Excited states in the adiabatic Holstein model

1998 J. Phys. A: Math. Gen. 31 10065

(http://iopscience.iop.org/0305-4470/31/50/005)

J. Phys. A: Math. Gen. **31** (1998) 10065–10085. Printed in the UK PII: S0305-4470(98)96536-2

Excited states in the adiabatic Holstein model

C Baesens† and R S MacKay

The Nonlinear Centre, DAMTP, Silver Street, University of Cambridge, Cambridge CB3 9EW, UK

Received 31 July 1998

Abstract. The adiabatic Holstein model describes interaction of electrons with classical phonons. Near the anti-integrable limit, where electron–phonon coupling dominates electron hopping, Aubry, Abramovici and Raimbault (1992 *J. Stat. Phys.* **67** 675–780) found many local minima of the energy, while at the opposite limit, called integrable, there is only one equilibrium for each choice of mean electronic density. To eliminate the excess local minima on passing from the anti-integrable to the integrable limit, there must be bifurcations with other critical points of higher index: *excited states*. In this paper, we find all the critical points of the energy at the two limits. We find that at the anti-integrable limit the excited states form submanifolds and stratified sets of various types, which we call resonances. We show that homology index theory implies that at least certain numbers of critical points from each resonance survive small perturbation from the anti-integrable limit. We calculate these numbers explicitly for some simple cases, and derive some general rules. The complete homology calculation in the general case and the study of the bifurcations on the route to the integrable limit are left for the future. We conclude by generalizing the approach to allow electron spin, magnetic fields and electron–electron interactions.

1. Introduction

The Holstein model [Ho] describes systems in which fermions can hop between sites of a network $\mathcal{S}$ and interact with phonon variables on the sites but not with each other except via Pauli's exclusion principle. In the adiabatic limit, the phonon variables are regarded as classical and motionless. The physical relevance of this limit can be questioned, but the large ratio between atomic and electronic masses makes it a reasonable starting point. We denote the phonon variable on site s by u_s. The Hamiltonian for the adiabatic Holstein model is

$$H = \sum_s \tfrac{1}{2}u_s^2 + h \tag{1}$$

with

$$h = \sum_s c n_s u_s - t\Delta. \tag{2}$$

Here

$$n_s = a_s^\dagger a_s \tag{3}$$

† On leave from: Centre de Dynamique de Systèmes Complexes, Université de Bourgogne, Dijon, France.

0305-4470/98/5010065+21$19.50

the number operator on site s for fermions, with creation and annihilation operators $a_s^\dagger, a_s$ respectively, satisfying the anticommutation relations $\{a_r, a_s\} = \{a_r^\dagger, a_s^\dagger\} = 0$, $\{a_r^\dagger, a_s\} = \delta_{rs}$, and Δ is an off-diagonal Hermitian 'hopping' operator, i.e. preserving the total number

$$\mathcal{N} = \sum_s n_s \tag{4}$$

but not the individual numbers n_s, e.g.

$$\Delta = \sum_{\langle r,s\rangle} a_r^\dagger a_s \tag{5}$$

over nearest-neighbour pairs $\langle r, s\rangle$ when $\mathcal{S}$ is a graph. For simplicity, in the first six sections of this paper we consider spinless fermions. To treat electrons, the effects of spin are included in section 7, where we also allow magnetic fields and electron–electron interactions.

There are two parameters: c represents fermion–phonon coupling and t represents fermion hopping amplitude. It is easily seen that only the ratio $c^2 : t$ plays a role. The limits $c^2 : t = 0$ and $t : c^2 = 0$ are called by [AAR] the 'integrable' and 'anti-integrable' limits, respectively, by analogy with problems in Hamiltonian dynamics.

A state of the Holstein model is specified by giving a function $u : \mathcal{S} \to \mathbb{R}$ and a one-dimensional subspace Ψ of the complex Hilbert space Ω spanned by all vectors of the form $a_{s_1}^\dagger \dots a_{s_k}^\dagger |\emptyset\rangle$ for any choice of set of distinct sites $s_1 \dots s_k$ with $k \leqslant S$, the number of sites in $\mathcal{S}$, where $|\emptyset\rangle$ is a 'vacuum state' with $a_s|\emptyset\rangle = 0$ for all $s \in \mathcal{S}$. The *energy* W of a state (u, Ψ) is the expectation of the Hamiltonian in that state, i.e.

$$W(u, \Psi) = \sum_s \frac{1}{2}u_s^2 + \frac{\langle\psi|h|\psi\rangle}{\langle\psi|\psi\rangle} \tag{6}$$

where ψ is any non-zero element of Ψ (note that the result is independent of the choice of representative $\psi \in \Psi$). For an infinite system, W might not be well defined, but its derivative with respect to the state (u, Ψ) is well defined, as is usual in variational problems. The *equilibrium states* are the critical points of the energy.

Aubry, Abramovici and Raimbault [AAR] found all the local minima of the energy for the adiabatic Holstein model at the anti-integrable limit ($t = 0$) and showed that they have locally unique continuations for small t (proof improved by [BM1] and extended to non-zero temperature electrons in [BM2]), and have finite coherence length (proof improved in [BM3]). In those papers, the term 'equilibrium state' was reserved for local minima. It is clear by topological arguments, however, that there must be other critical points of the energy, for example minimax points. In this paper, we find them all.

The Hamiltonian H commutes with the total number operator $\mathcal{N}$, so without loss of generality we can restrict attention to the subspaces with a definite number of fermions. We restrict attention to finite systems, with S sites and F fermions. It is an open question to what extent our continuation results from the anti-integrable limit can be made uniform in S and F.

First, note that the space of fermionic states is topologically equivalent to† $\mathbb{C}P^K$, where

$$K = \binom{S}{F} - 1 \tag{7}$$

† Complex projective K-space $\mathbb{C}P^K$ is the quotient of $\mathbb{C}^{K+1}\backslash\{0\}$ under the equivalence relation $\psi \sim \lambda\psi$ for all $\lambda \in \mathbb{C}\backslash\{0\}$.

because the space of antisymmetric F-fermion wavefunctions over S sites is linear with complex dimension

$$\binom{S}{F} = \frac{S!}{F!(S-F)!}. \tag{8}$$

Secondly, the energy grows quadratically as any of the u_n go to infinity. Hence the energy is a *proper* function from $\mathbb{R}^S \times \mathbb{C}P^K$ to $\mathbb{R}$, meaning that the inverse image of a compact set is compact. Furthermore, the negative gradient of the energy is inward on the boundary of a suitable large compact subset B of $\mathbb{R}^S \times \mathbb{C}P^K$; in fact, we know that all critical points must have $u \in [-c, 0]^S$, because they satisfy

$$u = -c\rho \tag{9}$$

where ρ_s is the *fermionic density* on site s, which lies *a priori* in $[0, 1]$[AAR]. So we can take B to have the form $[-c-\varepsilon, \varepsilon]^S \times \mathbb{C}P^K$ for some $\varepsilon > 0$. It follows that the numbers c_i of critical points of W of index† i, if all non-degenerate, must satisfy the Morse inequalities (e.g. [NS]):

$$\sum_{j=0}^{i} (-)^j c_{i-j} \geqslant \sum_{j=0}^{i} (-)^j \beta_{i-j} \qquad \text{for each } 0 \leqslant i \leqslant 2K \tag{10}$$

with equality for $i = 2K$, where β_i is the ith Betti number of B, which is the same as that for $\mathbb{C}P^K$, i.e.

$$\beta_i = \begin{cases} 1 & \text{if } 0 \leqslant i \leqslant 2K \text{ is even} \\ 0 & \text{if } i \text{ is odd.} \end{cases} \tag{11}$$

In particular, by adding the inequalities for $i = 2k$ and $2k-1$, there must be at least one critical point of each even index $2k$. Furthermore, since from [AAR] there are $K+1 = \binom{S}{F}$ critical points of index 0 at and near the anti-integrable limit, there must be at least K critical points of index 1, if non-degenerate. Even if the critical points are not all non-degenerate, topological arguments (such as the minimax principle or Liusternik–Schnirelman theory, e.g. [Sc], or homology index theory, which we will discuss in section 4) show that there must be other critical points.

In this paper we show where these other critical points are, for parameter values at and near the anti-integrable limit. We find that there are many degenerate critical points which are not local minima; they form submanifolds or, more generally, 'stratified sets'. Furthermore, we show that the theory of 'homology index' for an isolated critical set implies that at least certain numbers of critical points survive small perturbation. We also find where all the critical points are at and near the integrable limit. This picture will help in understanding how the equilibrium states bifurcate on passing from the anti-integrable to the integrable limit.

The plan of the paper is as follows. In section 2, we analyse a simple pedagogical case: one fermion on two sites in the whole of the (t, c) parameter space. Then in section 3 we address the general case of F fermions on S sites at the anti-integrable limit $t : c^2 = 0$. We review homology index theory in section 4, and compute it for a range of critical sets in section 5 and derive some general results. In section 6 we address briefly the opposite limit $c^2 : t$ zero or small. In section 7 we address the effects of spin, magnetic fields and electron–electron interaction. The paper concludes with a discussion in section 8.

† For a reminder of the definition of the index of a non-degenerate critical point, see the beginning of section 4.

2. A simple case: one fermion on two sites

We begin by analysing a simple case: $S = 2$, $F = 1$. We write ψ_s for the complex amplitude of the fermion on site s, and without loss of generality we suppose the wavefunction is normalized:

$$|\psi_1|^2 + |\psi_2|^2 = 1. \tag{12}$$

Then

$$W = \tfrac{1}{2}u_1^2 + \tfrac{1}{2}u_2^2 + cu_1|\psi_1|^2 + cu_2|\psi_2|^2 - t(\psi_1\bar{\psi}_2 + \bar{\psi}_1\psi_2). \tag{13}$$

Completing the square gives

$$W = \tfrac{1}{2}(u_1 + c|\psi_1|^2)^2 + \tfrac{1}{2}(u_2 + c|\psi_2|^2)^2 - \tfrac{1}{2}c^2(|\psi_1|^4 + |\psi_2|^4) - t(\psi_1\bar{\psi}_2 + \bar{\psi}_1\psi_2). \tag{14}$$

So the u-dependence is quadratic, and at all critical points the first two terms vanish. Thus we are left with

$$W = -\tfrac{1}{2}c^2(|\psi_1|^4 + |\psi_2|^4) - t(\psi_1\bar{\psi}_2 + \bar{\psi}_1\psi_2). \tag{15}$$

After identifying pairs (ψ_1, ψ_2) which differ by a global phase factor, the space of fermionic degrees of freedom becomes $\mathbb{C}P^1$, which is the Riemann sphere. One can use spherical coordinates (θ, φ) as follows:

$$\begin{aligned} &\psi_1 = \cos\frac{\theta}{2},\ \psi_2 = \mathrm{e}^{\mathrm{i}\varphi}\sin\frac{\theta}{2} \qquad \text{for } 0 \leqslant \theta < \pi, \varphi \in \mathbb{T}^1 = \mathbb{R}/2\pi\mathbb{Z} \\ &\psi_1 = \mathrm{e}^{-\mathrm{i}\varphi}\cos\frac{\theta}{2},\ \psi_2 = \sin\frac{\theta}{2} \qquad \text{for } 0 < \theta \leqslant \pi, \varphi \in \mathbb{T}^1 \end{aligned} \tag{16}$$

which agree in the overlap zone $0 < \theta < \pi$ up to a global phase factor. In these coordinates, W becomes

$$W = -\tfrac{1}{2}c^2(1 - \tfrac{1}{2}\sin^2\theta) - t\cos\varphi\sin\theta. \tag{17}$$

For $t = 0$, the critical points in $\mathbb{C}P^1$ are two minima at $\theta = 0$ and π respectively (the north and south poles), corresponding to putting the fermion on one site or the other, and a circle of maxima at $\theta = \pi/2$ (the equator), corresponding to making a 'resonant pair' where the fermion spends half its time on each site.

This circle of critical points is 'normally hyperbolic' [Fe] for the negative gradient flow of W. Following Bott [Bo] we define the *index* of a submanifold of critical points, if it is normally hyperbolic for the negative gradient flow of W, to be the dimension of the unstable manifold of any of its points, in other words, the dimension of the unstable manifold of the critical submanifold minus its own dimension. Thus this circle of critical points has index 1.

For t non-zero, the degeneracy is lifted. The equator is still invariant under the gradient flow, but only the points $\varphi = 0$ and π remain critical; they are of index 1 and 2 respectively. As t increases, the minima at the poles migrate along the semicircle $\varphi = 0$ according to

$$\sin\theta = \pm 2t/c^2 \tag{18}$$

until $t = \frac{1}{2}c^2$, when they coalesce with the index 1 point and it becomes of index 0. Thereafter, as t continues to increase, or equivalently as c decreases, the critical points remain just one of index 2 at $\theta = \pi/2, \varphi = \pi$ and one of index 0 at $\theta = \pi/2, \varphi = 0$. At the integrable limit, $u_1 = u_2 = 0$ for both critical points and they correspond to the two eigenstates for the fermion.

The sequence of events is sketched in figure 1.

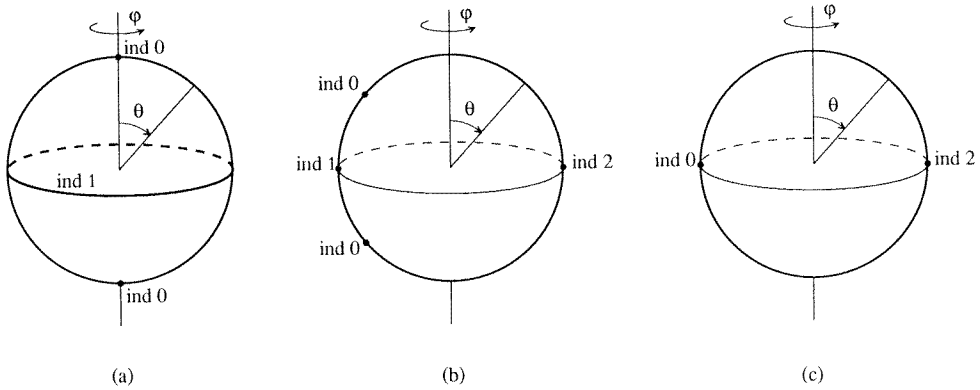

Figure 1. Critical points of the energy on the Riemann sphere for one fermion on two sites for three values of $t : c^2$. (*a*) $t = 0$ (anti-integrable limit), (*b*) $0 < t : c^2 < \frac{1}{2}$, (*c*) $t : c^2 > \frac{1}{2}$.

3. The anti-integrable limit

Having seen what happens for the special case of one fermion on two sites, we now proceed to the general case of F fermions on S sites. In this section we study the anti-integrable limit in parameter space, where $t : c^2 = 0$; without loss of generality we put $c = 1$.

At the anti-integrable limit, there are $\binom{S}{F}$ local minima, given by choosing F out of the S sites to be occupied and the remainder unoccupied. The main point of this paper is that there are also critical points of higher index. They come in isolated subsets, some of which are differentiable submanifolds but many of which are only 'stratified sets' (to be defined shortly). They arise by formation of groups of 'resonant' sites. We will proceed by examining a few cases and then formulate a general theorem.

The simplest case is resonant pairs, generalizing section 2. If a pair of sites is chosen to be 'resonant', with $\rho_s = -u_s = \frac{1}{2}$, and the remaining $F-1$ fermions are put on $F-1$ of the remaining $S-2$ sites as desired, then one obtains a critical circle of index 1. It is a circle of critical points because the relative phase of the fermion wavefunction on the resonant pair of sites is arbitrary. It has index 1 because at the anti-integrable limit there is no coupling between sites, so the resonant pair can be treated as in section 2. We call such a circle of critical points a '$\frac{1}{2}$-resonance'. The number of such critical circles is $\binom{S}{2 \quad F-1}$, using the multinomial notation

$$\binom{S}{r_1 \quad \cdots \quad r_k} = \frac{S!}{r_1! \ldots r_k! (S - \sum_i r_i)!}. \tag{19}$$

Similarly, critical two-tori of index 2 occur by choosing a triple of sites to be resonant, with $\rho_s = -u_s = \frac{1}{3}$ (or $\frac{2}{3}$), and putting the remaining $F-1$ (respectively, $F-2$) fermions on $F-1$ ($F-2$) of the remaining $S-3$ sites. This can be done in $\binom{S}{3 \quad F-1} + \binom{S}{3 \quad F-2}$ ways. Each choice gives a critical two-torus which we call a $\frac{1}{3}$ or $\frac{2}{3}$ resonance, respectively. The case of one fermion on three sites has index two because completing the square as in section 2 leaves $W = -\frac{1}{2}\sum \rho_s^2$ at the anti-integrable limit, with $\rho_s = |\psi_s|^2$, and the constraint $\sum \rho_s = 1$. Subject to this constraint, W is maximum when all three $\rho_s = \frac{1}{3}$, which defines a two-torus after identifying states

which differ only by a global phase shift, and has unstable manifold of dimension four, hence index two. The same result can be obtained in the case of two fermions on three sites by fermion–hole symmetry.

Critical manifolds of index 3 occur in several ways. One can put one or three fermions on a resonant quadruplet with $\rho_s = -u_s = \frac{1}{4}$ or $\frac{3}{4}$ respectively, the remaining fermions being put on $F-1$ (resp. $F-3$) of the remaining $S-4$ sites. Each choice gives a critical three-torus of index 3, called a $\frac{1}{4}$ or $\frac{3}{4}$ resonance.

Alternatively, one can put two fermions on a resonant quadruplet with $\rho_s = -u_s = \frac{1}{2}$, and the remaining $F-2$ on any choice of the remaining $S-4$ sites. There are $\binom{S}{4}\binom{}{F-2}$ ways of doing this. Each of these choices gives a connected set of critical points called a $\frac{2}{4}$-resonance, which is not a submanifold as it has some singularities, but nonetheless can be regarded as being seven-dimensional and of index 3, as follows.

To see the structure of the $\frac{2}{4}$-resonance, the space of two-fermion states on a set of four sites has complex dimension $\binom{4}{2} - 1 = 5$, with homogeneous coordinates ψ_{12}, ψ_{13}, ψ_{14}, ψ_{23}, ψ_{24}, ψ_{34}, say, where ψ_{ij} corresponds to applying the creation operator $a_i^\dagger a_j^\dagger$ to the vacuum. For convenience, we will impose the normalization condition $\sum |\psi_{ij}|^2 = 1$. The resonance condition imposes the restrictions

$$\rho_1 = |\psi_{12}|^2 + |\psi_{13}|^2 + |\psi_{14}|^2 = \tfrac{1}{2} \tag{20}$$

$$\rho_2 = |\psi_{12}|^2 + |\psi_{23}|^2 + |\psi_{24}|^2 = \tfrac{1}{2} \tag{21}$$

$$\rho_3 = |\psi_{13}|^2 + |\psi_{23}|^2 + |\psi_{34}|^2 = \tfrac{1}{2} \tag{22}$$

$$\rho_4 = |\psi_{14}|^2 + |\psi_{24}|^2 + |\psi_{34}|^2 = \tfrac{1}{2}. \tag{23}$$

Taking $\frac{1}{2}[(20) + (21) + (22) - (23)]$ yields

$$|\psi_{12}|^2 + |\psi_{13}|^2 + |\psi_{23}|^2 = \tfrac{1}{2}. \tag{24}$$

Subtracting (20) from this, we obtain

$$|\psi_{14}|^2 = |\psi_{23}|^2. \tag{25}$$

Similarly

$$|\psi_{34}|^2 = |\psi_{12}|^2 \tag{26}$$

and

$$|\psi_{24}|^2 = |\psi_{13}|^2. \tag{27}$$

Thus for non-zero choices of $\psi_{12}, \psi_{13}, \psi_{23}$ on the five-sphere (24) we have a three-torus given by (25)–(27) (with angles $\arg\psi_{14}, \arg\psi_{34}, \arg\psi_{24}$) which reduces to a two-torus on elimination of global phase rotation. Hence for $\psi_{12}, \psi_{13}, \psi_{23} \neq 0$, the $\frac{2}{4}$ resonance is locally a seven-dimensional submanifold. But near $\psi_{12} = 0$, for example, with $\psi_{13,}\psi_{23} \neq 0$, it has locally the structure of $K \times \mathbb{T}^3 \times I$, where K is a cone over a two-torus,

$$K = \{(\psi_{12}, \psi_{34}) \in \mathbb{C}^2 : |\psi_{12}|^2 = |\psi_{34}|^2\}, \tag{28}$$

$\mathbb{T}^d$ denotes a d-torus and I an interval. This is because given $(\psi_{12}, \psi_{34}) \in K$ and $|\psi_{13}| \in (0, 1/\sqrt{2})$, then (24) determines $|\psi_{23}|$ and (25), (27) give a four-torus, which reduces to a three-torus on elimination of global phase rotation.

Thus the $\frac{2}{4}$-resonance is *not* a submanifold. Instead it is a *stratified set*, meaning a locally finite disjoint union of smooth submanifolds of various dimensions (called the

'strata'), such that if U and V are strata and the closure $\bar{U}$ intersects V then $V \subset \bar{U}$, with for each stratum an algebraic local model for the neighbourhood of any point of the stratum (see [Mat] for a more general definition). In the case of the $\frac{2}{4}$-resonance, the strata are a seven-dimensional submanifold ($\psi_{12}, \psi_{13}, \psi_{23} \neq 0$), three four-dimensional submanifolds ($\psi_{12} = 0, \psi_{13}, \psi_{23} \neq 0$ and permutations) and three circles ($\psi_{12} = 0, \psi_{13} = 0, \psi_{23} \neq 0$ and permutations) (look ahead to figure 3). We define the *dimension* of a stratified set to be the maximum of the dimensions of its strata, giving seven for the $\frac{2}{4}$-resonance.

Bott's definition of index (section 2) does not generalize easily to stratified sets, as in general they are not normally hyperbolic (nor even submanifolds). In the case of a repelling stratified set, however, it is natural to define its index to be its codimension. Thus the $\frac{2}{4}$ resonance has index 3 because it corresponds to a maximum of $W = -\frac{1}{2}\sum_{s=1}^{4} \rho_s^2$ in the space where $\sum \rho_s = 2$, so is repelling, and the total fermionic space has (real) dimension 10. In the next section we will show a better generalization of the definition of index of a critical set.

Another type of critical set can be obtained by taking several resonant tuples with different fermionic densities. The simplest case is that of a resonant pair with $\rho = \frac{1}{2}$ and a resonant triple with $\rho = \frac{1}{3}$ ($\frac{2}{3}$ is similar). This gives a critical set of dimension seven which turns out to be a differentiable submanifold. We call it a $\frac{1}{2} \times \frac{1}{3}$ product resonance. To show that it is a submanifold, suppose without loss of generality that $\rho_1 = \rho_2 = \frac{1}{2}$ and $\rho_3 = \rho_4 = \rho_5 = \frac{1}{3}$. Its equations are (modulo global phase rotation)

$$\psi_{12} = \psi_{34} = \psi_{35} = \psi_{45} = 0 \tag{29}$$

and

$$\begin{array}{ccccccc} |\psi_{13}|^2 & + & |\psi_{14}|^2 & + & |\psi_{15}|^2 & = & \frac{1}{2} \\ + & & + & & + & & \\ |\psi_{23}|^2 & + & |\psi_{24}|^2 & + & |\psi_{25}|^2 & = & \frac{1}{2} \\ \| & & \| & & \| & & \\ \frac{1}{3} & & \frac{1}{3} & & \frac{1}{3} & & \end{array} . \tag{30}$$

This defines a two-dimensional hexagon in the space of $|\psi_{ij}|^2$, illustrated in figure 2. The set of points of the $\frac{1}{2} \times \frac{1}{3}$-resonance projecting to any disk in the interior of the hexagon is diffeomorphic to $D^2 \times \mathbb{T}^5$, where D^2 denotes a closed two-dimensional disk. The set of points projecting to a thin strip including an edge of the hexagon and neighbourhoods of its endpoints is a $\mathbb{T}^2$-bundle over $D^2 \times \mathbb{S}^3$, where $\mathbb{S}^d$ denotes a d-sphere, because without loss of generality choose the edge $|\psi_{13}|^2 = 0$; then for $\epsilon > 0$ we have $\{\psi_{13} : |\psi_{13}|^2 \leqslant \epsilon\} \cong D^2$, $|\psi_{23}|^2 = \frac{1}{3} - |\psi_{13}|^2$ gives a circle for each ψ_{13} (if $\epsilon < \frac{1}{3}$) which we reduce to a point by global phase rotation, $|\psi_{24}|^2 + |\psi_{25}|^2 = \frac{1}{6} + |\psi_{13}|^2$ gives an $\mathbb{S}^3$ for each ψ_{13}, and the equations $|\psi_{14}|^2 = \frac{1}{3} - |\psi_{24}|^2$ and $|\psi_{15}|^2 = \frac{1}{3} - |\psi_{25}|^2$ give a $\mathbb{T}^2$ for each $(\psi_{13}, \psi_{24}, \psi_{25}) \in D^2 \times \mathbb{S}^3$ (if $\epsilon < \frac{1}{6}$). Alternatively, the strip can be viewed as a $\mathbb{T}^3$-bundle over $D^2 \times \mathbb{S}^2$ by applying the global phase rotation to the $\mathbb{S}^3$ instead of the circle.

It is clear that these constructions can be continued, giving critical sets of a huge variety of types, consisting of collections of resonant Q-tuplets with density $\rho_s = -u_s = P/Q$ for any integers P, Q such that $2 \leqslant Q \leqslant M$ and $0 < P < Q$. In particular, for a large system, any sequence of rational densities can be obtained locally. Note that the cases $1/Q$ and $(Q-1)/Q$ give $(Q-1)$-tori of index $Q-1$.

It is also clear that formation of resonant sets is the only way to make critical points at the anti-integrable limit, other than the local minima. In outline, this is because all critical points must satisfy $\rho_s = -u_s$, and for any site s for which u_s is distinct from the values at

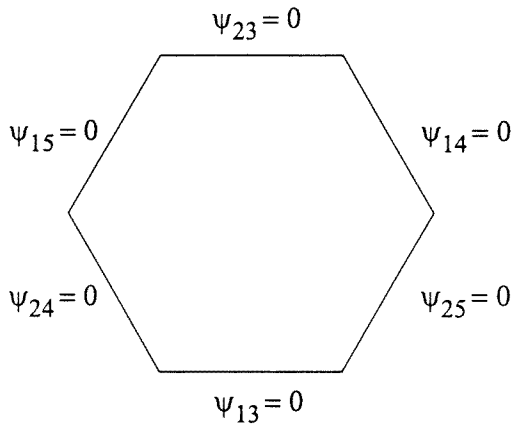

Figure 2. The space of squared amplitudes of the components of the wavefunction for the $\frac{1}{2}\times\frac{1}{3}$-resonance is a hexagon.

all other sites, the fermionic stationary states all have $\rho_s = 0$ or 1. We formulate this as a theorem.

Theorem 1. The critical set at the anti-integrable limit in the F-fermion subspace, is the disjoint union of compact isolated critical sets C_ρ over the set of functions $\rho : \mathcal{S} \to \mathbb{Q} \cap [0,1]$ such that

(1) $\sum_s \rho_s = F$,

(2) writing $S_{\frac{p}{q}}(\rho) := \{s \in \mathcal{S} : \rho_s = \frac{p}{q}\}$, $\forall \frac{p}{q} \in \mathbb{Q}$ in lowest terms, the cardinality $\sharp S_{\frac{p}{q}}(\rho)$ is a multiple of q, say $k_{pq}q$,

where C_ρ is given by the set of normalized F-fermion wavefunctions whose coordinates $\psi_{s_1\cdots s_F}$ satisfy

(1) $\psi_{s_1\ldots s_F} = 0$ if $\sharp\{i : s_i \in S_{\frac{p}{q}}(\rho)\} \neq k_{pq}p$ for some $\frac{p}{q} \in \mathbb{Q}$, and

(2) $\sum_{s_1\ldots s_F \ni s} |\psi_{s_1\ldots s_F}|^2 = \rho_s$, where '$\ni$' stands for 'containing'.

Proof. At the anti-integrable limit, after elimination of u we have

$$W(\Psi) = -\tfrac{1}{2}\sum_s \rho_s^2 \tag{31}$$

where, restricting to the F-fermion subspace,

$$\rho_s = \frac{\sum_{s_1\ldots s_F \ni s} |\psi_{s_1\ldots s_F}|^2}{\sum_{s_1\ldots s_F} |\psi_{s_1\ldots s_F}|^2}. \tag{32}$$

So

$$\frac{\partial W}{\partial \psi_{s_1\ldots s_F}} = -\sum_s \rho_s \frac{\partial \rho_s}{\partial \psi_{s_1\ldots s_F}} \tag{33}$$

$$= \left(\sum_{s=1}^{S} \rho_s^2 - \sum_{i=1}^{F} \rho_{s_i}\right)\bar{\psi}_{s_1\ldots s_F}. \tag{34}$$

Thus Ψ is a critical point iff for all $s_1 \ldots s_F$ either

$$\psi_{s_1\ldots s_F} = 0 \tag{35}$$

or

$$\sum_{i=1}^{F} \rho_{s_i} = M := \sum_s \rho_s^2. \tag{36}$$

These conditions oblige all the ρ_s to be equal (and hence to $\frac{F}{S}$), unless $\mathcal{S}$ decomposes into disjoint subsets S_j, $j = 1, \ldots, J$ with the property that $\exists p_j \in \mathbb{Z}$ for $j = 1, \ldots, J$ such that $\sum_j p_j = F$ and $\psi_{s_1 \ldots s_F} = 0$ if $\sharp\{i : s_i \in S_j\} \neq p_j$ for some j. If such a decomposition exists and is maximal, then condition (36) implies that ρ is constant on each S_j. But then from (32), $\sum_{s \in S_j} \rho_s = p_j$, so $\rho_s = \frac{p_j}{q_j}$ on S_j, where $q_j = \sharp S_j$, the size of S_j. Finally if two of the S_j have the same fermionic density then the corresponding critical set is a subset of a larger one where these two S_j are merged into one. So every critical point belongs to one of the sets C_ρ. Sets C_ρ with different functions ρ are disjoint. They are compact because they can be viewed as closed and bounded subsets of $\mathbb{R}^n$ for some n. □

Probably each set C_ρ is also connected, but we did not prove this.

Note that the energies of the excited states are easy to compute at the anti-integrable limit. The minima for F fermions on S sites have energy $-\frac{1}{2}F$. Formation of a resonant Q-tuplet with P fermions raises the energy by $\frac{P(Q-P)}{2Q}$. This is at least $\frac{1}{4}$, the minimum change being for the case $Q = 2$, $P = 1$.

4. Review of homology index theory

In order to see how many critical points from the critical sets at the anti-integrable limit survive general perturbation, and to make a good generalization of index to critical stratified sets, we need a rapid tutorial on 'homology index theory', due to [Co] (cf section 3 of [An]).

The *index* of a non-degenerate critical point of a C^2 function $W : M \to \mathbb{R}$ on a manifold M is the number of negative squares in the Lagrange canonical form for the second variation of W; equivalently it is the dimension of the unstable manifold for the negative gradient flow of W, with respect to any Riemannian metric on M. On $\mathbb{C}P^K$ we choose the natural Riemannian metric, defined at a point $\Psi \in \mathbb{C}P^K$ with homogeneous coordinates ψ_k, $k = 0, \ldots, K$ by the inner product

$$\langle \xi, \eta \rangle = \mathrm{Re}\left(\sum \bar{\xi}_k \eta_k - \sum \bar{\xi}_i \psi_i \sum \bar{\psi}_j \eta_j / \sum \bar{\psi}_k \psi_k \right). \tag{37}$$

Then the negative gradient flow of a function $W : \mathbb{C}P^K \to \mathbb{R}$ is represented by

$$\dot{\psi}_k = -\frac{\partial W}{\partial \bar{\psi}_k}. \tag{38}$$

The concept of index has various extensions. In particular each compact isolated set C of critical points has a '*homology index*' which is a function γ from $\mathbb{Z}_+$ to $\mathbb{Z}_+ : i \mapsto \gamma_i$, to be defined shortly. First we give some examples. If C is a non-degenerate critical point of index I then $\gamma_i = 0$ for all $i \neq I$ and $\gamma_I = 1$. If C is a normally hyperbolic submanifold for the negative gradient flow of W, of dimension d, and with unstable manifold of dimension $d + I$ then

$$\gamma_i = 0 \qquad \text{for } i < I \text{ or } i > d + I \tag{39}$$

$$\gamma_{I+j} = \beta_j \qquad \text{for } 0 \leqslant j \leqslant d \tag{40}$$

where β_j is the jth Betti number of C, i.e. the dimension of the homology group $H_j(C; \mathbb{R})$, e.g. for a d-torus, β_j is the binomial coefficient $\binom{d}{j}$. For an introduction to homology

groups, see [NS]. In a nutshell, *chains* are linear combinations of 'nice' subsets with the same dimension, and *cycles* are chains whose boundary is zero; every boundary is a cycle but not every cycle is a boundary; the ith *homology group* is the set of cycles of dimension i modulo boundaries of $(i+1)$-dimensional chains.

To give the general definition of homology index of a critical set, each compact isolated set C of critical points has a compact neighbourhood N (called an *isolating neighbourhood*) and an $\epsilon > 0$, such that C is the maximal invariant set in N under the positive and negative gradient flow of W, and if $x[0,t]$ is a trajectory of the negative gradient flow lying in N with $x(t) \in \partial N$, t maximal, then $x(t, t+\epsilon) \cap N = \emptyset$. The set E of such points $x(t)$ (considering all trajectories x) is called the *exit set* of N. Then γ_i is defined to be the dimension of the relative homology group $H_i(N, E; \mathbb{R})$. The relative homology group of a topological space N modulo a subset E is defined like the homology group of N but treating any subset of E as 0.

For example, to compute the homology index of the equator in the one fermion on two sites example of section 2, we can take as isolating neighbourhood the annulus between the tropics of Cancer and Capricorn. Then the exit set is the union of the two tropics; $H_2(N, E)$ is generated by the whole annulus, $H_1(N, E)$ by any arc joining the two tropics, and $H_0(N, E) = 0$. Thus $\gamma_0 = 0, \gamma_1 = 1, \gamma_2 = 1$.

For homology index we adopt the notational convention $\gamma = [\gamma_0, \gamma_1, \ldots, \gamma_n]$, where n is the least integer such that $\gamma_i = 0$ for all $i > n$. So, for example, the homology index of the equator in the one fermion on two sites problem is $\gamma = [0, 1, 1]$.

The significance of the homology index, for our purposes, is that it gives strong restrictions on what can happen to a set C of critical points on perturbation of W. Think of the perturbation as belonging to a continuous family W_ϵ with $W_0 = W$. Suppose ϵ_0 is small enough that an isolating neighbourhood N can be chosen for C such that no critical points lie on ∂N for any $\epsilon \in [0, \epsilon_0]$. If all the critical points of W_ϵ in N are non-degenerate (as is the generic case) then the numbers c_i with index i satisfy the Morse inequalities

$$\sum_{i=0}^{n} (-)^i c_{n-i} \geqslant \sum_{i=0}^{n} (-)^i \gamma_{n-i} \qquad \text{for all } n \tag{41}$$

with equality in the top dimension. In particular, by taking the sum of two successive inequalities,

$$c_n \geqslant \gamma_n \qquad \text{for all } n. \tag{42}$$

The difference between the left and right hand sides of (41) can be attributed to 'unpaired connecting orbits' [Fl] (see [Sa] for a review).

In fact, the Morse inequalities apply even if some critical points in N are degenerate, provided c_i is defined to be the sum of the ith homology indices over any decomposition of the set of critical points in N into isolated critical sets. This is useful to us because certain perturbations we shall consider possess symmetries which lead to some submanifolds of a submanifold of critical points being preserved.

The homology index also places restrictions on the result of interaction of two or more isolated critical sets on perturbation. The union U of the relevant critical sets plus all their connecting orbits under the gradient flow is an isolated invariant set for the gradient flow, so if we compute its homology index and it is not identically zero then at least some of the critical points will remain, the numbers being restricted by the Morse inequalities again.

To compute the homology index of an isolated critical set there are several possible procedures. For example, one can make a cell decomposition (as we will do for the $\frac{2}{4}$-resonance in section 5.2). One can sometimes shorten the calculation by using exact

sequences (e.g. [GH]), for example $H_i(A) \to H_i(X) \to H_i(X, A) \to H_{i-1}(A)$ where A is a well chosen subset of X and the maps are, respectively, those induced by inclusion $A \subset X$, reduction modulo A, and the boundary map; but a calculation of the rank or nullity of some boundary operators is usually still required, so the savings are not great. Or one can perturb to a Morse function and choose a Riemannian metric such that all connecting orbits are transverse and then compute the homology of the resulting Witten complex [Fl, Sa]; this may sound as if it defeats the purpose of the exercise, but the idea is that by computing the Floer homology for one perturbation of the function one obtains constraints on the numbers of critical points for all small perturbations.

Note two useful facts: for a product $X \times Y$ of topological spaces, $H_i(X \times Y) \cong \sum_{j=0}^{i} H_j(X) \oplus H_{i-j}(Y)$, taking homology over a field (Künneth formula), and for an orientable manifold M of dimension d, $H_i(M) \cong H_{d-i}(M)$ (Poincaré duality).

5. Near the anti-integrable limit

5.1. Simplest results

Near the anti-integrable limit, i.e. for $t : c^2$ small (we shall again take $c = 1$ throughout this section), all non-degenerate critical points persist, but we do not expect any critical subsets of dimension greater than zero to be preserved. However, by homology index theory, at least certain numbers of critical points must persist from each. For example, from a $(Q-1)$-torus at least 2^{Q-1} critical points persist, if non-degenerate.

Before analysing the details of how many critical points are preserved from each of the critical sets of theorem 1, we remark that close enough to the anti-integrable limit, there are no other critical points than those arising from anti-integrable critical sets. This is because at the anti-integrable limit all the critical subsets are locally maximal. There exists $t_0 > 0$ and a union U of isolating neighbourhoods for the set of the critical points at the anti-integrable limit such that for $t < t_0$ all critical points in U are obtained by the above procedure. There exists $t_1 \in (0, t_0)$ such that outside U the gradient of W is too large at $t = 0$ for critical points to be created there while $t < t_1$. Hence for $t < t_1$, all critical points are obtained by the above procedure.

To compute the homology index of an isolated critical set C_ρ we need a suitable isolating neighbourhood N. For a single P/Q resonance we can choose N to be the pre-image of a ball in ρ-space under the map $\Psi \mapsto \rho_s = \sum_{\underline{s} \ni s} |\psi_{\underline{s}}|^2$, about the centre of the simplex (the P/Q resonance corresponds to the single maximum in ρ-space in the centre). Then the exit set is the whole boundary ∂N, and it remains only to decide on the homology of N relative to ∂N. A minor simplification can be achieved by noting that by inverting the sign of W, the exit set can be made empty, and so

$$H_i(N, \partial N) \cong H_{d-i}(N) \qquad \text{where } d = \dim N = 2\left[\binom{Q}{P} - 1\right]. \tag{43}$$

Thus, for example (using the notation introduced in the previous section),

- a $\frac{1}{2}$ resonance has homology index $\gamma = [0, 1, 1]$, because $N \cong I \times \mathbb{T}^1$,
- a $\frac{1}{3}$ or $\frac{2}{3}$ resonance has homology index $\gamma = [0, 0, 1, 2, 1]$, because $N \cong D^2 \times \mathbb{T}^2$,
- a $\frac{1}{4}$ or $\frac{3}{4}$ resonance has homology index $\gamma = [0, 0, 0, 1, 3, 3, 1]$, because $N \cong B^3 \times \mathbb{T}^3$.

Here, B^d denotes a d-dimensional ball. The generalization to $1/Q$ and $(Q-1)/Q$ resonances, $Q > 4$, is obvious.

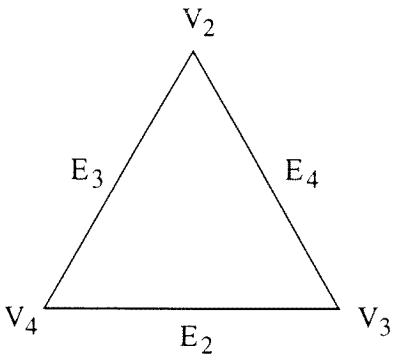

Figure 3. The triangle defined by (44) for the $\frac{2}{4}$-resonance Σ. The edges have the equations $E_2 : |\psi_{12}|^2 = 0$, $E_3 : |\psi_{13}|^2 = 0$, and $E_4 : |\psi_{14}|^2 = 0$, respectively, minus the vertices.

5.2. The $\frac{2}{4}$ resonance

The simplest non-trivial case of a single resonance is the $\frac{2}{4}$ resonance. As described in section 3 the equations for the $\frac{2}{4}$ resonance can be reduced to:

$$|\psi_{12}|^2 + |\psi_{13}|^2 + |\psi_{14}|^2 = \tfrac{1}{2} \tag{44}$$

$$|\psi_{34}|^2 = |\psi_{12}|^2 \tag{45}$$

$$|\psi_{24}|^2 = |\psi_{13}|^2 \tag{46}$$

$$|\psi_{23}|^2 = |\psi_{14}|^2 \tag{47}$$

modulo global phase shift. The first equation defines a triangle in the space of $|\psi_{12}|^2, |\psi_{13}|^2, |\psi_{14}|^2 \geqslant 0$, illustrated in figure 3. Above each interior point of the triangle is a five-torus, which collapses to a three-torus over each edge point and to a one-torus (circle) over each vertex. The resulting object Σ is not a manifold, because a local model near any edge point is $K \times \mathbb{T}^3 \times I$, as already remarked in section 3, where K denotes the cone over a two-torus.

To compute the homology of this seven-dimensional object, we use a cell decomposition. A *cell* in a (Hausdorff) topological space X is a subset homeomorphic to an open ball of some dimension, such that the homeomorphism has a continuous extension from the closed ball into X. A *cell decomposition* of X is a partition of X into cells such that the boundary of any cell is contained in the union of all cells of lower dimension (e.g. [Th]). The first step in obtaining our cell decomposition is to decompose the triangle of figure 3 into its interior Δ, its three edges E_2, E_3, E_4 and its three vertices V_2, V_3, V_4. Next, the five-torus above each point of Δ is slit along five four-tori: $\theta_{34} = \theta_{12}$, $\theta_{24} = \theta_{13}$, $\theta_{23} = \theta_{14}$, $\theta_{12} = \theta_{24}$ and $\theta_{13} = \theta_{23}$ (modulo global phase shift), and the resulting intersections, where we introduce the notation

$$\theta_{ij} = \arg \psi_{ij}. \tag{48}$$

Note that our choice does not preserve permutation symmetry, but to do so would require more subdivision. The three-torus above each point of each edge is slit along three two-tori and their intersections; to match the slicing of the interior, we slit $E_2 \times \mathbb{T}^3$ along $\theta_{24} = \theta_{13}$, $\theta_{23} = \theta_{14}$ and $\theta_{13} = \theta_{23}$, $E_3 \times \mathbb{T}^3$ along $\theta_{34} = \theta_{12}$, $\theta_{23} = \theta_{14}$ and $\theta_{12} = \theta_{13}$, and $E_4 \times \mathbb{T}^3$ along $\theta_{34} = \theta_{12}$, $\theta_{24} = \theta_{13}$ and $\theta_{12} = \theta_{23}$. This makes use of the slightly non-obvious fact that the limit of $\theta_{12} = \theta_{24} = \theta_{13}$ as $|\psi_{24}| \to 0$ is $\theta_{12} = \theta_{13}$. Finally, the circle above each

vertex is cut at one point; specifically, $V_3 \times \mathbb{T}^1$ is cut at $\theta_{24} = \theta_{13}$, $V_4 \times \mathbb{T}^1$ at $\theta_{14} = \theta_{23}$, and $V_2 \times \mathbb{T}^1$ at $\theta_{12} = \theta_{34}$. The cells are assigned orientations, which we do not list here†.

This leads to the table 1 of cells, boundaries, cycles, representatives of homology, and Betti numbers for Σ. Thus, using (43), the homology index of Σ is $\gamma = [0, 0, 0, 1, 5, 7, 1, 0, 1, 0, 1]$. In particular, the sum of the Betti numbers of Σ is 16, so perturbation leads to at least 16 critical points near Σ (if non-degenerate).

As a consistency check, table 2 displays the homology indices for all the critical sets for the problem of two fermions on four sites. There are six minima (e.g. with fermionic density distribution 0011), 12 circles of index 1 (e.g. $\frac{1}{2}\frac{1}{2}01$), eight two-tori of index 2 (e.g. $\frac{1}{3}\frac{1}{3}\frac{1}{3}1$ and $\frac{2}{3}\frac{2}{3}\frac{2}{3}0$), and the $\frac{2}{4}$ resonance Σ. The consistency check is that the Morse inequalities for this decomposition into critical subsets hold, i.e. that all elements of the final column are non-negative and the top one is 0.

5.3. The $\frac{1}{2} \times \frac{1}{3}$-resonance

This is given by the equations (29) and (30). To compute its homology index, note first that the submanifold R given by (29) is invariant under the gradient flow, because from (34) and (38),

$$\dot{\psi}_{ij} = -\frac{\partial W}{\partial \bar{\psi}_{ij}} = \left(\rho_i + \rho_j - \sum_{s=1}^{5} \rho_s^2\right)\psi_{ij} \tag{49}$$

which is zero when $\psi_{ij} = 0$. Furthermore, near the $\frac{1}{2} \times \frac{1}{3}$-resonance, R is normally hyperbolic, because at the resonance the normal components of the gradient flow evaluate to

$$\dot{\psi}_{12} = \tfrac{1}{6}\psi_{12} \tag{50}$$

$$\dot{\psi}_{34} = -\tfrac{1}{9}\psi_{34} \tag{51}$$

$$\dot{\psi}_{35} = -\tfrac{1}{9}\psi_{35} \tag{52}$$

$$\dot{\psi}_{45} = -\tfrac{1}{9}\psi_{45} \tag{53}$$

and their complex conjugates. This gives two unstable directions: ψ_{12} and $\bar{\psi}_{12}$. So it is enough to compute the homology index of the resonance in R and then shift the results up two dimensions.

In R, (29) imposes the restrictions

$$\rho_1 + \rho_2 = 1 \tag{54}$$

$$\rho_3 + \rho_4 + \rho_5 = 1. \tag{55}$$

Then the $\frac{1}{2} \times \frac{1}{3}$-resonance is the set of maxima of W in R, because $W = -\frac{1}{2}\sum \rho_s^2$ is maximized subject to these restrictions by $\rho_1 = \rho_2 = \frac{1}{2}$, $\rho_3 = \rho_4 = \rho_5 = \frac{1}{3}$. Thus it suffices to compute the Betti numbers β_i of the $\frac{1}{2} \times \frac{1}{3}$ resonance and then apply (43) to obtain its homology index in R. Now $\dim R = 10$, hence the homology index in the whole space is $\gamma_i = \beta_{12-i}$.

This reduces the problem to computing the Betti numbers of the resonance. We already showed that it is a submanifold and of dimension seven. It is also orientable because the subset where all $\psi_{ij} \neq 0$ is diffeomorphic to $D^2 \times \mathbb{T}^5$ and the rest is a union of cells of dimensions five or less. Thus we have $\beta_0 = \beta_7 = 1$ and $\beta_i = \beta_{7-i}$. The computation analogous to table 1 is tedious, however, and we were not able to determine any more Betti

† Many of the orientations can be inferred from table 1.

Table 1. Calculation of the Betti numbers β_i of the $\frac{2}{4}$ resonance Σ. i denotes 'dimension', 'Hom' denotes 'homology representatives', *pt* denotes 'point', and $\rightarrow$ means 'copy the entry from the same row of the preceding column'.

i	Cells C_i	Boundaries ∂C_i	Cycles	Hom	β_i
7	$\Delta \times I^5$	0	$\Delta \times I^5$	$\rightarrow$	1
6	$\Delta \times I^4\{\theta_{34} = \theta_{12}\}$	0	$\Delta \times I^4\{\theta_{34} = \theta_{12}\}$	$\rightarrow$	5
	$\Delta \times I^4\{\theta_{24} = \theta_{13}\}$	0	$\Delta \times I^4\{\theta_{24} = \theta_{13}\}$	$\rightarrow$	
	$\Delta \times I^4\{\theta_{23} = \theta_{14}\}$	0	$\Delta \times I^4\{\theta_{23} = \theta_{14}\}$	$\rightarrow$	
	$\Delta \times I^4\{\theta_{12} = \theta_{24}\}$	0	$\Delta \times I^4\{\theta_{12} = \theta_{24}\}$	$\rightarrow$	
	$\Delta \times I^4\{\theta_{13} = \theta_{23}\}$	0	$\Delta \times I^4\{\theta_{13} = \theta_{23}\}$	$\rightarrow$	
5	$\Delta \times I^3\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13}\}$	0	$\Delta \times I^3\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13}\}$	$\rightarrow$	7
	$\Delta \times I^3\{\theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\}$	0	$\Delta \times I^3\{\theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\}$	$\rightarrow$	
	$\Delta \times I^3\{\theta_{23} = \theta_{14}, \theta_{34} = \theta_{12}\}$	0	$\Delta \times I^3\{\theta_{23} = \theta_{14}, \theta_{34} = \theta_{12}\}$	$\rightarrow$	
	$\Delta \times I^3\{\theta_{34} = \theta_{12}, \theta_{13} = \theta_{23}\}$	0	$\Delta \times I^3\{\theta_{34} = \theta_{12}, \theta_{13} = \theta_{23}\}$	$\rightarrow$	
	$\Delta \times I^3\{\theta_{23} = \theta_{14}, \theta_{12} = \theta_{24}\}$	0	$\Delta \times I^3\{\theta_{23} = \theta_{14}, \theta_{12} = \theta_{24}\}$	$\rightarrow$	
	$\Delta \times I^3\{\theta_{12} = \theta_{24}, \theta_{13} = \theta_{23}\}$	0	$\Delta \times I^3\{\theta_{12} = \theta_{24}, \theta_{13} = \theta_{23}\}$	$\rightarrow$	
	$\Delta \times I^3\{\theta_{24} = \theta_{13} = \theta_{12}\}$	$E_3 \times I^3$	$\Delta \times I^3\{\theta_{24} = \theta_{13} = \theta_{12}\} - \Delta \times I^3\{\theta_{24} = \theta_{13} = \theta_{23}\}$	$\rightarrow$	
	$\Delta \times I^3\{\theta_{24} = \theta_{13} = \theta_{23}\}$	$E_3 \times I^3$			
	$\Delta \times I^3\{\theta_{23} = \theta_{14} = \theta_{13}\}$	$E_4 \times I^3$			
	$\Delta \times I^3\{\theta_{34} = \theta_{12} = \theta_{24}\}$	$E_2 \times I^3$			
4	$E_2 \times I^3$	0	$E_2 \times I^3$		1
	$E_3 \times I^3$	0	$E_3 \times I^3$		
	$E_4 \times I^3$	0	$E_4 \times I^3$		
	$\Delta \times I^2\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\}$	0	$\Delta \times I^2\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\}$	$\rightarrow$	
	$\Delta \times I^2\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13} = \theta_{23}\}$	$E_3 \times I^2\{\theta_{34} = \theta_{12}\}$			
	$\Delta \times I^2\{\theta_{34} = \theta_{12} = \theta_{24}, \theta_{23} = \theta_{14}\}$	$E_2 \times I^2\{\theta_{23} = \theta_{14}\}$			
	$\Delta \times I^2\{\theta_{34} = \theta_{12}, \theta_{23} = \theta_{14} = \theta_{13}\}$	$E_4 \times I^2\{\theta_{34} = \theta_{12}\}$			
	$\Delta \times I^2\{\theta_{24} = \theta_{13} = \theta_{12}, \theta_{23} = \theta_{14}\}$	$E_3 \times I^2\{\theta_{23} = \theta_{14}\}$			
	$\Delta \times I^2\{\theta_{24} = \theta_{13} = \theta_{23} = \theta_{12}\}$	$E_3 \times I^2\{\theta_{12} = \theta_{13}\}$			
	$\Delta \times I^2\{\theta_{34} = \theta_{12} = \theta_{24} = \theta_{13}\}$	$E_2 \times I^2\{\theta_{24} = \theta_{13}\} + E_3 \times I^2\{\theta_{34} = \theta_{12}\}$			
	$\Delta \times I^2\{\theta_{34} = \theta_{12} = \theta_{24}, \theta_{13} = \theta_{23}\}$	$E_2 \times I^2\{\theta_{13} = \theta_{23}\} + E_3 \times I^2\{\theta_{34} = \theta_{12}\}$			
	$\Delta \times I^2\{\theta_{24} = \theta_{13} = \theta_{23} = \theta_{14}\}$	$E_3 \times I^2\{\theta_{23} = \theta_{14}\} + E_4 \times I^2\{\theta_{24} = \theta_{13}\}$			
	$\Delta \times I^2\{\theta_{23} = \theta_{14} = \theta_{13}, \theta_{12} = \theta_{24}\}$	$E_3 \times I^2\{\theta_{23} = \theta_{14}\} + E_4 \times I^2\{\theta_{12} = \theta_{24}\}$			

Table 1. (Continued)

i	Cells C_i	Boundaries ∂C_i	Cycles	Hom	β_i
3	$E_2 \times I^2\{\theta_{24} = \theta_{13}\}$	0	$E_2 \times I^2\{\theta_{24} = \theta_{13}\}$		0
	$E_2 \times I^2\{\theta_{23} = \theta_{14}\}$	0	$E_2 \times I^2\{\theta_{23} = \theta_{14}\}$		
	$E_2 \times I^2\{\theta_{13} = \theta_{23}\}$	0	$E_2 \times I^2\{\theta_{13} = \theta_{23}\}$		
	$E_3 \times I^2\{\theta_{34} = \theta_{12}\}$	0	$E_3 \times I^2\{\theta_{34} = \theta_{12}\}$		
	$E_3 \times I^2\{\theta_{23} = \theta_{14}\}$	0	$E_3 \times I^2\{\theta_{23} = \theta_{14}\}$		
	$E_3 \times I^2\{\theta_{12} = \theta_{13}\}$	0	$E_3 \times I^2\{\theta_{12} = \theta_{13}\}$		
	$E_4 \times I^2\{\theta_{34} = \theta_{12}\}$	0	$E_4 \times I^2\{\theta_{34} = \theta_{12}\}$		
	$E_4 \times I^2\{\theta_{24} = \theta_{13}\}$	0	$E_4 \times I^2\{\theta_{24} = \theta_{13}\}$		
	$E_4 \times I^2\{\theta_{12} = \theta_{23}\}$	0	$E_4 \times I^2\{\theta_{12} = \theta_{23}\}$		
	$\Delta \times I\{\theta_{34} = \theta_{12} = \theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\}$	$E_2 \times I\{\theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\} + E_3 \times I\{\theta_{34} = \theta_{12}, \theta_{23} = \theta_{14}\}$			
	$\Delta \times I\{\theta_{24} = \theta_{13} = \theta_{23} = \theta_{14} = \theta_{12}\}$	$E_3 \times I\{\theta_{23} = \theta_{14}, \theta_{12} = \theta_{13}\} + E_4 \times I\{\theta_{24} = \theta_{13} = \theta_{12}\}$			
	$\Delta \times I\{\theta_{23} = \theta_{14} = \theta_{13}, \theta_{12} = \theta_{24} = \theta_{34}\}$	$E_2 \times I\{\theta_{23} = \theta_{14} = \theta_{13}\} + E_4 \times I\{\theta_{12} = \theta_{24} = \theta_{34}\}$			
	$\Delta \times I\{\theta_{12} = \theta_{24} = \theta_{13} = \theta_{23} = \theta_{34}\}$	$E_2 \times I\{\theta_{13} = \theta_{23} = \theta_{24}\} + E_3 \times I\{\theta_{12} = \theta_{23} = \theta_{34}\}$			
	$\Delta \times I\{\theta_{13} = \theta_{23} = \theta_{14} = \theta_{24}, \theta_{34} = \theta_{12}\}$	$E_3 \times I\{\theta_{34} = \theta_{12}, \theta_{23} = \theta_{14}\} + E_4 \times I\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13}\}$			
2	$E_3 \times I\{\theta_{34} = \theta_{12}, \theta_{23} = \theta_{14}\}$	0	$E_3 \times I\{\theta_{34} = \theta_{12}, \theta_{23} = \theta_{14}\}$	$\rightarrow$	1
	$E_2 \times I\{\theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\}$	0	$E_2 \times I\{\theta_{24} = \theta_{13}, \theta_{23} = \theta_{14}\}$		
	$E_4 \times I\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13}\}$	0	$E_4 \times I\{\theta_{34} = \theta_{12}, \theta_{24} = \theta_{13}\}$		
	$E_2 \times I\{\theta_{23} = \theta_{14} = \theta_{13}\}$	$V_3 \times I$	$E_2 \times I\{\theta_{23} = \theta_{14} = \theta_{13}\} + E_4 \times I\{\theta_{12} = \theta_{24} = \theta_{34}\}$		
	$E_4 \times I\{\theta_{12} = \theta_{24} = \theta_{34}\}$	$-V_3 \times I$			
	$E_3 \times I\{\theta_{12} = \theta_{13} = \theta_{34}\}$	$V_4 \times I$	$E_3 \times I\{\theta_{12} = \theta_{13} = \theta_{34}\} + E_2 \times I\{\theta_{13} = \theta_{23} = \theta_{24}\}$		
	$E_2 \times I\{\theta_{13} = \theta_{23} = \theta_{24}\}$	$-V_4 \times I$			
	$E_4 \times I\{\theta_{24} = \theta_{13} = \theta_{12}\}$	$V_2 \times I$	$E_4 \times I\{\theta_{24} = \theta_{13} = \theta_{12}\} + E_2 \times I\{\theta_{13} = \theta_{23} = \theta_{24}\}$		
	$E_2 \times I\{\theta_{13} = \theta_{23} = \theta_{24}\}$	$-V_2 \times I$			
	$\Delta \times pt$	$E_2 \times pt + E_3 \times pt + E_4 \times pt$			

Table 1. (Continued)

i	Cells C_i	Boundaries ∂C_i	Cycles	Hom	β_i
1	$V_3 \times I$	0	$V_3 \times I$		0
	$V_2 \times I$	0	$V_2 \times I$		
	$V_4 \times I$	0	$V_4 \times I$		
	$E_2 \times pt$	$V_3 \times pt - V_4 \times pt$	$E_2 \times pt + E_3 \times pt + E_4 \times pt$		
	$E_3 \times pt$	$V_4 \times pt - V_2 \times pt$			
	$E_4 \times pt$	$V_2 \times pt - V_3 \times pt$			
0	$V_3 \times pt$	0	$V_3 \times pt$	$\rightarrow$	1
	$V_2 \times pt$	0	$V_2 \times pt$		
	$V_4 \times pt$	0	$V_4 \times pt$		

Table 2. The Betti numbers for $\mathbb{C}P^5$, and the homology indices for all the critical subsets at the anti-integrable limit. The penultimate column gives the sum over the critical sets of the homology indices at dimension i, and the final column gives the excess of the alternating sum of the homology indices up to dimension i for the sum of the critical sets over that for the Betti numbers of $\mathbb{C}P^5$.

i	$\mathbb{C}P^5$	6 min	12 $\mathbb{T}^1$	8 $\mathbb{T}^2$	Σ	$\sum\gamma_i$	Excess
10	1	0	0	0	1	1	0
9	0	0	0	0	0	0	0
8	1	0	0	0	1	1	0
7	0	0	0	0	0	0	0
6	1	0	0	0	1	1	0
5	0	0	0	0	7	7	0
4	1	0	0	8	5	13	7
3	0	0	0	16	1	17	5
2	1	0	12	8	0	20	12
1	0	0	12	0	0	12	7
0	1	6	0	0	0	6	5

numbers with confidence. It would be particularly interesting to know them, as it could give insight into whether a product resonance can in some way be regarded as a (tensor) product.

5.4. Some general rules

The reduction of the computation of homology index of a product resonance to that of its homology generalizes in many cases.

Theorem 2. Given $r \in \mathbb{N}$ and positive integers P_j, Q_j, $j = 1, \ldots, r$, such that P_j/Q_j are distinct and $\sum n_j P_j/Q_j \neq 0$ whenever $n_j \in [-P_j, Q_j - P_j]$ is a collection of integers not all zero satisfying $\sum n_j = 0$, then the homology index γ of a $P_1/Q_1 \times \cdots \times P_r/Q_r$-resonance is given in terms of its Betti numbers β by

$$\gamma_i = \beta_{L-i} \tag{56}$$

where

$$L = 2\left(A + \prod_j \binom{P_j}{Q_j} - 1\right) \tag{57}$$

and A is the number of collections (n_j) above such that $\sum n_j P_j/Q_j > 0$.

Proof. Without loss of generality, let $F = \sum P_j$, $S = \sum Q_j$, and choose a decomposition of the network into subsets S_j of size Q_j, $j = 1, \ldots, r$. The product resonance lies in the submanifold R where $\psi_{s_1 \ldots s_F} = 0$ whenever $k_j := \sharp\{i : s_i \in S_j\} \neq P_j$ for some j. The submanifold R is invariant under the gradient flow because

$$\dot{\psi}_{s_1 \ldots s_F} = \left(\sum_i \rho_{s_i} - \sum_s \rho_s^2\right)\psi_{s_1 \ldots s_F}. \tag{58}$$

It is normally hyperbolic near the product resonance, because at the resonance

$$\sum_s \rho_s^2 = \sum_j P_j^2/Q_j \tag{59}$$

and

$$\sum_i \rho_{s_i} = \sum_j k_j P_j / Q_j. \tag{60}$$

Thus the 'exponent' from (58) is

$$\sum_i \rho_{s_i} - \sum_s \rho_s^2 = \sum_j (k_j - P_j) P_j / Q_j. \tag{61}$$

For a normal component to R, $k_j \neq P_j$ for some j; but $\sum_j k_j = \sum_j P_j$. Thus the number of unstable dimensions from R is $2A$.

The complex dimension of the space of $\psi_{s_1 \ldots s_F}$ with P_j suffices in S_j for each j is $\prod \binom{Q_j}{P_j}$. Thus R is $\mathbb{C}P^J$ with $J = \prod \binom{Q_j}{P_j} - 1$. Hence $L = 2(A + J)$. □

If there exists (n_j) not all zero such that $\sum_j n_j P_j / Q_j = 0$, then further analysis is required. The simplest example of this is the $\frac{1}{4} \times \frac{1}{2} \times \frac{1}{3}$-resonance.

We did not find a way yet to compute the homology of the general P/Q-resonance, nor the general product resonance.

6. At and near the integrable limit

In this section, we switch attention to the opposite limit in parameter space, namely the integrable limit $c^2 : t = 0$. Then all critical points have $u_s = 0$ for all s, and the energy is just a quadratic function of the fermionic variables. If $\{E_i\}_{i=1\ldots S}$ is the single fermion spectrum, then the F-fermion spectrum consists of all sums $\sum_k$ of F different E_i. There are $\binom{S}{F}$ choices. If all the $\sum_k$ are distinct, then arranging them in increasing order we obtain one critical point of the total energy for each choice, and the kth choice has index $2(k-1)$. This is because the energy is a Hermitian form, so the index increases by one complex dimension (i.e. by two real dimensions) at each critical point.

Often there are degeneracies. For example, for a one-dimensional chain, or ring with no preferred direction, the E_i come in degenerate pairs corresponding to wavenumbers $\pm k$ (except for the states with wavenumber 0 or π). This modifies the above picture, giving a critical $\mathbb{C}P^{m-1}$ if m F-fermion states are degenerate, instead of isolated critical points.

Non-degenerate critical points all persist on small perturbation, which includes the effect of adding small fermion–phonon coupling. Typical small perturbation breaks the degeneracy of critical $\mathbb{C}P^{m-1}$ manifolds ($m > 1$), however. Nonetheless, as they are normally hyperbolic for the gradient flow, they continue as invariant submanifolds of the gradient flow. Then the numbers of critical points on such a submanifold are constrained by the Morse inequalities for $\mathbb{C}P^{m-1}$. For $c^2 : t$ small, the simplest scenario is that precisely $\binom{S}{F}$ equilibrium states persist, one of each even index between 0 and $2\left(\binom{S}{F} - 1\right)$. This is the situation that we would expect unless there are symmetries which preserve some of the $\mathbb{C}P^{m-1}$ or some submanifolds within them.

7. Effects of spin, magnetic fields and electron–electron interaction

In this section, instead of spinless fermions we treat spin $\frac{1}{2}$ fermions, which we call *electrons*. This means that we have creation and annihilation operators $a^\dagger_{s\sigma}, a_{s\sigma}$, for each $s \in \mathcal{S}$ and

$\sigma \in \{\uparrow, \downarrow\}$, number operators

$$n_{s\sigma} = a^{\dagger}_{s\sigma} a_{s\sigma} \tag{62}$$

and in the Hamiltonian (1) we take

$$n_s = n_{s\uparrow} + n_{s\downarrow} \tag{63}$$

and Δ to be an off-diagonal Hermitian operator which allows electrons of both spins to hop, e.g.

$$\Delta = \sum_{\langle s,r\rangle,\sigma} a^{\dagger}_{r,\sigma} a_{s,\sigma} + a^{\dagger}_{s,\sigma} a_{r,\sigma} \tag{64}$$

where $\langle s, r\rangle$ represent nearest-neighbour pairs. The space of electron states for $F_\uparrow$ up-electrons and $F_\downarrow$ down-electrons on S sites has one complex dimension for each way of arranging them with at most one up-electron and one down-electron on each site. As usual, two states can be regarded as equivalent if they differ only by multiplication by a non-zero complex number.

In contrast to the spinless case, we can have degenerate local minima for spin $\frac{1}{2}$ electrons at the anti-integrable limit. For example, any anti-integrable configuration with $P > 1$ polarons (singly occupied sites) and $D := |F_\uparrow - F_\downarrow| \neq P$, has a spin degeneracy. There are

$$J = \begin{pmatrix} P \\ \frac{P-D}{2} \end{pmatrix} \tag{65}$$

ways of distributing the spins among the polarons. This gives a critical $\mathbb{C}P^{J-1}$ of index 0.

In [AAR] (and [BM1]), this spin degeneracy was not a problem, because it was possible to treat electronic configurations as equivalent if they give the same electronic density distribution, regardless of spin. Then the critical $\mathbb{C}P^{J-1}$ collapses to a non-degenerate critical point. Since this can be done for $t \neq 0$ too, the critical point has a locally unique continuation, and so the $\mathbb{C}P^{J-1}$ continues as a submanifold of critical points.

Various effects, however, can break the spin degeneracy. The simplest is the Zeeman effect in a magnetic field. This adds $\sum_s B_s(n_{s\uparrow} - n_{s\downarrow})$ to the Hamiltonian. If the magnetic field B_s is uniform, then [AAR] managed to circumvent the spin degeneracy again, by reducing to a problem in electronic density only, as the Zeeman effect is the same on all states with given numbers $F_\uparrow$ and $F_\downarrow$ of electrons. But we can treat the case of non-uniform field just as easily. For uniform magnetic field the critical $\mathbb{C}P^{J-1}$ of index 0 just discussed, is unchanged, except in energy. It continues for $t \neq 0$ and/or perturbations from uniform magnetic field, to an invariant $\mathbb{C}P^{J-1}$ for the gradient flow, but no longer consists entirely of critical points. But, by the same arguments as in the rest of this paper, we deduce that it contains at least J critical points if non-degenerate.

A particularly interesting and important effect which breaks the spin degeneracy is electron–electron interaction. For example, suppose we add a Hubbard interaction $\sum_s U n_{s\uparrow} n_{s\downarrow}$ to the Hamiltonian. This does not affect the positions of the critical points at the anti-integrable limit, apart from changing their energy (which is raised by U for each bipolaron). But for $t \neq 0$, the critical $\mathbb{C}P^{J-1}$ of index 0 above continues to an invariant $\mathbb{C}P^{J-1}$ for the gradient flow, and again we deduce that it contains at least J critical points if non-degenerate. However, the Hubbard interaction preserves the overall $SU(2)$ symmetry of simultaneous rotation of all spins. Thus, for example, the familiar triplet states degeneracy for two-electron states is preserved, which leads to critical $\mathbb{C}P^2$ submanifolds, and there are analogous degeneracies for more electrons.

Similar results hold for critical subvarieties of index greater than 0. In particular, we obtained by explicit calculation in June 1992 an equilibrium state which Aubry called a 'spin resonant bipolaron' [Au1].

Other forms of electron–electron interaction can also be treated, for example, a repulsion between electrons on neighbouring sites (even in the spinless case): $\sum_s U_2 n_s n_{s+1}$.

8. Discussion

We have located all the critical points of the energy for the adiabatic Holstein model at the integrable and anti-integrable limits, and shown how they behave near both these limits. The key idea was, in contrast to [AAR], first to minimize the energy with respect to the phonon variables and then to study the resulting function of the fermionic variables. We did this in detail in the case of spinless fermions, but showed that the case with spin is similar. The number of critical points at the anti-integrable limit vastly exceeds that at the integrable limit, many of them forming connected critical sets. We were able to show that many critical points survive perturbation to small fermion hopping, by studying the homology index of the isolated critical sets with respect to the gradient flow of the energy, though the calculation of the minimum number that must survive from each isolated critical set remains to be completed. As the electron hopping amplitude increases there must be a massive sequence of bifurcations annihilating the excess critical points (generically in pairs but pitchforks can also be expected under various symmetries). It would be very interesting to study the bifurcations numerically.

The use of symmetries would lead to simplifications in the analysis, which we did not discuss. In particular, the intersection of each of our critical sets at the anti-integrable limit with a symmetry class will have its own Betti numbers, and hence an associated number of critical points with that symmetry will persist. Sometimes restriction to a symmetry class lifts the degeneracy completely, so the anti-integrable critical sets with given symmetry are isolated non-degenerate critical points in their symmetry class, which persist straightforwardly by the implicit function theorem. This idea has been used to good effect by [Pr].

Our analysis has been for finite systems. An important open question is whether an isolating neighbourhood for a given resonance can be found for a range of electron hopping amplitude t which is uniform in the system size.

Instead of using the gradient flow of the energy function, an alternative approach to finding critical points persisting from an isolated critical submanifold is to choose a local coordinate system for which it is a graph, then prove persistence of each of its points to a point which is critical with respect to normal variations, which is easy to do uniformly in the size of the network (cf variational approach to Melnikov's method for continuation of periodic orbits from a manifold of degenerate ones, proposed in [Mac] and developed in [Au2, AMS]), and deduce that the resulting function on the unperturbed critical set has at least some number of critical points. To analyse the effect of perturbation on critical sets which are not submanifolds, however, we see no replacement for the gradient flow approach.

We have generalized the ideas to allow electron–electron interaction. An interesting question is whether our procedure extends to the case where the phonon variables are made quantum mechanical operators. If so, the results would provide a mathematical framework in which to develop Aubry's ideas on high T_c superconductors [Au1]. The quantum mechanical analogue of our elimination of the phonon variables is the Lang–Firsov transformation, as discussed by Aubry, and it would be interesting to see whether this route can be pursued.

Acknowledgments

We thank Serge Aubry and Christophe Golé for useful comments. This work was supported by the Nuffield Foundation, a British Council Alliance with the Laboratoire Léon Brillouin at Saclay where the paper was started in 1993, the Institut des Hautes Etudes Scientifiques at Bures-sur-Yvette, and the UK Science and Engineering Research Council. It also benefitted from participation in an EC network on Stability and universality in classical mechanics and a British Council Alliance between Cambridge and Dijon.

References

[AMS] Ahn T, MacKay R S and Sepulchre J-A 1998 *Dynamics of Relative Phases of Multi-Breathers* in preparation

[An] Anosov DV 1988 Smooth dynamical systems *Dynamical Systems I* part II, ed D V Anosov and V I Arnol'd (Berlin: Springer)

[Au1] Aubry S 1993 Bipolaronic charge density waves and polaronic spin density waves in the Holstein–Hubbard model *Proc. ECRYS-93 (Carry-le-Rouet, France)*

[Au2] Aubry S 1997 Breathers in nonlinear lattices: existence, linear stability and quantization *Physica* D **103** 201–50

[AAR] Aubry S, Abramovici G and Raimbault J L 1992 Chaotic polaronic and bipolaronic states in the adiabatic Holstein model *J. Stat. Phys.* **67** 675–780

[BM1] Baesens C and MacKay R S 1994 Improved proof of existence of chaotic polaronic and bipolaronic states for the adiabatic Holstein model and generalisations *Nonlinearity* **7** 59–84

[BM2] Baesens C and MacKay R S 1996 Effect of temperature on polaronic and bipolaronic structures of the adiabatic Holstein model *J. Stat. Phys.* **85** 471–88

[BM3] Baesens C and MacKay R S 1997 Finite coherence length for equilibria of the adiabatic Holstein model *J. Math. Phys.* **38** 2104–14

[Bo] Bott R 1954 Non-degenerate critical manifolds *Ann. Math.* **60** 249–61

[Co] Conley C 1978 Isolated invariant sets and the Morse index *Conf. Board Math. Sci.* **38** (Providence, RI: American Mathematical Society)

[Fe] Fenichel N 1971 Persistence and smoothness of invariant manifolds for flows *Indiana Univ. Math. J.* **21** 193–226

[Fl] Floer A 1989 Witten's complex and infinite dimensional Morse theory *J. Diff. Geom.* **30** 207–21

[GH] Greenberger M J and Harper J R 1981 *Algebraic Topology* (Reading, MA: Addison Wesley)

[Ho] Holstein T 1959 Studies of polaron motion, part II. The 'small' polaron *Ann. Phys.* **8** 343–89

[Mac] MacKay R S 1996 Dynamics of networks: features that persist from the uncoupled limit *Stochastic and Spatial Structures of Dynamical Systems* ed S J van Strien and S M Verduyn Lunel (Amsterdam: North Holland) pp 81–104

[Mat] Mather J 1973 Stratifications and Mappings *Dynamical Systems* ed M Peixoto (New York: Academic) pp 195–232

[NS] Nash C and Sen S 1983 *Topology and Geometry for Physicists* (New York: Academic)

[Pr] Proville L 1998 Structures polaroniques et bipolaroniques dans le modèle de Holstein–Hubbard adiabatique à deux électrons et ses extentions *PhD Thesis* Orsay, France

[Sa] Salamon D A 1990 Morse theory, the Conley index and Floer homology *Bull. Lond. Math. Soc.* **22** 113–40

[Sc] Schwartz J T 1969 *Nonlinear Functional Analysis* (New York: Gordon and Breach)

[Th] Thurston W P 1997 *Three-Dimensional Geometry and Topology* vol 1, ed S Levy (Princeton, NJ: Princeton University Press)