Journal of Statistical Physics, Vol. 54, Nos. 1/2, 1989

Statistical Mechanics of the Nonlinear Schrödinger Equation. II. Mean Field Approximation

Joel L. Lebowitz,[1] Harvey A. Rose,[2] and Eugene R. Speer[3]

Received July 5, 1988

We investigate a mean field approximation to the statistical mechanics of complex fields with dynamics governed by the nonlinear Schrödinger equation. Such fields, whose Hamiltonian is unbounded below, may model plasmas, lasers, and other physical systems. Restricting ourselves to one-dimensional systems with periodic boundary conditions, we find in the mean field approximation a phase transition from a uniform regime to a regime in which the system is dominated by solitons. We compute explicitly, as a function of temperature and density (L^2 norm), the transition point at which the uniform configuration becomes unstable to local perturbations; static and dynamic mean field approximations yield the same result.

KEY WORDS: Nonlinear Schrödinger equation; statistical mechanics; unbounded Hamiltonians; Gibbs measures; mean field.

1. INTRODUCTION

This paper is a continuation of the investigation begun in ref. 1 (referred to below as I) of continuum statistical mechanical models in which the Hamiltonian is not bounded below. The motivating physical systems are plasmas and lasers, which in appropriate regimes may be modeled by the nonlinear Schrödinger equation (NLSE) in three and two dimensions, respectively. (The Zakharov equations for the plasma case are also discussed briefly in I.) Our ultimate goal is a statistical theory of the excitations of these systems; we refer the reader to the introduction to I for more extensive discussion.

[1] Departments of Mathematics and Physics, Rutgers University, New Brunswick, New Jersey 08903.

[2] Los Alamos National Laboratory, Los Alamos, New Mexico 87545.

[3] Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903.

Here, as in I, we restrict our detailed study to the one-dimensional NLSE, defined on a finite interval $[0, L]$ and satisfying periodic boundary conditions. The Hamiltonian is

$$H(\phi) = \frac{1}{2} \int_0^L |\phi'(x)|^2 \, dx - \frac{1}{p} \int_0^L |\phi(x)|^p \, dx \tag{1.1}$$

where ϕ is a complex field. The exponent p, which measures the degree of nonlinearity, will be assumed to satisfy $p > 2$. If we regard the real and imaginary parts of ϕ as conjugate canonical variables, we find that the corresponding equation of motion for a time-dependent field $u(x, t)$ $\{= [\phi(x)](t)\}$ is the NLSE:

$$iu_t = -u_{xx} - |u|^{p-2} u, \qquad x \in [0, L] \tag{1.2}$$

Equation (1.2) is to be supplemented by boundary conditions

$$u(0, t) = u(L, t), \qquad t \in \mathbb{R} \tag{1.3}$$

and an initial condition

$$u(x, 0) = \phi_0(x), \qquad 0 \leqslant x \leqslant L \tag{1.4}$$

We remark that (1.2) conserves, in addition to the Hamiltonian H itself, the L^2 norm (analogous to the particle number) of the field:

$$N(\phi) = \int_0^L |\phi(x)|^2 \, dx \tag{1.5}$$

In I we posed the problem of constructing the Gibbs ensemble for this model, formally given by

$$Z^{-1} e^{-\beta H(\phi)} \prod_{x \in [0, L]} d\phi(x) \tag{1.6}$$

The lack of a lower bound on H implies that (1.6) is not normalizable. (As discussed in I, it also implies that the thermodynamic limit will not exist for these ensembles.) Since the L^2 norm is conserved by the dynamics, however, it is natural to modify the ensemble by introducing a cutoff on the L^2 norm of ϕ; thus, we define formally

$$d\nu_{\beta, f}(\phi) = Z^{-1} f(\|\phi\|_2^2) \, e^{-\beta H(\phi)} \prod_{x \in [0, L]} d\phi(x) \tag{1.7}$$

where f must decrease fairly rapidly at infinity.

In this paper we consider ensembles defined by several different cutoff functions f. We will refer to the conceptually simple choice

$$f(\|\phi\|_2^2)=\delta(\|\phi\|_2^2-N) \tag{1.8}$$

as the *canonical ensemble*. This is also, in most cases, the physically preferred choice, but is technically awkward to deal with in practice. For this reason, the measure actually constructed in I is a *modified canonical ensemble* defined by restricting the field to an L^2 ball rather than a sphere:

$$f(\|\phi\|_2^2)=\chi_{\{\phi \mid \|\phi\|_2^2 \leqslant N\}} \tag{1.9}$$

It is easy to see from the analysis of I that the essential features of this modified ensemble would be unchanged if f were chosen as an approximate delta function restricting the field to a spherical shell $N \leqslant \|\phi\|_2^2 \leqslant N+\delta N$.

The standard grand canonical ensemble, with $f(\|\phi\|_2^2)=\exp(-\beta\mu\,\|\phi\|_2^2)$, is not normalizable. It was noted in I, however, that a *modified grand canonical ensemble* (MGCE), with

$$f(\|\phi\|_2^2)=\exp(-\beta\mu\,\|\phi\|_2^{2r}) \tag{1.10}$$

is normalizable for $p<6$ as long as r satisfies $r \geqslant (p+2)/(6-p)$. Substituting (1.10) into (1.7) yields the usual Gibbs measure under the imposition of the mean constraint

$$\langle \|\phi\|_2^{2r} \rangle = N^r \tag{1.11}$$

On the other hand, numerical studies for the case $r=3$, described in I, suggest that this ensemble is concentrated on field configurations with either a rather high or a rather low L^2 norm and achieves (1.10) as a balance of these alternatives, and simple arguments suggest that this behavior will be more pronounced at higher values of r. The MGCE thus seems qualitatively quite different from the canonical or modified canonical ensembles.

For the modified canonical ensemble (1.9) we show in I that there exists a constant N_0 such that $\nu_{\beta,f}$ exists and is normalizable ($Z<\infty$) precisely in the cases $p<6$ and $p=6$, $N<N_0$. This result is closely related to the theory of (1.2) as an evolution equation. As has been shown,[1-3] for $p<6$ the initial value problem (1.2)–(1.4) has solutions existing globally in time for arbitrary smooth (H^1) initial data; for $p>6$ there are solutions which blow up in finite time for smooth initial data with arbitrarily small L^2 norm; for $p=6$ the value N_0 of the L^2 norm is critical: global existence is guaranteed for initial conditions with L^2 norm less than N_0.

One question we ask about these ensembles is the nature of typical field configurations. In I we noted that both numerical simulation of the dynamics of the (closely related) Zakharov equation and Monte Carlo investigation of the canonical ensemble on a lattice suggest that "typical" configurations at large N involve solitonlike concentrations of the field. At low N, on the other hand, typical configurations appear to be uniform. A natural question is then whether this change of behavior is smooth or involves some nonanalyticity, i.e., a phase transition. In I we could not definitely answer this question either analytically or numerically; this suggests that we try some approximate treatment.

There are various ways to derive "mean field" or "spherical model" approximations to the ensembles discussed above; we will consider one dynamic and several static approaches, all of which lead to different members of a one-parameter family of approximations. The most satisfactory static derivation, from a logical point of view, is a variational argument. Each of the measures (1.8)–(1.10) may be obtained by minimizing the free energy $\langle H\rangle_\rho - \beta^{-1}S(\rho)$ [where $S(\rho)$ is the entropy of the measure ρ] over a class of measures satisfying a suitable constraint: measures supported on the L^2-sphere and the L^2-ball for the canonical and modified canonical ensembles, respectively, and measures satisfying the mean constraint (1.11) for the MGCE. If we minimize instead over a restricted class, we obtain an approximate measure as well as an upper bound for the free energy of the exact ensemble.

In practice there are difficulties with this approach. To develop a tractable model we would like to work with Gaussian measures, a class which is inconsistent with the support constraints characterizing the canonical and modified canonical ensembles. Gaussian measures can be used to approximate the MGCE in the manner described above, but the resulting system is still quite complicated. This difficulty, and the qualitative differences noted above between the MGCE and the ensemble actually studied in I, lead us to use the MGCE primarily for technical and conceptual purposes in this paper. Instead, we will derive an approximate model by minimizing the free energy over the class of Gaussian measures which satisfy the mean constraint (1.11) in the "grand canonical" case $r=1$. Because the true grand canonical measure does not exist, we regard this procedure as yielding an approximation to the canonical or modified canonical ensemble.

This approximation, which we will refer to as the *mean field model*, has an attractive and natural structure. As we will see in Section 2, it is the grand canonical ensemble

$$Z^{-1} \exp[-\beta\tilde{H}(\phi)] \exp(-\beta\mu \, \|\phi\|_2^2) \prod_{x\in[0,L]} d\phi(x) \qquad (1.12)$$

for an "effective" quadratic Hamiltonian

$$\tilde{H}(\phi) = \frac{1}{2}\int_0^L |\phi'(x)|^2\,dx - \frac{1}{2}\int_0^L V(x)^{(p-2)/2}\,|\phi(x)|^2\,dx \qquad (1.13)$$

Comparison of (1.13) with the original Hamiltonian (1.1) suggests that the potential $V(x)$ should be an approximation to the square magnitude of the field, and this is in fact correct: for (1.12) to minimize the free energy, V must satisfy the condition

$$V(x) = \lambda\langle|\phi(x)|^2\rangle \qquad (1.14)$$

for a certain constant λ determined by the variational calculation. The expectation here is taken in the measure (1.12), and (1.14) may thus be regarded as a self-consistency condition on V. The Lagrange multiplier μ is of course adjusted to impose the mean constraint (1.11) in the case $r = 1$, which in view of (1.14) becomes

$$\int_0^L V(x)\,dx = \lambda N \qquad (1.15)$$

Formulas (1.14) and (1.15) are the basic equations which determine V and μ, and hence our static mean field model. We find in practice, however, that the model provides the best approximation to the true canonical and modified canonical ensembles if we set $\lambda = 1$ in these equations, rather than using the value from the variational calculation. This difficulty arises because we are approximating a variational problem for measures constrained by a support condition with one for Gaussian measures constrained in the mean; we have some justification for the preferred value $\lambda = 1$ from low-temperature considerations but lack a complete theory from which this approximation emerges naturally. In general, then, we will regard λ as a free parameter in our approximating model. In fact, other simple static approximation schemes can lead to (1.14) and (1.15) with different values of λ.

We note one other relation between the various classes of measures considered here. Our approximating measure (1.12) has $\langle\|\phi\|_2^{2r}\rangle < \infty$ for all r. Thus, the free energy for this measure is an upper bound for the free energy of an MGCE with appropriate constraint $\langle\|\phi\|_2^{2r}\rangle = \hat{N}^r$. We will use this argument in Section 3 to establish the boundedness of the free energy for (1.12).

An alternate approach to determining approximate equilibrium states for our system is first to approximate the dynamics, not of the fields themselves, but rather of their correlation functions, and then to look for stationary solutions of that approximate dynamics. Exact equations of motion for the two-point correlation functions, which follow from the NLSE (1.2), involve expectations of field products of degree higher than 2; we make an approximation in which such expectation values are replaced by products of powers of the correlation functions themselves. The mean field ensembles arising from solutions of (1.14) and (1.15) are indeed stationary states for the resulting dynamics, although other stationary states also exist.

In this paper, we investigate solutions of (1.14) and (1.15) analytically and also report some numerical results. We have the most complete picture of the solutions in the case $p=4$. Here we can show rigorously that in a region of high temperature and/or small N, (1.14) and (1.15) have a unique solution corresponding to a uniform state, i.e., to a $V(x)$ independent of x. As the temperature is lowered (for N not too small) we observe numerically that the system undergoes a first-order phase transition: (1.14) and (1.15) acquire a nonuniform, single-peaked solution which is thermodynamically preferred (and which does not branch continuously from the uniform solution). We also have considerable analytic evidence for this behavior. As the temperature is lowered further, the mean field configuration becomes more peaked; at zero temperature (if λ is chosen properly) it is equal to the square of the magnitude of the ground state of the true Hamiltonian (1.1). In fact, zero-temperature solutions of (1.14) and (1.15) may be shown to correspond to solutions of the variational problem $\delta H(\phi)/\delta\phi=0$ with constraint $\|\phi\|_2^2=N$.

We also investigate the stability against local perturbations of the state given by a uniform mean field, both thermodynamically and dynamically. We find that the regions of local stability in the static and dynamic senses coincide. At high temperature, where we have proved for $p=4$ and believe for $p<6$ that this state is absolutely thermodynamically preferred, local stability in the static sense will of course be satisfied, although as discussed in Section 3.4, local stability does not appear to imply global stability.

We have argued in I that the true model probably does not have a phase transition in the sense of any nonanalytic behavior in the parameters. Nevertheless, there is fairly good qualitative agreement between Monte Carlo simulations of that model and predictions of the mean field theory. We discuss the comparison in Section 3.4.

The paper is organized as follows. In Section 2 we derive (1.14) and (1.15) by the variational calculation discussed above and point out that they also arise from several simpler schemes for approximating the original model. These different derivations lead to different values of the constant λ. In Section 3 we describe our analytic results and report on numerical investigations; the combination of these approaches leads to a fairly complete picture of the solutions of (1.14)–(1.15), particularly in the case $p = 4$. Section 4 is devoted to a discussion of the dynamic mean field approximation and its consequences.

2. FORMULATION OF THE MODEL

2.1. Derivation from a Variational Principle

We here consider the problem, discussed in the introduction, of minimizing the free energy $\langle H \rangle_\rho - \beta^{-1} S(\rho)$ over Gaussian measures ρ which satisfy the mean constraint (1.11); we will shortly specialize to the case $r = 1$ of this constraint. As posed, however, this problem is purely formal, since the two terms in this free energy are both infinite for the measures we consider. To obtain a well-defined variational problem we may either carry out the minimization on a lattice and then take the continuum limit, or work directly in the continuum but use Wiener measure as a reference measure. These procedures yield the same result; we will follow the latter.

We begin by describing the relevant Wiener measure. Let C denote the set of continuous, complex-valued functions ϕ on $[0, L]$ which satisfy $\phi(0) = \phi(L)$; C is the space of field configurations for our theory. The Wiener measure on C, formally given by

$$d\mu_\beta(\phi) = Z_0^{-1} e^{-\beta(\phi, T\phi)/2} \prod_{x \in [0, L]} d\phi(x) \tag{2.1}$$

with

$$T\phi = -\phi''$$

may be defined more precisely as the measure with marginal distributions[4]

$$\text{Prob}(\{\phi \mid \phi(x_i) \in A_i,\ 1 \leqslant i \leqslant n+1\})$$
$$= p_\beta^L(0, 0)^{-1} \int_{y_i \in A_i} \prod_{i=1}^{n+1} p_\beta^{x_i - x_{i-1}}(y_i, y_{i-1}) \Big|_{y_0 = y_{n+1}} dy_1 \cdots dy_{n+1} \tag{2.2}$$

Here $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = L$, $A_1, \ldots, A_{n+1}$ are Borel subsets of $\mathbb{C}$, dy is Lebesque measure on $\mathbb{C}$, and for $y, y' \in \mathbb{C}$,

$$p_\beta^x(y, y') \equiv (\beta/2\pi x) \exp(-\beta\, |y - y'|^2/2x)$$

We remark that μ_β is invariant under complex conjugation and that μ_β is not normalizable [it is invariant under field translations $\phi(x) \to \phi(x) + \phi_0$].

If ρ is a measure on C which is absolutely continuous with respect to μ_β, so that $d\rho(\phi) = g(\phi)\, d\mu_\beta(\phi)$, we define the relative entropy and free energy of ρ to be, respectively,

$$S_\beta(\rho) = -\int_C g \log g \, d\mu_\beta$$

and

$$F_\beta(\rho) = \left\langle -\frac{1}{p}\int_0^L |\phi(x)|^p \, dx \right\rangle_\rho - \beta^{-1} S_\beta(\rho) \tag{2.3}$$

The modified canonical ensemble studied in I is given [see (1.7), (1.9)] by

$$Z^{-1} \exp\left[\frac{\beta}{p}\int_0^L |\phi(x)|^p \, dx\right] \chi_{\{\phi \mid \|\phi\|_2^2 \leq N\}} \, d\mu_\beta$$

and may be regarded as the measure supported on the L^2-ball which minimizes the relative free energy (2.3). As approximating ensembles we consider normalized Gaussian measures having the form

$$d\rho_A(\phi) = Z_A^{-1} e^{\beta(\phi, A\phi)/2} \, d\mu_\beta(\phi) \tag{2.4}$$

and satisfying

$$\langle \|\phi\|_2^{2r} \rangle_{\rho_A} = N^r \tag{2.5}$$

Here A is a bounded operator on $L^2([0, L])$ which commutes with complex conjugation and satisfies $(T - A) > \varepsilon > 0$. For such A the operator $R = (T - A)^{-1}$ has a continuous integral kernel $r(x, y)$; ρ_A is then the (normalizable) Gaussian measure for which the real and imaginary parts of ϕ, denoted ϕ^{R} and ϕ^{I}, are independent, identically distributed random variables with covariance given by

$$\langle \bar{\phi}(x)\, \phi(y) \rangle = 2\langle \bar{\phi}^{\mathrm{R}}(x)\, \phi^{\mathrm{R}}(y) \rangle = 2\beta^{-1} r(x, y) \tag{2.6}$$

(see, e.g., ref. 4; the continuity of r will be discussed below). We may solve the constrained minimization problem by introducing a Lagrange multiplier μ [differing by a factor from that of (1.12)] and seeking extremals of the functional

$$F(A, \mu) = F(\rho_A) - \frac{\mu}{2}(\langle \|\phi\|_2^{2r} \rangle - N^r) \tag{2.7}$$

From now on we will restrict our considerations to the case $r=1$. To simplify (2.7) we note first that any complex random variable X whose real and imaginary parts are identically distributed independent Gaussian variables satisfies

$$\langle |X|^p \rangle = \Gamma\left(\frac{p+2}{2}\right) \langle |X|^2 \rangle^{p/2} \tag{2.8}$$

and second that, for any bounded operator B on $L^2[0, L]$,

$$\langle (\phi, B\phi) \rangle_{\rho_A} = 2\beta^{-1} \operatorname{Tr}(BR)$$

as is easily verified by approximating B operators of finite rank. Thus, if we set

$$v(x) = \langle |\phi(x)|^2 \rangle = 2\beta^{-1} r(x, x) \tag{2.9}$$

(2.7) reduces to

$$F(A, \mu) = -\frac{1}{2} \Gamma\left(\frac{p}{2}\right) \int_0^L v(x)^{p/2}\, dx + \beta^{-1} \operatorname{Tr}[R(A-\mu)]$$
$$+ \frac{\mu}{2} N - \beta^{-1} \log Z_A \tag{2.10}$$

We now look for stationary points of F, and will use the following notation for functional derivatives: if X, Y, and Z are Banach spaces, $U \subset X \times Y$ is open, and $G: U \to Z$, then $D_x G(x, y): X \to Z$ denotes the Fréchet derivative of G with respect to x at the point (x, y). Since $D_A RB = -RBR$, (2.10) leads to

$$[D_A F(A, \mu)] B$$
$$= \beta^{-1} \Gamma\left(\frac{p+2}{2}\right) \int_0^L v(x)^{(p-2)/2}\, \hat{r}(x, x)\, dx - \beta^{-1} \operatorname{Tr}[RBR(A-\mu)]$$
$$= -\beta^{-1} \operatorname{Tr}[RBR(A-\mu-W)] \tag{2.11}$$

where $\hat{r}$ is the integral kernel of RBR and W denotes the operator of multiplication by $\Gamma((p+2)/2)\, v(x)^{(p-2)/2}$. Setting the variation (2.11) to zero, we finally conclude that at an extremal of $F(A, \mu)$, A is a multiplication operator

$$A = V(x)^q + \mu \tag{2.12}$$

with V a continuous potential satisfying

$$V(x) = \lambda_1 \langle |\phi(x)|^2 \rangle = 2\beta^{-1} \lambda_1 r(x, x) \tag{2.13}$$

and from (2.5) with $r=1$,

$$\int_0^L V(x)\,dx = \lambda_1 N \tag{2.14}$$

Here $q=(p-2)/2$ and $\lambda_1 = \Gamma((p+2)/2)^{1/q}$. Note that (2.12)–(2.14) recapture (1.13)–(1.15).

2.2. Reformulation of the Model

For later convenience we may reformulate the model as follows. Let $X = L^\infty([0, L]; \mathbb{R})$ be the space of measurable, essentially bounded, real-valued functions on $[0, L]$, and for $N \in \mathbb{R}$ let $X_N \subset X$ be the affine subspace defined by $X_N = \{V \in X \mid \int_0^L V\,dx = \lambda N\}$. For $V \in X$ let $R(\mu, V) = (T - V^q - \mu)^{-1}$ denote the resolvent of $T - V^q$. We will prove below (see Theorem 3.3) that [for μ not in the spectrum $\operatorname{spec}(T - V^q)$] $R(\mu, V)$ has nonnegative continuous integral kernel $r(\mu, V)(x, y)$ and that if $\mu < \inf \operatorname{spec}(T - V^q)$, then $r(\mu, V)(x, x) > 0$. Now, for $N > 0$ and $V \in X_N$, let $\kappa(V)$ be the unique real number which satisfies $\kappa(V) < \inf \operatorname{spec}(T - V^q)$ and for which

$$\beta^{-1} \operatorname{Tr} R(\kappa(V), V) \equiv \beta^{-1} \int_0^L r(\kappa(V), V)(x, x)\,dx = N/2 \tag{2.15}$$

Then, defining $J^\lambda_{\beta, N}\colon X_N \to X_N$ by

$$J^\lambda_{\beta, N}(V)(x) = 2\lambda\beta^{-1} r(\kappa(V), V)(x, x) \tag{2.16}$$

we see that solving (2.13) and (2.14), and thus determining the mean field at any β, N, reduces to solving the fixed-point equation on X_N:

$$V = J^{\lambda_1}_{\beta, N}(V) \tag{2.17}$$

We next define a generalized free energy functional on X_N by

$$\begin{aligned} \hat{F}^\lambda_{\beta, N}(V) = {} & \frac{1}{2}\int_0^L V(x)^q\, v(x)\,dx - \frac{\lambda^q}{p}\int_0^L v(x)^{p/2}\,dx \\ & + \frac{\kappa(V)}{2} N - \beta^{-1} \log \tilde{Z}_V \end{aligned} \tag{2.18}$$

where

$$\tilde{Z}_V = \int_C \exp\left\{\frac{\beta}{2}(\phi, [V^q + \kappa(V)]\phi)\right\} d\mu_\beta$$

Note that (1) the free energy $F(A, \mu)$ reduces to $\hat{F}^{\lambda_1}(V)$ when (2.12) is inserted in (2.10) and μ is replaced by $\kappa(V)$, and (2) (2.17) is the condition for an extremal of $\hat{F}^{\lambda_1}$. Thus, the construction of a solution to our mean field model corresponds to minimizing $\hat{F}^{\lambda_1}$ over X_N. In particular, if more than one solution of (2.17) exists, we choose the solution with minimum free energy $\hat{F}^{\lambda_1}(V)$.

2.3. Alternate Mean Field Models

As mentioned in the introduction, there are alternate derivations of mean field models which yield (1.14) and (1.15), or equivalently (2.17), with different values of the constant λ. We may, for example, simply replace the true Hamiltonian H of (1.1) by an approximate quadratic Hamiltonian and form the grand canonical ensemble, leading to (1.12)–(1.13); μ is again to be adjusted so that (1.15) holds. A natural prescription for the approximating Hamiltonian is to choose the potential $V(x)$ so that the potential energy density in $\tilde{H}$ is equal in mean to the nonlinear term in the original Hamiltonian (1.1):

$$\langle V(x)^q \, |\phi(x)|^2 \rangle = \frac{2}{p} \langle |\phi(x)|^p \rangle \tag{2.19}$$

where, as usual, expectations are computed with the approximating measure. Using (2.8) and (2.9), we may rewrite (2.19) as

$$V(x) = \lambda_2 \langle |\phi(x)|^2 \rangle = 2\lambda_2 \beta^{-1} r(\mu, V)(x, x) \tag{2.20}$$

where $\lambda_2 = \Gamma(p/2)^{1/q}$, recovering (1.14) or equivalently the fixed-point equation

$$V = J^{\lambda}_{\beta, N}(V) \tag{2.21}$$

with $\lambda = \lambda_2$.

Alternatively, we may ask that our approximate theory reproduce the correct behavior for the ensemble in the limit of zero temperature. In that limit the canonical measure (1.7)–(1.8) will be concentrated on the lowest energy solution ϕ_0 of the variational problem:

$$\frac{\delta}{\delta\phi} [H - \mu \, \|\phi\|_2^2]_{\phi = \phi_0} \equiv -\phi_0'' - |\phi_0|^{p-2} \phi_0 - \mu\phi_0 = 0 \tag{2.22}$$

while the approximate measure (1.12) will be concentrated on the lowest energy eigenfunction ϕ_1 of $\tilde{H}$, which satisfies

$$-\phi_1'' - V^q \phi_1 - \mu\phi_1 = 0 \tag{2.23}$$

For these measures to agree as closely as possible we must have $V = |\phi_1|^2 = \langle \|\phi_1\|^2 \rangle_{\beta=\infty}$. Thus, we are led back to (2.21), with $\lambda = \lambda_3 = 1$.

As discussed in the introduction, the existence of these alternative models leads us to regard λ in (2.21) as a free constant in our theory. In this context there are two possibilities for choosing among multiple solutions of (2.21). On the one hand, (2.21) is the condition for an extremal of the generalized free energy $\hat{F}^{\lambda}(V)$, and it is mathematically natural simply to reformulate our problem as that of minimizing this quantity. Physically, however, it is the solution which minimizes the true free energy $\hat{F}^{\lambda_1}(V)$ which should be preferred. There seems little difference in the practical conclusions of these two approaches.

Remark 2.1. Even with the choice $\lambda = 1$ the canonical and mean field ensembles are quite different at low temperature: the canonical ensemble will have small fluctuations around the ground state discussed above, but the mean field model, which is Gaussian with mean zero and macroscopic variance, will have large fluctuations at all temperatures. (This is discussed further in Section 3.4.) As an alternate approximation we could consider a Gaussian ensemble with nonzero mean, that is, we could replace (2.4) by

$$d\rho_{A,\psi}(\phi) = Z_A^{-1} e^{\beta(\phi-\psi, A(\phi-\psi))/2}\, d\mu_\beta(\phi)$$

and then minimize the free energy with respect to A and ψ. We would expect this to yield a better approximation than the model we actually study, but it is less tractable analytically.

3. PROPERTIES OF THE MODEL

In this section we describe in detail our knowledge of the mean field model, that is, of the problem of finding solutions of the fixed-point equation (2.21). Here we give a brief summary of our results. It is convenient to organize these into categories according to the values of the parameters β and N:

1. *General results.* We show that the behavior of the Gaussian free energy $\hat{F}^{\lambda}(V)$ is consistent with the behavior of the true free energy found in I: $\hat{F}^{\lambda}(V)$ is bounded below for $p < 6$ and is unbounded for $p > 6$. We also have partial results for $p = 6$. Finally, we show that mean field configurations are always smooth.

2. *High temperature and the uniform field configuration.* A uniform mean field is always a solution of (2.21); for the special case $p = 4$ we can show that this is the unique solution at high temperature or at low values

of N. We also discuss the local stability of this solution against small perturbations and determine exactly the curve in the β–N phase plane on which this stability is lost, as well as the nature of the bifurcation at this curve.

3. *Low-temperature solitions.* Zero-temperature solutions of (2.21) may be analyzed completely and shown to correspond to solutions of the time-independent NLSE

$$-\phi''(x) - \lambda^q \phi^{p-1}(x) - \mu\phi(x) = 0$$

Here again, as in Section 2.3, we see that the choice $\lambda = 1$ yields the best correspondence between the mean field and exact models. All solutions at low temperature lie on smooth curves passing through the zero-temperature solutions.

4. *Transition region.* Numerical investigations of the transition between high- and low-temperature regimes indicate that the loss of stability of the uniform mean field is via a first-order transition to a single-peaked field configuration, which changes smoothly to the configuration corresponding to the ground state of the NLSE as the temperature is lowered to zero.

Before describing our results in more detail, we must introduce some notation. For any $W \in X$ we write the eigenvalues of $T - W$ as $\{\omega_k(W)\}_{k=-\infty}^{\infty}$, numbering them to correspond with $\omega_k(0) \equiv \omega_k = (2\pi k/L)^2$:

$$\omega_0(W) \leqslant \omega_1(W) \leqslant \omega_{-1}(W) \leqslant \cdots$$

The eigenfunctions (normalized in L^2) will be denoted $\psi_k(W)$, with $\psi_k(0)(x) \equiv \psi_k(x) = L^{-1/2} e^{2\pi i k x/L}$. In this notation, the resolvent $R(\mu, W) = (T - W - \mu)^{-1}$ has integral kernel

$$r(\mu, W)(x, y) = \sum_{k=-\infty}^{\infty} \frac{\bar{\psi}_k(V^q)(x)\, \psi_k(V^q)(y)}{\omega_k(V^q) - \mu} \tag{3.1}$$

and the basic equations (2.21) and (2.15) of the mean field theory become

$$V(x) = 2\lambda\beta^{-1} \sum_{k=-\infty}^{\infty} \frac{|\psi_k(V^q)(x)|^2}{\omega_k(V^q) - \kappa(V)} \tag{3.2}$$

and

$$\sum_{k=-\infty}^{\infty} \frac{1}{\omega_k(V^q) - \kappa(V)} = \frac{\beta N}{2} \tag{3.3}$$

3.1. General Results

We first ask whether the free energy $\hat{F}^{\lambda}(V)$ is bounded below on X_N. Recall from I that the free energy for the true model in the modified canonical ensemble is bounded below if $p<6$ or if $p=6$ and $N>N_0$, and is not bounded below if $p>6$ or $p=6$ and $N>N_0$, where $N_0=\pi\sqrt{3}/2$. We have a similar but less complete result for the mean field model:

Theorem 3.1. The Gaussian free energy defined on X_N, $\hat{F}^{\lambda}(V)$: (a) is bounded below if $p<6$; (b) is not bounded below if $p>6$ or if $p=6$ and $N>N_0/\lambda$.

We suspect that $\hat{F}^{\lambda}$ is in fact bounded below for $p=6$, $N<N_0/\lambda$, but do not have a proof. Note that we again have the closest correspondence between the mean field and exact models if we choose $\lambda=1$.

Proof. (a) For the lower bound we give a thermodynamic argument. We discuss in detail only the true free energy $\hat{F}^{\lambda_1}$; any other value of λ may be treated by observing that $\hat{F}^{\lambda}$ is the true free energy for a Hamiltonian in which the coefficient p^{-1} of $\|\phi\|_p^p$ is replaced by $(\lambda/\lambda_1)^q p^{-1}$.

Let $\tilde{\nu}_N$ denote the modified grand canonical ensemble defined by

$$d\tilde{\nu}_N=\tilde{Z}^{-1}\exp[-\beta(H_I+\mu\,\|\phi\|_2^{2r})]\,d\mu_\beta$$

and the constraint (1.11). Here we choose r to be a fixed integer satisfying $r\geqslant(p+2)/(6-p)$ and suppress the dependence of the measure on β. We know from I that the partition function $\tilde{Z}$ is finite, and we may compute

$$\begin{aligned}F(\tilde{\nu}_N)&=\langle H_I\rangle_{\tilde{\nu}_N}+\beta^{-1}\langle-\beta H_I-\beta\mu\,\|\phi\|_2^{2r}\rangle_{\tilde{\nu}_N}-\beta^{-1}\log\tilde{Z}\\&=-\mu N^r-\beta^{-1}\log\tilde{Z}\end{aligned}\tag{3.4}$$

Note also that $\tilde{Z}<\infty$ implies $\mu>0$, so that

$$\frac{\partial F(\tilde{\nu}_N)}{\partial N}=-r\mu N^{r-1}<0\tag{3.5}$$

Now let ρ be any measure satisfying (1.11), with density (relative to μ_β) given by $Z_\rho^{-1}\exp(-\beta H_\rho)$, where $Z_\rho=\int\exp(-\beta H_\rho)\,d\mu_\beta$. From Jensen's inequality,

$$\begin{aligned}\tilde{Z}&=\int\exp[-\beta(H_I+\mu\,\|\phi\|_2^{2r})]\,d\mu_\beta\\&=Z_\rho\langle\exp[-\beta(H_I+\mu\,\|\phi\|_2^{2r}-H_\rho)\rangle_\rho\\&\geqslant Z_\rho\exp[-\beta(\langle H_I\rangle_\rho+\mu N^r-\langle H_\rho\rangle_\rho)]\end{aligned}$$

With (3.4), this yields a lower bound for the free energy of ρ:

$$F(\rho) = \langle H_I \rangle_\rho + \langle H_\rho \rangle_\rho - \beta^{-1} \log Z_\rho \geqslant F(\tilde{v}_N) \tag{3.6}$$

Finally, if $\hat{\rho}$ is the measure associated with a mean field V, so that

$$d\hat{\rho} = \hat{Z}^{-1} \exp\{\beta[(\phi, V^q \phi) + \kappa(V) \|\phi\|_2^2]\} \, d\mu_\beta$$

then by Wick's theorem

$$\langle \|\phi\|_2^{2r} \rangle_{\hat{\rho}} \leqslant r! \, \langle \|\phi\|_2^2 \rangle_{\hat{\rho}}^r = r! \, N^r$$

Thus, if $\hat{N} = (r!)^{1/r} N$, (3.5) and (3.6) imply

$$\hat{F}^{\lambda_1}(V) = F(\hat{\rho}) \geqslant F(\tilde{v}_{\hat{N}})$$

(b) We now turn to the case $p \geqslant 6$ of the theorem, where our proof is based directly on the definition

$$\hat{F}^\lambda(V) = \frac{1}{2} \int_0^L V(x)^q \, v(x) \, dx - \frac{\lambda^q}{p} \int_0^L v(x)^{p/2} \, dx + \frac{\kappa(V)}{2} N - \beta^{-1} \log \tilde{Z}_V$$

We first note that the inequality $ta + (1-t)b \geqslant a^t b^{1-t}$ $(0 \leqslant t \leqslant 1)$, applied with $a = (\lambda v)^{p/2}$, $b = V^{p/2}$, and $t = 2/p$, yields

$$\hat{F}^\lambda(V) \leqslant \frac{q}{p\lambda} \int_0^L V(x)^{p/2} \, dx + \frac{\kappa(V)}{2} N - \beta^{-1} \log \tilde{Z}_V \tag{3.7}$$

We will define a sequence of potentials V_n such that $\lim_{n\to\infty} \hat{F}^\lambda(V_n) = -\infty$. For these potentials, the significant terms on the right-hand side of (3.7) are the first two, each being of order $\|V_n\|_{p/2}^{p/2}$. The last term is of lower order and we relegate its treatment to Appendix A, where we show that for some constant B,

$$\beta^{-1} \log \tilde{Z}_V \geqslant B(\|V\|_{p/2}^{p/2})^{(p-2)/(p+2)} \tag{3.8}$$

In defining V_n we find it convenient to work on the interval $[-L/2, L/2]$ rather than on $[0, L]$. A key role will be played by the standard interpolation inequality on $\mathbb{R}$,

$$\|f\|_{p,\mathbb{R}}^p \leqslant C_p \, \|f'\|_{2,\mathbb{R}}^{(p-2)/2} \, \|f\|_{2,\mathbb{R}}^{(p+2)/2} \tag{3.9}$$

We note that $C_6 = (2/\pi)^2$ and that, for $p = 6$, equality in (3.9) is obtained for $f(x) = f_0(x) \equiv \sqrt{2}\,\mathrm{sech}^{1/2}\, x$. If we scale this function by defining $g_n(x) = nf_0(n^2 x)\,|_{[-L/2, L/2]}$, then for $\varepsilon > 0$ and n sufficiently large,

$$\|g_n\|_6^6 \geqslant (1-\varepsilon)\, n^4\, \|f_0\|_{6,\mathbb{R}}^6$$

$$\|g_n'\|_2^2 \leqslant n^4\, \|f_0'\|_{2,\mathbb{R}}^2$$

$$\|g_n\|_2^2 \leqslant \|f_0\|_{2,\mathbb{R}}^2$$

so that

$$\|g_n\|_6^6 \geqslant (1-\varepsilon)\, C_6\, \|g_n'\|_2^2\, \|g_n\|_2^4$$

Let $M_n = \|g_n\|_2^2$ and define $V_n(x) = \lambda N M_n^{-1} g_n^2(x)$; then $V_n \in X_N$. We estimate the lowest eigenvalue of $T - V_n^q$ by a variational calculation: taking $\phi_n(x) = M_N^{-1/2} g_n(x)$ yields

$$\begin{aligned}\omega_0(V_n^q) &\leqslant (\phi_n, (T - V_n^q)\,\phi_n)\\ &\leqslant [M_n^3(1-\varepsilon)\, C_6]^{-1}\, \|g_n\|_6^6 - (\phi_n, V_n^q \phi_n)\\ &= [(\lambda N)^3\,(1-\varepsilon)\, C_6]^{-1}\, \|V_n\|_3^3 - (\lambda N)^{-1}\, \|V_n\|_{p/2}^{p/2}\end{aligned}$$

Since $\kappa(V_n) \leqslant \omega_0(V_n^q)$, (3.7) and (3.8) give

$$\begin{aligned}\hat{F}^\lambda(V_n) \leqslant{} & -(\lambda p)^{-1}\, \|V_n\|_{p/2}^{p/2} + [2\lambda^3 N^2(1-\varepsilon)\, C_6]^{-1}\, \|V_n\|_3^3\\ & - B(\|V\|_{p/2}^{p/2})^{(p-2)/(p+2)}\end{aligned} \tag{3.10}$$

Because $\|V_n\|_{p/2}^{p/2}$ scales for large n as $n^{(p-2)}$, (3.10) shows that $\lim_{n\to\infty} \hat{F}^\lambda(V_n) = -\infty$ for $p > 6$ or $p = 6$, $\lambda N > (3/C_6)^{1/2}$, as claimed. ∎

We now turn to the question of smoothness of minimizing potentials, and introduce function spaces $X^{(k)}$, where $X^{(0)} \equiv X \equiv L^\infty[0, L]$, and, for $k > 0$, $X^{(k)}$ is the space of functions f on $[0, L]$ such that $f \in C^{(k-1)}$, $f^{(k-1)}$ is absolutely continuous, and $f^{(k)} \in X$. (All derivatives respect the periodic boundary conditions.) Resolvents for potentials lying in $X^{(k)}$ are smooth in the direction corresponding to translation of the system:

Lemma 3.2. Suppose that $W \in X^{(k)}$, that $\mu \notin \mathrm{spec}(T - W)$, and that $r(\mu, W)(x, y)$ denotes the integral kernel of the resolvent $R(\mu, W)$. Then for any j, $0 \leqslant j \leqslant k$,

$$(\partial_x + \partial_y)^j\, r(\mu, W)(x, y) \tag{3.11}$$

is continuous, and its first partial derivatives exist almost everywhere and are essentially bounded.

Proof. We abbreviate the resolvent $R(\mu, 0)$ by $R(\mu)$ and note that its kernel may be computed explicitly:

$$r(\mu)(x, y) = \begin{cases} \dfrac{\cosh \alpha(L/2 - x + y)}{2\alpha \sinh \alpha L/2} & \text{if} \quad x \geqslant y \\ \dfrac{\cosh \alpha(L/2 - y + x)}{2\alpha \sinh \alpha L/2} & \text{if} \quad y \geqslant x \end{cases} \tag{3.12}$$

where $\alpha^2 = -\mu$; $r(\mu)$ clearly is translation invariant and has essentially bounded first partial derivatives. For μ sufficiently large and negative we may express the kernel of $R(\mu, W)$ by a Neumann series:

$$r(\mu, W)(x, y) = \sum_{k=0}^{\infty} \int \cdots \int r(\mu)(x, z_1)\, W(z_1)\, r(\mu)(z_1, z_2) \times W(z_2) \cdots r(\mu)(z_n, y)\, dz_1 \cdots dz_k \tag{3.13}$$

We proceed by induction on j. For $j=0$ the conclusions of the lemma follow immediately from (3.13). For $j>0$ we use $\partial_\xi r(\mu)(\xi, \eta) = -\partial_\eta r(\mu)(\xi, \eta)$ and repeated integration by parts to obtain a representation of (3.11) as a sum of terms of the form

$$\int \cdots \int r(\mu)(x, z_1)\, W^{(m_1)}(z_1)\, r(\mu)(z_1, z_2)\, W^{(m_2)} \cdots r(\mu)(z_n, y)\, dz_1 \cdots dz_k$$

where $\sum m_i = j$. Again, the conclusions of the lemma follow immediately. Finally, we analytically continue to other points of the resolvent set by repeated use of

$$r(\mu', W)(x, y) = \sum_{k=0}^{\infty} (\mu' - \mu)^k \int \cdots \int r(\mu, W)(x, z_1) \cdots \times r(\mu, W)(z_n, y)\, dz_1 \cdots dz_k \tag{3.14}$$

We now argue as above but from (3.14) rather than (3.13); we no longer have translation invariance, but may write

$$\partial_\xi r(\mu, W)(\xi, \eta) = [(\partial_\xi + \partial_\eta) - \partial_\eta]\, r(\mu, W)(\xi, \eta)$$

and use the differentiability properties of $r(\mu, W)$ already established to verify the same properties for $r(\mu', W)$. ∎

Smoothness of minimizing potentials is a direct consequence:

Theorem 3.3. Any solution $V \in X_N$ of (2.21) is C^∞.

Proof. The lemma and the definition (2.16) of J imply immediately that, if $V \in X_N \cap X^{(k)}$, then $J(V) \in X_N \cap X^{(k+1)}$. Hence, any fixed point of J lies in $\bigcap_{k=0}^{\infty} X^{(k)} = C^{\infty}[0, L]$.

3.2. High Temperature and the Uniform Solution

We now turn to the description of the model in the high-temperature regime, where uniform field configurations are expected. As noted above, the uniform potential in X_N,

$$V_0(x) = \lambda N/L$$

is always a solution of (2.21). We remark that from (3.12),

$$r(\mu)(x, x) = \frac{\coth(\alpha L/2)}{2\alpha} \tag{3.15}$$

so that the normalization condition (2.15) implies

$$\kappa(V_0) = -\gamma^2 - V_0^q$$

with γ the unique positive solution of

$$\frac{L \coth(\gamma L/2)}{2\gamma} = \frac{\beta N}{2} \tag{3.16}$$

Basic estimates that we need are given in the following result.

Lemma 3.4. (a) Suppose that $\mu \in \mathbb{R}$ and $V \in X$ satisfy $\|V\|_1 \|r(\mu)\|_\infty \leqslant \theta < 1$. Then the kernel $r(\mu, V^q)$ satisfies

$$\|r(\mu, V^q)\|_\infty \leqslant (1-\theta)^{-1} \|r(\mu)\|_\infty$$

(b) For any $V_1, V_2 \in X_N$, $|\kappa(V_1) - \kappa(V_2)| \leqslant \|V_1^q - V_2^q\|_\infty$.

(c) For any $V \in X_N$ with $V(x) \geqslant 0$, $\kappa(V) \leqslant \kappa(V_0) + V_0^q$.

Proof. (a) We simply estimate all terms in the expansion (3.13).

(b) By the min-max principle,

$$|\omega_k(V_1^q) - \omega_k(V_2^q)| \leqslant \|V_1^q - V_2^q\|_\infty$$

But

$$\sum_k \frac{1}{\omega_k(V_1^q) - \kappa(V_1)} = \sum_k \frac{1}{\omega_k(V_2^q) - \kappa(V_2)} = \lambda\beta N$$

so that

$$\sum_k \frac{\kappa(V_1)-\kappa(V_2)}{[\omega_k(V_1^q)-\kappa(V_1)][\omega_k(V_2^q)-\kappa(V_2)]}$$
$$=\sum_k \frac{\omega_k(V_1^q)-\omega_k(V_2^q)}{[\omega_k(V_1^q)-\kappa(V_1)][\omega_k(V_2^q)-\kappa(V_2)]}$$

from which the result follows.

(c) Since $V \geqslant 0$, $\omega_k(V^q)<\omega_k$. If $\omega_0(V^q) \leqslant \kappa(V_0)+V_0^q$, the result is immediate; otherwise,

$$\begin{aligned}\sum_k [\omega_k(V^q)-\kappa(V_0)-V_0^q]^{-1} &\geqslant \sum_k [\omega_k-\kappa(V_0)-V_0^q]^{-1}\\ &=\beta N/2\\ &=\sum_k [\omega_k(V^q)-\kappa(V)]^{-1}\end{aligned}$$

so that $\kappa(V) \leqslant \kappa(V_0)+V_0^q$. ∎

We now show that, in the case $p=4$, the uniform field $V=V_0$ is the unique solution of (2.21) at high temperature or small N (the region given in the following theorem is presumably not optimal). Since by Theorem 3.1 the free energy $\hat{F}^\lambda(V)$ is bounded below, it has unique minimum at V_0 in this region.

Theorem 3.5. Suppose that $p=4$ $(q=1)$. If $\beta N^2<2L/3\lambda$, then the uniform potential V_0 is the unique solution to (2.21) in X_N.

Proof. From (3.12) and (3.16),

$$\|r(\kappa(V_0), V_0)\|_\infty = \|r(\kappa(V_0)+V_0)\|_\infty = \beta N/2L \tag{3.17}$$

Now suppose that $V \in X_N$ satisfies $J^\lambda_{\beta,N}(V)=V$; in particular, $V(x) \geqslant 0$ and hence $\|V\|_1=N$. The operator identity

$$\begin{aligned}&R(\kappa(V), V)-R(\kappa(V_0), V_0)\\ &\quad=R(\kappa(V), V)[V-V_0+\kappa(V)-\kappa(V_0)]\, R(\kappa(V_0), V_0)\end{aligned}$$

together with (3.17) and the estimates of Lemma 3.4, yields

$$\begin{aligned}\|V-V_0\|_\infty &= \|J_{\beta,N}(V)-V_0\|_\infty\\ &\leqslant 2\lambda\beta^{-1}L\,\|r(\kappa(V), V)\|_\infty\\ &\qquad\times \|V-V_0+\kappa(V)-\kappa(V_0)\|_\infty\, \|r(\kappa(V_0), V_0)\|_\infty\\ &\leqslant (\lambda\beta N^2/L)(1-\lambda\beta N^2/2L)^{-1}\,\|V-V_0\|_\infty\end{aligned}$$

Since $\beta N^2<2L/3\lambda$, we must have $V=V_0$. ∎

For $p>4$, or in regimes of β and N where multiple solutions of the fixed-point equation (2.21) exist, we have not been able to find all solutions or determine (analytically) the solution which minimizes the potential energy. We can always determine, however, whether or not V_0 is a *stable* fixed point of J: as described in the next theorem, V_0 is stable for small N or, for higher values of N, above an N-dependent transition temperature. This is relevant because the stability of V_0 as a fixed point corresponds to its being a *local* minimum of the corresponding potential energy F^λ (in particular, of the true free energy for $\lambda=\lambda_1$). Moreover, we can determine by a local analysis the nature of the bifurcation which occurs when stability is lost, and thus obtain information about other solutions of (2.21).

Theorem 3.6. (a) V_0 is a stable fixed point of J^λ if and only if $\beta<\beta_c$, where

$$\beta_c=\beta_c(N)=\begin{cases}\dfrac{L^2\coth f(N)}{2Nf(N)} & \text{if} \quad N>N_1\\ \infty & \text{if} \quad N\leqslant N_1\end{cases} \tag{3.18}$$

with $N_1=(L/\lambda)(2\pi^2/qL^2)^{1/q}$ and

$$f(N)=\tfrac{1}{4}[2qL^2(\lambda N/L)^q-4\pi^2]^{1/2}$$

(b) Fix $N>N_1$. Then there exist δ, $\varepsilon>0$ such that:

(i) For $\beta_c+\delta>\beta\geqslant\beta_c$ the only solution of (2.21) in the ball $\|V-V_0\|_\infty<\varepsilon$ is $V=V_0$.

(ii) For $\beta_c>\beta>\beta_c-\delta$ the only solutions of (2.21) in the ball $\|V-V_0\|<\varepsilon$ are V_0 and a family of unstable solutions $\{V^{(\beta,y)}\mid y\in[0,L)\}$ differing by translations: $V^{(\beta,y)}(x)=V^{(\beta,0)}(x+y)$. Moreover,

$$\lim_{\beta\to\beta_c-}(\beta_c-\beta)^{1/2}\,\|V-V_0\|_\infty$$

exists and is nonzero.

The bifurcation described in (b) of the theorem, sometimes called a "pitchfork of revolution,"[5] is obtained by rotating the pitchfork of Fig. 1 around its axis; solid and dashed curves in this figure represent stable and unstable solutions, respectively, of (2.21). We remark that there are other bifurcations from $V=V_0$, at higher values of β, which we have not investigated in detail.

Proof. (a) V_0 is stable if the spectrum of $DJ^\lambda(V_0)$ is contained in the unit disk. To find this spectrum, we differentiate $[T-V^q-\kappa(V)]$ $R(\kappa(V),V)=I$, yielding

$$DR(\kappa(V),V)\,W=R(\kappa(V),V)\{qV^{q-1}W+[D\kappa(V)\,W]\}\,R(\kappa(V),V)$$

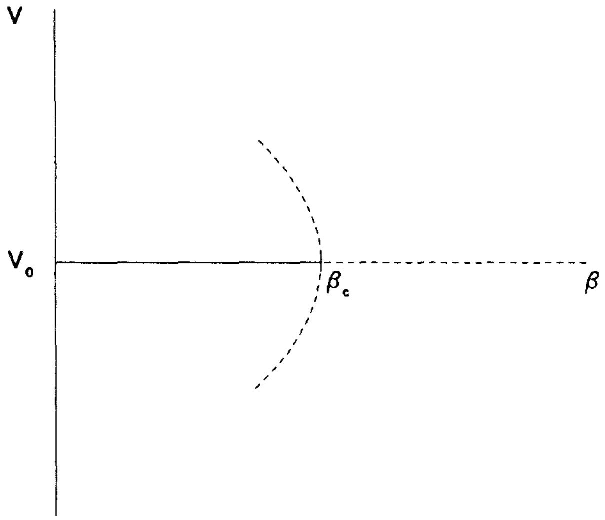

Fig. 1. Cross section of pitchfork bifurcation of the uniform solution of the mean field equation (2.21).

where $W \in X_0$, so that, from (2.16),

$$[DJ(V)\,W](x) = 2\lambda\beta^{-1} \int_0^L [r(\kappa(V), V)(x, y)]^2 \times \{qV(y)^{q-1}\,W(y) + [D\kappa(V)\,W]\}\,dy \qquad (3.19)$$

Now set $V = V_0$ in (3.19), note that $R(\kappa(V_0), V_0) = R(-\gamma^2)$, where γ is given by (3.16), and integrate over x. Using $DJ(V)(W) \in X_0$ and the fact that, from (3.12), $\int_0^L [r(\kappa(V_0), V_0)(x, y)]^2\,dx$ is independent of y and non-zero, we see that $D\kappa(V_0) = 0$, and hence DJ has kernel

$$DJ(V_0)(x, y) = 2q\lambda\beta^{-1}V_0^{q-1}[r(\kappa(V_0), V_0)(x, y)]^2$$

Explicit calculation from (3.12) shows that for, $k \neq 0$, $\psi_{\pm k}$ is an eigenvector of $DJ(V_0)$ with eigenvalue given by

$$\rho_k = \frac{qV_0^q}{2}\frac{1}{\gamma^2 + (k\pi/L)^2} \qquad (3.20)$$

and since the span of these ψ_k is dense in X_0, this exhausts the spectrum. The boundary of the stability region is determined by setting $\rho_1 = 1$; solving this equation for γ and substituting the result into (3.16) yields

$$\beta = \frac{L^2 \coth f(N)}{2Nf(N)}$$

Noting that ρ_k is increasing as a function of β and N leads to the stability region $\beta < \beta_c$.

(b) We study the bifurcation at the critical curve $\beta = \beta_c$, $N > N_1$, by the standard Liapunov–Schmidt reduction, adapted to respect the symmetries in the problem (we follow closely the procedure of ref. 5). We hold p, λ, and N fixed, and for convenience replace $J^{\lambda}_{\beta,N}$ by the map $\hat{J}_\beta$: $X_0 \to X_0$ given by

$$\hat{J}_\beta(W) = J^{\lambda}_{\beta,N}(V_0 + W) - V_0^q \tag{3.21}$$

$\hat{J}_\beta$ has 0 as a fixed point and satisfies $D\hat{J}(0) = DJ(V_0)$. From the above discussion we know that the eigenspace of $D\hat{J}(0)$ corresponding to the eigenvalue ρ_1 is the two-dimensional subspace Y_1 of X_0 spanned by $\cos(2\pi x/L)$ and $\sin(2\pi x/L)$; we let Y_2 be the complementary subspace spanned by the remaining (nonconstant) Fourier modes, let E_1, E_2 be the corresponding projections, and for $W \in X_0$ write $W_i \equiv E_i W$, $i = 1, 2$. Then (2.21) becomes

$$W_1 = E_1 \hat{J}_\beta(W) \tag{3.22a}$$

$$W_2 = E_2 \hat{J}_\beta(W) \tag{3.22b}$$

In a neighborhood of β_c the map $I - D\hat{J}(0)$ is invertible on Y_2 and hence (3.22b) may be solved for W_2 as $W_2 = H_\beta(W_1)$, so that (2.21) reduces to a fixed-point equation on the two-dimensional space Y_1:

$$W_1 = E_1 \hat{J}_\beta(W_1 + H_\beta(W_1)) \tag{3.23}$$

We may reduce the complexity of the problem further by utilizing the covariance of $\hat{J}_\beta$ under the group $O(2)$ of translations and reflections in the variable x; in particular, we write S_a: $X_0 \to X_0$ for the reflection given by $S_a W(x) = W(a - x)$. By a translation we may assume that the fixed point we seek has $W_1 = w \cos(2\pi x/L)$, i.e., W_1 invariant under S_L; then $H_\beta(W_1)$ must also be S_L invariant, since otherwise $S_L H_\beta(W_1)$ would be a second solution to (3.22b). Thus, we may replace each Y_i throughout the above argument by $Y_i^s = \{W \in Y_i \mid W = S_L W\}$, reducing (3.23) to a one-dimensional fixed-point problem

$$w = j_\beta(w) \tag{3.24}$$

and simplifying the calculation of H_β. [This is the problem we would have encountered by applying the Liapunov–Schmidt procedure to the original problem (3.21) posed for the space X_0^s of S_L-invariant potentials.] The set of solutions of (2.21) consists of all translates of solutions found by this procedure.

Equation (3.24) must have the form

$$(\beta-\beta_c)w+cw^3+\text{h.o.t.}=0 \tag{3.25}$$

since the reduced problem is still covariant under $S_{L/2}$, and this implies that there is no term in (3.25) proportional to w^2. A direct computation, sketched in Appendix B, shows that $c>0$. Standard theory then implies that (3.24) has a pitchfork bifurcation, with the trivial solution $w=0$ for all β and with two additional solutions $w_\pm \sim \pm[(\beta_c-\beta)/c]^{1/2}$ for $\beta<\beta_c$. Translates of these solutions form the solution set described in the theorem. ∎

3.3. Low-Temperature Solutions

We now turn to the low-temperature regime ($\beta=\infty$), where a complete analysis of the model is possible. Here we use $T=\beta^{-1}$ as a parameter, and hence investigate the map $K_N^\lambda\colon (0,\infty)\times X_N \to X_N$ defined by $K_N^\lambda(T,V)\equiv J_{T^{-1},N}^\lambda(V)$.

Theorem 3.7. K_N extends to a smooth map $K_N^\lambda\colon U\to X_N$, where $U\subset\mathbb{R}\times X_N$ is an open neighborhood of $[0,\infty)\times X_N$. Moreover:

(i) Each solution $V(x)$ of the fixed-point problem

$$V=K_N^\lambda(0,V) \tag{3.26}$$

is of the form $V(x)=\lambda\,|\phi(x)|^2$, where $\phi(x)$ is a nonvanishing solution of the time-independent NLSE: for some $\mu\in\mathbb{R}$,

$$-\phi''(x)-\lambda^q\,|\phi(x)|^{p-2}\,\phi(x)-\mu\phi(x)=0 \tag{3.27}$$

(ii) For each solution $V(x)$ of the zero-temperature problem there is a smooth curve of solutions $V(T,x)$ of (3.26) with $V(0,x)=V(x)$.

Proof. We calculate $K_N(T,V)$ for $V\in X_N$ and $T\approx 0$. Then $\kappa(V)$ is determined from (3.3):

$$\sum_k [\omega_k(V^q)-\kappa(V)]^{-1}=N/2T$$

with $\kappa(V)<\omega_0(V^q)$. For small T this yields

$$\kappa(V)=\omega_0(V^q)-(2T/N)\,h(T)$$

with $h(T)$ analytic at $T=0$ and $h(0)=1$. Then, from (3.1),

$$K(T,V)(x)=\lambda N\,|\psi_0(V^q)(x)|^2+T\hat{K}(T,V)(x)$$

with $\hat{K}$ analytic in T at $T=0$. This proves the first statement of the theorem; in particular,

$$K(0, V)=\lambda N\,|\psi_0(V^q)|^2 \tag{3.28}$$

Now if V satisfies (3.26), we define $\phi(x)=N^{1/2}\psi_0(V^q)(x)$; then $V=\lambda\,|\phi|^2$ and (3.27) is simply the eigenvalue equation for $\psi_0(V)$. This verifies (i). Part (ii) will follow from the implicit function theorem once we verify that $D_V K(0, V)$ does not have 1 as an eigenvalue. But $D_V K(0, V)$ may be calculated from (3.28) and standard first-order perturbation theory via

$$D_V\psi_0(V^q)(W)=\sum_{k\neq 0}\frac{(\psi_k(V^q),\, W\psi_0(V^q))}{\omega_0(V^q)-\omega_k(V^q)}\,\psi_k(V^q)$$

which leads to

$$D_V K(0, V)(x, y)=\sum_{k\neq 0}\left[\frac{\bar{\chi}_k(x)\,\chi_k(y)}{\lambda_0(V)-\lambda_k(V)}+\text{c.c.}\right]$$

where $\chi_k=\psi_k(V^q)\,\psi_0(V^q)$. Thus $D_V K(0, V)$ is negative semidefinite. ∎

We remark that, for $\phi(x)$ real, (3.27) is the equation of a classical particle moving in a potential

$$U(\phi)=\frac{\lambda^q}{p}\phi^p+\frac{\mu}{2}\phi^2$$

Nonvanishing, nonconstant real solutions of (3.27) on $[0, L]$ arise for $\mu<0$ and correspond to oscillations of this classical particle, with negative energy, in one well of the resulting double-well potential. The oscillations must have period L/m for some integer $m\geqslant 1$, leading to solutions $\phi(x)$ with m maxima in the interval $[0, L]$. The single-peaked solution $\phi_0(x)$ (corresponding to $m=1$) varies with N to form a continuous family of "solitons" of different peak heights and widths; the threshold (almost uniform) soliton has L^2 norm N_1 (see Theorem 3.6).

3.4. Transition Region

To interpolate between the low- and high-temperature regimes we have investigated stable fixed points of (2.17) numerically, by iteratively applying $J^{\lambda}_{\beta,N}$ to various initial potentials. We will discuss here only the case $p=4$.

At very low temperature there is only one stable fixed point (up to translation), whose nature depends on the value of N relative to N_1. For

$N < N_1$ it is the uniform solution V_0; this is consistent with the local stability result of Theorem 3.6 and with the fact, noted above, that N_1 is the threshold for the existence of nonuniform solutions of (3.27). For $N > N_1$ the stable fixed point corresponds, via $V = \lambda |\phi_0|^2$ as in Theorem 3.7, to the single-peaked soliton solution ϕ_0 of (3.27).

To describe the situation at higher temperatures we imagine that N is kept fixed, satisfing $N > N_1$, and that β is decreased from ∞. As this is done we continue to observe a unique stable solution, which is again solitonlike and which varies continuously with β, the height decreasing as β is decreased. At β_c, the uniform solution becomes stable as well; at a lower value $\hat{\beta}_c$ the solitonlike solution, still with a positive height, abruptly vanishes. Below $\hat{\beta}_c$ only the uniform solution is stable. This behavior is consistent with the nature of the bifurcation at β_c as described in Theorem 3.6; apparently, the unstable paraboloid of nonuniform solutions described there, opening toward small values of β, curves back at $\hat{\beta}_c$ to form the stable branch of solutions which persist to $\beta = \infty$. A qualitative picture of the solution set is obtained by rotating the curve of Fig. 2 around the $V = V_0$ axis.

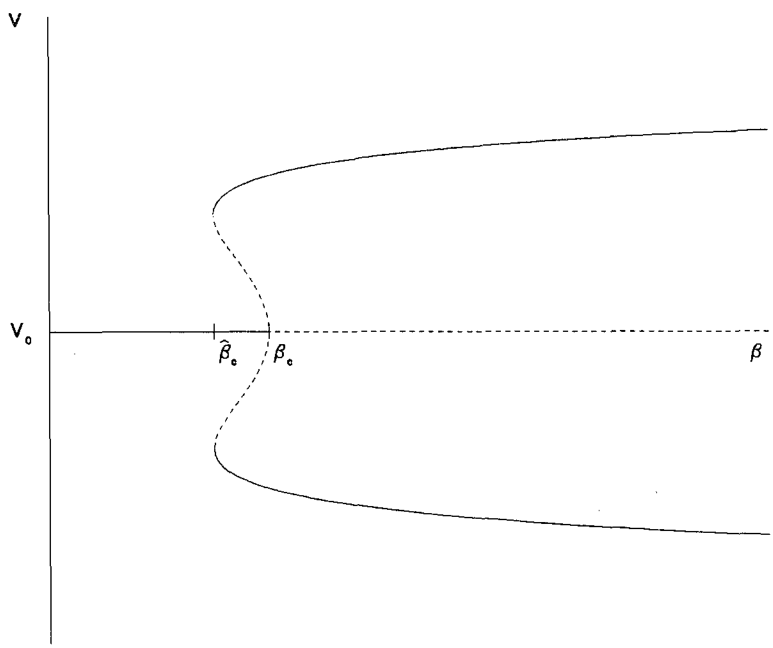

Fig. 2. Qualitative global picture of solutions of the mean field equation (2.21). The uniform solution loses global stability at $\hat{\beta}_c$.

In the region $\hat{\beta}_c < \beta < \beta_c$ there are two locally stable solutions of (2.21); we may compute their free energies to see which is thermodynamically preferred. The exact answer depends, but not very sensitively, on the value of λ and on which free energy we use (see discussion at the end of Section 2). For $\lambda = 1$ and with the true free energy $\hat{F}^{\lambda_1}$, for example, we find that the nonuniform solution is preferred whenever it exists; the model thus has a first-order phase transition at $\hat{\beta}_c$. (Use of the generalized free energy $\hat{F}^{\lambda}$ produces a first-order transition at a value of β between β_c and $\hat{\beta}_c$.)

Finally, we compare the predictions of the mean field model with the results of Monte Carlo studies of the canonical ensemble, described in I. Specifically, for certain observables $Q(\phi)$, we will compare the mean field expectation value $\langle Q \rangle$ with the corresponding expectation $\langle Q \rangle_c$ in the canonical ensemble, calculated as a time average in the Monte Carlo simulation.

Remark 3.8. Two obvious differences in the ensembles—in translation invariance and in low-temperature fluctuations—make it clear that $\langle Q \rangle_c$ will be well approximated by $\langle Q \rangle$ at best only for appropriately chosen Q. We illustrate the point with two examples:

(a) It would seem natural to compare $\langle |\phi(x)|^2 \rangle_c$ with $\langle |\phi(x)|^2 \rangle \equiv \lambda^{-1} V(x)$. Unfortunately, translation invariance of the canonical measure implies that the former quantity is identically equal to N, while the mean-field ensemble breaks translation invariance at low temperature.

(b) In I we took $\langle M_2(\phi) \rangle_c$ as a measure of the nonuniformity of typical field configurations in the canonical ensemble, where

$$M_r = N^{-r} \int_0^L |\phi(x)|^{2r}\, dx$$

How do $\langle M_r \rangle_c$ and $\langle M_r \rangle$ compare? At low temperatures the typical field configuration in the canonical ensemble will be, up to small fluctuations, the ground state, that is, the unique (except for translation and overall change of phase) single-peaked soliton solution $\phi_0(x)$ of the classical equation (3.27), so that (again at low temperatures)

$$\langle M_r \rangle_c \approx N^{-r} \|\phi_0\|_{2r}^{2r} = N^{-r} \| V_\infty^1 \|_r^r \tag{3.29}$$

(Here we indicate the parameter dependence of the mean field potential by writing V_β^λ.) In the mean field ensemble, on the other hand, the field $\phi(x)$ is Gaussian, with mean 0 and variance $\lambda^{-1} V(x)$, so that fluctuations persist at all temperatures. In particular, (2.8) implies

$$\langle M_r \rangle \approx \Gamma(r+1)(N\lambda)^{-r} \| V_\infty^\lambda \|_r^r \tag{3.30}$$

For fixed λ, (3.29) and (3.30) can agree for at most one value of r. The natural choice $\lambda = 1$ leads to the (trivial) agreement at $r = 1$.

In I we proposed to avoid the problem of translation invariance and to extract a typical field profile by computing the quantity $F(x) = \langle |\tilde{\phi}(x)|^2 \rangle_c$, where $\tilde{\phi}$ is the translate of ϕ for which the maximum magnitude occurs at a fixed point, say $L/2$. Although we cannot compute the corresponding expectation in the mean field model exactly, we can compare $F(x)$ directly with $\lambda^{-1} V(x)$. (As compared with the computation of $\langle M_2 \rangle$ and $\langle M_2 \rangle_c$ as described above we are here taking the expectation before squaring the field amplitude, thus avoiding the disparity due to low-temperature fluctuations.) It is easy to see, as above, that these quantities will agree well at low temperatures with the choice $\lambda = 1$, which we make from now on. Comparisons at various temperatures are presented in Figs. 3–5. In Fig. 3 we plot the quantities $N^{-2} \|F\|_2^2$ and $N^{-2} \|V\|_2^2$ for $N = 40$ and $N = 80$ at a range of temperatures. The coincidence at low temperature is built into our choice of observables, but the good agreement

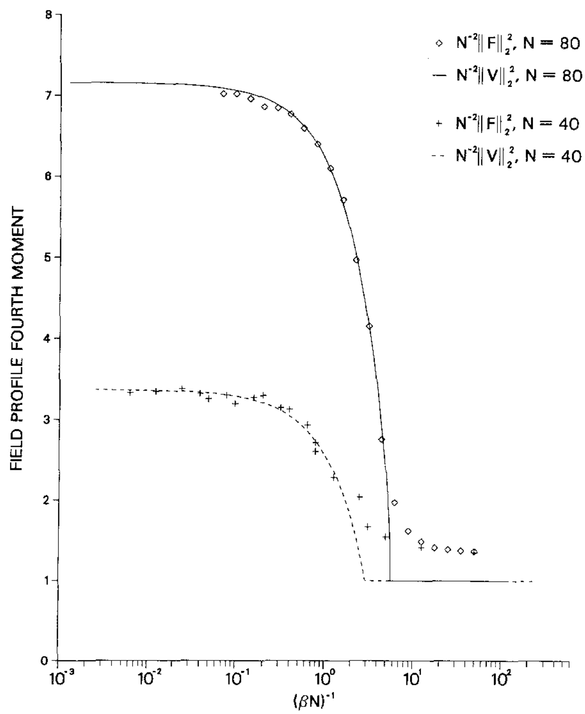

Fig. 3. Comparison of the fourth moments of the field as obtained from Monte Carlo simulations ($\|F\|_2^2$) with the mean field prediction ($\|V\|_2^2$).

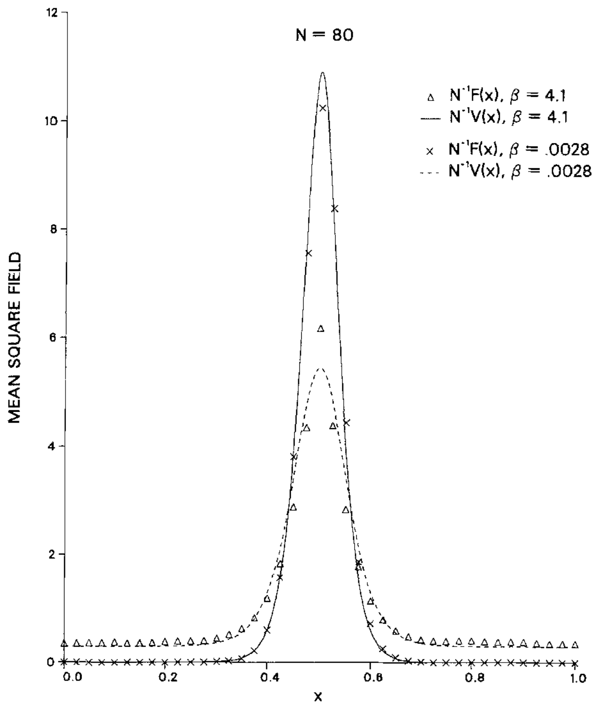

Fig. 4. Comparison of the square magnitude of the field $F(x)$ from Monte Carlo simulations with the mean field prediction $V(x)$, at $N = 80$.

near the phase transition in the mean field model indicates that our approximation captures the properties of the true model surprisingly well. The disagreement at high temperatures is an artifact of our choice of the quantity $F(x)$; the fluctuations present at high temperature and the centering involved in the definition of F produce a profile for F with a narrow spike at $L/2$, increasing the L^2 norm relative to the mean N. In Figs. 4 and 5 we compare, for $N = 80$ and $N = 40$ respectively, normalized profiles $N^{-1}F$ and $N^{-1}V$ at low temperature and at a temperature just below the mean field phase transition; again, the good agreement in the latter case is a signal of the quality of our approximation. The peak enhancement due to centering, discussed above, can be seen in the elevation of $F(L/2)$ in the higher-temperature data in these figures.

4. DYNAMICAL CONSIDERATIONS

In this section we define an approximate dynamics for the correlation functions of fields of our theory. This dynamics has as stationary solution

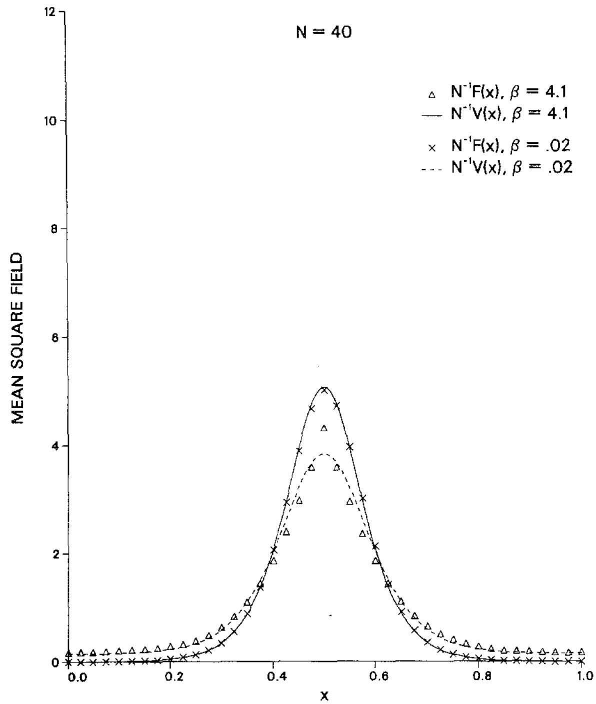

Fig. 5. Comparison of the square magnitude of the field $F(x)$ from Monte Carlo simulations with the mean field prediction $V(x)$, at $N=40$.

the mean field approximation for the Gibbs measures developed in earlier sections. Using this dynamics enables us to compute approximate values for the correlation functions in the mean-field ensembles at unequal times. It also permits us to investigate the dynamical stability of the mean field ensemble with uniform field, finding a stability region which coincides with the region of thermodynamic stability found in Theorem 3.6.

As in Section 1, we let $u(x, t)$ denote a time-dependent field, i.e., a solution of the NLSE (1.2), and we write $G(x_1, t_1, x_2, t_2)$ for the product $u(x_1, t_1)\, u^*(x_2, t_2)$. By (1.2), G satisfies

$$i\partial_{t_1} G(x_1, t_1, x_2, t_2) = -\partial_{x_1}^2 G(x_1, t_1, x_2, t_2) - |G(x_1, t_1, x_1, t_1)|^{p-2}\, G(x_1, t_1, x_2, t_2) \quad (4.1)$$

as well as a similar equation for $\partial_{t_2} G$. Let us suppose now that at time $t=0$ the fields are distributed according to some ensemble ρ; we want then to study the evolution of the two-point correlation function $C(x_1, t_1, x_2, t_2) \equiv \langle G(x_1, t_1, x_2, t_2)\rangle_\rho$. Taking the expectation of (4.1), we obtain an equation for C, which, however, depends on the higher moments of the field. We close this equation by making the approximation

$$\langle |G(x_1, t_1, x_1, t_1)|^{p-2} G(x_1, t_1, x_2, t_2)\rangle_\rho \approx \lambda^q \langle G(x_1, t_1, x_1, t_1)\rangle_\rho^{p-2} \langle G(x_1, t_1, x_2, t_2)\rangle_\rho \tag{4.2}$$

which leads to

$$i\partial_{t_1} C(x_1, t_1, x_2, t_2) = -\partial_{x_1}^2 C(x_1, t_1, x_2, t_2) - \lambda^q C^q(x_1, t_1, x_1, t_1)\, C(x_1, t_1, x_2, t_2) \tag{4.3}$$

The product approximation (4.2) may be regarded as a dynamical mean field approximation. As in earlier sections, we regard λ as a free parameter; note that, from (2.8), (4.2) becomes exact for $\lambda = \lambda_1$ if we take $x_1 = x_2$ and $t_1 = t_2$, and assume that the ensemble at time t_1 is Gaussian.

To derive the results of our previous approach from this dynamical model, we obtain an evolution equation for the equal time covariance $\tilde{C}(x_1, x_2; t) \equiv C(x_1, t, x_2, t)$ by adding to (4.3) the corresponding equation for $i\partial_{t_2} C$, yielding

$$i\partial_t \tilde{C} = I(\tilde{C}) \tag{4.4}$$

with

$$I(\tilde{C})(x_1, x_2) = -(\partial_{x_1}^2 - \partial_{x_2}^2)\, \tilde{C}(x_1, x_2) - \lambda^q [\tilde{C}^q(x_1, x_1) - \tilde{C}^q(x_2, x_2)]\, \tilde{C}(x_1, x_2)$$

Equation (4.4) may be regarded as defining an evolution for Gaussian ensembles, a Gaussian measure being determined by its covariance. [We note that the linearized equation of motion obtained from (4.4) is identical in form to the random phase approximation.[7]] It is now easy to verify that the measure

$$d\rho_V(\phi) = Z_V^{-1} \exp\{\beta[(\phi, V^q \phi) + \kappa(V)\, \|\phi\|_2^2]\}\, d\mu_\beta(\phi)$$

is invariant under this dynamics; that is, the covariance $\tilde{C}_V(x_1, x_2; t) = \langle \phi(x_1)\, \phi^*(x_2)\rangle_{\rho_V}$ satisfies $I(\tilde{C}_V) = 0$, whenever V is a solution of the mean field fixed-point equation (2.21), which here becomes

$$V(x) = \lambda \tilde{C}_V(x, x)$$

We may thus think of the mean field measure ρ_V as an equilibrium state for our approximate dynamics. Other invariant states exist, however; in particular, any translation-invariant covariance $\tilde{C}(x_1, x_2) = \hat{C}(x_1 - x_2)$ is stationary (including $\tilde{C}_{V_0}$, the covariance corresponding to the uniform mean field).

With this in mind it is natural to compute the unequal time correlation function in a stationary mean field measure ρ_V,

$$C'_V(x_1, x_2, t) = \langle G(x_1, t + t_2, x_2, t_2) \rangle_{\rho V}$$

from (4.3) by replacing the equal time covariance $C^q(x_1, t_1, x_1, t_1)$ occurring there with its stationary value $\lambda^{-1} V(x_1)$. Using the notation introduced in the introduction to Section 3 for the eigenvalues and eigenfunctions of $T - V^q$, we can readily solve (4.3) to give

$$C'_V(x_1, x_2, t) = 2\beta^{-1} \sum_{k=-\infty}^{\infty} \frac{\bar{\psi}_k(V^q)(x)\, \psi_k(V^q)(y) \exp[i\omega_k(V^q) t]}{\omega_k(V^q) - \mu} \tag{4.5}$$

We remark that by subtracting from (4.3) the corresponding equation for $i\partial_{t_2} C$ we obtain a more symmetric evolution equation for C'_V:

$$\begin{aligned} i\partial_t C'_V(x_1, x_2, t) = &-\tfrac{1}{2}(\partial^2_{X_1} + \partial^2_{X_1})\, C'_V(x_1, x_2, t) \\ &- \tfrac{1}{2}[V^q(x_1) + V^q(x_2)]\, C'_V(x_1, x_2, t) \end{aligned}$$

We now turn to the question of the dynamical stability of the mean field equilibrium ensembles. Suppose again that V is a solution of (2.21) with corresponding Gaussian ensemble ρ_V; to investigate the stability of ρ_V under our dynamics we must calculate the spectrum of the linearization $DI(\tilde{C}_V)$ of I at $\tilde{C}_V$. We carry out this computation for the covariance $\tilde{C}_0$ corresponding to the mean field $V = V_0$. As we will see, the spectrum is invariant under complex conjugation, so that stability will correspond to the existence of the real spectrum only.

Because $\tilde{C}_0(x, x) = \lambda^{-1} V_0$ is independent of x, we have

$$\begin{aligned} [DI(\tilde{C}_0) A](x_1, x_2) = &-(\partial^2_{x_1} - \partial^2_{x_2})\, A(x_1, x_2) \\ &- q\lambda V_0^{q-1} [A(x_1, x_1) - A(x_2, x_2)]\, \tilde{C}_0(x_1, x_2) \end{aligned} \tag{4.6}$$

Since A is L-periodic in x_1 and x_2, it may be written as $A_1(x, z) + A_2(x, z)$, where $x = (x_1 + x_2)/2$, $z = x_1 - x_2$, A_1 is $L/2$-periodic in x and L-periodic in z, and A_2 is $L/2$-antiperiodic in x and L-antiperiodic in z. Correspondingly, we seek eigenfunctions of $DI(\tilde{C}_0)$ of the form

$$A(x, z) = e^{2\pi i k x/L} H(z) \tag{4.7}$$

where either k is even and H is L-periodic or k is odd and H is L-antiperiodic. Note that for $k = 0$ any function (4.7) satisfies $DI(\tilde{C}_0) A = 0$, so that all eigenvalues in this case are zero; we thus assume $k \neq 0$ in what follows. Of course, a perturbation $hC = \hat{C}_0 + A$, for A of the form (4.7) with $k = 0$, would preserve the normalization $\int_0^L C(x, x) = N$ only if $H(0) = 0$.

From (2.6) and (3.12) we have

$$\tilde{C}_0(x_1, x_2) \equiv \hat{C}(z) = \frac{\cosh \gamma(L/2 - z)}{\beta\gamma \sinh \gamma L/2} \tag{4.8}$$

where $0 \leqslant z \leqslant L$ and γ is given by (3.16) (note that $\hat{C}$ is extended by periodicity to other values of z and is not differentiable at $z = 0$ and $z = L$). Substitution of (4.7) and (4.8) into (4.6) leads to an equation for the eigenvalue η:

$$-(4\pi ik/L)\, H'(z) - 2i \sin(\pi kz/L)\, q\lambda V_0^{q-1} \hat{C}(z)\, H(0) = \eta H(z) \tag{4.9}$$

Setting $\hat{\eta} = (4\pi k)^{-1} L\eta$, we may solve (4.9) to obtain

$$H(z) = H(0) \left[e^{i\hat{\eta}z} - \frac{Lq\lambda V_0^{q-1}}{2\pi k} \int_0^z e^{i\hat{\eta}(z-\zeta)/L} \sin\left(\frac{\pi k\zeta}{L}\right) \hat{C}(\zeta)\, d\zeta \right] \tag{4.10}$$

Imposing periodicity $H(L) = (-1)^k H(0)$ then leads to the eigenvalue conditions

$$\begin{aligned} &q\lambda V_0^{q-1} D^{-1} [(\gamma^2 + (\pi k/L)^2 - \hat{\eta}^2) \cosh(\gamma L/2) \cos(\hat{\eta}L/2) \\ &\quad + 2\gamma\hat{\eta} \sinh(\gamma L/2) \sin(\hat{\eta}L/2)] - 2\gamma\beta \sinh(\gamma L/2) \cos(\hat{\eta}L/2) = 0 \end{aligned} \tag{4.11a}$$

for k odd, and

$$\begin{aligned} &q\lambda V_0^{q-1} D^{-1} [(\gamma^2 + (\pi k/L)^2 - \hat{\eta}^2) \cosh(\gamma L/2) \sin(\hat{\eta}L/2) \\ &\quad - 2\gamma\hat{\eta} \sinh(\gamma L/2) \cos(\hat{\eta}L/2)] - 2\gamma\beta \sinh(\gamma L/2) \sin(\hat{\eta}L/2) = 0 \end{aligned} \tag{4.11b}$$

for k even. Here

$$D = D(\hat{\eta}) = [(\gamma^2 + (\pi k/L)^2 - \hat{\eta}^2)^2 + 4\gamma^2\hat{\eta}^2]$$

It is convenient to rewrite (4.11) in the form

$$P_k(\hat{\eta}) = \begin{cases} \hat{\eta} \tan(\hat{\eta}L/2) & \text{if } k \text{ is odd} \\ -\hat{\eta} \cot(\hat{\eta}L/2) & \text{if } k \text{ is even} \end{cases} \tag{4.12}$$

where $P_k(\hat{\eta})$ is the fourth-order polynomial defined by

$$P_k(\hat{\eta}) = \beta D(\hat{\eta})(q\lambda V_0^{q-1})^{-1} - (2\gamma)^{-1}\, [\gamma^2 + (\pi k/L)^2 - \hat{\eta}^2] \coth(\gamma L/2)$$

We should be aware, however, that in the transition from (4.11) to (4.12) we have introduced four spurious roots [the solutions of $D(\hat{\eta})=0$] and have lost the root $\hat{\eta}=0$ in the case k even. We may find the real or purely imaginary solutions of (4.12) simply by graphing both sides as functions of $\hat{\eta}$ or $i\hat{\eta}$, respectively; we describe the results of this analysis separately for odd and even k.

When k is odd, (4.12) has real roots $\hat{\eta}_j$ and $\hat{\eta}_{-j}=-\hat{\eta}_j$, $j=2, 3,...,$ which satisfy $(2j-3)\pi/L<\hat{\eta}_j<(2j-1)\pi/L$. For $P_k(0)>0$ there are two additional real roots $\hat{\eta}_1$ and $\hat{\eta}_{-1}=-\hat{\eta}_1$ with $0<\hat{\eta}_1<\pi/L$; these become pure imaginary for $P_k(0)<0$. Since from (3.16), $\gamma^{-1}\coth(\gamma L/2)=\beta N/L$ and hence

$$P_k(0)=\beta(q\lambda V_0^{q-1})^{-1}\,[\gamma^2+(\pi k/L)^2][\gamma^2+(\pi k/L)^2-qV_0^q/2] \quad (4.13)$$

the stability region for the modes with wave number k is given by

$$\gamma^2+(\pi k/L)^2>qV_0^q/2 \quad (4.14)$$

This is in precise agreement with the condition $\rho_k>1$ [see (3.20)] derived in the proof of Theorem 3.6 for the local thermodynamic stability of the uniform mean field ensemble against perturbations of wave number k. From (4.14) it is clear that the stability regions increase with k and hence that the $k=1$ condition provides the actual stability boundary. We remark that, in the limit $L\to\infty$, (4.14) coincides with the stability condition for the so-called broadband modulational instability of plasma physics.[8]

When k is even, a similar analysis of (4.12) finds roots $\hat{\eta}_j$ for $j\pm 1, \pm 2,...$ and gives $P_k(0)>-1$ as the condition for reality of $\hat{\eta}_{\pm 1}$ and hence for dynamic stability against perturbations of wave number k. (Recall that $\hat{\eta}_0=0$ is also an eigenvalue.) This condition is weaker than the condition $\rho_k>1$ [equivalently, $P_k(0)>0$] for thermodynamic stability against these perturbations; we do not have an explanation for this difference. However, since for each even wave number k the region of thermodynamic stability is contained in that of dynamic stability, and in turn contains the $k=1$ stability region, it is again (4.14) for $k=1$ which gives the true region of dynamical stability. Thus we have the following result.

Theorem 4.1. The covariance $\tilde{C}_{V_0}$ is linearly stable under the dynamics (4.4) if $\beta<\beta_c$, where β_c is as defined in Theorem 3.6.

Proof. To complete the proof, we sketch the verification that the computation above has determined the complete spectrum of $DI(\tilde{C}_{V_0})$. If we regard this operator as acting on $L^2([0, L]\times[0, L])$, then the subspace of functions of the form (4.7) is invariant, and it suffices to determine the spectrum of the operator I_k [essentially the restriction of $DI(\tilde{C}_{V_0})$ to the subspace],

$$(I_k H)(z) = -(4\pi ik/L)\, H'(z) - F(z)\, H(0) \tag{4.15}$$

which occurs in (4.9). But the second term in (4.15) is a relatively degenerate (actually rank-one) perturbation of the first, so that the spectrum of I_k is completely determined by a Weinstein–Aronszajn (WA) determinant.[6] In fact, this WA determinant is given by the left-hand side of (4.11a) divided by $\cos(\hat{\eta}L/2)$ for k odd, and by the left-hand side of (4.11b) divided by $\sin(\hat{\eta}L/2)$ for k even. The conclusion then follows immediately from the treatment in ref. 6.

APPENDIX A

Our goal is to prove an estimate, valid for any positive potential $V \in X$, that is needed in the proof of Theorem 3.1:

$$-\beta^{-1} \log \tilde{Z}_V \leqslant B(1 + Q^{(p-2)/(p+2)}) \tag{A.1}$$

Here $Q = \| V \|_{p/2}^{p/2}$; we will write κ for $\kappa(V)$ and $\tilde{\omega}_k$ for $\tilde{\omega}_k(V^q)$. Now in finite dimension we could use the identity

$$\int_{\mathbb{R}^n} \exp[-(z, Az)/2]\, d^n z = (2\pi)^{n/2} (\det A)^{-1/2}$$

to write the partition function as the ratio of determinants of T and $(T - V^q - \kappa)$; the corresponding equation in infinite dimension is

$$-\log \tilde{Z}_V = \log(2\pi\beta) + \log(\tilde{\omega} - \kappa) + \sum_{k \neq 0} \log \frac{\tilde{\omega}_k - \kappa}{\omega_k} \tag{A.2}$$

(the zero modes get special treatment because $\omega_0 = 0$). In fact, however, (A.2) is rigorously correct; see, e.g., ref. 4, Theorem 3.11.

We next obtain bounds on the eigenvalues $\tilde{\omega}_k$ by variational calculations. The min-max theorem, together with the positivity of V^q, leads immediately to

$$\tilde{\omega}_k \leqslant \omega_k \tag{A.3}$$

To derive lower bounds we will need the form of the interpolation inequality (3.9) that is valid on the interval $[0, L]$: for any $\varepsilon > 0$ there is a constant $K_\varepsilon > 0$ such that

$$\| f \|_p^p \leqslant (C_p + \varepsilon)\, \| f' \|_2^{(p-2)/2}\, \| f \|_2^{(p+2)/2} + K_\varepsilon\, \| f \|_2^p \tag{A.4}$$

Then for any $\phi \in L^2$ we have, by Hölder's inequality and (A.4),

$$\begin{aligned}(\phi, V^q\phi) &\leqslant Q^{(p-2)/p} \|\phi\|_p^2 \\ &\leqslant Q^{(p-2)/p}[(C_p+\varepsilon)^{2/p} \|\phi'\|_2^{(p-2)/p} \|\phi\|_2^{(p+2)/p} + K_\varepsilon^{2/p} \|\phi\|_2^2]\end{aligned}$$

and hence, if $\|\phi\|_2 = 1$,

$$(\phi, (T-V^q)\phi) \geqslant (\phi, T\phi) - d(\phi, T\phi)^\gamma - e \tag{A.5}$$

with $\gamma = q/p$, $d = (C_p+\varepsilon)^{2/p} Q^{(p-2)/p}$, and $e = K_\varepsilon^{2/p} Q^{(p-2)/p}$. We extract two consequences from (A.5). First, minimizing the right-hand side over possible values of $(\phi, T\phi)$ leads to

$$\tilde{\omega}_0 \geqslant -(1-\gamma)\,\gamma^{\gamma/(1-\gamma)} d^{1/(1-\gamma)} + e \geqslant -C(1+Q^{2(p-2)/(p+2)}) \tag{A.6}$$

(Here and below C represents a generic constant.) Second, let k_0 be the integer part of $(L/2\pi)\{[2(d+e)]^{p/(p+2)}+1\}$ and take $|k| > k_0$, so that $\omega_k \geqslant 1$ and $\omega_k \geqslant [2(d+e)]^{1/(1-\gamma)}$. Then, since any subspace $\mathscr{M} \subset L^2$ of dimension $2k$ (for k positive) or $2\,|k|+1$ (for k negative) contains a vector $\phi_{\mathscr{M}}$ of norm one with $(\phi_{\mathscr{M}}, T\phi_{\mathscr{M}}) \geqslant \omega_k$,

$$\begin{aligned}\tilde{\omega}_k &= \inf_{\mathscr{M}} \sup_{\substack{\phi \in \mathscr{M} \\ (\phi,\phi)=1}} (\phi, (T-V^q)\phi) \\ &\geqslant \omega_k - d\omega_k^\gamma - e \\ &\geqslant \omega_k[1-(d+e)\,\omega_k^{\gamma-1}] \\ &\geqslant \omega_k/2\end{aligned} \tag{A.7}$$

Now, from (3.3) and (A.7),

$$\begin{aligned}\frac{\beta N}{2} &\leqslant \frac{2k_0+1}{\tilde{\omega}_0 - \kappa} + \sum_{|k|>k_0} \frac{1}{\tilde{\omega}_k - \kappa} \\ &\leqslant \frac{2k_0+1}{\tilde{\omega}_0-\kappa} + C\,|\kappa|^{-1/2}\end{aligned}$$

which implies (e.g., by separate consideration of $\kappa \geqslant 1$ and $\kappa \leqslant 1$)

$$\tilde{\omega}_0 - \kappa \leqslant C(k_0+1) \leqslant C(1+Q^{(p-2)/(p+2)}) \tag{A.8}$$

From (A.6) and (A.8) we have

$$|\kappa| \leqslant C(1+Q^{2(p-2)/(p+2)})$$

Thus, from (A.2) and (A.3),

$$-\beta^{-1}\log \tilde{Z}_V \leqslant C\left\{1+\log Q+\sum_{k>0}\log[1+Ck^{-2}(1+Q^{2(p-2)/(p+2)})]\right\}$$

Equation (3.7) now follows from

$$\begin{aligned}
&\sum_{k\geqslant 1}\log(1+Ak^{-2})\\
&\qquad=\left(\sum_{k^2\leqslant A}+\sum_{k^2>A}\right)\log(1+Ak^{-2})\\
&\qquad\leqslant \hat{k}\log A-2\log\hat{k}!+\sum_{1}^{\hat{k}}\log(1+A^{-1}k^2)+\sum_{\hat{k}+1}^{\infty}\log(1+Ak^{-2})\\
&\qquad\leqslant CA^{1/2}
\end{aligned}$$

where $\hat{k}$ denotes the integer part of $A^{1/2}$, and in the last step we have used $\log(1+x)\leqslant x$ and Stirling's formula.

APPENDIX B

In this Appendix we compute the constant c appearing in (3.25) and show that it is positive. Without the introduction of further notation we regard $\hat{J}$ as a map on X_0^s, the space of even potentials. If $g(w,\beta)\equiv j_\beta(w)-w$, then (3.24) becomes

$$0=g(w,\beta)=g_{\beta w}(0,\beta_c)(\beta-\beta_c)w+(1/6)\,g_{www}(0,\beta_c)w^3+\text{h.o.t.}$$

and we will show that $g_{\beta\omega}$ and g_{www} are both positive at $(0,\beta_c)$. We introduce the notation $D^k\hat{J}_\beta(W)(W_1,\ldots,W_k)$ for the kth derivative of $\hat{J}_\beta$ at W as a k-linear form on X_0^s, and define $\phi_k(x)=\cos(2\pi kx/L)$ and $M=D\hat{J}_{\beta_c}(0)-I$. For convenience we will make use of formulas for the partial derivatives of g tabulated by Golubitsky and Schaeffer [ref. 5, (1.3.23)]; with $\Phi(W,\beta)=\hat{J}_\beta(W)-W$, $v_0=\phi_1$, and $v_0^*=2L^{-1}\phi_1$ these become

$$g_{\beta w}(0,\beta_c)=2L^{-1}(\phi_1,\partial_\beta D\hat{J}\phi_1) \tag{B.1a}$$

$$g_{www}(0,\beta_c)=2L^{-1}(\phi_1,D^3\hat{J}(\phi_1,\phi_1,\phi_1)-3D^2\hat{J}(\phi_1,M^{-1}E_2^sD^2\hat{J}(\phi_1,\phi_1))) \tag{B.1b}$$

Here and in the balance of this Appendix, unless specified otherwise, all derivatives of $\hat{J}$ are evaluated at $\beta=\beta_c$, $W=0$. We remark that (B.1a) is already simplified from the form in ref. 5, using $\hat{J}_\beta(0)\equiv 0$.

Next, recall that

$$\hat{J}_\beta(W) = 2\lambda\beta^{-1}\mathscr{S}[T - (V_0 + W)^q - \kappa(V_0 + W)]^{-1}$$

where $\mathscr{S}$, acting on an operator, yields the restriction of its kernel to the diagonal. The derivatives of $\hat{J}$ (we record only the diagonal elements for simplicity) are then

$$D\hat{J}(W) = 2\lambda\beta^{-1}\mathscr{S}[RA_1R]$$
$$D^2\hat{J}(W, W) = 2\lambda\beta^{-1}\mathscr{S}[RA_2R + 2R(A_1R)^2]$$
$$D^3\hat{J}(W, W, W) = 2\lambda\beta^{-1}\mathscr{S}[RA_3R + 3RA_1RA_2R + 3RA_2RA_1R + 6R(A_1R)^3]$$

with

$$R = R(-\gamma^2) = [T - V_0 - \kappa(V_0)]^{-1}$$
$$A_n = n!\, b_n W^n + D^n\kappa(W, \ldots, W), \qquad b_n = \binom{q}{n} V_0^{q-n}$$

Now in fact we calculated $D\hat{J}$ in the proof of part (a) of Theorem 3.6, finding

$$D\hat{J}_\beta(0)\,\phi_k = \rho_k(\beta)\,\phi_k \tag{B.2}$$

[see (3.20)]. From (B.1a), then,

$$g_{\beta w}(0, \beta_c) = (\partial\rho_1/\partial\beta)_{\beta=\beta_c} > 0$$

It remains to check the sign of $g_{www}(\beta_c, 0)$. Note that, from (B.2), M is given by $M\phi_k = [\rho_k(\beta_c) - 1]\,\phi_k$.

Higher derivatives of $\hat{J}$ could be evaluated by the method used to derive (B.2), but calculations with the explicit kernel (3.12) are quite complicated. Instead, we return to the formula (3.1):

$$r(x, y) = \sum_{k=-\infty}^{\infty} \frac{\bar{\psi}_k(x)\,\psi_k(y)}{\omega_k + \gamma^2} \tag{B.3}$$

by means of which the various terms in (B.1b) are readily computed. From (B.3) it follows by induction that

$$[R\phi_{j_1}R \cdots R\phi_{j_n}R](x, y)$$
$$= 2^{-n} \sum_{\sigma_1, \ldots, \sigma_n = \pm 1} \sum_{k=-\infty}^{\infty} \frac{\bar{\psi}_k(x)\,\psi_{k+\sum_1^n \sigma_i j_i}(y)}{(\lambda_k + \nu)(\lambda_{k+\sigma_1 j_1} + \nu)\cdots(\lambda_{k+\sum_1^n \sigma_i j_i} + \nu)}$$

from which

$$[R\phi_{j_1}R\cdots R\phi_{j_n}R](x,x)$$
$$=L^{-1}2^{-(n-1)}\sum_{\sigma_2,\dots,\sigma_n=\pm 1} S_n\left(j_1, j_1+\sigma_2 j_2,\dots, j_1+\sum_2^n \sigma_i j_i\right)\phi_{j_1+\Sigma_2^n \sigma_i j_i}$$

where

$$S_n(l_1, l_2,\dots, l_n)=\sum_{k=-\infty}^{\infty}\frac{1}{(\omega_k+\nu)(\omega_{k+l_1}+\nu)(\omega_{k+l_2}+\nu)\cdots(\omega_{k+l_n}+\nu)} \quad \text{(B.4)}$$

Using this notation, for example, we may recalculate

$$D\hat{J}\phi_k=2\lambda(\beta_c L)^{-1}\, b_1 S_1(k)\,\phi_k$$

which agrees with (B.2).

Similarly, recalling that $D\kappa=0$ and using $\phi_j\phi_k=2^{-1}(\phi_{j+k}+\phi_{j-k})$, we find

$$\begin{aligned}D^2\hat{J}(\phi_j,\phi_k)=2\lambda(\beta L)^{-1}\,\{&b_1^2[S_2(j,j+k)\,\phi_{j+k}(x)+S_2(j,j-k)\,\phi_{j-k}(x)]\\ &+b_2[S_1(j+k)\,\phi_{j+k}(x)+S_1(j-k)\,\phi_{j-k}(x)]\\ &+S_1(0)\,D^2\kappa(\phi_j,\phi_k)\}\end{aligned} \quad \text{(B.5)}$$

Applying $\int_0^L D^k\hat{J}(W_1,\dots,W_k)(x)\,dx=0$ leads to

$$D^2\kappa(\phi_j,\phi_k)=-\delta_{jk}(b_2+b_1^2 S_2(j,0)\,S_1(0)^{-1}) \quad \text{(B.6)}$$

Equations (B.5) and (B.6) determine the second term in (B.1b); in particular,

$$M^{-1}E_2^s D^2\hat{J}(\phi_1,\phi_1)=2\lambda(\beta L)^{-1}\,[b_1^2 S_2(1,2)+b_2 S_1(2)][\rho_2(\beta_c)-1]\,\phi_2 \quad \text{(B.7)}$$

and, since $S_2(1,-1)=S_2(1,2)=S_2(2,1)$,

$$2L^{-1}(\phi_1, D^2\hat{J}(\phi_1,\phi_2))=2\lambda(\beta L)^{-1}\,[b_1^2 S_2(2,1)+b_2 S_1(1)] \quad \text{(B.8)}$$

Thus, the value of the second term in (B.1b) is

$$(2\lambda/\beta L)^2\,[1-\rho_2(\beta_c)]^{-1}\,[b_1^2 S_2(1,2)+b_2 S_1(2)][b_1^2 S_2(1,2)+b_2 S_1(1)] \quad \text{(B.9)}$$

The first term in (B.1b) is calculated similarly; (B.6) is needed, as well as the relation $D^3\kappa(\phi_1, \phi_1, \phi_1) = 0$ (which follows from the $S_{L/2}$ symmetry). We find a contribution to (B.1b) which is the sum of

$$3\lambda(\beta L)^{-1} [3b_3 S_1(1) + 4b_1 b_2 S_2(1, 2) + b_1^3 S_3(1, 2, 1)] \qquad \text{(B.10)}$$

and

$$3\lambda(\beta L)^{-1} b_1^3[S_3(1, 2, 1) + S_3(1, 0, 1) - 2S_2(0, 1)^2 S_1(0)^{-1}] \qquad \text{(B.11)}$$

It remains only to show that the sum of (B.9)–(B.11) is positive.

Observing that $b_1 > 0$, we first show that (B.11) is positive. We write $a_k = (\omega_k + \gamma^2)^{-1}$; then, from (B.4),

$$S_3(1, 0, 1)\, S_1(0) = \frac{1}{4} \sum_{j,k} a_j^2 a_k^2 (a_{k+1}^2 + a_{j+1}^2 + a_{k-1}^2 + a_{j-1}^2) \qquad \text{(B.12)}$$

$$S_3(1, 2, 1)\, S_1(0) = \frac{1}{4} \sum_{j,k} a_j^2 a_k^2 (2a_{k+1} a_{k-1} + 2a_{j+1} a_{j-1}) \qquad \text{(B.13)}$$

and

$$S_2(0, 1)^2 = \frac{1}{4} \sum_{j,k} a_j^2 a_k^2 (2a_{k+1} a_{j+1} + 2a_{k+1} a_{j-1} + 2a_{k-1} a_{j+1} + 2a_{k-1} a_{j-1}) \qquad \text{(B.14)}$$

Combining (B.12)–(B.14) shows that

$$[S_3(1, 0, 1) + S_3(1, 2, 1)]\, S_1(0) - 2S_2(0, 1)^2$$
$$= \frac{1}{4} \sum_{j,k} a_j^2 a_k^2 (a_{k+1} + a_{k-1} - a_{j+1} - a_{j-1})^2$$

which completes the verification.

To show that (B.9) plus (B.10) is positive we evaluate explicitly the S_i which appear, for example, by the residue calculus. Setting $d_1 = \omega_1^2 + 4\gamma^2$ and $d_2 = \omega_2 + 4\gamma^2$, and using (3.16), we find

$$S_1(j) = \beta N d_j^{-1}, \qquad j = 1, 2$$
$$S_2(1, 2) = 3\beta N(d_1 d_2)^{-1}$$
$$S_3(1, 2, 1) = \beta N(d_1^2 d_2)^{-1} [10 + (\gamma/\omega_1)^2] + L^2\omega_1^2(8\gamma^2 d_1)^{-1} \operatorname{csch}^2(\gamma L/2)$$
$$\geq 10\beta N(d_1^2 d_2)^{-1}$$

From $\rho_k(\beta_c) = 1$ we have $2qV_0^q = d_1$. Then (B.9) is

$$V_0^{-2}(12d_2)^{-1}(q+2)[(4q-1) + (q + 2(\gamma/\omega_1)^2)] \geq V_0^{-2}(12d_2)^{-1}(q+2)(4q-1)$$

and (B.10) is bounded below by

$$3V_0^{-2}(4d_2)^{-1}\,[(q-1)(q-2)\,d_2+6(q-1)\,d_1+5d_1]$$

The sum of these two equations,

$$V_0^{-2}(12d_2)^{-1}\,(40q^2-47q+61)\,\omega_1^2+(13q^2+34q+9)\,\gamma^2]$$

is positive for all q and γ, which completes the proof.

ACKNOWLEDGMENTS

This work was supported in part by AFOSR grant 87-0010 (to J. L. L.) and by the U.S. DOE and in part by ONR contract DP-F-85-28 (to H. A. R.).

REFERENCES

1. J. L. Lebowitz, H. A. Rose, and E. R. Speer, *J. Stat. Phys.* **50**:657–687 (1988).
2. J. Ginibre and G. Velo, *Ann. Inst. Henri Poincaré* **28**:287–316 (1978).
3. M. Weinstein, *Commun. Math. Phys.* **87**:567–576 (1983).
4. B. Simon, *Functional Integration and Quantum Physics* (Academic Press, New York, 1979).
5. M. Golubitsky and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory* (Springer-Verlag, New York, 1985).
6. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, New York, 1984).
7. L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962).
8. D. F. Dubois and H. A. Rose, *Phys. Rev. A* **24**:1476–1504 (1981).