-bonded neighbor n is of the form

$$A_{\mathrm{h}}^{n} = \left\langle -\frac{1}{\beta}\ln\left\{\int_{\tau}\exp[-\beta(V_{\mathrm{h}}^{0} + v_{n})]\,\mathrm{d}\tau\right\}\right\rangle + \frac{1}{\beta}\ln\left\{\int_{\tau}\exp[-\beta V_{\mathrm{h}}^{0}]\,\mathrm{d}\tau\right\}$$

where the second quantity in braces is simply Q_i, v_n is the potential energy of interaction of the neighbor n in a specific orientation with the central molecule, and the averaging is over the nine possible relative orientations. The subscript h refers to the number of hydrogen bonds. V_{h} is averaged as mentioned above.

We evaluate these integrals numerically assuming all molecules entering into the potential energy expressions except the central molecule to be fixed at their lattice positions and orientations. The integration is performed over the coordinates of the central molecule and is discussed in greater detail in Appendix III. This expression is summed over the values of n consistent with the given basic cell configuration and then averaged over the three possible orientations of the central molecule relative to the basic cell configuration. The treatment of the various other quantities is strictly analogous.

The degree of approximation in the above development may seem excessive. In truth, we believe that the only possible source of significant error is the assumption that the positions and orientations (in the sense of hydrogen-bonding directions) of the nonhydrogen neighbors are those of an ideal lattice. We discuss this further in section XII. Certain detailed considerations which we will not go into here indicate that the error introduced by the other approximations is insignificant in light of the relatively small magnitude of the perturbation terms.

The form of the potential function makes it apparent that interactions between molecules whose centers are further separated than the 3.3-Å maximum distance for which our cell treatment calculates interactions must be included in the internal energy. A simple model of these interactions is based on the assumption that each molecule interacts with a uniform dielectric medium (with the bulk dielectric constant) which begins just outside of its basic cell. We assume the radius of this spherical cavity to be 4 Å at 300 K, to vary with temperature in the same manner as the cell size, and simply treat the dielectric medium as a conductor. This is valid, since the difference between dielectric–vacuum and conductor–vacuum interface properties in systems of this sort is of the order of $1/\epsilon$, which is negligibly small for water whose ϵ has values in the range 55–88 for the temperatures dealt with here. The conducting medium–molecule interaction is trivially

---

(31) J. J. Fox and A. E. Martin, *Proc. Roy. Soc., Ser. A*, **174**, 234 (1940).

(32) We wish to note that the spectra of the librational momentum autocorrelation functions about the three principal inertial axes calculated by Rahman and Stillinger[12] do not disagree nearly as much with the available data for the liquid as does our calculation of the average BNS librational Einstein frequency with the data for ice. We believe this discrepancy to be due to a cancellation of the curvature error by the various other errors in the molecular dynamics calculation which are discussed in this paper.

soluble and generates a long-range energy contribution of -1.184 kcal/mol at 300 K.[33]

No molecular symmetry effects are considered, since their ultimate source is the rotational partition function of a freely rotating molecule and, therefore, they are irrelevant to our model. We also ignore nuclear spin statistics in both our calculation and the presentation of the experimental data.

VI. Communal Effects

By this we mean the thermodynamic effects associated with the correlation of motions of neighboring molecules which are ignored in the cell theory of fluids. In the language of lattice dynamics, communal effects constitute the difference between the Einstein model and the fully developed Born–Huang[34] treatment.

In light of our general approach, the best way of calculating communal effects would be to calculate the full intermolecular frequency spectrum in the quasiharmonic approximation and subtract out the spectrum of the uncorrelated motions of single molecules. Since this is, at best, an impractical proposition, we immediately seek simplifications. First, we drop the correlation of librational motions with each other and with translational motions. This is reasonable, since the librational portion of the spectrum is at such high frequencies that minor modifications of it are probably not very significant thermodynamically and it probably does not couple very strongly with the considerably lower frequency translational modes.

This reduces the problem to a purely translational calculation. We originally attempted to apply the translational spectrum approximation technique described in section IX to it, but quickly found that the communal effects are so small that the inaccuracy of our approximation made this approach useless. We then calculated the communal effects in a fully occupied ice I (c) (diamond-like) lattice at 283 K and the corresponding lattice dimensions using Smith's treatment for the diamond lattice[35] to give us an estimate of the quantities involved. The energy effect turned out to be negligibly small and the entropy effect turned out to be 0.391 eu. We felt that this magnitude made any attempt at precision pointless and simply chose to approximate the entropy effect as 0.391 eu times the fraction of hydrogen bonds intact at any given temperature. The assumption of proportionality is probably valid, since intermolecular motion correlations may well be expected to be transmitted mainly through the hydrogen bond network.

VII. Minimization of the Free Energy

All properties of the model are easily expressed in terms of the $\{P_i\}$ which must all have values between zero and one and must be normalized to unity. To ensure that they are nonnegative and normalized at all steps of the minimization, we reexpress the free energy in terms of the variables $\{x_i\}$ which are related to the $\{P_i\}$ by

$$P_i = x_i/\sum_j \exp(x_j)$$

and are defined only to within an arbitrary constant term.

Besides the normalization requirement, other constraints must be introduced to ensure that the structure corresponding to any attainable set of values of the $\{x_i\}$ is consistent with our fundamental assumptions and physically meaningful in the sense discussed in Appendix II. First, there is the constraint introduced in section III that the average number of second neighbors be three-fourths of the average number of first neighbors. Second, of the various configurations of the overlap regions of first and second neighbor cells, fourteen may occur in dissimilar orientations in the basic cell in the sense of the discussion in Appendix II. Upon expressing the necessary related constraints in linear algebraic form, we found that two of them were redundant in the linear algebraic sense and thus could be dropped from further consideration.

The free energy expression is too complicated to directly solve the minimization conditions to yield the equilibrium values of the $\{x_i\}$. Therefore, we use the gradient descent technique taking each of the 13 abovementioned conditions into account by simply adding a term of the form $\alpha_j C_j^2$ to the free energy for each one where α_j is a suitable positive constant and C_j is the linear algebraic expression which must equal zero in order to satisfy condition j. Although this approach may seem somewhat inelegant and bulky, we found it to be quite reliable and efficient. We discuss it in greater detail in Appendix III.

VIII. Heat Capacity and Kirkwood Correlation Factor

In computing the heat capacity, we have chosen to simply numerically differentiate the calculated values of the enthalpy. This is actually a more accurate procedure than using the greatest possible degree of analytical differentiation because the latter would involve separate numerical differentiations of all 374 basic cell probabilities and introduce the possibility of cumulative error. We wish to point out that the values of the heat capacity calculated from the enthalpy may be expected to differ from those calculated from the entropy for two reasons. First, we employ an explicitly derived expression for the entropy rather than the temperature derivative of the free energy. (We do this to keep the inaccuracies of our enthalpy expression from affecting the calculated values of the entropy too much and *vice versa*.) Second, even if we did use the free energy derivative as our expression for the entropy and the corresponding expression for the enthalpy, the two expressions for the heat capacity would not agree because of treatment of the constraints in minimizing the free energy.

This problem would not arise if the derivatives of the free energy with respect to $\{P_i\}$ were all zero at equilibrium. In fact, they are equal to minus the derivatives of the constraint terms which we insert into the free energy expression in the minimization step. Fortunately, the two errors together produce a total discrepancy of no more than about 2 cal/(°K mol) throughout the temperature range. We believe the enthalpy derived

---

(33) This problem is trivial because of the symmetry of the positions of the four charges on the central molecule about the center of the spherical cavity. The four necessary image charges form another regular tetrahedron whose four vertices are on the four lines passing through the center of the cavity and the charges of the central molecule. The high symmetry of the two tetrahedrons allows the remainder of the calculation to be reduced to little more than an evaluation of two charge–charge distances.

(34) M. Born and K. Huang, "Dynamical Theory of Crystal Lattices," Oxford University Press, London, 1954.

(35) H. M. J. Smith, *Phil. Trans. Roy. Soc., London*, **241**, 105 (1948).

values to be more fundamental because the enthalpy expression contains the intramolecular zero-point energy and long range interaction effects explicitly. (The values calculated from the enthalpy are the only ones which we present.)

The Kirkwood correlation factor[18,36] appears in the microscopic theory of dielectrics and is generally approximated as

$$g = 1 + \sum_j \langle \cos \gamma_j \rangle$$

where j ranges over the near neighbors of any given molecule and γ_j is the angle between the dipole moment of the central molecule and the dipole moment of molecule j.

In our calculation we take the average over phase, cell type, and orientations, and include those neighbors which appear in the basic cell in the summation to get an expression of the type

$$g = 1 + \sum_i P_i g_i$$

where g_i is the sum of the average cosines of the angles between the dipole moment of the central molecules and that of each of its neighbors in cell type i. We use only classical phase integrals in this calculation.

Since the noncentral molecules are assumed to be rigidly fixed in their positions and orientations in our cell integral evaluations, we may express g_i in terms of the averaged components of the central molecule's dipole moment along three perpendicular axes. We calculated average cosines between the central molecule and its hydrogen bonded neighbors and the effects of the nonhydrogen neighbors upon these quantities in the same way we calculated the various thermodynamic perturbation terms. We also calculated the average cosines between the central molecule and the various non-hydrogen-bonded neighbors separately for each neighbor in each relative orientation and added these quantities on with suitable averaging.

We wish to point out, however, that our operational definition of g is by no means the only possible one. Various other cut-offs of the domain of "near neighbors" and treatments of both direct and indirect long range interaction effects, which we completely ignore, are possible. (For an example of a completely different approach, see ref 12.) For this reason we urge some skepticism on the part of the reader regarding our treatment of this problem.

IX. Translational Frequency Spectrum Estimate

In another publication[37] we describe a technique for generating the even moments of the frequency spectrum of an arbitrary amorphous substance in terms of the intermolecular potential energy function and the concentration of certain types of molecular clusters in it. This publication should be consulted in order to render this section intelligible.

In this case, the obvious choice of lattice vectors to work with are those connecting first and second neighbors. We have derived expressions for the 0th through 10th moments using the approximation discussed at the end of ref 37. These expressions involve contributions from clusters of five or fewer molecules only. Further moments could not be evaluated exactly, because the expressions for them would involve the concentrations of molecular clusters which cannot be fitted into our basic cell. This means, of course, that their concentrations are not expressible in terms of $\{P_i\}$. Also, the number of distinct closed lattice vector sequences which it is possible to construct out of six or more nonzero lattice vectors increases very rapidly with the number of vectors.

Unfortunately, we found that six moments were not sufficient for our purposes. Therefore, we were forced to seek an appropriate approximation to the 12th and 14th moments. Upon examining the relative magnitudes of the contributions to the 10th moment we found that the largest were those of the vector sequences associated with the following clusters of molecules: the single molecule, the hydrogen bonded pair; the hydrogen bonded triplet; and the non-hydrogen-bonded first neighbor pair. Of the remaining vector sequences, we found that the smaller the number of nonzero vectors, the larger the contribution. On this basis we chose to approximate the 12th and 14th moments by the sums of the contributions of all vector sequences associated with the four types of clusters specifically enumerated above and of those sequences constructed by adding null vectors to all other sequences which contribute to the 10th moment. We estimate the relative magnitude of the error introduced into the values of the two higher moments by this approximation to be of the order of 10^{-3}.

We then inserted the equilibrium values of $\{P_i\}$ into these expressions to obtain the actual values of the moments and generated an approximation to the spectrum by fitting a polynomial in the second through ninth powers of the frequency to the values of the moments by means of an inverted Hilbert matrix. One further assumption, that of a certain maximum frequency, is necessary. The most reasonable estimate for this value is the maximum vibrational frequency of a fully hydrogen-bonded ice I(c) network which we obtained as a by-product of our study of communal effects which is described in section VI. This value is 316 cm^{-1}.

We also generated an approximation to the spectrum using values for the moments which include only the contributions to the higher moments from the four major molecular clusters enumerated above. The approximation obtained by this procedure differs only insignificantly from that obtained from the full treatment described above. This result tends to confirm the usual assumption that the hydrogen-bonding network is the over-ridingly important structural characteristic of water.

X. Sources and Preparation of the Experimental Data

The thermodynamic standard states which we employ in this work are as follows: for the enthalpy, the infinitely dilute vapor at absolute zero; for the entropy, the one relative to which the entropy of ice at absolute zero is just the residual entropy with the nuclear spins ignored. Most tabulations of experimental data are, however, relative to water at the normal melting point and, therefore, the experimental values of the enthalpy, entropy, and free energy must

---

(36) D. Eisenberg and W. Kauzmann, ref 14, pp 105–112, 191–194.
(37) O. Weres, *Chem. Phys. Lett.*, 14, 155 (1972).

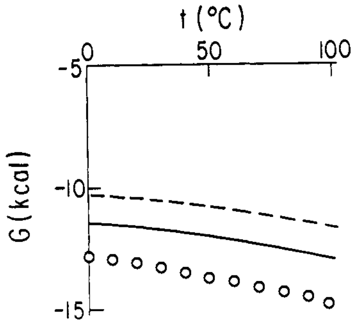

Figure 2. Calculated and experimental free energies. The points indicate experimental values. The dashed line indicates unmodified BNS values and the solid line indicates librational curvature rescaled values.

be suitably converted in order to be comparable with our calculated values.

The above considerations do not apply to the heat capacity and we simply use Osborne's values as quoted by Dorsey.[38]

To obtain the correct entropy values we added on the residual entropy and the calorimetric entropy of the liquid at the melting point, as calculated by Giauque and Stout,[39] to the values of Osborne, Stimson, and Ginnings quoted by Dorsey.[40]

To convert the values of the enthalpy to our standard state, we first calculated the enthalpy of the ideal vapor at 673 K relative to the liquid at the melting point by using Havliček and Miškovsky's value for the real vapor quoted by Dorsey[41] and the second virial coefficient for that temperature calculated from the formula given by Keyes.[42] Friedman and Haar[43] theoretically calculated the various thermodynamic functions of the ideal vapor using the same standard state we do. We obtained the value of the enthalpy at 673 K from their tables by interpolation and subtracted the quantity described immediately above from it. This gave us the enthalpy of the liquid at the melting point relative to our chosen standard state. Finally, we added this quantity to the Osborne, Stimson, and Ginnings values of the enthalpy quoted by Dorsey.[44]

We generated the free energy values from the entropy and enthalpy values and used the International Skeleton Steam Tables density values quoted by Dorsey.[45]

XI. Results

The greater bulk of our results is presented in Figures 2–8 and Tables I–III. It is apparent that rescaling the librational curvature of the potential leads to reasonably good agreement with experiment in the enthalpy and entropy results and improved agreement with experiment in the heat capacity results. All liquid properties other than the translational spectrum were calculated at ten degree intervals.

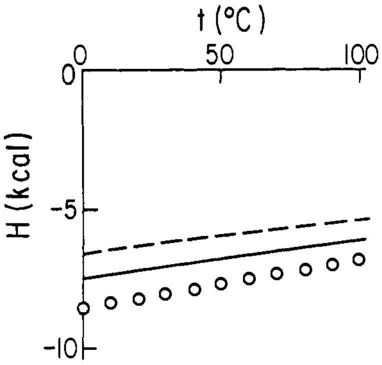

Figure 3. Calculated and experimental enthalpies. Same representation as in Figure 2.

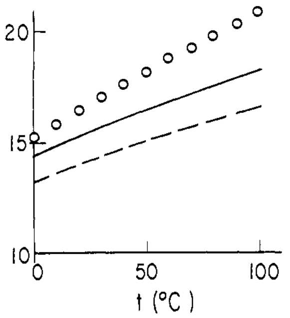

Figure 4. Calculated and experimental entropies. Same representation as in Figures 2 and 3.

Table II. An Analysis of the Values of Enthalpy and Entropy Generated by the Rescaled Curvature Calculation

	t = 0°	t = 100°
Enthalpy, kcal/mol		
Lattice	−8.296	−8.088
Translational[a]	1.729	2.290
Librational[a]	2.517	2.858
Non-hydrogen-bonded neighbor[b]	−1.325	−1.271
Long range	−1.190	−1.166
Intramolecular zero point[c]	−0.900	−0.694
Total	−7.465	−6.071
Experimental total	−8.594	−6.791
Entropy, eu		
Configurational	4.48	4.48
Orientational	1.70	1.69
Translational	7.12	9.28
Librational	2.00	3.44
Non-hydrogen-bonded neighbor[b]	−1.17	−0.95
Vibrational	0.26	0.26
Total	14.39	18.20
Experimental total	15.17	20.79

[a] Includes zero point energy. [b] Refers to perturbation terms. [c] Relative to vapor.

Table III. Experimental and Calculated Thermodynamic Constants for Melting

	ΔH, cal	ΔS, eu	ΔC_p, cal/(mol °K)
Experimental[a]	1463	5.26	8.91
BNS	2822	6.20	3.11[b]
Curvature rescaled BNS	2864	6.29	4.43[b]

[a] Reference 27. [b] Unreliable because ice values do not include intramolecular zero-point energy change contributions.

We will restrict further discussion to the rescaled curvature results unless otherwise indicated.

(38) N. E. Dorsey, "Properties of Ordinary Water Substance," Reinhold, New York, N. Y., 1940, p 258. This is the most extensive compilation of water data available but is weak on spectroscopic data due to the date of publication.

(39) W. F. Giaque and J. W. Stout, *J. Amer. Chem. Soc.*, **58**, 1144 (1936).

(40) N. E. Dorsey, ref 38, p 587.

(41) N. E. Dorsey, ref 38, p 114.

(42) F. G. Keyes, *Trans. ASME*, **78**, 555 (1958).

(43) A. S. Friedman and L. Haar, *J. Chem. Phys.*, **22**, 2051 (1954).

(44) N. E. Dorsey, ref 38, p 585.

(45) N. E. Dorsey, ref 38, pp 200–202.

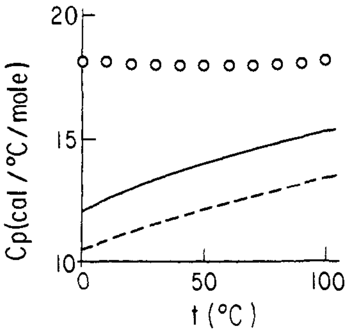

Figure 5. Calculated and experimental heat capacities. Same representation as in Figures 2-4.

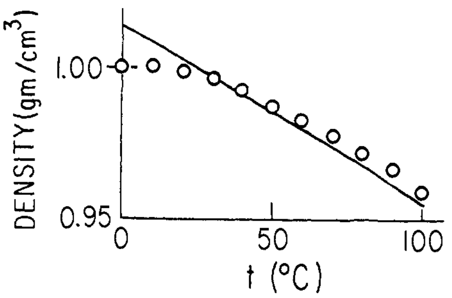

Figure 6. Calculated and experimental densities. The points represent experimental values and the solid line represents librational curvature rescaled potential values. The unmodified BNS values differ from the latter only insignificantly and are not plotted.

Throughout the temperature region studied there is very little variation in the equilibrium basic cell probabilities. The four most important basic cell configurations are shown in Figure 7 along with their concentrations at the melting and boiling points. The number of hydrogen bonds per molecule, ρ, and average number of nearest neighbors are nearly constant at about 1.35, 0.587, and 4.7, respectively. The last is in excellent agreement with the experimental value 4.5 reported by Narten, *et al.*,[2a] and certainly well within the limit of error of the interpretation of the experimental data. The fractions of molecules with two, three or four hydrogen-bonded neighbors are nearly constant throughout the temperature range at about 46, 38, and 16%, respectively. Our calculated value for the Kirkwood correlation factor is about 1.87 throughout the temperature range. The position of the main peaks in the translational frequency spectrum approximation presented in Figure 8 agree well with the ir and inelastic neutron scattering data quoted by Eisenberg and Kauzmann.[46] The negative portions of the spectrum and its high end divergence are numerical artifacts arising from our technique for fitting the spectrum to the moments. We obtained the same structural and spectral results from the unmodified potential calculation.

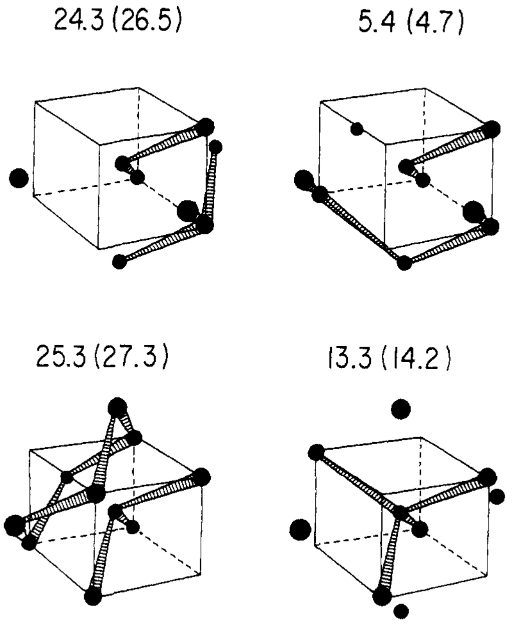

Figure 7. The four most important basic cell configurations. The numbers above each represent its percentage concentration at t = 0 and t = 100°, the latter in parentheses.

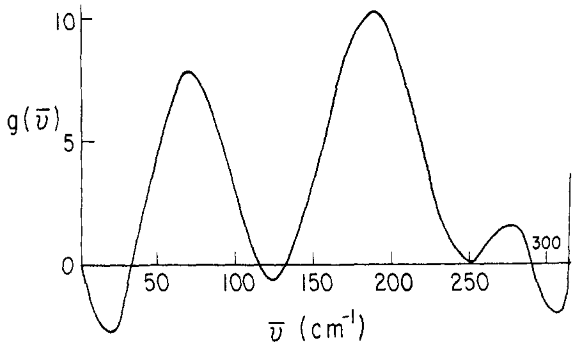

Figure 8. Approximation to the translational frequency spectrum at t = 10°. The ordinate scale refers to g/N represented as a function of the frequency divided by the maximum frequency.

We believe that a goodly portion of the error in the entropy at the lower temperatures is due to the oversimplified treatment of the effects of the correlation of motions of neighboring molecules. Although these results include an approximation to the effect of translational correlations, they completely ignore the correlations of librational motions both among themselves and with translational motions (see, however, ref 26). If intermolecular correlation affects the librational spectrum the same way that it affects the translational spectrum of most three-dimensional systems, *i.e.*, causes a considerable number of modes to appear at about one-third or one-half the Einstein frequency, this alone would be sufficient to explain most or all of the observed entropy error. Another related effect is that of possible anharmonicity in the vibrational modes at the lower end of the translational spectrum which would cause them to contribute more to the entropy than purely harmonic degrees of freedom would. This could be tested by measuring the temperature dependence of the 70 cm^{-1} spectral peak in ice. If a rapid rate of decrease of the maximum frequency with temperature were found at $h\nu/kT \sim 1$, corresponding to closer spacing between higher levels, it would be evidence for the existence of such an anharmonicity effect. (This suggested use of low temperature ice data to study liquid water at normal temperatures is valid, because the information sought has to do with the intermolecular potential energy function which

(46) D. Eisenberg and W. Kauzmann, ref 14, p 228.

must be defined to be the same in both ice and water if the concept is to be meaningful.)

If the above mentioned correlational effects were the only source of error in the entropy calculation, one would expect that the error in the calculated entropy of water at the melting point would be about two-thirds that of ice, since our results indicate that about two-thirds of the hydrogen bonds remain intact in water. Actually, the observed ratio is only about one-third. This indicates the existence of a partially compensating error of about 0.8 eu/mol in one of the other entropy terms. The probable source of this discrepancy becomes evident when the calculated values of the enthalpies of water and ice at the melting point are examined. The calculated enthalpy of ice is in almost perfect agreement with experiment while the calculated enthalpy of water is about 1.2 kcal insufficiently negative. This enthalpy error is probably due largely to the averaging over orientations of the effects of non-hydrogen-bonded near neighbors which are assumed to be at the same lattice position and charge orientations relative to the central molecule in each cell regardless of the relative dipole orientation, as discussed in section V. In reality, the generally attractive non-hydrogen-bonded neighbor pairs must be closer together than the generally repulsive pairs, leading to lower enthalpy and entropy contributions from the cell terms. Another error, probably of the same sign, is introduced into the values of the enthalpy and entropy generated by the rescaled curvature calculation by the use of the unmodified BNS potential and classical phase integrals in the evaluation of the non-hydrogen-bonded neighbor effect perturbation terms. Assuming a total value of 0.8 eu/mol for these errors also accounts for the error in the calculated entropy change associated with melting.

We believe that the rapidly increasing entropy error at the higher temperatures studied is due largely to the gradual breakdown of the validity of our microstructure hypothesis with increasing temperature. This view is supported by the rapid breakdown of intermolecular correlation beyond the first neighbor distance evident in the experimental radial distribution function above about 50°C.[2a] However, we feel that this deficiency should not hurt our model in its most important potential applications, those dealing with biological systems, since biological processes all take place well below this temperature.

A quick comparison of Figures 5 and 6 reveals that the calculated density and heat capacity err most in the lower portion of the temperature range. We believe that these low temperature errors are due to the inapplicability of our configurational entropy expression to highly hydrogen-bonded structures. This inapplicability is indicated by the fact that the value of the configurational entropy of ice predicted by it is -1.38 eu/mol, instead of zero as it should be. Undoubtedly this error carries over to highly hydrogen-bonded water structures, causing them to appear unrealistically thermodynamically unstable relative to less highly hydrogen-bonded structures. The general opinion about the thermal expansion coefficient anomaly is that it is due to a decrease in the extent of the hydrogen bonding overcoming normal thermal expansion. If this were the case such an error in the configurational entropy would tend to mask it, as does indeed happen in our calculations. (We wish to stress that the calculated basic cell configuration probabilities and all quantities derived from them depend on the error in the derivatives of the free energy expression near the minimum much more than on its values.)

Before proceeding to an analysis of the calculated heat capacity it is necessary to discuss the much referred to but rarely scrutinized distinction between the configurational and vibrational components. In terms of the simple cell model, the vibrational contribution refers only to the portion of the heat capacity arising from the individual molecules "climbing the walls" of the potential surfaces of their cells with fixed basic cell structures and probabilities. We calculate it by using the quantum mechanical harmonic oscillator heat capacity expression and the various Einstein frequencies of translation and libration. These contributions cannot be greater than 11.92 cal/(mol°K) ($6R$). The remaining contributions all appear as part of the configurational heat capacity. They include the effects of both intramolecular and intermolecular zero-point energy changes, basic cell structural changes, basic cell configuration probability changes and, in the case of our calculation, the effect of long-range electrostatic interaction changes. In our calculations the sum of the contributions from all of these sources averaged over the temperature interval dealt with is 2.97 cal/(mol °K). Since the difference between the total average heat capacity of 13.96 cal/(mol °K) and this quantity is 11 cal/(mol °K), very nearly the maximum possible value of the vibrational component, we are forced to conclude that most or all of the heat capacity error arises from the configurational component.

An immediate corollary which the reader could draw from the above conclusion is that our result of no variation of the extent of hydrogen bonding is seriously in error. However, we ask him to consider the following. The average heat capacity error of about 4 cal/(mol °K) corresponds to an enthalpy change error over the hundred degree interval of about 400 cal. On the basis of our hydrogen bond energy and hydrogen-bond associated intermolecular zero-point energy results we estimate the enthalpy change associated with the formation of a mole of hydrogen bonds to be about -5 kcal. This means that the heat capacity error indicates an error of only 0.08 hydrogen bond per molecule in the calculated numbers of hydrogen bonds at the boiling point relative to the melting point. This is only about 6% of the total. In fact, we believe that some part of the error arises from the artificially constrained basic cell structure employed and from a breakdown of the microstructure hypothesis. It is quite evident that the cell structure must change with temperature in ways other than simple thermal expansion. (This may also be treated as an effect of the anharmonicity of the intermolecular potential, from which it ultimately arises.) We do, however, believe that some change in the number of hydrogen bonds does occur as argued above in the discussion of the calculated low temperature heat capacity and density errors.

The neutron scattering data indicate a decrease of the peak librational frequency from 580 cm^{-1} at the melting point to about 400 cm^{-1} near the boiling point.[46]

This also seems to contradict our result of very little structural change with temperature. However, we wish to argue against this interpretation. Part of this variation is undoubtedly due to the increasing contributions from "hot" state excitations, which can be expected to occur at lower energies due to anharmonicity effects. Another, and probably more important, origin is that the librational force field is related to the translational motions in the same way that the intramolecular vibrational force field is related to the various intermolecular motions as discussed in section V.

Upon examining the most probable basic cell configurations shown in Figure 7, it is evident that our initial assumption of an ice VIII like space lattice has not at all constrained our resulting structure to resemble this substance. Of the four configurations shown, only the first appears to be part of an interpenetrated network structure and even here the continuation of the structure does not necessitate it. This makes our results consistent with the observation made by Rahman and Stillinger[12] that interpenetration does not appear to be an important structural feature of their molecular dynamics generated sample structures. Interpenetration exists, but is of limited extent.

Our calculated value of 1.87 for g corresponds to a value 8.82 for the quantity $\mu^2 g$, which is more easily experimentally determined. The value of $\mu^2 g$ at 27°C reported by Harris, *et al.*,[47] is about 12.08. There are two possible sources of error which probably tend to lower the calculated value. First, our basic cells are fairly small. Second, our constrained cell structure tends to overweigh lower g relative orientations which appear to generally be energetically less favorable.

XII. Possibilities for Future Research

On the basis of our results we feel that the major problems which must now be dealt with in the realm of theoretical water research are the previously discussed error in the BNS potential and quantum effects.

One possible approach to developing an improved potential function would be to use more than four charges. Their positions and magnitudes could be adjusted to duplicate the electrostatic potential field of the real molecule by using quantum chemical results.

The quantum effects problem appears to be unavoidable in all possible *ab initio* or nearly *ab initio* treatments and models. It will probably also exist in most future treatments of strongly interacting molecular systems. In this case it was possible to deal with them by suitable lattice dynamics-like approximations. More accurate or more complex work will, however, probably require the development of more sophisticated and general techniques for dealing with them. Possible approaches would be to develop tractable quantum correction terms to the Percus-Yevick equation or to the classical equations of motion (for use in molecular dynamics calculations). Another possibility would be to develop a general technique for deriving an effective classical or purely harmonic potential corresponding to any given actual quantum mechanical potential.

Other general problems concern the interpretation of experimental spectra in terms of microstructure and the analysis of intermolecular effects upon intramolecular vibrations. Although numerous attempts have been made in these directions, we have found the generality and reliability of the results to be too low to contribute to the construction of our model in any but the crudest ways.

We will now detail some specific possible improvements to our model. We wish to caution the reader that it probably would not be worthwhile to attempt their implementation until our model is further justified, because the amount of labor involved in each of them is considerable.

As mentioned in section III, we originally envisaged using the Hijmans-De Boer[19] approach to the Ising model in our treatment of the configurational entropy. We see two possible advantages to this approach. First, Bell's[10] results indicate that this sort of treatment is capable of reproducing the density anomaly, which our treatment seems incapable of. We feel that this also indicates the possibility of reducing or eliminating the minor but irritating problems we encountered in calculating the configurational heat capacity and the number of hydrogen bonds. Second, the various constraints can be dealt with through Lagrange multipliers. In other words, our 374 minimization requirements can be reduced to as many simultaneous equations as there are constraints. Further, these simultaneous equations are of a form which is particularly well suited for solution *via* Newton-Raphson iteration. The Newton-Raphson technique would allow arbitrarily high accuracy solutions for the equilibrium values of $\{P_i\}$ to be obtained. More specifically, the nonvanishing free energy derivative problem discussed in section VIII would be eliminated. This would remove the related portion of the ambiguity in the calculated values of the heat capacity and keep similar problems from developing in treating other higher derivative quantities. However, we wish to note that a goodly number of constraints arise in the Hijmans-De Boer treatment which are not present in our treatment. Also our preliminary tests of the Hijmans-De Boer treatment indicate that the Newton-Raphson iteration is unsuitable for obtaining an initial estimate of the solution. This difficulty could, however, be circumvented by obtaining the initial solution by a steepest descent method such as the one employed by us and then using the Newton-Raphson technique to obtain solutions of any degree of accuracy required.

At each temperature, each cell contribution is expressed as a properly weighted sum of contributions from seven harmonic oscillators and ten perturbation terms. The number of independent terms necessary to completely describe each property at each temperature, therefore, is only 33, indicating that most of the 374 variables employed by the present authors are redundant. We feel that a rederivation of the full model in terms of considerably fewer than 374 independent variables is certainly possible, but not necessarily practical. This possibility should, however, be kept in mind by future investigators.

Probably, the most important improvement of the cell contribution results would arise from allowing cell structural variations with temperature other than simple thermal expansion. An especially useful modification of this type would be to allow different relative

---

(47) F. E. Harris, E. W. Haycock, and B. J. Alder, *J. Chem. Phys.*, 21, 1943 (1953).

displacements of the equilibrium configurations for pairs of non-hydrogen-bonded neighbors of different relative orientations. Probably the best way of determining the magnitude of such cell distortions would be to minimize the free energy in respect to them at each temperature. It would also be useful to develop a tractable technique to deal with nonadditive perturbation effects. This would improve the accuracy of our non-hydrogen-bonded neighbor interaction results.

As noted in section V, our treatments of the longer range interactions and the Kirkwood factor are rather crude. A possible alternate approach to these related problems would be to develop a general method of describing long range intermolecular correlation in terms of short range correlation with intermediate molecules. This, of course, would also be of great value in calculating the long range portions of molecular distribution functions.

The obvious way to improve the treatment of the effects of the correlation of molecular motions would be to carry out the correct cluster expansion.[37,48] A standard cluster expansion using the actual potential is, however, not practical due to the enormous numerical complexity involved. Also, the magnitude of the quantum effects prohibits the use of classical cluster integrals. A more reasonable alternative would be to use the quasiharmonic approximation to the actual potential, including the coupling between all six coordinates of each molecule. Anharmonicity effects could be calculated by using an appropriate perturbation technique. Fortunately, our calculated spectrum's good agreement with experiment indicates that the clusters which must be dealt with need not be large. Also, the close agreement of our two translational spectrum estimates indicates that probably only fully hydrogen-bonded clusters need be considered.

We believe that our model is easily extendable to deal with certain types of solutions; in particular, those of inert gases and those of certain types of macromolecules. Solutions of inert gases should be easy to treat because here one basically needs only to introduce basic cells which contain solute atoms into the calculation. Evaluating the appropriate cell integrals and perturbation terms and extending the constraints and configurational entropy expressions should involve no great difficulty. One should, however, be careful to appropriately modify the structure of those cells which contain solute atoms to be consistent with the lack of solvent–solute hydrogen bonds. Other small solutes such as methane, ammonium, hydroxyl, and hydronium ions could also be studied. (In dealing with ionic solutions one should, of course, be careful to account for the long range interionic forces, but it seems that an appropriate dielectric continuum approximation should be easy to derive.) Most other solutions with solutes between the extremes of the inert gases and macromolecules appear to be intractable because of the difficulty of identifying the structures and calculating the properties of the basic cells of the solute molecules due to the size of the solute and the complexity of the solute–solvent potential interaction. The case of macromolecular solutions is exceptional in that the solute–solvent mass and concentration disparities make it possible to ignore the solute cell contributions. Therefore, the problem of calculating the thermodynamic effects of the presence of the solute molecules reduces itself to calculating the solute molecules' effects upon the environments of the neighboring water molecules. This, of course, would not solve the formidable polymer conformational problem, but it would allow solvent effects to be calculated accurately which is a precondition for a full solution of the problem. In the case of rigid macromolecules the conformational problem does not exist, thereby allowing a complete treatment.

Acknowledgments. This research has been supported by the Directorate of Chemical Sciences, Air Force Office of Scientific Research. We have also benefited from the use of facilities provided by the Advanced Research Projects Agency for materials research at the University of Chicago. Oleh Weres wishes to thank the Hertz Foundation for fellowship support.

Appendix I. Bell's Model

Bell[10] postulates a body centered cubic structure and treats the liquid as a lattice gas just as we do. He employs the Guggenheim–McGlashan[21] treatment of the lattice statistical problem using the smallest distorted tetrahedron which appears in the body centered cubic lattice as the basic figure. He includes the orientations of the molecules in the definition of the figure configurations, thereby eliminating the necessity for a separate orientational entropy calculation. The only calculated results which he presents are state surfaces, graphs of isothermal compressibilities, and tables of critical point properties, numbers of nearest neighbors, and ice lattice energies corresponding to three sets of potential energy parameters.

The intermolecular potential is represented by an expression which contains three adjustable parameters. These represent the energy of formation of a hydrogen bond, the interaction energy of two nearest neighbor molecules arising from interactions other than hydrogen bonding, and a "repulsion" parameter which is associated with the fully occupied configuration of the smallest triangle of sites occurring in the body centered cubic lattice.

Bell's general approach was to develop the simplest possible plausible lattice-statistical model and exploit this simplicity to obtain the maximum possible range of results. Our approach, in contrast, was to concentrate primarily on the thermodynamic functions in the normal liquid range of temperatures and introduce as many fine details as seemed necessary to produce reasonable agreement with experiment. For this reason we consider the two models to be complementary rather than competitive. We feel that together they constitute an extensive test of the validity of our common basic assumptions.

The major drawback caused by the extreme simplicity of Bell's model is that it allows only qualitative agreement with experiment in most cases. On the other hand, he was able to calculate state surfaces, a task which our version is nearly incapable of. We feel that the major accomplishment of Bell's work was to demonstrate that the body centered cubic lattice-statistical model is capable of generating a density maximum and an isothermal compressibility minimum. Something which we find upsetting but

---

(48) J. De Boer, *Physica (Utrecht)*, 20, 655 (1954).

not unexpected is that Bell's results show no trace of a liquid–solid phase transition. Instead of a distinct transition between two phases, there is only a gradual variation of density with temperature such as one would expect in a glassy substance. We feel that this also bodes ill for any attempt to extend our approach in that direction.

We also wish to make some comments regarding Bell's approach itself. One of his main criteria for adjusting the values of the energy parameters is the accuracy with which they duplicate critical point properties. We believe this to be an unfortunate procedure because the lattice gas model is rather unsuitable for depicting the gas phase. Hence, parameters fixed in this way must contain a considerable error from this source alone. Also, we have strong reservations about the use of the repulsion parameter, whose purpose appears to be primarily to ensure that the fully occupied lattice structure (*i.e.*, ice VII like structure) is thermodynamically unstable relative to the cubic ice (diamond like) structure at low temperatures and pressures. We do not deny that this is a valid consideration, but believe that it could be dealt with more realistically. As the model stands, this contribution adds an unrealistically large positive term to the free energy of considerably less dense structures. We believe that this is the cause of the rather low densities generated in the solid temperature range at low pressures. The relative instability of the high density structure is obviously due to the fact that neighboring molecules in it are squeezed together by all the other molecules around them; *i.e.*, it is a cooperative effect. If a small region of that structure, say on the order of three or four sites, borders on less dense regions, it is most probably much more stable than if the structure were uniformly dense. We conclude this because the presence of nearby regions of lower density must allow for some distortions which release the intermolecular repulsive forces. We suggest that this parameter be replaced by another term, probabilistic in form, added on to the internal energy to account for the communal repulsive interactions in somewhat larger regions of high density without spuriously destabilizing lower density structures.

Finally, we strongly disagree with Bell's interpretation of the actual appearance of the structure of the liquid. He suggests that it consists of fairly large regions of dense ice VIII like structure and less dense cubic or hexagonal ice like structure. This interpretation is, of course, invalidated by X-ray scattering data.[2] Fortunately, however, this interpretation is not used in any further stage of his work and may, therefore, be disregarded without invalidating any of his derivations or results.

Appendix II. Lattice Statistics

The models which we chose to emulate in our treatment of the configurational entropy problem are the theory of disordered crystalline alloys (which is isomorphous with the Ising model of ferromagnets) and Pauling's calculation of the residual entropy of ice.[17] In both the calculation of the entropy starts with the concept of a certain spatial lattice over which the relevant structural entities are distributed at random. In Pauling's calculation, these entities are protons and free electron pairs. In the Ising model, they are basic cell configurations of up and down spins distributed over all basic cells in the lattices with given probabilities. (In general, the concept of basic cell, also referred to as the basic figure, is defined considerably more flexibly than in our work, occasionally referring to such simple groups of lattice sites as pairs and triplets.) A random distribution is initially assumed as a matter of calculational convenience.

This initial estimate involves a large degree of overcounting because not all of the configurations counted therein are physically possible. In the Pauling calculation, the initial estimate includes configurations in which some molecules have fewer or more than two protons which is, of course, a physical impossibility. In the Ising model, neighboring basic cells overlap, and it is necessary that the overlapping parts of each cell have identical configurations. Otherwise, certain sites would have both up and down spins assigned to them simultaneously, again a physical impossibility. Other factors are multiplied in to compensate for this overcounting. Multiplied together, they constitute an estimate of what fraction of the distributions in the original random ensemble are physically possible. In the Ising model, the typical derivation of this quantity is as follows. Each basic cell is considered separately. It is assigned a certain configuration and the configurations of the remainder are treated as an ensemble of all possible combinations consistent with the original random distribution. The problem is to calculate what fraction of the distributions of the configurations in this ensemble are consistent with the initially assigned configuration of the given basic cell. In the Pauling calculation, the basic cell of a given configuration is replaced by an oxygen atom with one of the six arrangements of protons around it consistent with a neutral molecule. The consistent overlap configuration requirement is replaced by the requirement that a proton must be hydrogen bonded to a free-electron pair and *vice versa*. The total restriction factor is $(^6/_{16})^N$. The $(^1/_{16})^N$ arises because there are 2^4 possible arrangements of protons on the four hydrogen bonds a given molecule is engaged in and only one of them is consistent with neutrality and the given orientation. The factor of 6^N accounts for the existence of six possible orientations of each molecule.

In the case of the Ising model, the equilibrium basic cell configuration probabilities must still be found. This is done by deriving a suitable expression for the internal energy in terms of the basic cell probabilities, combining the energy and entropy expressions into an expression for the free energy, and minimizing. This minimization must be performed under suitable constraints. First of all, the basic cell probabilities must be normalized. (The same applies to any sets of configurational probabilities of lattice sections smaller than the basic cell which might appear in the calculation, but this possibility does not concern us here.) Another type of constraint arises because some of the configurations of the various basic cell overlap regions might be asymmetric in respect to the group of symmetries of the basic cell. Let us assume that a given overlap region configuration may occur in two nonsimilar orientations in a basic cell. The requirement of overlap region consistency demands that its concentration in both of the orientations be equal. To see that this is so, consider the set of all basic cells distributed over the lattice with a certain number of overlap region configurations of the given type within them in the first orientation. These overlap configurations must, of course, appear an equal number of times in the full set of basic cells in the other orientation if the distribution of basic cell configurations is to be physically acceptable. Since the concentrations of the overlap configurations in both orientations can always be expressed as linear combinations of the basic cell probabilities, these constraints are expressible as linear conditions on the basic cell probabilities. (Certain other constraints of this type arise in some treatments of the Ising model but, again, this does not concern us here.)

Appendix III. Computational Details

All of the numerical work reported was executed on the IBM 360/65 installation at the University of Chicago using the Fortran IV G level compiler.

The calculations which were the most expensive in their final form were those which dealt with the generation and characterization of the basic cell configurations, the evaluation of the cell property integrals, and the final minimization of the free energy.

The generation and characterization of the basic cell configurations algorithms were coded using the two byte integers zero and one to represent vacant and occupied positions in the basic cell. The configurations were initially generated by forming all possible combinations of the eighteen allowed nonsimilar first neighbor configurations with each of the fifty allowed second neighbor configurations. Each resulting configuration was then tested to eliminate those that did not satisfy the first-hydrogen-bonded neighbor minimum hydrogen bonding requirement. Finally, each of them was compared with the others to eliminate those that were actually the same except for being in different orientations within the cell. The generation of the 374 configurations plus an evaluation of the degree of symmetry of each one (most often none) took about 5 min of machine time. An independent program generated the expressions for the concentrations of the various molecular clusters which appear in the spectral analysis and the minimization constraints. This step required about 3 min of computer time.

In numerically evaluating the various cell integrals, we restricted the integration to the most important phase region, that in which the oxygen atom is within a cube of one-quarter the edge length of the edge length of the basic cell at the center of the basic cell, the first Euler angle in less than 30°, and the sum (modulo 360°) of the second and third Euler angles is between -30 and 30°. We employed the Conroy algorithm[49] using 1004 points. We also performed some numerical tests on selected integrals to estimate the integration error. The error introduced by restricting the range of integration appears to be less than 1%. Increasing the number of integration ponts to 4822 changes the values of the integrals by only a few percent. Therefore, we may safely conclude that the numerical error in this calculation is not significant because it is almost certainly overwhelmed by that introduced by our various derivational approximations. (It appears, however, that the numerical error may be significant when the heat capacity contributions are calculated in the analytically differentiated form which requires taking the difference of combinations of numerically evaluated integrals.)

A large increase in computational efficiency can be realized if it is recognized that the only change that occurs in the BNS potential under a transformation belonging to the tetrahedral group is a permutation of the charges' signs. The charge–charge distances, the calculations of which comprise the better part of the integral evaluations, do not change. This allows the electrostatic potential energies of a pair of molecules in all 36 relative orientations to be calculated simultaneously. Furthermore, there are only nine relative orientations which differ in the absolute magnitude of the electrostatic interaction. The other 27 may be obtained from these nine by total charge inversions of one or both molecules. The dispersion component of the potential energy is, of course, the same for all 36. (In practice we used only three relative orientations averaging over different orientations of the given basic cell configuration to account for the three possible orientations of the central molecule. A further reduction of computational time is possible if one chooses to eliminate cell structural variations with temperature, thus allowing the same set of potential energy values to be used to evaluate the phase integrals at all tempertures.) As mentioned in section VIII, the major task involved in calculating the Kirkwood correlation factor is evaluating the average projections of the central molecule's moment on the three spatial axes. These may be expressed as very simple combinations of the matrix of rotation relating the central molecules' actual angular position to the ideal lattice axis.

The time required to generate the full set of potential energy values for one temperature is about 3 min; that required to evaluate all of the various phase integrals, including those entering into the evaluation of the analytically differentiated form of the heat capacity and the Kirkwood correlation factor, is about 40 sec.

Our minimization program employs a variable step length algorithm. In this case, we define the step length in terms of the reduction in the free energy per step. We start out with an initial value of the free energy (including the constraint terms), its gradient with respect to $\{P_i\}$ at the given point, and an initial step length. We then multiply the gradient vector by the appropriate factor so that a translation in P_i space along the gradient would lead to a free energy reduction equal to the step length if the free energy were purely linear in $\{P_i\}$. Next, we perform this translation and calculate a new value of the free energy. The two values of the free energy and knowledge of the initial step length provide three pieces of information which are then used to generate an approximation parabola between the two points along the initial gradient direction. We approximate the actual minimum in the given projection of $\{P_i\}$ space as being the minimum of the approximation parabola. If this point is at between one-half of the modified gradient and its

---

(49) H. Conroy, *J. Chem. Phys.*, **47**, 5307 (1967).

full length, we take it as the new initial point and start over. If this point is outside of the stated interval, we temporarily halve or double the initial steplength and start over at the initial point. If either of these procedures is performed in the course of two consecutive iterations the characteristic quantity is either halved or doubled as appropriate. Suitable optional courses of action are built in to allow for the possibility of a negatively concave approximation parabola. The iteration is terminated when the "characteristic quantity" has been reduced to less than $2^{-10}RT$.

The square of the density self-consistency constraint is weighted by 10^3 and those of each of the remaining constraints by 10^2. In order for these quantities to be meaningful, we must specify the exact algebraic forms of the constraints which we employed. The first one we specify by the statement: the filled fraction of all first neighbor sites in all basic cells is equal to the filled fraction of all second neighbor sites in all basic cells. The overlap configuration consistency constraints we specify as: the number of times a given overlap configuration occurs per molecule in one orientation is equal to the number per molecule in the other orientation. We began the first minimization (for 303 K) with the "step length" set equal to RT and each of the basic cell probabilities set equal to $^1/_{374}$th. Obtaining the first result required about 450 iterations and 4 min. Subsequent runs employed the results obtained at a neighboring temperature as starting point and were initiated with the "step length" set at $0.25RT$. They each required no more than about 20 iterations, *i.e.*, about the minimum number required to reduce the "step length" by eight factors of 2. In each case the magnitudes of the errors in the self-consistency conditions were in the range 8×10^{-4}–8×10^{-3}. These errors could, of course, have been reduced by employing larger weighing factors, but only at the cost of much longer minimization runs.

We also wish to note that the reason we were limited to an eight moment formulation of our frequency spectrum approximation technique is that the Hilbert matrices required for larger moment number formulations are not possible to invert accurately using the double precision arithmetic available on the IBM 360/65. This problem might be avoided by using a non-Hilbert matrix dependent formulation or by carrying out the necessary inversions by means of pure integer or multiple precision techniques.

Calculation of H/D, $^{12}C/^{13}C$, and $^{12}C/^{14}C$ Fractionation Factors from Valence Force Fields Derived for a Series of Simple Organic Molecules

S. R. Hartshorn[1] and V. J. Shiner, Jr.*

Contribution from the Department of Chemistry, Indiana University, Bloomington, Indiana 47401. *Received February 7, 1972.*

Abstract: The computer programs of Schachtschneider and Snyder have been used to fit force fields to published, observed fundamental frequencies of vibration for a series of small organic molecules and their deuterated analogs. These force fields have been used to calculate H/D, $^{12}C/^{13}C$, and $^{12}C/^{14}C$ fractionation factors for specific positions in each molecule. The fractionation factors of H attached to carbon are affected in a regular way by variations in the nature of the other groups attached to the same carbon but within the accuracy of the calculations do not seem to be significantly affected by the nature of substituents one atom further removed. Similarly the carbon fractionation factors are affected by the nature of the atoms directly attached but are hardly influenced by more remote structural changes. The use of these factors in the interpretation of equilibrium and kinetic isotope effects is illustrated for several reactions.

The calculation of fractionation factors for isotope exchange reactions has been of interest to chemists for a number of years and has had important applications in the development of isotope separation technology and in the study of geological and biological processes.[2,3]

Since the elucidation of the fundamental theory by Bigeleisen and Mayer[4] and by Melander,[5] the calculation of these factors has been a straightforward problem for molecular species for which the values of the fundamental vibration frequencies were known. However, until the recent development of computer programs for the Wilson FG matrix method of molecular vibrational analysis, only a relatively few molecules could be treated with sufficient accuracy for the results to be of interest. In recent years, complete vibrational analyses of an increasing variety of small organic molecules have been published. It therefore appeared useful to us to use the generalized programs developed by Schachtschneider and Snyder[6] for the vibrational analysis problem and the Wolfsberg and Stern program[7] for the calculation of isotope effects *via* the Bigeleisen equa-

---

(1) Postdoctoral Fellow.

(2) H. C. Urey, *J. Chem. Soc.*, 562 (1947).

(3) J. Bigeleisen, *Science*, 147, 463 (1965).

(4) J. Bigeleisen and M. Mayer, *J. Chem. Phys.*, **15**, 261 (1947).

(5) L. Melander, "Isotope Effects on Reaction Rates," Ronald Press, New York, N. Y., 1960.

(6) J. H. Schachtschneider and R. G. Snyder, *Spectrochim. Acta*, **19**, 117 (1963).

(7) M. Wolfsberg and M. J. Stern, *Pure Appl. Chem.*, **8**, 225 (1964).