PHYSICAL REVIEW B VOLUME 24, NUMBER 1 1 JULY 1981

Phase transitions in heterogeneous films

J. G. Dash
Department of Physics, University of Washington, Seattle, Washington 98195

R. D. Puff
Department of Physics and Institute for Nuclear Theory, University of Washington, Seattle, Washington 98195

(Received 21 November 1980)

The thermodynamics of monolayer films adsorbed on nonuniform substrates is analyzed in terms of arbitrary but well characterized substrate heterogeneity. It is shown that a combination of vapor-pressure isotherms of a two-phase film and high-resolution diffraction can yield the distribution functions of binding energy and of crystallite size. These distributions can be used as a basis for characterizing all films subject to similar adsorption forces. The effects of these distributions are discussed, with reference to single-phase regimes, phase boundaries of two-phase regions, triple-point melting, and second-order transitions. Several specific examples drawn from recent literature are discussed. A method is developed for unfolding ideal film behavior from measurements on well-characterized nonuniform substrates.

I. INTRODUCTION

Current interest in phase transitions of surface films[1] heightens the need for improved substrate characterization and for greater understanding of the effects of heterogeneity. Although observations of several types of surface phases and phase transitions indicates that many films are quite uniform, imperfections that would otherwise be considered minor in kind and number can be important in the neighborhood of phase transitions, where compressibilities become divergent. Heterogeneity may be an important factor in some questions of lively concern, e.g., continuous versus first-order melting in specific systems, whether certain films have two-dimensional (2D) liquid phases, and whether commensurate-incommensurate transitions are first order.

Several types of imperfection play significant roles in physical adsorption, principally crystalline disorder, chemical impurities, surface roughness, microporosity, and orientational disorder. The manner and degree to which they affect an individual film depend on its composition and phase. Presumably, a complete characterization of the surface together with a detailed microscopic model of the film would allow one to calculate the perturbations. Present knowledge and techniques are far from this stage of sophistication. However, for certain general questions it is not necessary to be so specific. In dealing with questions of phases and phase transitions one is primarily concerned with perturbations of a small number of thermodynamic variables, such as the density and structure of the film. The various types of substrate imperfection affect the density and structure in distinctive ways, but their mechanisms are primarily limited to just two; variations in substrate attractive potential and effects due to finite sizes. Variations in substrate potential introduce lateral fields, which cause density variations. Small crystallite and island sizes introduce boundary effects and limit long-range interactions. The two parameters can be mapped, in principle, by a detailed topography of the substrate-adatom potential at all positions along the surface, which would show the boundaries of the domains and the locations of nucleation centers, as well as the more uniform regions in between. For some purposes it is important to know the spatial distribution, but for many applications a less detailed characterization, specifying only a statistical distribution, is sufficient. We limit ourselves to such cases, where it is enough to know what fraction of the surface has a certain range of potential, and how many crystallites and islands there are of a given size.

In the next section of this paper we propose an experimental procedure to obtain these distributions. In Sec. III we discuss the effects of heterogeneity on specific thermodynamic regimes, with applications to published data, and draw tentative conclusions concerning several current questions. In Sec. IV we show how it is possible to unfold the unperturbed properties of a film from measurements on the heterogeneous system.

II. SUBSTRATE CHARACTERIZATION

A. Binding energy heterogeneity, without size effects

Nonuniform substrate binding energy has been recognized for many years as a principal source of film heterogeneity.[2] A technique for measuring the energy distribution from vapor-pressure isotherms dates from the same time.[3] The technique is based on the direct dependence of the pressure of a low coverage gas phase film on the substrate-adatom binding energy averaged over the population. The method works best for obtaining the number of sites with higher than average binding energies. It has been employed to obtain the high-energy side of the He-graphite binding energy distribution.[4] This classic method is not adequate, however, for surveying the intermediate and weak binding portions of the distribution, since these are appreciably occupied only at relatively high coverage, where adatom interactions are both more important and change with coverage. The method also suffers in that it has limited sensitivity to small variations of potential, of magnitudes characteristic of the most uniform regions of well ordered samples. But these drawbacks can be eliminated by a simple modification of the technique: use of a two-phase film in place of a low-density gas.

In a regime of first-order phase condensation a film has infinite two-dimensional compressibility, which makes it sensitive to arbitrarily weak lateral fields. On an ideal substrate the vapor pressure of a two-phase film is a function of T but not of coverage, the isotherm showing a vertical (constant pressure and constant chemical potential μ) step in the coexistence region. On a nonideal surface the isotherm has a finite slope. The slope is a measure of the ideality, i.e., of the constancy of the substrate binding. In the two-phase region the denser phase is preferentially adsorbed on the more attractive parts of the surface, and as the coverage is increased the interface between the two surface phases moves upward in potential. As the potential energy at the interface varies, so does the chemical potential of the system, hence the vapor pressure of the film. Thus the steepness of the isotherm gauges the ideality of the surface at the immediate locations of the interface. Since the capacity of the regions having the same binding energy is gauged by the quantity adsorbed, the capacity and perfection can be measured simultaneously as a function of substrate potential. A direct measure can be obtained in terms of a density of states $G(\mu)$, defined by

$$G(\mu) \equiv \frac{1}{N}\left(\frac{\partial N}{\partial \mu}\right)_{T,A} = \left(\frac{\partial \ln N}{\partial \mu}\right)_{T,A} , \qquad (1)$$

where N is the number of molecules adsorbed.

As it stands, $G(\mu)$ describes a specific film and thermodynamic regime on a particular substrate. To be most useful, a characterization should describe the substrate alone, so that it might then be used to predict the heterogeneity of different films and their phases. Such a generalization can be obtained from $G(\mu)$, as follows.

In the case of an ideal uniform film the substrate potential ϵ is, to first order, an additive term in the chemical potential, so that the equilibrium relation between the film and gas can be written

$$\mu \equiv \mu_{\text{vapor}}(P,T) = \mu_{\text{film}}(\epsilon, \phi, T) \simeq -\epsilon + \mu_2(\phi, T) , \qquad (2)$$

where μ_2 is the "two-dimensional equation of state" of the film in terms of T and the spreading pressure ϕ. We make here the "local equilibrium" assumption that Eq. (2) applies also in cases where the substrate potential ϵ varies locally over the film; i.e., where one can treat ϵ as a function of position (x,y) on the film, and assume that equilibrium is established between vapor and film for each local region in the neighborhood of (x,y). Specifically, we are assuming that there are no terms in μ_{film} involving gradients of ϵ. The relationship between variations of potential energy and chemical potential is particularly simple in a two-phase regime. In a two-phase regime at constant T the spreading pressure is fixed at some value $\phi = \phi_{eq}$. On a substrate of varying ϵ the interface between the two phases is always located at the same value of ϕ_{eq} and hence the same value of μ_2. Variations in chemical potential of the system directly reflect the variations of substrate potential at the position of the interface between the two film phases, i.e., $d\mu_{\text{vapor}} = -d\epsilon_{\text{interface}}$. Therefore, the density of states with respect to energy is simply related to the slope of the isotherm in the two-phase region

$$G(\mu) = -\left(\frac{\partial \ln N}{\partial \epsilon_{\text{int}}}\right)_{T,A} . \qquad (3)$$

Changes in the adsorbed quantity can be related to changes in the adsorption area swept out by the shifting interface. Specifying the populations and areas in the two phases (1,2),

$$\begin{aligned} N &= N_1 + N_2, \quad A = A_1 + A_2 , \\ n_1 &= N_1/A_1, \quad n_2 = N_2/A_2 , \end{aligned} \qquad (4)$$

the changes δN can be related to the changes in area, with a correction term involving the two-dimensional compressibilities of the two phases:

$$\delta N = (n_2 - n_1)\delta A_2 + (A_2 n_2 K_2 + A_1 n_1 K_1)\delta\phi , \qquad (5)$$

where the compressibility K is obtainable from thermodynamic data such as vapor-pressure isotherms by[5,6]

$$K \equiv \frac{1}{n}\left(\frac{\partial n}{\partial \phi}\right)_T = \frac{1}{n^2}\left(\frac{\partial n}{\partial \mu}\right)_T . \qquad (6)$$

Equating $d\phi = n d\mu$ (Ref. 6) at constant T and differentiating with respect to μ we obtain

$$\left(\frac{\partial N}{\partial \mu}\right)_{T,A} = (n_2 - n_1)\left(\frac{\partial A_2}{\partial \mu}\right)_T + (A_2 n_2^2 K_2 + A_1 n_1^2 K_1) . \qquad (7)$$

The second term on the right is relatively unimportant on good substrates, where $(\partial A_2/\partial\mu)_T$ is large, and we shall therefore neglect it in the remaining analysis. The dominant term involves a distribution function over area, which we term the "areal density of states" $g_A(\epsilon)$,

$$g_A(\epsilon)\equiv-\frac{1}{A}\frac{dA_2}{d\epsilon_{\text{int}}}\simeq\frac{1}{A}\left(\frac{\partial A_2}{\partial\mu}\right)_T , \tag{8a}$$

where A is the total area. $g_A(\epsilon)$ is characteristic of the film-substrate combination, independent of density and temperature. Formally, A_2 is the area covered by the adatoms subject to the condition that the local substrate interaction energy $\epsilon_l(x,y)$ be less than the energy ϵ (choosing phase 2 arbitrarily as the more dense phase). Thus,

$$A_2(\epsilon)=\int_A\int dxdy\,\theta[\epsilon-\epsilon_l(x,y)] ,$$

where θ is a unit step function, and

$$\begin{aligned}g_A(\epsilon)&\equiv-\frac{1}{A}\frac{dA_2}{d\epsilon_{\text{int}}}=\frac{1}{A}\frac{dA_2}{d\epsilon}\\&=\frac{1}{A}\int_A\int dxdy\,\delta[\epsilon-\epsilon_l(x,y)] .\end{aligned} \tag{8b}$$

It is the expression on the right-hand side of Eq. (8b) which appears naturally in the formal treatment (see Appendix). This characterization, in terms of $g_A(\epsilon)$, can be further broadened by a simple physical ansatz.

We assume that the energy heterogeneity of different adatoms having the same class of substrate interaction can be scaled proportional to their average binding energy ϵ_0. The rationale for the assumption is that most substrate imperfections can be modeled as distortions in their spatial arrangements, and therefore should produce the same fractional changes in the interactions of the substrate with different adatoms. Subsequently in this paper we shall bring evidence to bear supporting the assumption. We therefore expect that the functional form of $g_A(\epsilon)$ is

$$g_A(\epsilon)=\frac{1}{\epsilon_0}g_A^*(\epsilon/\epsilon_0) , \tag{9}$$

where the "reduced areal density of states" g^* is the same function for different adatoms on the same substrate. This form can be derived directly from Eq. (8b) if the only difference between adatom-substrate energy functions $\epsilon_l(x,y)$ is the strength of the interaction, i.e., if there is a common length scale for all combinations. In this case ϵ_0 is proportional to that interaction strength. However, we expect the form (9) to be more general than this derivation would indicate, if ϵ_0 is taken to be the average binding energy.

Combining Eqs. (6), (7), and (9), we obtain the reduced areal density of states g_A^* in terms of the isotherm slope:

$$g_A^*(\epsilon/\epsilon_0)\simeq\frac{N\epsilon_0}{(n_2-n_1)A}G(\mu) . \tag{10}$$

B. Size effects

Finite-size effects include a variety of distinct phenomena affecting films. Limited crystallite and platelet dimensions truncate long-range interactions and spatial coherence. Grain boundaries and crystallite edges disturb the substrate potential, making it locally larger or smaller than in the interior of a large perfect domain. The local anomalies are condensation nuclei for phase condensation and the growth of two-dimensional islands. Crevices and regions of negative curvature at grain boundaries and steps cause capillary condensation of phases at pressures below their saturation value.

Quantitative estimates of island and domain sizes have been obtained from measurements of heat capacity,[7–10] vapor pressure,[11,12] two-dimensional pressure,[13] neutron scattering,[14] low-energy electron diffraction,[12] and x ray diffraction.[15] In one study[12] combining vapor pressure and low-energy electron diffraction (LEED) measurements it was possible to estimate the relative importance of energy and size effects, and to conclude for one sample that size effects were dominant. In all of the remaining studies the possibility of energy heterogeneity was neglected at the outset. For the majority the analysis was limited to an estimate of the average size L of a typical domain, but one detailed comparison between experiment and model calculations yielded an empirical distribution of domain sizes. It has been comforting that the estimates of L obtained by different techniques on similar substrates are within an order of magnitude of each other. However on closer inspection we note that some estimates differ by as much as a factor of 4 or 5, which may indicate the inadequacies of the basic models.

A substantial advance in substrate characterization can be made with detailed direct measurements of the size distributions, in various film phases and at different stages of growth. The size distribution can then be combined with isotherm measurements to yield a density of states independent of size effects. The most direct technique would seem to be diffraction, of neutrons, x rays, or electrons, where line broadening and intensity changes can be caused by small 2D grain size. Finer details than the average L are difficult to obtain except under conditions of extremely high instrumental resolution, but they can in principle be obtained. In this paper we will assume that the size distribution is known. Its form will be taken as the distribution function, defined by

$$h(l)\equiv\left(\frac{\partial\ln N}{\partial l}\right)_{T,A} , \tag{11}$$

where N is the number of atoms enclosed in domains of diameter l or greater.

In the analysis of line broadening to obtain $h(l)$ it is important to note that line broadening can also be caused by energy heterogeneity, which causes variation in film density. It does not affect registered phases however, since the structure is locked to the substrate, hence for these phases the line shapes can yield $h(l)$ directly. Energy heterogeneity plays a minor role in line broadening of typical incommensurate solid phases, where two-dimensional compressibilities are small. Where necessary, first-order corrections to account for this contribution to linewidth can be made following the analysis of isotherms for $g_A^*(\epsilon/\epsilon_0)$.

The size distribution is to some extent independent of the magnitudes of atom-substrate interaction and is therefore directly applicable to a wide class. However, it cannot be presumed to be the same for films of distinctly different states of registry, as for instance commensurate versus incommensurate phases. These regimes probably differ strongly in their reaction to substrate crystallite order, are likely to have markedly different $h(l)$'s. Therefore the characterization for size should be carried out with films of the same type of registry as the phase under study.

Finite crystallite size plays a role in the thermodynamics of two-phase regimes, and it therefore affects the isotherms used to characterize the substrate. If the interfacial line tension between the two surface phases is σ, the difference $\delta\mu$ between the chemical potentials of the finite crystallite regime and an infinite system is[12,16]

$$\delta\mu = 2\sigma[l(n_2-n_1)]^{-1} \ . \tag{12}$$

This shift must be added to μ_{film} in Eq. (2), so that the equilibrium condition at the interface becomes

$$\mu_{\text{vapor}}(P,T) = -\epsilon_{\text{int}} + \mu_2(\phi_{\text{int}},T) + 2\sigma[l(n_2-n_1)]^{-1} \ . \tag{13}$$

Differentiating with respect to N, we obtain a relation between $G(\mu)$, the density of sites and $h(l)$.

$$[g_A(\epsilon)]^{-1} = (n_2-n_1)A[NG(\mu)]^{-1} + 2\sigma A[Nl^2 h(l)]^{-1} \ . \tag{14}$$

Thus, isotherm and line-broadening measurements combined can yield the size and energy distributions. These can serve to characterize the substrate for a class of adsorbates and regimes, i.e., for all adsorbates subject to the same type of long-range attractions to the substrate, in thermodynamic regimes similar to those used for characterization.

III. EFFECTS OF HETEROGENEITY

In this section we examine the effects of heterogeneity on different thermodynamic regimes. It will be assumed that the distribution functions g^* and h appropriate to the regime are known. The discussion is largely given in general terms, but some specific examples are treated in detail.

A. Homogeneous phases

The simplest regime is one which ideally is an infinite uniform phase. On a real substrate the film has nonuniform density and is effectively divided into small domains. These perturbations can in principle affect many thermodynamic properties.

The density variations can be described by means of a distribution $F(\bar{n},n,T)$ such that the fractional number of atoms whose local density lies between n and $n+dn$, when the average density N/A is $\bar{n}$ and the temperature is T, is given by

$$\frac{dN}{N} = F(\bar{n},n,T)\,dn \ . \tag{15}$$

The local density variation $dn = n^2 K d\mu_2$ at constant temperature is related to the local change in substrate energy, i.e., $d\mu_2 = -d\epsilon$. Therefore, $F = -(1/Nn^2K) \times (\partial N/\partial\epsilon)_T$. If the film is relatively incompressible, $dN = \bar{n}dA$, so that $(\partial N/\partial\epsilon)_T = \bar{n}(\partial A/\partial\epsilon)_T = -Ng_A(\epsilon)$, according to Eq. (8). This, together with Eq. (9) gives

$$F(\bar{n},n,T) = g^*(\epsilon/\epsilon_0)/\epsilon_0 n^2 K(n,T) \ , \tag{16}$$

where $K(n,T)$ is the local isothermal compressibility, and ϵ is explicitly given, as a function of $\bar{n}$, n, and T, by

$$\epsilon = \mu(\bar{n},T) - \mu_2(n,T) \ . \tag{17}$$

Typical local thermodynamic properties are functions of n and T. For a local property $X(n,T)$ the heterogeneous sample average is

$$\bar{X}(\bar{n},T) = \int dn F(\bar{n},n,T) X(n,T) \ , \tag{18}$$

with

$$\bar{n} = \int dn F(\bar{n},n,T)\,n \ ,$$
$$1 = \int dn F(\bar{n},n,T) \ . \tag{19}$$

Although our heuristic discussion appears to apply only to the case of homogeneous phases, it is shown in the Appendix that, at least for the case of energy heterogeneity alone, the distribution F given by Eq. (16) is quite general. Some care must be taken in expressing thermodynamic derivatives with respect to $\bar{n}$ and T. From the detailed treatment in the Appendix, it can be shown that the averaged compressibility

$K^{\mathrm{av}} \equiv (1/\bar{n}^2)(\partial\bar{n}/\partial\mu)_T$, thermal expansion coefficient $\alpha^{\mathrm{av}} \equiv -(1/\bar{n})(\partial\bar{n}/\partial T)_{\bar{\phi}}$, and specific heat at constant area $C_A^{\mathrm{av}} \equiv (T/\bar{n})(\partial\bar{s}/\partial T)_{\bar{n}}$ are given by

$$\begin{aligned} K^{\mathrm{av}} &= \frac{1}{\bar{n}^2} \int dn F(\bar{n},n,T)\, n^2 K(n,T) \ , \\ \frac{\alpha^{\mathrm{av}}}{K^{\mathrm{av}}} &= \bar{s} - \bar{n} \int dn \left(\frac{\partial F}{\partial \bar{n}}\right)_{n,T} s(n,T) \ , \\ C_A^{\mathrm{av}} &= \frac{1}{\bar{n}} \int dn F(\bar{n},n,T)\, n C_A(n,T) \\ &\quad + \frac{T}{\bar{n}} \int dn \left(\frac{\partial F}{\partial T}\right)_{n,\bar{n}} s(n,T) \ , \end{aligned} \tag{20}$$

where the local compressibility, entropy density, and specific heat appear in the integrals. Certain linear combinations of these sample-averaged thermodynamic derivatives can be expressed directly in terms of F itself, rather than in terms of temperature and density derivatives of F (see Appendix).

Of particular interest is the heat capacity. The additional term in Eq. (20) can be interpreted physically as the effect of density redistribution. That is, redistribution among the regions of differing binding energy can produce an appreciable contribution, as discussed in a recent paper.[17] This heterogeneity term is unimportant in relatively incompressible dense phases, but may become important at low densities. Recently it has been shown[17] that a first-order correction for the effect in the heat capacities of low-coverage He on Grafoil makes a significant change in the deduced values of the virial coefficient.

Size effects can occur in various properties of homogeneous phases. In addition to diffraction line broadening, which has been discussed, limited sizes may affect transport phenomena: spin[18] and mass diffusion, and sound and heat transport are all limited by imperfections and crystallite boundaries. Phases with especially large coherence lengths are particularly sensitive. Outstanding examples of such systems are films in the vicinity of higher-order phase transitions. We postpone discussion of this topic to Sec. III D below, where it will be taken up in detail. A different regime sensitive to size limitation is an incommensurate solid at low temperature. At progressively lower T the principal phonons contributing to the heat capacity have longer wavelengths, varying roughly as $\lambda \sim n^{1/2}\theta/T$, where θ is the two-dimensional Debye temperature. As T is decreased λ sweeps through the size distribution; as more modes become lost to the frequency distribution their contribution to the heat capacity is lost, and the heat capacity falls below the 2D Debye law.

B. First-order phase transitions

Among the most serious effects of substrate heterogeneity are their blurring of phase transitions. On an ideal substrate a first-order phase change is marked by discontinuities in the first derivatives of the thermodynamic state functions, but on a nonuniform surface the transition is rounded into a continuous change. The rounding is caused by both binding energy variations and size effects.

The rounding of the ideally sharp corners of vapor pressure isotherms has already been discussed, in the section on characterization. A related thermodynamic property, the specific heat, will be analyzed at some length, with application to experimental results.

The heat capacity of a two-phase system can be expressed as the sum of three terms: the contributions of each phase and a conversion term. The conversion term is[6,19]

$$\begin{aligned} C_{\mathrm{conv}} = T &\left[\left(\frac{\partial S_1}{\partial N_1}\right)_{T,A_1} - \left(\frac{\partial S_2}{\partial N_2}\right)_{T,A_2}\right]\left(\frac{\partial N_1}{\partial T}\right)_{N,A} \\ &+ T\left[\left(\frac{\partial S_1}{\partial A_1}\right)_{T,N_1} - \left(\frac{\partial S_2}{\partial A_2}\right)_{T,N_2}\right]\left(\frac{\partial A_1}{\partial T}\right)_{N,A} \ , \end{aligned} \tag{21}$$

where S_1, S_2 are the total entropies of the individual phases. This term is not present on the single-phase side of the boundary line on the phase diagram, hence the signature of the first-order transition would be a heat capacity discontinuity due to the sudden appearance or disappearance of C_{conv}. The discontinuity can be expressed as follows. Let the function $T'(n)$ represent the phase boundary in the local system. The boundary can, for example, be the line dividing the homogeneous vapor phase from coexisting vapor and liquid, as shown in Fig. 1. As one crosses this

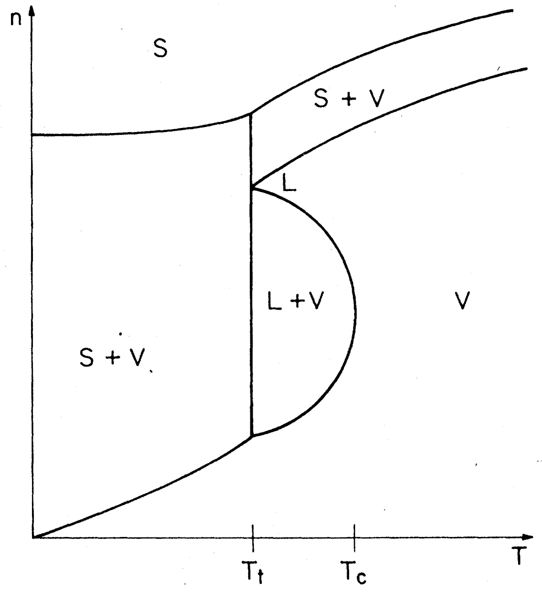

FIG. 1. Phase diagram of a monolayer with first-order transitions between solid, liquid, and vapor phases.

phase boundary from slightly below $T'(n)$ to slightly above $T'(n)$ (at fixed n), the local heat capacity exhibits a discontinuity given by C_{conv}. Explicitly, the discontinuity is

$$\left(\frac{C}{A}\right)_{T=T'(n)-0^+} - \left(\frac{C}{A}\right)_{T=T'(n)+0^+} = \frac{T'(n)}{K^{(+)}} \frac{1}{\left(n\frac{dT'(n)}{dn}\right)^2} , \quad (22)$$

where $K^{(+)}$ is the isothermal compressibility $K(n,T'(n)+0^+)$ on the single-phase side of the boundary. The calculated magnitude of this term in typical systems is large compared to other terms. Nevertheless, no discontinuities have been observed, even in systems in which vapor-pressure isotherms give strong indications that first-order phase boundaries do exist. Their absence can be explained by heterogeneity.

We consider an approach to the transition from either the single or the two-phase side. The boundary line in the ideal system is a specific function $T'(n)$ of the density. In a system with average density $\bar{n}$ and local density variations (described by the distribution F discussed previously) there will be a spread in boundary temperatures around $T'(\bar{n})$. Depending on the shape of the local density variations, i.e., on the nature of F, the transition will show a reduced and/or rounded anomaly in the neighborhood of $T'(\bar{n})$; if F is sufficiently broad (i.e., gives significant weight to n values far from $\bar{n}$) there will be no observable remnant of the ideal discontinuity. The shapes of anomalies can be calculated if the transition line slope and $g_A(\epsilon)$ are known. Although complete characterizations in the manner described here have not been made, some information on the heterogeneity of commonly used substrates is available, and can be applied to the problem. However first we calculate the sensitivity of different transitions to energy heterogeneity.

We define the specific sensitivity to heterogeneity as $S(T')$,

$$S(T') \equiv k_B \left| \frac{dT'}{d\epsilon} \right| , \quad (23)$$

i.e., the shift of transition temperature with respect to variations in binding energy, measured in units of $\delta\epsilon/k_B$. With Eq. (6) and $n\,d\mu = -n\,d\epsilon \cong d\phi$ we can relate $S(T')$ to the compressibility and the slope of the boundary:

$$S(T') = +k_B n^2 K \left| \frac{dT'}{dn} \right| . \quad (24)$$

Using Eq. (24) one can calculate $S(T')$ of systems for which detailed vapor-pressure isotherms are available. Two particularly interesting systems are Kr/graphite and Xe/graphite, both of which appear to have solid-vapor and solid-hypercritical fluid regimes according to vapor-pressure isotherms,[20,21] yet neither of the systems have displayed the expected heat-capacity discontinuities, nor even rounded anomalies (except in one case, discussed below), at their evaporation or melting phase boundaries.[22,23] Part of the reason for the discrepancy is the fact that the isotherm measurements were made with uncompressed graphite powder, which appears to have exceptionally good uniformity, while the heat-capacity measurements were made with Grafoil, a compressed powder of lower quality. Another contributing factor is the high sensitivity of the transition temperatures to energy heterogeneity. We have calculated $S(T')$ for each of the transitions, and list the results in Table. I. The results range from 0.05 to 1; at the upper value a spread of 1 K in transition temperature would be caused by a variation of about 1 k_B in the binding, i.e. less than 0.1% variation in ϵ.[19] Therefore the detection of experimental signals approaching the form of ideal transitions demands extreme substrate uniformity; if not of the entire sample, at least an appreciable part of it. The one exception to the complete absence of anomalies, alluded to earlier, is the melting transition of Kr. Heat capacities at coverages from 0.8 to ~1 layer show broad anomalies[22] at temperatures lying along a line consistent with the boundary traced out by vapor-pressure isotherms.[21] The appearance of these experimental signals, as contrasted with the absence of corresponding signals in Xe or with anomalies due to condensation in either system, is consistent with the appreciably lower sensitivity of the Kr melting transition. It is particularly interesting that, although there are alternative explanations for the shape of the Kr peak,[10,22] the experimental anomaly has a shape which would result from the smearing of an ideal mesalike signal[6] due to the crossing of a first-order melting regime.

Also listed in Table I are the sensitivities of the condensation and incommensurate solid melting transitions of ^{4}He on graphite. (For this discussion we

TABLE I. Specific sensitivity to energy heterogeneity of phase transitions $S(T') = k_B dT'/d\epsilon$, various films on graphite.

Film	Condensation	Melting
Kr	0.2	0.05
Xe	1.0	0.3
^{4}He	0.08	0.07

do not distinguish whether the phase change at melting is first order or continuous.) These values, calculated from the phase boundary slopes[7] and compressibilities,[4] are within the range covered by the Kr and Xe values, yet they contrast with the heavier atoms in that the ^{4}He heat capacities on Grafoil[7,24] do show pronounced anomalies even along the condensation line (see Fig. 12 of Ref. 7). The difference can be due to a combination of two effects. First, as seen in Table I, the sensitivity of the helium condensation is appreciably lower. Also important is the much lower total binding energy of helium, approximately one-tenth of the Kr value.[4] A common $g_A^*(\epsilon)$ for all of the systems on the same type of substrate produces lateral fields only one-tenth as large in helium; it is these fields which cause the density variations that blur the transitions. Combining factors, and scaling from the experimental widths ~0.5 K of the ^{4}He condensation anomalies, predicts that the widths of the Kr and Xe peaks would be 25 and 50 K, respectively: it is therefore not surprising that no evidence has been found of those transitions in the heat-capacity studies. Scaling the melting transitions in the same way, a width of ~0.5 K of the ^{4}He peaks indicates a corresponding width of 3.5 K for Kr and over 20 K for Xe, which is consistent with the observation of a broad Kr anomaly but nothing apparent for Xe.

The comparisons lend support to the assumption that the reduced density of states can characterize a substrate for classes of adsorbates, and that $g_A(\epsilon)$ is equal to $g_A^*(\epsilon/\epsilon_0)/\epsilon_0$.

Finally, we note that the experimental width of the ^{4}He condensation specific-heat anomaly, together with the calculated $S(T')$ of the transition, indicates that the binding energy variation of He on Grafoil is on the order of $6k_B$.

C. Triple points

"Two-dimensional triple points" have been deduced from studies of several monolayer systems, by vapor-pressure isotherms,[20,21,25–28] neutron scattering,[29,30] x ray diffraction,[15] and calorimetry.[23,31–33] The specific-heat results are particularly striking: Ne, O_2, and Xe films on graphite showing strong sharp peaks with widths as narrow as $\delta T \simeq 0.003$,[23] limited by instrumental resolution. These sharp peaks are in contrast with absent or broad anomalies at the two-phase boundaries in the same films. We believe that the explanation involves the spatial distribution of heterogeneity coupled with incomplete equilibrium in the experiments.

Our discussion is given with reference to the ideal phase diagram of Fig. 1, which describes the situation in a uniform film having a normal triple point. Although the temperature T_t is independent of the average film density, this alone is not enough to explain the insensitivity of T_t to heterogeneity. The triple point is a fixed point, independent of the relative amounts of the three surface phases. This independence has nothing to do with the stability of n against the lateral fields of a heterogeneous substrate. Fields cause density variations in any phase according to its compressibility, whether or not there is another, coexisting phase. Thus, the triple point of the heterogeneous film can be experienced by only that part of the film in zero binding-energy gradient, where the densities of the coexisting phases correspond to their values in the uniform system. Since the experiments show that major fractions of the films undergo sharp triple point melting, it follows that at T_t at least, substantial parts of the systems are on an effectively uniform surface. The conclusion appears to contradict the evidence for strong heterogeneity at other phase boundaries. We propose the following explanation.

Let us assume that the spatial distribution of binding energy is in the form of terraces, of uniform regions, their dimensions being consistent with size effects. Such patchwise-uniform or "homotattic" structures have been employed as models for calculations of nonuniform films.[34,35] In a uniform phase the density of a film in equilibrium on such a substrate will vary from patch to patch according to the energy variation and the 2D equation of state. The density profile follows the spatial variation of ϵ on the patches, and the density distribution $F(n)$ is the same (except for size effects) as that on any substrate having the same $g(\epsilon)$. If a film on such a substrate is cooled below T_c it divides into two phases on each patch with relative amounts according to the local average density. On further cooling each uniform two-phase region passes through the ideal triple point.

However, this explanation requires an additional condition: that the equilibrium between different regions be relatively slow at or near the triple point. If the system were in complete thermodynamic equilibrium in the two-phase regime, some atoms would have distilled from regions of weaker to stronger binding, until in the complete equilibrium state the dense phase would cover all of the strong-binding regions, with $\infty > \epsilon > \epsilon_{\text{int}}$, and the low density phase covering all of the rest. The interface would be some relatively short boundary, the number of atoms on the boundary being on the order of $N^{1/2}$ for a continuous $g(\epsilon)$. In these circumstances the only part of the film capable of triple point melting would be the very small fraction at the interface. The rest of the film would be at lower or higher spreading pressures, and hence lower or higher densities than the vapor and liquid in two-phase equilibrium along the interface. These phases would undergo condensation and melting transitions broadened by heterogeneity in the manner described in the preceding section.

Therefore the explanation requires that, although the film is in equilibrium on each patch, the equilibrium is incomplete between patches. This situation is plausible due to the typically low vapor times the long diffusion times at T_t in the absorbents used in calorimetry.

In contrast to the Ne, O_2, and Xe results, other films have not shown resolution limited heat capacities. Experimental widths, at intermediate coverages where the triple point might be expected, range from $\delta T/T \simeq 0.02$ for Kr (Ref. 22) and N_2 (Ref. 36) to 0.3 for Ar (Ref. 37) (see Fig. 2 and Table II). The broad anomalies were obtained with the same substrates as used for the Ne, O_2, and Xe films.

While it is possible that the broad anomalies are due to continuous melting, as has been suggested on the basis of several different mechanisms,[10,22,38] it seems possible that some of the broad transitions are intrinsically first order, but seriously affected by heterogeneity. Neutron diffraction shows appreciable size broadening in the films on Grafoil substrates. In N_2 the mean domain size of the registered solid[14] is consistent with the values deduced from the specific-heat transition widths due to line tension and a size distribution.[36] In Ar films the linewidths are much greater.[39] Since the Ar solid is incommensurate the width must be due to a combination of size broadening and density variation. Density variation may play a much more important role in films having a very limited liquid density range. In the case of Kr, for example, whose phase diagram is quite detailed,[21] the liquid range is quite narrow. Assuming that the liquid does exist, a moderate spread of density caused by energy heterogeneity could reach well beyond the entire coexistence region. Under these circumstances a sharp triple point signal due to the minor fraction of the film could be masked by the discontinuities due to transitions in the low-density and compressed film region. N_2 films appear to have a similar phase diagram. The phase diagram of Ar is not sufficiently known to test the mechanism.

TABLE II. Abruptness of solid-fluid transitions at "triple points" of experimental monolayers on Grafoil and Papyex graphite (see Fig. 2).

Atom	$\delta T_{\text{width}}/T$	$\delta T_{\text{shift}}/T$
Xe[a]	≤ 0.003	≤ 0.001
Ne[b]	0.006	≤ 0.006
O_2[c]	0.012	0.004
CD_4[d]	0.013	
N_2[e]	0.02	0.04
Kr[f]	0.02	0.024
Ar[g]	0.3	0.04

[a] Reference 23.
[b] Reference 31.
[c] Reference 32.
[d] Reference 29.
[e] Reference 36.
[f] Reference 22.
[g] Reference 37.

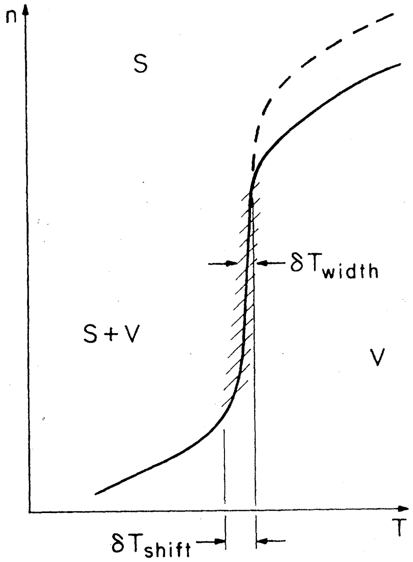

FIG. 2. Experimental phase diagram of monolayers with diffuse solid-vapor transitions (see Table II). The dotted line corresponds to a line of heat-capacity peaks seen in Kr/graphite films.

D. Second-order phase transitions

The study of second-order phase transitions in adsorbed films is a very active field of experiment and theory. Several types of phase change are known or presently speculated to be second order: the gas-liquid critical point, order-disorder transitions, magnetic ordering, commensurate-incommensurate transitions, melting of incommensurate solids, and superfluid transitions. In some of these cases the question of whether the transition is first or second order is of prime interest; in others the determination of critical exponents is a principal objective. To the extent that heterogeneity can change the apparent critical exponents and even the apparent order of the transition, the careful characterization of substrates is crucial to the field. A detailed prescription for removing the effects of heterogeneity from experimental data is given in Sec. IV, and it can be applied to critical regions as well as others. Here we give a general discussion, and investigate some specific examples.

In all second-order transitions the correlation radius r_c and generalized susceptibility χ diverge as the critical point is approached[16]:

$$r_c \propto |t|^{-\nu}; \quad \chi \propto |t|^{-\gamma} , \tag{25}$$

where $t \equiv (T - T_c)/T_c$ and the critical exponents $\nu, \gamma > 0$. The increase of r_c as $T \rightarrow T_c$ is truncated by the finite sizes of crystallites, so that the temperature region unaffected by size limitation does not extend all the way to T_c. Energy heterogeneity also limits critical behavior, but sensitivity to energy heterogeneity depends on the specific type of transition. Where the relevant susceptibility is the density compressibility K, as in the gas-liquid critical point, energy heterogeneity couples directly to the critical behavior. In these cases the excluded critical region becomes a combined result of the divergence of both χ and r_c. The range of temperatures excluded from ideal behavior can be scaled in terms of the characteristics of the substrate. For the cases where $\delta\epsilon$ is the appropriate field for the divergent susceptibility, a general scaling law allows us to relate $\delta\epsilon$ to t_{ex}, the limiting temperature of ideal behavior. The general scaling law is[16]

$$\delta\epsilon \propto |t_{\text{ex}}|^{\beta+\gamma} \ , \tag{26}$$

where β is the critical exponent of the order parameter. Finite size of crystallites plays an independent role; from Eq. (25) we obtain for a typical dimension l,

$$l \propto |t_{\text{ex}}|^{-\nu} \ . \tag{27}$$

The combination of energy heterogeneity and finite size produces a scaling law of the form

$$t_{\text{ex}} = a\,|\delta\epsilon|^{1/(\beta+\gamma)} + b\,|l|^{-1/\nu} \ , \tag{28}$$

where a and b are nonuniversal constants. Comparisons between different well characterized substrates can presumably yield the kind of detailed information that will allow the determination of the coefficients in Eq. (28); about all that can be concluded at this time is that no such detailed comparisons have yet been made.

In transitions for which the relevant susceptibility does not couple to the energy heterogeneity, variations in ϵ can still cause appreciable blurring of the transition. The mechanism has been described in Sec. III B, whereby induced density variations produce a distribution of transition temperatures. The blurring of the transition in these cases depends on the value of K and the slope of the phase boundary. In some of the transitions, e.g., the gas-liquid critical point and the order-disorder transition of He/graphite, the slope $dT_c/dn = 0$ at $n = n_c$. In these cases the spread in T_c is due only to the curvature of the boundary, and is markedly reduced from the typical spread of transition temperatures seen in other transitions, e.g., condensation. This appears to be one of the factors contributing to the striking sharpness of the specific-heat peaks of the He/graphite transition.[7,9,24,40]

The peak rounding and finite heights of the specific heats of He on various forms of graphite have been interpreted as due solely to size effects.[7,9,40] However, from our earlier discussion, we consider it virtually certain that energy heterogeneity must also contribute to the rounding. There is in fact evidence that in the very uniform ZYX graphite[41] size effects are less important than energy heterogeneity. In recent high-resolution x-ray studies of the registered Kr solid on ZYX graphite, diffraction linewidths indicate surface coherence lengths of more then 2000 Å,[15,42] as much as 10 times the lengths deduced from the He order-disorder transition heat capacities on the same type of substrate.[9] While it is conceivable that the quality of the two different ZYX samples is so different, we think it much more likely that the discrepancy is a matter of interpretation. The x-ray linewidth of a registered phase is subject to size broadening but not to energy heterogeneity, while the transition temperature rounding is subject to both energy and size effects. The discrepancy between the two interpretations indicates that, contrary to common belief, it is energy heterogeneity and not size limitation that is dominant.

The most familiar second-order transition is the gas-liquid critical point. Several films have phase diagrams obtained from vapor-pressure isotherms that indicate liquid-vapor coexistence regions and hence 2D critical points. However in none of these systems has any singular behavior been seen, as for example, divergent heat capacities. Their absence can be explained by heterogeneity, specifically energy heterogeneity, coupled with divergent compressibilities. Even if very much more uniform substrates could be obtained the divergence of K as $T \rightarrow T_c$ makes the critical region particularly difficult to observe. Apart from the critical region, the dependence of $G(\mu)$ on the densities of the liquid and vapor phases helps to explain some features seen in vapor-pressure isotherms. In Eq. (10) we see that $G(\mu)$ is proportional to $(n_2 - n_1)$. Since $g_A^*(\epsilon)$ is a constant, if follows that the slope of the isotherms in the two-phase region should decrease monotonically with increasing temperature. This is observed in many of the systems.[43–45]

IV. INVERSION PROBLEM

Experimental quantities will in general be related to the averages $\langle X \rangle$, defined in the Appendix [Eqs. (A16)–(A18)]. In cases where a two-phase region exists locally, one must make a distinction between the average $\bar{X}$ and the average $\langle X \rangle$. As described in the Appendix, the two averages are equal for local quantities X which are finite in the two-phase region, but $\bar{X}$ is not appropriate for all quantities when the film is in a two-phase coexistence regime. Instead, it is the average

$$\langle X\rangle = \int_0^{n_{\max}} dn F(\bar{n},n,T)X^{\rm loc}(n,T) \qquad (29)$$

over all densities (including the interface region) which is appropriate in all cases. The function F is given explicitly by

$$F(\bar{n},n,T) \equiv \frac{g[\mu(\bar{n},T)-\mu_2(n,T)]}{n^2K(n,T)} \quad . \qquad (30)$$

We would like to construct a simple inversion scheme for extracting the local quantity from the average; i.e., for obtaining information about the local variables from the experimental averages. We suppose here that the functional form of $g(\epsilon)$ is known, that $\mu(\bar{n},T)$ has been determined experimentally, and that $\mu_2(n,T)$ (and K) are supplied by some simple local model for the film. The form discussed in Sec. III, $g(\epsilon)=(1/\epsilon_0)g^*(\epsilon/\epsilon_0)$, may be used, but is not a necessary condition for the analysis below.

Every "average" property can be related to an "F" average of a local property. Specifically, if we regard $C_A^{\rm av}$, $\alpha^{\rm av}$, and $K^{\rm av}$ as an appropriate set of average quantities, then it follows from Eqs. (A24), (A25), and (A26) that

$$\bar{n}^2K^{\rm av}(\bar{n},T) = \langle n^2K\rangle \quad ,$$

$$\bar{n}[\bar{S}K^{\rm av}(\bar{n},T)-\alpha^{\rm av}(\bar{n},T)] = \left\langle n^2K(n,T)\left(\frac{\partial S}{\partial n}\right)_T\right\rangle \quad ,$$

$$\bar{n}C_A^{\rm av}(\bar{n},T)+\frac{T[\bar{S}K_T^{\rm av}(\bar{n},T)-\alpha^{\rm av}(\bar{n},T)]^2}{K^{\rm av}(\bar{n},T)} = \left\langle nC_A+Tn^2K\left[\left(\frac{\partial S}{\partial n}\right)_T\right]^2\right\rangle \quad . \qquad (31)$$

We see that the left-hand side of Eq. (31) is some combination $A^{\rm av}(\bar{n},T)$ of average properties, while the quantity in the $\langle\ \rangle$ on the right-hand side is some local function $B^{\rm loc}(n,T)$. In general, then, the relation between average and local properties is always of the form

$$A^{\rm av}(\bar{n},T) = \langle B^{\rm loc}(n,T)\rangle \equiv \int_0^{n_{\max}} F(\bar{n},n,T)B^{\rm loc}(n,T) \quad . \qquad (32)$$

The traditional derivatives representing constant area specific heat, thermal expansion coefficient, and isothermal compressibility, are not the simplest choices of average and corresponding local quantity. For example, the choices $A^{\rm av}=\bar{n}$, $B^{\rm loc}=n$ or $A^{\rm av}=\bar{s}$, $B^{\rm loc}=s(n,T)$ are obviously simpler cases of the general structure exhibited by Eq. (32).

It is a simple matter to invert Eq. (32) by supposing a grid of experimental densities $\bar{n}$ (all at the same T) with their associated values of $A^{\rm av}$. Let there be Z experimental densities n_i and define the $Z\times Z$ matrix

$$S_{ij} \equiv \left(\frac{n_j}{n_{\max}}\right)^{i-1} \quad , \quad i,j=1,2,\ldots,Z \quad . \qquad (33)$$

The inverse of this matrix provides quadrature weights for the approximate integration required in Eq. (32). Thus,

$$W_i \equiv \sum_{j=1}^{Z}(S^{-1})_{ij}\frac{1}{j} \qquad (34)$$

defines quadrature weights so that the integral in Eq. (32) is approximately given by

$$A_{\rm av}(n_i) = \sum_{j=1}^{Z} Q_{ij}B^{\rm loc}(n_j) \qquad (35)$$

with

$$Q_{ij} \equiv n_{\max}F(n_i,n_j)\,W_j \quad . \qquad (36)$$

The temperature dependence of $A^{\rm av}$, $B^{\rm loc}$, and F is implied. A further matrix inversion of Q now provides explicitly the value of $B^{\rm loc}$. Thus,

$$B^{\rm loc}(n_i) = \sum_{j=1}^{Z}(Q^{-1})_{ij}A^{\rm av}(n_j) \quad , \qquad (37)$$

$$i=1,2,\ldots,Z \quad .$$

It should be noted that simple checks on both the approximate integral inversion and on the validity of the original $F(\bar{n},n)$ are the requirements

$$1=\sum_{j=1}^{Z}Q_{ij} \quad , \quad n_i=\sum_{j=1}^{Z}Q_{ij}n_j \quad . \qquad (38)$$

The first of these follows from the normalization on F, and the second follows from the fact that $A^{\rm av}=\bar{n}$ corresponds to $B^{\rm loc}=n$.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Energy and the National Science Foundation, under Grant No. DMR-7822697.

APPENDIX: AVERAGING PROCESS FOR HETEROGENEOUS SUBSTRATES

For purposes of illustration and motivation, first consider the total number of particles per unit area, as an integral over the local film density

$$\bar{n} = \frac{N}{A} = \frac{1}{A}\int_A\int n[x,y]\,dxdy \quad , \qquad (A1)$$

We regard the film as a two-dimensional system, whose local density $n[x,y]$ is the 2D thermodynamic function $n(\mu_2,T)$, with μ_2 the 2D chemical potential. Thus,

$$n[x,y] = n(\mu_2,T) \tag{A2}$$

with the functional form determined by the 2D equation of state. The temperature T is uniform. The film is locally in equilibrium with 3D vapor, so that

$$\mu = \mu_{\text{vapor}} = \mu_{\text{film}} = \mu_2 + \epsilon_l(x,y) \quad , \tag{A3}$$

where $\epsilon_l(x,y)$ is the locally varying substrate potential energy. The vapor chemical potential μ is a constant, determined by the 3D gas pressure P and temperature T. Thus, the spatial dependence of $\mu_2 = \mu - \epsilon_l(x,y)$ determines the local variation in film density.

Equation (A1) can therefore be written in the form

$$\bar{n} = \int_{-\infty}^{\infty} d\epsilon\, n(\mu-\epsilon,T)\, g_A(\epsilon) \quad , \tag{A4}$$

where

$$g_A(\epsilon) \equiv \frac{l}{A} \int_A \int dxdy\, \delta[\epsilon - \epsilon_l(x,y)];$$

$$\int_{-\infty}^{\infty} d\epsilon\, g_A(\epsilon) = 1 \quad . \tag{A5}$$

By employing Eq. (A2), with a value of μ_2 which is independent of total area, we are implicitly assuming the 2D thermodynamic limit ($A \to \infty$) for the film. Thus, we have in mind here energy-averaging effects only, as distinct from explicit size effects; the limit $g(\epsilon) = \lim_{A\to\infty} g_A(\epsilon)$ is therefore appropriate and will be assumed in the following discussion.

We may rewrite Eq. (A4) in the form

$$\bar{n}(\mu,T) = \int_{-\infty}^{\infty} d\mu_2\, g(\mu-\mu_2)\, n(\mu_2,T) \quad . \tag{A6}$$

The "natural" variables are μ and T. For systems undergoing first-order phase changes, there are regimes of temperature and density where two phases coexist. In these regimes the average density is composed of contributions from each phase, i.e.,

$$\bar{n}(\mu,T) = \int_{-\infty}^{\mu_2^0(T)} d\mu_2\, g(\mu-\mu_2)\, n_1(\mu_2,T) + \int_{\mu_2^0(T)}^{\infty} d\mu_2\, g(\mu-\mu_2)\, n_2(\mu_2,T) \quad , \tag{A7}$$

where the subscript 1 refers to the lower density phase, and $\mu_2^0(T)$ is the common 2D chemical potential of the two phases (the value of μ_2 at the interface).

Now, any local thermodynamic quantity $X(\mu_2,T)$, which is a function of chemical potential and temperature, can be averaged in the same manner. Thus

$$\bar{X} = \int_{-\infty}^{\mu_2^0(T)} d\mu_2\, g(\mu-\mu_2)\, X_1(\mu_2,T) + \int_{\mu_2^0(T)}^{\infty} d\mu_2\, g(\mu-\mu_2)\, X_2(\mu_2,T) \tag{A8}$$

and the derivatives with respect to μ and T become

$$\begin{aligned} \left(\frac{\partial \bar{X}}{\partial T}\right)_{\mu} &= \frac{d\mu_2^0}{dT} g(\mu-\mu_2^0)[X_1(\mu_2^0(T),T) - X_2(\mu_2^0(T),T)] + \overline{\left(\frac{\partial X}{\partial T}\right)_{\mu_2}} \quad , \\ \left(\frac{\partial \bar{X}}{\partial \mu}\right)_{T} &= -g(\mu-\mu_2^0)[X_1(\mu_2^0(T),T) - X_2(\mu_2^0(T),T)] + \overline{\left(\frac{\partial X}{\partial \mu_2}\right)_{T}} \quad , \end{aligned} \tag{A9}$$

and we note that $d\mu_2^0/dT = -(s_1-s_2)/(n_1-n_2)$ for a first-order transition.

Consider now the averages of density n, entropy density s, spreading pressure ϕ, and chemical potential μ_2. For ϕ and μ_2, there is no discontinuity (x_1-x_2) along the phase boundary. Since the 2D variables satisfy the Gibbs-Duhem relation, $n d\mu_2 = -sdT + d\phi$, we have

$$\left(\frac{\partial\bar{\phi}}{\partial T}\right)_{\mu} = \bar{S} \, , \quad \left(\frac{\partial\bar{\phi}}{\partial \mu}\right)_{T} = \bar{n} \, , \quad \left(\frac{\partial\bar{\mu}_2}{\partial T}\right)_{\mu} = 0 \quad ,$$

$$\left(\frac{\partial\bar{\mu}_2}{\partial \mu}\right)_{T} = 1 \quad .$$

Consequently, $\bar{\mu}_2 = \mu + \text{constant}$, $d\mu = d\bar{\mu}_2$, and

$$d\bar{\phi} = \bar{n} d\mu + \bar{S} dT \quad . \tag{A10}$$

In other words, there is a simple Gibbs-Duhem relation between the average quantities $\bar{\phi}$, $\bar{n}$, $\bar{s}$, $\bar{\mu}_2$. The constant relating $\bar{\mu}_2$ and μ is simply the average substrate potential energy

$$\bar{\mu}_2 = \mu - \bar{\epsilon} \quad ,$$

where

$$\bar{\epsilon} = \frac{1}{A} \int_A \epsilon_l(x,y)\, dxdy = \int_{-\infty}^{\infty} \epsilon g(\epsilon)\, d\epsilon \quad . \tag{A11}$$

Note that $\bar{\epsilon}$ is not a function of either temperature or chemical potential.

The local 2D internal energy density (not including substrate energy) is

$$e - \mu_2 n = Ts - \phi \tag{A12}$$

and these local variables also satisfy the thermodynamic identity $de = \mu_2 dn + Tds$. Because the average $\bar{e}$ contains $\overline{\mu_2 n}$, we define an average quantity $\bar{u}$ by

$$\bar{u} \equiv (\bar{e} - \overline{\mu_2 n} + \mu\bar{n}) = \mu\bar{n} + T\bar{S} - \bar{\phi} \quad ,$$

where the second equality follows by inserting the average of e. The differential $d\bar{u}$ becomes, with the use of Eq. (A10), $d\bar{u} = \mu d\bar{n} + Td\,\bar{s}$. The thermodynamic identity thus involves the new variable $\bar{u}$. In summary, we have

$$\begin{aligned} \bar{u} &= \bar{e} - \overline{\mu_2 n} + \mu\bar{n} \ , \\ \bar{u} &= \mu\bar{n} + T\,\bar{s} - \bar{\phi} \ , \\ d\bar{u} &= \mu d\bar{n} + Td\,\bar{s} \ . \end{aligned} \tag{A13}$$

On an ideal substrate [with $\epsilon_l(x,y) = \text{const} = -\epsilon_0$ and $g(\epsilon) = \delta(\epsilon + \epsilon_0)$], we have $\overline{\mu_2 n} = \bar{\mu}_2\bar{n} = (\mu + \epsilon_0)\bar{n}$. Hence $\bar{u} = \bar{e} - \epsilon_0\bar{n}$ on the ideal substrate. More important is the fact that $\bar{u}$ is the *total* energy per unit area. That is, $u(x,y) = e(x,y) + \epsilon_l(x,y)n(x,y)$ is the full local energy density. Since $\epsilon_l(x,y) = \mu - \mu_2(x,y)$, we see that $u = e - \mu_2 n + \mu n$, and its average $\bar{u}$ is the total energy per unit area, including the substrate potential energy. Equations (A13) and (A10) provide a "complete" set of thermodynamic relations for the averaged quantities, identical with the usual relations.

In fact, the "average" relations are those most evident from a purely thermodynamic point of view. Our derivation, starting with the assumption of local equilibrium and locally defined thermodynamic quantities, merely shows in a pedestrian fashion that these assumptions are consistent with the thermodynamics which must be correct for the entire film.

The important thermodynamic derivatives can be constructed from averages of local quantities via Eqs. (A9). For example, the average density derivative with respect to chemical potential is

$$\left(\frac{\partial\bar{n}}{\partial\mu}\right)_T = g(\mu - \mu_2^0(T))[n_2(T) - n_1(T)] + \overline{n^2K} \ , \tag{A14}$$

where K is the 2D compressibility. According to the definition of the averaging process [Eq. (A8)], $\overline{n^2K}$ includes contributions from each phase, and both contributions are small. It is the first term in Eq. (A14) which on the ideal substrate gives a δ function at $\mu - \mu_2^0(T) = -\epsilon_0$, weighted by the difference between phase densities on the phase boundary. On a real surface, the finite slope of the isotherm provides a direct measure of $g(\epsilon)$.

In fact, we may define an average isothermal compressibility by the expression

$$\begin{aligned} K^{\text{av}} \equiv -\frac{1}{A}\left(\frac{\partial A}{\partial\phi}\right)_{N,T} &= \frac{1}{\bar{n}}\left(\frac{\partial\bar{n}}{\partial\bar{\phi}}\right)_T = \frac{1}{\bar{n}^2}\left(\frac{\partial\bar{n}}{\partial\mu}\right)_T \\ &= \frac{1}{\bar{n}^2}(\Delta n g_0 + \overline{n^2K}) \ , \end{aligned} \tag{A15}$$

where $\Delta n = n_2(T) - n_1(T)$ and $g_0 \equiv g(\mu - \mu_2^0(T))$. K^{av} is directly measurable. Furthermore, observe that the average defined by Eq. (A8) can be converted to an integral over density, recognizing that $\mu_2 \to -\infty$ corresponds to $n = 0$ and $\mu_2 \to +\infty$ will correspond to some maximum possible density $n_{\max}$ (assuming a "hard disk" part in the interaction between film atoms). $\mu_2^0(T)$ corresponds to $n_1(T)$ or $n_2(T)$. Thus, the average $\bar{X}$ in the two-phase regime is

$$\begin{aligned} \bar{X} = &\int_0^{n_1(T)} dn \frac{g[\mu - \mu_2(n,T)]}{n^2K(n,T)} X_1(n,\ T) \\ &+ \int_{n_2(T)}^{n_{\max}} dn \frac{g[\mu - \mu_2(n,T)]}{n^2K(n,T)} X_2(n,T) \ . \end{aligned} \tag{A16}$$

This expression excludes explicitly any contribution from the region $\int_{n_1}^{n_2}$, i.e., any contribution due to the existence of the two-phase interface. We anticipate, and will show explicitly below, that an average defined so as to include such contributions will be useful. Therefore, we define the average

$$\langle X\rangle \equiv \int_0^{n_{\max}} dn \frac{g[\mu - \mu_2(n,T)]}{n^2K(n,T)} X(n,T) \tag{A17}$$

so that

$$\langle X\rangle = g_0 \int_{n_1(T)}^{n_2(T)} dn \frac{X(n,T)}{n^2K(n,T)} + \bar{X} \ . \tag{A18}$$

We have used the fact that $\mu_2 = \mu_2^0(T)$ in the two-phase region, so that $g[\mu - \mu_2(n,T)] = g[\mu - \mu_2^0(T)] = g_0$.

Now, it is clear that for any local quantity X which is *finite* in the two-phase region, $\langle X\rangle = \bar{X}$, since $K \to \infty$ in the two-phase region. This includes all the quantities in Eqs. (A10) and (A13). If $X = n^2K$, as in Eq. (A15), we see that $\langle n^2K\rangle = \Delta n g_0 + \overline{n^2K}$. Consequently, Eq. (A15) can be rewritten as

$$\begin{aligned} \bar{n}^2K^{\text{av}} = \left(\frac{\partial\bar{n}}{\partial\mu}\right)_T &= \langle n^2K\rangle \\ &= \int_0^{n_{\max}} dn g[\mu - \mu_2(n,T)] \ . \end{aligned} \tag{A19}$$

If $\mu - \mu_2^0(T)$ is close to the energy most heavily weighted in $g(\epsilon)$, then only the two-phase contribution is important; i.e., the first term in Eq. (A14). If there is no two-phase region, then the original definition of $\bar{X}$ would have included all n values. In other words, the average $\langle X\rangle$ given in Eq. (A17) is appropriate in all cases. Because we have *defined* K^{av} [Eq. (A15)] in analogy with the local definition, K^{av} is $\langle n^2K\rangle/\bar{n}^2$ rather than simply $\langle K\rangle$.

The complete thermodynamic description requires three derivatives. We may take the others, in addition to Eq. (A19) to be

$$\begin{aligned} \left(\frac{\partial\bar{n}}{\partial T}\right)_\mu &= \left(\frac{\partial\bar{S}}{\partial\mu}\right)_T = -\Delta S g_0 + \overline{nK(S - \alpha/K)} \ , \\ \left(\frac{\partial\bar{S}}{\partial T}\right)_\mu &= \frac{(\Delta S)^2}{\Delta n} g_0 + \frac{1}{T}\overline{nC_A} + \overline{K(S - \alpha/K)^2} \ , \end{aligned} \tag{A20}$$

where α is the thermal expansion coefficient and C_A the specific heat at constant area (nC_A is the local heat capacity per unit area) for the 2D film. We have used the definitions $\Delta s = s_1(T) - s_2(T)$ and $\Delta n = n_2(T) - n_1(T)$, together with the fact that $d\mu_2^0(T)/dT = \Delta s/\Delta n$. The fact that each of these derivatives involves a g_0 term plus a $(\bar{\ })$ term follows from Eqs. (A9).

The right-hand side of Eq. (A20) can be expressed most simply in terms of the $\langle\ \rangle$ average. To show this, we observe that in the two-phase region

$$S = [S_2(T)(n - n_1(T)) + S_1(T)(n_2(T) - n)]/\Delta n \ , \tag{A21}$$

while, in general

$$\left(\frac{\partial S(n,T)}{\partial n}\right)_T = \frac{1}{n}\left(S - \frac{\alpha}{K}\right) \tag{A22}$$

so that in the two-phase region

$$\frac{1}{n}\left(S - \frac{\alpha}{K}\right) = -\frac{\Delta S}{\Delta n} \ . \tag{A23}$$

That is, although both α and K diverge in the two-phase region, their ratio is finite and such that $(1/n)(s - \alpha/K)$ is independent of n. From Eq. (A17), we see that the interface term contributes $-\Delta s g_0$ to $\langle nK(s - \alpha/K)\rangle$ and $[(\Delta s)^2/\Delta n]g_0$ to $\langle K(s - \alpha/K)^2\rangle$. Therefore, using Eq. (22), our complete set of derivatives becomes

$$\bar{n}^2 K^{\mathrm{av}} \equiv \left(\frac{\partial \bar{n}}{\partial \mu}\right)_T = \langle n^2 K\rangle = \int_0^{n_{\max}} dng[\mu - \mu_2(n,T)] \ ,$$

$$\begin{aligned} \left(\frac{\partial \bar{n}}{\partial T}\right)_\mu = \left(\frac{\partial \bar{S}}{\partial \mu}\right)_T &= \left\langle nK\left(s - \frac{\alpha}{K}\right)\right\rangle \\ &= \int_0^{n_{\max}} dng[\mu - \mu_2(n,T)]\frac{\partial S(n,T)}{\partial n} \ , \end{aligned} \tag{A24}$$

$$\begin{aligned} \left(\frac{\partial \bar{S}}{\partial T}\right)_\mu &= \frac{1}{T}\langle nC_A\rangle + \left\langle K\left(s - \frac{\alpha}{K}\right)^2\right\rangle \\ &= \frac{1}{T}\langle nC_A\rangle + \int_0^{n_{\max}} dng[\mu - \mu_2(n,T)]\left(\frac{\partial s}{\partial n}\right)^2 \ . \end{aligned}$$

These derivatives are then the most directly related to averages of local quantities, and the first one (related to K^{av}) is also directly accessible experimentally. Notice that it is only the basic derivatives with respect to μ and T which will satisfy

$$\left(\frac{\partial \bar{X}}{\partial \mu}\right)_T = \left\langle\left(\frac{\partial X}{\partial \mu_2}\right)_T\right\rangle ; \quad \left(\frac{\partial \bar{X}}{\partial T}\right)_\mu = \left\langle\left(\frac{\partial X}{\partial T}\right)_{\mu_2}\right\rangle \ .$$

That is, the derivative of the average is not always the average of the derivative. This is obviously true for $K^{\mathrm{av}} = (1/\bar{n}^2)\langle n^2K\rangle$ but in this case the difference is a result of definition; i.e., $(\partial\bar{n}/\partial\mu)_T$ is still equal to the average $\langle(\partial n/\partial\mu_2)_T\rangle$. A more interesting example is the thermal expansion coefficient, which we define by

$$\begin{aligned} \alpha^{\mathrm{av}} \equiv -\frac{1}{\bar{n}}\left(\frac{\partial \bar{n}}{\partial T}\right)_{\bar{\phi}} &= \frac{\bar{S}}{\bar{n}^2}\left(\frac{\partial \bar{n}}{\partial \mu}\right)_T - \frac{1}{\bar{n}}\left(\frac{\partial \bar{n}}{\partial T}\right)_\mu \\ &= \frac{\bar{S}}{\bar{n}^2}\langle n^2K\rangle - \frac{1}{\bar{n}}\left\langle nK\left(s - \frac{\alpha}{K}\right)\right\rangle \\ &= \frac{1}{\bar{n}}\langle n\alpha\rangle + \frac{1}{\bar{n}^2}[\bar{S}\langle n^2K\rangle - \bar{n}\langle snK\rangle] \ . \end{aligned} \tag{A25}$$

The fact that $(1/\bar{n})\langle n\alpha\rangle$ appears instead of $\langle\alpha\rangle$ is again a result of definition. But in addition, there is a term which appears because the product of s and n^2K averages is not equal to the product of n and snK averages.

Another derivative, of more direct interest than α^{av} is the heat capacity, defined by

$$C_A^{\mathrm{tot}} = \left(\frac{\partial A\bar{u}}{\partial T}\right)_{\bar{n},A} = A\left(\frac{\partial \bar{u}}{\partial T}\right)_{\bar{n}} = AT\left(\frac{\partial \bar{S}}{\partial T}\right)_{\bar{n}} \ .$$

The properly averaged specific heat at constant area is then given by

$$\begin{aligned} C_A^{\mathrm{av}} \equiv \frac{T}{\bar{n}}\left(\frac{\partial \bar{S}}{\partial T}\right)_{\bar{n}} &= \frac{T}{\bar{n}}\left[\left(\frac{\partial \bar{S}}{\partial T}\right)_\mu - \left\{\left(\frac{\partial \bar{n}}{\partial T}\right)_\mu\right\}^2 \Big/ \left(\frac{\partial \bar{n}}{\partial \mu}\right)_T\right] \\ &= \frac{1}{\bar{n}}\langle nC_A\rangle + \frac{T}{\bar{n}}\left[\left\langle n^2K_T\left\{\left(\frac{\partial s}{\partial n}\right)_T\right\}^2\right\rangle\right. \\ &\qquad \left. - \left\langle n^2K_T\left(\frac{\partial s}{\partial n}\right)_T\right\rangle^2 \Big/ \langle n^2K_T\rangle\right] \ . \end{aligned} \tag{A26}$$

This expression also shows an explicit in square brackets term analogous to the one in α^{av}. Equation (A26) actually does *not* show a divergence in the ideal substrate limit. The terms containing g_0 appear only in the ratio $\Delta n g_0 \overline{n^2K}/(\Delta n g_0 + \overline{n^2K})$, which is finite even when g_0 is a δ function.

The temperature dependence of the heat capacity, and the blurring effect of energy heterogeneity on a first-order local heat capacity discontinuity, has been discussed in Sec. III B. A related question, which is somewhat simpler, concerns the density dependence when T is low enough so that the ideal film is in two phase equilibrium. In the two-phase region, the local heat capacity per unit area satisfies

$$\frac{C_A}{A}=-nT\frac{d}{dT}\left(\frac{d\mu_0(T)}{dT}\right)+T\frac{d}{dT}\frac{d\phi(T)}{dT}$$

$$=-nT\frac{d}{dT}\left(\frac{\Delta S}{\Delta n}\right)+T\frac{d}{dT}\frac{d\phi(T)}{dT}$$

so that

$$\frac{\partial}{\partial n}\left(\frac{C_A}{A}\right)_T=-T\frac{d}{dT}\left(\frac{\Delta S}{\Delta n}\right) \tag{A27}$$

is a function of T only. This (linear in n) property has been used to characterize the two-phase region. However, the linearity in n is strictly correct only in the case of an ideal substrate.

The effects of energy heterogeneity on the density dependence are most easily seen by taking the $\bar{n}$ derivative of the total heat capacity per unit area. For the general case, we have $C_A^{\text{tot}}/A=T(\partial\bar{s}/\partial T)_{\bar{n}}$ so that

$$\frac{\partial}{\partial\bar{n}}\left(\frac{C_A^{\text{tot}}}{A}\right)_T=T\frac{\partial}{\partial T}\frac{\partial\bar{S}}{\partial\bar{n}}=T\frac{\partial}{\partial T}\left[\left(\frac{\partial\bar{S}}{\partial\mu}\right)_T\bigg/\left(\frac{\partial\bar{n}}{\partial\mu}\right)_T\right]_{\bar{n}}$$

or, from Eqs. (A15) and (A20)

$$\frac{\partial}{\partial\bar{n}}\left(\frac{C_A^{\text{tot}}}{A}\right)_T=-T\frac{\partial}{\partial T}\left(\frac{\Delta S g_0-\overline{n^2K\left(\frac{\partial S}{\partial n}\right)_T}}{\Delta n g_0+\overline{n^2K}}\right)_{\bar{n}}. \tag{A28}$$

We have argued, in applying Eq. (A14) or Eq. (A15), that the gas and liquid phase contributions to $\overline{n^2K}$ are small compared to $\Delta n g_0$ if we are in the region where $\mu-\mu_0(T)$ is close to the energy most heavily weighted in $g(\epsilon)$. $\{g_0=g[\mu-\mu_2^0(T)]$ would be infinite at this energy on the ideal substrate.$\}$ If we can use the same argument for the numerator in Eq. (A28), then the right-hand side becomes $-T(d/dT)(\Delta s/\Delta n)$, identical with the local expression. The point here is that the $\bar{n}$ dependence of C_A^{tot}/A in the approximately linear region will not provide immediate useful information about $g(\epsilon)$. Furthermore, each of the derivatives K^{av}, α^{av}, and C_A^{av} are given by distribution averages over $g(\epsilon)$, but the quantity to be averaged generally involves a mixture of local thermodynamic derivatives. Only K^{av} is simple.

---

[1] For recent reviews and conference proceedings, see *Proceedings of the Colloque sur les Phases Bidimensionelles,* edited by M. Bienfait [J. Phys. (Paris) Colloq. 38, C4-1 (1977)]; *Phase Transitions in Surface Films,* edited by J. G. Dash and J. Ruvalds (Plenum, New York, 1980); *Ordering in Two Dimensions,* edited by S. K. Sinha (American Elsevier, New York, 1981); O. E. Vilches, Ann. Rev. Phys. Chem. 31, 463 (1980).

[2] G. D. Halsey, Adv. Catal. 4, 259 (1952).

[3] L. E. Drain and J. A. Morrison, Trans. Faraday Soc. 48, 316 (1952).

[4] R. L. Elgin and D. L. Goodstein, Phys. Rev. A 9, 2657 (1974).

[5] G. A. Stewart, S. Siegel, and D. L. Goodstein, in *Proceedings of the 13th International Conference on Low Temperature Physics, Boulder, 1972,* edited by R. H. Kropschut and K. D. Timmerhaus (Plenum, New York, 1974).

[6] J. G. Dash, *Films on Solid Surfaces* (Academic, New York, 1975).

[7] M. Bretz, J. G. Dash, D. C. Hickernell, E. O. McLean, and O. E. Vilches, Phys. Rev. A 8, 1589 (1973).

[8] T. T. Chung and J. G. Dash, Surf. Sci. 66, 559 (1977).

[9] M. Bretz, Phys. Rev. Lett. 38, 501 (1977).

[10] D. J. E. Callaway and M. Schick, Phys. Rev. B 23, 3494 (1981).

[11] Y. Larher, Mol. Phys. 38, 789 (1979).

[12] U. Bardi, A. Glachant, and M. Bienfait, Surf. Sci. 97, 137 (1980).

[13] J. G. Dash, J. Suzanne, H. Schechter, and R. E. Peierls, Surf. Sci. 60, 411 (1976).

[14] J. K. Kjems, L. Passell, H. Taub, A. D. Novaco, and J. G. Dash, Phys. Rev. B 13, 1446 (1976).

[15] R. J. Birgenau, E. M. Hammons, P. Heiney, P. W. Stephens, and P. M. Horn, in *Ordering in Two Dimensions,* edited by S. K. Sinha (American Elsevier, New York, 1981).

[16] L. Landau and E. M. Lifshitz, *Statistical Physics,* 3rd ed. (Pergamon, New York, 1980), Part I. We have adapted the analysis of surface-tension effects in bulk phases to two dimensions. In Ref. 12 the equation given for $\delta\mu$ neglects n_1, the density of the less-dense phase. The error is unimportant for 2D gas phases at relatively low temperatures, but can be important under other circumstances.

[17] A. F. Silva-Moreira, J. Codona, and D. L. Goodstein, Phys. Lett. 76A, 324 (1980).

[18] M. G. Richards, in *Phase Transitions in Surface Films,* edited by J. G. Dash and J. Ruvalds (Plenum, New York, 1980).

[19] The second term in Eq. (21) was omitted in Ref. 6.

[20] A. Thomy and X. Duval, J. Chem. Phys. 66, 1966 (1967); 67 1101 (1970).

[21] Y. Larher, J. Chem. Soc. Faraday Trans. 1 70, 320 (1974); Y. Larher and A. Terlain, J. Chem. Phys. 72, 1052 (1980).

[22] D. M. Butler, J. A. Litzinger, G. A. Stewart, and R. B. Griffiths, Phys. Rev. Lett. 42, 1289 (1979); D. M. Butler, J. A. Litzinger, and G. A. Stewart, *ibid.* 44, 466 (1980).

[23] J. A. Litzinger and G. A. Stewart, in *Ordering in Two Dimensions,* edited by S. K. Sinha (American Elsevier, New York, 1981).

[24] S. V. Hering, S. W. Van Sciver, and O. E. Vilches, J. Low Temp. Phys. 25, 793 (1976).

[25] M. Matecki, A. Thomy, and X. Duval, Surf. Sci. 75, 142 (1978).

[26] C. Bockel, A. Thomy, and X. Duval, Surf. Sci. 90, 109 (1979).

[27] A. Enault and Y. Larher, Phys. Lett. 48A, 213 (1974); Surf. Sci. 62, 233 (1977).

[28] C. Tessier and Y. Larher, in *Ordering in Two Dimensions*, edited by S. K. Sinha (American Elsevier, New York, 1981).

[29] A. Glachant, J. P. Coulomb, M. Bienfait, and J. G. Dash, J. Phys. (Paris) Lett. 40, L-543 (1979).

[30] J. P. Coulomb, J. Suzanne, M. Bienfait, M. Matecki, A. Thomy, B. Croset, and C. Marti, J. Phys. (Paris) 41, 1155 (1980).

[31] G. B. Huff and J. G. Dash, J. Low Temp. Phys. 24, 155 (1976).

[32] J. Stoltenberg and O. E. Vilches, Phys. Rev. B 22, 2920 (1980).

[33] R. E. Rapp, E. P. deSouza, and E. Lerner (unpublished).

[34] S. Ross and J. P. Olivier, *On Physical Adsorption* (Interscience, New York, 1964).

[35] A. D. Novaco, J. Low Temp. Phys. 9, 457 (1972).

[36] T. T. Chung and J. G. Dash, Surf. Sci. 66, 559 (1977).

[37] T. T. Chung, Surf. Sci. 87, 348 (1979).

[38] F. E. Hanson, M. J. Mandell, and J. P. McTaque, J. Phys. (Paris) 38, C076 (1977).

[39] H. Taub, K. Carneiro, J. K. Kjems, L. Passell, and J. P. McTague, Phys. Rev. B 16, 4551 (1977).

[40] M. J. Tejwani, O. Ferreira, and O. E. Vilches, Phys. Rev. Lett. 44, 152 (1980).

[41] UCAR/ZYX graphite is a product of Union Carbide Corporation.

[42] R. J. Birgeneau, G. S. Brown, P. M. Horn, D. E. Moncton, and P. W. Stephens, J. Phys. C 14, L49 (1981).

[43] M. Matecki, A. Thomy, and X. Duval, Surf. Sci. 75, 142 (1978).

[44] P. Dolle, M. Matecki, and A. Thomy, Surf. Sci. 91, 271 (1980).

[45] A. Terlain, M. Matecki, and A. Thomy, Surf. Sci. 91, 271 (1980).