J Stat Phys (2009) 137: 1166–1204
DOI 10.1007/s10955-009-9834-x

Phase Diagram of the ABC Model on an Interval

A. Ayyer · E.A. Carlen · J.L. Lebowitz · P.K. Mohanty · D. Mukamel · E.R. Speer

Received: 18 June 2009 / Accepted: 16 September 2009 / Published online: 14 October 2009

Abstract The three species asymmetric ABC model was initially defined on a ring by Evans, Kafri, Koduvely, and Mukamel, and the weakly asymmetric version was later studied by Clincy, Derrida, and Evans. Here the latter model is studied on a one-dimensional lattice of N sites with closed (zero flux) boundaries. In this geometry the local particle conserving dynamics satisfies detailed balance with respect to a canonical Gibbs measure with long range asymmetric pair interactions. This generalizes results for the ring case, where detailed balance holds, and in fact the steady state measure is known, only for the case of equal densities of the different species: in the latter case the stationary states of the system on a ring and on an interval are the same. We prove that in the limit $N \to \infty$ the scaled density profiles are given by (pieces of) the periodic trajectory of a particle moving in a quartic confining potential. We further prove uniqueness of the profiles, i.e., the existence of a single phase, in all regions of the parameter space (of average densities and temperature) except at low temperature with all densities equal; in this case a continuum of phases, differing by translation, coexist. The results for the equal density case apply also to the system on the ring, and there extend results of Clincy et al.

A. Ayyer
Institut de Physique Théorique, IPhT, CEA Saclay, and URA 2306, CNRS, 91191 Gif-sur-Yvette Cedex, France

E.A. Carlen · J.L. Lebowitz · E.R. Speer (✉)
Department of Mathematics, Rutgers University, Piscataway, NJ 08854-8019, USA
e-mail: speer@math.rutgers.edu

J.L. Lebowitz
Department of Physics, Rutgers University, Piscataway, NJ 08854, USA

P.K. Mohanty
TCMP Division, Saha Institute of Nuclear Physics, 1/AF Bidhan Nagar, Kolkata 700064, India

D. Mukamel
Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel

Keywords ABC model · Reflection asymmetric mean field model · Weakly asymmetric dynamics · Exact phase diagram · Elliptic functions

1 Introduction

One dimensional systems play an important role in statistical mechanics. In addition to their intrinsic interest as models of physical systems in confined quasi-linear geometries, they are in many cases exactly solvable and in fact are, with few exceptions, the only exactly solvable many body systems [1–4]. These exact solutions provide insights into both equilibrium and non-equilibrium collective behavior in higher dimensions.

An interesting connection between equilibrium and non-equilibrium phase transitions in one dimension is provided by the ABC model introduced by Evans et al. [5, 6] (a model with similar behavior was discussed in [7]): a one dimensional system consisting of three species of particles, labeled A, B, C, on a ring containing N lattice sites (one may equivalently regard it as a model with two species and empty sites, or as a "clock" model with three states). We will here typically let $\alpha = A$, B, or C denote a particle type, and make the convention that $\alpha + 1, \alpha + 2, \ldots$ denote the particle types which are successors to α in the cyclic order ABC. The system evolves by particle conserving nearest neighbor exchanges with asymmetric rates: $\alpha\gamma \to \gamma\alpha$ (clockwise) with rate $q_{\alpha,\gamma}$, where $q_{\alpha,\alpha+1} < q_{\alpha+1,\alpha}$; the total numbers N_α of particles of each species are conserved and satisfy $\sum_\alpha N_\alpha = N$. Evans et al. argued that this system will, in the limit $N \to \infty$ with $N_\alpha/N \to r_\alpha$ and with $r_\alpha > 0$ for all α, segregate into pure A, B, and C regions, with rotationally invariant distribution of the phase boundaries. For the case $q_{AB} = q_{BC} = q_{CA} = q < 1$ and $q_{BA} = q_{CB} = q_{AC} = 1$, the only case we will consider here, they showed further that when $N_A = N_B = N_C = N/3$ the dynamics satisfies detailed balance with respect to the Gibbs measure of a certain Hamiltonian having a long range pair interaction, so that the stationary state is an equilibrium state.

In a later development Clincy et al. [8] considered a weakly asymmetric version of this model in which $q = e^{-\beta/N}$; the stationary state for the equal density case $N_\alpha = N/3$ then becomes a Gibbs measure of the form $\exp\{-\beta \hat{E}_N\}$, with energy $\hat{E}_N$ equal, up to a constant, to the Hamiltonian of [5, 6] divided by N. The parameter β thus plays the role of an inverse temperature: $T = \beta^{-1}$. (The energy is written as E_N in [8], but here we adopt the convention that quantities with and without a circumflex refer respectively to the ring geometry and to the interval geometry introduced below.) Clincy et al. obtained the Euler-Lagrange equations for the minimizers of the free energy functional, which here is equivalent to the large deviation functional (LDF), of the rescaled particle densities $\rho_\alpha(x)$, in the limit $N \to \infty$ with $N_\alpha/N \to 1/3$ and $i/N \to x$, and showed that the uniform density profiles $\rho_\alpha(x) = r_\alpha = 1/3$ are always a solution of these equations, but that for temperatures T below the critical temperature $T_c = \beta_c^{-1} = (2\pi\sqrt{3})^{-1}$ there is also a nonuniform solution of the linearized equations which has a lower free energy than the uniform state. They interpreted this as a second order transition at T_c from the uniform to the nonuniform state, and confirmed by numerical simulations that for $T > T_c$ the system is in a single phase with essentially no correlations (they are $O(1/N)$) between the locations of particles of different species, while for $T < T_c$ there is segregation of the different species. They further argued that this transition will persist (possibly becoming first order) for unequal densities. In this case the stationary state is no longer an equilibrium state and is in fact unknown; despite this, they were able to obtain the LDF to order β^2 for all densities (see also [9]).

In the present work we consider the weakly asymmetric ABC system on a one-dimensional lattice of sites $i = 1, \ldots, N$ with zero flux boundary conditions: the dynamics

are the same as above, except that a particle at site $i = 1$ (respectively $i = N$) can only jump to the right (respectively left). We shall refer to this geometry as an interval. In contrast to the situation on the ring, there is for this system *always*, whatever the values of N_A, N_B, and N_C, an energy function E_N (see (2.1) below) such that the dynamics satisfies detailed balance with respect to an equilibrium measure $\nu_\beta = Z^{-1} \exp\{-\beta E_N\}$.

When $N_A = N_B = N_C = N/3$ the energy E_N for the system on the interval agrees with the energy $\hat{E}_N$ on the ring [8], in the sense that if we (mentally) connect site N to site 1 clockwise, and thus identify each configuration $\underline{\zeta}$ in the interval with a corresponding configuration $\underline{\hat{\zeta}}$ on the ring, then $E_N(\underline{\zeta})$ and $\hat{E}_N(\underline{\hat{\zeta}})$ agree up to a constant which depends only on the N_α. Thus the probabilities of the configurations $\underline{\zeta}$ and $\underline{\hat{\zeta}}$ (at the same β) are the same, and the invariance under rotations of $\hat{E}_N$ implies a rather surprising "rotation" invariance of E_N and of the Gibbs measure on the interval. It follows also that all the results obtained in the present paper for the interval model in the case of equal densities give corresponding results on the ring. When the N_α are not all equal one might of course use E_N similarly to define an energy of a ring configuration, but this energy would then depend on which site is chosen for the origin. The resulting Gibbs measure on the ring would be neither rotationally invariant nor invariant under the time evolution of the ABC dynamics.

Using the Gibbsian nature of the invariant measure and following [8] we can obtain the free energy functional $\mathcal{F}(\rho_A, \rho_B, \rho_C)$ of the density profiles $\rho_\alpha(x)$ in the scaling limit $N \to \infty$, $i/N \to x$. The parameters of the model in this limit are the inverse temperature β and the mean densities $r_\alpha = \int_0^1 \rho_\alpha \, dx$. Our main goal is to determine, for given values of these parameters, whether $\mathcal{F}$ has a unique minimizing profile or whether, conversely, there is more than one minimizer and thus coexistence of phases, and in either case the form of the minimizing profile(s). We are able to establish uniqueness at high temperatures by a direct study of the functional $\mathcal{F}$. We further show that minimizers must satisfy the Euler-Lagrange equations for $\mathcal{F}$, and by a study of the solutions of these equations we establish the nonuniqueness of minimizers when $r_A = r_B = r_C = 1/3$ and $T < T_c$; the nonuniqueness corresponds to the lack of rotational invariance for the minimizers on the ring. Conversely, we establish uniqueness of the minimizers (on the interval) whenever the densities are not all equal. In the process we also obtain the form of the minimizing density profiles as elliptic functions corresponding to pieces of the periodic trajectory of a particle moving in a quartic confining potential.

We remark that work on the ABC model on the ring has been carried out recently by Bodineau et al. [9]. They investigated analytically the pair correlations, and thus also the local density fluctuations, in the uniform state, and showed that the latter diverge as $T \searrow T_c$. They also suggested a possible formula for the large deviation functional and discussed the possibility of deriving this from the macroscopic fluctuation theory [10]. The model on the ring was also studied, in a different context, by Fayolle and Furtlehner [11, 12], who obtained results which agree, for the case in which the system on the ring has the same stationary Gibbs measure as does the system on the interval, with those obtained here. Results concerning the relation between the hydrodynamic equations for the ABC model and the LDF were obtained by Bertini et al. [10].

The outline of the rest of the paper is as follows. In Sect. 2 we define the Gibbs measure with respect to which the dynamics satisfy detailed balance. We then study some properties of this measure; in particular, we describe its ground states, which are nonunique whenever two species have equal mean densities which are greater than or equal to $1/3$. We investigate also some properties of the microscopic correlation functions at finite-temperature: the mean field nature of the interactions in this system leads in the $N \to \infty$ limit to local measures which are exchangeable.

In Sect. 3 we consider the scaling limit of the model and briefly discuss the nature of possible limiting density profiles. In Sect. 4 we obtain the Helmholtz free energy $\mathcal{F}$ as a functional of the scaled densities $\rho_\alpha(x),\ x \in [0,1]$ (this is equivalent to obtaining the LDF) as well as the Euler-Lagrange equations satisfied by minimizers of $\mathcal{F}$; the proof that minimizers exist and must satisfy these equations is postponed until Sect. 10. In Sect. 5 we investigate the minimizers of $\mathcal{F}$ by studying in detail the solutions of the Euler-Lagrange equations. These solutions, which describe all stationary points of $\mathcal{F}$, are given by periodic (elliptic) functions describing the motion of a particle in a one dimensional quartic potential. There are many such solutions for large β; despite this, we prove uniqueness of the globally minimizing density profiles for all β so long as the mean densities are not all equal. Details of the uniqueness proof are given in Sects. 6 and 7. In Sect. 8 we discuss the phase diagram of the model and describe a perturbation expansion around the uniform state. In Sect. 9 we show that $\mathcal{F}$ is a convex functional of the $\rho_\alpha(x)$'s (with fixed r_α's) at high temperature, which gives an alternate proof of the uniqueness of minimizers of $\mathcal{F}$ in this regime. A concluding discussion is given in Sect. 11.

2 The Steady State of the Model

A configuration $\underline{\zeta}$ of the ABC model on the interval is an N-tuple $(\zeta_1, \ldots, \zeta_N)$, with $\zeta_i = A$, B, or C. We will let $\eta_\alpha(i)$ be a random variable which specifies whether a particle of species α is present at site i: $\eta_\alpha(i) = 1$ if $\zeta_i = \alpha$ and $\eta_\alpha(i) = 0$ otherwise, so that $\eta_A(i) + \eta_B(i) + \eta_C(i) = 1$ and $N_\alpha = \sum_{i=1}^N \eta_\alpha(i)$. The time invariant measure for the weakly asymmetric dynamics described in Sect. 1 is a canonical Gibbs measure,

$$\nu_\beta(\underline{\zeta}) = Z^{-1} \exp[-\beta E_N(\underline{\zeta})], \tag{2.1}$$

where

$$E_N(\underline{\zeta}) = \frac{1}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} [\eta_C(i)\eta_B(j) + \eta_A(i)\eta_C(j) + \eta_B(i)\eta_A(j)] \tag{2.2}$$

and Z is the usual equilibrium normalization factor, that is, the (canonical) partition function with fixed particle numbers N_α. To verify this invariance one checks detailed balance for the dynamics: if $\underline{\zeta}$ is a configuration with particles of types α and $\alpha + 1$ on sites i and $i + 1$, respectively (recall the convention that α, $\alpha + 1, \ldots$ run cyclically through A, B, and C), and $\underline{\zeta}^{i,i+1}$ is the configuration with these particles interchanged, then the transition rates are that $\underline{\zeta} \to \underline{\zeta}^{i,i+1}$ at rate $e^{-\beta/N}$ and $\underline{\zeta}^{i,i+1} \to \underline{\zeta}$ at rate 1, and the detailed balance condition $e^{-\beta/N} \cdot \nu_\beta(\underline{\zeta}) = 1 \cdot \nu_\beta(\underline{\zeta}^{i,i+1})$ follows from (2.1) and (2.2). For $\beta < \infty$, ν_β is the unique stationary measure; this follows from the transitivity of the dynamics on the set of all $N!/(N_A!N_B!N_C!)$ configurations consistent with $\sum_i \eta_\alpha(i) = N_\alpha$. When $\beta = 0$ all these configurations are equally likely.

One can rewrite (2.2) in various forms which differ from each other only by functions of N_A, N_B, and N_C, and this does not affect ν_β when the N_α are fixed. As an example, an equivalent energy is

$$E_N^* = \frac{1}{N} \left\{ \sum_{j=1}^{N} j[\eta_B(j) - \eta_A(j)] + 3 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \eta_B(i)\eta_A(j) \right\}. \tag{2.3}$$

We shall generally use the form $E_N(\underline{\zeta})$ given in (2.2), as it clearly exhibits the cyclic symmetry between the different species.

Remark 2.1 (a) The energy E_N^* of (2.3) is of particular interest when the B species is not present, i.e., when $N_B = 0$. In this case we may call the A particles just particles and the C particles holes, and the model reduces to the weakly asymmetric simple exclusion process (WASEP). Then $E_N^* = -N^{-1}\sum_j j\eta_A(j)$, which is the energy arising from an external field of magnitude N^{-1} pushing the particles to the right. See [13] and Sect. 7.2.2 of the review [14] for discussions of the partially asymmetric simple exclusion process on an interval.

(b) On the ring, the stationary measure for the WASEP gives equal weight to all $\binom{N}{N_A}$ configurations. In contrast to the situation on the interval, however, the dynamics is not reversible, i.e., does not satisfy detailed balance, with respect to this stationary measure. We thus have a true nonequilibrium stationary state (NESS).

2.1 Ground States

For $\beta = \infty$ there are many configurations stationary for the dynamics: all those of the form $\cdots\mathbf{ABCAB}\cdots$, where a boldface letter denotes a block of particles of the corresponding species. These configurations are local minimizers of the energy $E_N(\underline{\zeta})$, in the sense that any interchange of adjacent particles of different species will raise the energy, but not all of them are obtained as limits of ν_β when $\beta \to \infty$. In this limit the measure will be concentrated on the ground states of E_N, that is, on the global minimizers of $E_N(\underline{\zeta})$ [15]. As we will see, the energy per particle in the ground state $\underline{\zeta}$ is $e_0 = N^{-1}E_N(\underline{\zeta}) = \min\{r_A r_B, r_A r_C, r_B r_C\}$, where $r_\alpha \equiv N_\alpha/N$.

We describe the ground states by considering several cases. It is helpful to note that (2.2) may be summarized as saying that there is a contribution of $1/N$ to the energy each time a B particle lies to the left of an A particle, a C to the left of a B, or an A to the left of a C. We also remark that if a local minimizer contains a sequence **BCAB** of four blocks with a total of k_α particles of type α, then one may lower the energy by regrouping these particles into three blocks, **BCA**, **ABC**, or **CAB**, unless $k_A = k_C \geq k_B$ (and similarly for **ABCA** and **CABC**).

(i) If one species is not present, say $N_\alpha = 0$, then the ground state energy is zero and there is a unique ground state configuration with all particles of type $\alpha + 1$ to the left of all particles of type $\alpha + 2$. This is the WASEP discussed in Remark 2.1. In this case there are no local minimizers other than the global minimizer.

From now on we suppose that none of the N_α vanish.

(ii) If one of the N_α is greater than the other two, say $N_B > \max\{N_A, N_C\}$ (other cases can then be found by cyclic permutation), then it is easy to see that there is a unique ground state $\underline{\zeta}$ consisting of three blocks ordered as **ABC**, with $e_0 = r_A r_C$.

(iii) If two of the N_α are equal and the third is smaller, say $N_A = N_C > N_B$, the ground state will be $N_B + 1$ fold degenerate, consisting of three or four blocks ordered as **BCA**, **CAB**, or **BCAB**, that is, some of B's will be at the left side of the interval, followed by all the C's, then all the A's, and finally the remaining B's. Here $e_0 = r_A r_B = r_B r_C$.

(iv) If all of the N_α are equal, $N_A = N_B = N_C$, then the ground states are the shifts, by an arbitrary number of sites, of the configuration with three blocks ordered as **ABC**; see the discussion in Sect. 1 of the correspondence in this case with the model on the ring. The ground state is thus N-fold degenerate, and $e_0 = 1/9$.

The ground state phase diagram of the system in the right triangle where $r_A, r_B \geq 0$ and $r_A + r_B \leq 1$ is given in Fig. 1 (remember that $r_C = 1 - r_A - r_B$). The ground state is unique everywhere except on the three line segments originating from $r_A = r_B = 1/3$ and

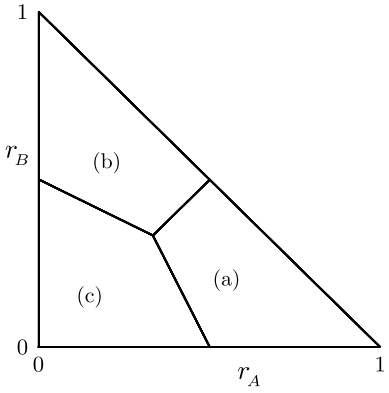

Fig. 1 The (r_A, r_B) ground state phase diagram. It exhibits three phases separated by three first order lines. In region (a), the A density is larger than that of the B's and C's and hence the order is such that the A-block is in the middle. Similarly (b) and (c) correspond to regions where B and C densities are largest

terminating at the midpoints of the sides of the triangle, corresponding to cases (iii) and (iv) above. On the edges of the triangle the system reduces to the WASEP.

2.2 Finite Temperature Correlations and Local States

The correlation functions at finite β are

$$\langle \eta_{\alpha_1}(i_1)\eta_{\alpha_2}(i_2)\cdots\eta_{\alpha_n}(i_n)\rangle, \tag{2.4}$$

where $\langle\cdot\rangle$ denotes the expectation with respect to the Gibbs measure ν_β. We point out here some simple relations that these satisfy, both in finite volume and in the limit $N\to\infty$, $N_\alpha/N\to r_\alpha$. (With some abuse of notation we will use r_α to denote both the density N_α/N in a finite system and the limiting value of the density in the $N\to\infty$ limit.)

First, it is easy to check from (2.1) and (2.2) that, whenever all the N_α are positive,

$$\frac{\langle\eta_\alpha(N)\rangle}{\langle\eta_\alpha(1)\rangle} = e^{\beta[r_{\alpha+2}-r_{\alpha+1}]}, \tag{2.5}$$

and this implies that for all N,

$$\prod_{\alpha=1}^{3}\langle\eta_\alpha(N)\rangle = \prod_{\alpha=1}^{3}\langle\eta_\alpha(1)\rangle. \tag{2.6}$$

Note that when $r=(1/3,1/3,1/3)$ and the system on the interval coincides with that on the ring, (2.6) is a consequence of the translation invariance of the latter.

Next we note that there is a constant K, which depends only on β and the r_α, such that for fixed $x\in(0,1)$, j_1 and j_2 integers with $0<j_1<j_2$, and α_0, α_1, and α_2 distinct particle species,

$$\lim_{N\to\infty}\langle\eta_{\alpha_0}(\lfloor xN\rfloor)\eta_{\alpha_1}(\lfloor xN\rfloor+j_1)\eta_{\alpha_2}(\lfloor xN\rfloor+j_2)\rangle = K, \tag{2.7}$$

where $\lfloor xN\rfloor$ denotes the largest integer less or equal to xN. To see that the limit in (2.7) is independent of x, j_1, j_2, α_0, α_1, and α_2 we first write $i_0=\lfloor xN\rfloor$, $i_1=\lfloor xN\rfloor+j_1$, and $i_2=\lfloor xN\rfloor+j_2$, and observe that if $\underline{\zeta}$ is a configuration with $\zeta_{i_0}=\alpha_0$, $\zeta_{i_1}=\alpha_1$, and $\zeta_{i_2}=\alpha_2$, and $\underline{\xi}$ is the configuration on $N-3$ sites which is obtained from $\underline{\zeta}$ by omitting the sites i_1, i_2, and i_3 and renumbering, then $E_N(\underline{\zeta})=E_{N-3}(\underline{\xi})+1+\Delta$; here the terms $1+\Delta$ arise

from those terms in the energy (2.2) involving the particles on sites i_1, i_2, and i_3, and Δ, which depends on ξ, the i_k, and the α_k, satisfies $|\Delta| \leq j_2/N$. Thus from (2.1) and (2.2),

$$Z^{-1}\langle \eta_{\alpha_0}(i_0)\eta_{\alpha_1}(i_1)\eta_{\alpha_2}(i_2)\rangle = \sum_{\xi} e^{-\beta E_{N-3}(\underline{\xi})+1+\Delta}. \tag{2.8}$$

The corresponding expression for a different choice of the i_k and α_k will differ from (2.8) only in the replacement of Δ by some Δ', from which the result follows.

The results (2.6) and (2.7) are closely related: each shows that a product of three occupation numbers, or their expectations, is site independent. We will see another such relation in Sect. 5: for all scaled density profiles $\rho_\alpha(x)$ which are stationary points of the free energy functional (see Sect. 4) the product $\rho_A(x)\rho_B(x)\rho_C(x)$ is in fact independent of x.

Similar identities can help us understand the nature of local states [16] in this model. By the *local state* $\mu_{(x)}$ at position $x \in (0, 1)$ we mean the infinite volume state, if it exists, whose correlation functions for particles of type $\alpha_1, \ldots, \alpha_h$ at sites $\lfloor xN \rfloor + j_1, \ldots, \lfloor xN \rfloor + j_h$ are given by

$$\lim_{N\to\infty} \left\langle \prod_{i=1}^{h} \eta_{\alpha_i}(\lfloor xN \rfloor + j_i) \right\rangle. \tag{2.9}$$

(If such a local state does not exist then one may identify subsequences N_k along which all the limits (2.9) do exist, and thus obtain a family of local states.) An argument similar to that establishing (2.7) shows that for $x \in (0, 1)$, $j_1, \ldots, j_h$ fixed, and P any permutation of $\{1, \ldots, h\}$,

$$\lim_{N\to\infty} \frac{\left\langle \prod_{i=1}^{h} \eta_{\alpha_i}(\lfloor xN \rfloor + j_i) \right\rangle}{\left\langle \prod_{i=1}^{h} \eta_{\alpha_{P(i)}}(\lfloor xN \rfloor + j_i) \right\rangle} = 1. \tag{2.10}$$

This implies [17] that $\mu_{(x)}$ is a product measure or a superposition of product measures, which of course reflects the mean field nature of the interaction. Dynamically this corresponds to the fact that when $N \to \infty$, the exchanges become symmetric and the invariant measures are just superpositions of extremal product measures [16]. (In fact this last argument applies also on the ring with unequal densities r_α.)

3 The Scaling Limit

We now turn to the properties of the model in the macroscopic scaling limit

$$N \to \infty, \qquad N_\alpha/N \to r_\alpha, \qquad i/N \to x. \tag{3.1}$$

We say that there is a unique (macroscopic) density profile in this limit when there exists a function $\rho(x) = (\rho_A(x), \rho_B(x), \rho_C(x))$ such that [18]

$$\lim_{N\to\infty} \text{Prob}\left\{ \left| \frac{1}{N} \sum_{i=1}^{N} \eta_\alpha(i)\varphi_\alpha\left(\frac{i}{N}\right) - \int_0^1 \rho_\alpha(x)\varphi_\alpha(x)\,\mathrm{d}x \right| > \delta \right\} = 0, \tag{3.2}$$

for all α, all $\delta > 0$, and all piecewise smooth functions $\varphi_\alpha(x)$ on the interval $[0, 1]$. Such $\rho_\alpha(x)$ must, of course, satisfy the conditions

$$0 \le \rho_\alpha(x) \le 1 \quad \text{and} \quad \int_0^1 \rho_\alpha(x)\,\mathrm{d}x = r_\alpha \quad \text{for } \alpha = A, B, C, \qquad \sum_\alpha \rho_\alpha(x) = 1. \tag{3.3}$$

Equation (3.2) expresses the convergence, in a certain sense, of the random profile $N^{-1}\sum_i \eta_\alpha(i)\delta_{i/N}$ to the non-random profile $\rho_\alpha(x)$: specifically, for each φ_α the random variable $N^{-1}\sum_i \eta_\alpha(i)\varphi_\alpha(i/N)$ converges in the sense of (3.2) to the non-random quantity $\int \rho_\alpha(x)\varphi_\alpha(x)\,\mathrm{d}x$. It may happen, on the other hand, that the limiting profile is itself random, in the sense that there is some family of limiting profiles $\rho(\omega) = \{\rho_\alpha(\omega; x)\}$, indexed by the variable ω lying in some sample space Ω and each satisfying (3.3), and a measure κ on Ω, such that for each test function $\varphi_\alpha(x)$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \eta_\alpha(i)\varphi_\alpha\left(\frac{i}{N}\right) = \int_0^1 \rho_\alpha(\cdot; x)\varphi_\alpha(x)\,\mathrm{d}x \tag{3.4}$$

in the sense of convergence in distribution. (In fact, (3.2) is also convergence in distribution, to a non-random limit.) We would then say that the limiting profile is not unique and that the system can exist in more than one phase, i.e., is in a phase transition region. This is exactly what happens to the ferromagnetic Ising model in two or more dimensions in a periodic box at temperatures below the critical temperature and fixed magnetization $m \in (-m^*, m^*)$, where m^* is the spontaneous magnetization [19]. In our system this happens, as we shall see, if and only if $r_A = r_B = r_C$ and $T < (2\pi\sqrt{3})^{-1}$. More generally, it may be that the limit in (3.4) will not exist; in that case, one would expect to obtain a limit in the given form by passing to a subsequence.

4 The Free Energy

The question of whether or not there exists a limiting profile(s) as given by (3.2) or (3.4), and of which of these alternatives applies, is related to the question of the existence and uniqueness of minimizers of the LDF for this system [20]. If $n(x) = (n_\alpha(x))_{\alpha=A,B,C}$ denotes a general continuum density profile satisfying (3.3) (written in terms of n rather than ρ) then the logarithm of the probability of finding the profile $n(x)$ in the scaling limit is asymptotically $-N\mathcal{F}_{\mathrm{LD}}(\{n(x)\})$, where $\mathcal{F}_{\mathrm{LD}}$ is the LDF. Possible limiting density profiles are thus those which minimize $\mathcal{F}_{\mathrm{LD}}$.

To obtain from the microscopic measure the probability of finding the profile $n(x)$ it is necessary (roughly speaking) to sum $\nu_\beta(\zeta)$, as given in (2.1), over all (microscopic) configurations ζ consistent with $n(x)$; this summation will yield a restricted partition function divided by the full partition function Z. Taking the logarithm of this ratio, dividing by N, and taking the scaling limit of the result will then yield

$$\mathcal{F}_{\mathrm{LD}}(\{n(x)\}) = \mathcal{F}(\{n(x)\}) - \inf_{n(x)} \mathcal{F}(\{n(x)\}), \tag{4.1}$$

where $\beta^{-1}\mathcal{F}(\{n(x)\})$ is the Helmholtz free energy of the system restricted to having a density profile $n(x)$:

$$\mathcal{F}(\{n(x)\}) = \beta\mathcal{E}(\{n(x)\}) - \mathcal{S}(\{n(x)\}). \tag{4.2}$$

Here the macroscopic energy $\mathcal{E}(\{n(x)\})$ and entropy $\mathcal{S}(\{n(x)\})$ are the limits of the microscopic energy and entropy per site; as indicated, these limits depend only on the profile $n(x)$. For the energy, this is due to the mean field nature of the microscopic energy E_N of (2.2). For the entropy, which is the logarithm of the number of microscopic configurations consistent with $n(x)$, this is due to the Bernoulli nature of each component of the local measure, which implies that all microscopic configurations consistent with $n(x)$ have the same energy and hence the same weight. Thus

$$\begin{aligned}\mathcal{F}(\{n\}) &= \beta \int_0^1 \mathrm{d}x \int_x^1 \mathrm{d}z \, [n_A(x)n_C(z) + n_B(x)n_A(z) + n_C(x)n_B(z)] \\ &\quad + \int_0^1 \mathrm{d}x \, [n_A(x)\ln n_A(x) + n_B(x)\ln n_B(x) + n_C(x)\ln n_C(x)] \\ &= \sum_\alpha \int_0^1 \left[\beta \int_0^1 \Theta(z-x) n_\alpha(x) n_{\alpha+2}(z)\,\mathrm{d}z + n_\alpha(x)\ln n_\alpha(x) \right] \mathrm{d}x. \end{aligned} \tag{4.3}$$

The profiles $n(x)$ minimizing (4.3)—that is, the candidates for limiting density profiles—will represent a compromise between the entropy, which wants to keep all the densities uniform, and the energy, which wants to keep the different particle types segregated.

The next theorem establishes the existence and properties of these minimizers.

Theorem 4.1 *Suppose that $0 < \beta < \infty$. Then given positive numbers r_A, r_B, and r_C with $r_A + r_B + r_C = 1$, let $F(r_A, r_B, r_C)$ denote the infimum of the values of $\mathcal{F}(\{n\})$ over (measurable) profiles $n(x)$ satisfying the constraints (3.3). Then there exist infinitely differentiable functions $\rho_A(x)$, $\rho_B(x)$ and $\rho_C(x)$, possibly non-unique, that satisfy these constraints and achieve the minimum*:

$$\mathcal{F}(\{\rho\}) = F(r_A, r_B, r_C). \tag{4.4}$$

Moreover, there is a $\delta > 0$ so that each ρ_α satisfies $\delta < \rho_\alpha(x) < 1 - \delta$ for all x; thus the minimizer is an interior point with respect to the constraint of taking values in $[0, 1]$ *and so satisfies the Euler-Lagrange equations (ELE) obtained from $\mathcal{F}$.*

The proof of this theorem is given in Sect. 10. We note here, however, that the minimizers lie in the interior of the constraint region essentially because the entropy term has infinite normal derivative at the boundary. For $\beta = \infty$, when only the energy counts, the densities minimizing $\lim_{\beta\to\infty} \beta^{-1}\mathcal{F} = \mathcal{E}$ will be the continuum limit of the ground state configurations described in Sect. 2.1 and need not satisfy the ELE.

To obtain the Euler-Lagrange equations satisfied by the minimizing profiles we must take the variational derivatives of $\mathcal{F}(\{n\})$ with respect to two of the density profiles while maintaining $\sum_\alpha n_\alpha = 1$ and $\int_0^1 n_\alpha(x)\,\mathrm{d}x = r_\alpha$; here we will treat $n_A(x)$ and $n_B(x)$ as independent, with $n_C(x) = 1 - n_a(x) - n_B(x)$. Defining $\mathcal{F}_\alpha(x) = \delta\mathcal{F}/\delta n_\alpha(x)$ to be the variational derivative taken as if the profiles $n_A(x)$, $n_B(x)$, and $n_C(x)$ were independent, we will then have

$$\left.\frac{\delta\mathcal{F}}{\delta n_A}\right|_{n_C = 1-n_A-n_B} = \mathcal{F}_A - \mathcal{F}_C, \qquad \left.\frac{\delta\mathcal{F}}{\delta n_B}\right|_{n_C = 1-n_A-n_B} = \mathcal{F}_B - \mathcal{F}_C. \tag{4.5}$$

Imposing the constraints $\int_0^1 n_\alpha(x)\,\mathrm{d}x = r_\alpha$ leads to the ELE

$$\mathcal{F}_A - \mathcal{F}_C = \text{constant}, \qquad \mathcal{F}_B - \mathcal{F}_C = \text{constant}, \tag{4.6}$$

where

$$\mathcal{F}_\alpha(x) = \log n_\alpha(x) + \beta \int_0^x [n_{\alpha+1}(z) - n_{\alpha+2}(z)]\,\mathrm{d}z + 1 + \beta r_{\alpha+2}. \tag{4.7}$$

Simple manipulations then show that under the constraint $\sum_\alpha n_\alpha(x) = 1$ the derivatives $\mathcal{F}_\alpha$ satisfy

$$\sum_{\alpha=A,B,C} n_\alpha \frac{\partial}{\partial x}\mathcal{F}_\alpha = 0, \tag{4.8}$$

which implies, rather surprisingly, that solutions of (4.6) will satisfy $\mathcal{F}_\alpha(x) =$ constant for all α, i.e., that the functional derivatives of $\mathcal{F}$ can be taken as if the n_α were independent. From (4.7), then, the minimizing profile $\rho(x)$ must satisfy the ELE

$$\rho_\alpha(x) = \rho_\alpha(0) e^{\beta \int_0^x [\rho_{\alpha+2}(y) - \rho_{\alpha+1}(y)]\,\mathrm{d}y}, \tag{4.9}$$

for all $x \in [0, 1]$. It follows that

$$\rho_\alpha(1) = \rho_\alpha(0) e^{\beta(r_{\alpha+2} - r_{\alpha+1})}, \tag{4.10}$$

which is consistent with (2.5). The ELE may also be written in differential form:

$$\frac{\mathrm{d}\rho_A}{\mathrm{d}x} = \beta\rho_A(\rho_C - \rho_B), \tag{4.11a}$$

$$\frac{\mathrm{d}\rho_B}{\mathrm{d}x} = \beta\rho_B(\rho_A - \rho_C), \tag{4.11b}$$

$$\frac{\mathrm{d}\rho_C}{\mathrm{d}x} = \beta\rho_C(\rho_B - \rho_A). \tag{4.11c}$$

These equations were derived in [8] for the case $r = (1/3, 1/3, 1/3)$. One obtains minimizing profiles by solving these equations with the constraints

$$\int_0^1 \rho_\alpha(x)\,\mathrm{d}x = r_\alpha. \tag{4.12}$$

Remark 4.2 As indicated above, one may consider $\mathcal{F}$ as a function of n_A and n_B alone, with $n_C(x) = 1 - n_A(x) - n_B(x)$. The corresponding variational derivatives $\lambda_A(x)$ and $\lambda_B(x)$ are the local chemical potentials for the density profiles,

$$\begin{aligned} \lambda_A &= \mathcal{F}_A - \mathcal{F}_C, \\ \lambda_B &= \mathcal{F}_B - \mathcal{F}_C, \end{aligned} \tag{4.13}$$

which are constant at a stationary point of $\mathcal{F}$ (see (4.6)). These are given by

$$\begin{aligned} \frac{\delta\mathcal{F}}{\delta n_A(x)} &= \lambda_A(\{n_A\}, \{n_B\}) \\ &= \log\left[\frac{n_A(x)}{n_C(x)}\right] - \beta \int_0^x (1 - 3n_B(z))\,\mathrm{d}z + \beta(r_C - r_B), \end{aligned} \tag{4.14}$$

$$\frac{\delta \mathcal{F}}{\delta n_B(x)} = \lambda_B(\{n_A\}, \{n_B\})$$
$$= \log\left[\frac{n_B(x)}{n_C(x)}\right] - \beta \int_0^x (1 - 3n_A(z))\, \mathrm{d}z + \beta(r_A - r_C), \tag{4.15}$$

with the replacements $n_C(x) = 1 - n_A(x) - n_B(x)$ and $r_C = 1 - r_A - r_B$.

5 Solutions of the ELE Equations

We now turn to a detailed study of the solutions of the Euler-Lagrange equations (4.11) with the constraints (4.12). As already noted in Theorem 4.1, for any given β and r_A, r_B, r_C there must exist at least one solution of these, corresponding to a minimizer of $\mathcal{F}$. There may, however, be many solutions; these will all correspond to stationary points of $\mathcal{F}$, and more than one of these stationary points may be a minimizer. In this section we determine all the minimizers.

5.1 The WASEP

We first consider briefly the case in which one of the densities is zero, say $r_B = 0$, when the system reduces to the WASEP (see Remark 2.1). This system, which can be solved explicitly on the interval, has in the scaling limit a unique limiting density profile—i.e., minimizer of the free energy—given by the solution of (4.11) with $\rho_B(x) = 0$:

$$\rho_A(x) = \frac{De^{\beta x}}{1 + De^{\beta x}} = 1 - \rho_C(x), \quad x \in [0, 1]. \tag{5.1}$$

The constant D is determined by the constraint $\int_0^1 \rho_A(x)\, \mathrm{d}x = r_A = 1 - r_C$. When $\beta \to \infty$, (5.1) reduces to the ground state configuration in which all the A particles are pushed to the right:

$$\rho_A(x)\big|_{\beta=\infty} = \begin{cases} 0, & \text{if } 0 \le x \le 1 - r_A, \\ 1, & \text{if } 1 - r_A \le x \le 1. \end{cases} \tag{5.2}$$

The local measure $\mu_{(x)}$ in this model is, as expected, the Bernoulli measure with density $\rho_A(x)$. (On the ring the stationary state of the WASEP is a nonequilibrium one in which all configurations have equal weight and the limiting scaled density profile is constant.)

5.2 Properties of Solutions

Through the remainder of this section we suppose that all the r_α are strictly positive. Let us begin by discussing the properties of some given solution $\rho(x) = (\rho_A(x), \rho_B(x), \rho_C(x))$ of the ELE equations (4.11) with the constraints (4.12). By differentiating $\log(\rho_A(x)\, \rho_B(x)\, \rho_C(x))$ with respect to x we see that the ELE imply that there is a constant K such that

$$\rho_A(x)\, \rho_B(x)\, \rho_C(x) = K, \quad 0 \le x \le 1, \tag{5.3}$$

and that the condition

$$\rho_A(x) + \rho_B(x) + \rho_C(x) = 1, \quad 0 \le x \le 1, \tag{5.4}$$

is preserved by the equations. Equation (5.3) is the scaling limit version of (2.7).

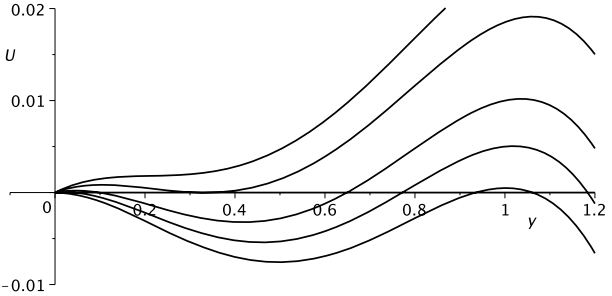

Fig. 2 Plots of $U_K(y)$ for $K = 1/20, 1/27, 1/50, 1/100,$ and $1/1000$

Since $\rho_A(x)$ and $\rho_C(x)$ have sum $1 - \rho_B(x)$ and product $K/\rho_B(x)$ they must be the two roots of the equation $r^2 + (1 - \rho_B(x))r + K/\rho_B(x) = 0$, from which

$$\rho_A(x) - \rho_C(x) = \pm\sqrt{\frac{\rho_B(x)(1-\rho_B(x))^2 - 4K}{\rho_B(x)}}. \tag{5.5}$$

Then squaring (4.11b) and using (5.5) we find that $\rho_B(x)$ is a solution of

$$\rho'(x)^2 + 8\beta^2 U_K(\rho(x)) = 0, \tag{5.6}$$

where

$$U_K(\rho) \equiv \frac{1}{2}K\rho - \frac{1}{8}\rho^2(1-\rho)^2. \tag{5.7}$$

Of course, the same argument shows that $\rho_A(x)$ and $\rho_C(x)$ must also be solutions of (5.6).

To study (5.6) we write $t = 2\beta x$ and let $y(t) = \rho(t/2\beta)$; then y satisfies

$$\frac{1}{2}y'(t)^2 + U_K(y(t)) = 0. \tag{5.8}$$

This is the equation of the zero energy solution of a mass 1 particle moving in a potential U_K. Because U_K is quartic in y the solutions are elliptic functions; see Appendix A. For $K > 1/27$, $U_K(y)$ is strictly positive for $0 < y < 1$ and so no solutions with y in this range exist. For $K = 1/27$, $U_K(y)$ has a local minimum with value 0 at $y = 1/3$, and (5.8) has constant solution $y(t) = 1/3$. For $0 < K < 1/27$, U_K has four zeros, 0, $a(K)$, $b(K)$, and $c(K)$, where $0 < a < b < 1 < c$ and $U_K(y) < 0$ for $a < y < b$. See Fig. 2. Since we are interested in solutions of (4.11) which satisfy $0 < \rho_\alpha(x) < 1$ we consider only the solutions of (5.8) which oscillate between a and b. Let y_K denote the solution of (5.8) which satisfies $y_K(0) = a$; $y_K(t)$ then has period

$$\tau_K = 2\int_a^b \frac{\mathrm{d}y}{\sqrt{-2U_K(y)}}. \tag{5.9}$$

For $0 \le t \le \tau_K/2$, $y_K(t)$ is determined by inverting the relation

$$t = t(y) = \int_a^y \frac{\mathrm{d}w}{\sqrt{-2U_K(w)}}; \tag{5.10}$$

$y_K(t)$ is then obtained for all t by extending to an even function of period τ_K.

Remark 5.1 We note some properties of this solution; those listed in (a) are elementary, and we relegate the proofs of (b) and (c) to Appendix B.

(a) $y_K(t)$ is even, is periodic with minimal period τ_K, has local minima at integer multiples of τ_K and local maxima at half integer multiples of τ_K, and is monotonic between these points. Moreover, $y_K(t) = y_K(s)$ if and only if either $s+t$ or $s-t$ is an integer multiple of τ_K.

(b) $\lim_{K\nearrow 1/27} \tau_K = 4\pi\sqrt{3} = 2\beta_c$ and $\lim_{K\searrow 0}(\tau_K/\ln(1/K)) = 6$. Moreover, for any $\epsilon > 0$,

$$\lim_{K\to 0} \frac{1}{\ln(1/K)} \int_a^\epsilon \frac{\mathrm{d}y}{\sqrt{-2U_K(y)}} = 2, \tag{5.11}$$

and

$$\lim_{K\to 0} \frac{1}{\ln(1/K)} \int_{b-\epsilon}^b \frac{\mathrm{d}y}{\sqrt{-2U_K(y)}} = 1. \tag{5.12}$$

This means that in the limit $K \to \infty$ the trajectory $y_K(t)$ will spend two-thirds of its time near the origin and one-third near $y = 1$; see for example the third trajectory shown in Fig. 4.

(c) τ_K is a strictly monotonic decreasing function of K for $0 < K \le 1/27$.

Now for $\alpha = A, B, C$ there must be a phase shift t_α such that

$$\rho_\alpha(x) = y_K(2\beta(x - 1/2) + t_\alpha), \quad 0 \le x \le 1, \tag{5.13}$$

that is, each $\rho_\alpha(x)$ is obtained by looking at the solution $y_K(t)$ within a window of length 2β centered at some value t_α, and rescaling from t to x. The phase shifts are not independent; in fact,

$$t_A = t_B + \tau_K/3 \quad \text{and} \quad t_C = t_B - \tau_K/3 \tag{5.14}$$

(see (27) of [8]). In verifying (5.14) the requirement (4.12) of certain average densities is not relevant. What matters is that we consider three solutions of the ELE (4.11) or equivalently (with the rescaling $t = 2\beta x$) three solutions $y_{K,\alpha}(t)$ of (5.8) satisfying

$$\begin{aligned} y'_{K,A} &= y_{K,A}(y_{K,C} - y_{K,B})/2, \\ y'_{K,B} &= y_{K,B}(y_{K,A} - y_{K,C})/2, \\ y'_{K,C} &= y_{K,C}(y_{K,B} - y_{K,A})/2, \end{aligned} \tag{5.15}$$

with $y_{K,\alpha}(t) = y_K(t + t_\alpha)$. We may assume that these are defined for all t and without loss of generality that $t_B = 0$. It is helpful to view the trajectories in the y–y' phase plane; see Fig. 3. At time $t = 0$ the B trajectory lies at the point on the y axis marked B1. Since the velocity on the B trajectory is zero at this point, it follows from (5.15) that $y_{K,A}(0) = y_{K,C}(0)$, that is, the A and C trajectories lie at points such as those marked A1 and C1, respectively; again from (5.15) we know that A1 must be in the upper half of the phase plane, since $y'_{K,A}(0) > 0$. Now follow the motion for a time period Δt until $y_{K,A}$ reaches the y axis at point A2; since now $y'_{K,A}(\Delta t) = 0$ it follows that the B and C trajectories are at points B2 and C2 with equal y coordinates, and it follows immediately (using the symmetry around the y axis) that the travel time along each of the six orbital segments in the figure is Δt. Thus $\Delta t = \tau_K/6$, $t_A = t_B + 2\Delta t = t_B + \tau_K/3$, and $t_C = t_B - \tau_K/3$.

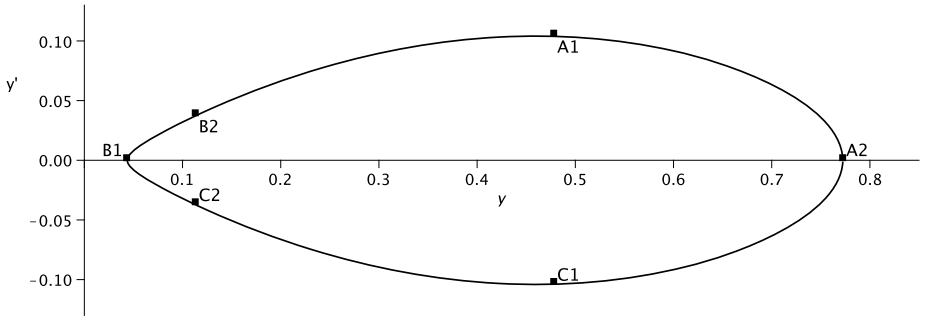

Fig. 3 Phase plane orbit of $y_K(t)$ for $K = 1/100$. The time intervals between the six marked points are all equal

Let us introduce the notation

$$\langle z \rangle_I = \frac{1}{|I|} \int_I z(t)\, \mathrm{d}t, \tag{5.16}$$

where $|I| = \int_I \mathrm{d}t$, for the average of the function z over the set I (which will always be a union of intervals), and define $Y_K(t) = \langle y_K \rangle_{[t-\beta, t+\beta]}$. Then with (5.13) and (5.14), (4.12) becomes

$$Y_K(t_B) = r_B, \qquad Y_K(t_B + \tau_K/3) = r_A, \qquad Y_K(t_B - \tau_K/3) = r_C. \tag{5.17}$$

The problem of solving (4.11) and (4.12) is now the problem of finding K and t_B satisfying (5.17). If either $K = 1/27$ so that $y_K(t) = 1/3$ is constant, or 2β is an integer multiple of τ_K, then $Y_K(t) = 1/3$ is constant and (5.17) has a solution (with arbitrary t_B) if and only if $r_A = r_B = r_C = 1/3$.

Figure 4 shows the curves $y_K(t)$ and $y_K(t \pm \tau_K/3)$ for several values of K. To obtain a solution of (5.17) for some K one views the corresponding three curves in a window of length 2β (corresponding to the full lattice, i.e., to the original unit interval after the variable change in (5.13)) centered at t_B. In this context we label a solution by an integer which is one more than the number of full periods (plus, perhaps, a fraction of a period) fit into the window: we say that a solution $\rho(x)$ of (5.17) with $K < 1/27$ is of *type* n, for $n = 1, 2, \ldots$, if

$$(n-1)\tau_K < 2\beta \le n\tau_K. \tag{5.18}$$

We do not assign a type to the constant solution $y_K = 1/3$, which exists for $K = 1/27$, $r_A = r_B = r_C = 1/3$, and every value of β, as discussed above.

5.3 Uniqueness of Solutions

The next theorem summarizes our results about solutions of the ELE and answers the questions posed in the first paragraph of this section. We consider the problem of finding y_K and t_B satisfying (5.17).

Theorem 5.2 (a) *If* $r_A = r_B = r_C = 1/3$ *then there exist* (i) *the constant solution*, (ii) *for* $\beta > n\beta_c = 2\pi n\sqrt{3}$, $n = 1, 2, \ldots$, *a solution of type* n *unique up to translation*, *and* (iii) *no other*

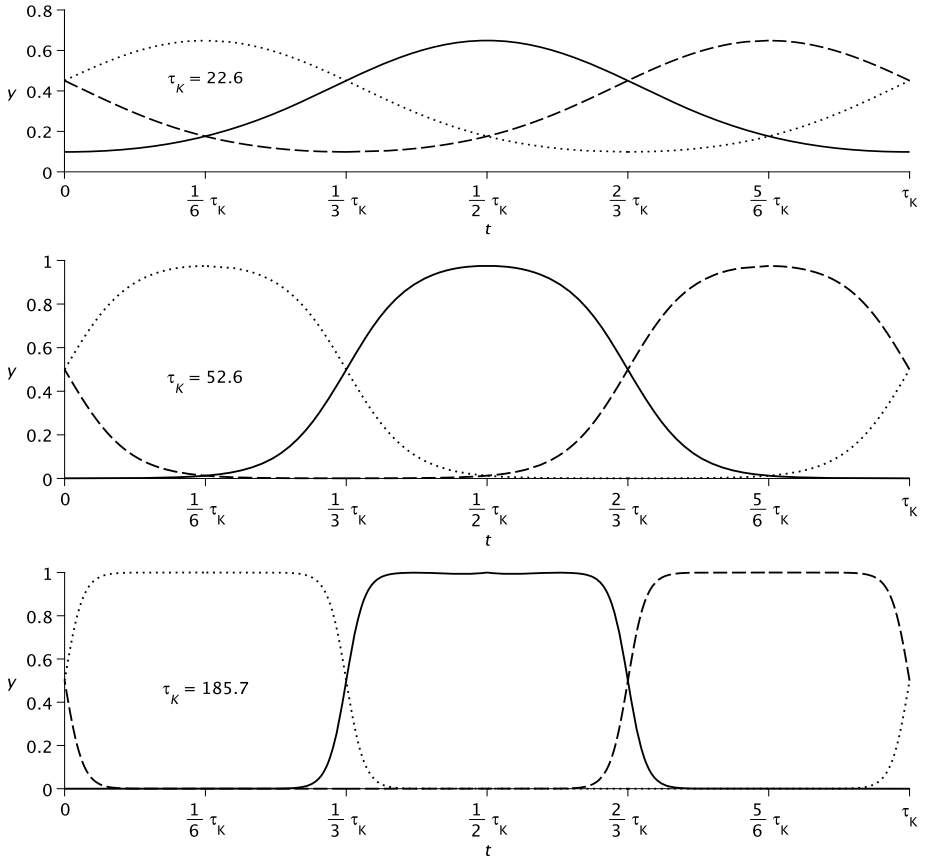

Fig. 4 Plots of $y_K(t)$ (*solid*), $y_K(t+\tau_K/3)$ (*dotted*), and $y_K(t-\tau_K/3)$ (*dashed*) for $K = 1/50$ (*top*), $K = 1/6400$ (*middle*), and $K = 5^{-2}2^{-40} = 1/27487790694400$ (*bottom*)

solutions. The minimizer of the free energy is, for $\beta \le \beta_c$, *the* (*unique*) *constant solution and, for* $\beta > \beta_c$, *the type* 1 *solution.*

(b) *For values of* r *other than* $(1/3, 1/3, 1/3)$ *there exists for all* β *a unique type* 1 *solution which is a minimizer of the free energy.*

In the remainder of this section we prove most of part (a) of this theorem and give an overview of a portion of the proof of part (b). More technical parts of the proof are given in Sects. 6 and 7. Before beginning, we summarize some simple properties of the function Y_K of (5.17).

Remark 5.3 (a) As remarked above, if either $K = 1/27$ or 2β is a multiple of τ_K then $Y_K(t) = 1/3$ for all t.

(b) For all K, β other than those of (a) a plot of $Y_K(t)$ and $Y_K(t \pm \tau_K/3)$ will look much like one of the sets of curves in Fig. 4. In particular, since y_K is even and τ_K-periodic, so is Y_K, and from $Y_K'(t) = (2\beta)^{-1}[y_K(t+\beta) - y_K(t-\beta)]$ it follows that Y_K is strictly monotonic between integer multiples of $\tau_K/2$. From this we see that if $Y_K(t) = Y_K(s)$ then either $s+t$ or $s-t$ must be an integer multiple of τ_K.

(c) To find all triples (r_A, r_B, r_C) for which (5.17) has a solution for a given K it suffices to consider values of t_B satisfying $0 \le t_B \le \tau_K/6$. For such t_B the possible triples satisfy $r_B \le r_C \le r_A$; other orderings of the same sets of values are found for other ranges of t_B.

Proof of Theorem 5.2(a) When $r = (1/3, 1/3, 1/3)$, (5.17) becomes

$$Y_K(t_B) = Y_K(t_B + \tau_K/3) = Y_K(t_B - \tau_K/3) = 1/3. \tag{5.19}$$

For every β, one solution of (5.19) is the constant $y_K(t)(= \rho_\alpha(x)) = 1/3$ corresponding to $K = 1/27$. Moreover, if $2\pi n\sqrt{3} < \beta \le 2\pi(n+1)\sqrt{3}$ then by Remark 5.1 there exist n one-parameter families of nonconstant solutions (indexed by t_B) obtained by solving

$$2\beta = j\tau_{K_j}, \quad j = 1, \ldots, n, \tag{5.20}$$

for K_j. Note that the solution (5.20) is of type j. No other solutions are possible, because (5.19) is inconsistent with the final observation in Remark 5.3(b).

To complete the proof we note that, as is easily seen, perturbing the uniform solution via $\rho_\alpha(x) \to 1/3 + \epsilon\cos(2\pi(x + x_\alpha))$, where $x_A = x_B + 1/3$ and $x_C = x_B - 1/3$, decreases the free energy when $\beta > \beta_c$, so that this solution cannot be a minimizer, and that we will prove in Sect. 7 (see Theorem 7.1) that no solution of type n with $n \ge 2$ can minimize the free energy. □

Proof of Theorem 5.2(b) It was shown in Theorem 4.1 that for any positive r_A, r_B, and r_C which satisfy $\sum_\alpha r_\alpha = 1$ there exists at least one minimizer of the free energy $\mathcal{F}$, and that this minimizer satisfies the ELE. By Theorem 7.1 no solution of type n with $n \ge 2$ can minimize the free energy, so this minimizer must be of type 1. It remains to prove uniqueness of this solution; we will give the full proof in Sect. 6 (see Theorem 6.1), but here we illustrate the idea of the proof by sketching the argument for the special case $r_B < 1/3$, $r_A = r_C > 1/3$, in which symmetry considerations considerably simplify the discussion.

We suppose then that for some such r and some β there are two distinct type 1 solutions of (5.17), and derive a contradiction. It follows from $r_A = r_C$ and Remark 5.3(b) that the phase shift t_B must be zero for each solution, so that we are looking at some $z_1 = y_{K_1}$ and $z_2 = y_{K_2}$ on the interval $J = [-\beta, \beta]$ which satisfy (see (5.16))

$$\langle z_1\rangle_J = \langle z_2\rangle_J = r_B. \tag{5.21}$$

We take $K_2 < K_1$ and write $\theta_i = \tau_{K_i}/2$, so that (because these are type 1 solutions) $\beta < \theta_1 < \theta_2$. The situation is as in Fig. 5; the qualitative features of this figure, which we use below, are obtained from $z_2(0) < z_1(0)$ (which holds because $y_K(0) = a(K)$, the smallest positive zero of the potential $U_K(\rho)$ (5.7), decreases with K), the symmetry, and the fact that (5.21) implies that the curves z_1 and z_2 must cross in the interval J. In the full proof below we will verify that the curves cross exactly twice, as shown.

We write $I_1 = [-\theta_1, \theta_1]$ and $I_2 = [-\theta_2, \theta_2]$ and observe that

$$\langle z_1\rangle_{I_1} = \langle z_2\rangle_{I_2} = 1/3. \tag{5.22}$$

Since $\langle z_2\rangle_{I_2}$ is a (weighted) average of $\langle z_2\rangle_J = r_B < 1/3$ and $\langle z_2\rangle_{I_2\setminus J}$, necessarily

$$\langle z_2\rangle_{I_2\setminus J} > r_B. \tag{5.23}$$

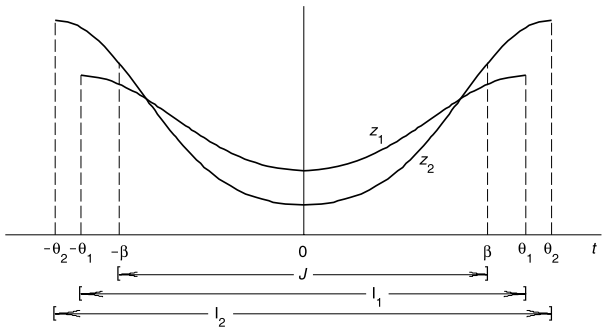

Fig. 5 Configurations of two solutions z_1, z_2 which are candidates to satisfy $\langle z_1\rangle_J = \langle z_2\rangle_J = r_B$

Moreover,

$$\langle z_2\rangle_{I_2\setminus J} \geq \langle z_2\rangle_{I_1\setminus J} > \langle z_1\rangle_{I_1\setminus J}, \tag{5.24}$$

where the first inequality holds because z_2 is decreasing to the left of the origin and increasing to the right, and the second because $z_2 > z_1$ on $I_1 \setminus J$. But then

$$\begin{aligned}\langle z_1\rangle_{I_1} &= \frac{\beta}{\theta_1} r_B + \frac{\theta_1-\beta}{\theta_1}\langle z_1\rangle_{I_1\setminus J}\\ &= r_B + \frac{\theta_1-\beta}{\theta_1}(\langle z_1\rangle_{I_1\setminus J} - r_B)\\ &< r_B + \frac{\theta_1-\beta}{\theta_1}(\langle z_2\rangle_{I_2\setminus J} - r_B)\\ &< r_B + \frac{\theta_2-\beta}{\theta_2}(\langle z_2\rangle_{I_2\setminus J} - r_B)\\ &= \langle z_2\rangle_{I_2},\end{aligned} \tag{5.25}$$

which contradicts (5.22). □

6 Uniqueness of Type 1 Solutions

Our goal in this section is to prove

Theorem 6.1 *For any $\beta > 0$ and any positive r_A, r_B, and r_C which satisfy $\sum_\alpha r_\alpha = 1$ and are not all equal to $1/3$ there exists at most one type 1 solution of (5.8) satisfying (5.14)–(5.17).*

Throughout the section we will consider two solutions y_{K_1}, y_{K_2} of (5.8) with $0 < K_2 < K_1 < 1/27$; we write τ_i rather than τ_{K_i}, $i = 1, 2$, for the corresponding periods, and note that $\tau_1 < \tau_2$ by Remark 5.1(c). We begin with a preliminary result.

Lemma 6.2 (a) *For any $t_0 \in \mathbb{R}$ and $n \in \mathbb{Z}$ the curves $y_{K_1}(t + t_0)$ and $y_{K_2}(t)$ intersect exactly once in the interval $(n\tau_2/2, (n+1)\tau_2/2)$.*

(b) (i) *If $0 \leq t \leq \tau_2/6$ then*

$$y_{K_2}(2\tau_2/3 - t) > y_{K_1}(2\tau_1/3 - t), \tag{6.1}$$

and (ii) *if in addition* $\tau_2 \le 3\tau_1$ *then*

$$y_{K_2}(\tau_2 - t) < y_{K_1}(\tau_1 - t). \tag{6.2}$$

Proof (a) Suppose that n is even so that y_{K_2} is increasing on the interval $(n\tau_2/2, (n+1)\tau_2/2)$; the proof for n odd is similar. For all t and t_0 the zeros $a(K)$ and $b(K)$ of U_K (see (5.7)) satisfy

$$a(K_2) < a(K_1) \le y_{K_1}(t + t_0) \le b(K_1) < b(K_2), \tag{6.3}$$

and since $y_{K_2}(n\tau_2/2) = a(K_2)$ and $y_{K_2}((n+1)\tau_2/2) = b(K_2)$, the existence of at least one intersection follows from the intermediate value theorem. On the other hand, if $y_{K_1}(t+t_0) = y_{K_2}(t) = y$ for some $t \in (n\tau_2/2, (n+1)\tau_2/2)$ then from (5.7) and (5.8), $|y'_{K_1}(t + t_0)| = \sqrt{-2U_{K_1}(y)} < \sqrt{-2U_{K_2}(y)} = y'_{K_2}(t)$, and this is inconsistent with the existence of two such intersection points.

(b) It follows from (a) that for any t_1, t_2, $y_{K_1}(t + t_1)$ and $y_{K_2}(t + t_2)$ can intersect at most once on any interval of monotonicity of $y_{K_2}(t + t_2)$. Thus it suffices to verify that (6.1) and (6.2) hold for $t = 0$ and $t = \tau_2/6$. For (6.1) this follows from $y_{K_2}(\tau_1/2) = b(K_2)$ (see (6.3)) and

$$y_{K_2}\left(\frac{2\tau_2}{3}\right) = \frac{1 - a(K_2)}{2} > \frac{1 - a(K_1)}{2} = y_{K_1}\left(\frac{2\tau_1}{3}\right). \tag{6.4}$$

For (6.2) it follows from $y_{K_2}(\tau_2) = a(K_2)$ and

$$y_{K_2}\left(\frac{5\tau_2}{6}\right) = \frac{1 - b(K_2)}{2} < \frac{1 - b(K_1)}{2} = y_{K_1}\left(\frac{5\tau_1}{6}\right) < y_{K_1}\left(\tau_1 - \frac{\tau_2}{6}\right), \tag{6.5}$$

where at the last step we have used monotonicity of y_{K_1} on $[\tau_1/2, \tau_1]$ and the condition $\tau_2 \le 3\tau_1$. □

In the remainder of this section we suppose that, given appropriate phase shifts, the solutions y_{K_1} and y_{K_2} considered above provide two type 1 solutions satisfying (5.14)–(5.17) for the given r_A, r_B, and r_C, and from this derive a contradiction. It is convenient to view these solutions in the same interval, which we take to be $J = [-\beta, \beta]$. This means that there are phase shifts t_1 and t_2 such that if $z_i(t) = y_{K_i}(t + t_i)$ and $w_i(t) = y_{K_i}(t + t_i + \tau_i/3)$ for $i = 1, 2$, then

$$\langle z_1 \rangle_J = \langle z_2 \rangle_J = r_B \quad \text{and} \quad \langle w_1 \rangle_J = \langle w_2 \rangle_J = r_A. \tag{6.6}$$

By Remark 5.3(c) we may again assume that $r_A \ge r_C \ge r_B$ and that

$$0 \le t_i \le \frac{\tau_i}{6}, \quad i = 1, 2. \tag{6.7}$$

The fact that these are type 1 solutions, and our assumption that $K_1 > K_2$, imply that

$$\beta < \frac{\tau_1}{2} < \frac{\tau_2}{2}. \tag{6.8}$$

The key idea in the proof is the same as that for the simple case considered in Sect. 5.3, but the general case presents two additional difficulties. First, without guidance from symmetry the choice of the intervals I_1 and I_2 is more delicate, and in fact we must consider

two distinct cases in which they are chosen by different prescriptions. Second, one must in some cases apply the reasoning to z_1 and z_2, and in others to w_1 and w_2, and one must show that one or the other is possible. Sorting out these cases requires some detailed analysis of the geometry of the curves.

Proof of Theorem 6.1 Equation (6.6) implies that the curves $z_1(t)$ and $z_2(t)$ must intersect at least once in the interior of the interval J, and since the length 2β of J is less than τ_2, Lemma 6.2(a) implies that they cannot intersect more than three times. On the other hand, from (4.10) it follows that $z_1(\beta)/z_2(\beta) = z_1(-\beta)/z_2(-\beta)$. It is convenient then to consider two possible alternatives:

z1: $z_2(\beta) > z_1(\beta)$, $z_2(-\beta) > z_1(-\beta)$, and the curves z_1 and z_2 intersect precisely twice within J;

z2: $z_2(\beta) \le z_1(\beta)$, $z_2(-\beta) \le z_1(-\beta)$, and either both inequalities are strict and the curves intersect precisely twice within J, or both are equalities and the curves have one intersection within J.

A similar analysis leads to corresponding alternatives for w_1 and w_2:

w1: $w_2(\beta) < w_1(\beta)$, $w_2(-\beta) < w_1(-\beta)$, and the curves w_1 and w_2 intersect precisely twice within J;

w2: $w_2(\beta) \ge w_1(\beta)$, $w_2(-\beta) \ge w_1(-\beta)$, and either both inequalities are strict and the curves intersect precisely twice within J, or both are equalities and the curves have one intersection within J.

We will first show that if **z1** (respectively **w1**) occurs then the argument of Sect. 5.3 may be applied to z_1 and z_2 (respectively w_1 and w_2) to derive a contradiction. Then we show that either **z1** or **w1** or must occur.

Now observe that necessarily $\beta > t_2$. For otherwise, $0 \le t_2 - \beta < t_2 + \beta \le \tau_2/3$, so that $z_2(t) = y_{K_2}(t + t_2)$ is increasing on J and by Lemma 6.2(a) cannot intersect $z_1(t)$ more than once there. It is then helpful to further subdivide the situation into two cases in which we can explicitly identify intervals on which z_2 is monotonic:

Case 1: $\beta + t_2 > \tau_2/2$. In this case we write $J = J_1 \cup J_2 \cup J_3$, where $J_1 = [-\beta, -t_2]$, $J_2 = [-t_2, \tau_2/2 - t_2]$, and $J_3 = [\tau_2/2 - t_2, \beta]$.

Case 2: $\beta + t_2 \le \tau_2/2$. Here we write $J = J_1 \cup J_2$, where $J_1 = [-\beta, -t_2]$ and $J_2 = [-t_2, \beta]$.

Then z_2 is decreasing on J_1 and J_3 and increasing on J_2.

Suppose now that **z1** occurs. Since $z_2(-\beta) > z_1(-\beta)$ and $z_2(-t_2) = y_{K_2}(0) < z_1(-t_2)$ it follows from Lemma 6.2(a) that $z_1(t)$ and $z_2(t)$ intersect once on J_1. We now consider separately the two cases introduced above.

Case 1: Define $I_1 = [\beta - \tau_1, \beta]$ and $I_2 = [\beta - \tau_2, \beta]$, so that $J \subset I_1 \subset I_2$. Since I_i has length τ_i, we again have (5.22) and therefore (5.23). Moreover, (5.24) follows from the monotonicity of z_2 on $I_2 \setminus J$ and then the fact that $z_2 > z_1$ on $I_1 \setminus J$, which follows from this monotonicity, Lemma 6.2(a), and the fact that z_2 intersects z_1 on J_1. But then we may deduce a contradiction between (6.6) and (5.22), just as we did in (5.25).

Case 2: The argument is similar. Let $I_2 = [-t_2 - \tau_2/2, \tau_2/2 - t_2] = I_2^\ell \cup I_2^r$, where $I_2^\ell = [-t_2 - \tau_2/2, -t_2]$ and $I_2^r = [-t_2, \tau_2/2 - t_2]$. Let $I_1 = I_1^\ell \cup I_1^r$, where $I_1^\ell = [-a, -t_2]$ and $I_2^r = [-t_2, b]$, be an interval of length τ_1 satisfying $J \subset I_1 \subset I_2$, with a and b chosen so that

$$\frac{|I_1^\ell \setminus J|}{|I_2^\ell \setminus J|} = \frac{|I_1^r \setminus J|}{|I_2^r \setminus J|}. \tag{6.9}$$

Now by arguing as in Case 1 we find that (5.24) holds with I_2 and I_1 replaced either by I_2^ℓ and I_1^ℓ or by I_2^r and I_1^r. Averaging these two equations (the weights for this averaging,

$$\frac{|I_i^\ell \setminus J|}{|I_i^\ell \setminus J| + |I_i^r \setminus J|} \quad \text{and} \quad \frac{|I_i^r \setminus J|}{|I_i^\ell \setminus J| + |I_i^r \setminus J|} \tag{6.10}$$

are independent of i by (6.9)) we again find that (5.24) itself holds, and the argument proceeds to a contradiction as in Case 1.

If **w1** occurs then we may derive a contradiction similarly: we need only replace z_i by w_i throughout and change the sign of some of the inequalities.

It remains to show that either **z1** or **w1** must occur; to do so we assume that **z2** occurs and show that then **w1** must also. The occurrence of **z2** implies that we must be in Case 1, i.e., that

$$\beta + t_2 > \frac{\tau_2}{2}, \tag{6.11}$$

since if $z_2(-\beta) < z_1(-\beta)$ there can be no intersection of z_1 and z_2 on J_1 and hence the two intersections of these curves must occur on J_2 and J_3, while if $z_2(-\beta) = z_1(-\beta)$ then there are three intersections which must occur on J_1, J_2 and J_3. We claim further that necessarily

$$t_2 - \frac{2\tau_2}{3} > t_1 - \frac{2\tau_1}{3}. \tag{6.12}$$

For if (6.12) does not hold then

$$z_2(\beta) = y_{K_2}(\beta + t_2) \geq y_{K_2}\left(\beta + t_1 + \frac{2(\tau_2 - \tau_1)}{3}\right) > y_{K_1}(\beta + t_1) = z_1(\beta). \tag{6.13}$$

Here for the first inequality we have used (6.7), (6.8), (6.11), and the falsity of (6.12), which together imply that

$$\frac{\tau_2}{2} < \beta + t_2 \leq \beta + t_1 + \frac{2(\tau_2 - \tau_1)}{3} < \frac{2\tau_2}{3}, \tag{6.14}$$

and then the monotonicity of y_{K_2} on $[\tau_2/2, 2\tau_2/3]$. For the second inequality we use Lemma 6.2(b.i), which is applicable by (6.14). But now

$$\begin{aligned} w_2(\beta) &= y_{K_2}\left(\beta + t_2 + \frac{\tau_2}{3}\right) \\ &< y_{K_1}\left(\beta + t_2 + \tau_1 - \frac{2\tau_2}{3}\right) \\ &< y_{K_1}\left(\beta + t_1 + \frac{\tau_1}{3}\right) = w_1(\beta), \end{aligned} \tag{6.15}$$

which implies **w1**. Here the first inequality is from Lemma 6.2(b.ii), applicable because (6.7), (6.8), and (6.11) imply

$$0 \leq 2\tau_2/3 - \beta - t_2 \leq \tau_2/6, \tag{6.16}$$

and (6.11) implies that $\beta > \tau_2/3$ and so $\tau_2 < 3\tau_1/2$, and the second from (6.12) and the monotonicity of y_{K_1} on $[2\tau_1/3, \tau_1]$, which is applicable by (6.16) and the observation that (6.11) implies that $\beta > \tau_2/3 > \tau_1/3$. □

7 Solutions of Type n, $n \geq 2$

It is easy to see that type n solutions of (5.17) with $n \geq 2$ do exist for certain but not all values of β and r. The next result, however, shows that these solutions are not of physical interest.

Theorem 7.1 *No type n solution of* (5.17) *with* $n \geq 2$ *minimizes the free energy.*

We begin by rewriting the free energy in terms of the variable $t = 2\beta x$. Consider then triples $z(t) = (z_A(t), z_B(t), z_C(t))$ of functions, defined in some interval $[c, d]$ of length 2β and satisfying $0 < z_\alpha(t) < 1$ and $\sum_\alpha z_\alpha(t) = 1$. The free energy $\mathcal{F}(\{z(t)\})$ of z is (see (4.3))

$$\mathcal{F}(\{z(t)\}) = (4\beta)^{-1}(\mathcal{E}(\{z(t)\}) - \mathcal{S}(\{z(t)\})), \tag{7.1}$$

where

$$\mathcal{E}(\{z\}) = \int_c^d \mathrm{d}t \int_t^d \mathrm{d}s\, [z_A(t)z_C(s) + z_B(t)z_A(s) + z_C(t)z_B(s)] \tag{7.2}$$

is the energy and

$$\mathcal{S}(\{z\}) = -2\int_c^d \mathrm{d}t\, [z_A \log z_A + z_B \log z_B + z_C \log z_C] \tag{7.3}$$

the entropy. We will use a *rearrangement* procedure for these triples.

Let $c = t_0 < t_1 < \cdots < t_m = d$ be a partition of $[c, d]$ and for $k = 1, \ldots, m$ let $I_k = (t_{k-1}, t_k]$. Given a permutation σ of $\{1, 2, \ldots, m\}$ we may rearrange the intervals I_k in $[c, d]$, together with the restrictions of z to each interval, into the order $I_{\sigma(1)}, \ldots, I_{\sigma(m)}$, thus defining in the obvious way a new function w on $[c, d]$, the *rearrangement* of z. A formal definition of w (which we will not use in the sequel) may be given as follows: The permuted intervals arise from the partition $c = s_0 < s_1 < \cdots < s_m = d$, with $s_j = \sum_{\{i|\sigma(i)\leq j\}}(t_i - t_{i-1})$, and this permutation is implemented by the map $\psi : [c, d] \to [c, d]$ defined by

$$\psi(c) = c, \qquad \psi(x) = x - t_{k-1} + s_{k-1} \quad \text{if } x \in I_k. \tag{7.4}$$

The rearrangement of z is then $w = z \circ \psi^{-1}$, i.e., $w_\alpha(\psi(t)) = z_\alpha(t)$.

The entropy of z and w are the same. The energy of z is

$$\mathcal{E}(\{z(t)\}) = \sum_{i=1}^m \mathcal{E}_i + \sum_{1\leq i<j\leq m} \mathcal{E}_{ij} \tag{7.5}$$

where

$$\mathcal{E}_i = \int_{t\in I_i} \mathrm{d}t \int_{s\in I_i, s>t} \mathrm{d}s\, [z_A(t)z_C(s) + z_B(t)z_A(s) + z_C(t)z_B(s)], \tag{7.6}$$

and for $i < j$, with $\bar{z}_{i\alpha} = \int_{I_i} z_\alpha(t)\,\mathrm{d}t$,

$$\begin{aligned}\mathcal{E}_{ij} &= \int_{t\in I_i} \mathrm{d}t \int_{s\in I_j} \mathrm{d}s\, [z_A(t)z_C(s) + z_B(t)z_A(s) + z_C(t)z_B(s)] \\ &= \bar{z}_{i,A}\bar{z}_{j,C} + \bar{z}_{i,B}\bar{z}_{j,A} + \bar{z}_{i,C}\bar{z}_{j,B}.\end{aligned} \tag{7.7}$$

It is natural to think of (7.7) as expressing $\mathcal{E}(\{z(t)\})$ as a sum of "self energies" of the restrictions of z to the intervals I_i and "interaction energies" between the portions of z in different intervals. The energy $\mathcal{E}(\{w(t)\})$ is obtained similarly; the self energy contribution will be the same and the interaction energies will differ only for interval pairs I_i, I_j whose order is interchanged by the rearrangement; thus

$$\begin{aligned}\mathcal{F}(\{w(t)\}) - \mathcal{F}(\{z(t)\}) &= \mathcal{E}(\{w(t)\}) - \mathcal{E}(\{z(t)\}) \\ &= \sum_{\{i,j|i<j,\,\sigma(j)<\sigma(i)\}} \Delta\mathcal{E}_{i,j}, \end{aligned} \tag{7.8}$$

where

$$\Delta\mathcal{E}_{i,j} = [\bar{z}_{i,C}\bar{z}_{j,A} + \bar{z}_{i,A}\bar{z}_{j,B} + \bar{z}_{i,B}\bar{z}_{j,C} - \bar{z}_{i,A}\bar{z}_{j,C} - \bar{z}_{i,B}\bar{z}_{j,A} - \bar{z}_{i,C}\bar{z}_{j,B}]. \tag{7.9}$$

Proof of Theorem 7.1 Suppose that the triple $z(t)$ as above gives a solution of (5.17) on the interval $[c, d]$; specifically, this means that for some K and some phase shift t_B, $c = t_B - \beta$, $d = t_B + \beta$, and $z_A(t) = y_K(t + \tau_K/3)$, $z_B(t) = y_K(t)$, and $z_C(t) = y_K(t - \tau_K/3)$. Suppose also that the solution is of type n for some $n \geq 2$, which implies in particular that $\tau = \tau_K < 2\beta$. We will show that for some rearrangement w of z the change in free energy (7.8) is negative. Since clearly $\langle w_\alpha\rangle_{[c,d]} = \langle z_\alpha\rangle_{[c,d]}$ for all α, w will satisfy the same constraints as z; thus z cannot be a minimizer of the free energy under those constraints.

The desired rearrangement will be defined in terms of a partition $t_0, \ldots, t_m$ with $t_0 = c$, $t_k = (k_0 + k)\tau/3$ for $k = 1, \ldots, m-1$, and $t_m = d$; here k_0 is the largest integer such that $k_0\tau/3 \leq c$ and m is the smallest integer such that $(k_0 + m)\tau/3 \geq d$. Note that here I_1 and I_m have length at most $\tau/3$ and I_j, $j = 2, \ldots, m-1$, has length exactly $\tau/3$. It is easy to see that $(m-2)\tau/3 < 2\beta \leq m\tau/3$, so that from (5.18), $3n - 2 \leq m \leq 3n + 1$ and in particular $m \geq 4$.

Now for each j, $j = 1, \ldots, m$, there will be an $\alpha \in \{A, B, C\}$ such that $z_\alpha(t) \geq z_{\alpha\pm1}(t)$ for all $t \in I_j$ (see Fig. 4); we will then say that I_j is a *full α-interval* if it has length $\tau/3$ and a *partial α-interval* if it has length less than $\tau/3$ (which is possible only if $j = 1$ or m). The types of the intervals are in cyclic order (again see Fig. 4): if I_j is an α-interval then I_{j+1} is an $(\alpha + 1)$-interval. Recalling that $m \geq 4$ we focus on four consecutive intervals $I_i, \ldots, I_{i+3}$. Without loss of generality we may assume that I_i is an A-interval, and we will then write $I_i = A_l$, $I_{i+1} = B$, $I_{i+2} = C$, and $I_{i+3} = A_r$, i.e., the initial configuration is $A_l B C A_r$. Consider then the permutation of these intervals in which we first exchange B and C, then interchange A_l with C (they are now adjacent) and A_r with B, and finally exchange A_l and A_r, leading to C, A_r, A_l, B. We will show that this rearrangement lowers the energy, completing the proof of the theorem.

The calculation is simplest when all four of the intervals are full, so we consider that special case first. For a full α-interval I_j we have $\bar{z}_{j,\alpha} = \zeta$ and $\bar{z}_{j,\alpha\pm1} = \hat{\zeta}$, where

$$\zeta = \int_0^{\tau/3} z_A(t)\,\mathrm{d}t, \quad \text{and} \quad \hat{\zeta} = \int_0^{\tau/3} z_B(t)\,\mathrm{d}t = \int_0^{\tau/3} z_C(t)\,\mathrm{d}t, \tag{7.10}$$

and so $\zeta > \hat{\zeta}$. We may then compute from (7.9) and (7.10) the energy difference $\Delta\mathcal{E}_{ij}$ arising from reversing the order of a full α_i-interval I_i and a full α_j-interval I_j, where $i < j$: if $\alpha_j = \alpha_i$ then $\Delta\mathcal{E}_{ij} = 0$, while if $\alpha_j = \alpha_i \pm 1$ then $\Delta\mathcal{E}_{ij} = \pm e$, where $e = (\zeta - \hat{\zeta})^2 > 0$. Rearranging the intervals as above thus first increases the energy by e, then decreases it

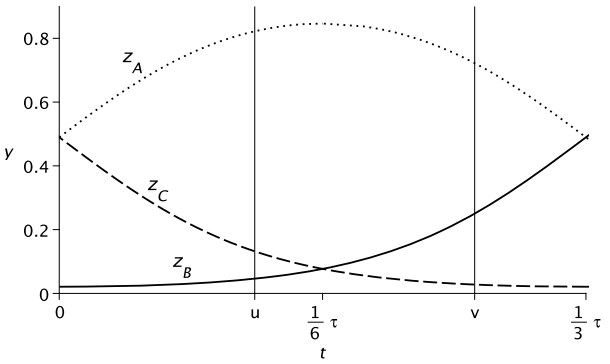

Fig. 6 Plots of z_A, z_B, and z_C in the interval $[0, \tau_K/3]$ for $K = 0.005$, with typical values of u and v

by $2e$; the final exchange of A_l and A_r leaves the energy unchanged. The net change of energy is $-e$.

We now consider the general case in which either A_l or A_r or both may be partial intervals. After a translation, if necessary, of $[c, d]$ by a multiple of τ we may assume that $A_l = (u, \tau/3]$, $B = (\tau/3, 2\tau/3]$, $C = (2\tau/3, \tau]$, and $A_r = (\tau, \tau + v]$ for some u, v with $0 \le u < v \le \tau/3$; the condition $u < v$ comes from the fact that if A_r is full then $u = 0$ and if A_l is full then $v = \tau/3$, while if both are partial then $v + \tau - u = d - c = 2\beta > \tau$. The energy change from the rearrangement is the sum, with $j = i + 3$, of

- $(\zeta - \hat{\zeta})^2$ from the exchange of C and B,
- $-(\zeta - \hat{\zeta})(\bar{z}_{i,A} - \bar{z}_{i,B})$ from the exchange of A_r and C,
- $-(\zeta - \hat{\zeta})(\bar{z}_{j,A} - \bar{z}_{j,C})$ from the exchange of B and A_r, and
- $\bar{z}_{i,C}\bar{z}_{j,A} + \bar{z}_{i,A}\bar{z}_{j,B} + \bar{z}_{i,B}\bar{z}_{j,C} - \bar{z}_{i,A}\bar{z}_{j,C} - \bar{z}_{i,B}\bar{z}_{j,A} - \bar{z}_{i,C}\bar{z}_{j,B}$ from the exchange of A_l and A_r.

This energy may conveniently be expressed in terms of

$$\eta_\alpha^{(1)} = \int_0^u z_\alpha(t)\,\mathrm{d}t, \quad \eta_\alpha^{(2)} = \int_u^v z_\alpha(t)\,\mathrm{d}t, \quad \text{and} \quad \eta_\alpha^{(3)} = \int_v^{\tau/3} z_\alpha(t)\,\mathrm{d}t, \tag{7.11}$$

using

$$z_{i,\alpha} = \eta_\alpha^{(2)} + \eta_\alpha^{(3)}, \qquad z_{j,\alpha} = \eta_\alpha^{(1)} + \eta_\alpha^{(2)}, \qquad \zeta = \eta_A^{(1)} + \eta_A^{(2)} + \eta_A^{(3)}, \tag{7.12}$$

and

$$\hat{\zeta} = \eta_B^{(1)} + \eta_B^{(2)} + \eta_B^{(3)} = \eta_C^{(1)} + \eta_C^{(2)} + \eta_C^{(3)} \tag{7.13}$$

(see Fig. 6). Finally, if we introduce the strictly positive quantities $\theta_B^{(k)} = \eta_A^{(k)} - \eta_B^{(k)}$ and $\theta_C^{(k)} = \eta_A^{(k)} - \eta_C^{(k)}$, then the total energy change may be written, using the identity displayed in (7.13), in the manifestly negative form

$$-\theta_B^{(2)}\theta_C^{(3)} - \theta_C^{(2)}\theta_B^{(1)} - \theta_B^{(2)}\theta_C^{(2)}. \tag{7.14}$$

□

Remark 7.2 Energy considerations for full intervals correspond closely to those of single particles: cyclic order ABC is energetically favored, the energy cost of a reversed pair BA,

CB, or AC is independent of the species involved, etc. The conclusion above that $CAAB$ is favored over $ABCA$ parallels the conclusion of Sect. 2.1 that the ground state of a particle system with a majority of A's has a block containing all the A's between blocks of the C's and the B's. For partial intervals, energy considerations are more subtle; the final step in the rearrangement above, the interchange of A_l and A_r, is irrelevant when these intervals are full but is needed when they are partial because A_r is richer in species C than in B and the reverse is true for A_l.

8 Phase Diagram of the ABC Model

Our results about the minimizers of $\mathcal{F}$ for specified β, r_A, and r_B provide the following picture of the canonical phase diagram of the model (see Fig. 7). The three first order lines in the $T = 0$ plane, discussed in Sect. 2, do not extend to finite temperatures. There is a first order line for $r_A = r_B = r_C = 1/3$ which starts at $T = 0$ and terminates at the critical temperature $T_c = (2\pi\sqrt{3})^{-1}$, i.e., as one crosses this line by varying the overall densities r_α, the density profiles change discontinuously for $T < T_c$. On the line of equal densities the transition at T_c is second order; the divergence of the variance in local density fluctuations as one approaches T_c from above on this line was investigated by Bodineau et al. [9] via a study of the $1/N$ dependence of the correlation function.

To see how the density profiles change as a function of the average densities near the $r_\alpha = 1/3$ line we take $r_A = r_C$ and vary r_B. The density profiles change continuously at high temperatures $\beta < \beta_c$ as the profiles become uniform at $r_\alpha = 1/3$. However, when $\beta > \beta_c$, the B particles prefer to be in the middle section of the interval for $r_B > 1/3$ and to be symmetrically arranged at both ends for $r_B < 1/3$, and this change of profile is discontinuous. To see this discontinuity we define an order parameter,

$$\pi_B = 2\int_{1/4}^{3/4} \rho_B(x)\,\mathrm{d}x, \tag{8.1}$$

and calculate it in the steady state. Figure 8 shows the order parameter π_B as a function of r_B for $\beta = 10$, 12 and 15. Clearly, for low temperatures, $\beta > \beta_c$, π_B shows a discontinuity at density $r_B = 1/3$.

8.1 Perturbation Expansion

In this section we apply a perturbation expansion to calculate the density profiles for average densities which deviate slightly from $1/3$. In particular we carry out a perturbation expansion of the profiles for average densities $r_A = 1/3 + \epsilon + \eta$, $r_B = 1/3 - 2\epsilon$ and $r_C = 1/3 + \epsilon - \eta$ with small parameters ϵ and η. For $\epsilon = \eta = 0$ the density profiles are homogeneous (namely,

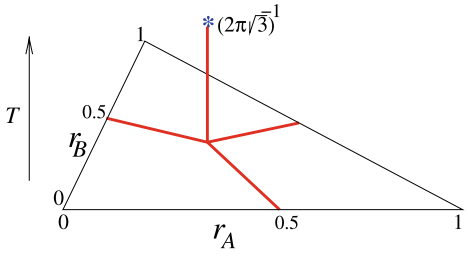

Fig. 7 The (T, r_A, r_B) phase diagram: The temperature is added as a vertical axis to the right triangle. The $T = 0$ phase transition lines do not extend to $T \neq 0$, except at the equal density point $r_\alpha = 1/3$ where it extends as a first-order line at low T and ends at the critical point $\beta = 2\pi\sqrt{3}$

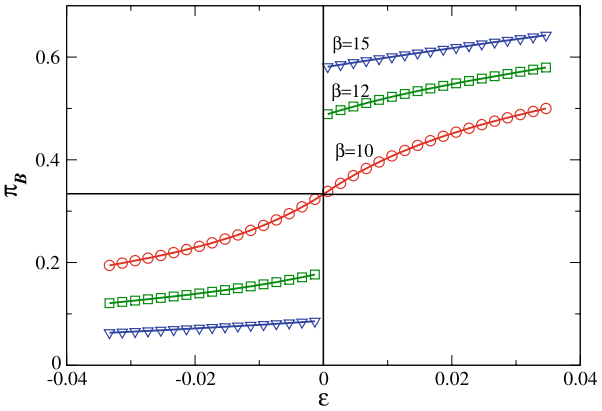

Fig. 8 The order parameter π_B defined in (8.1) is shown as function of $\epsilon \equiv (r_B - 1/3)/2$ for different temperatures $\beta = 10$, 12 and 15. Other parameters are, $r_A = r_C = (1 - r_B)/2$. Data were obtained via a mean-field dynamical solution in a system of 200 sites

$\rho_\alpha(x) = 1/3$) at high temperatures, and they become non-homogeneous at $\beta < \beta_c = 2\pi\sqrt{3}$. We now expand the density profiles around the homogeneous densities:

$$\rho_A(x) = \frac{1}{3} + a(x), \quad \rho_B(x) = \frac{1}{3} + b(x), \quad \text{and} \quad \rho_C(x) = \frac{1}{3} + c(x) \tag{8.2}$$

with small deviations $a(x)$, $b(x)$ and $c(x)$ which satisfy $a(x) + b(x) + c(x) = 0$. As a result of this constraint, one is left with two independent functions which we choose to be $b(x)$ and $s(x) \equiv a(x) - c(x)$. The average densities are then fixed by,

$$\int_{-1/2}^{1/2} b(x)\,\mathrm{d}x = -2\epsilon \quad \text{and} \quad \int_{-1/2}^{1/2} s(x)\,\mathrm{d}x = 2\eta, \tag{8.3}$$

where, for convenience, the interval is taken to be $-1/2 \le x \le 1/2$. In terms of $b(x)$ and $s(x)$ the equations for the minimizers (4.11) reduce to

$$\begin{aligned} \frac{\mathrm{d}b}{\mathrm{d}x} &= \frac{\beta}{3}s + \beta b s, \\ \frac{\mathrm{d}s}{\mathrm{d}x} &= -\beta b + \frac{3\beta}{2}b^2 - \frac{\beta}{2}s^2. \end{aligned} \tag{8.4}$$

We proceed by expanding $b(x)$ and $s(x)$ in terms of two small parameters u and v, which, at the end of the calculation, will be determined by ϵ and η. To be explicit, let us write

$$b(x) = \sum_{k=1}^{\infty} b_k(x), \quad \text{and} \quad s(x) = \sum_{k=1}^{\infty} s_k(x), \tag{8.5}$$

where $b_k(x)$ and $s_k(x)$ are functions of x which are of order k in the small parameters u and v. To first order in u and v (8.4) yields the following equations for $b_1(x)$ and $s_1(x)$:

$$\frac{\mathrm{d}b_1}{\mathrm{d}x} = \frac{\beta}{3}s_1 \quad \text{and} \quad \frac{\mathrm{d}s_1}{\mathrm{d}x} = -\beta b_1.$$

Solving these equations one finds

$$\begin{aligned} b_1(x) &= u\cos(\alpha x) + v\sin(\alpha x), \\ s_1(x) &= -u\sqrt{3}\sin(\alpha x) + v\sqrt{3}\cos(\alpha x) \end{aligned} \tag{8.6}$$

with $\alpha = \beta/\sqrt{3}$. Expanding $b(x)$ and $s(x)$ to second order in u and v the following equations are obtained for $b_2(x)$ and $s_2(x)$,

$$\begin{aligned}\frac{\mathrm{d}b_2}{\mathrm{d}x} &= \frac{1}{\sqrt{3}}\alpha s_2 - \frac{3\alpha}{2}(u^2 - v^2)\sin(2\alpha x) + 3\alpha uv\cos(2\alpha x),\\ \frac{\mathrm{d}s_2}{\mathrm{d}x} &= -\sqrt{3}\alpha b_2 + \frac{3\sqrt{3}\alpha}{2}(u^2 - v^2)\cos(2\alpha x) + 3\sqrt{3}\alpha uv\sin(2\alpha x).\end{aligned} \tag{8.7}$$

These equations can be solved to yield

$$\begin{aligned}b_2(x) &= \frac{1}{2}(u^2 - v^2)\cos(2\alpha x) + uv\sin(2\alpha x),\\ s_2(x) &= \frac{\sqrt{3}}{2}(u^2 - v^2)\sin(2\alpha x) - \sqrt{3}uv\cos(2\alpha x).\end{aligned} \tag{8.8}$$

The equations for the third order terms, $b_3(x)$ and $s_3(x)$, are then

$$\begin{aligned}\frac{\mathrm{d}b_3}{\mathrm{d}x} &= \frac{1}{\sqrt{3}}\alpha s_3 + \frac{3}{2}\alpha uw^2\sin(\alpha x) - \frac{3}{2}\alpha vw^2\cos(\alpha x),\\ \frac{\mathrm{d}s_3}{\mathrm{d}x} &= -\sqrt{3}\alpha b_3 + \frac{3\sqrt{3}}{2}\alpha uw^2\cos(\alpha x) + \frac{3\sqrt{3}}{2}\alpha vw^2\sin(\alpha x),\end{aligned} \tag{8.9}$$

where $w^2 = u^2 + v^2$. To proceed with the analysis one eliminates $s_3(x)$ from the equation for $b_3(x)$ to obtain the following second order differential equation for $b_3(x)$

$$\frac{\mathrm{d}^2 b_3}{\mathrm{d}x^2} + \alpha^2 b_3 = 3\alpha^2 uw^2\cos(\alpha x) + 3\alpha^2 vw^2\sin(\alpha x). \tag{8.10}$$

This equation may be solved to yield

$$b_3(x) = \frac{3}{2}\alpha uw^2 x\sin(\alpha x) - \frac{3}{2}\alpha vw^2 x\cos(\alpha x). \tag{8.11}$$

Inserting this solution in (8.9) results in the following solution for s_3

$$s_3(x) = \frac{3\sqrt{3}}{2}\alpha uw^2 x\cos(\alpha x) + \frac{3\sqrt{3}}{2}\alpha vw^2\sin(\alpha x). \tag{8.12}$$

In summary, to third order in u and v the density profiles are given by

$$\begin{aligned}b(x) &= u\cos(\alpha x) + v\sin(\alpha x) + \frac{1}{2}(u^2 - v^2)\cos(2\alpha x) + uv\sin(2\alpha x)\\ &\quad + \frac{3}{2}\alpha uw^2 x\sin(\alpha x) - \frac{3}{2}\alpha vw^2 x\cos(\alpha x),\\ s(x) &= -u\sqrt{3}\sin(\alpha x) + v\sqrt{3}\cos(\alpha x) + \frac{\sqrt{3}}{2}(u^2 - v^2)\sin(2\alpha x)\\ &\quad - \sqrt{3}uv\cos(2\alpha x) + \frac{3\sqrt{3}}{2}\alpha uw^2 x\cos(\alpha x) + \frac{3\sqrt{3}}{2}\alpha vw^2 x\sin(\alpha x).\end{aligned} \tag{8.13}$$

To third order in u and v these equations may be reexpressed as

$$\begin{aligned} b(x) &= u\cos(\gamma x) + v\sin(\gamma x) + \frac{1}{2}(u^2 - v^2)\cos(2\gamma x) \\ &\quad + uv\sin(2\gamma x) + O(u^4, v^4), \\ s(x) &= -u\sqrt{3}\sin(\gamma x) + v\sqrt{3}\cos(\gamma x) + \frac{\sqrt{3}}{2}(u^2 - v^2)\sin(2\gamma x) \\ &\quad - \sqrt{3}uv\cos(2\gamma x) + O(u^4, v^4), \end{aligned} \tag{8.14}$$

with

$$\gamma = \alpha - \frac{3}{2}\alpha w^2. \tag{8.15}$$

Note that $s(x) = b(x + 2\pi/3\gamma) - b(x - 2\pi/3\gamma)$, which is consistent with (5.13)–(5.14).

The expansion parameters u and v may be expressed in terms of ϵ and η using (8.3). This leads to

$$\begin{aligned} -2\epsilon &= \left[\frac{2}{\gamma}\sin(\gamma/2)\right]u + \left[\frac{1}{2\gamma}\sin\gamma\right](u^2 - v^2), \\ 2\eta &= \left[\frac{2\sqrt{3}}{\gamma}\sin(\gamma/2)\right]v - \left[\frac{\sqrt{3}}{\gamma}\sin\gamma\right]uv. \end{aligned} \tag{8.16}$$

For $\epsilon = \eta = 0$ a nonzero solution of (8.16) exists only if $\sin\gamma/2 = 0$, and thus by (8.15) only if $\alpha > 2\pi$, i.e., $\beta > \beta_c$. For $\beta \le \beta_c$, then, all profiles are homogeneous. For $\beta > \beta_c$ there exist type n solutions with $\gamma = 2n\pi$ and w small whenever β is just above $n\beta_c$. Their amplitude $w = (u^2 + v^2)^{1/2}$ is determined by β via (8.15) but u and v are otherwise undetermined, corresponding to the translation invariance of the set of solutions.

For non-vanishing ϵ or η it is convenient to expand (8.16) to third order:

$$\begin{aligned} -2\epsilon &= \left[\frac{2}{\alpha}\sin(\alpha/2)\right]u + \left[\frac{1}{2\alpha}\sin\alpha\right](u^2 - v^2) \\ &\quad + \frac{3}{2}\left[-\cos(\alpha/2) + \frac{2}{\alpha}\sin(\alpha/2)\right]uw^2, \\ 2\eta &= \left[\frac{2\sqrt{3}}{\alpha}\sin(\alpha/2)\right]v - \left[\frac{\sqrt{3}}{\alpha}\sin\alpha\right]uv \\ &\quad + \frac{3\sqrt{3}}{2}\left[-\cos(\alpha/2) + \frac{2}{\alpha}\sin(\alpha/2)\right]vw^2. \end{aligned} \tag{8.17}$$

Then (8.17) has one or more non-vanishing solutions for u or v at any temperature. The equilibrium (minimizing) solution is the one which is of type 1. The minimizing density profiles, as has been shown, are non-homogeneous both above and below β_c; they vary continuously with the temperature β, and no phase transition takes place.

The small u and v expansion may be used to verify the first order nature of the transition at $\epsilon = \eta = 0$ just below the critical point. To demonstrate this point we take, for simplicity, $\eta = 0$ and consider small ϵ and $\alpha = 2\pi + \Delta\alpha$ with $\Delta\alpha > 0$. In this case $v = 0$, the amplitude u satisfies

$$2\epsilon = \frac{1}{2\pi}(\Delta\alpha - 3\pi u^2)u, \tag{8.18}$$

and the density profile is given by

$$b(x) = u\cos(2\pi + \Delta\alpha - 3\pi u^2)x. \tag{8.19}$$

At $\epsilon = 0$, equation (8.18) has three solutions,

$$u = 0, \qquad \pm\sqrt{\frac{\Delta\alpha}{3\pi}}. \tag{8.20}$$

Since the stable solution satisfies $\Delta\alpha - 3\pi u^2 < 0$, it follows from (8.18) that for the stable solution one has $\epsilon/u < 0$. Thus for $\epsilon > 0$ the negative u solution is stable and $\lim_{\epsilon\searrow 0} u = -\sqrt{\Delta\alpha/3\pi}$, while for $\epsilon < 0$ the positive u solution is stable and $\lim_{\epsilon\nearrow 0} u = \sqrt{\Delta\alpha/3\pi}$. The density profile is therefore discontinuous at $\epsilon = 0$; see Fig. 8.

9 Convexity of the Free Energy at High Temperature

It follows from Theorem 5.2 that if $\beta < \beta_c = (2\pi\sqrt{3})$ then for any positive r_A, r_B and r_C with $\sum_\alpha r_\alpha = 1$ there is a unique solution of the ELE (4.11) with the constraints (4.12), that is, a unique stationary point of the free energy $\mathcal{F}$ under those constraints. One would then like to know also the nature of the fluctuations about this equilibrium profile, and in particular to show that these fluctuations are Gaussian with a finite covariance. To establish this it is necessary to show that the free energy has a positive definite second variation at the minimizer in the space of density profiles satisfying the above constraints, which implies that it is a strictly convex functional in a neighborhood of the minimizer. We believe that this is indeed true generally, with the obvious exception of the line segment $\beta \geq \beta_c$, $r_A = r_B = r_C = 1/3$. In this section we show a global convexity property in a restricted temperature range: for $\beta < 4\pi/3 = (4/3\sqrt{3})\beta_c$, $\mathcal{F}$ (with the mean density of species α constrained to be r_α) is globally strictly convex.

To establish the convexity we compute the second variation. Consider the free energy functional $\mathcal{F}(\{n(x)\})$ on the convex set of nonnegative densities satisfying

$$n_A(x) + n_B(x) + n_C(x) = 1 \quad \text{and} \quad \int_0^1 n_\alpha(x)\,\mathrm{d}x = r_\alpha. \tag{9.1}$$

To make the variations, let ϕ_A and ϕ_B be bounded continuous functions with

$$\int_0^1 \phi_A(x)\,\mathrm{d}x = \int_0^1 \phi_B(x)\,\mathrm{d}x = 0. \tag{9.2}$$

Then for small values of t,

$$(n_A, n_B, n_C) \to (n_A + t\phi_A, n_B + t\phi_B, n_C - t[\phi_A + \phi_B]) \tag{9.3}$$

is an admissible variation, and all admissible variations are of this form.

The order t^2 contribution from the entropy is

$$\frac{1}{2}\int_0^1 \mathrm{d}x\left[\frac{1}{n_A(x)}\phi_A^2(x) + \frac{1}{n_B(x)}\phi_B^2(x) + \frac{1}{n_C(x)}(\phi_A(x) + \phi_B(x))^2\right]$$
$$\geq \int_0^1 \mathrm{d}x\big[\phi_A^2(x) + \phi_B^2(x)\big], \tag{9.4}$$

where we have used the fact that for any positive a, b, and c with $a+b+c=1$ the matrix

$$\begin{bmatrix} a^{-1}+c^{-1}-2 & c^{-1} \\ c^{-1} & b^{-1}+c^{-1}-2 \end{bmatrix} \tag{9.5}$$

is positive semidefinite, since its diagonal entries are nonnegative and its determinant is $(abc)^{-1}[(a-b)^2(1-c)+c(1-c/2)^2+3c^3/4]$. Because the energy is quadratic, the order t^2 contribution from the energy is independent of (n_A, n_B, n_C); it is

$$\begin{aligned}\beta \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y\, \Theta(y-x)\big[&-\phi_A(x)(\phi_A(y)+\phi_B(y)) \\ &+\phi_B(x)\phi_A(y)-(\phi_A(x)+\phi_B(x))\phi_B(y)\big].\end{aligned} \tag{9.6}$$

However, since for functions f and g with $\int_0^1 f(x)\,\mathrm{d}x = \int_0^1 g(x)\,\mathrm{d}x = 0$,

$$\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y\, \Theta(y-x) f(x) g(y) = -\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y\, \Theta(y-x) g(x) f(y), \tag{9.7}$$

(9.6) reduces to

$$3\beta \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y\, \Theta(y-x)\phi_B(x)\phi_A(y). \tag{9.8}$$

Now let $\mathcal{H}$ be the Hilbert space of square integrable functions f on $[0,1]$ with $\int_0^1 f(x)\,\mathrm{d}x = 0$, and define on operator K on $\mathcal{H}$

$$Kf(x) = P\left(\int_x^1 f(y)\,\mathrm{d}y\right), \tag{9.9}$$

where P is the orthogonal projection onto $\mathcal{H}$ in $L^2([0,1])$. Then it is easy to see that

$$\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y\, \Theta(y-x)\phi_B(x)\phi_A(y) = \langle \phi_B, K\phi_A\rangle_{\mathcal{H}}. \tag{9.10}$$

Combining the entropy bound (9.4) with (9.8) and (9.10), we see that the second variation is bounded below by

$$\int_0^1 \mathrm{d}x(\phi_A^2(x)+\phi_B^2(x)) + 3\beta\langle \phi_B, K\phi_A\rangle_{\mathcal{H}}. \tag{9.11}$$

We need to show that this quadratic form is nonnegative definite for $\beta < 4\pi/3$, which we shall do with a spectral calculation.

To do the calculation, define

$$\varphi_n(x) = \sqrt{2}\sin(n2\pi x) \quad \text{and} \quad \psi_n(x) = \sqrt{2}\cos(n2\pi x). \tag{9.12}$$

Then one has

$$K\varphi_n(x) = \frac{1}{n2\pi}\psi_n(x) \quad \text{and} \quad K\psi_n(x) = -\frac{1}{n2\pi}\varphi_n(x). \tag{9.13}$$

If we now write

$$\phi_A = \sum_{n=1}^{\infty}(a_n\varphi_n + b_n\psi_n) \quad \text{and} \quad \phi_B = \sum_{n=1}^{\infty}(c_n\varphi_n + d_n\psi_n), \tag{9.14}$$

then (9.11) becomes

$$\sum_{n=1}^{\infty}\left[(a_n^2 + b_n^2 + c_n^2 + d_n^2) + \frac{3\beta}{n2\pi}(-c_n b_n + d_n a_n)\right]. \tag{9.15}$$

This is nonnegative as long as $\beta \le 4\pi/3$, which proves the strict positivity of the second variation, and hence the convexity of $\mathcal{F}$ for such β

We also remark that if in the equal density case $r_A = r_B = r_C = \frac{1}{3}$ one considers quadratic variations around the constant profile then one can improve the lower bound (9.4) on the second variation of the entropy to

$$3\int_0^1 (\phi_A^2 + \phi_B^2 + \phi_A\phi_B)\,\mathrm{d}x. \tag{9.16}$$

Some modification of the remainder of the argument then shows that the constant profile is indeed a local minimum all the way to β_c.

10 Existence of Minimizers

In this section we prove Theorem 4.1, that is, we show that there exist profiles $\rho(x)$ which minimize the free energy functional $\mathcal{F}(\{n(x)\})$ of (4.3) and that these satisfy the corresponding Euler Lagrange equations. The latter cannot be taken for granted, as there may be situations in which the minimizing density profile $\{\rho(x)\}$ is on the boundary of the permissible domain, e.g., one might have $\rho_A(x) = 0$ for some values of x, and when that happens the minimizing profile need not satisfy the ELE. In the proof of existence we will consider the weak L^1 topology on profiles, in which a sequence $\{f_k\}$ converges to f, where $f_k, f \in L^1([0,1])$, if and only if

$$\lim_{k\to\infty}\int_0^1 \mathrm{d}x\, f_k(x)\phi(x) = \int_0^1 \mathrm{d}x\, f(x)\phi(x) \tag{10.1}$$

for all bounded and measurable functions ϕ. We will use the fact that, by the Dunford-Pettis Theorem [21], the subset $\mathcal{K} := \{f \mid 0 \le f(x) \le 1 \text{ a.e.}\}$ of $L^1([0,1])$ is compact in this topology (since it is clearly uniformly integrable). We need also one other preliminary result.

Lemma 10.1 *The free energy $\mathcal{F}$ is lower semicontinuous in the weak L^1 topology, that is, if $n = (n_A, n_B, n_C)$ and $n_k = (n_{k,A}, n_{k,B}, n_{k,C})$, $k = 1, 2, \ldots$, are profiles with each sequence $\{n_{k,\alpha}\}$ converging to n_α in the sense of (10.1) then* $\liminf_{k\to\infty}\mathcal{F}(\{n_k\}) \ge \mathcal{F}(\{n\})$.

Proof It suffices to show that if $\lim_{n\to\infty} f_n = f$ and $\lim_{n\to\infty} g_n = g$ in the weak L^1 topology then

$$\lim_{n\to\infty}\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}z\,\Theta(z-x) f_n(x) g_n(z) = \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}z\,\Theta(z-x) f(x) g(z) \tag{10.2}$$

and

$$\liminf_{n\to\infty} \int_0^1 \mathrm{d}x\, f_n(x) \ln f_n(x) \geq \int_0^1 \mathrm{d}x\, f(x) \ln f(x). \tag{10.3}$$

To verify (10.2), define $h_n(x) = \int_0^1 \mathrm{d}z \Theta(z-x) f_n(x)$ and $h(x) = \int_0^1 \mathrm{d}z \Theta(z-x) f(x)$, and notice that by the definition of weak convergence,

$$\lim_{n\to\infty} h_n(x) = h(x) \tag{10.4}$$

almost everywhere. Then as each h_n takes its values in $[0, 1]$, the dominated convergence theorem implies that h_n converges to h strongly in the L^1 norm. Fix $\epsilon > 0$, and pick N so that $\|h_n - h\|_1 < \epsilon$ for all $n \geq N$. Then

$$\begin{aligned}&\left|\int_0^1 \mathrm{d}x\, g_n(x)h_n(x) - \int_0^1 \mathrm{d}x\, g(x)h(x)\right| \\ &\quad\leq \left|\int_0^1 \mathrm{d}x\, (g_n(x) - g(x))h(x)\right| + \left|\int_0^1 \mathrm{d}x\, g_n(x)(h_n(x) - h(x))\right|.\end{aligned} \tag{10.5}$$

But $|\int_0^1 \mathrm{d}x\, g_n(x)(h_n(x) - h(x))| \leq \|g_n\|_\infty \|h_n - h\|_1 \leq \epsilon$ since $0 \leq g_n \leq 1$. Likewise, since $0 \leq h \leq 1$, $\lim_{n\to\infty} \int_0^1 \mathrm{d}x\, (g_n(x) - g(x))h(x) = 0$ by the definition of weak convergence. Thus for all n sufficiently large, $|\int_0^1 \mathrm{d}x\, g_n(x)h_n(x) - \int_0^1 \mathrm{d}x\, g(x)h(x)| < 2\epsilon$, and (10.2) follows.

Equation (10.3) is an easy consequence of the convexity of the function $-S(f) = \int_0^1 \mathrm{d}x\, f(x) \ln f(x)$ on $\mathcal{K}$. Indeed, for any real number a, the set $\mathcal{K}_a = \{f \in \mathcal{K} : -S(f) \leq a\}$ is convex. It is also closed in the L^1 norm topology, for if $\{f_n\}$ is a norm convergent sequence in $\mathcal{K}_a$ with limit f then there is a subsequence converging almost everywhere, and passing to the subsequence and using the dominated convergence theorem, with the obvious bound $|f_n| \leq 1/e$, yields $-S(f) = \lim_{n\to\infty} -S(f_n) \leq a$, so that $f \in \mathcal{K}_a$ as well. But by a theorem of Mazur [22], a convex set in a Banach space is closed if and only if it is weakly closed, and so $\mathcal{K}_a$ is weakly closed. It follows easily from this that if $\{f_n\}$ converges weakly to f, then

$$-S(f) \leq \liminf_{n\to\infty} -S(f_n), \tag{10.6}$$

which is (10.3). □

Proof of Theorem 4.1 Let $\{n_{k,A}, n_{k,B}, n_{k,C}\}$ be a minimizing sequence for $\mathcal{F}$, with $n_{k,\alpha} \in \mathcal{K}$, $\sum_\alpha n_{k,\alpha} = 1$, and $\int_0^1 n_{k,\alpha}(x)\, \mathrm{d}x = r_\alpha$, so that

$$\lim_{k\to\infty} \mathcal{F}(\{n_k(x)\}) = F(r_A, r_B, r_C). \tag{10.7}$$

By the weak compactness of $\mathcal{K}$ we can choose a subsequence along which each $n_{k,\alpha}$ converges to some $\rho_\alpha \in \mathcal{K}$. From (10.1) it follows immediately that $\int_0^1 \mathrm{d}x\, \rho_\alpha(x) = r_\alpha$, from $\sum_\alpha n_{k,\alpha} = 1$ that $\sum_\alpha \rho_\alpha = 1$, and from Lemma 10.1 that

$$\mathcal{F}(\{\rho(x)\}) \leq \liminf_{n\to\infty} \mathcal{F}(\{n_k(x)\}) = F(r_A, r_B, r_C). \tag{10.8}$$

Thus, $\mathcal{F}(\{\rho(x)\}) = F(r_A, r_B, r_C)$.

We next show that each ρ_α is uniformly bounded below by some constant $\delta > 0$. Suppose for example that this is not the case for ρ_A, so that for each $\delta > 0$, the set $D_\delta = \{x \mid \rho_A(x) < \delta\}$ has strictly positive measure $|D_\delta|$. We can then "fill in the hole" in ρ_A: we transfer some mass from the larger of ρ_B and ρ_C to ρ_A, in a way that preserves the constraint $\rho_A(x) + \rho_B(x) + \rho_C(x) = 1$, and then transfer mass in the opposite direction on a certain "safe" set M to restore the constraints $\int_0^1 \rho_\alpha \, dx = 1$. Because "entropy abhors a vacuum", these transfers will, for δ sufficiently small, strictly lower the value of $\mathcal{F}$, thus contradicting the optimality of $\{\rho_\alpha\}$.

On the set M, ρ_A should be strictly bounded below; we take $M = \{x \mid \rho_A(x) > r_A/2\}$ and note that $|M| > 0$ because ρ_A has mean value r_A. Next, let E_δ be any subset of D_δ with $0 < |E_\delta| \le |M|$, let F_δ be the subset of E_δ on which $\rho_B \ge \rho_C$, and let $G_\delta = E_\delta \backslash F_\delta$. Finally, define $\{\widetilde{\rho}_\alpha\}$ by

$$\widetilde{\rho}_A(x) = \rho_A(x) + \delta \, 1_{E_\delta}(x) - \delta \frac{|E_\delta|}{|M|} \, 1_M(x), \tag{10.9}$$

$$\widetilde{\rho}_B(x) = \rho_B(x) - \delta \, 1_{F_\delta}(x) + \delta \frac{|F_\delta|}{|M|} \, 1_M(x), \tag{10.10}$$

$$\widetilde{\rho}_C(x) = \rho_C(x) - \delta \, 1_{G_\delta}(x) + \delta \frac{|G_\delta|}{|M|} \, 1_M(x). \tag{10.11}$$

Using the fact that, on D_δ, $\max\{\rho_A(x) \ \ \rho_B(x)\} \ge (1 - \delta)/2$, and on M, $\rho_B(x), \rho_C(x) \le 1 - \rho_A/2 < 1$, one sees easily that if $0 < \delta < \min\{r_A/2, 1/3\}$ then $\widetilde{\rho}_\alpha$ satisfies $0 \le \widetilde{\rho}_A(x) \le 1$.

It is now easy to verify that for all δ sufficiently small,

$$-S(\widetilde{\rho}_A) = -S(\rho_A) + \big(\delta \ln(\delta) + O(\delta)\big)|E_\delta|. \tag{10.12}$$

Moreover, changes in $S(\rho_B)$, $S(\rho_C)$, and the energy components of $\mathcal{F}$ are of order $O(\delta)|E_\delta|$, and thus, for δ sufficiently small, $\mathcal{F}(\{\widetilde{\rho}_\alpha\}) < \mathcal{F}(\{\rho_\alpha\})$, which contradicts the optimality of $\{\rho_\alpha\}$. Hence for some $\delta > 0$, it must be the case that D_δ is a null set. The same argument applies of course to ρ_B and ρ_C, and hence each of the $\{\rho_\alpha\}$ is uniformly bounded below by a strictly positive constant. Then from the fact that $\rho_A(x) + \rho_B(x) + \rho_C(x) = 1$ almost everywhere, it follows that $\rho_\alpha(x) < 1 - 2\delta$ for almost every x and each α.

Finally, since we have an interior minimum, $\{\rho_\alpha\}$ must satisfy the ELE for $\mathcal{F}$, that is, $\mathcal{F}_A - \mathcal{F}_C$ and $\mathcal{F}_B - \mathcal{F}_c$ must be constant (see (4.6)), which from (4.7) implies that ρ_A/ρ_C and ρ_B/ρ_C are differentiable. From this the constraint $\sum_\alpha \rho_\alpha(x) = 1$ implies readily that all the ρ_α are differentiable, and the argument of Sect. 3 them shows that they satisfy the ELE in the form (4.9). This implies infinite differentiability. □

11 Discussion and Conclusion

The dynamics defining the weakly asymmetric ABC model on the interval are entirely local and identical to those for the model on a ring, except for the insertion of a barrier between sites N and 1. Remarkably, when $N_A = N_B = N_C = N/3$ this has no effect on the stationary state of the system, which (as was known before for the ring) is given by a Gibbs measure with mean field type asymmetric long range pair interactions (see (2.1), (2.2)). When the species numbers are unequal, however, the stationary state on the ring is a nonequilibrium one, with a net current, while the state on the interval is again given by a canonical Gibbs measure with the same interactions.

In the scaling limit $N \to \infty$, $N_\alpha/N \to r_\alpha$, the system on the ring with $r_A = r_B = r_C = 1/3$ was known to have a second order transition at $\beta = \beta_c = 2\pi\sqrt{3}$ from a macroscopic state with uniform density profiles $\rho_\alpha(x) = 1/3$ to one in which the density profiles are periodic, with period one. This transition, which by the isomorphism between the ring and interval systems at equal densities carries over to the interval, corresponds to a broken symmetry in which the phase of the typical density profile in the scaling limit is uniformly distributed over [0, 1]. This corresponds in (3.4) to $\Omega = [0, 1]$, with the measure κ uniform on Ω.

In the case of the ring there is strong evidence [9], but no proof so far, that there will be a transition from uniform to nonuniform typical configurations also for a range of unequal, but strictly positive, densities. It might then be natural to conjecture that something similar happens on the interval. Our results here, however, prove that this is not the case: for unequal densities the change from $\beta = 0$ to $\beta = \infty$ is smooth, with unique density profiles $\rho_\alpha(x)$ for each $\beta < \infty$. In the limit $\beta \to \infty$ the system will segregate into three or four blocks of different species; see the discussion in Sect. 2.1 of the ground states of the finite system. When $0 < r_B < r_A = r_C$ the limiting configuration has four blocks of particles, **BCAB**, with the B particles evenly divided between the left and right ends of the interval, that is, the degeneracy of the ground state of the finite system, in which the B particles could be arbitrarily divided between the two ends, is broken in the scaling limit when $\beta \to \infty$.

A striking characteristic of the resulting (T, r_A, r_B) phase diagram (Fig. 7) is that one observes a phase transition at precisely one value of the external parameters (here r_A and r_B) as the temperature is lowered. Such a feature is typically encountered for grand canonical ensembles of systems in the presence of an ordering field (e.g. the (T, H_x, H_y) phase diagram of the XY model in a magnetic field $\vec{H} = (H_x, H_y)$). Note, however, that here we are considering a mean field model and permitting only those changes in the system in which the densities r_A, r_B, and r_C are kept strictly fixed. The question of what happens for more general variations, that is, if we consider a grand canonical ensemble, has to be investigated separately (see note 5 below). This is unlike the case of systems with short range interactions, where one has an equivalence of ensembles. What is interesting in this model compared with the usual mean field situation is the rich spatial structure of the equilibrium states, which is due entirely to the directional asymmetry of the interactions. Thus if we modified the energy function in (2.2) to involve a sum over all $j \neq i$, not just $j > i$, we would obtain $E_N^{\mathrm{mf}} = N_A N_B + N_A N_C + N_B N_C$ which, in the canonical ensemble, would give uniform minimizing density profiles at all temperatures, since the interaction would not depend on the spatial structure and the entropy term prefers the uniform state.

As we have shown, the minimizing profiles are given by pieces of the graphs of elliptic functions y_K describing the τ-periodic trajectory of a particle with zero energy in a quartic confining potential. It follows from the general analysis of the Euler-Lagrange equations that these elliptic functions have the property that $y_K(t) + y_K(t + \tau/3) + y_K(t - \tau/3) = 1$ while $y_K(t)y_K(t + \tau/3)y_K(t - \tau/3) = K$ is independent of t. It is not clear to us whether or not this property of these elliptic functions $y_K(t)$ is known in the literature.

We note that, as pointed out in Sect. 2, the local measures μ_x must be either product measures or superpositions of product measures. Clearly when (3.2) holds, as it will wherever the minimizers are unique, $\mu_x = \nu_{\rho(x)}$ will be a product measure with density $\rho(x)$, while for $r_A = r_B = r_C = 1/3$ and $\beta > \beta_c$ the μ_x will be a superposition of the measures $\mu_\rho(z)$ as z varies uniformly over [0, 1].

We end this section with several open problems.

1. It follows from our analysis that for large β there will be solutions of the ELE of type $n > 1$ which are stationary points, but not global minimizers, of $\mathcal{F}$. We have not determined whether or not these correspond to local minima.
2. It follows from the proof (see Sect. 9) of the strict convexity of $\mathcal{F}\{n(x)\}$ for small β that the fluctuations about the minimizing $\rho(x)$ at fixed r_α are (constrained) Gaussian. We know that this is not true for $\beta > \beta_c$ and $r_A = r_B = r_C = 1/3$, due to the existence of a one parameter family of minimizing profiles. What of the fluctuations about minimizing densities for other values of r_α when β is not very small?
3. As mentioned in the introduction, Evans et al. considered the ABC model on the ring with general rates $q_{\alpha,\gamma}$ $(\alpha \neq \gamma)$ and found that the stationary state is Gibbsian whenever $N_\alpha^{-1} \log(q_{\alpha+1,\alpha+2}/q_{\alpha+2,\alpha+1})$ is independent of α. This is possible, of course, only when the ratio of any two of these logarithms is rational, and then only when N is a multiple of some smallest possible system size, just as for the model considered in the body of this paper the stationary state can be Gibbsian only if N is divisible by 3. To study the weakly asymmetric scaling limit one would take $q_{\alpha,\alpha+1} = e^{-\beta c_\alpha/N} q_{\alpha+1,\alpha}$ with some normalization, say $\sum_\alpha c_\alpha = 1$. The energy

$$E_N = \frac{1}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{\alpha} c_\alpha \eta_{\alpha+1}(i) \eta_\alpha(j) \tag{11.1}$$

   would then be translation invariant when $N_\alpha = c_\alpha N$. As in the current paper, however, one may use (11.1) to obtain a Gibbs measure $Z^{-1} \exp[-\beta E_N]$ on the interval for any values of the N_α, and then study the scaling limit (3.1). It would be interesting to carry out an analysis of the free energy minimizing profiles also for this model.
4. The ABC model can be extended to higher dimensions either on a torus or in a box [23]. There does not, however, seem to be any natural way to extend the Gibbs measure to these systems for any values of the r_α's.
5. Our analysis of $\mathcal{F}(\{n\})$ has been carried out entirely for the canonical ensemble, the natural one for the exchange dynamics. From the point of view of the Gibbs measure (2.1) it is natural to consider also the grand canonical ensemble in which the N_α are not fixed. In that case one would have to consider variations in the grand canonical partition function, the pressure, for fixed chemical potentials. It turns out that in this case there is a phase transition at values of β less than β_c. We have also found a natural generalization of the ABC dynamics which leads to a grand canonical Gibbs measure: one considers the system on a ring of $N+1$ sites, with one special particle (of type X) replacing the boundaries in the interval model, and permits exchanges $AX \leftrightarrow XB$, and cyclic permutations of these, at rates which satisfy detailed balance with the given chemical potentials. This will be the subject of future work.

Acknowledgements We thank Lorenzo Bertini, Thierry Bodineau, Bernard Derrida, Erel Levine and Errico Presutti for helpful discussions. The work of J.L.L. and A.A. was supported by NSF Grant DMR-0442066 and AFOSR Grant AF-FA9550-04. Support of the Israel Science Foundation (ISF), the Minerva Foundation with funding from the Federal Ministry for Education and Research and of the Albert Einstein Center for Theoretical Physics is gratefully acknowledged.

Appendix A: The Profiles as Elliptic Functions

The function $t(y)$ defined in (5.10) is an elliptic function, since $U_K(y)$ is a quartic polynomial in y. To identify this function explicitly we introduce a fractional linear or Möbius

transformation which sends the four roots 0, a, b, and c of $U_K(y)$ to the four roots of the polynomial which appears in the incomplete elliptic integral of the first kind [24],

$$F(x,k)=\int_0^x \frac{\mathrm{d}t}{\sqrt{(1-x^2)(1-k^2x^2)}}. \tag{A.1}$$

We choose the parameter k to satisfy $k\geq 1$ and map $(0,a,b,c)\to(-1,-1/k,1/k,1)$. The formula for the transformation is

$$y\to z=f(y)=\frac{\alpha_+ y-1}{\alpha_- y+1}, \tag{A.2}$$

where

$$\alpha_\pm=\frac{\pm ab+\sqrt{ab(c-b)(c-a)}}{abc}, \tag{A.3}$$

which implies

$$k=\frac{1+\alpha_- a}{1-\alpha_+ a}. \tag{A.4}$$

Writing the inverse Möbius transformation of (A.2) as

$$z\to y=g(z)=\frac{1+z}{\alpha_+-\alpha_- z}, \tag{A.5}$$

we can write (5.10) as

$$\begin{aligned} t&=\int_{-1/k}^{f(y)}\frac{(\alpha_++\alpha_-)\,\mathrm{d}z}{(\alpha_+-\alpha_- z)^2\sqrt{-2U_K(g(z))}}\\ &=\varkappa\int_{-1/k}^{f(y)}\frac{\mathrm{d}z}{\sqrt{(1-z^2)(1-k^2z^2)}}\\ &=\varkappa\left[F\left(\frac{1}{k},k\right)+F(f(y),k)\right], \end{aligned} \tag{A.6}$$

where

$$\varkappa=\frac{2(\alpha_++\alpha_-)}{\sqrt{(1-\alpha_+c)(1-\alpha_+b)(1-\alpha_+a)}}. \tag{A.7}$$

The period of oscillation (5.9) is then

$$\tau=\frac{4}{\varkappa}F\left(\frac{1}{k},k\right). \tag{A.8}$$

The inverse of the elliptic integral $F(x,k)$ is the Jacobi elliptic function $\mathrm{sn}(x,k)$, that is, $F(\mathrm{sn}(x,k),k)=x$. Then $f(y(t))=\mathrm{sn}(\varkappa t-F(\frac{1}{k},k),k)$, so that

$$y(t)=\frac{1+\mathrm{sn}(\varkappa t-F(\frac{1}{k},k),k)}{\alpha_+-\alpha_-\,\mathrm{sn}(\varkappa t-F(\frac{1}{k},k),k)} \tag{A.9}$$

is the required solution to (5.8). The density profiles are then

$$\begin{aligned} n_A(x) &= \frac{1+\mathrm{sn}(2\varkappa\beta x - \frac{7}{3}F(\frac{1}{k},k),k)}{\alpha_+ - \alpha_- \,\mathrm{sn}(2\varkappa\beta x - \frac{7}{3}F(\frac{1}{k},k),k)}, \\ n_B(x) &= \frac{1+\mathrm{sn}(2\varkappa\beta x - F(\frac{1}{k},k),k)}{\alpha_+ - \alpha_- \,\mathrm{sn}(2\varkappa\beta x - F(\frac{1}{k},k),k)}, \\ n_C(x) &= \frac{1+\mathrm{sn}(2\varkappa\beta x + \frac{1}{3}F(\frac{1}{k},k),k)}{\alpha_+ - \alpha_- \,\mathrm{sn}(2\varkappa\beta x + \frac{1}{3}F(\frac{1}{k},k),k)}. \end{aligned} \tag{A.10}$$

As a check we have derived expressions for values of the elliptic functions at the special points $\frac{1}{3}F(\frac{1}{k},k)$ and $-\frac{7}{3}F(\frac{1}{k},k)$ in a way analogous to computing $\sin 5\pi/12$ (i.e., by showing that the desired value is the root of a polynomial obtained using trigonometric addition formulae), and then used the addition formulae of the elliptic functions [24] to verify directly the constancy of the sum (5.4) and the product (5.3) of the densities. These computations were very intensive and were assisted with the help of the computer algebra package, Maple™.

Appendix B: Proofs for Remark 5.1

Proof of Remark 5.1 (b) It is convenient to change variables in the integrals occurring in (5.9) and (5.11)–(5.12) by writing $y = a + (b-a)s$, so that for example (5.9) becomes

$$\tau_K = 4\int_0^1 \frac{\mathrm{d}s}{\sqrt{a+s(b-a)}\sqrt{s(1-s)}\sqrt{c-(a+s(b-a))}}. \tag{B.1}$$

Then as $K \nearrow 1/27$ the points a, b, and c approach $1/3$, $1/3$, and $4/3$, respectively, so τ_K approaches $4\sqrt{3}\int_0^1 \mathrm{d}s/\sqrt{s(1-s)} = 4\pi\sqrt{3}$. Under this change of variable the integral in (5.11) becomes (assuming that K is so small that $a < \epsilon$)

$$I_1 \equiv 2\int_0^{\epsilon(K)} \frac{\mathrm{d}s}{\sqrt{s(s+a)}} \sqrt{\frac{s+a}{(a+s(b-a))(1-s)(c-(a+s(b-a)))}}, \tag{B.2}$$

where $\epsilon(K) = (\epsilon - a)/(b-a)$. As $K \searrow 0$, $a = 4K + O(K^2)$ and $b_\pm = 1 \pm 2\sqrt{K} + O(K)$, so that $e(K) = \epsilon + O(K)$ and the second square root in (B.2) is $\eta(s) = 1 + O(K)$ uniformly on the integration range, from which

$$\lim_{K\searrow 0} \frac{I_1}{\ln(1/K)} = \lim_{K\searrow 0} \frac{2}{\ln(1/K)} \int_0^{\epsilon} \frac{\mathrm{d}s}{\sqrt{s(s+a)}} = 2. \tag{B.3}$$

The argument for the limit in (5.12) is similar; the difference in the limiting values arises because $1-b$ is of order $\sqrt{K}$ for K small. Finally, a similar analysis shows that

$$\int_\epsilon^{1-\epsilon} \frac{\mathrm{d}y}{\sqrt{-2U_K(y)}} \tag{B.4}$$

is bounded as $K \searrow 0$ which, with (B.1), shows that $\lim_{K\searrow 0}(\tau_K/\ln(1/K)) = 6$.

(c) The proof is long, computational, and unilluminating. We show that $\mathrm{d}\tau_K/\mathrm{d}K > 0$; as mentioned in Sect. 5.2, it is most convenient to consider $a \in (0, 1/3)$ as the fundamental parameter and show that $\mathrm{d}\tau_{K(a)}/\mathrm{d}a > 0$ (recall that $K(a) = a(1-a)^2/4$ so $\mathrm{d}K/\mathrm{d}a > 0$). Our starting point is (B.1); writing

$$h(a,s) \equiv (a + s(b-a))(c - (a + s(b-a))) \tag{B.5}$$

and $R = \sqrt{a(4-3a)}$ we find that

$$\frac{\mathrm{d}\tau_K}{\mathrm{d}a} = 4\int_0^1 \frac{\mathrm{d}\,h(a,s)}{\mathrm{d}s}\frac{\mathrm{d}s}{h(a,s)^{3/2}\sqrt{s(1-s)}} = \frac{4}{R}\int_0^1 \frac{g(a,s)\,\mathrm{d}s}{h(a,s)^{3/2}\sqrt{s(1-s)}}, \tag{B.6}$$

where

$$g(a,s) = 3a(1-a)(1+s) + (9a^2 - 12a + 2)s^2 + (1 - 3a - (5-9a)s + (2-3a)s^2)R. \tag{B.7}$$

Unfortunately, $g(a,s)$ is not nonnegative for all a, s, but the sum of g and its reflection around the point $s = 1/2$ is. This follows from the next lemma, whose proof is given immediately below.

Lemma B.1 *For* $0 \le a \le 1/3$ *and* $0 \le s \le 1/2$, (i) $h(a,s) \le h(a,1-s)$; (ii) $g(a,s) \ge 0$, *and* (iii) $g(a,s) + g(a,1-s) \ge 0$.

Using the lemma, we have from (B.6)

$$\begin{aligned}
\frac{\mathrm{d}\tau_K}{\mathrm{d}a} &= \frac{4}{R}\int_0^{1/2}\left[\frac{g(a,s)}{h(a,s)^{3/2}} + \frac{g(a,1-s)}{h(a,1-s)^{3/2}}\right]\frac{\mathrm{d}s}{\sqrt{s(1-s)}} \\
&\ge \frac{4}{R}\int_{\substack{0\le s\le 1/2 \\ g(a,1-s)\le 0}}\left[\frac{g(a,s)}{h(a,s)^{3/2}} + \frac{g(a,1-s)}{h(a,1-s)^{3/2}}\right]\frac{\mathrm{d}s}{\sqrt{s(1-s)}} \\
&\ge \frac{4}{R}\int_{\substack{0\le s\le 1/2 \\ g(a,1-s)\le 0}}\left[g(a,s) + g(a,1-s)\right]\frac{\mathrm{d}s}{h(a,s)^{3/2}\sqrt{s(1-s)}} \\
&\ge 0,
\end{aligned} \tag{B.8}$$

where the three inequalities in (B.8) are justified by parts (ii), (i), and (iii) of the claim, respectively. Note also that the first inequality is strict unless $g(a,s) \le 0$ for all $s \in [1/2, 1]$, in which case the last inequality will be strict (here we use the fact that neither $g(a,s)$ nor $g(a,s) + g(a,1-s)$ vanish identically in s for any a); thus (B.8) actually implies $d\tau_K/da > 0$. □

Proof of Lemma B.1 For (i), note that

$$h(a,1-s) - h(a,s) = (b-a)(1-2s)(R-a); \tag{B.9}$$

each of the factors on the right hand side is easily seen to be nonnegative for the range of variables in question. Before proceeding to (ii) and (iii) we make two preliminary observations. First, the coefficient

$$\lambda(a) \equiv (9a^2 - 12a + 2) + (2-3a)R \tag{B.10}$$

of s^2 in $g(a,s)$ (which is a quadratic polynomial in s) is positive for all $a \in [0,1/3)$, for from $\lambda(0)=2$, $\lambda(1/3)=0$, and $\lambda'(1/3)=-8$ we see that a zero of λ in the open interval $(0,1/3)$ would imply two roots of λ'' there, contradicting $\lambda'''(a)=48/R^5$ strictly positive. Second, $\kappa(a)=g(a,1/2)$ is positive on $[0,1/3)$, for $\kappa'''(a)=-6/(a^2R(4-3a))$ is negative on $(0,1/3)$ and since, as one easily checks, $\kappa''(1/2)=\kappa'(1/2)=\kappa(1/2)=0$, we conclude that $\kappa''(a)$, $-\kappa'(a)$, and $\kappa(a)$ are all positive. These observations immediately imply (iii) of the claim, since they imply that $g(a,s)+g(a,1-s)$ is for any a a parabola in s with positive minimum at $s=1/2$.

We turn finally to (ii) of the claim. For fixed a the minimum of $g(a,s)$ occurs at

$$s_0(a)=\frac{3a(1-a)+(5-9a)R}{2(9a^2-12a+2+(2-3a)R)}. \tag{B.11}$$

For those a for which $s_0(a)\geq 1/2$, the observations in the preceding paragraph imply $g(a,s)\geq 0$, $s\in[0,1/2]$. But s_0 is an increasing function of a, since

$$s_0'(a)=\frac{5(2-3a)+18a^2(2-a)+3R}{R(9a^2-12a+2+(2-3a)R)^2}; \tag{B.12}$$

moreover, if $a_0=2/25$ then $s_0(a_0)=(107\sqrt{47}-69)/(88\sqrt{47}+686)>1/2$. Thus it suffices to show that

$$g(a,s_0(a))=\frac{a(-22+45a-36a^2+9a^3)+(4-9a+18a^2-9a^3)R}{2(9a^2-12a+2+(2-3a)R)} \tag{B.13}$$

is nonnegative for $a\in(0,a_0)$. The denominator in (B.13) is easily seen to be positive in this range. We denote the numerator by $\phi(a)$; then $\phi(0)=0$ and $\phi(a_0)>0$, and positivity will follow once we prove that there is an a_1 with $0<a_1<a_0$ such that $\phi'(a)>0$ for $a\in(0,a_1)$ and $\phi'(a)<0$ for $a\in(a_1,a_0)$. But if $\psi(a)=R\phi'(a)$ then

$$\psi(a)=108a^4-288a^3+234a^2-66a+8+(36a^3-108a^2+90a-22)R; \tag{B.14}$$

then, as is easily checked, $\psi(0)>0$, $\psi(a)<0$, and $\psi'(a)<0$ on $(0,a_0)$, since $\psi'(a)=(\psi_1(a)+\psi_2(a)R)/R$, where

$$\psi_1(a)=-22+303a-810a^2+738a^3-216a^4 \tag{B.15}$$

and

$$\psi_2(a)=-33-432a^2+234a+216a^3, \tag{B.16}$$

when written as polynomials in $a-a_0$, have negative constant term and alternating signs, and hence are negative on $(0,a_0)$. □

References

1. Lieb, E.H., Mattis, D.C. (eds.): Mathematical Physics in One Dimension. Academic Press, New York (1966)
2. Percus, J.K.: Exactly solvable models of classical many-body systems. In: Lebowitz, J.L. (ed.) Simple Models of Equilibrium and Nonequilibrium Phenomena. Amsterdam, North-Holland (1987)
3. Privman, V. (ed.): Nonequilibrium Statistical Mechanics in One Dimension. Cambridge University Press, Cambridge (1997)

4. Schütz, G.M.: Exactly solvable models for many-body systems far from equilibrium. In: Domb, C., Lebowitz, J.L. (eds.) Phase Transitions and Critical Phenomena, vol. 19. Academic Press, London (2000)
5. Evans, M.R., Kafri, Y., Koduvely, H.M., Mukamel, D.: Phase separation in one-dimensional driven diffusive systems. Phys. Rev. Lett. **80**, 425–429 (1998)
6. Evans, M.R., Kafri, Y., Koduvely, H.M., Mukamel, D.: Phase separation and coarsening in one-dimensional driven diffusive systems: Local dynamics leading to long-range Hamiltonians. Phys. Rev. E **58**, 2764–2778 (1998)
7. Lahiri, R., Barma, M., Ramaswamy, S.: Strong phase separation in a model of sedimenting lattices. Phys. Rev. E **61**, 1648–1658 (2000)
8. Clincy, M., Derrida, B., Evans, M.R.: Phase transition in the ABC model. Phys. Rev. E **67**, 066115 (2003)
9. Bodineau, T., Derrida, B., Lecomte, V., van Wijland, F.: Long range correlations and phase transition in non-equilibrium diffusive systems. J. Stat. Phys. **133**, 1013–1031 (2008)
10. Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Towards a nonequilibrium thermodynamics: a self-contained macroscopic description of driven diffusive systems. J. Stat. Phys. **135**, 857–872 (2009)
11. Fayolle, G., Furtlehner, C.: Stochastic deformations of sample paths of random walks and exclusion models. In: Drmota, M., Flajolet, P., Gardy, D., Gittenberger, B. (eds.) Mathematics and Computer Science III: Algorithms, Trees, Combinatorics and Probabilities. Trends in Mathematics. Birkhäuser, Basel (2004)
12. Fayolle, G., Furtlehner, C.: Stochastic dynamics of discrete curves and multi-type exclusion processes. J. Stat. Phys. **127**, 1049–1094 (2007)
13. Sandow, S., Schütz, G.: On $U_q[SU(2)]$-symmetric driven diffusion. Europhys. Lett. **26**, 7–12 (1994)
14. Blythe, R.A., Evans, M.R.: Nonequilibrium steady states of matrix-product form: a solver's guide. J. Phys. A **40**, R333 (2007)
15. Aizenmann, M., Lieb, E.: The IIIrd law of thermodynamics and the degeneracy of ground states in lattice systems. J. Stat. Phys. **24**, 279–297 (1981)
16. Liggett, T.M.: Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Springer, Berlin (1999)
17. Hewitt, E., Savage, L.J.: Symmetric measures on Cartesian products. Trans. Am. Math. Soc. **80**, 470–501 (1955)
18. Kipnis, C., Landim, C.: Scaling Limits of Interacting Particle Systems. Grundlehren der Mathematischen Wissenschaften, vol. 320. Springer, Berlin (1999)
19. Lebowitz, J.L.: Coexistence of phases in Ising ferromagnets. J. Stat. Phys. **16**, 463–476 (1977)
20. Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics. Grundlehren der Mathematischen Wissenschaften, vol. 271. Springer, Berlin (1985)
21. Dunford, N., Pettis, B.J.: Linear operators on summable functions. Trans. Am. Math. Soc. **47**, 323–392 (1940)
22. Mazur, S.: Über konvexe Menge in linearen normierten Raumen. Stud. Math. **4**, 70–84 (1933)
23. Kafri, Y., Biron, D., Evans, M.R., Mukamel, D.: Slow coarsening in a class of driven systems. Eur. Phys. J. B **16**, 669–676 (2000)
24. Armitage, J.V., Eberlein, W.F.: Elliptic Functions. London Mathematical Society Student Texts, vol. 67. Cambridge University Press, Cambridge (2006)