Surface layers on ice

Charles A. Knight
National Center for Atmospheric Research, Boulder, Colorado

Abstract. There is a widespread opinion that ice surfaces are covered by a liquid layer at equilibrium, and several treatments of the ice surface describe it as a homogeneous layer with a definable thickness. Arguments are presented that the ice surface cannot have a homogeneous surface layer and that representations of it that use such a layer are not useful approximations. No part of the transition layer on ice can be homogeneous except at a triple point without violating the phase rule. Some experiments suggest the existence of a surprisingly thick transition region at ice surfaces, but others do not. Water apparently does not completely wet ice, since its contact angle on ice close to the triple point appears not to be zero, and sharply faceted growth forms of ice from the vapor close to the melting point also suggest the absence of a thick, liquid, surface transition region.

Introduction

The nature of ice surfaces, meaning ice-vapor or ice-air interfaces in this article, has been discussed widely and for a long time with regard to their degree of disorder, their likeness to liquid water. Hobbs [1974] summarized the older literature, but significant theoretical and observational studies have appeared since that time. Prominent among recent theoretical treatments have been representations of ice as having a melted surface layer below the melting point, a liquid layer that is treated as if it had definable, bulk thermodynamic properties [e.g., Lacmann and Stranski, 1972; Nenow and Trayanov, 1986; Dash, 1989a, b]. Recently, such a model has been applied in cloud physics to charge transfer between colliding and separating ice particles [Baker and Dash, 1989, 1994], and it has also been used in ice crystal growth calculations [Kuroda and Lacmann, 1982; see Kobayashi and Kuroda, 1987] and in calculations of ice evaporation rates [Chen and Crutzen, 1994]. Its consequences have been reviewed [Dash et al., 1995].

It is argued here that this kind of representation of the ice surface is fundamentally incorrect. It is also argued that this is not just a technicality. This fallacy leads to qualitatively incorrect reasoning about surface behavior in situations important in cloud physics and chemistry.

All real surfaces are transition regions. Many are very abrupt but none are perfectly so, and some can have important structural changes over distances of several or many molecular spacings. These changes may involve the amount of molecular order, the nature of that order, or both. However, as this writer sees it, it is a fallacy to represent surface transition regions as homogeneous layers of discretely different structure from the phases on either side. The published treatments that share this fallacy are similar in principle but they differ in details. For the sake of exposition, the recent article on thunderstorm electrification [Baker and Dash, 1994, hereinafter referred to as B-D] is chosen for specific reference of the criticisms. The original idea that the ice surface might really be melted at subzero temperatures and that the layer can be treated as a phase dates at least to Lacmann and Stranski [1972], who used it to explain snow dendrites.

The usual (classical) thermodynamic representation of surfaces [Gibbs, 1878] is discussed first, because once its fundamentals are understood, the fallacy in the surface layer representation quickly becomes clear. A one-component system (here, H2O) is assumed for simplicity, since none of the surface layer proponents requires more than one. Other conditions are equilibrium and no surface curvature.

The Thermodynamic Surface

Figure 1 is a sketch of a system composed of ice and vapor, and equation (1) is a formula for its free energy.

\[ G = \mu_s N_s + \mu_v N_v + A \gamma_{sv}, \]

where \( G \) is the free energy of the system, \( \mu \) the chemical potential per molecule, \( N \) the number of molecules, subscripts \( s \) and \( v \) represent ice (solid) and water vapor, \( A \) is interfacial area, and \( \gamma_{sv} \) is the surface energy. Surface energy is defined as the work (or energy) required to create unit area of new surface reversibly. One might think of slowly deforming the system in Figure 1, maintaining temperature and pressure constant, until the surface area is one unit larger and determining how
much energy was required. Another way would be to recognize that $\mu_0$ equals $\mu_s$ at equilibrium across a flat interface, so equation (1) is equivalent to

$$G = \mu N + A\gamma_{sv},$$

where $N$ is the total number of molecules in the system. One might then count the molecules, determine $\mu$, determine the total free energy of the system, and $A\gamma_{sv}$ is the remainder; $\gamma_{sv}$ is measurable in either of these ways in principle to any degree of accuracy, with no reference to actual surface structure.

Equation (2) emphasizes that surface energy is an excess quantity: it is not the energy of any material, but the extra energy that results from the presence of the surface. It is not referred to any volume of material. The question of where it "belongs" does not arise in considerations of the total energy of the system in Figure 1 or in many practical applications. It does matter exactly where the surface plane in the representation is located, however, because the relative volumes of ice and vapor depend upon it. The volume and density of the total system are measurable to any degree of precision, and the densities of the individual phases are known. There is only one value of $y$ (Figure 1) for which the actual mass of the system is the sum of the products of volume and density of the phases, so it is natural to ascribe the surface energy to that plane in this one-component system. (In systems with more than one component the issue of where the surface plane of the representation is located is critical in treating adsorption phenomena consistently.) Gibbs called the surface plane the "dividing surface."

The important point is that the system in Figure 1 is fully, formally specified, and all the variables have operational, meaningful, and consistent definitions. However, there is no information in this representation about the nature of the actual surface transition region between ice and vapor. This region has whatever structure minimizes the surface free energy [e.g., Fletcher, 1968], because that is the condition of equilibrium; but in its thermodynamic representation it is a mathematical plane.

### The Surface Layer Representation

In several works noted above, the identical ice surface is represented as a discrete layer. In B-D, equation (1) is replaced by equation (3) below (which is equation (1) in B-D), which implies the sketch of Figure 2.

$$G = \mu_s N_s + \mu_{QLL} N_{QLL} + \mu_v N_v + A\gamma_{sv}(h),$$

where $QLL$ is the quasi-liquid layer and $\gamma_{sv}(h)$ is a surface energy term representing both of the two new interfaces shown in Figure 2.

Equation (3) implies two zero-volume interfaces instead of one, since presumably $N_s + N_v + N_{QLL} = N$, the total number of molecules in the system; but as with the classical interface of Figure 1, this is a representation, not a reality. As before, the actual, physical surfaces must really be transition regions. (Numerical modeling [Karim and Haymet, 1988] suggests that the water-ice interface itself may be exceptionally thick.) Thus a convention is needed to locate the interfaces in the representation. Also, neither of the ways of determining interfacial energy that can be used to specify $\gamma_{sv}$ in (1) is available now. There is no way of creating new
surface area for one of these interfaces without simultaneously creating the other and the QLL as well, if this is the equilibrium structure of the ice surface. There is no evident way to assign numbers to these variables.

Actually, we need to discuss two different cases because it is not clear what the users of this kind of representation consider the real interface to be. They might think either that it is a truly homogeneous layer with two interfaces as in Figure 2, or that it is a single, thick transition region, and Figure 2 is an approximation. The arguments here will be that the first option is impossible in a one-component system away from the triple point and that a thick transition region is not usefully represented as a homogeneous layer.

The first option posits an actual, homogeneous surface layer. This means physically that the center portion of the layer is unaffected by the two interfaces, for if the interfaces influenced the state of this material, it is inconceivable that it would be homogeneous: that the strength of this influence would not depend upon distance. Being homogeneous, this interface layer is a phase and since it must be in equilibrium with the phases on either side, its existence immediately violates the phase rule. Gibbs' phase rule, one of the most fundamental theorems of thermodynamics, states that the maximum possible number of phases at equilibrium equals the number of components in a system minus the degrees of freedom plus two. Thus if the pressure and temperature of the one-component, water system at any place except the triple point, with ice in equilibrium with water vapor, there is one degree of freedom. Either pressure or temperature can be varied independently but not both, and ordinary liquid water cannot exist at equilibrium. Likewise, if a "quasi-liquid" phase has any equilibrium existence, it could only coexist with ice and vapor at some other triple point.

Dash [1989b] argued that there exists a "thermo-molecular pressure" such that this layer is at a pressure lower than that of the vapor and ice on either side, making its chemical potential equal to that of the vapor and the ice. It would be liquid water at the appropriately lower pressure. The logic here appears to be circular in terms of the main question of whether such a layer can exist. If it exists, it has to be a phase, and its chemical potential has to equal that of the other phases, and therefore its pressure has to be lower. But if the material is really homogeneous, its state must be independent of the surfaces and vice versa, so $\gamma_s(h)$ must also be independent of the thickness of the layer. In any case, its existence would still violate the phase rule.

(It may be worth remarking that the pressure difference at equilibrium across a curved surface, the basis for capillary phenomena, derives entirely from the work needed to create surface. It exists because the system geometry requires that the volume of a phase cannot change without the surface area changing as well. It is irrelevant whether the interface transition zone is less or more dense than the bulk material, whether one might consider this material at a higher or lower "pressure." The discussion in this article concerns flat interfaces, and any surface influence upon pressure would have to be an entirely different kind of phenomenon from that across a curved surface.)

Li and Neumann [1991] argued in a somewhat different context (they treated only multicomponent systems) that a surface layer thickness $h$ can provide another degree of freedom that allows the existence of a true surface phase. The thickness would have to be defined by some convention which they did not specify, but even if this were possible, it could not be an independent variable. One cannot imagine a constraint that would allow $h$ to vary at constant pressure and temperature while retaining equilibrium. Thus $h$ cannot fill this role even if it could be defined. The same error is committed in B-D, both in general and in the specific case where the free energy of "the dry solid vapor interface without a surface layer" is used in defining $\gamma_e(h)$ in equation (3). This means setting $h$ equal to zero and claiming that the resulting interfacial energy is a meaningful quantity. If the equilibrium surface really is "wet," what constraint could possibly be applied, either in actuality or in a molecular-dynamic model [e.g., Kroes, 1992] to keep it "dry" while a meaningful energy is determined?

In the second interpretation of the liquid layer treatments, the quasi-liquid layer representation, equation (3) and Figure 2, is an approximation to a thick transition region, nowhere homogeneous in the direction normal to the interface. There is no real, bulk material that can be said to correspond to "the material of the QLL," so the phase rule is obeyed. This must be the actual case according to the argument above, and the issue now is whether the QLL representation can serve any purpose. If this new representation is to be useful, formal conventions are needed to define and arrive at values of the density and chemical potential of the QLL material in the representation, and to define the locations of both of the new, zero-volume interfaces. Otherwise the numerical values of most of the variables in equation (3) are indeterminate. The need for such conventions is not even mentioned in B-D or any of the other articles using this kind of representation. Thus there appears to be no known way to apply equation (3) to an ice surface or to any other surface.

The Surface Layer Representation As an Approximation

According to the argument above, the surface layer representation contains unspecified parameters that are crucially important. Part of the arbitrariness is removed in B-D by assigning to the surface layer the prop-
erties of ordinary liquid water at the temperature and pressure of the ice and vapor. They claim that this is a reasonable approximation because it would be correct "in the limit of very thick layers," but the theme of how good the approximation might be is not pursued further. If "in the limit of very thick layers" means at the triple point, any water layer thickness is in equilibrium if the contact angle of water on ice is zero. At any lower temperature, however, the thicker the layer of water (necessarily supercooled) on the surface, the farther from equilibrium the system would be. Yet the surface layer representation is needed in B-D down to at least -15°C to be useful for its application to charge generation in thunderstorms.

Regardless of application, it is difficult to see how the surface layer representation could be shown to be a good approximation anywhere, when prescriptions are not supplied to quantify the theory being approximated.

Are Difficulties with Surface Layer Representation Just Technicalities or Are They Important?

This writer is uncomfortable with some very common word usage applied to surface structure. "Surface melting" and "liquid layer" are unfortunate phrases because they imply that the surface is a phase, which it certainly is not in the usual sense. "Surface disordering" and "partially disordered transition region" would be better. "Surface layer," "liquidlike layer," and "quasi-liquid layer" can also be unfortunate if they carry the inappropriate mental picture of a homogeneous layer rather than a transition region. These phrases are commonly used without intending the controversial implications, and making a major issue of their inappropriateness would be belaboring a technicality. However, the formal liquid layer representation is another matter in that it leads to intuitive thinking about surface behavior that is clearly incorrect.

For example, ice particle collisions are discussed in B-D as follows. "When the particles make contact, the liquid layers merge. Then, as the particles move away from each other, the bridging liquid narrows and breaks, leaving a fraction of the combined thickness on each particle." This is intuitive thinking about the surface transition zone exactly as if it were a phase: as if the ice-ice collision is just like a collision between (say) glass particles covered with liquid water. Referring to Figure 2, this is thinking of the two fictitious interfaces in the representation as if they were independent, which they clearly are not.

It has not been argued here that it is impossible for ice surfaces to have some structure that allows for exceptional charge transfer in a collision: that is a separate issue. In fact, there might be a liquid, surface phase if sufficient solute is present [e.g., Conklin and Bales, 1993], and then the description of an ice-ice collision in B-D might have application. The argument here is just that the discussion of ice surfaces in B-D is both invalid and badly misleading.

The Real Ice Surface

So far the discussion has concerned the validity of representations of ice surfaces. The question of whether or not the actual ice surface has a "thick," "quasi-liquid," transition layer with little long-range order has importance in geophysics, but it is a separate issue. Observations and numerical modeling are the two ways of investigating this question, and both are problematic. Numerical modeling still entails approximations that may well be oversimplifications, and the extensive experimental studies to date do not present a clear and consistent picture of the nature and extent of the surface transition.

Two current reviews from very different perspectives are by Dash et al. [1994] and by Petrenko [1994]. The former takes the existence of a liquid layer as an established fact. The latter includes in a small subsection titled "Existence of a special layer" the statement "All authors admit to its existence. Let us join them, keeping in mind an impressive listing of experimental evidence considered above." There is indeed evidence for a "special layer," but all surfaces are special layers. The problem for the scientist is just how special this particular "layer" is, and the body of experimental evidence leaves a wide latitude for interpretation.

Most students of ice surface structure probably form a personal opinion of what the surface is like, in spite of the existence of some contrary experimental evidence. This writer's opinion is that the ice surface transition zone probably is neither very thick nor completely disordered. The evidence that points this way comes from observations of surface phenomena rather than direct measurements of surface structure: the contact angle of liquid water on ice and ice crystal growth habits. These are discussed in the Appendix, along with ellipsometric results which provide some contrary evidence but also provide one example of inconsistencies in the published studies.

Conclusions

There are a number of related topics that could have been considered in this article. Disjoining pressure, the pressure between two rigid walls held a very small distance apart with another phase in between, does depend upon that distance but is a phenomenon not relevant to the present discussion because of the extra constraint imposed by the rigid walls. Likewise, the two-dimensional phase changes that occur in adsorbed monolayers are not relevant, because such monolayers can be represented realistically as two dimensional, whereas the ice surface cannot. There is a lot of lit-
temperature on "unfrozen water" at subzero temperature in frozen soils and how it relates to frost-heaving. The phase arguments presented in discussing the free surface of ice apply equally well to the interface between ice and an insoluble silicate mineral.

The thesis here is that the ice surface cannot have a homogeneous, liquid layer below the triple point without it being caused by a second component in the system, because such an occurrence would violate the phase rule. It might have a thick, transition layer with little or no longrange order over an appreciable distance, but representing such a layer as if it were homogeneous only introduces confusion into thinking about surface behavior. It is hoped that this article will sharpen the terms of the discussion.

The actual nature of the surface region of ice is not settled. Experimental results even using closely related techniques of optical ellipsometry are not easily reconciled with each other (see Appendix), and further investigation clearly is needed, including repetitions of some of the past experiments.

Appendix: Contact Angle of Water on Ice, Ellipsometric Measurements of the Ice Surface, and Ice Growth Habits

This writer has suspected for many years that the ice surface is not completely disordered even at the triple point, because of observations of the contact angle of water on ice. It may be impossible to measure this unequivocally, but indications are that it is not zero [Knight, 1967, 1971; Elbaum et al., 1993], and the temperatures at which these observations were made were much less than 1°C from the triple point. It is difficult to envision a nonzero contact angle of liquid water on a thick layer of liquid that is nearly identical. One cannot say exactly what "thick layer" means in this statement but certainly more than 100 Å would qualify. (These are inevitably qualitative statements, since the surface is really a transition layer. As long as one does not write down an equation that involves the thickness, one can avoid defining it precisely. It might be defined operationally in terms of one property or another, but different experimental techniques would not be expected to agree.) Several authors have suggested that layers considerably thicker than 20 or even 100 Å exist close to the freezing point, but this is difficult to reconcile with observations that water does not completely wet ice. While the contact angle measurements are subject to uncertainty because of both imperfect equilibrium and inadequate proof of purity, considerable care was taken in the referenced investigations.

Three ellipsometric studies of the ice surface have been performed specifically to investigate the quasi-liquid layer. Beaglehole and Nason [1980] found no measurable layer on the basal face below a temperature unresolvably close to 0°C, within about a tenth of a degree, but they did find a layer on the prism face that became thick above about -5°C. Furukawa et al. [1987] deduced liquid layers above -2°C on the basal face and above -4°C on the prism, both thickening without evidence of an identifiable limit as the temperature approached 0°C. The prism face results were not quantitatively reproducible. The experiments of both Beaglehole and Furukawa involved ice-air interfaces, but the surfaces were prepared differently and the ellipsometric techniques were different. Elbaum et al. [1993] studied ice in contact with both pure vapor and air, finding that in contact with vapor, "...neither the basal nor the prismatic facet undergoes complete surface melting," while with air in the chamber, a thick layer formed on the basal facet and on noncrystallographic surfaces, close to the melting point.

These results are inconsistent with each other, and both of the latter two studies reported their own results as not well reproducible in some respects. As Elbaum et al. [1993] remark, it does seem clear that air is important, and the results, in general, suggest that impurities may be an important problem. (The freezing-point lowering of a saturated solution of air in water is less than 0.01°C, so air cannot furnish a true liquid layer at a temperature significantly below the true freezing point.) The water system is one in which impurity effects may be nearly impossible to rule out, so this difficulty may remain in future experiments.

The striking and familiar changes of ice crystal shape as a function of temperature in growth from the vapor [e.g., Hobbs, 1974] have been interpreted in terms of a liquid layer model [Kuroda and Lacmann, 1982; Kobayashi and Kuroda, 1987]. These treatments are formal, and it seems worth discussing the issue the other way around: does ice growth behavior from the vapor shed any light on the nature of the ice surface? This is a potentially valuable source of evidence because the orientation-dependence of ice growth rates, the habit, might be expected to be a rather sensitive indicator of surface structure. Furthermore, the ice growth habit changes have been reproduced in a number of laboratories without especially careful purity precautions (using laboratory air, for instance), so the surface changes that the growth habit changes reflect must not themselves be supersensitive to accidental impurities. (There are systematic effects of a number of gaseous organic compounds upon ice growth habit, however [e.g., Hobbs, 1974].)

The habit changes at -4° and -8°C are quite abrupt, and they certainly reflect relative changes in ice surface properties between the prism and the basal faces. However, the evidence is that the growth mechanism is surface nucleation, not growth at repeatable steps [e.g. Keller et al., 1980; Frank, 1982]. The nucleation rate is so sensitive to the edge energy of a growth layer that gradual changes of surface structure could easily pro-
duce very abrupt changes of growth habit. Thus abrupt growth habit changes need not imply abrupt or drastic changes of surface structure.

A more significant feature in the present context is the presence or absence of facets. If ice has a nearly amorphous surface transition region that is much like liquid water (it cannot be homogeneous, but it could be completely devoid of long-range order for an appreciable distance), then one would expect the growth of ice from the vapor at a temperature close to 0°C to be much like growth from liquid water. The process would be condensation of molecules into the amorphous region followed by crystallisation in the region of transition to long-range order. The more like liquid water the amorphous region is, the more the crystal growth should be like that from slightly supercooled liquid water.

Interpretation of crystal growth habits is complicated by the presence of two influences: the growth shape is often completely determined by orientation-dependent growth rates, but there is also a tendency toward equilibrium form, the shape with the lowest total surface energy per unit volume. The latter might dominate in slow growth of small crystals. There is strong evidence that the equilibrium shape of ice crystals in vapor or air at moderate subzero temperatures is completely rounded, with no facets at all. Colbeck [1985] obtained this result, and the observations by Nakaya [1956] of the spontaneous, isothermal breakup of vapor figures within ice crystals into much smaller, completely rounded bubbles appear convincing in this regard. A rounded equilibrium form is evidence for a rough surface, but neither for nor against an amorphous layer on the surface, so a rounded shape in very slow growth from the vapor also carries no strong implications about the nature of the ice surface.

The growth of ice from slightly supercooled water is in the form of flat, thin, circular disks at first, which develop scalloped edges at a diameter of a few millimeters and eventually attain a dendritic but completely rounded form at larger sizes [Arakawa, 1955; see Hobbs, 1974]. The growth rate is extremely anisotropic at very low supercooling, with the rate parallel to the c axis being nearly zero, but typically, there is little or no indication of prism facets. Ice growth from liquid water is interpreted as involving growth in the c axis direction by two-dimensional nucleation but growth normal to the c axis by a continuous mechanism, with surfaces parallel to the c axis being rough [Jackson, 1958; Jackson et al., 1967]. If the ice-air interface has a thick, very waterlike layer at, say -0.5°C, then one would expect ice growth from the vapor at -0.5°C and slight undersaturation with respect to supercooled water to have this same habit: basal facets with an otherwise curved interface.

Since the observations reported in the literature were not perfectly clear on whether ice growth from the vapor is completely faceted in these circumstances or not, a simple experiment was performed. The apparatus consisted of a 20 cm³ glass vial with a screw cap, with 2 cm³ of sugar solution with freezing point -0.1°C in the bottom and a microliter pipet tip in the lid (Figure 3). The pipet tip was filled with pure water and apparatus sealed and put into a constant temperature chamber at -0.4°C. After temperature equilibration the supercooled water in the pipet tip was nucleated. The water froze slowly, and the crystal exposed at the end within the vial was nearly invariably a single crystal. It then grew from the vapor in the constant-temperature environment for many hours in conditions subsaturated with respect to liquid water, controlled by the sugar solution. Figure 4 shows one result, from 87 hours of growth at -0.4 ± 0.1°C, with basal and prism facets intersecting in sharp edges.

This looks like evidence against thick, quasi-liquid layers on either the basal or the prism face of ice at -0.4°C. As was the case with the optical ellipsometric studies, however, there are a number of observations of ice growth habits in the literature and the results here too are not completely consistent. For instance, Sei and Gonda (1992) and Sakamoto and Shichiri (1992) reported and illustrated prism facets on small ice crystals grown in water at rather small supercooling, but the ice crystals were evidently floating at the water-air interface. Furukawa and Shimada (1992) present photos that show no traces of such facets in growth in bulk.
water, and Knight (1968) showed crystals grown over many days in liquid water at a few hundredths of a degree supercooling with no prism facets. The writer has found that any contact of water with the plastic polyvinyl chloride, which is used as a foam to make ice buckets and in solid form as a common tubing material, causes prism facets on ice crystals grown slowly. The effect appears to be due to small amounts of a substance related to polyvinyl alcohol in solution. Thus it is possible for prism facets to be caused by minute amounts of a rather common impurity, but one can not know whether that explains the faceted forms reported above.

The observations of ice growth from the vapor at temperatures very close to 0°C are also interestingly diverse. For instance, in contrast to the result shown in Figure 4, Furukawa and Kohata [1993] found no prism facets on negative crystals grown at temperatures above -2°C (this is the ice-vapor interface, without air) and Keller et al. [1980] found no prism faces on small ice discs growing from the vapor at -5°C and 2.5% ice supersaturation, until the disks thickened to about 40 μm, at which time prism facets developed.

The reasons for the diversity in these and other such observations on ice are not clear, and the nature of the ice surface itself is not clear. One of the strongest pieces of direct evidence for a thick, liquidlike layer on ice is the ellipsometric study by Furukawa et al. [1987], which suggests a surface layer with the index of refraction of liquid water that is as thick as 1000 Å at -0.5°C. On the one hand, it seems unreasonable to suggest that enough impurities might be present to account for this as a layer of solution. Yet on the other hand, a layer this thick is hard to envision without it being a true phase. It also seems unreasonable both that the contact angle of water on a layer this thick would not be zero (especially much closer to 0°C, where they suggest a layer much thicker even than 1000 Å), and that the growth habit at this temperature, with such a layer, would be sharply faceted.

Acknowledgments. The National Center for Atmospheric Research is supported by the National Science Foundation. Discussions with a number of people, notably A. D. J. Haymet, B. Baker, and J. Nelson, have shaped the form of this paper, as has correspondence with J. G. Dash and M. B. Baker, and comments from W. A. Cooper and J. E. Dye on a semifinal version were very helpful.

References


