

Smectic-A to bilayer evolution in concentrated surfactant solutions: The role of spontaneous curvature

Rony Granek^{a)} and William M. Gelbart

Department of Chemistry and Biochemistry, University of California, 405 Hilgard Avenue, Los Angeles, California 90024-156905

Yardena Bohbot and Avinoam Ben-Shaul

Department of Physical Chemistry and The Fritz Haber Research Center, The Hebrew University, Jerusalem 91904, Israel

(Received 7 February 1994; accepted 29 March 1994)

We study the two-dimensional (2-D) structural and thermodynamic changes in smectic-A/lamellar phases of self-assembling surfactant systems, in which the rim associated with a bilayer edge has a preferred curvature. This property was not considered in previous studies of 2-D aggregation, where an infinite bilayer emerges already at very low concentrations. A lattice Hamiltonian is used to describe the bending energy of the rim: An occupied lattice site corresponds to a minimum, disklike, micelle, and a bending energy penalty is associated with corners and straight edges depending on the value of the spontaneous curvature. When the spontaneous radius of curvature of the rim is small and the bending modulus is large, we find that the “condensation” transition—i.e., the “collapse” of the finite aggregates into a continuous bilayer—is postponed to high concentrations. At low concentrations the bending energy leads to an effective *repulsive* interaction between the aggregates, which in turn can result in ordered (modulated) structures for not too large ratios of thermal energy to bending energy (which is the expected situation in most systems of interest). Our model should be applicable to the systems of decylammonium chloride and cesium perfluorooctanoate studied by Boden and co-workers (NMR and conductivity measurements) and Zasadzinski and co-workers (freeze fracture), where monodisperse micellar disks are observed to layer in stacked planes. In the latter system a 2-D order of disk-shaped aggregates appears within the smectic-A layers, which is also consistent with our theory. Experimental studies of the structural evolution under *further* condensation of the system are not yet available.

I. INTRODUCTION

Self-assembly in surfactant solutions has been studied extensively in the past two decades.^{1–5} This field evolved separately from the older, broader, area of “ordinary” colloids and liquid crystals, because the *self-assembled* aggregates—micelles—do not maintain their size and shape upon concentration of the system. Moreover, a simple argument, first introduced by Israelachvili *et al.*⁶ and further elaborated by others,^{7,8} shows that two-dimensional (2-D) aggregation leads to the formation of an infinite bilayer at concentrations close to the critical micelle concentration (CMC), via a first-order phase transition. Indeed, smectic-A—*layered*—phases typically appear in surfactant solutions in the form of lamellar sheets of bilayers.

It has therefore been surprising to find that in some systems, studied in particular by Boden *et al.*⁹ and Zasadzinski and co-workers,¹⁰ disklike micelles, which appear at the CMC, *retain approximately their size and shape* at relatively large concentrations and *persist even into the nematic and smectic-A phases*. The smectic-A phase thus appears to consist of disklike micelles organized in planes. Nevertheless, the origin of this phenomenon—and the apparent inconsis-

tency with the “general” argument mentioned above—has not been addressed theoretically (although a few related theoretical studies have been described^{11,12}).

In this paper, we argue that when a bending energy of the rim, namely a curvature dependent “line tension” which includes a preferred (“spontaneous”) curvature, is introduced, it is possible to obtain micellar disk phases with rather wide stability. We develop a lattice Hamiltonian formulation specifically designed to describe the smectic-A phase. This model allows us to study both the self-assembly structure within the planes and the first-order transition to a continuous bilayer which ultimately must occur upon concentration increase. Starting with a dilute 2-D system of aggregates with a strongly preferred curvature, one expects a few scenarios to compete with each other upon further concentration. First, the disklike aggregates can “crystallize” with (say) an hexagonal order which increases their packing entropy, much as is found in the freezing of a system of hard core disks. Second, the aggregates might elongate their shape into ellipses or ribbons, which also will increase their packing efficiency, while paying some penalty in bending energy. These ribbonlike aggregates, which are likely to be polydisperse in size, may then organize into 2-D smectic (“stripe”) or nematic phases. And third, the free energy can be lowered by a 2-D phase separation into a *continuous*—i.e., infinite—bilayer phase and a “dilute” micellar disk phase. The advan-

^{a)}Present and permanent address: Department of Materials and Interfaces, Weizmann Institute of Science, Rehovot 76100, Israel.

tage of the approach we develop is that it allows us to study the competition between these several scenarios using a *single* Hamiltonian.

II. BENDING ENERGY IN THE SELF-ASSEMBLY APPROACH

Simple theories of self-assembly often start with the free energy F written in the form⁸

$$F = k_B T \sum_s n_s (\ln n_s + E_s - 1), \quad (1)$$

where n_s is the number of aggregates with s molecules (s aggregates), and $k_B T E_s$ is the free energy of an s aggregate. The total number of molecules N is fixed by the constraint

$$\sum_s s n_s = N. \quad (2)$$

The *equilibrium* distribution, which we still denote by n_s , is obtained by minimizing $F - k_B T \mu N$ (with $k_B T \mu$ being the molecular chemical potential) with respect to n_s to obtain

$$n_s = e^{-(E_s - \mu s)}. \quad (3)$$

For 1-D aggregation $E_s \approx \tilde{E}_{\text{cyl}} s + E_{\text{caps}}$, where \tilde{E}_{cyl} is the energy (in units of $k_B T$) per molecule in the cylindrical body of the rodlike micelle, and $k_B T E_{\text{caps}}$ is the energy of its (hemispherical) caps. Clearly one can never have $\tilde{E}_{\text{cyl}} = \mu$ because the sum in (2) would diverge. This suggests that it is not possible to get an infinitely long rod coexisting with finite rods, and reflects the fact that there is no phase transition in 1-D.^{13,14} However, for 2-D aggregation the situation is qualitatively different. More explicitly, for disk-shaped aggregates and for large s we have (dropping numerical factors of order unity)

$$E_s \approx \tilde{E}_{\text{bil}} s + \tilde{e}_{\text{rim}} p s^{1/2}, \quad (4)$$

with

$$\tilde{e}_{\text{rim}} = \tilde{E}_{\text{rim}} - \tilde{E}_{\text{bil}} \quad (5)$$

and where \tilde{E}_{bil} and \tilde{E}_{rim} are essentially the energies (standard chemical potentials) of a surfactant molecule in bilayer and rim geometries, respectively, and p is (roughly) the number of surfactant molecules in a rim cross section: $p s^{1/2}$ is the total number N_{rim} of molecules in the rim of a (large) disk. Thus an “infinite” bilayer appears when $\mu = \tilde{E}_{\text{bil}}$ and the sum

$$\sum_s s n_s = \sum_s s \exp[-\tilde{e}_{\text{rim}} p s^{1/2}] \quad (6)$$

reflects the number of molecules in the finite aggregates coexisting with the bilayer.

Suppose now that the rim has a spontaneous (preferred) curvature for which the rim energy per unit length is a minimum. For an arbitrarily shaped rim and for sufficiently small curvatures, the rim energy can be described by the bending Hamiltonian^{15,16}

$$H_b = \int d\zeta \left[\gamma + \frac{1}{2} \kappa \left(\frac{1}{R(\zeta)} - \frac{1}{R_0} \right)^2 \right]. \quad (7)$$

Here ζ is the position along the rim, κ is a 1-D bending modulus (with dimensions of “energy \times length”), $R(\zeta)$ is the local radius of curvature, and R_0 is the spontaneous radius of curvature; γ is the rim energy per unit length (relative to the bilayer energy) when the local curvature is identical with the spontaneous one. Thus, for a disk-shaped aggregate, \tilde{e}_{rim} in Eq. (4) takes the form (with s_0 the number of molecules in a disk of radius R_0)

$$\tilde{e}_{\text{rim}} \approx E_0 + \frac{\kappa}{2l} \left(\frac{1}{s^{1/2}} - \frac{1}{s_0^{1/2}} \right)^2, \quad (8)$$

where $E_0 = \gamma r^2 / 2l$ (r being the radius of the head group) and $2l$ is the bilayer thickness. Inserting this expression in Eq. (6) we see that the largest contribution to the sum comes from terms with $s \approx s_0$ rather than from $s \approx 0$ (which latter is the case for constant, s independent, \tilde{e}_{rim}). Thus, if $E_0 \ll \kappa/l$ one can have in this case many more molecules belonging to the finite disks rather than to the infinite bilayer—see (6). This implies that the “dilute” phase, namely the phase which corresponds to finite aggregates, has a wider concentration range of stability.

The self-assembly approach described above, although being rather primitive, can be relatively easily generalized to more complicated situations (e.g., adding aggregate–aggregate interactions,⁸ or incorporating 3-D configurational entropy contributions^{8,17} to E_s). But it suffers from the clear disadvantage of being unable to account *on the same footing* for *both* long range order *and* fluctuations of aggregate size and shape, and, in particular, for the case of infinite bilayer. A more flexible, lattice, formulation, which can overcome this difficulty, is described next.

III. ISING LATTICE-GAS HAMILTONIAN

It is possible to describe without modification self-assembly systems in the language of the well-known lattice-gas model,¹⁴ so long as one does not need to include 3-D configurational entropy for flexible objects or 3-D interaggregate interactions beyond those of the short-range type. (Such effects can be described using field theoretical formalisms, e.g., as in the study by Huse and Leibler¹⁸ of defects in the sponge phase.) In the Ising lattice-gas model the Hamiltonian is given by¹⁴

$$H = -\epsilon_{\text{pair}} \sum_{\langle ij \rangle} S_i S_j, \quad (9)$$

where $S_i = 1$ (vs 0) describes the occupation of a site by a particle, and the sum is over all nearest neighbors (NN). A particle here corresponds to a micellar aggregate so it must be at least as large as a spherical or minimum disk micelle in 3-D and 2-D situations, respectively. (Note, however, that the 3-D Ising model is not relevant to surfactant aggregates.) For the 2-D case of interest to us we shall see that the energy parameter ϵ_{pair} plays the role of the rim energy, i.e., when two aggregates become nearest neighbors they fuse into *one* particle and convert “rim” molecules into “body” molecules (see 2-D estimate given immediately below for the case where $\epsilon_{\text{pair}} > 0$). Note that this model exhibits a particle–hole

symmetry: Under the transformation $S_i \rightarrow 1 - S_i$ the Hamiltonian is invariant except for terms which are linear in the average concentration $c = \langle S_i \rangle$.

For a 2-D square lattice—and for $\epsilon_{\text{pair}} > 0$ —this model exhibits a first-order condensation transition for $T < T_c$ where $^{14} k_B T_c \approx \epsilon_{\text{pair}}$ (the equality holds in the mean-field approximation). The coexistence and spinodal curves are symmetric around $c = 1/2$, as a consequence of the particle-hole symmetry. Since $\epsilon_{\text{pair}} \approx N_{\text{rim}} \tilde{e}_{\text{rim}}$ then with $\tilde{e}_{\text{rim}} \geq 0.1 k_B T$ and $N_{\text{rim}} \geq 100$ for a small disk we have that $\epsilon_{\text{pair}} \geq 10 k_B T$ and so the temperature regime which is applicable here corresponds to $T \ll T_c$. (For \tilde{e}_{rim} negative, on the other hand, no condensation will occur, unless additional interactions—such as the many-body terms discussed below—are incorporated.) Since the coexistence curve is very steep near $c = 0$ (or $c = 1$) the concentration at which phase separation (from micellar to bilayer phase) begins to occur is vanishingly small. In this respect there is a complete correspondence between this viewpoint and the more naive approach described in the previous section, i.e., infinite bilayer appears at a very low concentration of surfactant. It should be kept in mind however that the third dimension could be crucial: the phase separation in 2-D allows in principle, for example, that in 3-D a single phase consisting of infinite bilayer(s) and finite aggregates would be stable.

IV. THE BENDING LATTICE-GAS HAMILTONIAN

A. Coupling constants

In order to describe the stability of concentrated solutions of disklike micelles, we want to follow the ideas presented in Sec. II and assume that the rim is associated with a spontaneous (nonzero) curvature for which the rim energy is a minimum. This special curvature is taken as a parameter of the theory, which in turn depends on the experimental conditions, such as salt concentration, temperature, and chain length. To treat these effects, the Ising lattice gas is no longer appropriate, and one has to construct a lattice Hamiltonian which can distinguish between different curvatures. More specifically, we distinguish on a 2-D square lattice between “positive,” “negative,” and zero curvature configurations (see Fig. 1) and assign them different energies: ϵ_p , ϵ_n , and ϵ_f , respectively. Now, in order to minimize further the number of free parameters, we can relate these energies to the bending Hamiltonian Eq. (7), noting that a corner of an occupied site (surrounded by empty sites) corresponds to a quarter of a circle. Ignoring the difference in the length of a quarter of a circle ($\pi a/4$) and a quarter of a square (a), where a is the diameter (equal to the lattice constant), we find

$$\epsilon_p = (K - b)^2, \quad (10)$$

$$\epsilon_n = (K + b)^2, \quad (11)$$

$$\epsilon_f = b^2, \quad (12)$$

where

$$K \equiv \sqrt{2} \sqrt{\kappa/a}, \quad (13)$$

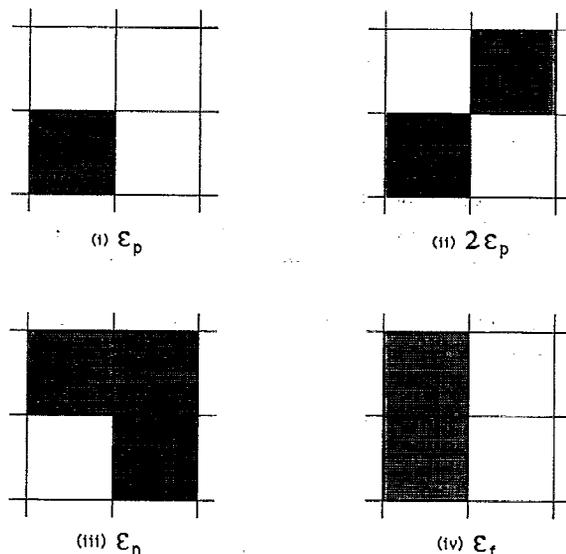


FIG. 1. Schematic representation of the possible four-site configurations on a square lattice, with their associated energies. Note that configuration (ii) corresponds to two separate disks and thus to an energy penalty $2\epsilon_p$.

$$b \equiv \frac{\sqrt{\kappa a}}{\sqrt{2}R_0} = \frac{a}{2R_0} K. \quad (14)$$

For the Ising energy parameter ϵ_{pair} [cf., Eq. (9)], we find similarly

$$\epsilon_{\text{pair}} = 2\gamma a, \quad (15)$$

where γ is again the optimum-curvature rim energy per unit length, relative to bilayer. If ϵ_{pair} is neglected, e.g., if γ is set equal to zero, we are left with only two energy constants, K and b , one corresponding to the bending energy ($K \sim \sqrt{\kappa}$) and the other to the spontaneous bending curvature ($b \sim 1/R_0$). Our lattice model is similar to those used by Safran and co-workers¹⁹ to describe microemulsion phases, but differs from theirs in several ways due to the different underlying physics.²⁰

B. Mean-field free energy

The free energy per lattice site $f_0 = F/N_0$ is easily obtained in the mean-field approximation as the sum of the entropy of mixing contribution and the energies ϵ_{pair} , ϵ_p , ϵ_n , and ϵ_f ; each weighted by their random mixing probabilities:

$$f_0 = 2\epsilon_{\text{pair}}c(1-c) + 4(K-b)^2c(1-c)^2 + 4(K+b)^2c^3(1-c) + 4b^2c^2(1-c)^2 + k_B T [c \ln c + (1-c) \ln(1-c)], \quad (16)$$

with

$$c \equiv \langle S_i \rangle \equiv N_{\text{av}}/N_0. \quad (17)$$

It can easily be checked that under the transformation $c \rightarrow 1 - c$, f_0 is not symmetric unless $b = 0$. This is more readily seen when one transforms to the “magnetization” variable $m = 2c - 1$ which leads to (omitting constants and terms linear in m which do not effect any physical quantity)

$$f_0 = -\frac{1}{2}[\epsilon_{\text{pair}} + (K-b)^2 + b^2]m^2 - 2bKm^3 - \frac{1}{4}(K^2 + 2Kb)m^4 + k_B T \times \left[\frac{1-m}{2} \ln\left(\frac{1-m}{2}\right) + \frac{1+m}{2} \ln\left(\frac{1+m}{2}\right) \right]. \quad (18)$$

From this representation it is clear that if $b=0$, the coefficient of m^3 vanishes and f_0 is a symmetric function of m . Thus the spontaneous curvature parameter b tunes the asymmetry between holes and particles in this model.

We now proceed to calculate the spinodal and coexistence lines for this free energy. The chemical potential can be obtained from $\mu = (\partial F / \partial N_{\text{av}})_{N_0, T} = [\partial(F/N_0) / \partial(N_{\text{av}}/N_0)]_T \equiv (\partial f / \partial c)_T = 2(\partial f / \partial m)_T$ and we get

$$\mu = k_B T \ln\left[\frac{1+m}{1-m}\right] - 2[\epsilon_{\text{pair}} + (K-b)^2 + b^2]m - 12Kbm^2 - 2(K^2 + 2Kb)m^3. \quad (19)$$

This chemical potential shows a twiddle shaped curve (“van der Waals loop”) below a certain critical temperature T_c , marking the onset of a (first-order) condensation phase transition. The spinodal curve for this transition is given as usual by $(\partial\mu/\partial c)_T = 0$, from which we find

$$4c(1-c) \frac{\epsilon_{\text{eff}}}{k_B T} = 1, \quad (20)$$

where

$$\epsilon_{\text{eff}} = \epsilon_{\text{pair}} + (K-b)^2 + b^2 + 12Kbm + 3(K^2 + 2Kb)m^2. \quad (21)$$

The coexistence curve is obtained in the usual way by equating both the chemical potentials and the osmotic pressures

$$\Pi = c\mu - f_0 \quad (22)$$

of the two phases.

C. Phase behavior

The resulting phase diagrams are shown in Figs. 2(A)–2(C) for the special case where $\epsilon_{\text{pair}}=0$ (i.e., $\gamma=0$), for three different choices of spontaneous curvature b , $0 \leq b \leq K$. For $b=0$ the phase diagram is symmetric, but it is not Ising-like: see Fig. 2(A). There are two critical points (of equal temperature T_c), and a stable middle phase appears below T_c in a narrow temperature regime. The two coexistence curves (solid, vs the dashed spinodals) join in the middle in a eutecticlike point. Note that $k_B T / K^2 \approx 1$ corresponds to $T \approx T_{\text{room}}$, since we estimate $K^2 \approx k_B T_{\text{room}}$.²¹

As b increases from zero the right-hand (RH) T_c increases, while the left-hand (LH) T_c gets lower and moves to lower concentrations. For b greater than some value b^* the two critical temperatures are no longer connected by a single spinodal curve: see Fig. 2(B) for $b=K/3$. ($b^* \approx 0.3K$ —corresponding to a spontaneous radius of curvature of $R_0 \approx \frac{3}{2}a$ —is the value of b for which the middle minimum of the spinodal curve lies at $T=0$.) These two spinodals give rise to the two coexistence curves shown. As can be seen, the small LH coexistence curve is metastable with respect to the large condensation transition (resulting from the RH spin-

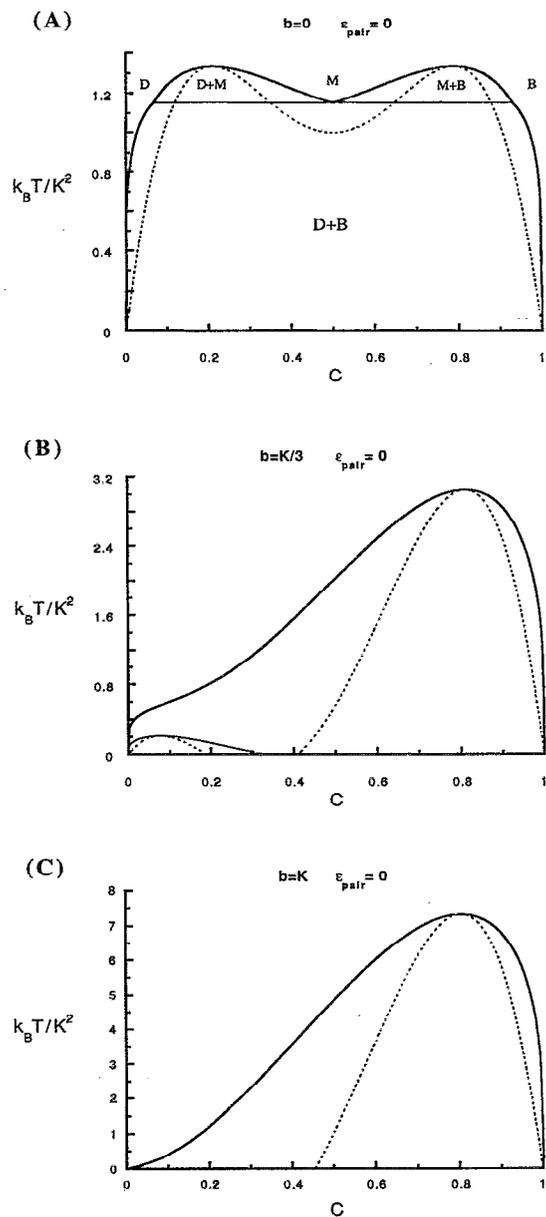


FIG. 2. Phase diagrams for three different values of the asymmetry parameter b : (A) $b=0$, (B) $b=K/3$, and (C) $b=K$. ($\epsilon_{\text{pair}}=0$ in all cases.) The dashed lines denote the spinodals, the full-thick lines are the coexistence curves. In (A) D denotes the dilute phase, B denotes the bilayer phase, and M denotes a middle phase. In (B) note that the coexistence regime on the bottom-left side is metastable with respect to the larger coexistence curve.

odal) into a bilayerlike phase. In this metastable sense, the LH coexistence curve can be (partially) viewed as describing a first-order “fusing” transition from single-site “disks,” into “disks” which are a few times larger. (However, as described in the next section, the low temperature part of this transition corresponds to condensation into an ordered stripe phase, rather than into an isotropic disklike phase.)

As b is further increased, there is a value of b ($\approx 0.5K$) for which the LH spinodal disappears completely. For the extreme case of $b=K$, which is the case when disks of the size of one lattice site are preferred [recall from Eq. (14) that

$b=K$ only when $a=2R_0$], the phase diagram is even more markedly nonsymmetric around $c=0.5$: see Fig. 2(C). ($b=K$ is in fact the largest that b can get in this model, since $2R_0 \geq a$.) Moreover the spinodal curve [cf. Eq. (20)] now ends at a finite concentration $c \approx 0.46$ [for which ϵ_{eff} of Eq. (21) vanishes] rather than at zero. Thus, in summary, for temperatures of interest ($T \approx T_{\text{room}}$), increasing concentration leads to either: successive “condensations” (if R_0 is large enough) from small disks to bigger objects and *then* to infinite bilayer; or a single disk-to-bilayer transition.

V. FLUCTUATIONS AND LONG-RANGE ORDER

A. The Hamiltonian

Next we look at the free energy which incorporates spatial variations of the concentration within a mean-field approximation. To derive this free energy, let us first write out the single-site Hamiltonian H_i , i.e., the sum of all terms in the total Hamiltonian of the system that include site i explicitly. This Hamiltonian is straightforwardly obtained from our assignment of energies in Sec. IV to the different four-site configurations depicted in Fig. 1. For convenience (and clarity) we transform the representation of H_i from the lattice-gas variable $S_i=0,1$ to the Ising variables $\sigma_i=-1,1$ (omitting constants and terms linear in σ_i). The resulting H_i can be written as a sum of products of two, three, and four spin variables:

$$H_i = H_i^{(2)} + H_i^{(3)} + H_i^{(4)}, \quad (23)$$

where

$$H_i^{(2)} = -\sigma_i \left[\frac{\epsilon_{\text{pair}} + 2(K-b)^2}{4} (\sigma_{i+x} + \sigma_{i-x} + \sigma_{i+y} + \sigma_{i-y}) + \frac{b^2 - (K-b)^2}{4} (\sigma_{i-x-y} + \sigma_{i+x-y} + \sigma_{i-x+y} + \sigma_{i+x+y}) \right], \quad (24)$$

$$H_i^{(3)} = -\sigma_i \frac{Kb}{2} [\sigma_{i-x}\sigma_{i-y} + \sigma_{i+x}\sigma_{i-y} + \sigma_{i-x}\sigma_{i+y} + \sigma_{i+x}\sigma_{i+y} + \sigma_{i-x}(\sigma_{i-x-y} + \sigma_{i-x+y}) + \sigma_{i-y}(\sigma_{i-x-y} + \sigma_{i-x+y}) + \sigma_{i+x-y} + \sigma_{i+x}(\sigma_{i+x-y} + \sigma_{i+x+y}) + \sigma_{i+y}(\sigma_{i-x+y} + \sigma_{i+x+y})], \quad (25)$$

$$H_i^{(4)} = -\sigma_i \frac{K^2 + 2Kb}{4} [\sigma_{i-x}\sigma_{i-y}\sigma_{i-x-y} + \sigma_{i+x}\sigma_{i-y}\sigma_{i+x-y} + \sigma_{i-x}\sigma_{i+y}\sigma_{i-x+y} + \sigma_{i+x}\sigma_{i+y}\sigma_{i+x+y}]. \quad (26)$$

Once again, we see that if $b=0$, $H_i^{(3)}$ vanishes and the Hamiltonian is symmetric under the transformation $\sigma_j \rightarrow -\sigma_j$. Note that the total Hamiltonian H_{tot} is *not* simply the sum of H_i because the two-spin terms would be counted twice, the three-spin terms thrice, and so on; more explicitly

$$H_{\text{tot}} = \frac{1}{2} \sum_i H_i^{(2)} + \frac{1}{3} \sum_i H_i^{(3)} + \frac{1}{4} \sum_i H_i^{(4)} \quad (27)$$

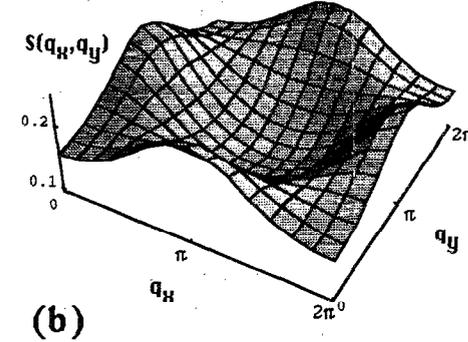
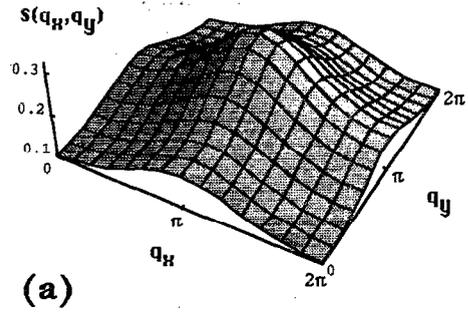


FIG. 3. 2-D structure factor $S(q_x, q_y)$ for two values of the asymmetry parameter b showing: (a) a checkerboard instability, and (b) a stripe instability. Both cases are for $\epsilon_{\text{pair}}=0$ and $c=0.2$. Other parameters: (a) $b^2=K^2=k_B T/3$; (b) $b=K/3$, $K^2=k_B T$.

(We note that this treatment of curvature energy is similar in spirit to the multispin models considered for *microemulsions* by Widom,²² Schick and co-workers,²³ Chen *et al.*,²⁴ Dawson *et al.*,²⁵ Matsen and Sullivan,²⁶ and others.²⁷)

B. The free energy

Unlike in Sec. IV we now allow for spatial variation of the average spin value $\langle \sigma_i \rangle \equiv m_i$. The mean-field free energy is obtained by introducing the *local* entropy of mixing s_i

$$s_i = -k_B \left[\frac{1-m_i}{2} \ln \left(\frac{1-m_i}{2} \right) + \frac{1+m_i}{2} \ln \left(\frac{1+m_i}{2} \right) \right], \quad (28)$$

whereupon the (averaged) entropy per site is given by

$$\frac{S(\{m_i\})}{N} = \frac{1}{N} \sum_i s_i \quad (29)$$

and the mean-field free energy (per site) is

$$f = \frac{1}{N} [H_{\text{tot}}(\{m_i\}) - TS(\{m_i\})], \quad (30)$$

where $H_{\text{tot}}(\{m_i\})$ is simply the H_{tot} given by Eqs. (24)–(27), but with each fluctuating spin variable (σ_j) replaced by its local average (m_j). Since this is still a complicated expression, we analyze it using a Landau–Ginzburg type

expansion.^{28,29} Specifically, we separate the local average concentration $c_i = (1 + m_i)/2$ into two contributions, the global average c and a local fluctuation ϕ_i : $c_i = c + \phi_i$ (or $m_i = m + 2\phi_i$). We assume that the $\{\phi_i\}$'s are sufficiently small to justify a Taylor expansion of f in their powers, with higher-than-sixth-order terms neglected. To further simplify

the result, we introduce a (lattice) Fourier transform whereby

$$\phi_i = \sum_{\mathbf{q} \neq 0} \phi_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{R}_i}$$

f can then be written in the form (all \mathbf{q} 's are nonzero, here and henceforth)

$$\begin{aligned} f = & f_0 + \frac{1}{2} k_B T \sum_{\mathbf{q}} S(\mathbf{q})^{-1} \phi_{\mathbf{q}} \phi_{-\mathbf{q}} + \frac{1}{3!} \sum_{\mathbf{q}_1, \mathbf{q}_2} \Gamma_3(\mathbf{q}_1, \mathbf{q}_2) \phi_{\mathbf{q}_1} \phi_{\mathbf{q}_2} \phi_{-(\mathbf{q}_1 + \mathbf{q}_2)} \\ & + \frac{1}{4!} \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} \Gamma_4(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \phi_{\mathbf{q}_1} \phi_{\mathbf{q}_2} \phi_{\mathbf{q}_3} \phi_{-(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3)} + \frac{1}{5!} \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4} \Gamma_5(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4) \phi_{\mathbf{q}_1} \phi_{\mathbf{q}_2} \phi_{\mathbf{q}_3} \phi_{\mathbf{q}_4} \phi_{-(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4)} \\ & + \frac{1}{6!} \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4, \mathbf{q}_5} \Gamma_6(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4, \mathbf{q}_5) \phi_{\mathbf{q}_1} \phi_{\mathbf{q}_2} \phi_{\mathbf{q}_3} \phi_{\mathbf{q}_4} \phi_{\mathbf{q}_5} \phi_{-(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4 + \mathbf{q}_5)} + \dots, \end{aligned} \quad (31)$$

where we have omitted sums over reciprocal lattice vectors.²⁸ (These latter sums give rise to "umklapp" terms,²⁸ but are clearly a lattice artifact for our purposes.) In Eq. (31) f_0 is the homogeneous part of the free-energy as given in Eq. (16) or (18). The other terms are the contributions from the spatial variations in density, and we now focus on these separate terms in turn.

C. The structure factor

According to the equipartition theorem, and within the Gaussian approximation, the coefficient of the ϕ^2 term is proportional to the reciprocal of the structure factor $S(\mathbf{q})$ and for this, *simple square lattice*, model, is given by

$$\frac{k_B T}{S(\mathbf{q})} = a_0(T - T^*), \quad (32)$$

with

$$a_0 = a_0(c) = \frac{k_B}{c(1-c)} \quad (33)$$

and (with all wave numbers henceforth in units of $1/a$)

$$\begin{aligned} T^* = T^*(c; \mathbf{q}) = & \frac{4c(1-c)}{k_B} \left\{ \frac{\epsilon_1}{2} [\cos(q_x) + \cos(q_y)] \right. \\ & \left. + \epsilon_2 \cos(q_x) \cos(q_y) \right\}. \end{aligned} \quad (34)$$

Here ϵ_1 and ϵ_2 are *effective* NN and NNN (respectively) coupling constants and are given by

$$\epsilon_1 = \epsilon_{\text{pair}} + 2(K-b)^2 + 8Kbm + 2[K^2 + 2Kb]m^2 \quad (35)$$

and

$$\epsilon_2 = -K^2 + 2Kb + 4Kbm + [K^2 + 2Kb]m^2. \quad (36)$$

This is precisely the structure factor of a lattice gas with only *pair* (two-site) NN (ϵ_1) and NNN (ϵ_2) interactions, except that here—due to the three- and four-spin interactions—these

coefficients are functions of the concentration c . If one puts $q_x = q_y = 0$, then this structure factor diverges on the spinodal curve given by Eq. (20) [since $\epsilon_1 + \epsilon_2 = \epsilon_{\text{pair}} + b^2 + (K-b)^2 + 12Kbm + 3[(K+b)^2 - b^2]m^2 \equiv \epsilon_{\text{eff}}$], as generally required from it being equal to the compressibility $\partial c / \partial \mu$. [See Figs. 2(A)–2(C) for examples of these spinodals with $\epsilon_{\text{pair}} = 0$, $0 < b \leq K$.]

We now focus on the extreme case of $b = K$, again taking $\epsilon_{\text{pair}} = 0$. Recall that $\epsilon_{\text{pair}} = 0$ corresponds to $\gamma = 0$, i.e., to there being no energy change upon moving molecules between bilayer and optimum-rim micellar environments, and that the choice $b = (a/2R_0)K = K$ implies $R_0 = a/2$, i.e., the spontaneous (and therefore optimum) rim curvature equals that of the minimum disk. For these values of b and ϵ_{pair} , the effective NN and NNN interactions given by the preceding equations take the form $\epsilon_1(c)/K^2 = 24(c - \frac{1}{2})(c + \frac{1}{6})$ and $\epsilon_2(c)/K^2 = 12c(c - \frac{1}{3})$, so that ϵ_1 is positive for $c > 0.5$, and negative for $c < 0.5$. This implies an *effective repulsive interaction* between NN particles, at low concentrations. Since in our model NN particles are considered as fused (to make, say, a piece of ribbon) this means that in this "low" concentration regime particles like to keep their integrity and *not* fuse, consistent with the large value of b ($=K$). For $c < \frac{1}{3}$, ϵ_2 also becomes negative. This can be understood in terms of a three body effect: If two particles stay close enough—i.e., as NNN's—then the appearance of a mutual NN particle will create a configuration of higher bending energy. Obviously, this should be a weaker effective interaction than the NN one, as indeed it is. This will be more clearly demonstrated below, when various *ordered* mesophases are considered.

Thus we conclude that for the case $b = K$ and $\epsilon_{\text{pair}} = 0$ there are three regimes of concentration on which it is natural to focus: $0 < c < \frac{1}{3}$ where both ϵ_1 and ϵ_2 are negative; $\frac{1}{3} < c < \frac{1}{2}$ where $\epsilon_1 < 0$ while $\epsilon_2 > 0$; and $\frac{1}{2} < c < 1$ where both are positive. For the high concentration $\epsilon_1, \epsilon_2 > 0$ regime it is clear that T^* —and hence $S(\mathbf{q})$ —is a maximum for

$\mathbf{q}=\{q_x, q_y\}=\{0,0\}$. For all lower concentrations, however (both $0 < c < \frac{1}{3}$ and $\frac{1}{3} < c < \frac{1}{2}$), it turns out that $S(\mathbf{q})$ has its maximum at $\mathbf{q}=\{\pi, \pi\}$ [see Fig. 3(a)], corresponding to a “checkerboard” pattern with wavelength $\sqrt{2}a$. Accordingly, for $0 < c < \frac{1}{2}$, the coefficient $a_0(T - T^*)$ of the quadratic term in the free energy expansion (in powers of $\phi_{\mathbf{q}}$) vanishes when T approaches T^* ($c; \mathbf{q}=\{\pi, \pi\}$) from above: at this temperature the uniform system becomes unstable with respect to the evolution of a (“checkerboard”) spatial inhomogeneity with wavelength $\sqrt{2}a$.

For other values of the elastic constants K , b and of ϵ_{pair} , more complicated scenarios arise for this instability from uniform to modulated states. For b small enough, for example, there are concentration regimes where the maximum in $S(\mathbf{q})$ —and hence in T^* —occurs [see Fig. 3(b)] for $\mathbf{q}=\{0, \pi\}$ (or $\{\pi, 0\}$), corresponding to a “stripe” pattern along y (or x), with wavelength $2a$. As an example, consider the case where $b = \frac{1}{3}K$ —i.e., $R_0 = \frac{2}{3}a$ —with $\epsilon_{\text{pair}}=0$ still. Then for $0.122 < c < 0.478$ a stripe instability occurs at T^* , whereas for higher concentrations the maximum in $S(\mathbf{q})$ appears again at $\mathbf{q}=\{0,0\}$.

More generally, the competition between stripe and checkerboard instabilities is determined according to the difference $|\epsilon_1| - 2|\epsilon_2| = 2b^2 - 4(K - b)^2 - \epsilon_{\text{pair}}$ (with both ϵ_1 and ϵ_2 assumed negative). When this difference is positive, the checkerboard instability dominates; when it is negative—stripe dominates. Note that this difference is independent of concentration; it is determined only by the energy parameters of the model. Intuitively, one could guess that it is the energy penalty for corners vs flat edges that should determine whether stripe or checkerboard instability dominates. Indeed, equating ϵ_p with ϵ_f gives $b^{**} = K/2$ for the value of b at this crossover while solving (say, for $\epsilon_{\text{pair}}=0$) $\epsilon_1 = 2\epsilon_2$ yields—as the prediction of our mean-field theory $-b^{**} = (2 - \sqrt{8}/2)K = 0.59K$. The stripe instability persists also for positive ϵ_1 so long as ϵ_2 is negative and the difference $\epsilon_1 - 2|\epsilon_2| = \epsilon_{\text{pair}} + 2b^2 + 16Kbm + 4(K^2 + 2Kb)m^2$, which is now a function of concentration, is negative. [When this latter difference becomes positive, and ϵ_1 is also positive, the maximum of $S(\mathbf{q})$ is back at $\mathbf{q}=\{0,0\}$.]

We have been concerned here with the \mathbf{q} positions of the $S(\mathbf{q})$ maxima, for different concentrations and for different spontaneous curvatures, because in the following sections we will argue that the relevant/dominant spatial orderings (fluctuations) are those with \mathbf{q} maximizing the structure factor.

D. Higher-order terms

The higher-order coefficients in the Landau–Ginzburg expansion Eq. (31) are given by

$$\Gamma_3(\mathbf{q}, \mathbf{q}') = k_B T \left[\frac{1}{(1-c)^2} - \frac{1}{c^2} \right] - \epsilon_3 \gamma_3(\mathbf{q}, \mathbf{q}'), \quad (37)$$

$$\Gamma_4(\mathbf{q}, \mathbf{q}', \mathbf{q}'') = 2k_B T \left[\frac{1}{(1-c)^3} + \frac{1}{c^3} \right] - \epsilon_4 \gamma_4(\mathbf{q}, \mathbf{q}', \mathbf{q}''), \quad (38)$$

$$\Gamma_5 = 6k_B T \left[\frac{1}{(1-c)^4} - \frac{1}{c^4} \right], \quad (39)$$

$$\Gamma_6 = 24k_B T \left[\frac{1}{(1-c)^5} + \frac{1}{c^5} \right], \quad (40)$$

where

$$\epsilon_3 = 48[2Kb + (K^2 + 2Kb)m], \quad (41)$$

$$\epsilon_4 = 96(K^2 + 2Kb), \quad (42)$$

and where (not bothering to symmetrize)

$$\begin{aligned} \gamma_3(\mathbf{q}, \mathbf{q}') &= \cos(q'_y)\cos(q_x) + \cos(q'_x)\cos(q_x + q'_x) \\ &\quad + \cos(q_x)\cos(q_y + q'_y), \end{aligned} \quad (43)$$

$$\gamma_4(\mathbf{q}, \mathbf{q}', \mathbf{q}'') = \cos(q_x + q'_x)\cos(q_y + q''_y). \quad (44)$$

The coefficients Γ_3 and Γ_4 reflect the competition between energy and entropy and may, in principle, be either positive or negative. Indeed, a negative Γ_4 requires that we expand up through the sixth-order term (for stability). The coefficients Γ_5 and Γ_6 are independent of wave vectors since they arise purely from the entropy, with no direct contributions from energy terms (which latter are nonlocal in the ϕ_i 's, thereby leading to wave vector dependence of the Γ 's).

E. Ordered phases

The analysis of the relative stability of various ordered phases is clearly limited by our use of a specific lattice. Nevertheless, the square lattice does permit us to examine the stability of ordered phases vs the isotropic phase, and even the stability of 2-D order—e.g., the checkerboard phase—vs 1-D order—e.g., the stripe phase (along the x or y axes). These ordered structures are natural for the square lattice, i.e., no frustrations are introduced.

For simplicity we allow only the first harmonic for each type of order. In general this is written as (here e is a dimensionless amplitude)

$$\phi(\mathbf{r}) = e \sum_{\mathbf{Q}_n} \exp[i\mathbf{Q}_n \cdot \mathbf{r}], \quad (45)$$

where \mathbf{Q}_n is the n th NN reciprocal lattice vector. More specifically, for the checkerboard phase we have $\mathbf{Q}_1 = k(\hat{x} + \hat{y})$ and $\mathbf{Q}_2 = -k(\hat{x} + \hat{y})$, implying

$$\phi(\mathbf{r}) = 2e_c \cos[k(x + y)] \quad (46)$$

and, similarly, for the stripe phase (along the x axis, say)

$$\phi(\mathbf{r}) = 2e_s \cos(kx). \quad (47)$$

In Fourier space these fluctuations correspond to

$$\phi_{\mathbf{q}} = e_c [\delta(q_x - k)\delta(q_y - k) + \delta(q_x + k)\delta(q_y + k)] \quad (48)$$

and

$$\phi_{\mathbf{q}} = e_s \delta(q_y) [\delta(q_x - k) + \delta(q_x + k)]. \quad (49)$$

The strategy is then to insert these forms into the Landau–Ginzburg expansion (31),²⁹ and minimize over both the wave number k and the amplitudes e_c and e_s . But if the analysis is limited to regions near the “disorder line” (see below) where the amplitudes are small, one can neglect the higher than quadratic terms (in ϕ , or—equivalently—in e) in minimizing over k . This corresponds to looking for the maxi-

mum in $S(\mathbf{k})$ [see above discussion of $S(\mathbf{q})$]: the preferred wave vector is just the one that maximizes the structure factor. (For the checkerboard and stripe instabilities we have verified that, in fact, the higher-order terms in the expansion do not alter at all the value of k obtained from the structure factor; these terms merely produce new maxima, rather than minima, in the plot of f vs k). Thus, for example, the checkerboard phase can appear (i.e., necessary but *not* sufficient condition) only for $\epsilon_1 < 0$.

For the regime where both ϵ_1 and ϵ_2 are negative, consider first the case where $2b^2 - 4(K-b)^2 - \epsilon_{\text{pair}} = |\epsilon_1| - 2|\epsilon_2|$ is positive (e.g., when $b=K$, $\epsilon_{\text{pair}}=0$ and $c < \frac{1}{3}$). For this case $k=\pi$ for both the checkerboard and stripe phases, but the checkerboard phase should be the stable one since its $S(\mathbf{q})$ value (above the transition) is higher. After summation over the \mathbf{q} 's we obtain for the corresponding free-energies

$$\delta f_c \equiv f_c - f_0 = \frac{1}{2} \tau_c e_c^2 + \frac{\lambda_c}{4} e_c^4 + \frac{\delta_c}{6} e_c^6, \quad (50)$$

$$\delta f_s \equiv f_s - f_0 = \frac{1}{2} \tau_s e_s^2 + \frac{\lambda_s}{4} e_s^4 + \frac{\delta_s}{6} e_s^6, \quad (51)$$

where

$$\tau_c = \frac{2k_B T}{c(1-c)} - 8(\epsilon_2 - \epsilon_1), \quad (52)$$

$$\tau_s = \frac{2k_B T}{c(1-c)} + 8\epsilon_2 \quad (53)$$

and, with $\mathbf{k} \equiv \pi(\hat{x} + \hat{y})$ (checkerboard) or $\mathbf{k} \equiv \pi\hat{x}$ (stripe),

$$\lambda_c = \lambda_s = \Gamma_4(\mathbf{k}, \mathbf{k}, -\mathbf{k}) = 2k_B T \left[\frac{1}{(1-c)^3} + \frac{1}{c^3} \right] - \epsilon_4, \quad (54)$$

$$\delta_c = \delta_s = 4k_B T \left[\frac{1}{(1-c)^5} + \frac{1}{c^5} \right]. \quad (55)$$

Because $\tau_s > \tau_c$ [i.e., $2b^2 - 4(K-b)^2 - \epsilon_{\text{pair}} > 0$], and the rest of the coefficients (λ and δ) are the same for the stripe and checkerboard modulations, *the stripe free energy lies below the checkerboard free energy, and the two free energies never cross one another*. Thus only the checkerboard phase should appear. This particular detail is clearly an artifact of our choice of a checkerboard modulation to represent the 2-D order. (Indeed, if the checkerboard order were replaced by an hexagonal order, regimes of first-order coexistence of isotropic hexagonal and hexagonal stripe should appear, as further discussed in Sec. VI.)

Note that both τ_c and λ_c (or τ_s and λ_s) can become negative. For small c ($c < 0.225$ in the $b=K$ case), the line $\tau_c=0$ lies above the line $\lambda_c=0$, so that the transition from the isotropic to checkerboard phase (the "disorder line") is second order and is given by $\tau_c=0$. On the other hand, above a certain value of c ($c > 0.225$ in the $b=K$ case) λ_c becomes negative before τ_c (as the temperature is lowered) and this transition is first order; the checkerboard-order amplitude e_c jumps from zero to a finite value [$e_c^2 = -3\lambda_c/(4\delta_c)$] just below this line (although the free energy is continuous, of course). The transition is determined by $\lambda_c^2 = (16/3)\tau_c\delta_c$. The point for which both τ_c and λ_c vanish ($c=0.225$,

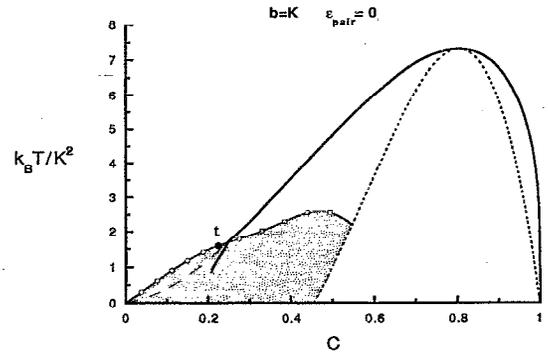


FIG. 4. Phase diagram for the $b=K$ case. Full and short-dashed lines denote coexistence and spinodal curves involving isotropic phases. The low temperature/low concentration continuation of the coexistence curve (full line) cannot be calculated within the approximations used in the text. The shaded area denotes a checkerboard phase, which is partly metastable. The line with circles denotes a second-order transition between a (dilute) isotropic phase and the checkerboard phase. The line with squares denotes a similar first-order transition. (Note that most of this line is metastable with respect to condensation.) t denotes a tricritical point. The long-dashed line is the coexistence curve calculated with the assumption of no ordering [i.e., identical to the full line in Fig. 2(C)], and is shown only for comparison.

$k_B T / K^2 = 1.60$ in the $b=K$ case) is thus a tricritical point. Note that whenever the transition between isotropic and checkerboard phases is first order, there is a discontinuity as well in the concentration (c), but we have chosen not to consider explicitly the associated coexistence curves here. (Details of this kind are treated, for example, by Andelman, Brochard and Joanny,³⁰ in their study of spatially modulated phases of Langmuir monolayers.)

This isotropic-checkerboard transition line is shown in Fig. 4; the line with circles corresponds to the second order transition, and the line with squares represents the first order transition, with the tricritical point denoted by t . We also see (Fig. 4) that the coexistence curve for the condensation transition, calculated with the assumption of an *isotropic* dilute phase [as in Fig. 2(C)], enters (see long-dashed line) into the ordering-checkerboard regime. Therefore, this regime of the coexistence curve has been recalculated (see full line) applying a Maxwell construction to the free energies of the checkerboard and of the condensed-isotropic phases. We see that the *equilibrium regime of checkerboard ordering* has become wider, which is due to the fact that the checkerboard free energy is lower than the previously used free energy corresponding to isotropic structure. The part of the checkerboard order which is inside the coexistence curve is of course only metastable.

A different scenario occurs when $|\epsilon_1| - 2|\epsilon_2|$ is *negative* (and both ϵ_1 and ϵ_2 are negative), as can happen for small enough b or large enough ϵ_{pair} . From the maximum of the structure factor we can anticipate that stripe ordering should be preferred over checkerboard ordering. For the stripe modulation the preferred k is again π so that τ_s is still given by Eq. (53). For the checkerboard the preferred k is now given by $\cos k = -\epsilon_1/(2\epsilon_2)$ (corresponding to wavelengths longer than the lattice constant). The values of τ_c and λ_c in Eqs. (52)–(54) are changed accordingly³¹ so that now we

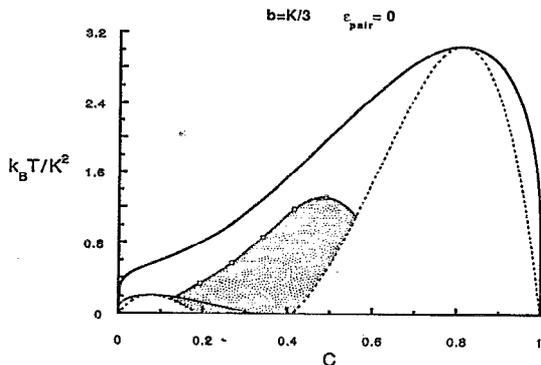


FIG. 5. Phase diagram for the $b=K/3$ case. The shaded area is a metastable stripe phase. The line with squares denotes a first-order transition between a (metastable) isotropic phase and the stripe phase. Other lines are the same as in Fig. 2(B).

have $\tau_s < \tau_c$ and $\lambda_s < \lambda_c$. This indeed means that the free energy of stripe order is lower than the free energy for checkerboard order (as explained above), and below the isotropic-stripe transition line we can find *only a stripe phase*, as anticipated. This transition can be either first or second order, as explained above.

In Fig. 5, we consider the case $\epsilon_{\text{pair}}=0$ and $b=K/3$ [as in Fig. 2(B)]. The stripe phase is seen to be stabilized—only in a *metastable* sense—by the bending Hamiltonian with small spontaneous curvature. The isotropic-stripe transition is found to be purely first order for this case [calculated according to $\lambda_s^2=(16/3)\tau_s\delta_s$]; there is no regime of second-order transition. But the stripe phase appears between the two spinodals (for condensation) and is entirely metastable with respect to global condensation; there is no *equilibrium* stripe phase! Of course, this does not rule out the possibility of an equilibrium stripe phase for *other* values of ϵ_{pair} —see Sec. VI A.

VI. DISCUSSION

A. Estimate of parameters

So far we have hardly given any direct physical estimates²¹ for the model parameters b , K , and ϵ_{pair} (or, equivalently, for R_0 , κ and γ) for systems of experimental interest. An *a priori* molecular theory is barely feasible even for very simple surfactants and thus is not so attractive. An alternative approach is to try to express these parameters in terms of other independently measurable quantities, for example, relating them to the elastic moduli describing the bending energy of a 2-D surfactant film (in a 3-D space). This energy is given by^{8,16}

$$H_b^{(2D)} = \int d^2\zeta \left[\frac{1}{2} k \left(\frac{1}{R_1} + \frac{1}{R_2} - C_0 \right)^2 + \frac{\bar{k}}{R_1 R_2} - \frac{1}{2} k C_0^2 \right], \tag{56}$$

where k and \bar{k} are the mean and Gaussian bending moduli (respectively), and C_0 is the 2-D spontaneous curvature of the film. In Eq. (56) we have subtracted $kC_0^2/2$ in order to obtain the rim energy relative to that of the bilayer. (Note that we have used R_0 to denote the 1-D spontaneous radius

of curvature. The 2-D spontaneous radius of curvature is obviously $2/C_0$.) Assuming this bending energy to hold even for the high curvature associated with the rim, one can use it to calculate the energy of an “ideal” disk whose thickness is $2l$ and whose diameter ($2R$) is much larger than its thickness. Comparing this to a calculation based on the 1-D elasticity theory given by Eq. (7), we obtain the following relations:

$$\kappa = (\pi/2)kl, \tag{57}$$

$$\frac{1}{R_0} = \frac{4}{\pi} \left(C_0 - \frac{1}{l} \right) - \frac{4\bar{k}}{\pi kl} \tag{58}$$

$$\gamma = \frac{\pi}{2} kl \left(C_0 - \frac{1}{l} \right)^2 - \frac{\pi}{2} k l C_0^2 - \frac{\pi kl}{4 R_0^2}. \tag{59}$$

Note that stability of a 2-D film^{16,32} requires a negative \bar{k} in the range $-2k < \bar{k} < 0$. In addition, molecular-level calculations^{33,34} suggest that $|\bar{k}| \ll k$ in most circumstances.

Equation (57) indicates that the length relating the 1- and 2-D bending moduli is a molecular one, as expected. Equation (58) simply shows how one can pass from (energetic) preference for disks (positive R_0) to preference for holes (negative R_0), of any size, by either tuning the 2-D spontaneous curvature of a hypothetical film C_0 or changing \bar{k} . Most important for our estimate is Eq. (59): For (the rather large) values of C_0 corresponding to $C_0 \approx l$, we see that γ is *negative*; for typical values of R_0 (say, $R_0 \approx 5l$) $\gamma \approx -\kappa/l^2$ so that (taking, say, $a \approx 2l$) $\epsilon_{\text{pair}} \approx -K^2$. [This can lead to other interesting results, such as ribbon-(rod)-like micellar growth and stripe structures, but these will not be discussed here.] On the other hand, for smaller, more typical values of C_0 corresponding to $C_0 \ll 1/l$ (and for large, atypical, values of \bar{k} corresponding to $|\bar{k}| \approx k$ —which can make again a positive R_0) ϵ_{pair} is *positive* and its magnitude is again of the order of K^2 . (Obviously, such values of ϵ_{pair} can suppress completely the checkerboard phase and lead to an early condensation, as in the usual Ising case.) This suggests that ϵ_{pair} in our model can assume in fact any value between $\sim -K^2$ to $\sim K^2$; values of $\epsilon_{\text{pair}} \approx 0$ (i.e., $|\epsilon_{\text{pair}}| \ll K^2$), such as the one used for the numerical results discussed earlier, are thus reasonable choices for application to experimental situations. Finally we note that, in addition to the connections outlined above between our lattice Hamiltonian coupling constants and 1- and 2-D elasticity properties, it should also be possible to relate them to the considerations introduced by Israelachvili, Mitchell, and Ninham⁶ in their discussion of micellar shapes in terms of surfactant volume-to-length ratios and relative sizes of molecular heads and tails. But these latter types of argument would require generalization to include biaxiality of molecular shapes, in order to account for different preferred curvatures along the orthogonal directions of micellar surface.

B. Lattice and mean-field artifacts

An important problem arising in lattice models, which serve to approximately describe continuum phenomena, is the question of lattice artifacts. The calculation described in the previous section, for example, shows that for sufficiently large spontaneous curvature, spatial ordering (“crystalliza-

tion") will occur before condensation. Furthermore, the transition from the isotropic to the checkerboard phase was found to be second order for most relevant (equilibrium) regimes: This is because the ϕ^3 term vanishes for this particular symmetry. However, an off-lattice calculation is likely to change this transition to a *first-order* transition between isotropic and *hexagonal* (triangular) phase. The results of our lattice-gas model can therefore apply only in a qualitative way to real surfactant systems.

One tentative way to eliminate these lattice artifacts is to perform suitable angular averagings. This corresponds to approximating the sums in (31) as in, say, the following way

$$\sum_{\mathbf{q}} S(\mathbf{q})^{-1} \phi_{\mathbf{q}} \phi_{-\mathbf{q}} \approx \sum_{\mathbf{q}} \langle S(\mathbf{q})^{-1} \rangle_{\theta} \phi_{\mathbf{q}} \phi_{-\mathbf{q}}. \quad (60)$$

The structure factor $S(q)^{-1} = \langle S(\mathbf{q})^{-1} \rangle_{\theta}$ now becomes isotropic. Following this procedure, we have searched again for q_{\max} —the wave number that maximizes $S(q)$ —and have found that it is only weakly dependent on c in the range $0 < c < 0.5$, with all values lying near $q_{\max} \approx 4$. We have also calculated the new disorder line marking the onset of a *first-order* transition from an isotropic to *hexagonal* phase. But this calculation cannot be regarded as reliable, mainly because all the transition lines (including the coexistence curve for condensation) can be quite different for a continuum version of the model. (Recall, for comparison, that hard-core disks freeze *off* lattice into hexagonal order at $c \approx 0.69$ – 0.73 and that a closely packed structure is obtained at $c \approx 0.91$.³⁵)

In addition, we have used here only a combined mean-field/Landau–Ginzburg approach, which assumes small concentration fluctuations. Thus, even within the lattice model, exact critical temperatures and transition lines are expected to lie at lower temperatures than those obtained here. For example, the termination of the transition line from the isotropic to the checkerboard phase at vanishing concentration as $T \rightarrow 0$ (see Fig. 4) is evidently wrong. Specifically, one expects that this line should reach (as $T \rightarrow 0$) the critical concentration for *random* percolation of the checkerboard sublattice, which should be at³⁶ $c = 0.59/2 = 0.30$.

In principle, one could resort to more sophisticated theoretical tools in studying our lattice Hamiltonian by using, say, the quasichemical (Bethe) approximation to calculate the free energy. Additional fluctuations, which are controlled by the higher-than-quadratic-order terms in the Landau–Ginzburg expansion, can be also taken into account following the approach of Brazovskii.³⁷ He used a self-consistent “Hartree approximation” to obtain a renormalized propagator (structure factor). This renormalization was shown to be important in determining the disorder line.³⁷ Indeed, we may expect that such a calculation will shift the disorder line to *negative* values of τ_c .³⁷ Hence the new transition line (from isotropic to checkerboard) will lie below the mean-field line, and this transition will become first-order. However, these corrections cannot make this line cut the $T=0$ axis at a finite concentration, as we have just argued that it should.

Preliminary Monte Carlo simulations³⁸ suggest indeed that the checkerboard phases do not appear until the concentration is high enough for percolation to occur, i.e., for the

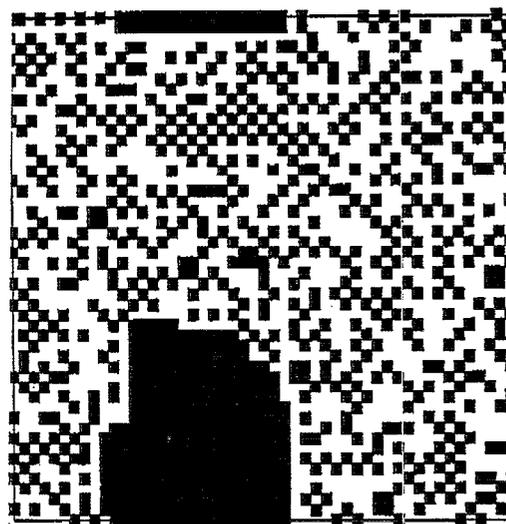


FIG. 6. Typical Monte-Carlo configuration from equilibrated 100×100 system at a temperature of $k_B T/K^2 = 1.0$ and a concentration (area fraction) of 0.4 . Here $K=b$, and $\epsilon_{\text{pair}}=0$, as in most of our mean-field examples in the text.

checkerboard clusters to become infinite on one sublattice or the other; furthermore, the temperature must be low enough for (minimum) disks to remain the dominant micellar shape. At a temperature of $k_B T/K^2 = 0.5$, for example, and for $b=K$ and $\epsilon_{\text{pair}}=0$, the Fourier-transformed density does not have a peak at $\mathbf{q}=\{\pi, \pi\}$ until the area fraction c exceeds about a third. Upon increasing c further one finds coexistence with a bilayer phase, i.e., with “condensed”/“fused” disks which at this low temperature have essentially no “holes” or “cracks” in them; the relative number of occupied sites associated with the bilayer phase increases with c in the usual way according to the level rule. At slightly higher temperature, one sees essentially the same scenario as the concentration is increased, but now the dilute phase involves some “fusion” of disks since the energy cost associated with having other-than-optimum rim curvature becomes more tolerable: Figure 6 shows this situation for $k_B T/K^2 = 1.0$ and $c=0.4$, again for $b=K$ and $\epsilon_{\text{pair}}=0$. At still higher temperatures (in particular, above the critical temperature for the condensation to bilayer, tentatively located at $k_B T/K^2 \approx 4$), the low- c phases show much more fusion (e.g., appearance of ribbons and big disks and even some highly ramified micellar shapes), while the high- c /bilayer states become riddled with holes and cracks.

C. 3-D coexistence

So far the discussion has been restricted to interactions and phase transitions in two dimensions. But when one concentrates the smectic-A phase, the system can choose between *either* increasing the concentration *within* each layer, or *decreasing* the separation d between layers. The relation between the three variables—the total volume fraction of surfactant in the system Φ , the in-layer concentration (area fraction) c , and the interlayer spacing d —follows from geometry:

$$c(\delta/d) = \Phi, \quad (61)$$

with $\delta=2l$ the bilayer (disks) thickness. Using freeze-fracture experiments (like those reported by Sammon *et al.*¹⁰), it should be possible, to a certain extent, to obtain directly the in-plane concentration. This information can also be deduced from x-ray measurements of the interlayer spacing d as a function of the overall surfactant volume fraction Φ .³⁹ Nevertheless, it remains a theoretical challenge to obtain the 3-D phase diagrams, where the only inputs are the composition and temperature. For such a calculation it is crucial to know the free-energy density of the (presumably repulsive) interaction between layers, which we denote in the following as $f_{\text{int}}(c,d)$. The total free-energy density is thus given by

$$f_{3\text{D}}(d,\Phi) = \frac{f(c)}{da^2} + f_{\text{int}}, \quad (62)$$

where $f(c)$ is the free energy per site that was obtained for the 2-D model and a is the lattice constant (as before).

For the condensed bilayer phase one can expect that f_{int} follows the power law associated with, say, Helfrich undulation repulsions, which scales as d^{-3} but is otherwise independent of concentration.⁴⁰ On the other hand, for the “dilute” checkerboard phase, for example, we expect an explicit dependence of the interaction on the in-layer concentration c ,⁴¹ and this can complicate the 3-D phase diagram. In general, in order to be able to calculate the 3-D phase diagram one needs to relate the chemical potential μ and osmotic pressure Π in the 3-D system to those obtained in the 2-D model ($\mu_{2\text{D}}$ and $\Pi_{2\text{D}}/a^2$). First, we note that the thermodynamic relation between c and Φ can be determined by minimizing $f_{3\text{D}}(d,\Phi)$ with respect to d at constant Φ , and using the relation (61).⁴² (This thermodynamic relation⁴² shows that, in equilibrium, the 2-D osmotic pressure has to be equal to an effective pressure arising from the interlayer interaction.) Then, if f_{int} is only a function of d , straightforward manipulations show that⁴³ $\mu = \mu_{2\text{D}}$ (as obviously required if the system is in equilibrium) and that the 3- and 2-D osmotic pressures are related by

$$\Pi = \frac{\Pi_{2\text{D}}}{da^2} - f_{\text{int}} + c \frac{\partial f_{\text{int}}}{\partial c} = -d \frac{\partial f_{\text{int}}}{\partial d} - f_{\text{int}}. \quad (63)$$

Recall that we are interested in the possibility of a coexistence between two different (3-D) smectic states, each characterized by its own values of interlayer spacing d and in-plane concentration c [and hence Φ —see Eq. (61)]. Accordingly, we equate the osmotic pressures of the two phases, using the second equality in Eq. (63). If $-d\partial f_{\text{int}}/\partial d - f_{\text{int}}$ is monotonic, as in the case of a power-law decay (resulting, say, from purely repulsive interactions), this equation implies $d_1 = d_2$, i.e., *the two smectic-A phases have the same interlayer spacing*. Consider now again the case where f_{int} is only a function of d . Then, since the “3-D” chemical potential is equal to the “2-D” one, one can map directly the 2-D coexistence curve to the 3-D one. The phases seen in the 2-D phase diagram will also appear in the 3-D one, since the d spacing is not changed through the transition. Such calculations are, however, beyond the scope of this paper, especially because in practice the explicit dependence of f_{int} on c must be taken into account.

D. The experimental situation

In freeze-fracture experiments on decyl-ammonium chloride (DACl) systems carried out by Sammon, Zasadzinski, and Kuzma,¹⁰ one can clearly see disklike micelles *ordered* in the lamellar planes, as suggested in the present theoretical study. The authors, however, describe their freeze-fracture pictures in terms of a “row”-like order, namely 1-D order, rather than the 2-D order suggested by our theoretical work. It is possible to explain this apparent discrepancy in a few ways. First, we note that the *real*-space pictures and their Fourier transforms (e.g., Fig. 2 in Ref. 10) appear to suggest a 2-D order. Second, it is possible that the system is not truly in equilibrium, which will then account for a (metastable) stripe phase. And third, it may be that our choice for the parameter ϵ_{pair} ($\epsilon_{\text{pair}} \ll K^2$) is not appropriate to this system, although our estimates do imply so. As both experimental and theoretical situations are not yet certain, we cannot rule out any of these possibilities. In any case, it is quite compelling that an ordered structure of *any* kind, involving layered disks, is observed, which is itself an interesting and nontrivial result.

VII. CONCLUSIONS

In this paper we have studied, for the first time, the effects of a line (rim) bending energy, i.e., a curvature-dependent—rather than constant—line tension. We have focused on the case of *nonzero spontaneous curvature*, mapping the continuous bending Hamiltonian onto a two-dimensional lattice-gas model, which has been solved using the mean-field approximation and a Landau–Ginzburg approach. We have shown that, in the absence of spontaneous curvature, the conventional self-assembly condensation to an infinite bilayer emerges already at vanishingly small concentrations; this is qualitatively the same result obtained by using a bare (noncurvature-dependent) line tension. However, for large spontaneous curvature, this condensation is postponed to significantly larger concentrations and can even be preempted by freezing of the disklike aggregates into an ordered structure. The calculation described in Sec. V suggests in particular that the disklike phase would *freeze* into a 2-D ordered structure, exemplified by a checkerboard modulation, before ultimately condensing into an infinite bilayer phase. This is a highly nontrivial scenario which merits further investigation. Simulations to test the predictions made herein are currently underway.³⁸

ACKNOWLEDGMENTS

We are grateful to Neville Boden, Robijn Bruinsma, Joe Zasadzinski, Dan Schwartz, and Heinz Hoffmann for very useful discussions. W.M.G. acknowledges the National Science Foundation, through Grant No. CHE-92-01166, for partial funding of this research. A.B.S. would like to thank the Yeshaya Horowitz Association for financial support of this work. R. G. is an incumbent of the William T. Hogan and Winifred T. Hogan Career Development Chair.

- ¹C. Tanford, *The Hydrophobic Effect* (Wiley, New York, 1980).
- ²*Physics of Amphiphiles: Micelles, Vesicles and Microemulsions*, edited by V. Degiorgio and M. Corti (North-Holland, Amsterdam, 1985).
- ³J. N. Israelachvili, *Intermolecular and Surface Forces* (Academic, London, 1992), 2nd ed..
- ⁴*Physics of Amphiphilic Layers*, edited by J. Meunier, D. Langevin, and N. Boccaro (Springer, Berlin, 1987).
- ⁵*Micelles, Membranes, Microemulsions and Monolayers*, edited by W. M. Gelbart, A. Ben-Shaul, and D. Roux (Springer, New York, 1994).
- ⁶J. N. Israelachvili, D. J. Mitchell, and B. W. Ninham, *J. Chem. Soc., Faraday Trans. 2* **72**, 1525 (1976).
- ⁷H. Wennerström and B. Lindman, *Phys. Rep.* **52**, 1 (1979).
- ⁸A. Ben-Shaul and W. M. Gelbart in Ref. 5, Chap. 1.
- ⁹N. Boden, S. A. Corne, and K. W. Jolley, *J. Phys. Chem.* **91**, 4093 (1987); N. Boden, R. J. Bushby, K. W. Jolley, M. C. Holmes, and F. Sixl, *Mol. Cryst. Liq. Cryst.* **152**, 37 (1987); N. Boden in Ref. 5, Chap. 3.
- ¹⁰M. J. Sammon, J. A. N. Zasadzinski, and M. R. Kuzma, *Phys. Rev. Lett.* **57**, 2834 (1986); J. A. N. Zasadzinski, M. J. Sammon, and M. R. Kuzma, in *Ordering and Organization in Ionic Solutions*, edited by N. Ise and I. Sogami (World Science, Singapore, 1988), p. 355.
- ¹¹M. P. Taylor and J. Herzfeld, *J. Phys.: Condens. Matter* **5**, 2651 (1993).
- ¹²C. K. Bagdassarian, D. Roux, A. Ben-Shaul, and W. M. Gelbart, *J. Chem. Phys.* **94**, 3030 (1991).
- ¹³L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, New York, 1980).
- ¹⁴R. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (Cambridge University, Cambridge, 1949); G. Parisi, *Statistical Field Theory* (Addison-Wesley, New York, 1988).
- ¹⁵L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, New York, 1986), p. 70.
- ¹⁶W. F. Helfrich, *Z. Naturforsch* **28c**, 693 (1973).
- ¹⁷Z.-G. Wang, M. E. Costas, and W. M. Gelbart, *J. Phys. Chem.* **9**, (1993); M. E. Cates and S. J. Candau, *J. Phys.: Condens. Matter* **2**, 6869 (1990).
- ¹⁸D. A. Huse and S. Leibler, *Phys. Rev. Lett.* **66**, 437 (1991).
- ¹⁹P. Chandra and S. A. Safran, *Europhys. Lett.* **17**, 691 (1992); S. T. Milner, S. A. Safran, D. Andelman, M. E. Cates, and D. Roux, *J. Phys. France* **49**, 1065 (1988); D. Andelman, M. E. Cates, D. Roux, and S. A. Safran, *J. Chem. Phys.* **87**, 7229 (1987); S. A. Safran, in *Structure and Dynamics of Strongly Interacting Colloids and Supramolecular Aggregates in Solution*, edited by S. H. Chen, J. S. Huang, and P. Tartaglia (Kluwer, Boston, 1992), p. 237.
- ²⁰The two main differences are: (i) We do not conserve the total length of line (or rim) in our model, since the same component is assumed to build both bilayer and rim, unlike in microemulsions where the total amount of interface is fixed by the surfactant concentration, and (ii) Our allowed configurations are different from those described by Safran and co-workers¹⁹ in that in our model "fusion" of two particles to make a ribbon occurs only when the two particles become nearest neighbors (in which case they must fuse). This allows us to conserve the total area fraction of particles under such topological transformations (being always equal to a given concentration c) and to incorporate into the lattice model the effect of producing "free space" under these transformations.
- ²¹Note that κ has dimensions of energy \times length and should thus be related to the usual 2-D bending modulus k (dimensions of energy) by a factor of an appropriate length, which turns out to be a molecular one (l)—see Sec. VI A and Eq. (57). Thus, taking $k \approx 10k_B T$ (Ref. 32) and $l \approx 10A$, we estimate $\kappa \approx kl \approx 10^2 k_B T A$, thereby implying from Eq. (13), with typical disk size a on the order of $10^2 A$, that $K^2 \approx \kappa/a \approx k_B T$. We note that a comparable estimate for the 1-D modulus κ is obtained from the dependence (on κ) of the typical persistence length l_p of the disk rim, $l_p \approx \kappa/(k_B T)$ (see Chap. 2 in Ref. 5 by G. Porte), by regarding the rim as (half of) a wormlike cylindrical micelle, which implies (assuming that, for the disklike micellar systems under discussion and for the wormlike micellar systems, the values of κ are comparable) $l_p \approx 100A$.
- ²²B. Widom, *J. Chem. Phys.* **84**, 6943 (1986).
- ²³M. Schick and W.-H. Shih, *Phys. Rev. B* **34**, 1797 (1986); G. Gompper and M. Schick, *Phys. Rev. B* **41**, 9148 (1990).
- ²⁴K. Chen, C. Ebner, C. Jayaprakash, and R. Pandit, *Phys. Rev. A* **38**, 6240 (1988).
- ²⁵K. A. Dawson, B. Walker, and A. Berera, *Physica A* **165**, 320 (1990); K. A. Dawson, *Phys. Rev. A* **35**, 1766 (1987).
- ²⁶M. W. Matsen and D. E. Sullivan, *Phys. Rev. A* **41**, 2021 (1990); M. W. Matsen, M. Schick, and D. E. Sullivan, *J. Chem. Phys.* **98**, 2341 (1993).
- ²⁷T. P. Stockfisch and J. C. Wheeler, *J. Phys. Chem.* **92**, 3292 (1988); *Monte Carlo Study of a Microscopic Lattice Model for Microemulsions*, preprint; A. Ciach, J. S. Hoye, and G. S. Stell, *J. Phys. A* **21**, L777 (1988).
- ²⁸C. S. O. Yokoi, M. D. Coutinho-Filho, and S. R. Salinas, *Phys. Rev. B* **24**, 4047 (1981).
- ²⁹L. Leibler, *Macromolecules* **13**, 1602 (1980).
- ³⁰D. Andelman, F. Brochard, and J. F. Joanny, *J. Chem. Phys.* **86**, 3673 (1987).
- ³¹More explicitly, for this case we get
- $$\tau_c = \frac{2k_B T}{c(1-c)} \frac{2\epsilon_1^2}{|\epsilon_2|}$$
- and
- $$\lambda_c = 2k_B T \left[\frac{1}{(1-c)^3} + \frac{1}{c^3} \right] - \left(\frac{\epsilon_1^2}{2\epsilon_2} - 1 \right) \epsilon_4.$$
- ³²See, for example, the discussion in Ref. 33.
- ³³I. Szeleifer, D. Kramer, A. Ben-Shaul, W. M. Gelbart, and S. A. Safran, *J. Chem. Phys.* **92**, 6800 (1990).
- ³⁴R. S. Cantor, *J. Chem. Phys.* **99**, 7124 (1993).
- ³⁵W. G. Hoover and F. H. Ree, *J. Chem. Phys.* **49**, 3609 (1968).
- ³⁶See, for example, Table I in H. Scher and R. Zallen, *J. Chem. Phys.* **53**, 3759 (1970).
- ³⁷S. A. Brazovskii, *Sov. Phys. JETP* **41**, 85 (1975); G. H. Fredrickson and E. Helfand, *J. Chem. Phys.* **87**, 697 (1987).
- ³⁸Y. Bohbot, A. Ben-Shaul, R. Granek, and W. M. Gelbart, unpublished.
- ³⁹P. O. Quist, K. Fontell, and B. Halle, *Microstructure and Thermodynamics of a Lamellar Phase with Disrupted Surfactant Bilayers*, preprint.
- ⁴⁰D. Sornette and N. Ostrowsky in Ref. 5, Chap. 5.
- ⁴¹If the disks in any given layer are assumed to move in the third dimension independently of one another, f_{int} should have the form (for $d \gg \delta$)
- $$f_{\text{int}} = -k_B T \frac{c}{a^2 \delta} \ln \left(\frac{d}{\delta} \right) \quad (64)$$
- ($\delta=2l$) describing the entropy penalty for squeezing a particle in one dimension.
- ⁴²Explicitly, this minimization leads to the result
- $$\frac{\Pi_{2D}(c)}{a^2} = -d^2 \frac{df_{\text{int}}}{dd} = -d^2 \left[\left(\frac{\partial f_{\text{int}}}{\partial d} \right)_c + \frac{\Phi}{\delta} \left(\frac{\partial f_{\text{int}}}{\partial c} \right)_d \right],$$
- with $\Pi_{2D}(c)$ being the 2-D dimensionless osmotic pressure used throughout this paper. (The true osmotic pressure is $\Pi_{2D}(c)/a^2$.)
- ⁴³More generally, the relation between the 2- and 3-D chemical potentials is given by
- $$\mu = \mu_{2D} + da^2 \left(\frac{\partial f_{\text{int}}}{\partial c} \right)_d.$$