On the dynamics of topological solitons

Speight, James Martin

On the Dynamics of Topological Solitons

James Martin Speight

A thesis submitted for the degree of
Doctor of Philosophy.

Department of Mathematical Sciences
University of Durham
Durham DH1 3LE.

June 1995

Preface

This thesis derives from work done by the author between October 1992 and June 1995 at the University of Durham, financially supported by a Particle Physics and Astronomy Research Council research studentship. No part of it has been submitted previously for any degree at any university.

With the exception of the reviews of chapter 1 and section 5.1, it is believed that the material in this thesis is original work. Chapter 2 results from collaboration with Richard Ward, and most of its contents appeared in a joint paper [1] published in *Nonlinearity*. The rest is the author's work alone. Chapters 3 and 4 are based on papers published in *Physical Review D* [2] and the *Journal of Mathematical Physics* [3] respectively, while the material in chapter 5 has not previously appeared.

It is my pleasure to thank Richard Ward, my supervisor, for his guidance and encouragement, and for getting my research off to a brisk start by collaborating with me on my first project. I am also greatly indebted to Bernd Schroers who has always been very generous with his time, and with whom I have had many valuable conversations. There is a welcome tradition in Durham mathematics theses of thanking one's friends for their support. In my case, I would like to thank Alistair MacIntyre and Joseph Mulvey for their infectious enthusiasm, Darren Wilkinson for helping me out of numerous Unix crises, and Benjamin Jordan for organizing enjoyable diversions. Finally, I am grateful to Trevor Samols for reproducing his data on critical vortex scattering at very short notice.

published without his prior written consent, and information derived from it should be acknowledged.

Abstract

On the Dynamics of Topological Solitons, J.M. Speight.

This thesis investigates the dynamics of lump-like objects in non-integrable field theories, whose stability is due to topological considerations. The work concerns three different low dimensional ($(1+1)$- and $(2+1)$-dimensional) systems and addresses the questions of how the topology and metric structure of physical space, the quantum mechanics of the basic field quanta and intersoliton interactions affect soliton dynamics.

In chapter 2 a sine-Gordon system in discrete space, but with continuous time, is presented. This has some novel features, namely a topological lower bound on the energy of a kink and an explicit static kink which saturates this bound. Kink dynamics in this model is studied using a geodesic approximation which, on comparison with numerical simulations, is found to work well for moderately low kink speeds. At higher speeds the dynamics becomes significantly dissipative, and the approximation fails. Some of the dissipative phenomena observed are explained by means of a dispersion relation for phonons on the spatial lattice. Chapter 3 goes on to quantize the kink sector of this model. A quantum induced potential called the kink Casimir energy is computed numerically in the weak coupling approximation by quantizing the lattice phonons. The effect of this potential on classical kink dynamics is discussed. Chapter 4 presents a study of the low-energy dynamics of a $\mathbb{C}P^1$ lump on the two-sphere in the geodesic approximation. By considering the isometry group inherited from global

symmetries of the model, the structure of the induced metric on the unit-charge moduli space is so restricted that the metric can be calculated explicitly. Some totally geodesic submanifolds are found, and the qualitative features of motion on these described. The moduli space is found to be geodesically incomplete. Finally, chapter 5 contains an analysis of long range intervortex forces in the abelian Higgs model, a massive field theory, extending a point source approximation previously only used in massless theories. The static intervortex potential is rederived from a new viewpoint and used to model type II vortex scattering. Velocity dependent forces are then calculated, providing a model of critical vortex scattering, and leading to a conjecture for the analytic asymptotic form of the metric on the two-vortex moduli space.

Contents

Chapter 1

Introduction

1.1 Context

Topological solitons may provide the key to understanding phenomena observed in a diverse range of physical systems. Their applications span the length scale, from the structure of the atomic nucleus to cosmological matter distribution, taking in aspects of condensed matter physics (superconductivity for example) and biophysics in between. They are an inherently nonlinear phenomenon, so the study of their properties is a difficult but fascinating mathematical challenge.

The word "soliton" means different things to different people, a less than ideal situation for a mathematical term to find itself in. There are essentially two kinds, integrable and topological, the latter being the objects of study here. There is some common ground: all solitons are spatially localized, stable, lump-like solutions of nonlinear partial differential equations. Equally, there are some marked differences. Integrable solitons owe their stability to local details of the evolution equation. There is a trade-off between dispersive and nonlinear sharpening terms, allowing lumps which neither spread out nor collapse. This balance is quite delicate in the sense that perturbing the evolution equation by adding some arbitrary term will almost certainly destroy it. So integrable systems are very special and comparatively rare. They have

been intensely studied, and a rich if rather obscure mathematical structure has been found. Infinitely many independent quantities conserved under the time evolution can be constructed, and for each system there is a nonlinear analogue of the Fourier Transform, called the Inverse Scattering Transform, by means of which the initial value problem can, at least in principle, be solved. Consequently, a great number of nontrivial solutions are known explicitly, exactly. Unfortunately, all such exact solutions reveal very simple soliton dynamics. For example, on collision two solitons always scatter elastically, essentially suffering no more than a phase shift. Such solitons are of very limited relevance to particle theory because their behaviour is over-simple, but also, more importantly, because integrability appears incompatible with Special Relativity except in the very restrictive case of a one-dimensional Universe (that is $(1+1)$-dimensional space-time). In fact there are no known Lorentz invariant integrable field theories in higher dimensions.

All this should be contrasted with the generic properties of topological solitons. These owe their stability to the most global aspects of the system – its topology. When the conditions that the field (the dependent variable of the partial differential equation) be single-valued and have finite energy are imposed, we obtain restrictions on its boundary behaviour which imply that there must be a solitonic lump somewhere in space. Stability tends to be very robust and does not depend critically on the field's self interaction. Consequently, it is much easier to find systems possessing observed symmetries of nature, and in particular there is a wealth of relativistic models which have topological solitons in any number of dimensions. The price paid for this freedom is the loss of the rich mathematical structure found in integrable systems. There are only a few conserved quantities (energy, momentum, angular momentum – all those observed in nature) and no Inverse Scattering Transform. As a result, exact solutions are rather more scarce, usually restricted to static solutions and Lorentz transformations thereof. Often even the static problem cannot be solved exactly. Soliton dynamics can usually only be studied using numerical simulation and

analytic approximation techniques, but these reveal a richer diversity of behaviour than that exhibited by integrable solitons, as one would expect.

So, although they are harder to calculate with, topological solitons are substantially more relevant to particle theory than are integrable solitons, and it is this which motivates us to attempt to understand their underlying mathematics. They have two main areas of application. First, the Standard Model is a gauge theory with gauge group $SU(3) \times SU(2) \times U(1)$. An essential ingredient of the model is the Higgs mechanism [4] whereby one postulates the existence of scalar bosons whose vacua are continuously degenerate and constitute a nontrivial gauge orbit (no such particles have yet been observed). At low energies, the Higgs bosons must choose a particular vacuum, leaving the system invariant only under the $U(1)$ stabiliser of electromagnetism, a subgroup of the $SU(2) \times U(1)$ electroweak gauge group. This process is called spontaneous symmetry breaking, and is necessary in order to introduce masses for the weak gauge bosons ($W^{\pm}, Z^0$), because gauge symmetry is ostensibly incompatible with massive gauge quanta, and also for the fermions of the theory, because the electroweak gauge group acts chirally on the fermion fields. At high temperatures, thermal fluctuations produce a single effective vacuum, often called the "false vacuum," of the Higgs bosons, which is itself a completely degenerate gauge orbit, so that full gauge symmetry is recovered. The idea is, then, that as the Universe expanded and cooled it underwent a phase transition (the electroweak transition) from the hot symmetric phase to the cold asymmetric phase we now observe. Such a phase transition may be preceded by another where some Grand Unified gauge group is broken to the $SU(3) \times SU(2) \times U(1)$ of the Standard Model. If so, this GUT transition would occur at different times in different regions of space, causing bubbles of true vacuum to expand in the false vacuum. Assuming the vacuum in each bubble is chosen randomly, where several bubbles meet the Higgs field may wind its way around the space of degenerate vacua in some topologically nontrivial way, forming a topological defect (soliton) and trapping substantial amounts of energy. This process

is called the Kibble mechanism [5] and can lead to the formation of domain walls, cosmic strings and monopoles, depending on the details of the symmetry breaking. Such large conglomerations of energy would have very significant gravitational effects, providing a natural mechanism for the formation of large scale inhomogeneities in the cosmological matter distribution, as are observed, for example, in the cosmic microwave background.

The above is an example of an application where topological solitons were found in a system already proposed for quite independent reasons. On the other hand, topological solitons behave in such a particle-like fashion that it is natural to attempt to use them to model particle interactions directly, by construction. They have a conserved topological charge, which may be identified, for example, with baryon number, and there exist antisolitons with negative charges, the images of solitons under the parity transformation. A soliton-antisoliton pair may annihilate emitting an oscillating wavetrain, often called radiation, a process highly reminiscent of particle-antiparticle annihilation. Solitons undergo nontrivial scattering and may form bound states. All this behaviour is present at the level of *classical* field theory. Following standard procedure, small amplitude travelling waves become, on quantization, the basic particle quanta of the system. The solitons then represent new particle degrees of freedom, beyond the reach of perturbation theory. Skyrme's bold idea (which actually predates the Higgs mechanism) was that protons and neutrons are solitons in a field theory whose basic field quanta are pions ($\pi^{\pm}, \pi^{0}$), carriers of the strong nuclear force [6]. This is the second area of application, nuclear physics. Although the Skyrme model's predictions are in imperfect agreement with experimental data, it has one great advantage over more phenomenological approaches: its all embracing predictive nature. With only two free parameters it attempts to model all nucleon interactions. The last few years have seen a revival in the Skyrme model's fortunes due to the observation that it may well emerge as a low energy effective formulation of QCD [7]. Skyrme's idea represents a belief that topological solitons are to be found as real particles in

nature, that the particle-like properties constitute more than just an analogy. Of course, the term soliton was coined precisely because of these properties, in analogy with the fundamental particles (proton, electron etc.). Many mathematicians regard elastic scattering as a defining characteristic of solitons, thereby restricting its use to the integrable type. Lumps which scatter inelastically are then called "solitary waves." There is a certain irony in this terminology: the objects called solitons often do *not* possess key particle-like characteristics (antisolitons, pair annihilation etc.) but *do* describe wave phenomena (ultrashort optical pulses in fibre optic cables and water waves in shallow channels for example) while the objects called waves can be distinctly particle-like, and not behave remotely like waves (which do not tend to bind together, for instance). So this distinction is rather unhelpful from a particle theory standpoint, and henceforth will not be made.

Topological solitons are, then, relevant to real physics. However, as already explained, their dynamics is very difficult to study because modern techniques developed for solving nonlinear partial differential equations require the property of integrability, not present in the systems of interest. Even numerical simulation of the field equations of realistic models in $(3+1)$-dimensions is extremely difficult and costly in computer time. One is therefore led to study simpler systems which have the key topological features of more physical models, in the hope of understanding the underlying mechanisms involved in soliton motion and interaction. Skyrme himself studied the sine-Gordon model as a warm-up for the full Skyrme model [8]. In this spirit, this thesis addresses, by looking at tractable examples, a number of mathematical questions regarding the dynamics of topological solitons:

1. How does changing the metric structure and topology of physical space affect the motion of a soliton? In chapter 2 we study a solitonic system defined on a space consisting of a discrete collection of points $\mathbb{Z}$, while in chapter 4 physical space is a two-sphere S^2.

2. How does the quantum mechanics of the basic particle quanta of a model affect

its solitons (chapter 3)?

3. How can one analyze the interaction of well separated solitons when even the static soliton solution is not available (chapter 5)?

Of course, these questions are not new, and it is hard to imagine ever obtaining complete answers to them. The work concerns three different systems. In each case we are interested in the system purely for its mathematical properties, so possible physical applications receive minimal discussion.

1.2 Basic concepts

Each model considered will be defined using the Lagrangian formalism of classical field theory, so that the field equations may be generated from an action principle, and symmetries of the system are manifest. There is a canonical energy functional, the Noether charge associated with time translation invariance, which gives the field immediate physical significance. One can define canonical field momenta and transform to the Hamiltonian formalism, the standard starting point for quantizing the model. So once the action is specified there are well defined procedures for extracting the physical content of the model. Not all nonlinear partial differential equations can be derived from an action principle, but in the absence of such a principle it is unclear how to invest the dependent variable (the "field") with physical meaning. Such equations are not of interest to us, and we do not regard them as field equations in any meaningful sense.

We take this opportunity to define some of the terms which will recur throughout this thesis. Denote physical space X, and a general point in space $x \in X$, so that space-time is $X \times \mathbb{R}$, and a space-time event is denoted (x, t). Then a configuration is a map from physical space to Φ, some target space, $\phi : X \to \Phi$, while a field is a continuous time sequence of configurations, that is, a map from space-time to the target space, $\phi^\star : X \times \mathbb{R} \to \Phi$. So we imagine the time evolution of a field as

a smooth progression through a sequence of configurations. The action S is a real valued functional on the space of all fields, and the field equations (Euler-Lagrange equations) are obtained by demanding that $\phi^\star$ be a local extremal of this functional, $\delta S/\delta\phi^\star = 0$. All systems of interest are time translation invariant. Explicitly, if $\phi^\star$ is a field and $\phi^\star_\tau$ is its image under an active time translation by τ, $\phi^\star_\tau(x,t) := \phi^\star(x,t+\tau)$, then $S[\phi^\star] = S[\phi^\star_\tau]\ \forall\, \phi^\star, \tau$. This continuous symmetry leads, by Noether's theorem, to a conserved quantity canonically identified with energy, E. Let $\dot\phi^\star$ be the time derivative of $\phi^\star$, a map from $X \times \mathbb{R}$ to the tangent bundle of Φ, and ϕ_t and $\dot\phi_t$ be $\phi^\star$ and $\dot\phi^\star$ evaluated on the t time slice. Then the energy is a real valued, positive definite functional $E[\phi_t, \dot\phi_t]$, and is conserved in the sense that it is independent of t provided $\phi^\star$ is a solution of the field equations (a local extremal of S). In some field theories, E depends on higher time derivatives of $\phi^\star$, but such systems are not relevant to our considerations. From the energy one can extract the potential energy, a functional on the space of configurations, $V[\phi] = E[\phi, 0]$, and the kinetic energy, $T[\phi, \dot\phi] = E[\phi, \dot\phi] - V[\phi]$.

We finally arrive at the primary object of interest, configuration space Q, defined to be the space of configurations with finite potential energy. Each finite energy configuration is a point in this space and a solution of the field equations is a continuous sequence of such configurations, that is, a trajectory in Q. The time evolution generated by the field equations cannot lead the field to leave configuration space because, kinetic energy being positive definite, this would inevitably violate conservation of energy. The functional $V[\phi]$ may be thought of as defining a potential relief on configuration space. More subtly, the kinetic energy T is, for each $\phi \in Q$, a bilinear map from $\mathrm{T}_\phi Q \times \mathrm{T}_\phi Q$ to the positive real numbers (where $\mathrm{T}_\phi Q$ is the tangent space of Q at ϕ), and so defines a *metric* on Q, at least formally (this metric is not always well defined). Our picture of field dynamics is that of a notional point particle describing a trajectory in an infinite dimensional configuration space under the influence of a potential V and a metric inherited from T. Let us summarize what we have so far:

- Physical space X.
- Target space Φ.
- Potential energy V, a functional on the space of all maps $\phi : X \to \Phi$.
- Configuration space $Q = \{\phi \,|\, V[\phi] \text{finite}\}$.

It is the topology of Q which is crucial to the existence of solitons, in particular, whether Q is a connected space. To go further it is helpful to be a little more specific about X. For the purpose of discussion, let it be the d-dimensional Euclidean space $\mathbb{R}^d$. Then the functional V is an integral over $\mathbb{R}^d$ of some function of ϕ and its space derivatives $\nabla\phi$ called the potential energy density $\mathcal{V}(\phi, \nabla\phi)$,

$$V[\phi] = \int d^d\mathbf{x}\, \mathcal{V}(\phi, \nabla\phi). \tag{1.1}$$

For a configuration ϕ to have finite energy (and hence be an element of Q), $\mathcal{V}$ must vanish on the boundary of space, denoted ∂X, otherwise this integral will diverge[1]. This places restrictions on ϕ which, depending on $\mathcal{V}$, may lead to the existence of solitons. This can happen in two ways.

Let the vacuum manifold Φ_0 be the subspace of Φ on which $\mathcal{V}(\phi, \mathbf{0})$ attains its absolute minimum, always normalized to zero. Any constant map $\phi(\mathbf{x}) = \phi_0 \in \Phi_0$ is a trivial static solution of the model with vanishing energy, called a vacuum. Then the first way that solitons can arise is if Φ_0 has a nontrivial $(d-1)$-th homotopy group, $\pi_{d-1}(\Phi_0)$. The m-th homotopy group of a connected topological space Y is defined [9] as the set of equivalence classes of maps $y : S^m \to Y$, where two maps y_1 and y_2 are equivalent if one is continuously deformable into the other, that is, if there exists a continuous map $y_{12} : S^m \times [0,1] \to Y$ (called a homotopy) such that $y_{12}(0) = y_1$ and $y_{12}(1) = y_2$. The only exception is $\pi_0(Y)$ which concerns maps from

[1]Strictly, this is only true if one demands that the configuration should have a well-defined limit as $|\mathbf{x}| \to \infty$. We will henceforth always assume that this regularity condition is satisfied.

the space consisting of a single point $\{0\}$ to Y, rather than from S^0, which would more naturally be defined as $\mathbb{Z}_2 = \{-1, 1\}$. The group composition law is induced by the composition of maps, but the group structure shall not concern us. A trivial homotopy group consists of a single class so that all maps are deformable to the trivial map where the whole m-sphere is mapped to a single point in Y (precisely which point is irrelevant since Y is connected). In particular, π_0 of any connected space is trivial and so, again exceptionally, we extend its definition to disconnected spaces. Now, the boundary of space ∂X is homeomorphic to the $(d-1)$-sphere, and finite V demands that $\mathcal{V}$ vanishes on ∂X, so any finite energy configuration ϕ defines a map from S^{d-1} to Φ_0. If $\pi_{d-1}(\Phi_0)$ is nontrivial then Q is partitioned into disjoint sectors, one for each homotopy class. A configuration in one sector can never be deformed to one in another sector by any continuous process (time evolution, for example). Most field theories have trivial $\pi_{d-1}(\Phi_0)$, but some have $\pi_{d-1}(\Phi_0) = \mathbb{Z}$, and these are particularly interesting from our viewpoint. To see that this leads to the existence of solitons, that is, localized lumps of energy, we introduce a radial coordinate $r \in \mathbb{R}_+$, the positive real line (including 0) and decompose X as $X \cong \mathbb{R}_+ \times S^{d-1}$. Then the map $S^{d-1} \to \Phi_0$ defined by a finite energy configuration ϕ is $\phi_\infty := \lim_{r\to\infty} \phi$. Consider the case where ϕ_∞ does not lie in the trivial homotopy class (is not deformable to the trivial constant map). Then there is no way for ϕ to remain on Φ_0 throughout X because if it did, since ϕ is single valued at $r = 0$, the radial dependence of ϕ would define a homotopy between the map ϕ_∞ and the constant map $\phi|_{r=0}$, lying in a different homotopy class, a contradiction. Hence ϕ must leave Φ_0 at least in some region of X, and this generically will result in a lump, or a number of lumps. Such lumps might not be static solutions, they may expand indefinitely or collapse inwards whereupon a singularity forms. However, there are many systems in which there are stable lumps, the abelian Higgs and Yang-Mills-Higgs systems being examples.

The other way topology can lead to the existence of lumps arises as follows. It could be that in order for $V[\phi]$ to be finite, ϕ must tend to the same value in Φ on all

∂X. This could happen if Φ_0 consists of a discrete set of points and ∂X is connected, or at the opposite extreme, if $\Phi_0 \equiv \Phi$, that is the vacuum manifold is the entire target space, but the $\nabla\phi$ dependence of $\mathcal{V}$ (such a $\mathcal{V}$ is necessarily ϕ independent of course) forces $\nabla\phi$ to vanish as $r \to \infty$. Since ϕ takes a single constant value on ∂X, it may be regarded as a map from the one point compactification $\mathbb{R}^d \cup \{\infty\} \cong S^d$ to Φ, that is, one may identify the whole boundary of space as a single point $\{\infty\}$. The critical homotopy group is now $\pi_d(\Phi)$. Each finite energy map must lie in one of its classes and once again, a configuration is trapped within its class by continuity of time evolution. If ϕ does not lie in the trivial homotopy class, it is not deformable to a constant map, so $\nabla\phi \neq \mathbf{0}$ at least in some region of X and in that region $\mathcal{V} > 0$, leading to a lump or lumps. In addition, if Φ_0 is discrete, ϕ must also leave Φ_0 somewhere in X, also contributing to $\mathcal{V}$. The same caveat applies: such lumps might not be stable static solutions, but stable examples can be found (Skyrmions, $\mathbb{C}P^1$ lumps etc.).

Condensed matter theorists are careful to distinguish between these two types of topological defect. Those which owe their existence to the nontriviality of $\pi_{d-1}(\Phi_0)$ they call monopoles while those due to nontrivial $\pi_d(\Phi)$ they call textures. The distinction seems a useful one. In both cases the key feature is that finite energy configurations have a well defined degree (also called topological charge), the degree of the map $\phi_\infty : S^{d-1} \to \Phi_0$ for monopoles and the degree of the map $\phi : \mathbb{R}^d \cup \{\infty\} \to \Phi$ for textures. Configuration space Q is a disjoint union of homotopy classes Q_n labelled by this degree, an integer n, so that $\pi_0(Q) = \mathbb{Z}$. Configurations homotopic to the vacuum (or vacua) have zero degree, and Q_0 is called the vacuum sector of the theory. By convention Q_1 is called the soliton sector and Q_{-1} the antisoliton sector.

There is one final space of interest to us, the moduli space M, a subspace of Q. This is the space of minimals of the potential energy functional $V[\phi]$. The field variational problem reduces to the variational problem for V in the case of static fields, so M is the space of (perhaps marginally) stable static solutions of the field

equations. Since $M \subset Q$, it too is a disoint union of homotopy classes M_n, each one the equipotential surface of minimum energy in Q_n, so that M_n is a finite dimensional subspace of the infinite dimensional charge n sector. Note that $M_{\pm 1}$ are non-empty if stable solitons exist, but it is possible that $M_n = \emptyset$ if $|n| > 1$ even so. Given one static solution $\phi \in M_n$, one can generate a continuous family of other such solutions by acting on ϕ with continuous symmetries of the model (translations, rotations and internal symmetries) since these leave V invariant and, being continuous, cannot lead out of M_n. (Discrete symmetries, by contrast, may generate new static solutions outside M_n. An example is the parity transformation which maps M_n onto M_{-n}.) Such a symmetry orbit will usually[2] exhaust $M_{\pm 1}$, but often will *not* exhaust higher moduli spaces. So there are often zero modes (tangent vectors to M) not associated with any symmetry of the model, a theoretically very interesting situation. Thinking of V as providing Q_n with a potential relief, M_n is the finite dimensional flat valley floor on which V takes its absolute minimum value on Q_n. Consider the initial value problem with $\phi(0) \in M_n$ and $\dot{\phi}(0) \in \mathrm{T}_{\phi(0)} M_n$ with small norm, that is low kinetic energy. The subsequent trajectory in Q_n representing the solution of the field equations must remain close to M_n, because departure from the moduli space entails climbing the valley walls which costs kinetic energy, of which there is little. In the geodesic approximation [10], ϕ is confined to M_n for all time by construction, and the time evolution of the field consists of adiabatic progression through a succession of static solutions. The number of degrees of freedom is reduced from infinity (the dimension of Q_n) to the finite dimension of M_n. Since M_n is an equipotential surface, the trajectory is determined solely by the metric on M_n, induced by that on Q_n, the kinetic energy functional. The evolution equations are simply the geodesic equations with respect to this induced metric, hence the name of the approximation. Studying low-energy field dynamics using this approach is still a nontrivial task, the main difficulty being evaluation of the induced metric on M_n. Nevertheless, the geodesic

[2]There are exceptions: see chapter 2 for example

approximation is more amenable to analytic effort than is the original field variational problem and, more importantly, leads to significant qualitative insight into soliton dynamics.

1.3 A pedagogical example: the sine-Gordon model

The summary contained in the last section is rather dense, so the reasoning in favour of the existence of topological solitons will be outlined again for each system studied as it arises. It is also useful to see how the argument develops in the context of a simple, concrete example. The standard choice for such an exercise is the sine-Gordon model, a particularly convenient choice for our purposes since it will also provide relevant background information for the discrete version of the model, studied in chapters 2 and 3. The model is rather special because although the solitons are topologically stable, the field equation is integrable (so the classification of solitons described in section 1.1 does not define an equivalence relation!). The integrability of the model has been well studied [11] but is not relevant here, so we restrict ourselves to the model's topological aspects [12].

Both physical space X and target space Φ are the real line $\mathbb{R}$, and space-time is $(1+1)$-dimensional Minkowski space with metric $(g_{\mu\nu}) = \mathrm{diag}(1,-1)$, the convention being that $x^0 = t$, $x^1 = x$. The field is denoted $\psi(x,t)$. Definition of the model is completed by the action,

$$S := \frac{1}{4}\int dt\, dx\, [\partial_\mu\psi\partial^\mu\psi - F^2(\psi)], \tag{1.2}$$

where $F(\psi) = \sin\psi$. (The model is more usually defined in terms of a rescaled field $\phi = 2\psi$, but the above form is more convenient in this context.) The field equation follows immediately, by the calculus of variations, from demanding that ψ be a local

extremal of S,

$$\partial_\mu \partial^\mu \psi + F F' = 0, \tag{1.3}$$

where $F' = dF/d\psi$. Note that the action is manifestly invariant under rigid space-time translations since these may be absorbed into a redefinition of the dummy integration variables t and x. This observation leads us, via Noether's theorem [13], to the conserved stress-energy tensor,

$$T^{\alpha\beta} = \frac{1}{2}\partial^\alpha \psi \partial^\beta \psi - \frac{1}{4} g^{\alpha\beta}(\partial_\mu \psi \partial^\mu \psi - F^2). \tag{1.4}$$

Conservation is easily checked:

$$\begin{aligned} \partial_\alpha T^{\alpha\beta} &= \frac{1}{2}\partial_\alpha \partial^\alpha \psi \partial^\beta \psi + \frac{1}{2}\partial^\alpha \psi \partial_\alpha \partial^\beta \psi - \frac{1}{2}\partial^\mu \psi \partial_\mu \partial^\beta \psi + \frac{1}{2} F F' \partial^\beta \psi \\ &= 0, \end{aligned} \tag{1.5}$$

using the field equation (1.3) This tensor therefore defines two conserved currents, $T^{\alpha 0}$ associated with time translation invariance, and $T^{\alpha 1}$ associated with space translation invariance. These in turn define two conserved charges, the spatial integrals of T^{00} and T^{01}, canonically identified with the field's energy and momentum respectively. The energy is, explicitly,

$$\begin{aligned} E &= \int dx\, T^{00} \\ &= \frac{1}{4}\int dx \left[\left(\frac{\partial \psi}{\partial t}\right)^2 + \left(\frac{\partial \psi}{\partial x}\right)^2 + F^2 \right]. \end{aligned} \tag{1.6}$$

From this we extract the potential and kinetic energies,

$$\begin{aligned} V &= \frac{1}{4}\int dx \left[\left(\frac{\partial \psi}{\partial x}\right)^2 + F^2 \right] \\ T &= \frac{1}{4}\int dx \left(\frac{\partial \psi}{\partial t}\right)^2. \end{aligned} \tag{1.7}$$

The potential energy density is

$$\mathcal{V} = \frac{1}{4}\left[\left(\frac{\partial\psi}{\partial x}\right)^2 + F^2\right]. \tag{1.8}$$

Note that the action may be written as the time integral of $T - V$, motivating the definition of Lagrangian $L := T - V$ in analogy with Lagrangian mechanics. The spatial integrals may then be thought of as sums over a continuous internal index $x \in X$, rather like sums over particle degrees of freedom in ordinary mechanics. This is one way of arriving at the picture of field dynamics as the motion of a notional particle in an infinite dimensional space.

If a configuration $\psi(x)$ is to have finite energy, $\mathcal{V}$ must vanish as $x \to \pm\infty$. Thus ψ must tend to a constant at each end of space ($\partial\psi/\partial x \to 0$) and this constant must be a root of $F = \sin\psi$, a vacuum of the theory ($\psi = m\pi$, $m \in \mathbb{Z}$). It is not necessary for ψ to tend to the *same* constant at both ends however. So, finite V implies

$$\psi \to \begin{cases} n_+\pi & x \to \infty \\ n_-\pi & x \to -\infty \end{cases} \tag{1.9}$$

both n_+ and n_- being integers. The action S is manifestly invariant under the transformation $\psi(x,t) \mapsto \psi(x,t) - n_-\pi$, so we may assume without loss of generality that $n_- = 0$. Dropping the $+$ subscript, we define n to be the topological charge of the configuration, a constant which cannot change under any continuous deformation of ψ preserving finiteness of V, that is, remaining on configuration space Q. Thus Q is partitioned into disjoint sectors Q_n. The $n = 1$ sector is called the kink sector, topological solitons in one dimension conventionally being called kinks. Looked at from this viewpoint the partition of Q is due to the nontriviality of $\pi_{d-1}(\Phi_0)$, since $\Phi_0 = \mathbb{Z} \subset \mathbb{R} = \Phi$, and $d = 1$ (that is $\pi_{d-1}(\Phi_0) = \pi_0(\mathbb{Z}) = \mathbb{Z}$), so the kink is classified as a monopole. This is a somewhat degenerate case, because ∂X and Φ_0 are discrete. In particular, n cannot be identified with the degree of the map $\psi_\infty : \partial X \to \Phi_0$ since

this degree is undefined.

Alternatively, given the symmetry of S just noted (under $\psi \mapsto \psi + m\pi$, $m \in \mathbb{Z}$) we could perfectly well regard ψ as an angular variable, identifying ψ values differing by $m\pi$, and think of target space as S^1, the circle, rather than $\mathbb{R}$. In this language, all the vacua $\psi = m\pi$ are identified as one single vacuum on S^1, $\psi = 0$. The finite energy condition then requires that ψ tends to the same point in Φ on all ∂X, that is, both ends of space. So finite energy configurations are effectively maps from the one point compactification $\mathbb{R} \cup \{\infty\} \cong S^1$ to $\Phi = S^1$, and such maps fall into disjoint homotopy classes because $\pi_1(S^1) = \mathbb{Z}$, each class being labelled by its topological charge n. We arrive at the same conclusion, that Q is partitioned into disjoint sectors Q_n, but from this viewpoint the partition is due to the nontriviality of $\pi_d(\Phi)$, and the kink is classified as a texture. The winding number n *is* the degree of the map $\psi : X \cup \{\infty\} \to \Phi$, so this line of argument is less degenerate and rather more natural than the previous one. It was as the lowest dimensional analogue possible of the Skyrmion (a texture) that Skyrme studied the sine-Gordon kink.

Having established that $\pi_0(Q) = \mathbb{Z}$, the next task is to solve the field equation for the static kink solution. Substituting $\dot{\psi} = 0$ into (1.3) the sine-Gordon equation reduces to a second order nonlinear ordinary differential equation for $\psi(x)$,

$$-\frac{d^2\psi}{dx^2} + FF' = 0. \tag{1.10}$$

This equation is quite straightforward to solve with kink boundary conditions. However, the problem of finding the static soliton can be simplified still further by means of an ingenious argument of a kind which will repeatedly occur in subsequent chapters, called the Bogomol'nyi argument [14]. The analogous argument can always be made for kink bearing (one dimensional solitonic) models, but is a feature of only a few special higher-dimensional models. Recall that the static field variational problem is equivalent to extremizing $V[\psi]$, and that stable static solutions are minimals of V. For any static field $\psi(x)$ satisfyng the kink boundary conditions $\psi(-\infty) = 0$,

$\psi(\infty) = \pi$, we start from the trivial inequality

$$\begin{aligned} 0 &\leq \frac{1}{4}\int dx \left(\frac{d\psi}{dx} - \sin\psi\right)^2 \qquad (1.11) \\ &= V[\psi] - \frac{1}{2}\int dx \frac{d\psi}{dx}\sin\psi \\ &= V[\psi] + \frac{1}{2}\left[\cos\psi(x)\right]_{-\infty}^{\infty} \\ &= V[\psi] - 1 \\ \Rightarrow V[\psi] &\geq 1, \qquad (1.12) \end{aligned}$$

that is, the functional $V[\psi]$ is bounded below by 1 on Q_1. Moreover, in order for ψ to saturate this bound the integrand of (1.11) must vanish for all x,

$$V[\psi] = 1 \Leftrightarrow \frac{d\psi}{dx} = \sin\psi. \qquad (1.13)$$

Equation (1.13) is called the Bogomol'nyi equation. Note that it is a *first* order nonlinear ordinary differential equation. Its general solution is

$$\psi = 2\tan^{-1} e^{x-b}, \qquad (1.14)$$

where b is an arbitrary real constant of integration, which does indeed satisfy kink boundary conditions, the only other solution being $\psi = 0$, the vacuum. We could have arrived at (1.13) by multiplying the static field equation (1.10) by $d\psi/dx$ and integrating, but the procedure above gives us, without further calculation, the static kink energy and the information that (1.14) is a *global minimal* of V on Q_1 and thus is stable. One can repeat the Bogomol'nyi argument with

$$\left(\frac{d\psi}{dx} + \sin\psi\right)^2 \qquad (1.15)$$

as the integrand in (1.11) and antikink boundary conditions on ψ ($\psi(-\infty) = 0$,

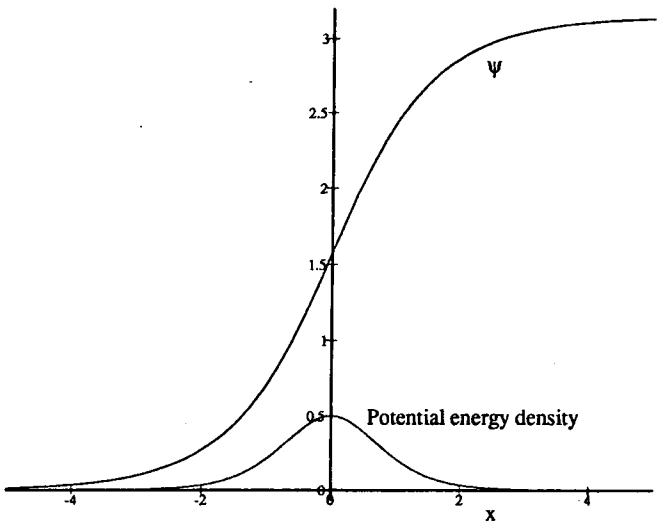

Figure 1.1: The static kink solution $\psi(x)$ and its potential energy density $\mathcal{V}(x)$.

$\psi(\infty) = -\pi$) to obtain the same bound $V[\psi] \geq 1$ on Q_{-1}, the antikink sector. The Bogomol'nyi equation is then

$$\frac{d\psi}{dx} = -\sin\psi, \tag{1.16}$$

and we see immediately that the antikink is the image of the kink under the parity transformation $x \mapsto -x$.

The $b = 0$ kink is plotted along with its energy density $\mathcal{V}(\psi(x))$ in figure 1.1. Changing b merely translates the kink along physical space, so it is natural to regard b, the point at which $\psi = \pi/2$, as the kink's position. Thinking of Φ as S^1, a general winding n (assumed positive) configuration must pass through $\psi = \pi/2$ at least n times, but may pass through more times if it winds back on itself. Let it pass through $\pi/2$ n_k times with positive $d\psi/dx$ and $n_{\bar{k}}$ times with negative $d\psi/dx$, so that $n = n_k - n_{\bar{k}}$. The configuration is naturally interpreted as containing n_k kinks and $n_{\bar{k}}$ antikinks located at the points where $\psi(x) = \pi/2$. Given the form of $\mathcal{V}$, such points will coincide with lumps of potential energy. If n is negative, the interpretation

is still valid, but antikinks outnumber the kinks. The existence of the zero mode[3] b is inevitable because of the model's translation invariance: the translation orbit of any configuration must be an equipotential curve in Q. In the case of the static kink, the Bogomol'nyi argument tells us that this equipotential curve is the absolute minimum of V on Q_1, the one dimensional valley floor of the potential relief, twisting through Q_1. This is the unit charge moduli space M_1. A static kink set in motion by a small Galilean boost has a trajectory in Q_1 starting at some point on M_1, with a small initial velocity tangential to M_1 (where small in the second sense means small with respect to the metric on Q_1, the kinetic energy). Such a trajectory must always stay close to M_1 by energetic considerations. The geodesic approximation to kink dynamics consists of confining the field for all time to move in this single dimension in accordance with the Lagrangian inherited from the full field theory. Explicitly, let b in (1.14) be time dependent $b(t)$. Then

$$\dot{\psi} = -\dot{b}\,\mathrm{sech}(x-b), \tag{1.17}$$

and so

$$\begin{aligned} T &= \frac{\dot{b}^2}{4}\int dx\,\mathrm{sech}^2(x-b) = \frac{\dot{b}^2}{4}\int \frac{dx}{\cosh^2 x} \\ &= \frac{\dot{b}^2}{4}\left[\tanh x\right]_{-\infty}^{\infty} = \frac{1}{2}\dot{b}^2, \end{aligned} \tag{1.18}$$

while $V \equiv 1$ since the Bogomol'nyi bound remains saturated. Note that T is the kinetic energy of a point particle moving geodesically on $M_1 \cong \mathbb{R}$ with a flat metric. Of course, since M_1 is one dimensional, it is inevitably intrinsically flat, but the key point is that the induced metric is flat when expressed in terms of b, the kink position. We are interested in the geometry of M_1 as a surface embedded in Q_1 viewed from

[3]Strictly speaking it is $\partial\psi/\partial b$, as a function of x calculated from the static kink solution, which is the zero mode rather than b itself, but little confusion can arise from treating the two interchangeably.

"outside" M_1. So, the inherited Lagrangian is

$$L = \frac{1}{2}\dot{b}^2 - 1, \tag{1.19}$$

that of a nonrelativistic free particle of unit mass (the static kink energy). The solution of the equation of motion generated by (1.19) is $b(t) = vt + b_0$, where $b(0) = b_0$, and $\dot{b}(0) = v$. The kink travels along at constant speed, and the approximate field solution is

$$\psi(x,t) = 2\tan^{-1} e^{x-vt-b_0}. \tag{1.20}$$

In this case, the geodesic approximation is a complicated way of deriving something simple. Note that the action is Lorentz invariant if ψ transforms as a Lorentz scalar. Hence from the static solution (1.14) we can generate moving kink solutions with arbitrary speed simply by performing a Lorentz boost:

$$\begin{aligned} t \mapsto \tilde{t} &= \frac{1}{\sqrt{1-v^2}}(t - vx) \\ x \mapsto \tilde{x} &= \frac{1}{\sqrt{1-v^2}}(x - vt) \\ \psi(x,t) \mapsto \tilde{\psi}(x,t) &= \psi(\tilde{x},\tilde{t}). \end{aligned} \tag{1.21}$$

This yields the exact solution

$$\tilde{\psi}(x,t) = 2\tan^{-1}\exp\left(\frac{x - vt - b_0}{\sqrt{1-v^2}}\right) \tag{1.22}$$

which agrees with (1.20) for small v. The trajectory corresponding to (1.22) is again a translation orbit in Q_1, travelled along at constant speed, and hence an equipotential curve, but with $V > 1$ because the relativistic contraction of the kink profile by $(1-v^2)^{-\frac{1}{2}}$ pushes the whole trajectory up the valley walls, away from M_1. There is a force pulling the notional particle towards the valley floor, but this is balanced by the inertial force due to the valley's twisting through Q_1. The geodesic approximation

is more useful when one is studying dynamics in higher moduli spaces, or when the original model is not Lorentz (or Galilei) invariant. In the simple case of the sine-Gordon model, however, all the higher moduli spaces are empty, as we shall shortly see, so there is no nontrivial demonstration of the approximation available in this system.

The model has some ungeneric properties, which we should point out. First, applying winding number n boundary conditions to the Bogomol'nyi argument (1.11) one obtains

$$V \geq \frac{1}{2}|\cos\psi(\infty) - \cos\psi(-\infty)| = |n| \bmod 2, \tag{1.23}$$

while the lower bound for V on Q_n so obtained is generically proportional to $|n|$ (the constant of proportionality being the single soliton energy). One can use the $n = 1$ Bogomol'nyi bound to find an optimal lower bound for V on Q_n as follows: if ψ has winding n then, by continuity, it must consist of a sequence of n unit winding sections (assuming $n > 1$); apply the $n = 1$ Bogomol'nyi argument piecewise to each of these sections in turn, yielding $V > n$. The inequality must be strict because if ψ were to saturate it, it would satisfy the Bogomol'nyi equation independently on each finite (or, at most, only semi-infinite) section, which is impossible given the general solution (1.14). The bound *is* optimal however since a winding n configuration can get arbitrarily close to it by spreading its n subkinks arbitrarily far apart. This suggests that kinks tend to repel one another, so one would suspect that there are no static n-kink solutions. This turns out to be true because, not only is a solution of the Bogomol'nyi equation a solution of the static field equation, the reverse is also true: all static solutions must satisfy either (1.13) or (1.16), so the only static solutions are kinks, antikinks and the vacuum. This statement can be seen to hold quite generally for Poincaré invariant sytems in one dimension. Consider the conservation law $\partial_\alpha T^{\alpha 1}$, arising from symmetry under spatial translations, in the case of a static field $\psi(x)$,

$$\partial_\alpha T^{\alpha 1} = 0 \Rightarrow \frac{dT^{11}}{dx} = 0$$

$$\Rightarrow \frac{d}{dx}\left[\left(\frac{d\psi}{dx}\right)^2 - F^2\right] = 0$$

$$\Rightarrow \left(\frac{d\psi}{dx}\right)^2 - F^2 = C, \tag{1.24}$$

where C is some constant, which, given finite energy boundary conditions, must be zero. Hence all static finite energy solutions must satisfy

$$\frac{d\psi}{dx} = \pm F, \tag{1.25}$$

the Bogomol'nyi equation. This holds for all $F(\psi)$. Now, the Bogomol'nyi equation has no solutions with winding $|n| > 1$, so while $M_{\pm 1} \cong \mathbb{R}$, all the higher moduli spaces are empty, $M_n = \emptyset$, $|n| > 1$. This does not generalize to higher dimensions. A higher dimensional system may or may not have a Bogomol'nyi bound, and a first order Bogomol'nyi equation for configurations saturating that bound. If it does, the bound may or may not be attainable, that is, the Bogomol'nyi equation may or may not have solutions with the correct boundary behaviour. However, if the bound is attainable in one topological sector then it is generically attainable in all sectors. So if a system has static single-soliton solutions saturating a Bogomol'nyi bound, it generically has static n-soliton solutions also, and moreover, the constituent solitons of such solutions can generically occupy any and all positions in X. One interprets this as meaning that such solitons do not exert any forces on one another when at rest. This contrasts with the kinks described above which always repel one another. Another aspect which does not generalize to systems in higher dimensions is the two-way equivalence of the Bogomol'nyi and static field equations. That is, there exist models with finite energy static solutions which do not satisfy the Bogomol'nyi equation, in the same topological sector as those that do. Such solutions cannot globally minimize the potential energy functional within their sector, and are generically saddle points rather than local minimals of V.

This concludes the introductory material, and the thesis proper begins in chapter

2. A spatially discrete system is presented which has topological properties closely related to those of its continuum limit, the sine-Gordon model, discussed above. We study the dynamics of a single kink propagating through space at both low and high speeds, using the geodesic approximation and numerical simulations. In chapter 3 we go on to quantize the kink sector of the discrete sine-Gordon system in the weak coupling approximation, and compute numerically a quantum induced potential called the kink Casimir energy. We return to the classical dynamics of a single soliton in chapter 4. This concerns the low-energy dynamics of a $\mathbb{C}P^1$ lump on the two-sphere. The induced metric on the unit charge moduli space (a six dimensional manifold, $(S^3 \times \mathbb{R}^3)/\mathbb{Z}_2$) is found explicitly, and the qualitative features of motion on totally geodesic submanifolds are described. Finally, chapter 5 presents an analysis of long range intervortex forces in the abelian Higgs model. A point source approximation is used to rederive the static intervortex potential from a new viewpoint, then velocity dependent forces are found by extending the method of linear retarded potentials. The results are used to study type II and critical vortex scattering, and to conjecture an analytic asymptotic form for the metric on the two-vortex moduli space. All of these chapters are self contained, so concluding remarks are presented at the end of each.

Chapter 2

A topological discrete sine-Gordon system

2.1 Introduction

There are two main reasons why solitons in discrete spaces (lattices) are of interest. First, many of the applications of topological solitons are in condensed matter and biophysics where the systems under study (a macroscopic sample of some solid material or a molecule of DNA for example) really *do* have a lattice structure. Much theoretical analysis of such systems has been carried out using continuum solitons, ignoring the discreteness of space, mainly as a matter of calculational convenience. However, it has become increasingly recognized that discreteness does introduce significant new effects into soliton dynamics, and that these have important phenomenological implications. Second, although solitons relevant to particle theory usually occur in models defined on a continuum, one must often resort to numerical simulations to study their dynamics. This inevitably leads to a fictitious discretization of space, because all mathematics on digital computers is inherently discrete. If one is to attempt to deduce facts about the behaviour of real continuum solitons from numerical simulations, one must be able to distinguish which of the effects seen are truly physical and

which are introduced by the fictitious discretization. By judicious choice of the means of discretization (inevitably rather arbitrary), one might then hope to partially eliminate the latter. In either case, one is motivated to try to understand the dynamics of solitons on lattices, and, as a first case, the dynamics of kinks on one dimensional lattices in particular.

This chapter deals with a discrete version of the sine-Gordon model, reviewed in section 1.3. We use the same notation (ψ for the field, T and V for kinetic and potential energies, L for Lagrangian etc.) but from now on x becomes a discrete variable taking values on a regular lattice of spacing h including the point $x = 0$, for brevity, $x \in h\mathbb{Z}$. A subscript + on a function $f(x)$ denotes forward shift on the lattice, $f_+(x) := f(x + h)$, and similarly $f_-(x) := f(x - h)$. The forward difference operator is denoted Δ and defined such that $\Delta f = (f_+ - f)/h$. Spatial integrals are replaced by sums over x. By a discrete sine-Gordon (DSG) system, we mean a Lagrangian system defined on this lattice whose continuum limit is the usual sine-Gordon model, already described. Note that the discretization process is highly non-unique: there are infinitely many inequivalent lattice systems all with the same continuum limit.

The conventional choice for discretizing the sine-Gordon model is simply to replace the partial derivative $\partial\psi/\partial x$ by the forward difference $\Delta\psi$, yielding the Frenkel-Kontorova model,

$$L_{FK} = \frac{h}{4} \sum_{x \in h\mathbb{Z}} [\dot{\psi}^2 - (\Delta\psi)^2 - \sin^2 \psi]. \tag{2.1}$$

The equation of motion generated from this, by extremizing the action $S_{FK} = \int dt\, L_{FK}$, is a second order differential-difference equation, of which no nontrivial solutions are known (the vacua $\psi = n\pi$ persist of course). The system has been extensively studied [15, 16, 17, 18, 19] using both numerical and perturbative techniques, and the qualitative features of kink dynamics are well known. There are two static kink solutions, modulo discrete translations ($x \mapsto x + nh$, $n \in \mathbb{Z}$, manifestly a symmetry of L_{FK}), one with the kink located directly on a lattice site and one with a kink located halfway between two lattice sites. As before, by kink location we mean

the point at which $\psi = \pi/2$, defined by linear interpolation between lattice sites if necessary. So the kink position takes values in $\mathbb{R}$, the real line containing the lattice, for a general configuration, but is restricted to the discrete subset $x \in (h/2)\mathbb{Z}$ for static kink solutions. Of these two static kinks, the one located at $x = nh$ has higher potential energy than the one located at $x = (n+\frac{1}{2})h$. The former is a saddle point of the potential and is an unstable solution, while the latter is a minimum and hence stable. The difference between the two static kink energies is called the Peierls-Nabarro barrier and has a crucial effect on kink dynamics. If a moving kink has insufficient kinetic energy, it cannot surmount the barrier and gets trapped between two lattice sites, a process called "pinning." It oscillates and emits radiation, settling down into the potential well. The motion of a soliton on the lattice is dissipative, there being no continuous translation symmetry to force momentum conservation. Thus a kink which starts off with enough kinetic energy to surmount the Peierls-Nabarro barrier and travel along the lattice dissipates its energy as radiation, slows down and eventually becomes trapped.

Such dynamics may make the model interesting from a physical viewpoint if analogous effects are seen in real condensed matter or biophysical systems, but is obviously radically different from the continuum dynamics, so the Frenkel-Kontorova model represents a bad choice for fictitious discretization. In the next section we will present an alternative discretization, specifically designed to preserve the topological (more precisely, the Bogomol'nyi) properties of the continuum model. There is a topological lower bound on the kink energy and a first order difference equation for kinks saturating this bound. These static solutions may be found explicitly. There is no Peierls-Nabarro barrier and consequently kink dynamics in this topological discrete sine-Gordon (TDSG) system is quite different from that in the Frenkel-Kontorova model.

2.2 Topological discretization

We wish to find D and F, lattice versions of $\partial\psi/\partial x$ and $\sin\psi$ respectively (that is, $D \to \partial\psi/\partial x$ and $F \to \sin\psi$ in the limit $h \to 0$) such that their product is

$$DF = -\Delta \cos\psi. \tag{2.2}$$

For then, defining the potential energy to be

$$V = \frac{h}{4} \sum_{x \in h\mathbb{Z}} [D^2 + F^2], \tag{2.3}$$

which has the correct continuum limit, we have a lattice version of the Bogomol'nyi argument. Let $\psi(x)$ have kink boundary conditions, $\psi(-\infty) = 0$, $\psi(\infty) = \pi$. Then,

$$\begin{aligned} 0 &\leq \frac{h}{4}\sum_x (D-F)^2 \qquad (2.4)\\ &= V - \frac{h}{2}\sum_x DF \\ &= V + \frac{1}{2}\sum_x (\cos\psi_+ - \cos\psi) \\ &= V - 1 \\ \Rightarrow V &\geq 1. \qquad (2.5) \end{aligned}$$

Once again, the energy of a kink configuration is bounded below by 1, and can attain this bound if and only if $D = F \ \forall x \in h\mathbb{Z}$, the lattice Bogomol'nyi equation.

Such a factorization of $\Delta \cos\psi$ is possible. The most natural choice is

$$\begin{aligned} D &= \frac{2}{h}\sin\frac{1}{2}(\psi_+ - \psi) \\ F &= \sin\frac{1}{2}(\psi_+ + \psi), \end{aligned} \tag{2.6}$$

although $\widetilde{D} := D/\alpha(h)$ and $\tilde{F} := \alpha(h)F$ would do just as well provided α is a strictly

positive, smooth function with $\alpha(0) = 1$. Rearranging the Bogomol'nyi equation,

$$\tan\frac{\psi_+}{2} = \frac{2+h}{2-h}\tan\frac{\psi}{2}, \tag{2.7}$$

we see that if $h > 2$ any solutions are inevitably oscillatory, which is incompatible with with kink boundary conditions. We therefore restrict h to the range $(0, 2)$. This upper bound on h could be eliminated by using the modified quantities $\widetilde{D}$ and $\widetilde{F}$ choosing α such that $h\alpha^2(h) < 2\ \forall h$ (for example, $\alpha = e^{-h}$), because the Bogomol'nyi equation $\widetilde{D} = \widetilde{F}$ then becomes

$$\tan\frac{\psi_+}{2} = \frac{2+\alpha^2 h}{2-\alpha^2 h}\tan\frac{\psi}{2}, \tag{2.8}$$

so no oscillatory behaviour can occur. There is no clear advantage in this, so we do not pursue it but choose the simplest version (2.6).

Given that $h \in (0, 2)$ the solution of (2.7) is

$$\psi(x) = 2\tan^{-1} e^{a(x-b)}, \tag{2.9}$$

where

$$a = \frac{1}{h}\log\left(\frac{2+h}{2-h}\right), \tag{2.10}$$

and b is an arbitrary real constant. This does indeed satisfy kink boundary conditions. Note that $a(h)$ is a monotonically increasing function on $(0, 2)$ and that $a \to 1$ as $h \to 0$, while a is unbounded as $h \to 2$. The static kink is therefore a sharpened version of the continuum kink (1.14) "sampled" on the lattice $x \in h\mathbb{Z}$. The sharpening factor disappears in the continuum limit and becomes infinite in the anticontinuum limit $h \to 2$. The potential energy of a kink is correspondingly highly localized. If $b = 0$ then the proportion of the energy (normalized to unity with our conventions) contributed by the two central links $\langle -h, 0\rangle$ and $\langle 0, h\rangle$ is $4h/(4 + h^2)$. If $h = 1$, this is 80% while in the limit $h \to 2$ it tends to 100%. The $h = 1$ static kink is plotted in figure 2.1. Perhaps the most surprising feature of (2.9) is the existence of

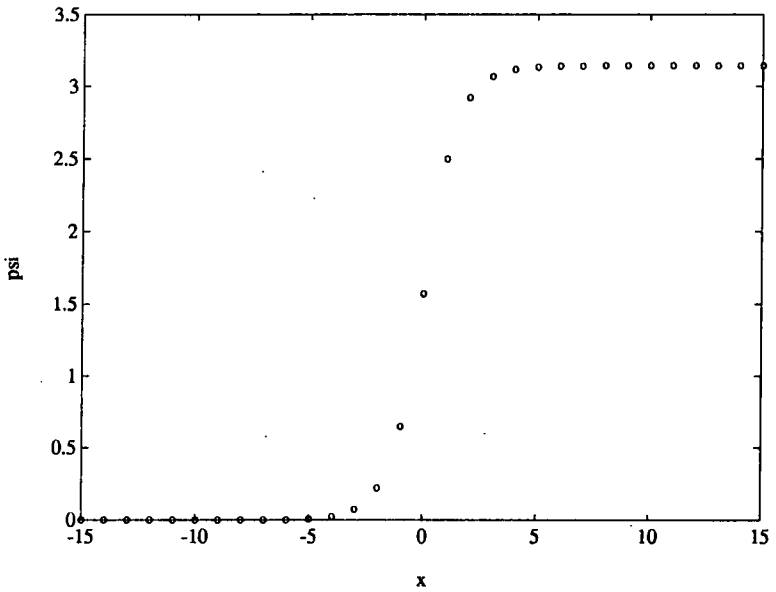

Figure 2.1: The static kink solution for $h = 1$ with $b = 0$.

the *continuous* zero mode b, which is not due to any symmetry of the model. Given that the model has only discrete translation symmetry, one would naively expect the static kink to be located either on the lattice $h\mathbb{Z}$ or $h(\mathbb{Z}+\frac{1}{2})$, but it turns out that the kink can take any position relative to the lattice. It follows that this system has no Peierls-Nabarro potential, and one would expect kink dynamics to be consequently much simpler than in the Frenkel-Kontorova model.

Since the configuration (2.9) satisfies $D = F$, it globally minimizes V within the class of configurations with kink boundary conditions (the kink sector) and hence is a stable static solution. Applying antikink boundary conditions and starting with summand $(D + F)^2$ in (2.4) we obtain the antikink Bogomol'nyi bound, $V \geq 1$, $V = 1 \Leftrightarrow D = -F$, so as one would expect, there are stable static antikinks, obtained from the kinks by mapping $x \mapsto -x$. So if ψ is a solution of $D = \pm F$, it must be a solution of the static field equation. It is interesting to consider the possibility of other nontrivial static solutions such as double kinks and kink-antikink pairs. Both exist in the Frenkel-Kontorova model because of the Peierls-Nabarro barrier: at long range kink-kink repulsion and kink-antikink attraction are not strong enough to overcome

the pinning force which traps the constituent kinks and antikinks of a configuration between lattice sites. It follows that there are static solutions consisting of well separated kinks and/or antikinks in stable equilibrium. Given the absence of a Peierls-Nabarro barrier in the TDSG system, one would expect no such solutions to exist. Recall that in the continuum model, nonexistence followed from Noether's theorem and translation invariance (equations (1.24) and (1.25)), by means of which we proved that all static solutions must satisfy the Bogomol'nyi equation. Remarkably, this result almost survives on the lattice, despite the loss of translation symmetry. From the explicit expression for the potential,

$$V = \frac{h}{4}\sum_x \left[\frac{4}{h^2}\sin^2\frac{1}{2}(\psi_+ - \psi) + \sin^2\frac{1}{2}(\psi_+ + \psi)\right], \tag{2.11}$$

we derive the static field equation, $\partial V/\partial\psi = 0$, where

$$\begin{aligned}
\frac{4}{h}\frac{\partial V}{\partial \psi} &= \frac{4}{h^2}\left[\sin\frac{1}{2}(\psi-\psi_-)\cos\frac{1}{2}(\psi-\psi_-) - \sin\frac{1}{2}(\psi_+-\psi)\cos\frac{1}{2}(\psi_+-\psi)\right] \\
&\quad + \sin\frac{1}{2}(\psi+\psi_-)\cos\frac{1}{2}(\psi+\psi_-) + \sin\frac{1}{2}(\psi_++\psi)\cos\frac{1}{2}(\psi_++\psi) \\
&= -\frac{2}{h^2}\left[\sin(\psi_+-\psi) + \sin(\psi_--\psi)\right] + \frac{1}{2}\left[\sin(\psi_++\psi) + \sin(\psi_-+\psi)\right] \\
&= -\frac{4}{h^2}\sin\frac{1}{2}(\psi_+ - 2\psi + \psi_-)\cos\frac{1}{2}(\psi_+-\psi_-) \\
&\qquad\qquad + \sin\frac{1}{2}(\psi_+ + 2\psi + \psi_-)\cos\frac{1}{2}(\psi_+-\psi_-) \\
&= -\cos\frac{1}{2}(\psi_+-\psi_-)\left[\frac{4}{h^2}\sin\frac{1}{2}(\psi_+ - 2\psi + \psi_-)\right. \\
&\qquad\qquad \left. - \sin\frac{1}{2}(\psi_+ + 2\psi + \psi_-)\right].
\end{aligned} \tag{2.12}$$

So each triple of lattice sites $\{\psi_-, \psi, \psi_+\}$ must satisfy either

$$\cos\frac{1}{2}(\psi_+ - \psi_-) = 0 \tag{2.13}$$

or

$$\frac{4}{h^2}\sin\frac{1}{2}(\psi_+ - 2\psi + \psi_-) - \sin\frac{1}{2}(\psi_+ + 2\psi + \psi_-) = 0. \tag{2.14}$$

Note that if $\{\psi_-, \psi, \psi_+\}$ satisfies (2.13) then $\psi_+ - \psi_- = \pi \bmod 2\pi$ but ψ is unconstrained, and that if the solution is to satisfy regular finite-energy boundary conditions, then (2.13) cannot hold for all x, but, at most, at a finite number of isolated sites. Along the rest of the lattice (2.14) must hold. Now, motivated by the continuum model, we consider the quantity $\Delta(D^2 - F^2)$:

$$\begin{aligned}
& [\Delta(D^2 - F^2)]_- \\
= \; & \frac{4}{h^2}\left[\sin\frac{1}{2}(\psi_+ - \psi) - \sin\frac{1}{2}(\psi_- - \psi)\right]\left[\sin\frac{1}{2}(\psi_+ - \psi) + \sin\frac{1}{2}(\psi_- - \psi)\right] \\
& - \left[\sin\frac{1}{2}(\psi_+ + \psi) - \sin\frac{1}{2}(\psi_- + \psi)\right]\left[\sin\frac{1}{2}(\psi_+ + \psi) + \sin\frac{1}{2}(\psi_- + \psi)\right] \\
= \; & \frac{16}{h^2}\cos\frac{1}{4}(\psi_+ - 2\psi + \psi_-)\sin\frac{1}{4}(\psi_+ - \psi_-)\sin\frac{1}{4}(\psi_+ - 2\psi + \psi_-)\cos\frac{1}{4}(\psi_+ - \psi_-) \\
& - 4\cos\frac{1}{4}(\psi_+ + 2\psi + \psi_-)\sin\frac{1}{4}(\psi_+ - \psi_-)\sin\frac{1}{4}(\psi_+ + 2\psi + \psi_-)\cos\frac{1}{4}(\psi_+ - \psi_-) \\
= \; & \sin\frac{1}{2}(\psi_+ - \psi_-)\left[\frac{4}{h^2}\sin\frac{1}{2}(\psi_+ - 2\psi + \psi_-) - \sin\frac{1}{2}(\psi_+ + 2\psi + \psi_-)\right].
\end{aligned} \tag{2.15}$$

Comparing (2.15) and (2.14) we see that if (2.14) holds everywhere, then $\Delta(D^2 - F^2) = 0$ and, applying finite energy boundary conditions, $D = \pm F$. We cannot, however, claim that there are no non-Bogomolnyi static solutions, because there may be isolated points where $\psi(x)$ satisfies (2.13) rather than (2.14). It is quite difficult to find such solutions without spoiling the finite-energy boundary behaviour. Nevertheless, they *do* exist, as may be illustrated by example.

Consider the configuration

$$\psi(x) = \begin{cases} \psi_K(x - b) & x \le 0 \\ \psi_K(x - b - 2h) + \pi & x \ge h, \end{cases} \tag{2.16}$$

where $\psi_K(x)$ is the kink located at zero, $\psi_K(x) = 2\tan^{-1}\exp(ax)$, and $b \in \mathbb{R}$. This is a static solution. Each of the triples $\{\psi_-(x), \psi(x), \psi_+(x)\}$ for $x \le -h$ satisfies (2.14)

because both pairs $\{\psi_-, \psi\}$ and $\{\psi, \psi_+\}$ satisfy the Bogomol'nyi equation. The same holds for the triples centred on $x \geq 2h$. So this configuration satisfies (2.14) for all x except $x = 0$ and $x = h$. However, at both these points (2.13) holds, by construction. Hence $\partial V/\partial\psi = 0 \quad \forall x$, and (2.16) is a static solution. This is a double kink, but the same method could be used to put more dislocations[1] in the lattice and produce a n-kink static solution. Note that the construction works for any value of b, and with the dislocation between any pair of neighbouring lattice sites – $x = 0$ and $x = h$ were chosen in this case. The static double kink and its energy distribution are plotted for four different values of b in figure 2.2. In each case there is a lump of energy spread over approximately four lattice sites caused by the kink, and a one-site spike caused by the dislocation, though when $b = -0.5$ they coincide. Since b can take any value there is a one-parameter curve of these solutions in configuration space Q_2. This should be an equipotential curve, since any variation of V along it would be in contradiction with the assertion that (2.16) is a static solution for all b. It is not hard to calculate V because the Bogomol'nyi argument gives the energies of the left and right hand tails ($x \leq 0$ and $x \geq h$), so one merely has to add the energy of the link $\langle 0, h\rangle$, the dislocation spike. The result is,

$$V = 1 - \frac{1}{2}[\cos\psi(0) + \cos\psi(h)] + \frac{h}{4}(D^2 + F^2)\Big|_{x=0} . \tag{2.17}$$

After much algebra, this reduces to

$$V = \frac{(2+h)^2}{4h}, \tag{2.18}$$

which is indeed independent of b. Note that $V > 2 \quad \forall h \in (0,2)$ leading us to conjecture that the solutions are saddle points of the potential: it is energetically favourable for the configuration to evolve into a conventional kink-kink pair arbitrarily

[1]This terminology is perhaps a little unfortunate since discrete kinks themselves are often used to model crystal dislocations. However, no apt alternative suggests itself.

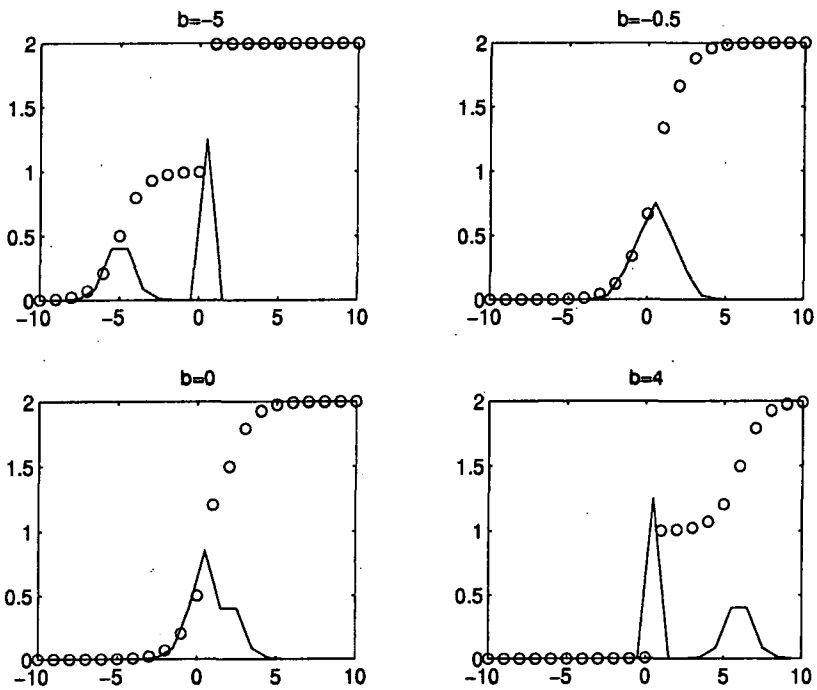

Figure 2.2: The static two-kink solution on the $h = 1$ lattice. The circles represent $\psi(x)/\pi$ while the solid line represents the energy distribution. The energy should be thought of as located in the links between lattice sites.

far apart so as to approach $V = 2$. One would suspect that any static n-kink solution constructed similarly would also be a saddle point of V on Q_n. Note also that $V \to \infty$ as $h \to 0$ so that the solution is lost in the continuum limit. In summary then, the static field equation implies the Bogomol'nyi equation, except in some rather pathological cases where $\psi(x)$ jumps by π mod 2π in a short interval. By exploiting this possibility one can construct *unstable* static n-kink solutions. It is unlikely that any stable non-Bogomol'nyi solutions exist.

So far we have only considered the static model. The dynamics is completed by defining the kinetic energy T. Again there is much freedom in this, the only constraint being that the continuum limit of T must be $\frac{1}{4}\int dx\,(\partial\psi/\partial t)^2$. The simplest and most

physical choice is

$$T = \frac{h}{4} \sum_{x \in h\mathbb{Z}} \dot{\psi}^2, \tag{2.19}$$

where $\dot{\psi} = d\psi/dt$. This is the kinetic energy of infinitely many identical point particles moving in one dimension (the lattice picture), or equally, of a single particle moving in an infinite dimensional Euclidean space (the configuration space picture). The equations of motion are obtained from the Lagrangian $L = T - V$ in the standard way. After rearranging, one obtains

$$\ddot{\psi} = \frac{4 - h^2}{4h^2} \cos\psi(\sin\psi_+ + \sin\psi_-) - \frac{4 + h^2}{4h^2} \sin\psi(\cos\psi_+ + \cos\psi_-), \tag{2.20}$$

a second order differential-difference equation. No exact time-dependent solutions of (2.20) are known. In particular, given the static kink solution, there is no Lorentz or Galilei symmetry by means of which one can generate kink solutions in uniform motion. (It is interesting to note, however, that with a different choice of T, one can generate equations of motion which *do* have an exact travelling kink solution, at least for one chosen velocity [20]. This velocity becomes, along with h, a parameter of the model.) Nevertheless, we do have a continuous moduli space of static kink solutions, so we can calculate a geodesic approximation to low speed kink dynamics. At higher speeds we must perform numerical simulations.

2.3 A geodesic approximation

Just as in the continuum model, the kink moduli space in the TDSG system is the real line parametrized by kink position b, $M_1 \cong \mathbb{R}$, contrary to one's naive expectation of a lattice model, $M_1 \cong \mathbb{Z}$. The Bogomol'nyi argument shows that V attains its topological minimum value on M_1, which is thus a level valley bottom twisting through the kink sector Q_1. We obtain the geodesic approximation to low speed kink dynamics by restricting ψ to M_1 by construction, but allowing its position on M_1 to vary with

time. The number of degrees of freedom is thus drastically cut, from infinity to one. The potential is constant, so the dynamics is determined by the metric $g_h(b)$ induced by the kinetic energy (the subscript is to emphasise the metric's parametric dependence on the lattice spacing). In contrast with the continuum model, we expect the metric to depend periodically on b, but not be constant. This will introduce new features into the slow motion of kinks.

Explicitly, let

$$\psi(x,t) = 2\tan^{-1} e^{a(x-b(t))}. \tag{2.21}$$

Then $V \equiv 1\ \forall t$, and $\dot{\psi} = -a\dot{b}\,\mathrm{sech}\, a(x-b(t))$, so

$$T = \frac{1}{2} g_h(b)\dot{b}^2, \tag{2.22}$$

where

$$g_h(b) = \frac{a^2 h}{2} \sum_{x \in h\mathbb{Z}} \mathrm{sech}^2 a(x-b) \tag{2.23}$$

is interpreted as a metric on M_1. The reduced Lagrangian is $L_{geo} = \frac{1}{2} g_h(b)\dot{b}^2 - 1$, from which the derived equation of motion is simply the geodesic equation for (M_1, g_h). Note that $g_h(b)$ is periodic with period h, is even, has a local maximum at $b = 0$ and a local minimum at $b = h/2$. Graphs of $g_h(b)$ for various values of h are shown in figure 2.3. One way to solve the geodesic equation is to transform $b \mapsto B_h$ to a coordinate on M_1 with respect to which the metric is constant. In this case

$$b \mapsto B_h := \int_0^b d\alpha \sqrt{\frac{g_h(\alpha)}{g_h(0)}}, \tag{2.24}$$

so that

$$L_{geo} = \frac{1}{2} g_h(0)\dot{B}_h^2 - 1. \tag{2.25}$$

If the initial data are $b = 0$, $\dot{b} = v$ (implying $B = 0$, $\dot{B}_h = v$) the solution is $B_h = vt$.

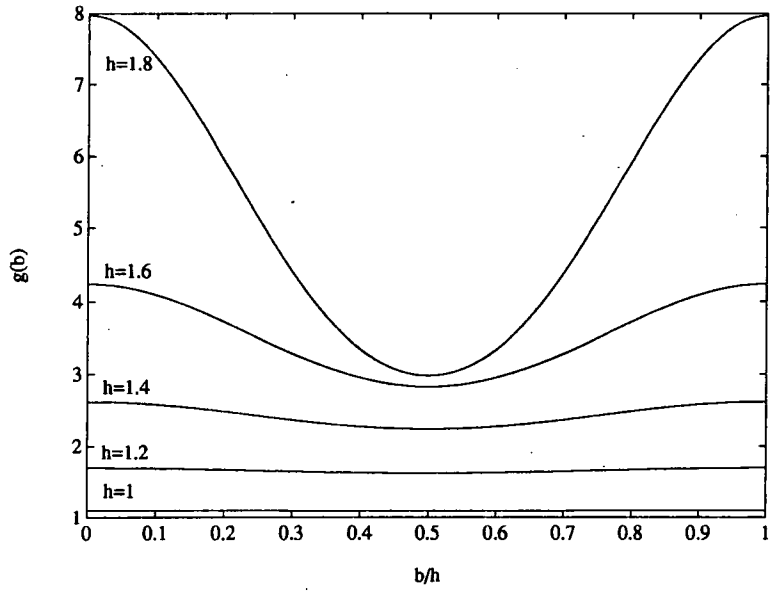

Figure 2.3: The metric $g_h(b)$.

To find the trajectory $b(t)$, one inverts the transformation (2.24),

$$b(t) = B_h^{-1}(vt). \tag{2.26}$$

So, in this approximation, finding the moving kink solution is reduced to evaluating the function $B_h(b)$ of equation (2.24). Note that this function is strictly increasing, and that

$$B_h(b + nh) = B_h(b) + nB_h(h) \tag{2.27}$$

for all integers n owing to the periodicity of $g_h(b)$, so it suffices to know B_h over a single period, $b \in [0, h]$. The integral in (2.24) is easily evaluated numerically, and one may then plot $B_h(b) = vt$ against b to obtain kink trajectories for different values of h (see figure 2.4). The main feature is that the kink wobbles as it moves through the lattice. This wobble is a dynamical effect: there is no potential (such as Peierls-Nabarro) causing it. Thinking of $g_h(b)$ as a position dependent kink mass, the kink is lightest between lattice sites ($b \in h(\mathbb{Z} + \frac{1}{2})$) and heaviest on lattice sites ($b \in h\mathbb{Z}$), and

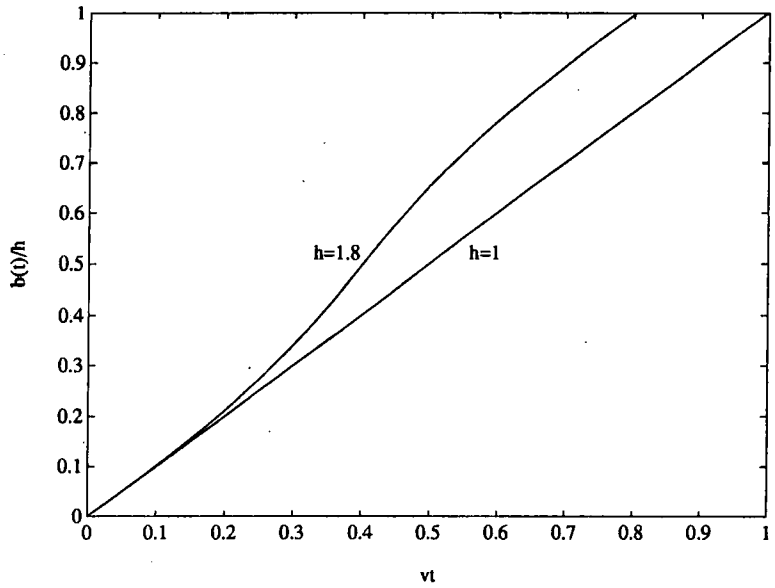

Figure 2.4: The trajectory $b(t)$ over one wobble period for $h = 1$ and $h = 1.8$

consequently, for a given conserved kinetic energy ($\frac{1}{2}g_h(0)v^2$), fastest between sites and slowest exactly on sites, when it moves with speed v. The time taken for the kink to travel from $x = 0$ to $x = h$ (the wobble period) is

$$\mathcal{T} = \frac{B_h(h)}{v}, \tag{2.28}$$

and $B_h(h)/h < 1$, approaching the upper bound in the continuum limit $h \to 0$. So the kink travels faster by a factor $h/B_h(h)$ than would be naively suggested by the initial velocity v. A graph of $B_h(h)/h$ shows that discreteness effects are small for $h \leq 1$, but grow large as h approaches 2 (see figure 2.5).

The accuracy of the approximation has been tested numerically using a fully-explicit fourth-order Runge-Kutta algorithm with fixed time-step 0.01. The initial condition was a Galilean-boosted static kink profile with initial velocity 0.01 lattice sites per unit time ($v = 0.01h$). Simulations of duration 1000 time units were performed for $h = 1, 1.2, 1.4, 1.6$ and 1.8. In every case the kink moves freely, without pinning, undergoing motion of the predicted periodicity. Furthermore, inspection of

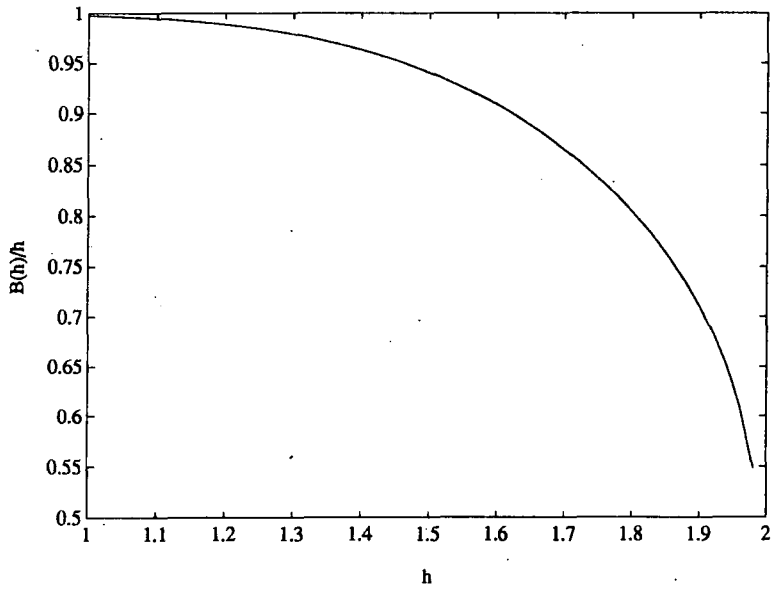

Figure 2.5: **Effect of discreteness on the wobble period, the function** $B_h(h)/h$.

the kink velocity over a single period reveals close agreement with $\dot{b}(t)$ calculated from (2.26) (see figure 2.6), although for $h = 1.2$, numerical errors rather swamp the very small theoretical wobble (note the scale on the velocity axis).

Collective-variable analyses of the Frenkel-Kontorova model (cf [19]) have introduced the kink position as an extra variable, accompanied by a constraint. But since the dynamics is much more complicated, all the degrees of freedom have to be kept in order to obtain accurate results. A truncation to one degree of freedom, as in the present case, does not work.

2.4 Fast-moving kinks

The geodesic approximation is expected to fail at high velocities. This is observed in the simulations as a gradual onset of kink deceleration as the initial velocity is increased. Failure occurs at lower velocities for coarser lattices—around $v = 0.02$ for $h = 1.8$ compared with $v = 0.15$ for $h = 1$. The kink energy is dissipated in the form of small amplitude oscillations ("radiation" or "phonons") emitted in its wake,

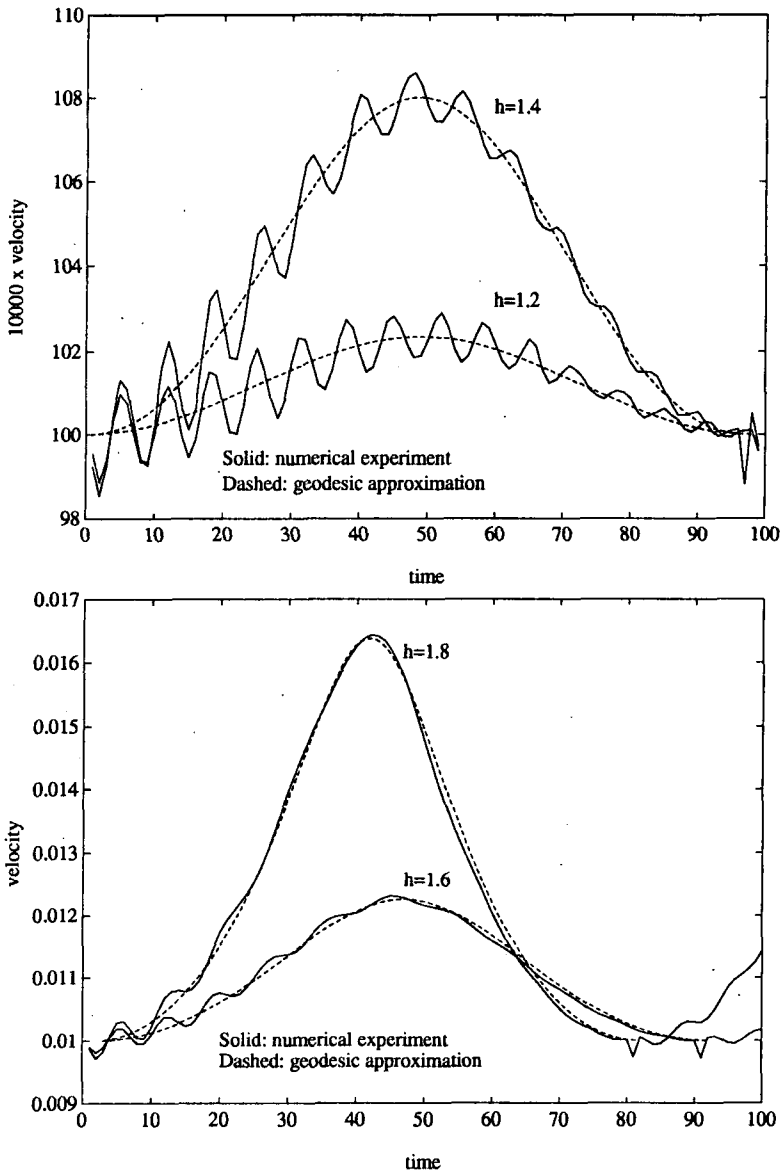

Figure 2.6: Kink velocity over a single wobble period.

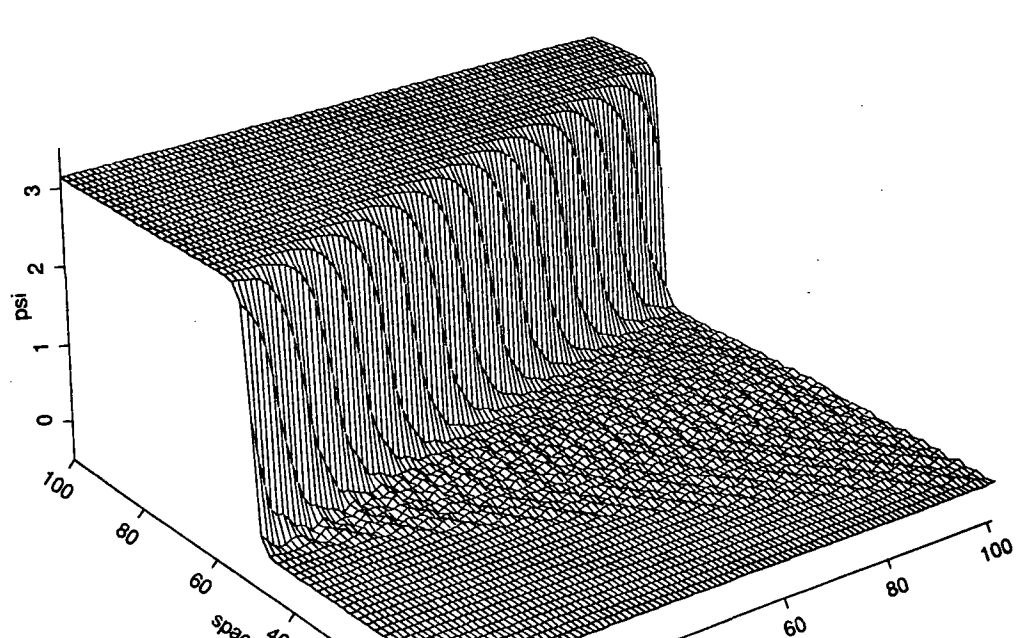

Figure 2.7: Radiation by a fast-moving kink.

propagating backwards (see figure 2.7).

The effect of this radiation on the kink velocity over a long time-scale can be seen in figure 2.8. The data were produced by the above-mentioned Runge-Kutta algorithm, run for 16000 time units with a time-step of 0.01. The initial configuration was a static kink Galilean-boosted to speed $v = 0.3$, on a lattice of unit spacing ($h = 1$). To cut reflexion of radiation from the fixed left-hand boundary, the first five lattice sites were damped. After an initial velocity drop of 0.02 in 10 time units as the kink assumes a more appropriate shape, it decelerates more slowly. The modulation of the amplitude of velocity oscillations is due to the velocity sampling (once every 10 time units) falling in and out of phase with the periodic wobble of the kink as it passes lattice sites.

The most interesting feature of figure 2.8 is the existence of a threshold velocity, $v \approx 0.16$, below which deceleration, and hence radiation, is much reduced. Some

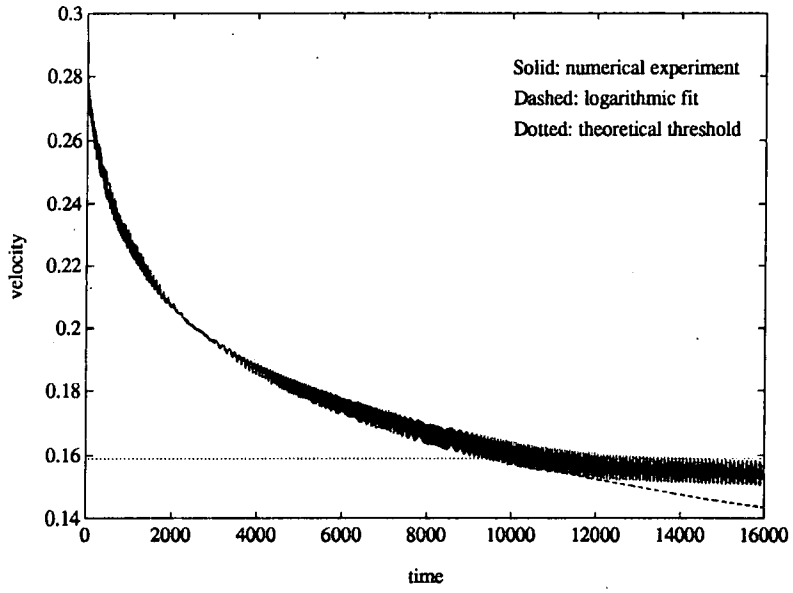

Figure 2.8: **Radiative kink deceleration.**

understanding of this phenomenon may be gained by an analysis (motivated by [18]) of the linearized equation of motion,

$$\ddot{\psi} = \frac{4-h^2}{4h^2}(\psi_+ + \psi_-) - \frac{4+h^2}{2h^2}\psi. \tag{2.29}$$

From this one derives a dispersion relation for small-amplitude travelling waves, namely,

$$\omega^2 = \frac{4+h^2}{2h^2} - \frac{4-h^2}{2h^2}\cos kh, \tag{2.30}$$

where k is the wave-number. The angular frequency ω ranges between 1 and $2/h$, the lower bound being responsible for the threshold velocity. As the kink travels along the lattice, it hits lattice sites with frequency v/h sites per unit time. Provided $v \geq h/2\pi$, the kink can excite radiation of the same frequency, that is $\omega = 2\pi v/h$. However, if $v < h/2\pi$, then $\omega < 1$, and the lattice cannot support such radiation. The kink can only excite higher harmonics, so the rate of energy dissipation is suddenly cut and the kink velocity becomes "quasi-stable." For $h = 1$, this threshold occurs at $v = 1/2\pi \approx 0.159$, in good agreement with the numerical data (figure 2.8).

We can use the dispersion relation (2.30) to find an upper bound on the group velocity $v_g := d\omega/dk$ of radiation of the lattice. It is convenient to use rescaled variables $\kappa := hk$ and $\Omega := h\omega/\sqrt{2 - h^2/2}$, and to define the constant $\gamma := (4 + h^2)/(4 - h^2) > 1$, so that the dispersion relation becomes

$$\Omega^2 = \gamma - \cos\kappa. \tag{2.31}$$

Differentiation with respect to κ (denoted by a prime) yields

$$2\Omega\Omega' = \sin\kappa, \tag{2.32}$$

and repeating gives

$$2\Omega\Omega'' + 2(\Omega')^2 = \cos\kappa. \tag{2.33}$$

At turning points of $v_g(k)$, $\Omega''(\kappa) = 0$, so

$$(\Omega')^2 = \frac{1}{2}\cos\kappa. \tag{2.34}$$

Note that this equation holds only at turning points of $\Omega' \propto v_g$, so that extremizing (2.34) to get $\Omega'_{max} = 1/2$ is spurious. Returning to (2.32), we find that

$$(\Omega')^2 = \frac{\sin^2\kappa}{4\Omega^2} = \frac{\sin^2\kappa}{4(\gamma - \cos\kappa)} \tag{2.35}$$

using (2.31). Thus, when κ satisfies

$$\frac{d}{d\kappa}\left\{\frac{\sin^2\kappa}{\gamma - \cos\kappa}\right\} = 0, \tag{2.36}$$

$\Omega' = 0$ (and hence $v_g = 0$) or $\Omega'' = 0$ and v_g is extremized. A little algebra gives three sets of roots of (2.36):

1. $\sin\kappa = 0 \quad \Rightarrow \quad \Omega' = 0$ from (2.32), $\quad \Rightarrow v_g = 0$

2. $\cos\kappa = \frac{2+h}{2-h} > 1$, false roots

3. $\cos\kappa = \frac{2-h}{2+h} \quad \Rightarrow \quad \Omega' \neq 0$ from (2.32), $\quad \Rightarrow \Omega'' = 0.$

Since $\Omega'' = 0$ in the third case, we can use (2.34) to find $(\Omega')^2$, whence, taking the rescalings into account, we find that

$$v_g^{max} = 1 - \frac{h}{2}. \tag{2.37}$$

In the continuum limit, $v_g^{max} \to 1$, the speed of light, while in the anticontinuum limit $v_g^{max} \to 0$ suggesting that radiation cannot travel along the lattice at all. In all the numerical simulations, the envelope of the radiative wake appeared to move with precisely this maximum speed. For example, the simulation depicted in figure 2.7 was performed on a lattice with $h = 1$, and thus $v_g^{max} = 1/2$. Close inspection reveals that the wake travels 50 length units backwards at apparently constant speed in 100 time units.

A theoretical understanding of the specific shape of the graph is more elusive. We make the ad-hoc assumption (motivated by figure 7 of [18]) that the energy ΔE lost by a kink in traversing a single lattice cell at speed v obeys an exponential law:

$$\Delta E = e^{pv-q}, \tag{2.38}$$

where p and q are positive constants, properties of the lattice. We further assume that the kinetic energy of the kink is

$$E_K = \frac{1}{2}cv^2 \tag{2.39}$$

where c is approximately constant provided h is not too large, as suggested by the results of section 3 (fig. 2). These two equations imply a first order differential equation

for $v(t)$, easily solved to give

$$v(t) = v(0) - A\log(Bt+1), \tag{2.40}$$

where

$$\begin{aligned} A &= \frac{1}{p}, \\ B &= \frac{p}{hc}e^{pv(0)-q}. \end{aligned} \tag{2.41}$$

The dashed curve in figure 2.8 is a fit of this formula to the numerical data, taking $A = 0.032, B = 0.004365, v(0) = 0.28$. Estimating $c = 1.0986$ by averaging the function $f(b)$ for $h = 1$, one deduces that $p = 31.3$, and $q = 17.5$. The fit is good for velocities greater than the radiation threshold at $v = 0.159$.

2.5 Conclusion

By preserving the Bogomol'nyi argument, one can find a discretization of the sine-Gordon model whose kink dynamics is significantly simpler than that of the Frenkel-Kontorova model. The Peierls-Nabarro barrier is eliminated, and although the dynamics is dissipative, so moving kinks suffer radiative deceleration, kinks are never pinned however slowly they travel. We conjecture that there are no nontrivial stable static solutions with regular finite-energy boundary conditions beyond those satisfying the Bogomol'nyi equation (kinks and antikinks). This is a first order difference equation whose solutions can be explicitly written down.

Several interesting questions remain. The existence of breathers (spatially localized, time periodic solutions) is comparatively rare in continuum field theories, and appears to be a property of integrable systems. This is not true of discrete systems: Aubry and MacKay [23] have proved the existence of breathers in a very wide class of one dimensional lattice models, at least in a neighbourhood of the anti-continuum

limit. The TDSG system does not fall inside this wide class due to its unconventional difference operator D (Aubry and MacKay essentially assume Δ, as in the Frenkel-Kontorova model). However the TDSG system does have the essential features needed for the proof, in particular, a trivial breather in the anti-continuum limit ($h \to 2$ in this case). Substituting $h = 2$ into the field equation (2.20), one obtains

$$\ddot{\psi} = -\frac{1}{2}\sin\psi(\cos\psi_+ + \cos\psi_-). \tag{2.42}$$

This supports a breather located at a single site, since a field of the form

$$\psi(x,t) = \begin{cases} 0 & x \neq 0 \\ \vartheta(t) & x = 0 \end{cases} \tag{2.43}$$

satisfies (2.42) provided

$$\ddot{\vartheta} + \sin\vartheta = 0, \tag{2.44}$$

that is, $\vartheta(t)$ satisfies the pendulum equation. One might therefore be able to extend the proof of [23] to this case.

One should be able to generalize "topological" discretization to any continuum model which has kinks. An example is the ϕ^4 model, also popular in condensed matter physics. The potential functional is

$$V_\phi = \frac{1}{4}\int dx \left[\left(\frac{\partial\phi}{\partial x}\right)^2 + (1-\phi^2)^2\right], \tag{2.45}$$

so the model has two degenerate vacua $\phi = \pm 1$, and a kink interpolating between them. If we define

$$\begin{aligned} d &:= \Delta\phi = \frac{\phi_+ - \phi}{h} \xrightarrow{h\to 0} \frac{\partial\phi}{\partial x} \\ f &:= 1 - \frac{1}{3}(\phi_+^2 + \phi_+\phi + \phi^2) \xrightarrow{h\to 0} 1 - \phi^2, \end{aligned} \tag{2.46}$$

then

$$df = -\Delta\left(\frac{1}{3}\phi^3 - \phi\right). \tag{2.47}$$

The lattice potential energy is

$$V := \frac{h}{4}\sum_{x\in h\mathbb{Z}}(d^2 + f^2) \xrightarrow{h\to 0} V_\phi, \tag{2.48}$$

as in the TDSG system. The Bogomol'nyi argument for kink configurations (boundary conditions $\lim_{x\to\pm\infty}\phi = \pm 1$) follows immediately:

$$\begin{aligned} 0 &\leq \frac{h}{4}\sum_x (d-f)^2 \\ &= V - \frac{h}{2}\sum_x \Delta\left(\frac{1}{3}\phi^3 - \phi\right) \\ \Rightarrow V &\geq \frac{2}{3} \\ \text{and} \quad V &= \frac{2}{3} \quad \Leftrightarrow \quad d = f. \end{aligned} \tag{2.49}$$

Unfortunately there is no explicit solution of the Bogomol'nyi equation $d = f$ (an ansatz of $\phi(x) = \tanh a(x-b)$ motivated by the continuum model does not work). It is therefore less easy to decide whether, as one might expect, the Peierls-Nabarro barrier has been eliminated. Recall that in the TDSG system this conclusion was reached when a *continuous* translation parameter was found in the kink solution. The essential point is that there exists a solution of $D = F$ satisfying kink boundary conditions for each $\psi(0) \in (0,\pi)$, that is, with the zeroth lattice site taking any value between the two vacua. The one-parameter curve in Q_1 covered by the solution ψ as $\psi(0)$ moves along the interval $(0,\pi)$ may then be defined to be the kink translation orbit. The analogous result can be shown to hold in the TDϕ^4 system described above, provided $h \leq 1$, using the boundedness and monotonicity properties of the iteration $\phi_+(\phi)$ derived from $d = f$. That is, there exists a unique monotonic solution of $d = f$ satisfying $\lim_{x\to\pm\infty}\phi = \pm 1$ for each $\phi(0) \in (-1,1)$. Given the absence of a

Peierls-Nabarro barrier, one would expect that there are no stable static multikink or kink-antikink solutions, but this is not certain. It is not clear whether there is any connexion between $\partial V/\partial\phi = 0$ and $\Delta(d^2 - f^2) = 0$; to answer this question requires the factorization of a quartic polynomial in three variables.

Generalizing topological discretization to higher dimensions is rather more difficult. In one dimension, continuum models have important discrete features even before discretization: the vacuum "manifold" and the boundary of space are both discrete, and consequently the Bogomol'nyi argument is already discrete. This is not so in higher dimensions. For example, it seems impossible to formulate a lattice $O(3)$ sigma model with a saturable Bogomol'nyi bound. One must be slightly less ambitious, and aim to find a model with as nearly saturable a bound as possible [24]. This will hopefully minimize, but not completely eliminate the Peierls-Nabarro barrier. The aim is to find lattice systems whose solitons closely mimic their continuum counterparts even on very coarse lattices, thereby saving considerable computational cost in numerical simulations.

Chapter 3

Some quantum aspects of the topological discrete sine-Gordon system

3.1 Introduction – the kink Casimir energy

Classically, a topologically stable field configuration may be regarded as lying in a potential well in an infinite dimensional configuration space. Two solutions in different sectors are separated by an infinite potential barrier by virtue of the topology. Quantum mechanically, a particle cannot sit at the bottom of a well: it always possesses a zero point energy dependent on the shape of the well bottom. The analogous situation applies to fields also.

We wish to find the zero point energy associated with a kink configuration in the discrete sine-Gordon system described in chapter 2. The calculation is performed in the weak coupling approximation, by performing a Taylor expansion of the potential about the kink configuration, truncating the expansion at quadratic order. We then find the normal modes of the truncated system, reducing the problem to that of an infinite set of decoupled harmonic oscillators. On quantization, each contributes a

zero point energy. Summing over all oscillators gives an infinite total.

If we similarly approximate the vacuum well-bottom we can derive the zero point energy associated with the trivial vacuum. This is also infinite. The "physical" quantity required is the energy difference between the kink and the vacuum, since we can always *define* the vacuum energy to be zero. This quantity, analogous to the Casimir energy of Quantum Electrodynamics [25], turns out to be finite and less than the classical energy of the kink. Since the model possesses only discrete translation symmetry, the normal mode frequencies, and hence the Casimir energy depend periodically on the kink position.

Calculation of the normal mode frequencies amounts to finding the eigenvalues of an infinite-order, tridiagonal, symmetric matrix. In practice this is not possible and the system must be truncated symmetrically about the kink centre, ignoring the large $|x|$ degrees of freedom. It is possible to show that the resulting quantum energy correction must be negative, for any size of truncated system. The correction may be calculated numerically.

Since we propose a perturbative quantum calculation we must introduce an explicit mass parameter m and a dimensionless coupling constant λ into the model. It is also convenient to change notation from $\psi(x)$, x taking discrete values, to ψ_j, the index j taking all integer values. In this notation the Lagrangian is

$$L = \frac{h}{4} \sum_{j \in \mathbb{Z}} \left[\dot{\psi}_j^2 - \frac{4}{\lambda^2 h^2} \sin^2 \frac{\lambda}{2}(\psi_{j+1} - \psi_j) - \frac{m^2}{\lambda^2} \sin^2 \frac{\lambda}{2}(\psi_{j+1} + \psi_j) \right], \tag{3.1}$$

where h is the lattice spacing just as before. Taking the $h \to 0$ limit one recovers the standard sine-Gordon model:

$$\lim_{h \to 0} L = \frac{1}{4} \int_{-\infty}^{\infty} dx \left[\left(\frac{\partial \psi}{\partial t} \right)^2 - \left(\frac{\partial \psi}{\partial x} \right)^2 - \frac{m^2}{\lambda^2} \sin^2 \lambda \psi \right], \tag{3.2}$$

where $x = hj$.

Let us briefly recapitulate the salient features of this TDSG system, using the new

notation. The model has an infinite set of discrete potential minima,

$$\psi = \frac{n\pi}{\lambda}, \tag{3.3}$$

where $n \in \mathbb{Z}$, and a static kink interpolating between neighbouring minima, which may be derived using a Bogomol'nyi argument. The energy of a static configuration is just the potential V, and

$$\begin{aligned} 0 &\leq \frac{h}{4}\sum_j \left(\frac{2}{\lambda h} \sin\frac{\lambda}{2}(\psi_{j+1} - \psi_j) - \frac{m}{\lambda}\sin\frac{\lambda}{2}(\psi_{j+1} + \psi_j) \right)^2 \\ &= V + \frac{m}{2\lambda^2}\sum_j (\cos\lambda\psi_{j+1} - \cos\lambda\psi_j) \\ \Rightarrow V &\geq \frac{m}{\lambda^2} \end{aligned} \tag{3.4}$$

when kink boundary conditions are imposed. This (the Bogomol'nyi) bound is saturated if and only if

$$\sin\frac{\lambda}{2}(\psi_{j+1} - \psi_j) = hm \sin\frac{\lambda}{2}(\psi_{j+1} + \psi_j). \tag{3.5}$$

The first order difference equation, (3.5), is called the Bogomol'nyi equation and, remarkably, has an explicit kink solution:

$$\psi_j = \frac{2}{\lambda}\tan^{-1}\left[\left(\frac{2+hm}{2-hm} \right)^{j-b/h} \right]. \tag{3.6}$$

The dimensionless parameter $hm \in (0,2)$ for sensible solutions. The arbitrary constant b may take any real value — the kink energy is not position dependent. The right hand side of (3.4) may be interpreted as the classical kink mass, M. Small velocity kink dynamics can then be approximated by geodesic motion of a point particle, mass M, on the manifold generated by b translation, with a natural induced metric.

3.2 The weak coupling approximation

We follow the method outlined in [12] adapted for the infinite lattice. We work in natural units, effectively having absorbed a factor $\sqrt{\hbar}$ into λ so that, by the standard argument, the weak coupling criterion $\lambda \ll 1$ yields a semi-classical approximation.

The momentum conjugate to ψ_j is

$$\pi_j = \frac{\partial L}{\partial \dot{\psi}_j} = \frac{h}{2}\dot{\psi}_j \tag{3.7}$$

Thus, the TDSG Hamiltonian is

$$H = \sum_j \frac{\pi_j^2}{h} + V(\psi) \tag{3.8}$$

where

$$V(\psi) = \frac{h}{4}\sum_j \left[\frac{4}{\lambda^2 h^2}\sin^2\frac{\lambda}{2}(\psi_{j+1} - \psi_j) + \frac{m^2}{\lambda^2}\sin^2\frac{\lambda}{2}(\psi_{j+1} + \psi_j)\right] \tag{3.9}$$

Let $\tilde{\psi}$ be a static configuration giving a local (in configuration space) minimum of the potential, $V(\psi)$. We treat motion about this stable configuration in the small λ approximation by Taylor expansion of V:

$$V(\psi) = V(\tilde{\psi}) + \frac{m}{2}\sum_{j,k} W_{jk}(\psi_j - \tilde{\psi}_j)(\psi_k - \tilde{\psi}_k) + \cdots \tag{3.10}$$

where

$$
\begin{aligned}
W_{jk} &= \frac{1}{m}\left.\frac{\partial^2 V}{\partial\psi_j\partial\psi_k}\right|_{\tilde{\psi}} \\
&= \frac{hm}{4}\Bigg(\delta_{j,k}\left\{\frac{2}{(hm)^2}\left[\cos\lambda(\tilde{\psi}_k-\tilde{\psi}_{k-1})+\cos\lambda(\tilde{\psi}_{k+1}-\tilde{\psi}_k)\right]\right. \\
&\quad \left.+\frac{1}{2}\left[\cos\lambda(\tilde{\psi}_k+\tilde{\psi}_{k-1})+\cos\lambda(\tilde{\psi}_{k+1}+\tilde{\psi}_k)\right]\right\} \\
&\quad +\delta_{j,k-1}\left[-\frac{2}{(hm)^2}\cos\lambda(\tilde{\psi}_k-\tilde{\psi}_{k-1})+\frac{1}{2}\cos\lambda(\tilde{\psi}_k+\tilde{\psi}_{k-1})\right] \\
&\quad +\delta_{j,k+1}\left[-\frac{2}{(hm)^2}\cos\lambda(\tilde{\psi}_{k+1}-\tilde{\psi}_k)+\frac{1}{2}\cos\lambda(\tilde{\psi}_{k+1}+\tilde{\psi}_k)\right]\Bigg). \qquad (3.11)
\end{aligned}
$$

Note that W is a real, symmetric, tridiagonal matrix. Note also that the next term in the expansion is

$$
\frac{1}{3!}\sum_{i,j,k}\left.\frac{\partial^3 V}{\partial\psi_i\partial\psi_j\partial\psi_k}\right|_{\tilde{\psi}}(\psi_i-\tilde{\psi}_i)(\psi_j-\tilde{\psi}_j)(\psi_k-\tilde{\psi}_k), \qquad (3.12)
$$

and that the three derivatives with respect to ψ introduce a factor of λ^3, leaving an overall factor of λ after cancelling the $1/\lambda^2$ of $V(\psi)$. Thus, higher corrections are at least of order λ, which is why one can truncate the series in the small λ limit.

Owing to the symmetry of the W matrix, there exists an orthogonal transformation R such that

$$
W_{jk} = \sum_{l,m} R^T_{jl} U_{lm} R_{mk} \qquad (3.13)
$$

where U is a diagonal matrix. We may reduce the system to a sequence of decoupled harmonic oscillators by transforming to the rotated coordinates (normal coordinates)

$$
\xi_j = \sum_k R_{jk}(\psi_k - \tilde{\psi}_k), \qquad (3.14)
$$

which have conjugate momenta

$$\eta_j = \sum_k R_{jk}\pi_k. \tag{3.15}$$

Then,

$$\frac{h}{2}\left(H - V(\tilde{\psi})\right) = \frac{1}{2}\sum_j \left(\eta_j^2 + \frac{hm}{2}\Omega_j^2\xi_j^2\right), \tag{3.16}$$

where Ω_j^2 are the eigenvalues of the W matrix, none of which can be negative since $\tilde{\psi}$ locally minimizes the potential.

We now quantize in standard canonical fashion, taking $\tilde{\psi}$ to be first the vacuum, then the kink located at $x = b$. The vacuum ground state energy is

$$E^0_{\{0\}} = \frac{m}{\sqrt{2hm}}\sum_j \Omega^0_j. \tag{3.17}$$

The kink, by virtue of the zero mode b of (3.6) lies not in a potential well, but in a level-bottomed valley meandering through configuration space. One of the normal modes is locally tangential to the valley bottom and consequently has vanishing corresponding eigenvalue (zero frequency). We shall treat this translation mode, b, classically because in the weak coupling approximation it is much heavier than the orthogonal modes (mass m/λ^2 compared with m).

While the translation orbit of the static kink (3.6) is an equipotential curve, neighbouring orbits are not: the potential varies periodically along them. So the eigenvalues of the W matrix $(\Omega^K_j)^2$ will be b dependent with period h. The ground state energy of a kink at b is

$$E^K_{\{0\}}(b) = \frac{m}{\lambda^2} + \frac{m}{\sqrt{2hm}}\sum_j \Omega^K_j(b). \tag{3.18}$$

where the sum may be taken over all eigenvalues since the one we wish to omit is zero anyway.

There is no reason to expect either $E^K_{\{0\}}(b)$ or $E^0_{\{0\}}$ to converge to a finite sum but

one would expect finite Casimir energy

$$\mathcal{E} = E^K_{\{0\}}(b) - E^0_{\{0\}} - \frac{m}{\lambda^2} \tag{3.19}$$

since the lattice spacing h has effectively cut off the ultra violet divergence problems of the continuum model. We must still take care when manipulating the divergent sums of (3.19) that they are suitably regulated before being combined. The method of regulation is determined by practical considerations: we calculate $E^K_{\{0\}}(b)$ and $E^0_{\{0\}}$ on a finite lattice, compute $\mathcal{E}$ and then allow the lattice size to grow large. In practice we must truncate the lattice to finite size anyway in order to solve the kink matrix eigenvalue problem.

3.3 The eigenvalue problem

We now address the problem of finding the eigenvalues of the vacuum and kink W matrices, W^0 and W^K. Let us first consider the vacuum matrix, obtained from (3.11) by substituting $\tilde{\psi}_k = 0$:

$$W^0_{jk} = \frac{hm}{4}\left[\left(\frac{4}{(hm)^2} + 1\right)\delta_{j,k} - \left(\frac{2}{(hm)^2} - \frac{1}{2}\right)(\delta_{j,k-1} + \delta_{j,k+1})\right]. \tag{3.20}$$

This is a very simple matrix. All the diagonal entries are equal, as are all the upper and lower diagonal entries. Consequently, the spectrum of the truncated matrix of order N is known exactly (see, for example, [26]):

$$(\Omega^0_j)^2 = \frac{4 + h^2m^2}{4hm} - \frac{4 - h^2m^2}{4hm}\cos\left(\frac{j\pi}{N+1}\right), \tag{3.21}$$

$j = 1, 2, \ldots, N$. In the $N \to \infty$ limit this is strikingly similar to the dispersion relation for phonons on the lattice (2.30) because the normal modes of oscillation about the vacuum *are* phonons.

The question of making sense of the $N \to \infty$ limit of $\sum_j \Omega_j^0$ is irrelevant because there is no hope of finding the exact spectrum of the kink W matrix, W^K. Substituting (3.6) into (3.11) we obtain an explicit expression for $W^K(b)$. Since the kink is highly localized (provided hm is not too small), at large $|j|$ the kink configuration rapidly approaches neighbouring vacuum minima so that $W^K_{jk}(b)$ tend to W^0_{jk} away from the matrix centre. The suggestion, then, is that we truncate the lattice, pinning all the large $|j|$ degrees of freedom, for which the kink and vacuum configurations are essentially identical, to their classical values. The resulting finite problem may be solved numerically.

One can prove that, whatever the order, N (an odd integer), of the symmetrically truncated system (that is, truncated symmetrically about the lattice site $j = 0$), the Casimir energy must be negative. The proof rests on the observation that, truncated to order $N = 2n + 1$,

$$W^K = W^0 + B + C, \tag{3.22}$$

where B is the symmetric, tridiagonal matrix,

$$B = \begin{pmatrix} -\beta_{-n} & b_{-n} & 0 & & & & \\ b_{-n} & -\beta_{-n+1} & b_{-n+1} & & & & \\ 0 & b_{-n+1} & -\beta_{-n+2} & & & & \\ & & & \ddots & & & \\ & & & & -\beta_{n-2} & b_{n-1} & 0 \\ & & & & b_{n-1} & -\beta_{n-1} & b_n \\ & & & & 0 & b_n & -\beta_n \end{pmatrix}, \tag{3.23}$$

$$\begin{aligned} b_j &= \frac{hm}{8}\left(1 - \cos\lambda(\tilde{\psi}_j + \tilde{\psi}_{j-1})\right) + \frac{1}{2hm}\left(1 - \cos\lambda(\tilde{\psi}_j - \tilde{\psi}_{j-1})\right) \qquad (3.24) \\ &\geq 0, \\ \beta_j &= b_j + b_{j+1}, \qquad (3.25) \end{aligned}$$

and C is the diagonal matrix with elements,

$$C_{ij} = \delta_{ij}\frac{hm}{4}\left[\cos\lambda(\tilde{\psi}_{j+1} + \tilde{\psi}_j) + \cos\lambda(\tilde{\psi}_j + \tilde{\psi}_{j-1}) - 2\right]. \tag{3.26}$$

Now, C manifestly has negative semi-definite eigenvalues, and B is a special case of a class of matrices whose spectra are known to be negative semi-definite [27]. Regarding B and C as perturbations to the matrix W^0, we apply a corollary of the minimax theorem [28]:

Theorem: If Θ, Υ, Γ are symmetric $N \times N$ matrices with eigenvalues $\theta_r, \upsilon_s, \gamma_t$ (all sets arranged in non-increasing order), and

$$\Theta = \Upsilon + \Gamma,$$

then

$$\theta_r \leq \upsilon_r + \gamma_1$$

for all $r = 1, 2, \ldots, N$.

Thus the eigenvalues of $W^0 + B$ are shifted down relative to those of W^0 (the greatest eigenvalue of B is negative or zero), and similarly, the eigenvalues of $W^0 + B + C = W^K$ are shifted down relative to those of $W^0 + B$. Hence,

$$\sum_j \Omega_j^K(b) \leq \sum_j \Omega_j^0, \tag{3.27}$$

and, substituting into (3.19),

$$\mathcal{E}(b) \leq 0, \tag{3.28}$$

the quantum mechanical effect must be to *lower* the kink energy.

We have assumed, on physical grounds, that

$$\mathcal{E}_N(b) = \frac{m}{\sqrt{2hm}}\sum_j \left(\Omega_j^K(b) - \Omega^0\right), \tag{3.29}$$

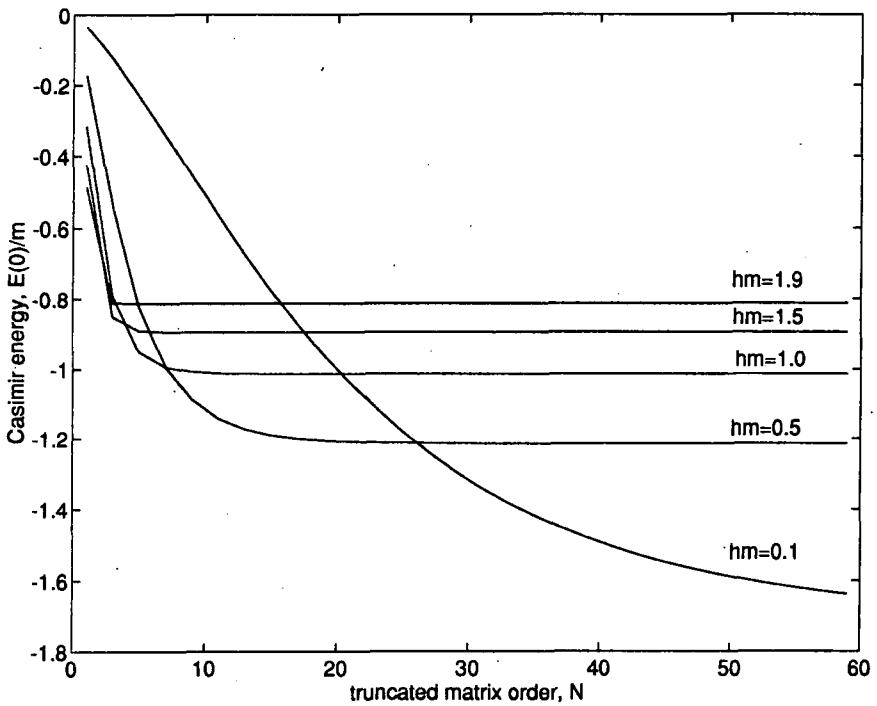

Figure 3.1: Convergence of the Casimir energy.

the Casimir energy of the order N truncated system converges to a constant value as N grows large. Numerical evidence for this is presented in figure 3.1, a graph of N against $\mathcal{E}_N(0)/m$ for various values of hm. Convergence is very fast for large hm (the $hm = 1.9$ curve is essentially flat for $N \geq 3$) but much slower for small hm. Of course, this happens because the more finely space is discretized, the more degrees of freedom the kink structure is spread over. There is also a much smaller exacerbating effect due to the hm dependence in the kink solution, (3.6) — large hm kinks are sharper in "real" x space than small hm kinks. Note that, as proved above, the Casimir energy is always negative and that its magnitude grows large for small hm as we expect from consideration of the unrenormalized continuum model [29]. Although $b = 0$ was chosen for these data, the rates of convergence are virtually independent of $b \in [-h/2, h/2]$.

The Casimir energy of suitably truncated systems is plotted as a function of kink

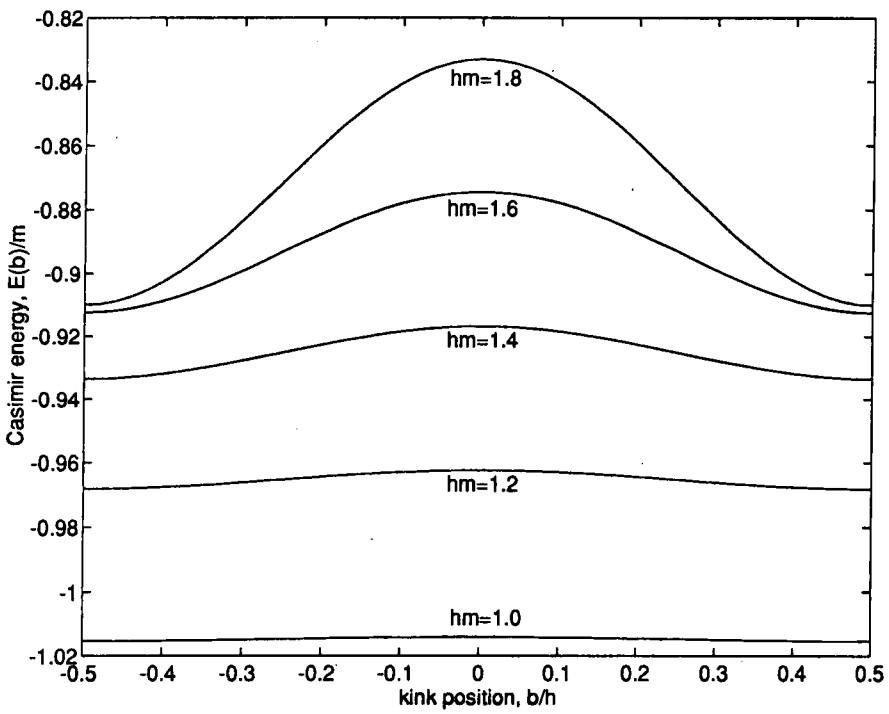

Figure 3.2: Position dependence of the Casimir energy.

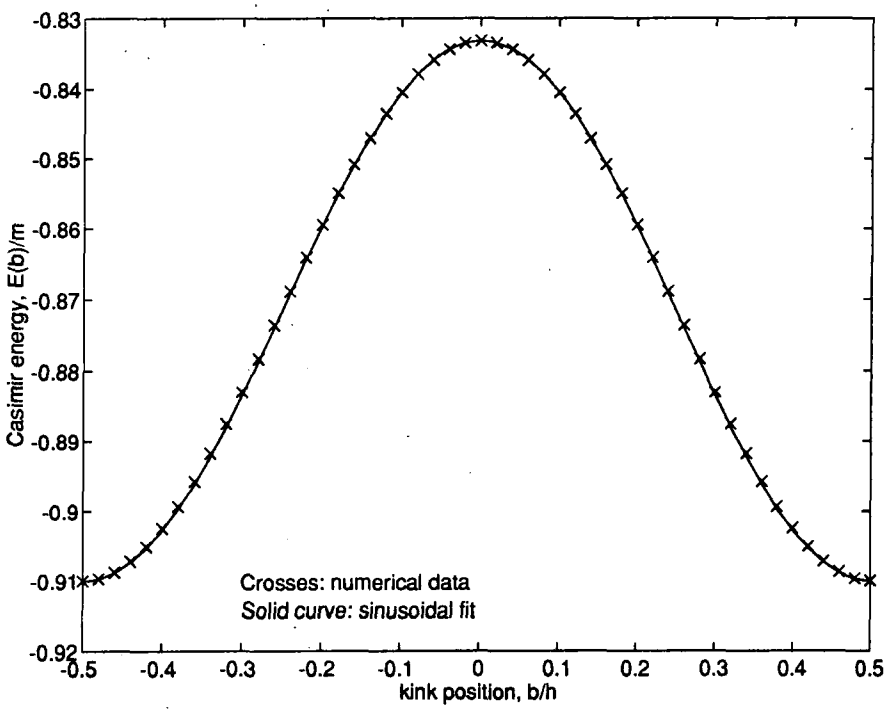

Figure 3.3: Sinusoidal fit to the $hm = 1.8$ Casimir energy.

position b in figure 3.2. It attains its maxima where the kink is located exactly on a lattice site and its minima where the kink is halfway between two sites. For large hm the resemblance to a sinusoidal curve is remarkable: figure 3.3 shows a sinusoidal fit for $hm = 1.8$. The resemblence deteriorates as hm decreases because numerical errors become relatively large as the Casimir energy barrier $[\mathcal{E}(0) - \mathcal{E}(h/2)]/m$ decreases (figure 3.4). The Casimir energy is probably not truly sinusoidal, however, because strong anharmonicity has been found to develop as hm approaches the anti-continuum limit ($hm \to 2$) very closely[1]. If one were to use $\mathcal{E}(b)$ in a phenomenological calculation, taking a truncated Fourier series would seem to be a convenient and justifiable approximation.

It is the Casimir energy barrier (figure 3.4) which is physically most relevant and which will directly affect classical kink dynamics. In the continuum limit full

[1]I am grateful to Adrian Kent for pointing this out to me

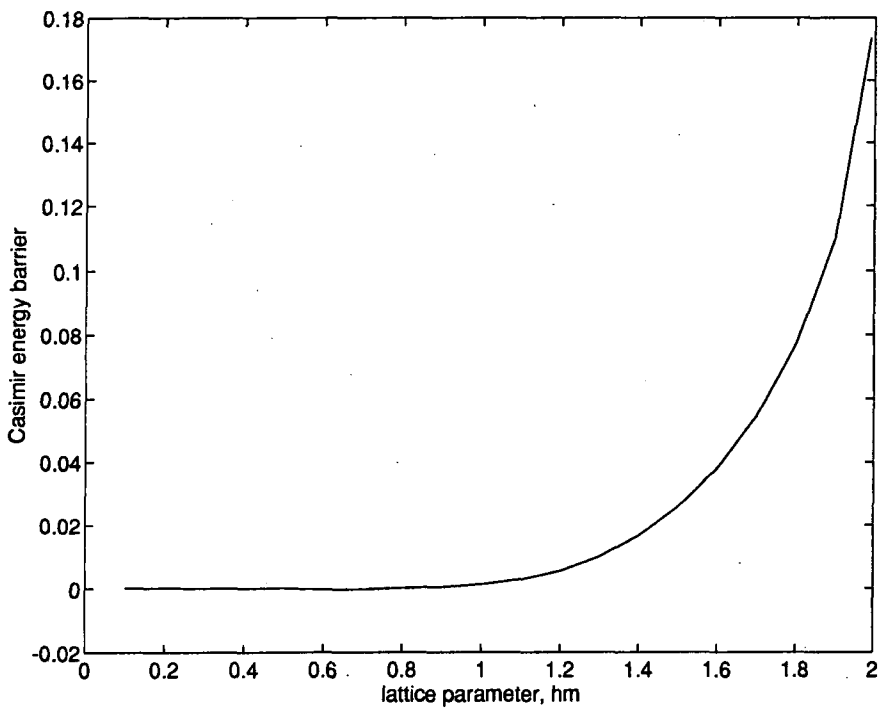

Figure 3.4: The effect of discreteness on the Casimir energy barrier.

continuous translation symmetry is recovered, so the barrier should disappear, as is indicated by the plot. In fact, in common with other (classical kinetic) discreteness effects of the TDSG system, the barrier is very small for all $hm < 1$. If $\mathcal{E}(b)/m$ for a given b/h is a strictly increasing function of $hm \in (0,2)$, as it appears to be, then given that $\mathcal{E}(b) < 0$, the barrier should approach a finite value in the $hm \to 2$ limit. Numerical evidence suggests that it does and that this value is around 0.17.

3.4 Concluding remarks

We have seen that quantum fluctuations around the kink configuration spoil the level kink valley bottom by introducing a Casimir energy which depends periodically on the (classical) kink position. This energy has been computed numerically in the weak coupling approximation and found to be approximately sinusoidal, maximum when $b = 0, \pm h, \pm 2h, \ldots$ and minimum when $b = \pm h/2, \pm 3h/2, \pm 5h/2, \ldots$, the difference between these extrema being large for large hm but rapidly vanishing in the continuum limit. It is superficially similar to the Peierls-Nabarro potential of the Frenkel-Kontorova model [18], but is entirely different in origin, being a purely quantum effect. Since the Casimir effect is a genuine physical phenomenon, experimentally verified in the context of Quantum Electrodynamics, we are led to the conclusion that classical kinks in this lattice model may be "pinned" by the quantum mechanics of the orthogonal modes.

Finally, following a suggestion by Gibbons and Manton [30] we could attempt to include the effect of variation of orthogonal mode frequencies on the quantized geodesic approximation of kink motion by including the Casimir energy as an extra potential term in the Hamiltonian:

$$\widehat{H}_{GA} = -\frac{1}{2}\Delta + \mathcal{E}(b) \tag{3.30}$$

where Δ is the covariant Laplacian on the submanifold of static kinks. Given the

periodic nature of $\mathcal{E}$, we would then expect band structure in the kink spectrum. There are two objections to this. First, the kink kinetic terms are of order λ^2 (due to the large kink mass), while the potential was expanded only up to order λ^0, so the perturbative expansion is not consistent. Second, picking out the $-\frac{1}{2}\Delta$ kinetic term at order λ^2 is also not consistent because there are several other kinetic terms at this order neglected in equation (3.30). This may be seen by reëxpressing the kinetic Hamiltonian of the full quantum field theory,

$$\hat{T} = \frac{1}{h}\sum_j \frac{\partial^2}{\partial \psi_j^2}, \tag{3.31}$$

(an infinite dimensional Laplacian) in terms of the kink translation mode b and the orthogonal modes ξ_j. The resulting formula is very messy, and includes order λ^2 cross terms with derivatives $\partial^2/\partial\xi_j\partial b$ and $\partial^2/\partial\xi_j\partial\xi_k$ in addition to the terms included in (3.30). The analogous expression in the continuum model [12] can be greatly simplified by boosting to the kink's rest frame, but this trick is not available to us here. These considerations cast doubt on the suggested procedure, at least in this case. However, one should note that a naive expansion of the Hamiltonian in λ may not be the most physically relevant procedure when considering quantum kink dynamics. For example, one could imagine making the demand that the kink's "speed" (in an appropriate quantum sense) be of order λ^0. Given the kink's large mass in the weak coupling approximation, one would then expect the kink kinetic term to *dominate* the Hamiltonian. So the legitimacy of the suggested procedure remains an open and somewhat controversial question.

Chapter 4

The dynamics of a $\mathbb{C}P^1$ lump on the sphere

4.1 Introduction

Sigma models have long been popular objects of study among theoretical physicists. They emerge as low energy effective theories of nuclear and particle physics and have applications in condensed matter physics. They also provide a useful testing ground for ideas whose eventual objective may concern more complicated theories. Probably the most studied sigma model is the $\mathbb{C}P^1$ model in $(2+1)$ dimensions. In this case the problem of soliton dynamics can only be attempted analytically using the geodesic approximation. However, as we shall see, the model defined on flat space suffers from singularities in this approximation which is thus ill-defined. So we are motivated to consider the model defined on the two-sphere. Soliton dynamics in curved space is, in any case, an interesting subject.

The $\mathbb{C}P^1$ model in flat space is a scalar field theory whose configuration space Q consists of finite energy maps from Euclidean $\mathbb{R}^2$ to the complex projective space $\mathbb{C}P^1$, the energy functional being constructed naturally from the Riemannian structures of the base and target spaces (that is, the model is a pure sigma model in the broad

sense). The requirement of finite energy imposes a boundary condition at spatial infinity, that the field approaches the same constant value, independent of direction in $\mathbb{R}^2$, so that the field may be regarded as a map from the one point compactification $\mathbb{R}^2 \cup \{\infty\} \cong S^2$ to $\mathbb{C}P^1$. Since $\mathbb{C}P^1 \cong S^2$ also, finite energy configurations are effectively maps $S^2 \to S^2$, the homotopy theory of which is well understood, and the configuration space is seen to consist of disconnnected sectors Q_n labelled by an integer n, the "topological charge" (degree),

$$Q = \bigcup_{n \in \mathbb{Z}} Q_n. \tag{4.1}$$

Each configuration is trapped within its own sector because time evolution is continuous.

The Lorentz invariant, time-dependent model is not integrable but complete solution of the static problem has been achieved by means of a Bogomol'nyi argument and the general charge n moduli space, the space of charge-n static solutions $\mathrm{M}_n \subset Q_n$, is known (that *all* static, finite energy solutions of the $\mathbb{C}P^1$ model saturate the Bogomol'nyi bound is a non-trivial result [31]). Each static solution within the charge-n sector has the same energy (minimum within that sector and proportional to n), and M_n is parametrized by $4n + 2$ parameters (the moduli), so such a moduli space may be thought of as the $(4n+2)$-dimensional level bottom of a potential valley defined on the infinite dimensional charge-n sector, Q_n. Low energy *dynamics* may be approximated by motion restricted to this valley bottom, a manifold embedded in the full configuration space, and thus inheriting from it a non-trivial metric induced by the kinetic energy functional. The approximate dynamic problem is reduced to the geodesic problem with this metric, and has been investigated by several authors [32, 33]. In the unit-charge sector one here encounters a difficulty: certain components of the metric are singular and the approximation is ill defined. For example, unit-charge static solutions are localized lumps of energy with arbitrary spatial scale, so one of the six

moduli of M_1 is a scale parameter. Motion which changes this parameter is impeded by infinite inertia in the geodesic approximation, a result in conflict with numerical evidence which suggests that lumps collapse under scaling perturbation [34].

This problem should not be present in the model defined on a compact two dimensional physical space. The obvious choice is the 2-sphere because the homotopic partition of the configuration space carries through unchanged. Also, S^2 with the standard metric is conformally equivalent to Euclidean $\mathbb{R}^2 \cup \{\infty\}$, and the static $\mathbb{C}P^1$ model energy functional is conformally invariant, so the whole flat space static analysis is still valid and all the moduli spaces are known. However, the kinetic energy functional *does* change and induces a new, well defined metric on the unit-charge moduli space. By means of the isometry group derived from the spatial and internal symmetries of the full field theory we can place restrictions on the possible structure of this metric, greatly simplifying its evaluation. The geodesic problem is still too complicated to be solved analytically in general, but by identifying totally geodesic submanifolds, it is possible to obtain the qualitative features of a number of interesting solutions. In particular, the possibilities for lumps travelling around the sphere are found to be unexpectedly varied.

4.2 The $\mathbb{C}P^1$ model on S^2

The $\mathbb{C}P^1$ model on the 2-sphere is defined by the Lagrangian

$$L[W] = \int_{S^2} dS \, \frac{\partial_\mu W \partial_\nu \bar{W}}{(1+|W|^2)^2} \eta^{\mu\nu} \tag{4.2}$$

where W is a complex valued field, dS is the invariant S^2 measure and $\eta^{\mu\nu}$ are the components of the inverse of the Lorentzian metric

$$\eta = dt^2 - d^2\Omega \tag{4.3}$$

on $\mathbb{R}(\text{time})\times S^2(\text{space})$, $d^2\Omega$ being the natural metric on S^2. Although the language of the $\mathbb{C}P^1$ model is analytically convenient, the homotopic classification and physical meaning of the field configurations are more easily visualized if we exploit the well known equivalence to the $O(3)$ sigma model [35, 36]. In the latter, the scalar field is a three dimensional isovector $\boldsymbol{\phi}$ constrained to have unit length with respect to the Euclidean $\mathbb{R}^3$ norm ($\boldsymbol{\phi}\cdot\boldsymbol{\phi}\equiv 1$), that is, the target space is the 2-sphere of unit radius with its natural metric, which we will denote S^2_{iso} for clarity. (The suffix refers to "isospace" in analogy with the internal space of nuclear physics models.) The $\mathbb{C}P^1$ field W is then thought of as the stereographic image of $\boldsymbol{\phi}$ in the equatorial plane, projected from the North pole, $(0,0,1)$. Explicitly,

$$\boldsymbol{\phi} = \left(\frac{W+\bar{W}}{1+|W|^2}, \frac{W-\bar{W}}{i(1+|W|^2)}, \frac{|W|^2-1}{1+|W|^2}\right) \tag{4.4}$$

and

$$W = \frac{\phi_1 + i\phi_2}{1-\phi_3}. \tag{4.5}$$

Then

$$L[W] \equiv L_\sigma[\boldsymbol{\phi}] = \frac{1}{4}\int_{S^2} dS\, \partial_\mu\boldsymbol{\phi}\cdot\partial_\nu\boldsymbol{\phi}\,\eta^{\mu\nu} \tag{4.6}$$

the familiar $O(3)$ sigma model Lagrangian. A W configuration, then, may be visualized as a distribution of unit length arrows over the surface of the physical 2-sphere S^2_{sp}. Each smooth map $S^2_{\text{sp}}\to S^2_{\text{iso}}$ falls into one of a discrete infinity of disjoint homotopy classes, each class associated with a unique integer which may be thought of as the topological degree of the map (see, for example [9]), so homotopic partition of the configuration space is built in to the model from the start.

We also choose stereographic coordinates (x,y) on S^2_{sp}, in terms of which,

$$(\eta_{\mu\nu}) = \text{diag}\left(1, \frac{-1}{(1+r^2)^2}, \frac{-1}{(1+r^2)^2}\right) \tag{4.7}$$

where $r=\sqrt{x^2+y^2}$, (x,y) takes all values in $\mathbb{R}^2$ and $x^0=t$, $x^1=x$, $x^2=y$. The

radius of S^2_{sp} has been normalized to unity. The invariant measure is,

$$dS = dx\,dy\,\sqrt{|\det(\eta_{\mu\nu})|} = \frac{dx\,dy}{(1+r^2)^2} \tag{4.8}$$

and so,

$$L[W] = \int dx\,dy\,\frac{1}{(1+|W|^2)^2}\left(\frac{|\dot{W}|^2}{(1+r^2)^2} - \left|\frac{\partial W}{\partial x}\right|^2 - \left|\frac{\partial W}{\partial y}\right|^2\right). \tag{4.9}$$

We identify kinetic energy,

$$T[W] = \int \frac{dx\,dy}{(1+r^2)^2}\frac{|\dot{W}|^2}{(1+|W|^2)^2} \tag{4.10}$$

and potential energy

$$V[W] = \int dx\,dy\,\frac{1}{(1+|W|^2)^2}\left(\left|\frac{\partial W}{\partial x}\right|^2 + \left|\frac{\partial W}{\partial y}\right|^2\right). \tag{4.11}$$

Note that the potential energy is identical to that for flat space by virtue of the conformal invariance of the static model (stereographic projection is a conformal transformation). Thus the familiar Bogomol'nyi argument [12] follows immediately and (i,j,k run over $1,2$ and $\times$ represents the $\mathbb{R}^3$ vector product in $\boldsymbol{\phi}$ space):

$$\begin{aligned}
0 &\leq \int dx\,dy\,(\partial_i\boldsymbol{\phi} \pm \epsilon_{ij}\boldsymbol{\phi}\times\partial_j\boldsymbol{\phi})\cdot(\partial_i\boldsymbol{\phi} \pm \epsilon_{ik}\boldsymbol{\phi}\times\partial_k\boldsymbol{\phi}) \\
&= 2\int dx\,dy\,[\partial_i\boldsymbol{\phi}\cdot\partial_i\boldsymbol{\phi} \mp \epsilon_{ij}(\partial_i\boldsymbol{\phi}\times\partial_j\boldsymbol{\phi})\cdot\boldsymbol{\phi}], \\
\Rightarrow V[W] &= \frac{1}{4}\int dx\,dy\,\partial_i\boldsymbol{\phi}\cdot\partial_i\boldsymbol{\phi} \\
&\geq \frac{1}{2}\left|\int dx\,dy\left(\frac{\partial\boldsymbol{\phi}}{\partial x}\times\frac{\partial\boldsymbol{\phi}}{\partial y}\right)\cdot\boldsymbol{\phi}\right| = 2\pi|n|,
\end{aligned} \tag{4.12}$$

where W is in the degree n homotopy class, equality holding if and only if

$$\partial_i\boldsymbol{\phi} = \mp\epsilon_{ij}\boldsymbol{\phi}\times\partial_j\boldsymbol{\phi}, \tag{4.13}$$

which, on substitution of (4.4) becomes the Cauchy-Riemann condition for W to be an analytic function of $z = x + iy$ (upper sign) or $\bar{z} = x - iy$ (lower sign). The former (latter) case corresponds to static solutions of positive (negative) degree, and if W is single valued with finite degree n, then it must be a rational map of degree n in z if $n \geq 0$ or in $\bar{z}$ if $n < 0$. We shall deal with the unit charge moduli space, consisting of all rational maps of degree 1 in z. Since the configuration space and moduli spaces of the flat space and spherical space models are diffeomorphic, we shall use the same notation (Q, Q_n, M_n etc.) in both cases.

4.3 The unit-charge moduli space

The simplest static unit-charge solution is

$$W(z) = z \tag{4.14}$$

which we shall call the symmetric hedgehog because its ϕ field points radially outwards at all points on S^2_{sp}.. Its energy density is uniformly distributed, so it is not really a lump. Since the static model is conformally invariant, any configuration obtained from this by a Möbius transformation must be another point on the moduli space. In fact the orbit of $W = z$ under the Möbius group *is* the space of degree 1 rational maps, each map being generated by one and only one group element. Thus we may identify the moduli space with the parameter space of the Möbius group.

There is a well known matrix representation of Möbius transformations [37] which we denote thus:

$$W(z) = \frac{az + b}{cz + d} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \odot z = M \odot z \tag{4.15}$$

where $M \in GL(2, \mathbb{C})$ so that $\det M \neq 0$. The last condition ensures the invertibilty of the transformation and fixes the degree of W at 1. The Möbius group product

becomes matrix multiplication,

$$M_2 \odot (M_1 \odot z) = (M_2 M_1) \odot z \tag{4.16}$$

where the left hand side means

$$M_2 \odot (M_1 \odot z) = \frac{a_2(M_1 \odot z) + b_2}{c_2(M_1 \odot z) + d_2} \tag{4.17}$$

in obvious notation. All matrices differing by a constant factor yield the same configuration, and $\det M \neq 0$ so when we divide by this scaling equivalence we can choose a unimodular matrix as the representative for each equivalence class. There are two such matrices possible for each distinct configuration because if M is unimodular, so is $-M$. Thus $SL(2,\mathbb{C})$ is a double cover of the moduli space, which we recover by dividing out the equivalence $M' \sim M \Leftrightarrow M' = \pm M$: the moduli space is $SL(2,\mathbb{C})/\mathbb{Z}_2$.

Coincidentally, $SL(2,\mathbb{C})$ is also a double cover of the proper orthochronous Lorentz group. The statement that any Lorentz transformation may be formed by a unique composition of a boost then a rotation (or *vice versa*) translates to the existence, for all $M \in SL(2,\mathbb{C})$, of $U \in SU(2)$ and H, a positive definite, unimodular, Hermitian 2×2 matrix (call this set $\mathcal{H}$), satisfying

$$M = UH \tag{4.18}$$

both U and H being unique [38]. It follows that the space $SL(2,\mathbb{C})$ is locally a product of S^3 (the group manifold of $SU(2)$) and $\mathbb{R}^3$ (the parameter space of $\mathcal{H}$), a result which generalizes globally, $SL(2,\mathbb{C}) \cong S^3 \times \mathbb{R}^3$.

We may choose local coordinates on $SL(2,\mathbb{C})$ by defining the standard Euler

angles (α, β, γ) on S^3,

$$U = \begin{pmatrix} \cos\frac{\alpha}{2} e^{i(\beta+\gamma)/2} & \sin\frac{\alpha}{2} e^{i(\beta-\gamma)/2} \\ -\sin\frac{\alpha}{2} e^{-i(\beta-\gamma)/2} & \cos\frac{\alpha}{2} e^{-i(\beta+\gamma)/2} \end{pmatrix} \tag{4.19}$$

and expanding H in terms of Pauli matrices $\boldsymbol{\tau}$,

$$H = \Lambda \mathbb{I} + \boldsymbol{\lambda} \cdot \boldsymbol{\tau} = \begin{pmatrix} \Lambda + \lambda_3 & \lambda_1 - i\lambda_2 \\ \lambda_1 + i\lambda_2 & \Lambda - \lambda_3 \end{pmatrix} \tag{4.20}$$

$\Lambda(\boldsymbol{\lambda})$ being chosen to ensure the unimodular and postive definite properties:

$$\Lambda = \sqrt{1 + \lambda^2}. \tag{4.21}$$

The 3-vector $\boldsymbol{\lambda}$ (modulus $|\boldsymbol{\lambda}| = \lambda$) takes all values in $\mathbb{R}^3$, while $\beta \in [0, 4\pi]$, $\gamma \in [0, 2\pi]$ and $\alpha \in [0, \pi]$. These ranges allow M to take all values in the double cover $SL(2, \mathbb{C})$. In analyzing the structure of the metric, it is convenient to work with $SL(2, \mathbb{C})$, checking that the metric is single valued under the identification of M with $-M$. The true moduli space $SL(2, \mathbb{C})/\mathbb{Z}_2$ is charted by the same coordinates but with β lying in the reduced range $[0, 2\pi]$, for U is then restricted to the "upper half" of S^3. The chart has a coordinate singularity at $\alpha = 0$ and at $\alpha = \pi$. The explicit connexion between a point in M_1 and the corresponding static solution will be made in section 4.5, below.

4.4 The induced metric and its isometry group

Field dynamics of the $\mathbb{C}P^1$ model may be visualized as the dynamics of a point particle with "position" $W : S^2_{\mathrm{sp}} \to S^2_{\mathrm{iso}}$ moving in an infinite-dimensional configuration space. A solution $W(t, x, y)$ of the field equations is thought of as a trajectory in this space, motion on which is determined by metric $T[W]$ and potential $V[W]$. In the

unit-charge sector, the Bogomol'nyi argument shows that there is a six-dimensional subspace on which the potential achieves its topological minimum value of 2π, and that any perturbation departing from this subspace must involve increasing V. If a configuration sitting at the bottom of this potential valley is given a small velocity tangential to it then we expect the ensuing time-evolved field to stay close to the valley bottom, for departure from it entails climbing up the valley walls. In the geodesic approximation [10] we restrict motion to the valley bottom, assuming that orthogonal modes are insignificant.

Thus, at all times $W(t,x,y)$ is a solution of the *static model*, but we allow the moduli $\{\boldsymbol{\lambda},\alpha,\beta,\gamma\}$, denoted collectively by $\{q^i : i = 1\ldots6\}$, to vary with time in accordance with the inherited action principle. So,

$$\dot{W} = \frac{\partial W}{\partial q^i}\dot{q}^i, \tag{4.22}$$

and the Lagrangian is

$$L = T - V = \int \frac{dx\,dy}{(1+r^2)^2}\frac{\partial W}{\partial q^i}\frac{\partial \bar{W}}{\partial q^j}\dot{q}^i\dot{q}^j - 2\pi. \tag{4.23}$$

Defining the induced metric,

$$g(q) = g_{ij}(q)\,dq^i\,dq^j = 2\int \frac{dx\,dy}{(1+r^2)^2}\frac{\partial W}{\partial q^i}\frac{\partial \bar{W}}{\partial q^j}\,dq^i\,dq^j \tag{4.24}$$

and ignoring the irrelevant constant, the Lagrangian is recast as that of a free particle moving on a Riemannian manifold with metric g:

$$L = \frac{1}{2}g_{ij}(q)\dot{q}^i\dot{q}^j. \tag{4.25}$$

The equations of motion are the geodesic equations. In principle all we need do is evaluate the integrals of (4.24), but these are 21 functions of 6 variables so as it stands this is intractable in practice. It is profitable to take a more circumspect approach,

using symmetries of the model to place restrictions on the structure of g.

Consider the rotation group $SO(3)$ acting on $S^2_{\rm sp}$ and $S^2_{\rm iso}$. The former is the group of spatial rotations (under which W, or equivalently ϕ, transforms as a scalar) while the latter is the group of global internal rotations (henceforth called "isorotations") of the ϕ field of which W is the stereographic image. Any such transformation $\mathcal{T}$ leaves invariant (a) the topological charge, so $\mathcal{T}$ is a bijection $\mathcal{T} : Q_n \to Q_n$, (b) the potential energy, so within Q_n static solutions are mapped to other static solutions, $\mathcal{T} : \mathrm{M}_n \to \mathrm{M}_n$, and (c) the kinetic energy, which induces the metric on M_n. Hence $\mathcal{T}$ is an isometry of (M_1, g). The $SU(2)$ subgroup of the Möbius group's double cover, $SL(2,\mathbb{C})$, acting via the operation $\odot$ defined by equation (4.15) is the double cover of the group of rotations of the 2-sphere [37] considered as operations on the projective plane (spatial or internal, ie acting on $z = x + iy$ or W). Thus we find that $(L, R \in SU(2), M \in SL(2,\mathbb{C}))$,

$$M \mapsto LM \Rightarrow W(z) \mapsto L \odot (W(z)) \tag{4.26}$$

produces an isorotation of the configuration $W(z) = M \odot z$, while

$$M \mapsto MR \Rightarrow W(z) \mapsto W(R \odot z) \tag{4.27}$$

produces a spatial rotation, both isometries of the induced metric.

The action of the isorotation on the moduli space is simple:

$$\begin{aligned} M = UH &\mapsto LUH \\ \Rightarrow U &\mapsto LU \\ H &\mapsto H. \end{aligned} \tag{4.28}$$

The isometry takes the $SU(2)$ left multiplication action on $S^3/\mathbb{Z}_2$ while leaving the $\mathbb{R}^3$ moduli $\boldsymbol{\lambda}$ unchanged. Using a technique standard in the analysis of isometries in

general relativity [39], we change from the coordinate basis on $S^3/\mathbb{Z}_2$, $\{d\alpha, d\beta, d\gamma\}$, to a non-coordinate basis, in this case the left-invariant 1-forms of the Lie group $SU(2)$. These may be found by expanding the left-invariant 1-form $U^{-1}dU$ in terms of a convenient basis of the Lie algebra $su(2)$, for example $i\boldsymbol{\tau}/2$. Explicitly,

$$U^{-1}dU = \boldsymbol{\sigma} \cdot \left(\frac{i}{2}\boldsymbol{\tau}\right) \tag{4.29}$$

where

$$\begin{aligned} \sigma_1 &= -\sin\gamma\, d\alpha + \cos\gamma \sin\alpha\, d\beta \\ \sigma_2 &= \cos\gamma\, d\alpha + \sin\gamma \sin\alpha\, d\beta \\ \sigma_3 &= \cos\alpha\, d\beta + d\gamma. \end{aligned} \tag{4.30}$$

If we evaluate the metric at one particular point on $S^3/\mathbb{Z}_2$, for all possible $\boldsymbol{\lambda} \in \mathbb{R}^3$, we can obtain the metric at all other points on $S^3/\mathbb{Z}_2$ because $S^3/\mathbb{Z}_2$ is the isorotation orbit of our base point, and isorotation is an isometry, so the metric must remain constant (for each $\boldsymbol{\lambda}$) over the entire orbit. "Constant" means unchanging when considered as a geometric object, not that the components with respect to the original coordinate basis are constant, because the basis vectors themselves transform non-trivially. The basis of (4.30) *is* invariant however, so the metric must be of the form

$$g = \mu_{ab}(\boldsymbol{\lambda})\sigma_a\sigma_b + \nu_{ab}(\boldsymbol{\lambda})\sigma_a d\lambda_b + \pi_{ab}(\boldsymbol{\lambda})d\lambda_a d\lambda_b, \tag{4.31}$$

where $a, b = 1, 2, 3$ and each of the component functions is independent of (α, β, γ).

Let us now consider the spatial rotations:

$$\begin{aligned} M = UH &\mapsto MR = URR^\dagger HR \\ \Rightarrow U &\mapsto UR \\ H &\mapsto R^\dagger HR \in \mathcal{H}. \end{aligned} \tag{4.32}$$

The latter in terms of coordinates is

$$H = \sqrt{1+\lambda^2}\,\mathbb{I} + \boldsymbol{\lambda}\cdot\boldsymbol{\tau} \mapsto \sqrt{1+\lambda^2}\mathbb{I} + R^\dagger(\boldsymbol{\lambda}\cdot\boldsymbol{\tau})R. \tag{4.33}$$

The action of conjugation of the Hermitian, traceless matrix $\boldsymbol{\lambda}\cdot\boldsymbol{\tau}$ by a unitary matrix R is well known [40] – it is equivalent to a $SO(3)$ rotation of $\boldsymbol{\lambda}$:

$$R^\dagger(\boldsymbol{\lambda}\cdot\boldsymbol{\tau})R = (\mathcal{R}\boldsymbol{\lambda})\cdot\boldsymbol{\tau} \tag{4.34}$$

where $\mathcal{R} \in SO(3)$ with components $\mathcal{R}_{ab} = \frac{1}{2}\mathrm{tr}(\tau_a R^\dagger \tau_b R)$. The action on the left-invariant 1-forms $\boldsymbol{\sigma}$ is similar. Under $U \mapsto UR$,

$$\begin{aligned} U^{-1}dU &\mapsto R^\dagger(U^{-1}dU)R \\ \Rightarrow \frac{i}{2}\boldsymbol{\sigma}\cdot\boldsymbol{\tau} &\mapsto \frac{i}{2}R^\dagger(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})R = \frac{i}{2}(\mathcal{R}\boldsymbol{\sigma}\cdot\boldsymbol{\tau}) \\ \Rightarrow \boldsymbol{\sigma} &\mapsto \mathcal{R}\boldsymbol{\sigma} \end{aligned} \tag{4.35}$$

where $\mathcal{R}$ is the same $SO(3)$ matrix defined above. Thus both $\boldsymbol{\lambda}$ and $\boldsymbol{\sigma}$ transform as 3-vectors under spatial rotations and as scalars under isorotations.

The metric must be invariant under spatial rotations also, so the task is to construct from $\boldsymbol{\lambda}$, $d\boldsymbol{\lambda}$ and $\boldsymbol{\sigma}$ the most general possible $(0,2)$ tensor which is scalar under these rotations. This is

$$\begin{aligned} g \;=\; & A\,d\boldsymbol{\lambda}\cdot d\boldsymbol{\lambda} + B(\boldsymbol{\lambda}\cdot d\boldsymbol{\lambda})^2 + C\,\boldsymbol{\sigma}\cdot\boldsymbol{\sigma} + D(\boldsymbol{\lambda}\cdot\boldsymbol{\sigma})^2 \\ & + E\,\boldsymbol{\sigma}\cdot d\boldsymbol{\lambda} + F(\boldsymbol{\lambda}\cdot d\boldsymbol{\lambda})(\boldsymbol{\lambda}\cdot\boldsymbol{\sigma}) + G(\boldsymbol{\lambda}\times\boldsymbol{\sigma})\cdot d\boldsymbol{\lambda} \end{aligned} \tag{4.36}$$

A—G being 7 unknown functions of $\lambda = |\boldsymbol{\lambda}|$ only.

The metric may be restricted still further on consideration of a discrete isometry. The kinetic energy is invariant under the discrete "parity" transformations $P_z : z \mapsto \bar{z}$ and $P_W : W \mapsto \bar{W}$. However, neither is an isometry of the moduli space because

each reverses the sign of the topological charge, mapping lumps to anti-lumps. The composite transformation $P_W \circ P_z$ *is* an isometry. Using the configuration of (4.15),

$$W(z) \overset{P_z}{\longmapsto} \frac{a\bar{z}+b}{c\bar{z}+d} \overset{P_W}{\longmapsto} \frac{\bar{a}z+\bar{b}}{\bar{c}z+\bar{d}} = \bar{M} \odot z. \tag{4.37}$$

In terms of the moduli, $M \mapsto \bar{M}$ is the transformation,

$$\begin{aligned} \boldsymbol{\sigma} &= (\sigma_1, \sigma_2, \sigma_3) \mapsto (-\sigma_1, \sigma_2, -\sigma_3) \\ \boldsymbol{\lambda} &= (\lambda_1, \lambda_2, \lambda_3) \mapsto (\lambda_1, -\lambda_2, \lambda_3). \end{aligned} \tag{4.38}$$

This isometry removes two of the terms in (4.36) because under it,

$$\boldsymbol{\sigma} \cdot d\boldsymbol{\lambda} \mapsto -\boldsymbol{\sigma} \cdot d\boldsymbol{\lambda}, \tag{4.39}$$

and

$$(\boldsymbol{\lambda} \cdot d\boldsymbol{\lambda})(\boldsymbol{\lambda} \cdot \boldsymbol{\sigma}) \mapsto -(\boldsymbol{\lambda} \cdot d\boldsymbol{\lambda})(\boldsymbol{\lambda} \cdot \boldsymbol{\sigma}) \tag{4.40}$$

so that $E(\lambda) \equiv F(\lambda) \equiv 0$.

The remaining five functions of λ are evaluated by choosing convenient orientations for $\boldsymbol{\lambda}$, positions on $S^3/\mathbb{Z}_2$ and tangent vectors (velocities), then calculating the kinetic energy and comparing with (4.36). Repeating this four times it is possible to extract the following (see figure 4.1):

$$\begin{aligned} A &= 4\pi S_2(\chi) \\ B &= \frac{4\pi}{\lambda^2}\left[\frac{4}{1+\lambda^2}S_1(\chi) - S_2(\chi)\right] \\ C &= \frac{\pi}{2} - 2\pi S_1(\chi) \\ D &= \frac{\pi}{\lambda^2}\left[6S_1(\chi) - \frac{1}{2}\right] \\ G &\equiv A \end{aligned} \tag{4.41}$$

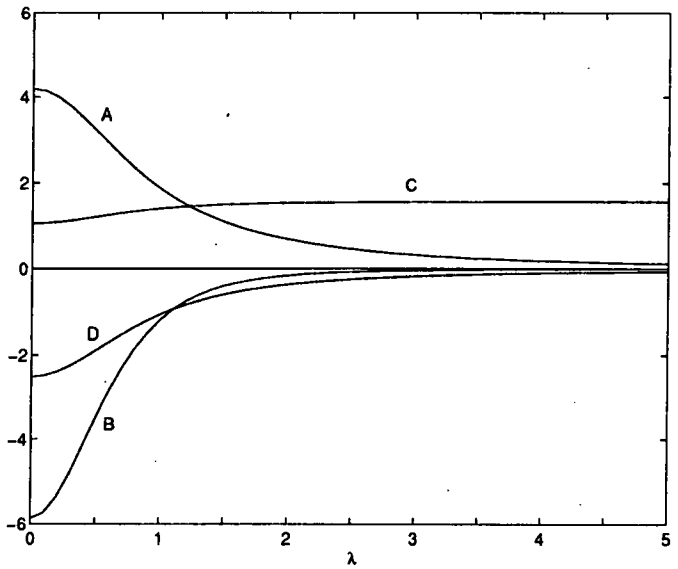

Figure 4.1: The metric functions A, B, C and D.

where,

$$\begin{aligned} \chi(\lambda) &= \frac{\sqrt{1+\lambda^2}+\lambda}{\sqrt{1+\lambda^2}-\lambda} \\ S_1(\chi) &= \frac{\chi^2}{2(\chi^2-1)^3}\left[(\chi^2+1)\log\chi^2 - 2\chi^2 + 2\right] \\ S_2(\chi) &= \frac{\chi}{(\chi^2-1)^3}\left[\chi^4 - 2\chi^2\log\chi^2 - 1\right]. \end{aligned} \tag{4.42}$$

Note that χ is a strictly increasing function of λ, and that $\chi : [0,\infty) \to [1,\infty)$. There appear to be divergences of the functions A—D at $\lambda = 0$, but these are in fact removable singularities, so all the limits of vanishing λ exist. Although B and D are negative it is straightforward to show that this metric is positive definite, as of course it must be. In summary then, the metric is

$$g = A\, d\boldsymbol{\lambda}\cdot d\boldsymbol{\lambda} + B(\boldsymbol{\lambda}\cdot d\boldsymbol{\lambda})^2 + C\,\boldsymbol{\sigma}\cdot\boldsymbol{\sigma} + D(\boldsymbol{\lambda}\cdot\boldsymbol{\sigma})^2 + A(\boldsymbol{\lambda}\times\boldsymbol{\sigma})\cdot d\boldsymbol{\lambda}. \tag{4.43}$$

4.5 Some totally geodesic submanifolds

Before discussing geodesics of the metric (4.43) we must describe the connexion between a point on the moduli space $(S^3/\mathbb{Z}_2)\times\mathbb{R}^3$ and its corresponding field configuration. Consider first the 3-dimensional submanifold defined by $U = \mathbb{I}$, parametrized by $\boldsymbol{\lambda} \in \mathbb{R}^3$. Any point in this subspace may be written as $\boldsymbol{\lambda} = \mathcal{R}\boldsymbol{\lambda}'$ where $\boldsymbol{\lambda}' = (0,0,\lambda)$, $\lambda \geq 0$ and $\mathcal{R} \in SO(3)$. The lump represented by $\boldsymbol{\lambda}'$ is

$$W'(z) = (\boldsymbol{\lambda}'\cdot\boldsymbol{\tau})\odot z = \left(\frac{\sqrt{1+\lambda^2}+\lambda}{\sqrt{1+\lambda^2}-\lambda}\right) z = \chi(\lambda)z. \tag{4.44}$$

This is a distorted hedgehog with the arrows pulled towards the North pole. The larger λ is, so the larger χ is and the greater is the distortion. Although it is usual to define the position of a lump as the position of maximum energy density, we shall refer to this as a lump of sharpness λ located at $\hat{\boldsymbol{\lambda}}' = (0,0,1)$, the antipodal point to the energy density peak which occurs where the arrows are stretched apart. Obviously the motion of any point is trivially mirrored by its antipodal image, so this terminology makes sense. The lump represented by $\boldsymbol{\lambda}$ is

$$\begin{aligned} W(z) &= (\mathcal{R}\boldsymbol{\lambda}'\cdot\boldsymbol{\tau})\odot z = [R^\dagger(\boldsymbol{\lambda}'\cdot\boldsymbol{\tau})R]\odot z \\ &= [R^\dagger(\boldsymbol{\lambda}'\cdot\boldsymbol{\tau})]\odot(R\odot z) = R^\dagger\odot[W'(R\odot z)]. \end{aligned} \tag{4.45}$$

This configuration is formed by first performing a spatial rotation – taking the $\boldsymbol{\phi}$ arrow at the old point z and placing it at the new point $R\odot z$ without changing its orientation – then performing the inverse isorotation. The result looks like the arrows have been fixed to $S^2_{\rm sp}$ which has then been rotated by $\mathcal{R}$ which, as defined, has the action on S^2 equivalent to $R^\dagger$, *not* R, acting on $\mathbb{C}P^1$ via $\odot$. That is, if we define P to be stereographic projection, P : $S^2 \to \mathbb{C}P^1$ so that P : $\boldsymbol{\phi} \mapsto W$, then P : $\mathcal{R}\boldsymbol{\phi} \mapsto R^\dagger\odot W$. So the lump at the North pole is shifted to $\hat{\boldsymbol{\lambda}} = \mathcal{R}\hat{\boldsymbol{\lambda}}'$.

All other points on the moduli space are on the isorotation orbit of this subman-

ifold, and isorotation, while changing the internal orientation of the lump, does not move the lump around on physical space. Thus we can always interpret $\hat{\boldsymbol{\lambda}}$ as the lump's position, and λ as parametrizing its sharpness. The symmetric hedgehog has $\lambda = 0$, and large λ lumps have taller, narrower energy density peaks than small λ lumps.

One way of attacking the geodesic problem is to reduce its dimension by identifying totally geodesic submanifolds, that is, choosing initial value problems whose solution is simplified by some symmetry. The easiest method for identifying such submanifolds is to find fixed point sets of discrete groups of isometries. Any isometry maps geodesics to geodesics, so if there were a geodesic starting off in the fixed point set of the isometry and subsequently deviating from it, this would be mapped under the isometry to another geodesic, identical to the first throughout its length in the fixed point set, but deviating from the set in a different direction. This violates the uniqueness of solutions of ordinary differential equations, so no such geodesic may exist. If the initial data are a point on the fixed point set and a velocity tangential to it, then the geodesic must remain on the fixed point set for all subsequent time.

Examining (4.43) we see that $\boldsymbol{\lambda} \to -\boldsymbol{\lambda}$ is an isometry. Its fixed point set is $S^3/\mathbb{Z}_2$, the isorotation orbit of the symmetric hedgehog, on which the metric is

$$g = \frac{\pi}{6}\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}. \tag{4.46}$$

The kinetic energy is the rotational energy of a totally symmetric rigid body, moment of inertia $\pi/6$. The solutions are just isorotations of the symmetric hedgehog at constant frequency about some fixed axis. In this case isorotation is equivalent to spatial rotation because $\boldsymbol{\lambda} = 0 \Rightarrow H = \mathbb{I}$.

A less trivial geodesic submanifold is the fixed point set of the parity transformation described above, $M \mapsto \bar{M}$. This is a 3-dimensional manifold, the product of the plane $\lambda_2 = 0$ in $\mathbb{R}^3$ with the circle $\{\alpha \in [0,\pi], \beta = \gamma = 0\} \cup \{\alpha \in [0,\pi], \beta = \gamma = \pi\}$ in $S^3/\mathbb{Z}_2$. The circle is more conveniently parametrized if we temporarily allow α the

domain $[0, 2\pi]$, for it is then $\{\alpha \in [0, 2\pi], \beta = \gamma = 0\}$. This space contains lumps of arbitrary sharpness located on a great circle through the poles of $S^2_{\rm sp}$, each lump having an internal phase, so certain of its geodesics may be candidates for "travelling lumps." Introducing spherical polar coordinates for $\boldsymbol{\lambda}$,

$$\boldsymbol{\lambda} = \lambda(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{4.47}$$

the plane $\lambda_2 = 0$ is parametrized by (λ, θ) where $\theta \in [0, 2\pi]$, again gluing two semi-circles together and extending the domain of θ to cover the whole circle in one go. The metric on this geodesic submanifold is

$$g = (A + \lambda^2 B)d\lambda^2 + \lambda^2 A\, d\theta^2 + C\, d\alpha^2 - \lambda^2 A\, d\theta\, d\alpha. \tag{4.48}$$

So the kinetic energy is

$$T = \frac{1}{2}\left[(A + \lambda^2 B)\dot{\lambda}^2 + \frac{p_\theta^2 C}{\lambda^2 A(C - \lambda^2 A/4)} + \frac{p_\alpha(p_\alpha + p_\theta)}{C - \lambda^2 A/4}\right] \tag{4.49}$$

where we have used the cyclicity of θ and α to eliminate $\dot{\theta}$ and $\dot{\alpha}$ in favour of their constant, canonically conjugate momenta,

$$\begin{aligned} p_\theta &= \lambda^2 A\left(\dot{\theta} - \frac{1}{2}\dot{\alpha}\right) \\ p_\alpha &= C\dot{\alpha} - \frac{1}{2}\lambda^2 A\dot{\theta}. \end{aligned} \tag{4.50}$$

Note that constant p_θ (p_α) does *not* imply constant $\dot{\theta}$ ($\dot{\alpha}$), nor does $p_\theta = 0$ ($p_\alpha = 0$) imply $\dot{\theta} = 0$ ($\dot{\alpha} = 0$).

This system can be visualized as a point particle of position dependent mass $(A + \lambda^2 B)$ moving in a potential. It is the form of the potential which determines the

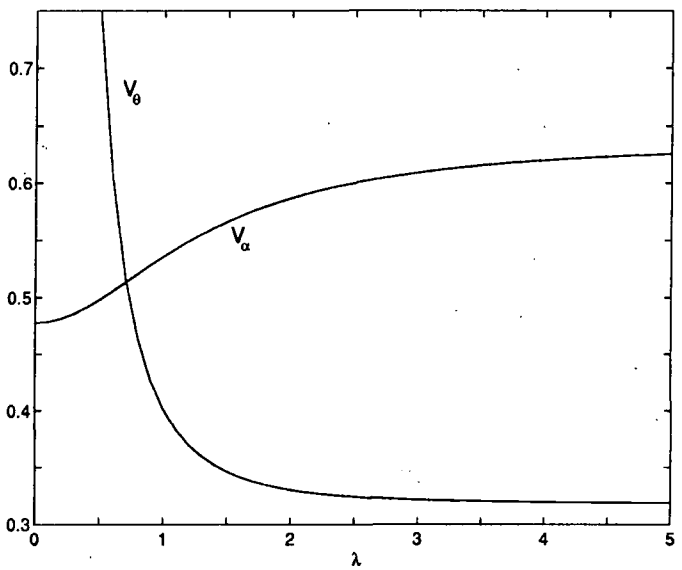

Figure 4.2: The potential functions $\mathcal{V}_\theta(\lambda)$ and $\mathcal{V}_\alpha(\lambda)$.

broad qualitative features of its behaviour:

$$\mathcal{V}(\lambda) = p_\theta^2 \mathcal{V}_\theta(\lambda) + p_\alpha(p_\alpha + p_\theta)\mathcal{V}_\alpha(\lambda). \tag{4.51}$$

As can be seen from figure 4.2 while $\mathcal{V}_\theta(\lambda)$ is monotonically decreasing, $\mathcal{V}_\alpha(\lambda)$ is monotonically increasing. This allows the possibility of potential minima where the forces $-p_\theta^2 \mathcal{V}'_\theta(\lambda)$ outwards (in the sense of increasing λ) and $p_\alpha(p_\alpha + p_\theta)\mathcal{V}'_\alpha(\lambda)$ inwards are in stable equilibrium. It certainly is *not* possible if $p_\alpha(p_\alpha + p_\theta) \leq 0$, for then $\mathcal{V}(\lambda)$ as a whole is monotonically decreasing. This region in the (p_α, p_θ) plane is shown shaded in figure 4.3. Whatever the initial conditions on λ, the lump always moves towards infinite λ without passing through $\lambda = 0$ (which would correspond to the lump swapping hemispheres), reaching the singularity $\lambda = \infty$, an infinitely tall, sharp spike, in finite time. Thus (M_1, g) is geodesically incomplete. This result follows from the rapid vanishing of the inertia to sharpening, $A + \lambda^2 B$, in the large λ limit (see figure 4.4). For example, consider the simple case $p_\alpha = p_\theta = 0$ and let $\lambda(0)$ and $\dot{\lambda}(0)$ be strictly positive. It is easily seen that t_∞, the time taken to reach the

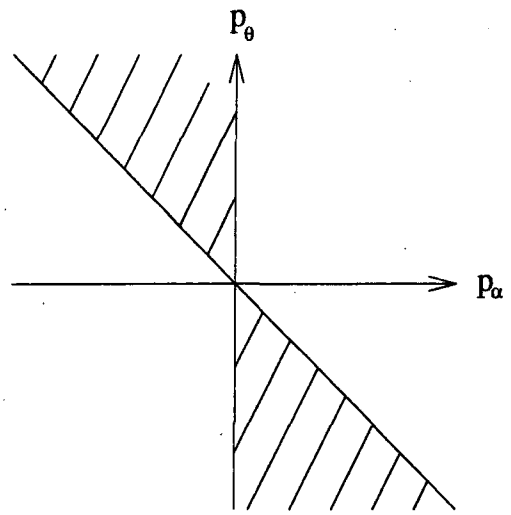

Figure 4.3: The (p_α, p_θ) plane. The instability region is shaded and includes boundaries.

singular spike is proportional to the following integral:

$$t_\infty \propto \int_{\lambda(0)}^{\infty} d\lambda \sqrt{A(\lambda) + \lambda^2 B(\lambda)}. \tag{4.52}$$

The integrand is finite over the integration range (even if $\lambda(0) = 0$), so if t_∞ diverges it can only be due to the large λ behaviour. But the integrand vanishes like $(\log \lambda)/\lambda^2$ at large λ, fast enough to ensure convergence. The inclusion of repulsive potentials can only make matters worse, so this singular behaviour extends to the rest of the shaded area.

In the unshaded region, one can define the positive constant $\kappa = p_\theta^2/(p_\alpha^2 + p_\alpha p_\theta)$ such that

$$\mathcal{V}(\lambda) = p_\alpha(p_\alpha + p_\theta)\left(\kappa \mathcal{V}_\theta(\lambda) + \mathcal{V}_\alpha(\lambda)\right). \tag{4.53}$$

Then the forms of the functions $-\mathcal{V}_\theta'$ and $\mathcal{V}_\alpha'$ (see figure 4.5) suggest that for each λ, there is one (and only one) value of κ (call it $\tilde{\kappa}$) for which $\mathcal{V}$ has a minimum at λ.

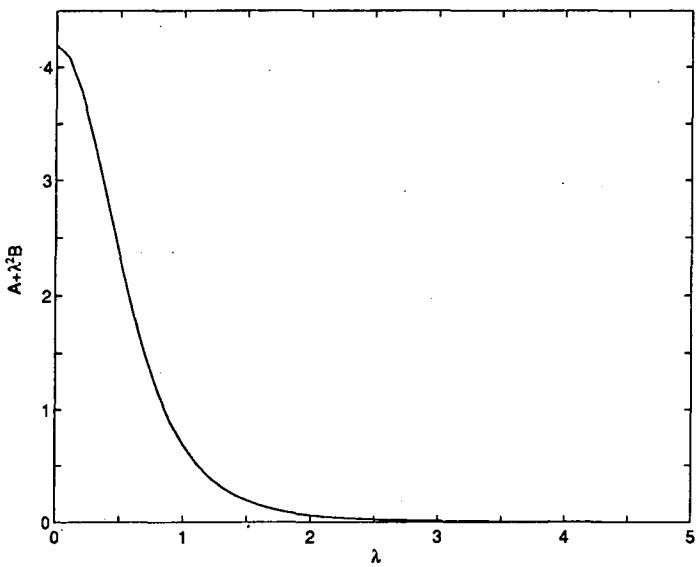

Figure 4.4: The inertia to sharpening, $A + \lambda^2 B$.

The equilibrium condition is $\mathcal{V}'(\lambda) = 0$, so

$$\tilde{\kappa}(\lambda) = -\frac{\mathcal{V}'_\alpha(\lambda)}{\mathcal{V}'_\theta(\lambda)}. \tag{4.54}$$

Inverting the definition of κ we find that there are two distinct values of p_α/p_θ for each $\tilde{\kappa}$. If p_α/p_θ takes one of these and $\dot{\lambda}(0) = 0$ then λ will not subsequently change and hence $\dot{\alpha}$ and $\dot{\theta}$ will also remain constant, allowing the lump to travel around a great circle on S^2_{sp} with constant speed and shape while undergoing constant frequency isorotation. The two values are

$$\nu_\pm(\lambda) \equiv -\frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{1}{\tilde{\kappa}(\lambda)}}. \tag{4.55}$$

Substituting (4.50) we can find the corresponding pair of stable ratios $\dot{\alpha}/\dot{\theta}$ as functions of λ,

$$\omega_\pm \equiv \frac{\lambda^2 A(\lambda)\,(2\nu_\pm(\lambda) + 1)}{2C(\lambda) + \lambda^2 A(\lambda)\nu_\pm(\lambda)}, \tag{4.56}$$

(see figure 4.6). Thus, for any lump sharpness λ and travel speed $\dot{\theta}$ there are two

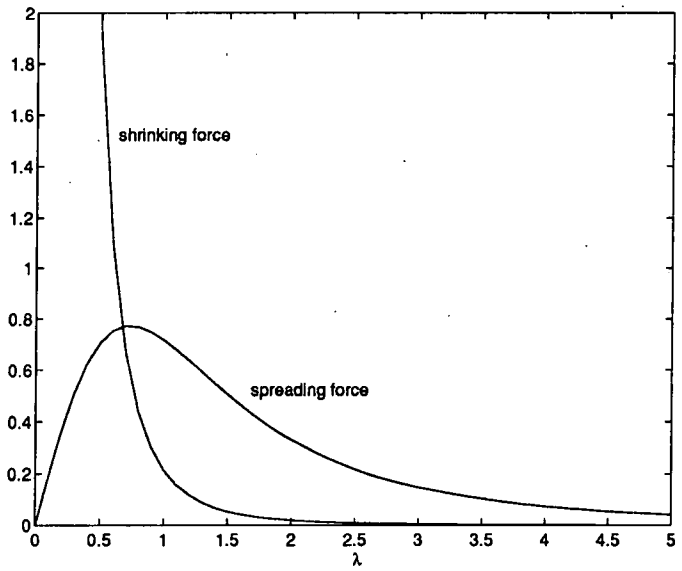

Figure 4.5: The outward (shrinking) force, $-\mathcal{V}'_\theta(\lambda)$ and the inward (spreading) force, $\mathcal{V}'_\alpha(\lambda)$. The vertical scales of the two curves are different.

possible isorotation frequencies $\dot{\alpha}$ which allow stable, uniform travel and these two stability "branches" never coincide. It is interesting to note that $\lim_{\lambda\to\infty}\omega_+(\lambda) = 1$ meaning that very tall, sharp lumps can travel uniformly with $\dot{\alpha} \approx \dot{\theta}$. Motion with constant λ and $\dot{\alpha} = \dot{\theta}$ is simply constant speed spatial rotation carrying the lump around a great circle. So when the extent of the lump's structure is negligible relative to the radius of curvature of $S^2_{\rm sp}$, it can travel in analogous fashion to a flat-space $\mathbb{C}P^1$ lump [32].

Since $\tilde{\kappa}(\lambda)$ takes all positive values, whatever value κ takes there is an equilibrium λ. If $\lambda(0)$ is near this value, then (assuming $|\dot{\lambda}(0)|$ is not too large) the shape of the lump will oscillate periodically about the preferred sharpness, and its speed of travel round the sphere will vary with the same period. If $|\dot{\lambda}(0)|$ *is* too large, or the lump is initially much too spread out for its κ, then it will escape to the singular spike in finite time.

Let us examine the concrete example $p_\theta = p_\alpha$. Figure 4.7 shows the potential $\mathcal{V}$ with its minimum and the lump travel speed $\dot{\theta}$ as functions of λ. We imagine a

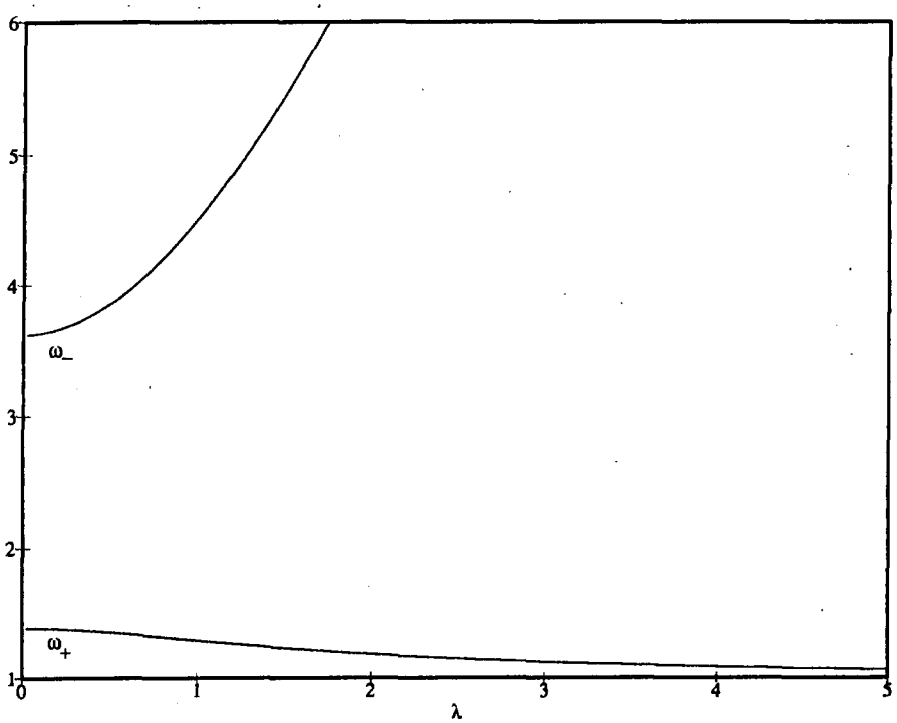

Figure 4.6: The stable frequency ratios, $\dot{\alpha}/\dot{\theta} = \omega_{\pm}(\lambda)$. The upper curve $\omega_{-}(\lambda)$ tends to infinity at large λ.

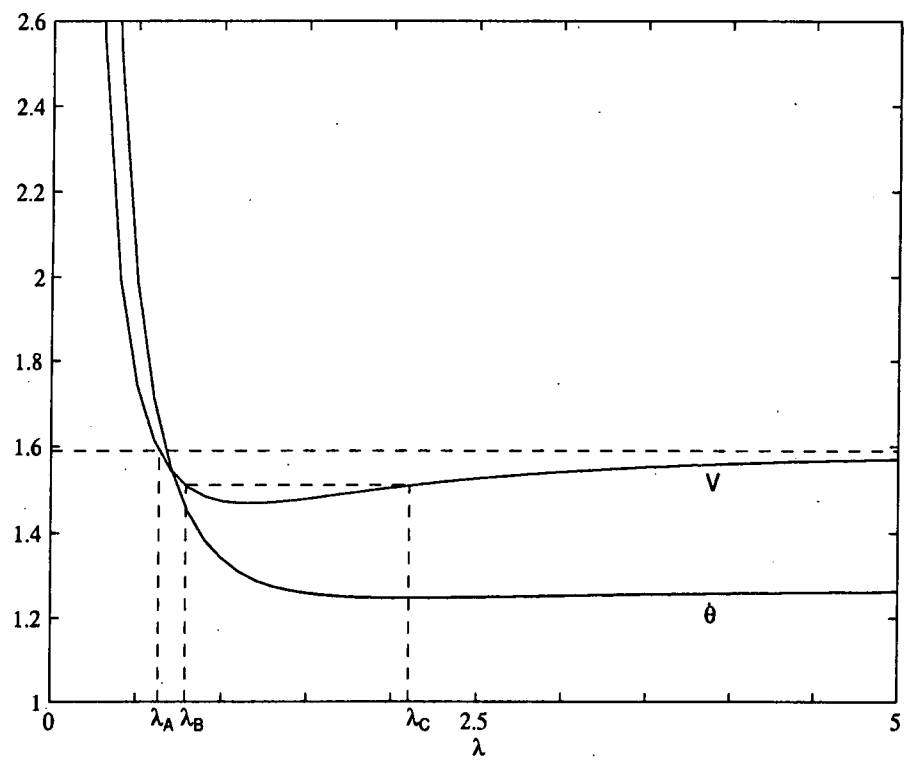

Figure 4.7: The $p_\theta = p_\alpha$ case: potential $\mathcal{V}$ and speed $\dot{\theta}$.

particle of position dependent mass moving in this potential and for simplicity take $\dot{\lambda}(0) = 0$. Clearly, if we release the particle with $\lambda(0) < \lambda_A \approx 0.626$, it will move off to infinity, the last case mentioned above. But if $\lambda(0) > \lambda_A$, oscillatory motion ensues. Even here there are two qualitatively different cases, because $\dot{\theta}(\lambda)$ has an absolute minimum at $\lambda_C \approx 2.096$, a turning point which is only reached if $\lambda(0) < \lambda_B \approx 0.789$ (or $\lambda(0) > \lambda_C$). If $\lambda_B < \lambda(0) < \lambda_C$ then the speed of travel oscillates in simple phase with the lump sharpness, going from fast, spread-out lump to slow, sharp lump and back again. But if $\lambda_A < \lambda(0) < \lambda_B$ or $\lambda(0) > \lambda_C$ the speed undergoes an extra wobble during the middle of the sharpness cycle, speeding up then slowing down again as it passes through its maximum sharpness. This case corresponds to lumps whose shape oscillates more acutely.

Other interesting geodesic submanifolds are generated by computing the fixed point sets Σ_{ab} of the isometries $M \mapsto (i\tau_a)^\dagger M(i\tau_b)$, simultaneous isorotation and spatial rotation by π about the a and b axes respectively:

$$\begin{aligned} U &\mapsto (i\tau_a)^\dagger U(i\tau_b), \\ H &\mapsto (i\tau_b)^\dagger H(i\tau_b) \\ \Rightarrow \lambda_c &\mapsto \begin{cases} \lambda_c & c = b \\ -\lambda_c & c \neq b. \end{cases} \end{aligned} \tag{4.57}$$

Thus if $q \in \Sigma_{ab}$ $\boldsymbol{\lambda}$ must point along the b-axis. On Σ_{bb}, $U = \exp(i\psi\tau_b/2)$ where $\psi \in [0, 2\pi]$, whereas if $a = b \pm 1 \bmod 3$ then $U = \exp(\pm i\pi\tau_c/4)\exp(i\psi\tau_b/2)$ where $c = b \mp 1 \bmod 3$. It follows that $\boldsymbol{\sigma}$ also points along the b-axis, independent of a. The $a \neq b$ submanifolds are the images of Σ_{bb} under $\pm\pi/2$ isorotations about the three axes, so it suffices to solve the geodesic problem on Σ_{bb} - geodesics on $\Sigma_{ab}, a \neq b$, are then obtained by acting with the appropriate isometry. The choice of b doesn't matter, and we choose to study the cylinder $\Sigma_{33} \cong S^1 \times \mathbb{R}$ consisting of lumps of every sharpness located at the North (South) pole if $\lambda_3 > 0$ ($\lambda_3 < 0$), arbitrarily rotated about the North-South axis. Note that $[U, H] = 0$ on Σ_{33} so "isorotated" and

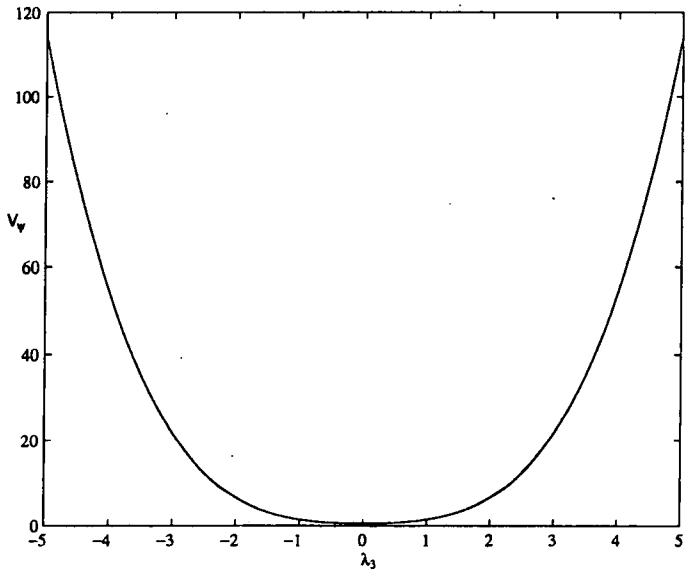

Figure 4.8: **Potential,** $\mathcal{V}_\psi(\lambda_3)$.

"spatially rotated" mean the same thing in this case.

The kinetic energy on Σ_{33} is

$$T = \frac{1}{2}\left[(A + \lambda_3^2 B)\dot{\lambda}_3^2 + \frac{p_\psi^2}{C + \lambda_3^2 D}\right] \tag{4.58}$$

where once again p_ψ is the momentum conjugate to ψ,

$$p_\psi = (C + \lambda_3^2 D)\dot{\psi}, \tag{4.59}$$

and is constant by virtue of the cyclicity of ψ. This looks like a particle in one dimension moving in a potential

$$p_\psi^2 \mathcal{V}_\psi = \frac{p_\psi^2}{2(C + \lambda_3^2 D)} = \frac{1}{2} p_\psi \dot{\psi} \tag{4.60}$$

with postion dependent mass. From the potential (figure 4.8) we see that all motion is oscillatory and that λ_3 periodically changes sign. Thus a lump set spinning about its own axis will spread out, its rotation slowing, until it is uniformly spread over

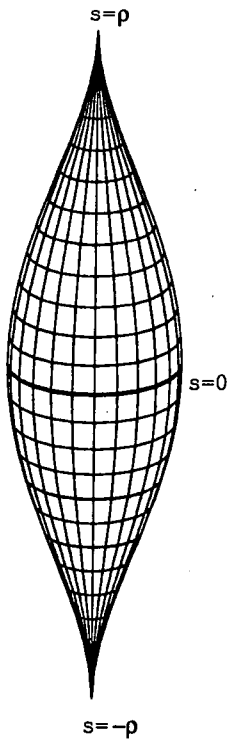

Figure 4.9: **A sketch of the geodesic submanifold Σ_{33} embedded as a surface of revolution in $\mathbb{R}^3$.**

the sphere, whereupon it will shrink to its mirror image in the opposite hemisphere, regaining its original spin speed as it does so. The process then reverses and the lump "bounces" between antipodal points indefinitely.

Defining the new coordinate

$$s(\lambda_3) = \int_0^{\lambda_3} d\mu \sqrt{A(\mu) + \mu^2 B(\mu)}, \tag{4.61}$$

which takes values in a finite open interval $(-\rho, \rho)$ symmetric about $s = 0$, the metric on Σ_{33} becomes

$$g = ds^2 + H^2(s) d\psi^2 \tag{4.62}$$

where $H^2\left(s(\lambda_3)\right) = C(\lambda_3) + \lambda_3^2 D(\lambda_3)$. Since $|dH/ds| < 1 \ \forall s$, the manifold may be embedded as a surface of revolution in $\mathbb{R}^3$ and geodesics on it can be visualized

directly. Figure 4.9 is a sketch of the embedded surface, which is of finite length and sausage-shaped with its ends pinched to infinitely sharp spikes, the tips of which are the points $\lambda_3 = \pm\infty$ and so are missing. The coordinates (s, ψ) are geodesic orthogonal coordinates: a curve of constant ψ is a geodesic along the length of the cylinder, lying in a plane containing the cylinder's axis, parametrized by arc length s, while a curve of constant s is a circle of radius $H(s)$, lying in a plane orthogonal to the cylinder's axis. Two such curves always intersect at right angles.

The spinning geodesics described above wind around the cylinder, never reaching the ends (this would violate conservation of "angular momentum") but winding back and forth between two circles $s = \pm\tilde{\rho}$, $\tilde{\rho} < \rho$ which they touch tangentially. The angle φ at which the geodesic intersects the circle $s = 0$ determines $\tilde{\rho}$. When $\varphi = 0$ the geodesic stays on the circle, $\tilde{\rho} = 0$ (a spinning symmetric hedgehog) and $\tilde{\rho}(\varphi)$ monotonically increases, tending to the supremum ρ as φ tends to $\pi/2$ ($\varphi = \pi/2$ is an irrotational geodesic between antipodal singular spikes). Note that the geodesic incompleteness already mentioned appears again, this time characterized by the finite length of the cylinder and the missing points $s = \pm\rho$ ($\lambda_3 = \pm\infty$).

4.6 Concluding remarks

The behaviour of isolated topological solitons in flat space is generally rather trivial, whereas, as we have seen, despite the homogeneity of S^2, the motion of a single lump on the sphere is surprisingly complicated. It does travel on great circles, but while doing so its shape may oscillate in phase with its speed, whose periodic variation is of one of two types depending on the violence of the shape oscillations, or it may collapse to an infinitely tall, thin spike in finite time. A lump sent spinning about its own axis spreads out then re-forms in the opposite hemisphere, endlessly commuting between antipodal points.

The infinities in the unit-charge metric in flat space can be attributed to the

lumps' polynomial tail-off: the kinetic energy needed to rigidly spin or scale-deform a lump diverges because such motions involve changing the field at spatial infinity. The $\mathbb{C}P^1$ model on any compact space should be free of this problem because the kinetic energy, being an integral over a space of finite volume, must be finite provided the kinetic energy density is non-singular. Conversely, one would expect the singularity to persist in the model defined on hyperbolic space.

The flat-space $\mathbb{C}P^1$ model can be made more "physical" by adding a $(2+1)$-dimensional version of the Skyrme term to stabilize against lump collapse, and a potential to stabilize against spread. The Bogomol'nyi bound remains valid but unsaturable. The potential is somewhat arbitrary, but one interesting possibility [41] gives a mass to small amplitude travelling waves of the ϕ field, termed pions in analogy with the Skyrme model, and gives the lump an exponential rather than polynomial tail. This allows the lump to rotate, a problem if one attempts a collective coordinate approximation to low-energy dynamics along the lines recently proposed in [42, 43]. The idea is to restrict the field to the "Bogomo'nyi regime" moduli space (in this case the space of static $\mathbb{C}P^1$ solutions), introducing a potential and a perturbed metric (in [42] but not [43]) to account for the new interactions, which are assumed to be weak. There seems little hope of perturbing the singular flat-space metric such that rotations become possible, but the problem does not arise on the sphere.

The geodesic approximation could be used to investigate the interaction of two lumps moving on S^2. Right angle scattering in head on collisions emerges naturally from the geodesic approximation of many flat-space models as a consequence of the classical indistinguishability of topological solitons. It would be interesting to see if there is some analogous behaviour on the sphere. However, evaluating the two-lump metric could be difficult since the action of the isometry group on the charge-2 moduli space is far less accessible than in the present case. Even in flat space [33], the scattering problem is sufficiently complicated to require considerable numerical effort.

Chapter 5

Intervortex forces

5.1 Introduction

The abelian Higgs model is a field theory in $(2+1)$ dimensional Minkowski space consisting of a complex-valued scalar field ϕ, coupled to a $U(1)$ gauge field A_μ. The scalar field is given a Higgs symmetry breaking self-interaction which allows topologically stable solitons, called vortices, to exist. This symmetry breaking property is of great interest to cosmologists, who study gauge symmetries (Grand Unified or electroweak) which are broken as the Universe undergoes a phase transition during expansion and cooling. The abelian Higgs model in $(3+1)$ dimensions provides the simplest example of a broken symmetry yielding cosmic strings, which may explain the inhomogeneity of large-scale cosmological matter distribution. The study below concerns vortices in the $(2+1)$ dimensional theory, which may be thought of as a cross-sectional slice through parallel, straight strings in the larger space. The model also has applications in superconductivity theory. We, however, shall treat vortices purely as solitons in their own right, independently of any applications.

Let us review the salient features of the abelian Higgs model [4, 44, 45]. We adopt the standard conventions for Minkowski space. Spacetime coordinates are written x^μ (collectively x), where μ (as any other Greek index) runs over $0, 1, 2$; time is the

zeroth component, $x^0 = t$, and we will often write the spatial pair (x^1, x^2) as a 2-vector $\mathbf{x}$; $(\partial_\mu) = (\partial/\partial x^\mu) = (\partial/\partial t, \nabla)$. Indices are raised and lowered by contraction with the Minkowski metric, $(\eta_{\mu\nu}) = \mathrm{diag}(1, -1, -1)$. The summation convention applies throughout. The Lagrangian density is

$$\mathcal{L} = \frac{1}{2} D_\mu \phi \overline{D^\mu \phi} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{\mu^2}{8} (|\phi|^2 - 1)^2, \tag{5.1}$$

where $D_\mu \phi$ is the gauge covariant derivative,

$$D_\mu \phi = (\partial_\mu + i A_\mu) \phi, \tag{5.2}$$

and $F_{\mu\nu}$ is the field strength tensor,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{5.3}$$

Note that that the electric charge and the vacuum magnitude of the Higgs field have been normalized to unity, leaving one parameter μ, the Higgs mass. The Lagrangian density is symmetric under $U(1)$ gauge transformations as follows: let χ be any smooth, real function on Minkowski space, $\chi : \mathbb{R}^{2+1} \to \mathbb{R}$; then the transformation

$$\begin{aligned} \phi &\mapsto e^{i\chi} \phi \\ A_\mu &\mapsto A_\mu - \partial_\mu \chi \end{aligned} \tag{5.4}$$

locally redefines the phase of ϕ, and maps

$$\begin{aligned} D_\mu \phi &\mapsto e^{i\chi} D_\mu \phi \\ F_{\mu\nu} &\mapsto F_{\mu\nu} \\ |\phi| &\mapsto |\phi|, \end{aligned} \tag{5.5}$$

so that $\mathcal{L} \mapsto \mathcal{L}$. This symmetry group is vast – the space of all smooth maps from

Minkowski space to the circle, $\mathbb{R}^{2+1} \to S^1$.

The Euler-Lagrange equations derived from $\mathcal{L}$ are a set of coupled, nonlinear, hyperbolic partial differential equations, of which no nontrivial solutions are known. To make progress we must dismantle the relativistic framework and concentrate on E, the energy rather than action functional. The kinetic and potential energy functionals are, respectively,

$$T = \frac{1}{2}\int d^2\mathbf{x}\left(|\dot{\phi}|^2 + |\dot{\mathbf{A}}|^2\right) \tag{5.6}$$

$$V = \frac{1}{2}\int d^2\mathbf{x}\left[\mathbf{D}\phi\cdot(\mathbf{D}\phi)^* + F_{12}^2 + \frac{1}{4}\mu^2(|\phi|^2-1)^2\right] \tag{5.7}$$

where gauge freedom has been used to set $A_0 \equiv 0$, leaving only time independent transformations available ($\mathbf{D}\phi$ denotes the 2-vector part of the *covector* $D_\mu\phi$, so $\mathbf{D}\phi = \nabla\phi - i\mathbf{A}\phi$; note the minus sign). The variational problem with action $\int d^3x(T-V)$ is then the same as that with action $\int d^3x\mathcal{L}$, provided we also impose the Euler-Lagrange equation associated with A_0 as a constraint:

$$\nabla\cdot\dot{\mathbf{A}} + \frac{1}{2}(\dot{\phi}\phi^* - \dot{\phi}^*\phi) = 0. \tag{5.8}$$

For a configuration $(\phi, \mathbf{A})$ to have finite total energy, the integrand of (5.7) must vanish at large $r = |\mathbf{x}|$. This imposes the following boundary conditions at spatial infinity:

$$|\phi| \to 1 \tag{5.9}$$

$$\mathbf{D}\phi \to \mathbf{0} \Rightarrow \mathbf{A} \to \frac{\nabla\phi}{i\phi}. \tag{5.10}$$

Introducing angular coordinate θ, we can rewrite (5.9),

$$\phi_\infty(\theta) := \lim_{r\to\infty}\phi(r,\theta) = e^{i\gamma(\theta)}, \tag{5.11}$$

that is, $\phi_\infty : S^1 \to S^1$. Single valuedness of ϕ demands that $\gamma(0) = \gamma(2\pi)+2n\pi, n \in \mathbb{Z}$, and the winding number n is invariant under any continuous deformation of ϕ_∞. It follows that each finite energy configuration has associated with it a winding number n, invariant under any continuous deformation preserving finiteness of energy (time evolution, for example). Thus configuration space is a disjoint union of sectors, each labelled by an integer n.

The winding number has very direct physical significance. We introduce a fictitious third spatial dimension orthogonal to physical $\mathbb{R}^2$, with unit vector $\hat{\mathbf{k}}$. Then $F_{12}\hat{\mathbf{k}} = -(\partial_1 A^2 - \partial_2 A^1)\hat{\mathbf{k}} = -\mathrm{curl}\mathbf{A}$, so we define $\mathbf{B} := -F_{12}\hat{\mathbf{k}}$ and call it the magnetic field. (In fact, the only reason for introducing $\hat{\mathbf{k}}$ is to allow the definition of the vector product between $\mathbb{R}^2$ vectors; it may be regarded as a purely algebraic device.) Continuing the analogy, we can define the magnetic flux through the physical $\mathbb{R}^2$ plane, denoted P,

$$\Phi := \int_P d\mathbf{S} \cdot \mathbf{B} = \int_P d\mathbf{S} \cdot \mathrm{curl}\mathbf{A} = \oint_{\partial P} d\mathrm{s} \cdot \mathbf{A}, \tag{5.12}$$

using Stokes' theorem. Substitute (5.11) into (5.10), the second finite energy condition:

$$\lim_{r\to\infty} r\hat{\boldsymbol{\theta}} \cdot \mathbf{A} = \gamma'(\theta). \tag{5.13}$$

Thus,

$$\Phi = \lim_{r\to\infty} \int_0^{2\pi} d\theta\, r\hat{\boldsymbol{\theta}} \cdot \mathbf{A} = \int_0^{2\pi} d\theta\, \gamma'(\theta) = 2n\pi. \tag{5.14}$$

So the magnetic flux is quantized in multiples of 2π, and the winding number of a configuration is the number of flux quanta associated with it, a topologically conserved quantity.

By continuity, a configuration with winding n must have exactly n zeros (of the Higgs field) in the plane, counted with multiplicity. A zero $\mathbf{x}_1 \in \mathbb{R}^2$ (so that $\phi(\mathbf{x}_1) = 0$) has multiplicity n_1 if on completing an infinitesimally small loop encircling $\mathbf{x}_1$ the Higgs field acquires a phase of $2n_1\pi$. Once again, single valuedness of ϕ requires that $n_1 \in \mathbb{Z}$. A configuration with $n = 1$, and with this winding centred on only one

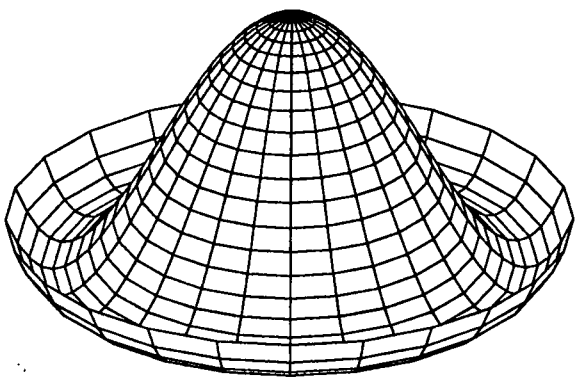

Figure 5.1: The Higgs symmetry breaking potential. Note that $\phi = 0$ is a local maximum.

simple zero, is called a vortex, and the $n = 1$ sector of configuration space is called the vortex sector. The image of a vortex under the parity bijection (a symmetry of the model) $x^1 \to x^1, x^2 \to -x^2$ is the analogous $n = -1$ configuration, called an antivortex. In fact entire negative winding sectors are paired with their positive winding counterparts in this way, so we may (and henceforth will), without loss of generality, consider only the $n \geq 0$ sectors. Examining the symmetry breaking potential (figure 5.1) we see that a zero of ϕ is likely to be a local maximum, or close to a local maximum, of the configuration's potential energy density (the integrand of (5.7)). So, a vortex has an energy lump located at its single zero, and is topologically stable – it is a soliton. It is natural to identify the zero as the vortex position. For more general configurations, a multiplicity n_1 zero, $\mathbf{x}_1$, is regarded as n_1 coincident vortices (antivortices if $n_1 < 0$). In this way any configuration can be assigned a vortex number and an antivortex number, and interpreted as a collection of vortices and antivortices located at its zeros. During time evolution the loci of these zeros are (anti-)vortex trajectories: where vortex meets antivortex they may unwind, annihilating and releasing their pent-up energy; where two vortices meet they may coalesce, forming a bound state, or undergo scattering. It is this last phenomenon which we will attempt to analyze.

Numerical experiments [46] show that vortex dynamics splits into three regimes, depending on the value of μ. If $\mu < 1$ (the so-called type I case), vortices attract and the only static n-vortex solution is n coincident vortices. By contrast, if $\mu > 1$ (type II) vortices repel. Again, the only static solution is n coincident vortices, but it is unstable unless $n = 1$. The critical case $\mu = 1$ is mathematically most interesting. Static vortices neither attract nor repel, and static solutions exist for all vortex positions. The link between vortex position and zero of ϕ has a precise expression in Taubes' theorem, which establishes a bijection between sets of zeros in physical $\mathbb{R}^2$ and static solutions.

Crucial to this analysis is the Bogomol'nyi argument [14]. Let $(\phi, \mathbf{A})$ be a static configuration. Since $\dot{\phi} = 0, \dot{\mathbf{A}} = \mathbf{0}$, the constraint (5.8) is trivially satisfied and $T = 0$, so the field variational problem reduces to extremizing (in fact, minimizing) the potential energy functional. The minimals within each sector can be found using a Bogomol'nyi argument somewhat more subtle than in the sine-Gordon and $\mathbb{C}P^1$ models. Assuming positive winding, we start from the trivial inequality,

$$\begin{aligned} 0 &\leq \frac{1}{2}\int d^2\mathbf{x}\left\{|D_1\phi + iD_2\phi|^2 + \left[-F_{12} + \frac{1}{2}(|\phi|^2 - 1)\right]^2\right\} \\ &= E + \frac{1}{2}\int d^2\mathbf{x}\left[i\left(D_2\phi\overline{D_1\phi} - D_1\phi\overline{D_2\phi}\right) - F_{12}(|\phi|^2 - 1)\right], \end{aligned} \tag{5.15}$$

since $E = V$ for static fields. Now,

$$\begin{aligned} D_2\phi\overline{D_1\phi} - D_1\phi\overline{D_2\phi} &= \partial_2\phi\overline{D_1\phi} - \partial_1\phi\overline{D_2\phi} + i\phi(A_2\overline{D_1\phi} - A_1\overline{D_2\phi}) \\ &= \partial_2(\phi\overline{D_1\phi}) - \partial_1(\phi\overline{D_2\phi}) \\ &\quad -\phi(\partial_2\overline{D_1\phi} - \partial_1\overline{D_2\phi} - iA_2\overline{D_1\phi} + iA_1\overline{D_2\phi}) \\ &= \partial_2(\phi\overline{D_1\phi}) - \partial_1(\phi\overline{D_2\phi}) + \phi(D_1D_2\phi - D_2D_1\phi)^* \\ &= \partial_2(\phi\overline{D_1\phi}) - \partial_1(\phi\overline{D_2\phi}) - i|\phi|^2F_{12}, \end{aligned} \tag{5.16}$$

since $[D_\mu, D_\nu] \equiv iF_{\mu\nu}$. Note that the first two terms of (5.16), above, may be written

in $\mathbb{R}^3$ language as $-\hat{\mathbf{k}} \cdot \mathrm{curl}(\phi\overline{\mathbf{D}\phi})$. Substituting into (5.15),

$$\begin{aligned} 0 &\leq E - \frac{i}{2}\int_P d\mathbf{S} \cdot \mathrm{curl}(\phi\overline{\mathbf{D}\phi}) + \frac{1}{2}\int d^2\mathbf{x}\, F_{12} \\ &= E - \frac{i}{2}\oint_{\partial P} d\mathbf{s} \cdot (\phi\overline{\mathbf{D}\phi}) - \frac{1}{2}\Phi, \end{aligned} \tag{5.17}$$

using Stokes' theorem. Finite energy requires that $\mathbf{D}\phi = \mathbf{0}$ on ∂P, so the boundary integral vanishes. Hence, by flux quantization (5.14),

$$E \geq n\pi, \tag{5.18}$$

equality holding if and only if

$$\begin{aligned} (D_1 + iD_2)\phi &= 0, \\ -F_{12} + \frac{1}{2}(|\phi|^2 - 1) &= 0. \end{aligned} \tag{5.19}$$

Of course, the bound is valid for time-varying fields also since kinetic energy cannot be negative, but it can only be saturated by static fields. Once again, the Bogomol'nyi equations are first order, in contrast with the Euler-Lagrange equations, which are second order.

No explicit, nontrivial solutions of the Bogomol'nyi equations are known. However, Taubes has proved [47] that for each unordered set of n points in the physical plane $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ there exists a solution, unique modulo gauge transformations, of the pair (5.19) with the correct boundary behaviour, and $\phi(\mathbf{x}_i) = 0, i = 1, 2, \ldots, n$, ϕ vanishing nowhere else in $\mathbb{R}^2$. If a point $\mathbf{x}_p$ appears in the set n_p times, then the solution has a zero of multiplicity n_p at $\mathbf{x}_p$. So, static critical vortices exert no net forces on each other, although vortices set off in relative motion do exert velocity dependent forces, causing deviation from straight-line motion, and scattering.

As an aside, it is interesting to note that the critical coupling Bogomol'nyi bound (5.18) can be used to place topological lower bounds on the energy of the system

at general coupling. Denote by E_{μ^2} the static energy functional of the model with Higgs mass μ, evaluated at a point in the winding $n \geq 1$ sector of configuration space $(\phi, \mathbf{A})$. So E_1 is the critical energy functional and $E_1 \geq n\pi$. First consider the type II case, $\mu^2 = 1 + |\epsilon| > 1$. Then from (5.7),

$$E_{\mu^2} = E_1 + \frac{|\epsilon|}{8} \int d^2\mathbf{x} \, (|\phi|^2 - 1)^2 > n\pi, \tag{5.20}$$

the inequality now being strict, and hence the bound unsaturable, because the vanishing of the integral above is incompatible with non-zero winding. Turning now to type I, $\mu^2 = 1/(1 + |\epsilon|) < 1$,

$$E_{\mu^2} = \frac{E_1}{1 + |\epsilon|} + \frac{|\epsilon|}{2(1 + |\epsilon|)} \int d^2\mathbf{x} \, \left[\mathbf{D}\phi \cdot \overline{\mathbf{D}\phi} + F_{12}^2 \right] > \mu^2 n\pi. \tag{5.21}$$

Again, the inequality is strict: the integral cannot vanish, because the configuration has non-zero total magnetic flux, so F_{12} cannot be everywhere zero. Numerical work [48] shows that the bounds (5.20), (5.21) are, further than unsaturable, not optimal. That is, $\mu^2 n\pi$ for $\mu^2 < 1$ and $n\pi$ for $\mu^2 > 1$ are neither minima nor infima of the energy functional E_{μ^2} evaluated on the winding n sector.

We return now to the main object of our study: the two-vortex scattering problem at arbitrary coupling. The starting point of our calculation is the assertion that the essentials of soliton dynamics are in the soliton degrees of freedom (vortex positions in this case) which constitute a finite subset of the infinitely many field degrees of freedom. We wish to find a Lagrangian mechanical system whose coordinates are the vortex positions and which describes, in a low energy approximation, two-vortex scattering. We seek to deduce this mechanical system from the field equations, then compare its predictions with numerical simulations.

Several techniques have been developed to achieve this. Most notable is the geodesic approximation, already described in previous chapters. This is valid when one has a saturable Bogomol'nyi bound, and has been successfully used [49] to analyze

vortex scattering at critical coupling. Another technique, first used to describe long range interactions of monopoles of the Yang-Mills-Higgs system in the BPS limit, is the method of linear retarded potentials. Far from a monopole's core, its fields are indistinguishable from those induced by a point (Dirac) monopole in a *linear* field theory. If physics is to be model independent, long range BPS monopole interactions should be the same as Dirac monopole interactions in the linear theory. Analyzing the latter system, Manton [50] re-derived the asymptotic form of the Atiyah-Hitchin metric [51, 52]. The assumed model independence is observed and has a precise mathematical expression. A similar approach is to use a field superposition ansatz for two-soliton configurations, patching together two well-separated, single solitons. This was the method used by Bettencourt and Rivers [53] to find *static* intervortex forces at arbitrary coupling. In fact, this and the method of linear retarded potentials are really two different ways of organizing the same calculations, as demonstrated explicitly for the case of Skyrmions (in the Skyrme model with *massless* pions) by Schroers [54]. We prefer the latter viewpoint over the superposition ansatz approach because the calculations are more compact, and at each stage there is contact with a concrete physical interpretation.

In outline then, the programme is:

1. Find the asymptotic expression for a single, static vortex solution.

2. Derive the linear field theory associated with the abelian Higgs model.

3. Identify the point source which replicates the vortex asymptotics in the linear theory.

4. Study intervortex forces by calculating an interaction Lagrangian for the point sources in the framework of the linear theory.

The novelty of the calculation is twofold: we will find *velocity dependent* forces (beyond the static analysis of [53]) in a *massive* theory. All previous applications of the

method have been to systems whose associated linear theories are massless – BPS monopoles [50], maximally charged black holes [55] and Skyrmions [54]. These have the simplifying feature that field disturbances travel uniformly at the speed of light and one may use retarded potentials to calculate intersource forces. In the vortex case, as in many other cases of physical interest, the linear theory is massive, so conventional retarded potentials are not available to us.

5.2 Asymptotics of the field equations

The first task is to find the asymptotic behaviour of vortex solutions at large r. It is convenient to obtain the Euler-Lagrange equations in general curvilinear coordinates, in terms of which the Minkowski metric is $g_{\mu\nu}$ and has (modulus of) determinant g. The action is

$$S[\phi, A] = \frac{1}{2}\int d^3x\,\sqrt{g}\left[D_\mu\phi g^{\mu\nu}\overline{D_\nu\phi} - \frac{1}{4}g^{\mu\alpha}g^{\nu\beta}F_{\mu\nu}F_{\alpha\beta} - \frac{1}{4}\mu^2(\phi\overline{\phi}-1)^2\right]. \qquad (5.22)$$

Variation with respect to $\overline{\phi}$ and A_ν yields, respectively,

$$\partial_\mu\left[\sqrt{g}D_\nu\phi g^{\nu\mu}\right] + \sqrt{g}\left[iA_\nu g^{\mu\nu}D_\mu\phi - \frac{1}{2}\mu^2\phi(|\phi|^2-1)\right] = 0, \qquad (5.23)$$

and,

$$-\partial_\mu\left[\sqrt{g}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}\right] - \sqrt{g}g^{\nu\beta}\left[\frac{i}{2}(\phi\partial_\beta\overline{\phi} - \overline{\phi}\partial_\beta\phi) + |\phi|^2A_\beta\right] = 0, \qquad (5.24)$$

if (ϕ, A) is an extremal of $S[\phi, A]$. The most natural coordinates in the vortex sector are plane polar. The Minkowski inverse metric is

$$(g^{\mu\nu}) = \begin{pmatrix} g^{00} & g^{0r} & g^{0\theta} \\ g^{r0} & g^{rr} & g^{r\theta} \\ g^{\theta 0} & g^{\theta r} & g^{\theta\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -\frac{1}{r^2} \end{pmatrix}, \qquad (5.25)$$

so that $g = r^2$. Substituting the ansatz ($\sigma \in \mathbb{R}$)

$$\begin{aligned} \phi &= \sigma(r)e^{i\theta}, \\ (A_0, A_r, A_\theta) &= (0, 0, -a(r)), \end{aligned} \tag{5.26}$$

the field equations reduce to

$$\frac{d^2\sigma}{dr^2} + \frac{1}{r}\frac{d\sigma}{dr} - \frac{1}{r^2}\sigma(1-a)^2 - \frac{1}{2}\mu^2\sigma(\sigma^2-1) = 0 \tag{5.27}$$

and

$$\frac{d^2a}{dr^2} - \frac{1}{r}\frac{da}{dr} + (1-a)\sigma^2 = 0, \tag{5.28}$$

the equations for A_0 and A_r being trivially satisfied. Note that this ansatz has unit winding by construction. The finite energy conditions (5.9),(5.10) become

$$\begin{aligned} |\phi| \to 1 &\Rightarrow \sigma \to 1, \\ D_\theta\phi \to 0 &\Rightarrow (\partial_\theta - ia)e^{i\theta} \to 0 \Rightarrow a \to 1, \end{aligned} \tag{5.29}$$

as $r \to \infty$. In addition, regularity demands that $\sigma(0) = a(0) = 0$. No explicit, exact solutions of (5.27),(5.28) with these boundary conditions are known. Numerical solutions suggest that both σ and a rise monotonically from 0 to 1 as r covers $[0, \infty)$, so $\mathbf{x} = \mathbf{0}$ is the only zero of ϕ, and (σ, a) truly represents a static vortex solution.

We are only interested in the asymptotic forms of σ and a, and for these analytic expressions *do* exist. Define the functions α and β such that

$$\begin{aligned} \sigma(r) &= 1 + \alpha(r), \\ a(r) &= 1 + \beta(r). \end{aligned} \tag{5.30}$$

At large r, (5.29) implies that α and β are small, so we substitute (5.30) into

(5.27),(5.28) and linearize in α and β:

$$\begin{aligned} \frac{d^2\alpha}{dr^2} + \frac{1}{r}\frac{d\alpha}{dr} - \mu^2\alpha &= 0 \\ \frac{d^2\beta}{dr^2} - \frac{1}{r}\frac{d\beta}{dr} - \beta &= 0. \end{aligned} \tag{5.31}$$

These equations hold asymptotically as $r \to \infty$. Now define $\tilde{r} := \mu r$ and $\tilde{\beta} := \beta/r$. Rewriting the pair (5.31),

$$\begin{aligned} \frac{d^2\alpha}{d\tilde{r}^2} + \frac{1}{\tilde{r}}\frac{d\alpha}{d\tilde{r}} - \alpha &= 0, \\ \frac{d^2\tilde{\beta}}{dr^2} + \frac{1}{r}\frac{d\tilde{\beta}}{dr} - \left(1 + \frac{1}{r^2}\right)\tilde{\beta} &= 0, \end{aligned} \tag{5.32}$$

we obtain the modified Bessel's equations of zeroth and first order respectively. It follows that

$$\begin{aligned} \alpha &\sim \frac{q}{2\pi} K_0(\mu r) \\ \beta &\sim \frac{m}{2\pi} r K_1(r) \end{aligned} \tag{5.33}$$

at large r, where $K_n, n = 0, 1$ is the n-th modified Bessel's function of the second kind [56]. Note that $K_1(r) = -K_0'(r)$. Since we have linearized the field equations, the asymptotic solutions contain unknown scale constants q and m which can only be fixed by solving (5.27), (5.28) numerically. It is the asymptotic behaviour, (5.33), that we shall replicate with appropriate sources in the associated linear field theory.

5.3 The point sources

As will be shown below, the linear theory associated with (5.1) consists of two uncoupled fields – a real scalar field with mass μ, and a real vector field of unit mass. The latter is *not* a gauge field, because the mass term breaks gauge invariance. We obtain a real scalar field by making a gauge choice. Since the scalar field is real, there is no

sense in which it can have winding at infinity. The winding number of a configuration of the nonlinear model is gauge invariant. Consider any smooth gauge transformation, $U(r,\theta) = \exp(i\chi(r,\theta))$. For each value of r this is a map from the circle at r, parametrized by θ, to $U(1) \cong S^1$. If this transformation changes the winding number n of ϕ, then $U(\infty,\theta)$ must itself have non-zero winding. But if the transformation is to be single valued at the origin, $\chi(0,\theta)$ must be a constant, so $U(0,\theta)$ must have zero winding. Then the r dependence of χ defines a homotopy between two maps $S^1 \to S^1$, $U(0,\theta)$ and $U(\infty,\theta)$, which lie in different homotopy sectors, a contradiction. Hence no such gauge transformation exists – n is gauge invariant. How, then, are we to make the comparison between a vortex, which has unit winding, and a solution of the linear theory, which has none? A key ingredient in the above argument is the demand of regularity at the origin, the vortex centre. However, we only require comparison at large r, well away from the origin, so for our purposes we can weaken this demand to regularity on $\mathbb{R}^2\backslash\{\mathbf{0}\}$. There exist gauge transformations regular on $\mathbb{R}^2\backslash\{\mathbf{0}\}$ which change the winding of a configuration. Applying one of these to the vortex, it can be unwound and compared with a linear theory configuration. Since we will introduce a singular point source into the linear theory, this is, from the outset, regular only on $\mathbb{R}^2\backslash\{\mathbf{0}\}$, not all $\mathbb{R}^2$.

Take the Lagrangian density (5.1) with gauge chosen so that $\phi \in \mathbb{R}$ (such a gauge choice is singular at vortex and antivortex positions, as described above). The vacuum is then $\phi = 1$, so we define ψ such that $\phi = 1 + \psi$, substitute into (5.1) and neglect terms of higher than quadratic order in ψ and A_μ, yielding

$$\mathcal{L}_{free} = \frac{1}{2}\partial_\mu\psi\partial^\mu\psi - \frac{1}{2}\mu^2\psi^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}A_\mu A^\mu. \tag{5.34}$$

Including the external source Lagrangian,

$$\mathcal{L}_{source} = \rho\psi - j_\mu A^\mu \tag{5.35}$$

with scalar density ρ and vector current j_μ, we obtain the following massive wave equations for ψ and A_μ,

$$(\Box + \mu^2)\psi = \rho \tag{5.36}$$

$$\Box A_\mu - \partial_\mu(\partial_\nu A^\nu) + A_\mu = j_\mu, \tag{5.37}$$

where $\Box = \partial_\nu \partial^\nu$, the d'Alembertian operator. Taking the divergence of (5.37), one finds

$$\partial_\mu A^\mu = \partial_\mu j^\mu, \tag{5.38}$$

a fact which would usually lead to the conclusion $\partial_\mu A^\mu = 0$, and the simplification of (5.37) to the Proca equation. In this case there is no global $U(1)$ symmetry of $\mathcal{L}_{free}$ with whose Noether current we can identify j_μ (because ψ is *real*), so there is no reason to assume that j_μ is a conserved current. On substitution of (5.38), (5.37) becomes

$$(\Box + 1)A_\mu = j_\mu + \partial_\mu(\partial_\nu j^\nu). \tag{5.39}$$

The vortex ansatz has $\phi = \sigma(r)e^{i\theta}$, so the gauge choice $\phi \in \mathbb{R}$ requires a singular gauge transformation with $\chi(r, \theta) = -\theta$. The unwound vortex Higgs field is

$$\phi = \sigma(r) = 1 + \alpha(r) \sim 1 + \frac{q}{2\pi} K_0(\mu r). \tag{5.40}$$

The gauge field transforms $A_\mu \mapsto A_\mu - \partial_\mu \chi$, so $A_0 \mapsto A_0$, $A_r \mapsto A_r$, $A_\theta \mapsto A_\theta + 1$. That is, the unwound vector field is

$$A_\theta = 1 - a(r) = -\beta(r) \sim -\frac{m}{2\pi} r K_1(r), \tag{5.41}$$

while $A_r = A_0 = 0$. In terms of the 2-vector field $\mathbf{A}$, the unwound asymptotic behaviour is

$$\mathbf{A} \sim -\frac{m}{2\pi} K_0'(r)\widehat{\boldsymbol{\theta}} = -\frac{m}{2\pi} \widehat{\mathrm{k}} \times \nabla K_0(r). \tag{5.42}$$

So, we need to find ρ such that the solution of (5.36) is

$$\psi = \frac{q}{2\pi} K_0(\mu r), \tag{5.43}$$

and j_μ such that the solution of (5.39) is

$$\mathbf{A} = -\frac{m}{2\pi} \hat{\mathbf{k}} \times \nabla K_0(r). \tag{5.44}$$

The static Klein-Gordon equation in two dimensions has Green's function K_0,

$$(-\Delta + \mu^2) K_0(\mu r) = 2\pi \delta(\mathbf{x}), \tag{5.45}$$

where $\Delta = \nabla \cdot \nabla$, the Laplacian on $\mathbb{R}^2$. Substituting (5.43) into (5.36), and using (5.45), one finds

$$\begin{aligned} (\Box + \mu^2) \left(\frac{q}{2\pi} K_0(\mu r) \right) &= \frac{q}{2\pi} (-\Delta + \mu^2) K_0(\mu r) = q\delta(\mathbf{x}) \\ \Rightarrow \rho &= q\delta(\mathbf{x}). \end{aligned} \tag{5.46}$$

Similarly, substitution of (5.44) into (5.39) yields

$$\begin{aligned} (\Box + 1) \left(-\frac{m}{2\pi} \hat{\mathbf{k}} \times \nabla K_0(r) \right) &= -\frac{m}{2\pi} \hat{\mathbf{k}} \times \nabla (-\Delta + 1) K_0(r) = -m\hat{\mathbf{k}} \times \nabla \delta(\mathbf{x}) \\ \Rightarrow \mathbf{j} - \nabla(\nabla \cdot \mathbf{j}) &= -m\hat{\mathbf{k}} \times \nabla \delta(\mathbf{x}). \end{aligned} \tag{5.47}$$

Taking the divergence of (5.47), one sees that $\nabla \cdot \mathbf{j}$ must be a solution of the *homogeneous* static Klein-Gordon equation, so if $\mathbf{j}$ is a localized point source, $\nabla \cdot \mathbf{j} = 0$. Thus, the unique point source satisfying (5.47) is

$$\mathbf{j} = -m\hat{\mathbf{k}} \times \nabla \delta(\mathbf{x}), \tag{5.48}$$

which does indeed have the property that $\nabla \cdot \mathbf{j} = 0$. Since $A^0 = 0$, we take $j^0 = 0$.

The physical interpretation of these expressions for ρ and $\mathbf{j}$ is that the point source consists of a scalar charge q and a magnetic dipole of moment m. We refer to this composite source as a point vortex. The monotonicity properties of σ and a, noted above, lead to the expectation that both q and m are negative. This is perhaps a little surprising in the case of m, because the magnetic field of a vortex points in the $+\hat{\mathbf{k}}$ direction on all $\mathbb{R}^2$, in the opposite sense to the point vortex's magnetic dipole. This is a peculiarity of *massive* electrodynamics in $(2+1)$ dimensions. If the photon mass, here normalized to unity, was zero, there would be no magnetic field away from $\mathbf{x} = \mathbf{0}$ at all. This is most easily seen by imagining the system embedded in $\mathbb{R}^3$ with translation symmetry along the $\hat{\mathbf{k}}$ direction. The source (5.48) then represents an infinitely long, thin solenoid perpendicular to the (x^1, x^2) plane, piercing it at $\mathbf{x} = \mathbf{0}$. We know that the magnetic flux induced by such a solenoid is confined entirely within the solenoid, that is, the line $x^1 = x^2 = 0$, so that $\mathbf{B} = \mathbf{0}$ except on this line. The sense of $\mathbf{B}$ on the line is given by the right hand screw rule. Extending this picture to the case with non-zero photon mass is dangerous. In fact, although the source (5.48) represents (for $m < 0$) a vanishingly small *clockwise* current loop, the resulting $\mathbf{B}$ field in the massive theory points in the $+\hat{\mathbf{k}}$ direction on all $\mathbb{R}^2 \backslash \{\mathbf{0}\}$,

$$\begin{aligned} \mathbf{A} &= -\frac{m}{2\pi}\hat{\mathbf{k}} \times \nabla K_0(r) = -\frac{|m|}{2\pi}\nabla \times (K_0(r)\hat{\mathbf{k}}) \\ \Rightarrow \mathbf{B} &= \nabla \times \mathbf{A} = \frac{|m|}{2\pi}\Delta K_0(r)\hat{\mathbf{k}} = \frac{|m|}{2\pi}K_0(r)\hat{\mathbf{k}}, \end{aligned} \tag{5.49}$$

using (5.45) with $\mathbf{x} \neq \mathbf{0}$. So, somewhat unexpectedly, we need a *clockwise* current loop in the linear theory to reproduce the correct asymptotic behaviour. One should also note an important difference between the present endeavour and the analogous analysis for BPS monopoles [50]. Both systems (abelian Higgs and Yang-Mills-Higgs) have topologically quantized magnetic flux Φ. In the monopole calculation, the Dirac monopole charge is assumed to have precisely the topological value, while we must fix the dipole moment of the point vortex by numerical solution of the field equations.

The difference lies in the fact that monopoles exist in $\mathbb{R}^3$, where magnetic flux is defined as flux through a spherical shell of large radius centred on the monopole – an asymptotic quantity, whereas vortices exist in $\mathbb{R}^2$, where Φ is flux through the whole plane, including the vortex interior. Such a quantity is not asymptotic, and so is not reflected in the properties of the point vortex.

The final stage in the identification of the point sources is determination of the numerical values of q and m. To do this one must numerically solve (5.27), (5.28) at each value of μ^2 required, and evaluate

$$\begin{aligned} q &= \lim_{r\to\infty} 2\pi \left(\frac{\sigma(r)-1}{K_0(\mu r)} \right), \\ m &= \lim_{r\to\infty} 2\pi \left(\frac{a(r)-1}{rK_1(r)} \right), \end{aligned} \tag{5.50}$$

approximately. Rather than solving the boundary value problem $\sigma(0) = a(0) = 0, \sigma(\infty) = a(\infty) = 1$, we solve the initial value problem $\sigma(0) = a(0) = 0, \sigma'(0) = \xi, a'(0) = \zeta$, using a fourth-order Runge-Kutta algorithm, where the constants ξ and ζ must be adjusted until the solution has the correct behaviour at large r, that is, until $\sigma(r_\infty) = a(r_\infty) = 1$, the "effective infinity" r_∞ being some large but finite radius (in this case, 10 or 8 depending on μ^2). The main difficulty with this shooting method is that there are two shooting parameters, ξ and ζ, to be adjusted. Shooting with only one parameter is much easier because if one finds a pair of parameter values one of which shoots too high, the other too low, then the true value must lie between them (barring unforeseen singular behaviour). We must aim for the point $(\sigma(r_\infty), a(r_\infty)) = (1,1)$ in $\mathbb{R}^2$, so the terms "too high" and "too low" have no meaning. The only systematic approach is to discretize shooting parameter space (call this space $\mathbb{R}^2_0$), (ξ, ζ) taking values on a lattice. We start with a coarse 10×10 lattice covering a large square in $\mathbb{R}^2_0$, outputing $(\sigma(r_\infty), a(r_\infty))$ for each point (ξ, ζ) on the lattice. Contour plots of $\sigma(r_\infty)$ and $a(r_\infty)$ give the $\sigma(r_\infty) = 1$ and $a(r_\infty) = 1$ contours in $\mathbb{R}^2_0$. These intersect at a unique point in $\mathbb{R}^2_0$, which we take as the central point of our

μ^2	q	m	r_c	r_s
0.4	-7.54	-14.92	4.23	3.05
0.6	-8.71	-12.61	3.75	2.73
0.8	-9.70	-11.31	3.43	2.53
0.9	-10.14	-10.89	3.22	2.45
1.0	-10.58	-10.57	-	2.39
1.1	-10.98	-10.31	2.98	2.34
1.2	-11.43	-10.06	3.07	2.29
1.3	-11.80	-9.85	2.96	2.25
1.4	-12.23	-9.66	2.95	2.21
1.6	-13.04	-9.34	2.88	2.15
1.8	-13.97	-9.09	2.87	2.10
2.0	-14.50	-8.86	2.72	2.04

Table 5.1: Numerical values of vortex scalar charge q and magnetic dipole moment m. The other data are critical radii: equilibrium separation r_c of the static intervortex potential and the metric singularity r_s at which the signature of the metric on moduli space changes.

new, finer lattice, still 10×10, but concentrated on a square of one hundredth the area. The procedure is repeated so that we "zoom in" on the correct parameter values (ξ, ζ), eventually searching a tiny area of $\mathbb{R}_0^2$ with a very fine lattice. This allows us to determine (ξ, ζ) to the considerable number of figures required for the numerical solution to behave as specified at large r. Searching all of $\mathbb{R}_0^2$ (i.e. the original guess patch) so finely would be prohibitively costly in computation time.

The results of this procedure may be seen in table 5.1 (the rightmost two columns will be explained in subsequent sections). They should be treated with some caution: at large r, the field equations (5.27), (5.28) collapse to Bessel's equations (5.32) which each have two independent solutions: one exponentially decaying, the other exponentially growing. We seek to pick out the former and completely exclude the latter, an impossible task. All numerical solutions blow up at large r, so, particularly when μ^2 is far from 1, r_∞ must be smaller than one would like (10 or 8). Comparison with the asymptotic forms must be made where the solution is barely asymptotic. Nevertheless, certain interesting features can be discerned. First, when $\mu = 1$, the

case of critical coupling, $q \approx m$. In fact one can prove that $q \equiv m$, because the $\mu = 1$ vortex satisfies the first order Bogomol'nyi equations (5.19), which one can use to *derive* the asymptotic behaviour of a from that of σ. Substituting the vortex ansatz into (5.19) yields,

$$r\frac{d\sigma}{dr} - (1-a)\sigma = 0, \tag{5.51}$$

$$\frac{2}{r}\frac{da}{dr} + (\sigma^2 - 1) = 0. \tag{5.52}$$

Define α and β as before, substitute into (5.51) and linearize:

$$\beta = -r\frac{d\alpha}{dr}. \tag{5.53}$$

Thus,

$$\alpha = \frac{q}{2\pi}K_0(r) \Rightarrow \beta = \frac{q}{2\pi}rK_1(r), \tag{5.54}$$

and so $m \equiv q$. To find q when $\mu = 1$, we can replace (5.51), (5.52) by a single second order ordinary differential equation in σ alone, then perform a *one parameter* (namely ξ) shooting method; a relatively simple task, as explained above. The correct value of ζ is fixed by ξ, so the point $(\xi, \zeta) \in \mathbb{R}^2_0$ for $\mu^2 = 1$ can be determined without the complicated lattice procedure. This point is placed at the centre of our initial lattice for the $\mu^2 = 1.1$ calculation. Similarly, the correct (ξ, ζ) for $\mu^2 = 1.1$ is placed at the centre of the initial $\mu^2 = 1.2$ lattice, and so on. In this way, we work piecemeal away from $\mu^2 = 1$ into both the type I and type II regimes.

The second striking feature of table 5.1 is that q increases monotonically, while m decreases monotonically with μ^2. We shall see that this behaviour leads to unexpected critical points of the type I and type II static intervortex potentials. Notice also that although $q \equiv m$ when $\mu^2 = 1$, q/m varies between 0.50 and 1.64 as μ^2 covers $[0.4, 2.0]$. So the assumption that $q \approx m$ away from critical coupling, made in [53], seems suspect.

5.4 Static intervortex forces

Having found the scalar charge and magnetic dipole moment carried by a point vortex, it is straightforward to calculate the force between two such vortices held at rest, in the framework of the linear theory. We begin by finding the interaction Lagrangian for two arbitrary (possibly time dependent) sources $(\rho_1, j_{(1)})$ and $(\rho_2, j_{(2)})$. Consider first the scalar interaction. The scalar Lagrangian is

$$L_S[\psi, \rho] = \int d^2\mathbf{x} \left[\frac{1}{2}\partial_\mu\psi\partial^\mu\psi - \frac{1}{2}\mu^2\psi^2 + \rho\psi\right]. \tag{5.55}$$

Let the field due to source ρ_i be ψ_i. Since the theory is linear, it obeys the superposition principle, so the field due to the combined source $\rho_1 + \rho_2$ is $\psi_1 + \psi_2$. Evaluating $L_s[\psi_1 + \psi_2, \rho_1 + \rho_2]$ we identify the cross terms as the interaction Lagrangian, L_ψ,

$$\begin{aligned} L_\psi &:= L_S[\psi_1 + \psi_2, \rho_1 + \rho_2] - L_S[\psi_1, \rho_1] - L_S[\psi_2, \rho_2] \\ &= \int d^2\mathbf{x} \left[\partial_\mu\psi_1\partial^\mu\psi_2 - \mu^2\psi_1\psi_2 + \rho_1\psi_2 + \rho_2\psi_1\right] \\ &= \int d^2\mathbf{x} \left[-\psi_1(\partial_\mu\partial^\mu\psi_2 + \mu^2\psi_2) + \rho_1\psi_2 + \rho_2\psi_1\right] \\ &= \int d^2\mathbf{x} \left[-\psi_1\rho_2 + \rho_1\psi_2 + \rho_2\psi_1\right] \\ &= \int d^2\mathbf{x}\, \rho_1\psi_2, \end{aligned} \tag{5.56}$$

using integration by parts and the inhomogeneous Klein-Gordon equation for ψ_2. Turning to the vector interaction, we find a similar expression. The vector Lagrangian is

$$L_V = \int d^2\mathbf{x} \left[-\frac{1}{2}(\partial_\mu A_\nu\partial^\mu A^\nu - \partial_\mu A_\nu\partial^\nu A^\mu) + \frac{1}{2}A_\mu A^\mu - j_\mu A^\mu\right]. \tag{5.57}$$

Again the superposition principle applies. Writing $A_{(i)}$ for the field due to $j_{(i)}$, the interaction Lagrangian is

$$L_A := L_V[A_{(1)} + A_{(2)}, j_{(1)} + j_{(2)}] - L_V[A_{(1)}, j_{(1)}] - L_V[A_{(2)}, j_{(2)}]$$

$$
\begin{aligned}
&= \int d^2\mathbf{x}\left[-(\partial_\mu A^{(1)}_\nu \partial^\mu A^\nu_{(2)} - \partial_\mu A^{(1)}_\nu \partial^\nu A^\mu_{(2)}) + A^{(1)}_\mu A^\mu_{(2)} - j^{(1)}_\mu A^\mu_{(2)} - j^{(2)}_\mu A^\mu_{(1)}\right] \\
&= \int d^2\mathbf{x}\left[A^{(1)}_\nu(\partial_\mu\partial^\mu A^\nu_{(2)} - \partial^\nu\partial_\mu A^\mu_{(2)} + A^\nu_{(2)}) - j^{(1)}_\mu A^\mu_{(2)} - j^{(2)}_\mu A^\mu_{(1)}\right] \\
&= -\int d^2\mathbf{x}\, j^{(1)}_\mu A^\mu_{(2)},
\end{aligned}
\tag{5.58}
$$

using the field equation for $A_{(2)}$.

We may now specialize to the case of two static point vortices. Let vortex 1 be at $\mathbf{y}$, vortex 2 be at $\mathbf{z}$. Then the scalar source for 1 is $\rho_{(1)} = q\delta(\mathbf{x}-\mathbf{y})$, while the scalar field due to source ρ_2 is $\psi_2 = qK_0(\mu|\mathbf{x}-\mathbf{z}|)/2\pi$. So,

$$
L_\psi = \int d^2\mathbf{x}\,\frac{q^2}{2\pi}\delta(\mathbf{x}-\mathbf{y})K_0(\mu|\mathbf{x}-\mathbf{z}|) = \frac{q^2}{2\pi}K_0(\mu|\mathbf{y}-\mathbf{z}|). \tag{5.59}
$$

The magnetic interaction is similar. Substituting $j^0_{(1)} = 0$, $\mathbf{j}_{(1)} = -m\hat{\mathbf{k}}\times\nabla\delta(\mathbf{x}-\mathbf{y})$, $A^0_{(2)} = 0$, $\mathbf{A}_{(2)} = -m\hat{\mathbf{k}}\times\nabla K_0(|\mathbf{y}-\mathbf{z}|)/2\pi$ into (5.58), we find

$$
\begin{aligned}
L_A &= \int d^2\mathbf{x}\,\frac{m^2}{2\pi}[\hat{\mathbf{k}}\times\nabla\delta(\mathbf{x}-\mathbf{y})]\cdot[\hat{\mathbf{k}}\times\nabla K_0(|\mathbf{x}-\mathbf{z}|)] \\
&= -\frac{m^2}{2\pi}\int d^2\mathbf{x}\,\delta(\mathbf{x}-\mathbf{y})\Delta K_0(|\mathbf{x}-\mathbf{z}|) \\
&= -\frac{m^2}{2\pi}\Delta_y K_0(|\mathbf{y}-\mathbf{z}|) \\
&= -\frac{m^2}{2\pi}K_0(|\mathbf{y}-\mathbf{z}|),
\end{aligned}
\tag{5.60}
$$

using (5.45) with $\mathbf{y}\neq\mathbf{z}$. The total interaction Lagrangian is $L_A + L_\psi =: -\mathcal{U}$. Since this is a function of the positions $\mathbf{y}$ and $\mathbf{z}$ only, we interpret $\mathcal{U}$ as the potential energy of the interaction,

$$
\mathcal{U} = \frac{1}{2\pi}\left[m^2K_0(r) - q^2K_0(\mu r)\right], \tag{5.61}
$$

having defined $\mathbf{r}$ to be the vortex displacement vector, $\mathbf{r} := \mathbf{y}-\mathbf{z}$, so that $r = |\mathbf{r}|$ is the vortex separation. This is the same potential as found in [53], but we arrived at it via a different route.

Note that (5.61) is consistent with the partition into type I, critical and type II

regimes. The (central) force due to $\mathcal{U}$ is

$$-\frac{d\mathcal{U}}{dr} = 2\pi[m^2 K_1(r) - \mu q^2 K_1(\mu r)]. \tag{5.62}$$

If $\mu < 1$ then $K_1(r) \to 0$ at large r faster than $K_1(\mu r)$, so that scalar attraction dominates over magnetic repulsion, and the force is negative, indicating vortex attraction. In the critical case $\mu = 1$, we have proved that $m \equiv q$, implying that $\mathcal{U} \equiv 0$, so there is no net force at all. When $\mu > 1$, magnetic repulsion dominates over scalar attraction at large r (because $K_1(\mu r)$ vanishes more quickly than $K_1(r)$) and type II behaviour emerges. This consistency at large r happens regardless of the specific values of q and m, and can be attributed to the inverse relationship between a field's mass and its range. However, the μ dependence of q/m greatly affects the qualitative nature of $\mathcal{U}$ at small to moderate values of r. Let us derive conditions on the function $\kappa(\mu) := q(\mu)/m(\mu)$ for there to exist critical points of the potential $\mathcal{U}$. If r_c is a critical point, then, from (5.62),

$$\mu\kappa^2(\mu) = \frac{K_1(r_c)}{K_1(\mu r_c)}. \tag{5.63}$$

Using the properties of Bessel's functions [56], one sees that the right hand side of (5.63) is a monotonically decreasing function (of r_c) with range $(0,1)$ if $\mu < 1$, and a monotonically increasing function with range $(1,\infty)$ if $\mu > 1$. It follows that there exists a unique critical point $r_c(\mu)$ of the potential for all $\mu \neq 1$ if and only if $\kappa(\mu) < 1/\sqrt{\mu}$ when $\mu < 1$ and $\kappa(\mu) > 1/\sqrt{\mu}$ when $\mu > 1$. We have proved that $\kappa(1) = 1$, and our numerical work suggests that $\kappa(\mu)$ is a monotonically increasing function, so the actual values of κ easily pass these criteria. The results of solving (5.63) numerically for the critical radius $r_c(\mu)$ are presented in table 5.1. Potentials for $\mu = 0.4$ (type I) and $\mu = 2.0$ (type II) are plotted in figure 5.2. Rebbi and Jacobs [48] have found approximate static intervortex potentials by numerically minimizing the potential energy functional (5.7) subject to the constraint that ϕ has two simple

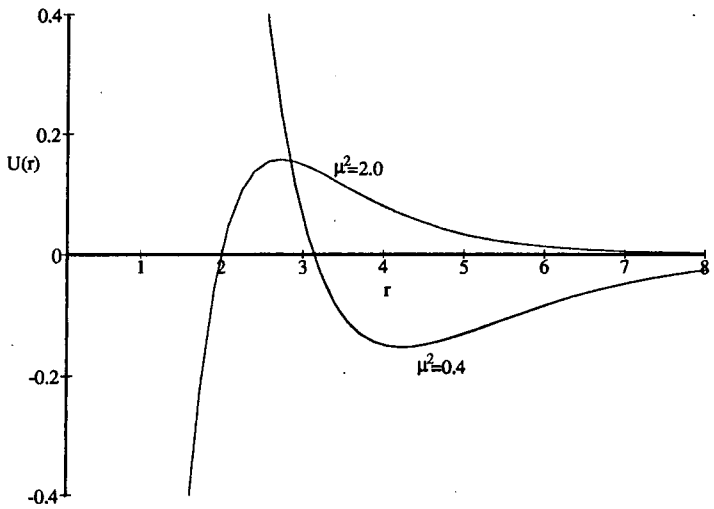

Figure 5.2: The potential function $\mathcal{U}(r)$ for $\mu^2 = 0.4$ and $\mu^2 = 2.0$.

zeros separated by a given distance d, for a range of values of d. In the type I case, they find that the dominance of scalar attraction over magnetic repulsion subsides as the vortex separation gets small, but not to the extent of producing a stable equilibrium at non-zero d. Similarly, they find that magnetic repulsion of type II vortices is increasingly counteracted by scalar attraction (though they do not use this terminology) as d becomes small. It would appear that $\mathcal{U}$ is in broad agreement with their results when $r > r_c$, but that the asymptotic approximation breaks down for $r < r_c$. Of course, this is to be expected – vortices are not point particles, as in our picture, and when they approach one another closely enough their overlap produces significant effects.

5.5 Type II vortex scattering

The interaction potential $\mathcal{U}$ provides a very simple model of two-vortex dynamics: the dynamics of two point particles each of mass M (the energy of a single vortex at rest, a μ^2 dependent quantity) interacting via the potential (5.61). Ignoring the

(trivial) centre of mass motion, the Lagrangian of such a mechanical system is

$$\begin{aligned} L &= \frac{1}{4}M|\dot{\mathbf{r}}|^2 - \mathcal{U}(r) \\ &= \frac{1}{4}M(\dot{r}^2 + r^2\dot{\vartheta}^2) - \mathcal{U}(r), \end{aligned} \tag{5.64}$$

since the reduced mass of the system is $M/2$. This is a manifestly bad model if $\mu = 1$, because it would predict that there is no scattering at all, in conflict with the results of numerical simulations [46] and the geodesic approximation [49]. Away from critical coupling, one might expect the potential $\mathcal{U}$ to dominate over velocity dependent corrections, at least at moderately low speeds, so the above model, although simple, may give a good quantitative account of vortex interactions. We choose to study type II vortices because these provide a simple, clear-cut dynamic problem: vortex scattering. Type I dynamics is slightly more complicated in that vortices can scatter or form bound states depending on the initial conditions. The coupling chosen for the type II numerical simulations of [46] is $\mu^2 = 2$, a choice which we follow for purposes of comparison. In the Lagrangian (5.64), the constants q and m are already known for $\mu^2 = 2$, but the vortex mass M is not. Rather than attempt to calculate M from our numerical solution, we use the careful numerical analysis of Rebbi and Jacobs [48]. Unfortunately, they found M for each of a regular sequence of μ values, rather than μ^2 values, so the $\mu = \sqrt{2}$ value is not quoted. However, a graph of μ against M is very nearly linear, so we use linear interpolation to estimate the $\mu = 1.41421\ldots$ mass from the $\mu = 1.4$ and $\mu = 1.5$ masses given. The result is $M = 1.51230\pi$ (note that this is consistent with the type II topological lower bound $E > \pi$ derived in section 5.1). The potential for this coupling is plotted in figure 5.2.

The problem we seek to solve is elementary and common in mechanics: the motion of a (notional) point particle of mass $M/2$ in the (central) potential $\mathcal{U}(r)$. From the plot of $\mathcal{U}(r)$ we see that all trajectories which do not encroach on the interior region $r < r_c$ (in which the model is not valid) are scattering trajectories, that is as $t \to \pm\infty$,

$r \to \infty$. At some point on the trajectory, the particle achieves its closest approach r_0 to the origin. By time-translation and rotational symmetries, we can, without loss of generality, take this point to lie on the $\vartheta = 0$ ray and occur at time $t = 0$. So, we wish to solve the equations of motion derived from (5.64) subject to initial conditions $r(0) = r_0$, $\vartheta(0) = 0$, $\dot{r}(0) = 0$, $\dot{\vartheta}(0) = \dot{\vartheta}_0$. The conserved energy associated with (5.64) is

$$\mathcal{E} = \frac{1}{2}\dot{r}^2 + \frac{2}{M}\mathcal{U}(r) + \frac{J^2}{2r^2}, \tag{5.65}$$

where the cyclicity of ϑ has been used to eliminate $\dot{\vartheta}$ in favour of the conserved angular momentum $J = r^2\dot{\vartheta}$. Note that $\mathcal{E}$ and J are determined by $(r_0, \dot{\vartheta})$, or, as is more useful in practice, $\mathcal{E}$ and $\dot{\vartheta}$ are determined by (r_0, J): substituting the initial data into (5.65),

$$\mathcal{E}(r_0, J) = \frac{2}{M}\mathcal{U}(r_0) + \frac{J^2}{2r_0^2} \tag{5.66}$$

and $\dot{\vartheta}(r_0, J) = J/r_0^2$ from the definition of J. Equating the right hand sides of (5.65) and (5.66), we can solve for $\dot{r}$ and reduce the time evolution $r(t)$ to quadratures. However, we are more interested in the geometry of scattering (see figure 5.3), that is the deflection angle $\Theta := \pi - 2\vartheta(\infty)$, where $\vartheta(\infty) = \lim_{t\to\infty}\vartheta(t)$. Now,

$$\begin{aligned} \frac{d\vartheta}{dr} &= \frac{\dot{\vartheta}}{\dot{r}} = \frac{J^2}{r^2}\left[\frac{4}{M}\left(\mathcal{U}(r_0) - \mathcal{U}(r)\right) + J^2\left(\frac{1}{r_0^2} - \frac{1}{r^2}\right)\right]^{-\frac{1}{2}} \\ \Rightarrow \vartheta(\infty) &= J^2\int_{r_0}^{\infty}\frac{dr}{r^2}\left[\frac{4}{M}\left(\mathcal{U}(r_0) - \mathcal{U}(r)\right) + J^2\left(\frac{1}{r_0^2} - \frac{1}{r^2}\right)\right]^{-\frac{1}{2}}, \end{aligned} \tag{5.67}$$

so the scattering problem is reduced to evaluating the above integral for all required initial conditions (r_0, J).

The scattering simulations described in [46] were parametrized not by (r_0, J) as is natural for the mechanical analysis above, but by (b, v_∞) where b is the impact parameter and v_∞ is the impact speed (the speed when $t \to -\infty$). (Note that b and v_∞ are asymptotic quantities, defined when the vortices are infinitely remote from one another, and that they provide a convenient parametrization of real scattering

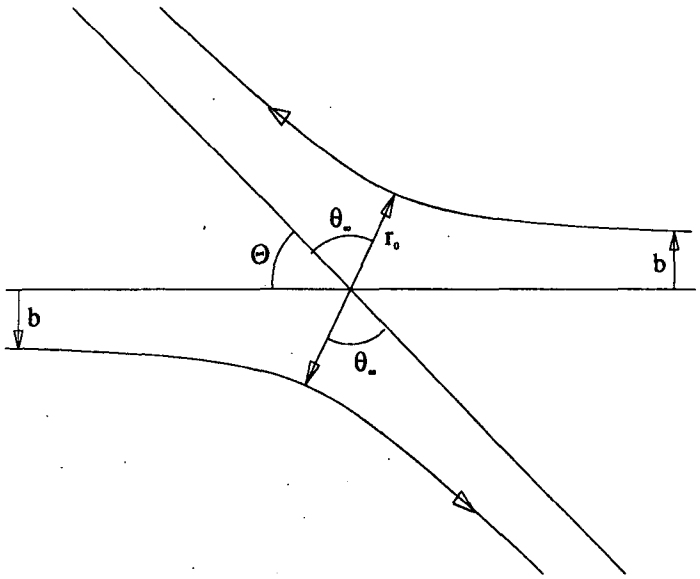

Figure 5.3: The geometry of vortex scattering.

events since they are natural experimental variables.) Four different values of v_∞ were chosen ($v_\infty = 0.1, 0.2, 0.3, 0.4$) and for each v_∞, scattering was simulated at a range of impact parameters in the interval $[0.5, 4]$. We execute the following procedure:

1. Choose a value of v_∞, one of the four above. This determines $\mathcal{E}$ because as $r \to \infty$ the energy is purely kinetic, $\mathcal{E} = v_\infty^2/2$.

2. Find a lower limit, r_{min}, on r_0. This corresponds to head on scattering ($b = 0$) for which $J = 0$, so we find it by solving (5.66) with $J = 0$,

$$\frac{1}{2}v_\infty^2 = \frac{2}{M}\mathcal{U}(r_{min}). \tag{5.68}$$

   In practice this must be done numerically.

3. Find an upper limit, r_{max}, on r_0. This corresponds to scattering with $b = 4$, which, since b is defined such that $J = 2v_\infty b$ implies $J = 8v_\infty$. Thus r_{max} is the

solution of (5.66) with $J = 8v_\infty$,

$$\frac{1}{2}v_\infty^2 = \frac{2}{M}\mathcal{U}(r_{max}) + \frac{32v_\infty^2}{r_{max}^2}. \tag{5.69}$$

Again this must be done numerically.

4. Calculate $\vartheta(\infty)$ for a range of r_0 values in the interval $[r_{min}, r_{max}]$. For each r_0, the corresponding J^2 (giving the correct v_∞) is found from (5.66),

$$J^2(r_0, v_\infty) = 2r_0^2\left(\frac{1}{2}v_\infty^2 - \frac{2}{M}\mathcal{U}(r_0)\right). \tag{5.70}$$

In this way we calculate Θ for fixed v_∞ and a range of values of r_0, this range being chosen to cover the impact parameter interval $[0, 4]$. For each (r_0, v_∞), we also calculate $b = J(r_0, v_\infty)/2v_\infty$. We then plot b against Θ. One might worry that at high speeds and low J, the vortices will penetrate the $r < r_c$ zone and become (unrealistically) captured. In fact the speed required for this is greater than 0.4, so it is not a problem.

Steps 2 and 3 of the above procedure are, as noted, numerically implemented. Not susurprisingly, given the form of $\mathcal{U}(r)$, the integration to evaluate $\vartheta(\infty)$ in step 4 is also necessarily numerical. Note the $(r-r_0)^{-\frac{1}{2}}$ singularity of the integrand of (5.67) at $r = r_0$. This presents no problem in principle because of the integration with respect to r, but must be treated carefully in any numerical algorithm. Schematically, we handle the integral as follows,

$$\begin{aligned}\vartheta(\infty) &= \int_{r_0}^{r_0+\delta} + \int_{r_0+\delta}^{\Delta} + \int_{\Delta}^{\infty} \\ &\approx \vartheta_\delta + \vartheta_{trap} + \vartheta_\Delta,\end{aligned} \tag{5.71}$$

where δ is small ($\delta = 0.1$) and Δ is large ($\Delta = 15$). The contribution ϑ_δ is calculated by Taylor expansion of the integrand about $r = r_0$, while ϑ_{trap} is evaluated using the trapezium rule. At large r the potential falls off exponentially, so for $r > \Delta$ we

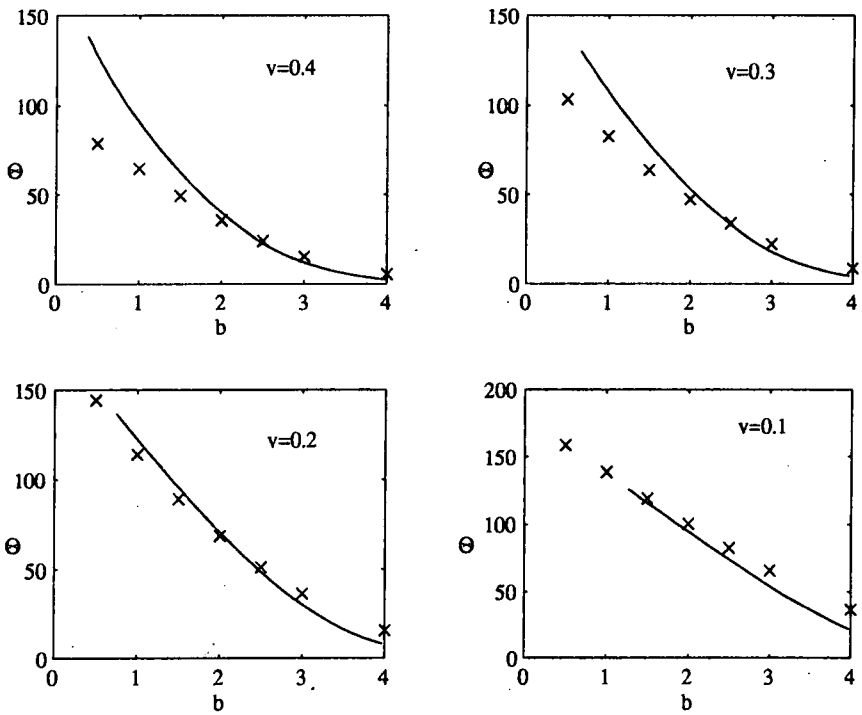

Figure 5.4: The scattering of $\mu^2 = 2$ (type II) vortices: deflection angle Θ versus impact parameter b at four different impact speeds. The solid curves were produced using the point source approximation, the crosses by numerical simulation of the full field equations [46].

set $\mathcal{U} \equiv 0$ and calculate ϑ_Δ in the free vortex approximation. The results of this algorithm are shown in figure 5.4. The fit to the numerical simulations of [46] could be improved by adjusting the values of q and m. Given the warning attached to these charges in section 5.3, this may well be justified. However, such adjustment corrupts the deductive nature of the model, so we prefer not to make it.

5.6 Velocity dependent forces

We next augment the very simple model described above by including velocity dependent corrections. In the case of critical coupling these corrections are leading, because $\mathcal{U} \equiv 0$, so they give the simplest model of critical vortex scattering in the

point source formalism. In the type II regime, the inclusion of velocity dependent corrections should improve on the scattering data of figure 5.4. The calculation is performed *a priori* in the centre of mass frame, so that $\mathbf{z} = -\mathbf{y}$ for all t. The first task is to find expressions for the sources ρ and j for a point vortex moving on some arbitrary trajectory $\mathbf{y}(t)$. We employ a quasi-adiabatic approximation: we assume that at each point $\mathbf{y}(t)$, the source is a static point vortex Lorentz boosted with velocity $\dot{\mathbf{y}}$. (An adiabatic approximation would be to assume that the point vortex always has the static form, and simply translates along the trajectory $\mathbf{y}(t)$.) The strategy is then to expand this moving source in powers of $|\dot{\mathbf{y}}|$, truncating at order $|\dot{\mathbf{y}}|^2$, calculate the time dependent fields it induces, truncating similarly, then use these to find an interaction Lagrangian.

5.6.1 Moving sources

We seek expressions for the scalar charge density ρ and vector current j of a point vortex moving along some curve $\mathbf{y}(t)$ in $\mathbb{R}^2$. At time $t = 0$, let the point vortex be at $\mathbf{x} = \mathbf{0}$, moving with velocity $\mathbf{u}$. Introduce rest frame coordinates η^μ, related to laboratory coordinates x^μ by a Lorentz boost (on x) with velocity $\mathbf{u}$. Explicitly,

$$\begin{aligned} \eta^0 &= \gamma(u)(x^0 - \mathbf{u}\cdot\mathbf{x}) = -\gamma\mathbf{u}\cdot\mathbf{x} \\ \eta^\parallel &= \gamma(u)(x^\parallel - ux^0) = \gamma x^\parallel \\ \eta^\perp &= x^\perp \end{aligned} \tag{5.72}$$

at time $t = 0 = x^0$, where $\gamma(u) = (1-u^2)^{-\frac{1}{2}}$. The spacelike components $\eta^\parallel$ and $\eta^\perp$ can be combined into a single 2-vector equation for $\boldsymbol{\eta}(\mathbf{x})$ by decomposing $\boldsymbol{\eta}$ into parallel and perpendicular components:

$$\boldsymbol{\eta} = \frac{(\boldsymbol{\eta}\cdot\mathbf{u})\mathbf{u}}{u^2} + \left[\boldsymbol{\eta} - \frac{(\boldsymbol{\eta}\cdot\mathbf{u})\mathbf{u}}{u^2}\right]$$

$$
\begin{aligned}
&= \frac{\gamma(\mathbf{x}\cdot\mathbf{u})\mathbf{u}}{u^2} + \left[\mathbf{x} - \frac{(\mathbf{x}\cdot\mathbf{u})\mathbf{u}}{u^2}\right] \\
&= \mathbf{x} + \frac{(\mathbf{x}\cdot\mathbf{u})\mathbf{u}}{u^2}\left[\frac{1}{\sqrt{1-u^2}} - 1\right] \\
&= \mathbf{x} + \frac{1}{2}(\mathbf{x}\cdot\mathbf{u})\mathbf{u} + \ldots
\end{aligned} \tag{5.73}
$$

where the ellipsis denotes discarded terms of order u^4 or greater (we shall not persist in so noting these discarded terms; henceforth an ellipsis will only be included where we have discarded further negligible terms in deriving the given expression). Now generalize to the situation of a point vortex located at $\mathbf{y}(t)$ moving with velocity $\dot{\mathbf{y}}(t)$. The rest frame coordinates of a general point $\mathbf{x}$ on the $\mathbf{x}^0 = t$ time-like surface are

$$
\boldsymbol{\eta}(t,\mathbf{x}) = \mathbf{x} - \mathbf{y} + \frac{1}{2}[(\mathbf{x}-\mathbf{y})\cdot\dot{\mathbf{y}}]\dot{\mathbf{y}} \tag{5.74}
$$

by mapping $\mathbf{x} \mapsto \mathbf{x} - \mathbf{y}(t)$, $\mathbf{u} \mapsto \dot{\mathbf{y}}(t)$ in equation (5.73).

We must now transform the sources, applying Lorentz boosts with velocity $-\dot{\mathbf{y}}$ to the rest frame distributions $\rho_{(0)}$ and $j_{(0)}$. Consider first the scalar distribution, which, from previous work, in the rest frame has the form

$$
\rho_{(0)}(\boldsymbol{\eta}) = q\delta(\boldsymbol{\eta}). \tag{5.75}
$$

Since ρ transforms as a Lorentz scalar, $\rho(x) = \rho_{(0)}(\boldsymbol{\eta}(\mathbf{x}))$ in the lab frame. Now,

$$
\int d^2\mathbf{x}\, f(\mathbf{x})\delta(\boldsymbol{\eta}) = \int d^2\boldsymbol{\eta} \left|\frac{\partial \mathbf{x}}{\partial \boldsymbol{\eta}}\right| f(\mathbf{x})\delta(\boldsymbol{\eta}) \tag{5.76}
$$

where $|\partial\mathbf{x}/\partial\boldsymbol{\eta}|$ is the determinant of the Jacobian of the transformation $\mathbf{x} \mapsto \boldsymbol{\eta}$. From (5.74),

$$
\begin{aligned}
\frac{\partial \boldsymbol{\eta}}{\partial \mathbf{x}} &= \mathbb{I} + \frac{1}{2}\dot{\mathbf{y}} \otimes \dot{\mathbf{y}} \\
\Rightarrow \frac{\partial \mathbf{x}}{\partial \boldsymbol{\eta}} &= \mathbb{I} - \frac{1}{2}\dot{\mathbf{y}} \otimes \dot{\mathbf{y}} + \ldots
\end{aligned}
$$

$$\Rightarrow \left|\frac{\partial \mathbf{x}}{\partial \boldsymbol{\eta}}\right| = 1 - \frac{1}{2}|\dot{\mathbf{y}}|^2 + \ldots \tag{5.77}$$

Substituting into (5.76),

$$\begin{aligned} \int d^2\mathbf{x}\, f(\mathbf{x})\delta(\boldsymbol{\eta}) &= \left(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2\right) \int d^2\boldsymbol{\eta} f(\mathbf{x})\delta(\boldsymbol{\eta}) \\ &= \left(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2\right) f(\mathbf{x})\Big|_{\boldsymbol{\eta}=\mathbf{0}} \\ &= \left(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2\right) f(\mathbf{y}) \end{aligned} \tag{5.78}$$

since $\boldsymbol{\eta} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{y}$. Thus,

$$\delta(\boldsymbol{\eta}) = \left(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2\right) \delta(\mathbf{x} - \mathbf{y}) \tag{5.79}$$

$$\Rightarrow \rho(\mathbf{x}) = \left(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2\right) q\delta(\mathbf{x} - \mathbf{y}). \tag{5.80}$$

The term scalar charge for q is something of a misnomer since, as shown by (5.80) it is not a scalar! A plate of area C in its rest frame, carrying uniform scalar charge density ρ has total scalar charge $q = C\rho$. Looked at from a boosted frame, the plate is squashed along the boost direction by a factor $1/\gamma$ due to Lorentz-Fitzgerald contraction, so the area of the plate in this frame is $C/\gamma \approx (1 - u^2/2)C$. The charge density is invariant, so the total scalar charge is $q' \approx (1 - u^2/2)q$ in agreement with the calculation above.

The moving magnetic dipole is rather more complicated, because j^μ transforms as a vector itself. The rest frame source is

$$j^0_{(0)} = 0, \quad \mathbf{j}_{(0)} = -m\hat{\mathbf{k}} \times \nabla_\eta \delta(\boldsymbol{\eta}). \tag{5.81}$$

Again, we perform a Lorentz boost on this with velocity $-\dot{\mathbf{y}}$, that is

$$j^\mu(x) = \Lambda^\mu_{\ \nu} j^\nu_{(0)}(\eta(x)), \tag{5.82}$$

where

$$\begin{aligned}\Lambda^0{}_0 = \Lambda^{\parallel}{}_{\parallel} &= \gamma \\ \Lambda^0{}_{\parallel} = \Lambda^{\parallel}{}_0 &= \gamma|\dot{\mathbf{y}}| \\ \Lambda^{\perp}{}_{\perp} &= 1, \end{aligned} \tag{5.83}$$

all other $\Lambda^{\mu}{}_{\nu} = 0$. Explicitly,

$$j^0(x) = |\dot{\mathbf{y}}|\gamma j^{\parallel}_{(0)}(\eta(x)) = \dot{\mathbf{y}} \cdot \mathbf{j}_{(0)}(\eta(x)) + \dots \tag{5.84}$$

$$j^{\parallel}(x) = \gamma j^{\parallel}_{(0)}(\eta(x)) = \left(1 + \frac{1}{2}|\dot{\mathbf{y}}|^2\right) j_{(0)}(\eta(x)) + \dots$$

$$j^{\perp}(x) = j^{\perp}_{(0)}(\eta(x)). \tag{5.85}$$

Repeating the algebraic trick of (5.73),

$$\begin{aligned}\mathbf{j}(\mathbf{x}) &= \frac{(\mathbf{j}(x)\cdot\dot{\mathbf{y}})\dot{\mathbf{y}}}{|\dot{\mathbf{y}}|^2} + \left[\mathbf{j}(x) - \frac{(\mathbf{j}(x)\cdot\dot{\mathbf{y}})\dot{\mathbf{y}}}{|\dot{\mathbf{y}}|^2}\right] \\ &= \mathbf{j}_{(0)}(\eta(x)) + \frac{1}{2}(\mathbf{j}_{(0)}(\eta(x))\cdot\dot{\mathbf{y}})\dot{\mathbf{y}} + \dots \end{aligned} \tag{5.86}$$

Now,

$$\begin{aligned}\nabla_\eta &= \left(\frac{\partial \mathbf{x}}{\partial \boldsymbol{\eta}}\right)\nabla \\ &= \left(\mathbb{I} - \frac{1}{2}\dot{\mathbf{y}}\otimes\dot{\mathbf{y}}+\right)\nabla \\ &= \nabla - \frac{1}{2}\dot{\mathbf{y}}(\dot{\mathbf{y}}\cdot\nabla), \end{aligned} \tag{5.87}$$

so, using (5.79), we find that

$$\begin{aligned}\mathbf{j}_{(0)}(\eta(x)) &= -m\hat{\mathbf{k}}\times\nabla_\eta\delta(\boldsymbol{\eta}) \\ &= -m\hat{\mathbf{k}}\times\left[\nabla - \frac{1}{2}\dot{\mathbf{y}}(\dot{\mathbf{y}}\cdot\nabla)\right]\left[\left(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2\right)\delta(\mathbf{x}-\mathbf{y})\right] \end{aligned}$$

$$= -m\left(1-\frac{1}{2}|\dot{\mathbf{y}}|^2\right)\hat{\mathbf{k}}\times\nabla\delta(\mathbf{x}-\mathbf{y}) + \frac{1}{2}m(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\dot{\mathbf{y}}\cdot\nabla\delta(\mathbf{x}-\mathbf{y})+\ldots \tag{5.88}$$

Substituting (5.88) into (5.86) and (5.84),

$$\mathbf{j}(x) = -m\left(1-\frac{1}{2}|\dot{\mathbf{y}}|^2\right)\hat{\mathbf{k}}\times\nabla\delta(\mathbf{x}-\mathbf{y}) + \frac{1}{2}m\left[\dot{\mathbf{y}}(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\cdot\nabla\delta(\mathbf{x}-\mathbf{y}) + (\hat{\mathbf{k}}\times\dot{\mathbf{y}})\dot{\mathbf{y}}\cdot\nabla\delta(\mathbf{x}-\mathbf{y})\right], \tag{5.89}$$

$$j^0(x) = m(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\cdot\nabla\delta(\mathbf{x}-\mathbf{y}), \tag{5.90}$$

the final result. One should note that

$$\partial_\mu j^\mu = \frac{\partial j^0}{\partial t} + \nabla\cdot\mathbf{j} = m(\hat{\mathbf{k}}\times\ddot{\mathbf{y}})\cdot\nabla\delta(\mathbf{x}-\mathbf{y}), \tag{5.91}$$

so that $\partial_\mu j^\mu \neq 0$ unless $\ddot{\mathbf{y}} = \mathbf{0}$ (in which case the rest frame is inertial and $\partial_\mu j^\mu = 0$ follows from the vanishing of $\nabla\cdot\mathbf{j}$ for a static point vortex). If $\ddot{\mathbf{y}} = \mathbf{0}$ the current is conserved, so we can visualize the current density of a vortex moving with constant velocity as a standard electric current. In analogy with ordinary electrodynamics, we identify $j^0 = \varrho$ as the electric charge density of the distribution. From (5.90) we see that $\varrho \neq 0$ for a moving magnetic dipole, but that ϱ corresponds to an electric dipole of moment $-m\hat{\mathbf{k}}\times\dot{\mathbf{y}} = |m|\hat{\mathbf{k}}\times\dot{\mathbf{y}}$. It is helpful to think of the magnetic dipole at rest as a small clockwise current loop consisting of interpervading gases of oppositely charged current carriers confined to a circle in $\mathbb{R}^2$, travelling in opposite senses, as in figure 5.5, so that the whole has $\varrho = 0$ everywhere. Consider such an object viewed in a frame in which it is moving with constant velocity $\dot{\mathbf{y}}$. In the figure, the positive gas in the upper half experiences greater Lorentz-Fitzgerald contraction than the negative due to its velocity of circulation (it is *charge* not charge *density* which is scalar here), and *vice-versa* in the lower half. Thus the upper half acquires a net positive charge density, while the lower half acquires an equal net negative charge density, this charge

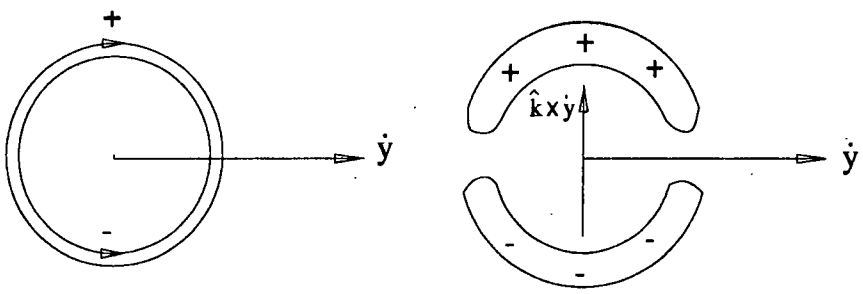

Figure 5.5: A moving current loop. The unit vector $\hat{\mathbf{k}}$ is directed out of the page. The left hand picture shows the senses of circulation of the charge carriers, while the right hand picture depicts the electric charge density as seen from the laboratory frame.

splitting being the origin of our electric dipole ϱ. Note the agreement of orientation. Turning now to the current density $\mathbf{j}$, the first term of (5.89) represents a current loop, while the second and third represent anisotropies due to the aforementioned charge splitting. The transport of the net positive charge (upper half) and net negative charge (lower half) with velocity $\dot{\mathbf{y}}$ produces the current represented by the second term in (5.89). Also, as the loop moves along, the charge must split in front of the loop and recombine behind it in order to create the electric dipole. The current due to this process is represented by the final term.

5.6.2 Interaction Lagrangians

In order to compute the interaction Lagrangians L_ψ and L_A in the case of arbitrarily moving vortices, one must find the fields ψ and A induced by time varying sources ρ and j. Were the linear theory massless, this would involve the use of retarded potentials, since disturbances of the fields due to time varying sources would propagate uniformly at the speed of light. For example, the potential induced by a moving point charge in classical electrodynamics has been well studied, and explicit formulae can be found in the literature [57]. Not surprisingly, the analogous problem in *massive* electrodynamics (or scalar field theory) has not received such attention: the only fundamental physical force transmitted by massive quanta is the weak force, which

has no rôle to play in the classical dynamics of point particles. Motivated by a technique used in the study of pion-nucleon models [58], we handle the problem by introducing formal temporal Fourier transforms, as follows. Let $\psi(t,\mathbf{x})$ be the field induced by time-varying source $\rho(t,\mathbf{x})$, according to the inhomogeneous Klein-Gordon equation. Define Fourier transforms $\tilde{\psi}$ and $\tilde{\rho}$ with variable ω dual to t. That is,

$$\begin{aligned} \psi(t,\mathbf{x}) &=: \int_{-\infty}^{\infty} d\omega e^{i\omega t}\tilde{\psi}(\omega,\mathbf{x}) \\ \rho(t,\mathbf{x}) &=: \int_{-\infty}^{\infty} d\omega e^{i\omega t}\tilde{\rho}(\omega,\mathbf{x}). \end{aligned} \tag{5.92}$$

Then,

$$(\Box+\mu^2)\psi = \rho \tag{5.93}$$

$$\begin{aligned} \Rightarrow (-\omega^2-\Delta+\mu^2)\tilde{\psi} &= \tilde{\rho} \\ \Rightarrow [-\Delta+(\mu^2-\omega^2)]\tilde{\psi} &= \tilde{\rho}, \end{aligned} \tag{5.94}$$

so $\tilde{\psi}(\omega,\mathbf{x})$ satisfies the static inhomogeneous Klein-Gordon equation with squared mass $\mu^2-\omega^2$ and source $\tilde{\rho}(\omega,\mathbf{x})$. Equation (5.94) is solved (at least formally) by convolution of $\tilde{\rho}$ with the Green's function $K_0(\sqrt{\mu^2-\omega^2}|\mathbf{x}-\mathbf{x}'|)/2\pi$,

$$\tilde{\psi}(\omega,\mathbf{x}) = \frac{1}{2\pi}\int d^2\mathbf{x}'\, K_0(\sqrt{\mu^2-\omega^2}|\mathbf{x}-\mathbf{x}'|)\tilde{\rho}(\omega,\mathbf{x}'). \tag{5.95}$$

Now expand the Green's function in ω/μ, truncating at order ω^2/μ^2,

$$\begin{aligned} K_0(\sqrt{\mu^2-\omega^2}|\mathbf{x}-\mathbf{x}'|) &= K_0(\mu(1-\frac{1}{2}\frac{\omega^2}{\mu^2}+\cdots)|\mathbf{x}-\mathbf{x}'|) \\ &= K_0(\mu|\mathbf{x}-\mathbf{x}'|) \\ &\quad -\frac{\omega^2}{2\mu^2}|\mathbf{x}-\mathbf{x}'|K_0'(\mu|\mathbf{x}-\mathbf{x}'|)+\cdots \end{aligned} \tag{5.96}$$

so that substitution into (5.95) yields

$$\tilde{\psi}(\omega,\mathbf{x}) = \frac{1}{2\pi}\int d^2\mathbf{x}'\, K_0(\mu|\mathbf{x}-\mathbf{x}'|)\tilde{\rho}(\omega,\mathbf{x}') \\ -\frac{\omega^2}{4\pi\mu}\int d^2\mathbf{x}'\,|\mathbf{x}-\mathbf{x}'|K_0'(\mu|\mathbf{x}-\mathbf{x}'|)\tilde{\rho}(\omega,\mathbf{x}'), \qquad (5.97)$$

whence we obtain $\psi(t,\mathbf{x})$ by (5.92),

$$\begin{aligned}\psi(t,\mathbf{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, e^{i\omega t}\int d^2\mathbf{x}'\, K_0(\mu|\mathbf{x}-\mathbf{x}'|)\tilde{\rho}(\omega,\mathbf{x}') \\ &\quad -\frac{1}{4\pi\mu}\int_{-\infty}^{\infty} d\omega\,\omega^2 e^{i\omega t}\int d^2\mathbf{x}'\,|\mathbf{x}-\mathbf{x}'|K_0'(\mu|\mathbf{x}-\mathbf{x}'|)\tilde{\rho}(\omega,\mathbf{x}') \\ &= \frac{1}{2\pi}\int d^2\mathbf{x}'\, K_0(\mu|\mathbf{x}-\mathbf{x}'|)\int_{-\infty}^{\infty} d\omega\, e^{i\omega t}\tilde{\rho}(\omega,\mathbf{x}') \\ &\quad +\frac{1}{4\pi\mu}\int d^2\mathbf{x}'\,|\mathbf{x}-\mathbf{x}'|K_0'(\mu|\mathbf{x}-\mathbf{x}'|)\frac{\partial^2}{\partial t^2}\int_{-\infty}^{\infty} d\omega\, e^{i\omega t}\tilde{\rho}(\omega,\mathbf{x}') \\ &= \frac{1}{2\pi}\int d^2\mathbf{x}'\, K_0(\mu|\mathbf{x}-\mathbf{x}'|)\rho(t,\mathbf{x}') \\ &\quad +\frac{1}{4\pi\mu}\int d^2\mathbf{x}'\,|\mathbf{x}-\mathbf{x}'|K_0'(\mu|\mathbf{x}-\mathbf{x}'|)\ddot{\rho}(t,\mathbf{x}'). \qquad (5.98)\end{aligned}$$

Note that truncating the expansion in ω is, in effect, the same as neglecting higher time derivatives of ρ in general, eventually acting on $\mathbf{y}(t)$ in our application. No claim of rigour is attached to the above Fourier transform manoeuvre. One should regard it as a convenient algebraic short-hand for obtaining a perturbative ansatz for (5.93). Substitution of (5.98) into (5.93) explicitly verifies that ψ is indeed a solution, up to higher derivative terms ($d^3\rho/dt^3$ etc.).

Using this procedure, we find ψ_2 , the field induced by time varying source ρ_2. The interaction of such a field with another time varying source ρ_1 is given by (5.56),

$$\begin{aligned}L_\psi &= \int d^2\mathbf{x}\,\rho_1\psi_2 \\ &= \frac{1}{2\pi}\int d^2\mathbf{x}\,d^2\mathbf{x}'\, K_0(\mu|\mathbf{x}-\mathbf{x}'|)\rho_1(t,\mathbf{x})\rho_2(t,\mathbf{x}') \\ &\quad -\frac{1}{4\pi\mu}\int d^2\mathbf{x}\,d^2\mathbf{x}'\,|\mathbf{x}-\mathbf{x}'|K_0'(\mu|\mathbf{x}-\mathbf{x}'|)\dot{\rho}_1(t,\mathbf{x})\dot{\rho}_2(t,\mathbf{x}'), \qquad (5.99)\end{aligned}$$

where a total time derivative has been discarded. The vector field calculation is essentially identical, yielding interaction Lagrangian

$$\begin{aligned} L_A &= -\frac{1}{2\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|) j^\mu_{(1)}(t,\mathbf{x}) J^{(2)}_\mu(t,\mathbf{x}') \\ &\quad +\frac{1}{4\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, |\mathbf{x}-\mathbf{x}'| K_0'(|\mathbf{x}-\mathbf{x}'|) \frac{\partial j^\mu_{(1)}}{\partial t}(t,\mathbf{x}) \frac{\partial J^{(2)}_\mu}{\partial t}(t,\mathbf{x}'), \end{aligned} \tag{5.100}$$

the only new feature being that one of the sources includes the "fictitious current" due to non-conservation of $j^{(2)}_\mu$. That is, the field equation for A_μ with source j_μ (5.39) looks like the Proca equation (a triplet of Klein-Gordon equations) with source $j_\mu + \partial_\mu(\partial_\nu j^\nu)$, so we define the pseudocurrent

$$J^{(2)}_\mu := j^{(2)}_\mu + \partial_\mu(\partial_\nu j^\nu_{(2)}), \tag{5.101}$$

and it is this source which appears in the algebra analogously to ρ_2 in the scalar calculation above. There is an apparent asymmetry in (5.100)– it looks asymmetric under the interchange of sources $1 \leftrightarrow 2$ – but this is easily removed, given the form of the pseudocurrent (5.101), by integration by parts; taking the first integral of (5.100), the extra term due to the fictitious current is (∂'_μ denotes $\partial/\partial x'^\mu$)

$$\begin{aligned} &-\frac{1}{2\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)\, j^\mu_{(1)}(x)\, \partial'_\mu(\partial'_\nu j^\nu_{(2)}(x'))\Big|_{t'=t} \\ &\quad = \frac{1}{2\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, j^\mu_{(1)}\, \partial'_\mu K_0\, \partial'_\nu j^\nu_{(2)}\Big|_{t'=t} \\ &\quad = \frac{1}{2\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, j^\mu_{(1)}\, \partial_\mu K_0\, \partial'_\nu j^\nu_{(2)}\Big|_{t'=t} \\ &\quad = -\frac{1}{2\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)\, \partial_\mu j^\mu_{(1)}(x)\, \partial'_\nu j^\nu_{(2)}(x')\Big|_{t'=t} \end{aligned} \tag{5.102}$$

where, as usual, all boundary integrals have vanished. A similar calculation for the second integral yields the extra term

$$\frac{1}{4\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, |x-\mathbf{x}'| K_0'(|\mathbf{x}-\mathbf{x}'|) \frac{\partial}{\partial t}\partial_\mu j^\mu_{(1)}(x) \frac{\partial}{\partial t'}\partial'_\nu j^\nu_{(2)}(x')\Big|_{t'=t} \tag{5.103}$$

so that $1 \leftrightarrow 2$ symmetry is recovered. Recall from (5.91) that $\partial_\mu j^\mu$ is of order $|\ddot{\mathbf{y}}|$, so neither term (5.102) nor (5.103) makes any contribution to L_A at the order to which we are calculating. We may thus discard them and work with the formula

$$\begin{aligned} L_A &= -\frac{1}{2\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|) j^\mu_{(1)}(t,\mathbf{x}) j^{(2)}_\mu(t,\mathbf{x}') \\ &\quad + \frac{1}{4\pi}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, |\mathbf{x}-\mathbf{x}'| K_0'(|\mathbf{x}-\mathbf{x}'|) \frac{\partial j^\mu_{(1)}}{\partial t}(t,\mathbf{x}) \frac{\partial j^{(2)}_\mu}{\partial t}(t,\mathbf{x}'). \end{aligned} \tag{5.104}$$

It remains to substitute the point vortex sources into (5.99) and (5.104) and evaluate the integrals, a rather lengthy calculation the details of which we present in appendix A. It is convenient to define the function $\Upsilon : \alpha \mapsto \alpha K_0'(\alpha)$. We eventually find that

$$L_\psi = \frac{q^2}{2\pi}\left[K_0(2\mu|\mathbf{y}|) - \frac{3}{2}|\dot{\mathbf{y}}|^2 K_0(2\mu|\mathbf{y}|) - \frac{1}{2}(\dot{\mathbf{y}}\cdot\hat{\mathbf{y}})^2 \Upsilon(2\mu|\mathbf{y}|)\right], \tag{5.105}$$

while

$$\begin{aligned} L_A &= \frac{m^2}{2\pi}\left[-K_0(2|\mathbf{y}|) + |\dot{\mathbf{y}}|^2\left(\frac{2K_0'(2|\mathbf{y}|)}{|\mathbf{y}|} - \frac{3}{2}K_0(2|\mathbf{y}|)\right)\right. \\ &\quad \left. + (\dot{\mathbf{y}}\cdot\hat{\mathbf{y}})^2\left(4K_0(2|\mathbf{y}|) + \frac{1}{2}\Upsilon(2|\mathbf{y}|) - \frac{4K_0'(2|\mathbf{y}|)}{|\mathbf{y}|}\right)\right]. \end{aligned} \tag{5.106}$$

We can now formulate a more refined mechanical model of two-vortex dynamics. As in section 5.4, let $\mathbf{r} := 2\mathbf{y} = r(\cos\vartheta, \sin\vartheta)$ and M be the single vortex mass at coupling μ^2. Also, it is henceforth implicit that a Bessel's function written without an argument is evaluated at r, unless it has a $\star$ superscript, in which case it is evaluated at μr. The Lagrangian of the model is

$$L = \frac{1}{2}\left(\frac{M}{2}\right)|\dot{\mathbf{r}}|^2 + L_\psi + L_A. \tag{5.107}$$

Substituting L_ψ from (5.105) and L_A from (5.106) and collecting terms yields

$$L = \frac{1}{2}\left(\frac{M}{2}\right)(G(r)\dot{r}^2 + r^2 H(r)\dot{\vartheta}^2) - \mathcal{U}(r), \tag{5.108}$$

where

$$\begin{aligned} G(r) &= 1 + \frac{1}{8\pi M}\left[m^2\left(5K_0 - rK_1 + \frac{8K_1}{r}\right) - q^2(3K_0^\star - \mu r K_1^\star)\right] \\ H(r) &= 1 - \frac{1}{8\pi M}\left[m^2\left(3K_0 + \frac{8K_1}{r}\right) + 3q^2 K_0^\star\right] \\ \mathcal{U}(r) &= \frac{1}{2\pi}(m^2 K_0 - q^2 K_0^\star). \end{aligned} \tag{5.109}$$

Note that $\mathcal{U}(r)$ is simply the static intervortex potential already found (5.61). We interpret this system as representing a particle of mass $M/2$ moving on a manifold (call it $\mathcal{M}$) with metric

$$g = G(r)dr^2 + r^2 H(r)d\vartheta^2 \tag{5.110}$$

under the influence of a potential $\mathcal{U}(r)$. Vortices are not classically distinguishable, that is, $-\mathbf{r}$ does not correspond to a configuration of vortices which is physically distinguishable from that corresponding to $\mathbf{r}$, so the points $\mathbf{r}$ and $-\mathbf{r}$ should be identified. Accordingly, $\vartheta \in [0, \pi]$, $\vartheta = 0$ and $\vartheta = \pi$ being identified, and g is a metric on the cone $\mathbb{R}^2/\sim$, where $\sim$ is the equivalence relation $\mathbf{r} \sim \mathbf{r}' \Leftrightarrow \mathbf{r} = \pm\mathbf{r}'$. Let us examine this metric more closely. Both $G(r)$ and $H(r)$ approach 1 exponentially fast as r grows large, so g is asymptotically flat. It is clear that $H(r)$ is a monontonically increasing function approaching 1 from below, and, given the blow-up of K_0 and K_1 at small r, that $H(r) \to -\infty$ as $r \to 0$, so $H(r)$ is negative for $r < r_s$, this critical radius depending on μ^2. The function $H(r)$ must behave in this way, regardless of $q(\mu^2)$ and $m(\mu^2)$, provided m and q never simultaneously vanish, which would be somewhat bizarre. We cannot be so definite about $G(r)$. Using the approximate values for $q(\mu^2)$ and $m(\mu^2)$ found in table 5.1, and a piecewise linear interpolant of the data quoted in [48] for $M(\mu^2)$, we can investigate its properties. It seems that

$G > 1$ for all r, that $G \to \infty$ as $r \to 0$ and that $G(r)$ is a monotonically decreasing function. So the only singular radius is r_s, at which the signature of g flips from Euclidean ($r > r_s$) to Lorentzian ($r < r_s$). This singularity is completely unphysical, and we make no attempt to interpret it (a similar singularity occurs in the BPS monopole calculation [50]). Given approximate (q, m, M) we can numerically solve $H(r_s) = 0$ to find $r_s(\mu^2)$. The results form the rightmost column of table 5.1. Note that $r_s(\mu^2)$ is a monotonically decreasing function, and that $r_s(\mu^2) < r_c(\mu^2)$ for all μ^2 (except $\mu^2 = 1$ for which r_c is not defined, there being no static potential). The singularity is therefore inside the "core region" and beyond the range of validity of the approximation.

The equations of motion derived from (5.108) are

$$\ddot{r}^i + \Gamma^i_{jk}\dot{r}^i\dot{r}^j + \frac{\partial \mathcal{U}}{\partial r_i} = 0 \tag{5.111}$$

where r^i, $i = 1, 2$ are some coordinates on $\mathcal{M}$ (for example $r^1 = r$, $r^2 = \vartheta$, the polar coordinates used above), and Γ is the Levi-Civita connexion derived from g. Consider the case of critical coupling $\mu^2 = 1$. Then $\mathcal{U} \equiv 0$, so if $r^i(t)$ is a solution of (5.111) it follows that

$$\ddot{r}^i = -\Gamma^i_{jk}\dot{r}^i\dot{r}^k, \tag{5.112}$$

and $|\ddot{\mathbf{r}}|$ is of order $|\dot{\mathbf{r}}|^2$ for all t. Differentiating (5.112) with respect to time, one sees that for each integer $n \geq 2$, there exists a set of (position dependent) coefficients $\Omega^{(n)\,i}_{j_1 j_2 \cdots j_n}$ such that

$$\frac{d^n r^i}{dt^n} = \Omega^{(n)\,i}_{j_1 j_2 \cdots j_n}\dot{r}^{j_1}\dot{r}^{j_2}\cdots\dot{r}^{j_n}, \tag{5.113}$$

so that $|d^n\mathbf{r}/dt^n|$ is of order $|\dot{\mathbf{r}}|^n$. This provides *a posteriori* justification for truncating the expansion in time derivatives of (ρ, j), in effect $\mathbf{y}(t)$. That is, although the assumption that higher time derivatives are negligible may turn out to be bad for real vortex dynamics, it is at least self consistent. Unfortunately, this is a property of purely geodesic motion and does not extend to off critical dynamics. Were the linear

theory massless, both $\mathcal{U}$ and g would fall off polynomially at large r, as is found in the case of Skyrmions [54]. Then (5.111) can be used to show self consistency of the expansion, provided one makes the extra assumption that $|\mathbf{r}|$ is always large (this is, after all, a long-range approximation). In our case, both $\mathcal{U}$ and g fall off exponentially, as do all their derivatives, so the situation is less clear. This problem will afflict any attempt to find a mechanical model for soliton dynamics in a massive field theory. The approximation must then be viewed as one assuming smooth motion, rather than low speed. We can still use (5.108) to model type II two-vortex dynamics, but the results may break the assumptions used to derive it. In fact, the Lagrangian (5.108) does lead to a very slight improvement in the scattering data presented in figure 5.4, derived using only the static interaction, but the improvement does not warrant the great complication introduced into the calculation of $\Theta(b, v_\infty)$ by the metric g, so we will not present these results here. Rather, we will concentrate on the case of critical coupling, comparing our results with those obtained by Samols [49] using the geodesic approximation.

5.7 Critical vortex scattering

A static vortex of the model at critical coupling saturates the Bogomol'nyi bound (5.18), so its mass is $M = \pi$. Substituting $\mu = 1$ and $q = m$ into (5.109) produces a great simplification in the metric $g = Gdr^2 + Hr^2d\vartheta^2$:

$$\begin{aligned} G(r) &= 1 + \frac{q^2}{4\pi}\left(K_0 + 4\frac{K_1}{r}\right) \\ H(r) &= 1 - \frac{q^2}{4\pi}\left(3K_0 + 4\frac{K_1}{r}\right). \end{aligned} \tag{5.114}$$

Samols [49] regards physical space as the complex plane $\mathbb{C} \cong \mathbb{R}^2$, and parametrizes the n-vortex moduli space by the n zeros of the Higgs field $z_1, z_2, \ldots, z_n$. He then goes on to prove, using properties of the Bogomol'nyi equations, that these are the natural

complex coordinates in which to describe the induced metric structure on the moduli space, in the sense that the metric is Hermitian in terms of them. The two-vortex moduli space is $\mathbb{C} \times \mathrm{M}$, where $\mathbb{C}$ is the space of centre of mass positions, and M is the space of relative positions (corresponding to $\mathcal{M}$ in our work). Defining relative coordinates (σ, ϑ) such that

$$\sigma e^{i\vartheta} = \frac{1}{2}(z_1 - z_2), \tag{5.115}$$

rotation and parity symmetries are sufficient to restrict the metric on M to the form

$$f_\sigma(\sigma) d\sigma^2 + f_\vartheta(\sigma)\sigma^2 d\vartheta^2, \tag{5.116}$$

while Hermiticity provides the extra restriction that $f_\sigma(\sigma) \equiv f_\vartheta(\sigma) =: F_S^2(\sigma)$, so that

$$g_S = F_S^2(\sigma)(d\sigma^2 + \sigma^2 d\vartheta^2). \tag{5.117}$$

Samols then goes on to compute $F(\sigma)$ numerically, and solve the scattering problem. We would like to identify our radial coordinate r with 2σ, but this would make (5.114) incompatible with (5.117). That is, given that $H \neq G$, our metric cannot be Hermitian in the complex coordinate $re^{i\vartheta}$. It follows that our vortex positions $\pm\mathbf{y}$ do not coincide with zeros of the Higgs field in the nonlinear theory, although coincidence is recovered asymptotically. Defining soliton positions is always somewhat arbitrary because solitons can really only be considered independent particles when they are infinitely remote from one another. For example, one could define vortex positions as local maxima of the potential energy density, which would only asymptotically coincide with zeros of ϕ. With hindsight, then, one cannot demand that the identification of r with 2σ should work for finite r, but the failure to do so is an unwelcome feature not seen in previous applications of the method. It might therefore be connected with the massive nature of the linear theory.

Nonetheless, this failing does not disqualify g from being the asymptotic form of g_S (or rather $4g_S$, given the different normalizations used). We can always *construct*

a radial coordinate s in terms of which g does take the Hermitian form. In comparing the metrics, we wish to discard the Lorentzian part of $\mathcal{M}$, that is the disc $r \leq r_s$, so let g_* be the metric g restricted to $\mathcal{M}_* = (\mathbb{R}^2 \backslash \mathrm{D})/\sim$ where D is the disc of radius $r_d > r_s$, excluding boundary, centred on $\mathbf{r} = \mathbf{0}$. We seek a coordinate transformation on $\mathcal{M}_*$ mapping $r \mapsto s(r)$ and leaving ϑ unchanged, such that there exists some function $F(s)$ satisfying

$$\begin{aligned} g_* = G(r)dr^2 + r^2 H(r)d\vartheta^2 &= 4F^2(s)(ds^2 + s^2 d\vartheta^2) \\ \Rightarrow 2sF(s) = r\sqrt{H(r)}, \quad \text{and} \quad 2F(s)\frac{ds}{dr} &= \sqrt{G(r)} \\ \Rightarrow \frac{d}{dr}[\log s(r)] &= \frac{1}{r}\sqrt{\frac{G(r)}{H(r)}} \qquad (5.118) \\ \Rightarrow \log s(r) - \log s_d &= \int_{r_d}^{r} \frac{d\tilde{r}}{\tilde{r}}\sqrt{\frac{G(\tilde{r})}{H(\tilde{r})}} \\ \Rightarrow s(r) &= s_d \exp\left[\int_{r_d}^{r} \frac{d\tilde{r}}{\tilde{r}}\sqrt{\frac{G(\tilde{r})}{H(\tilde{r})}}\right] \qquad (5.119) \end{aligned}$$

where $s_d := s(r_d)$. Note that if we take $r_d = r_s$ there is a $(\tilde{r} - r_s)^{-\frac{1}{2}}$ singularity in the integrand of (5.119) which, whilst no problem in principle, complicates the numerical evaluation of the integral, so for later convenience we choose to exclude a slightly larger disc D than is strictly necessary. It remains to choose the constant s_d. Since we wish to identify s with Samols' σ, this is fixed by the requirement that $\lim_{r\to\infty} s(r)/r = 1/2$:

$$s_d = \lim_{r\to\infty} \frac{r}{2} \exp\left[-\int_{r_d}^{r} \frac{d\tilde{r}}{\tilde{r}}\sqrt{\frac{G(\tilde{r})}{H(\tilde{r})}}\right]. \qquad (5.120)$$

Given the exponential fall off of G and H, we know that this limit exists, although we cannot calculate it exactly. Since $G(r) > H(r) \quad \forall r$,

$$s_d < \lim_{r\to\infty} \frac{r}{2} \exp\left[-\int_{r_d}^{r} \frac{d\tilde{r}}{\tilde{r}}\right] = \frac{r_d}{2}. \qquad (5.121)$$

So $s(r)/r$ is a function on $[r_d, \infty)$ with value $s_d/r_d < 1/2$ at r_d and tending to $1/2$ as $r \to \infty$. Does it have any turning points? Assume that such a point exists, $r = r_\star$. Then

$$\begin{aligned} \frac{1}{r_\star} s'(r_\star) &= \frac{s(r_\star)}{r_\star^2} \\ \Rightarrow \frac{d}{dr} \log s(r)\Big|_{r=r_\star} &= \frac{1}{r_\star}, \end{aligned} \tag{5.122}$$

but from (5.118) the left hand side is

$$\frac{1}{r_\star}\sqrt{\frac{G(r_\star)}{H(r_\star)}} > \frac{1}{r_\star} \tag{5.123}$$

since $G > H \ \forall r$, so no such $r_\star$ exists. Thus, $s(r)/r$ monotonically increases from minimum $s_d/r_d < 1/2$ to supremum $1/2$. It follows that a given point $(r, \vartheta) \in \mathcal{M}_*$ represents a two vortex configuration with inter-zero distance less than r, the difference vanishing exponentially at large r.

The key numerical task is the evaluation of the function

$$f(r) := \int \frac{d\lambda}{\lambda}\sqrt{\frac{G(\lambda)}{H(\lambda)}} \tag{5.124}$$

in the domain $[r_d, r_\infty]$, where r_∞ is some large value, the effective infinity. Then

$$s_d \approx \frac{r_\infty}{2} e^{-f(r_\infty)}, \tag{5.125}$$

and

$$s(r) = s_d e^{f(r)}. \tag{5.126}$$

We choose $r_d = 2.5$, $r_\infty = 52.5$ and use the trapezium rule, yielding $s_d \approx 0.94$. Having calculated $s(r)$ we can plot it against $F(s(r)) = r\sqrt{H(r)}/2s(r)$ and compare the results with Samols' numerically determined metric function $F_S(\sigma)$, as presented

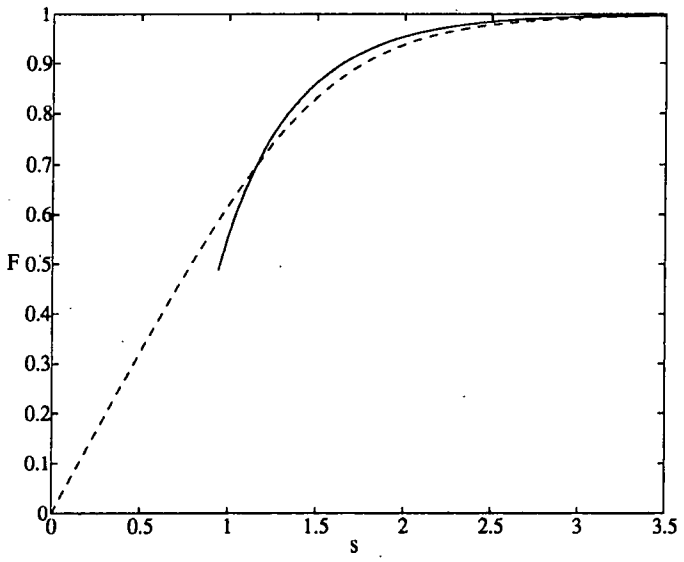

Figure 5.6: The metric profile function $F(s)$ of the metric g_* in Hermitian form (solid curve) compared with Samols' [49] numerically determined function (dashed curve).

in figure 5.6. We see that it is certainly plausible that g_* is the asymptotic form of $4g_S$. Given the metric functions G and H it is straightforward to check that the curvature

$$K = -\frac{1}{\sqrt{GH}}\frac{d}{dr}\left(\frac{1}{\sqrt{G}}\frac{d}{dr}(r\sqrt{H})\right) \tag{5.127}$$

is always positive on $\mathcal{M}_*$, and rapidly vanishes as $r \to \infty$. Thus we can interpret $(\mathcal{M}_*, g_*)$ as a rounded cone with its cap cut off, where the missing cap represents the forbidden core region in which our approximation breaks down.

The main object of this section is to model critical vortex scattering. To do this we must solve the geodesic problem on $(\mathcal{M}_*, g_*)$. We could use g_* in its Hermitian form, but this would introduce an extra layer of numerical approximation (the coodinate transformation $r \mapsto s(r)$), so we choose instead to use g_* in the original coordinates. The scattering problem is defined in terms of asymptotic parameters in any case: impact parameter b and impact speed v_∞ both of which are defined on the part of configuration space in which the term "vortex position" is unambiguous. In fact the geometry of geodesics does not depend on initial velocity, as may be seen by rescaling

the time variable in (5.112), so deflection angle Θ is independent of v_∞. This property is found to hold approximately for low to moderate v_∞ in numerical simulations of vortex scattering, but breaks down at very high speeds. So the scattering data $\Theta(b)$ provide a coordinate independent characterization of the metric structure on $\mathcal{M}_*$.

The kinetic energy associated with g_* is

$$\mathcal{E} = \frac{1}{2}\left(G(r)\dot{r}^2 + \frac{J^2}{r^2 H(r)}\right) \tag{5.128}$$

where, once again, the cyclicity of ϑ has been used to eliminate $\dot{\vartheta}$ in favour of the conserved angular momentum $J = r^2 H(r)\dot{\vartheta}$. Without loss of generality, we can solve the initial value problem $r(0) = r_0$, $\vartheta(0) = 0$, $\dot{r}(0) = 0$, $\dot{\vartheta}(0) = \dot{\vartheta}_0$, essentially parametrized by (r_0, J). Substituting the initial data into (5.128) we see that the pair (r_0, J) fixes $\mathcal{E}$, so that

$$\begin{aligned}
\frac{J^2}{r_0^2 H(r_0)} &= G(r)\dot{r}^2 + \frac{J^2}{r^2 H(r)} \\
\Rightarrow \dot{r} &= J\sqrt{\frac{1}{G(r)}\left(\frac{1}{r_0^2 H(r_0)} - \frac{1}{r^2 H(r)}\right)} \\
\Rightarrow \frac{d\vartheta}{dr} &= \frac{1}{r^2 H(r)}\left[\frac{1}{G(r)}\left(\frac{1}{r_0^2 H(r_0)} - \frac{1}{r^2 H(r)}\right)\right]^{-\frac{1}{2}} \qquad (5.129) \\
\Rightarrow \vartheta(\infty) &= \int_{r_0}^{\infty} \frac{dr}{r^2 H(r)}\left[\frac{1}{G(r)}\left(\frac{1}{r_0^2 H(r_0)} - \frac{1}{r^2 H(r)}\right)\right]^{-\frac{1}{2}}. \qquad (5.130)
\end{aligned}$$

Note that the absence of J in (5.129) implies that the geodesic $\vartheta(r)$ is independent of initial speed, as claimed. It remains to find the connexion between r_0 and b. Recall that $J = 2v_\infty b$, so evaluating $\mathcal{E}$ at r_0 and ∞ and equating expressions gives

$$\begin{aligned}
\frac{2v_\infty^2 b^2}{r_0^2 H(r_0)} &= \frac{1}{2}v_\infty^2 \\
\Rightarrow b(r_0) &= \frac{1}{2}r_0\sqrt{H(r_0)}. \qquad (5.131)
\end{aligned}$$

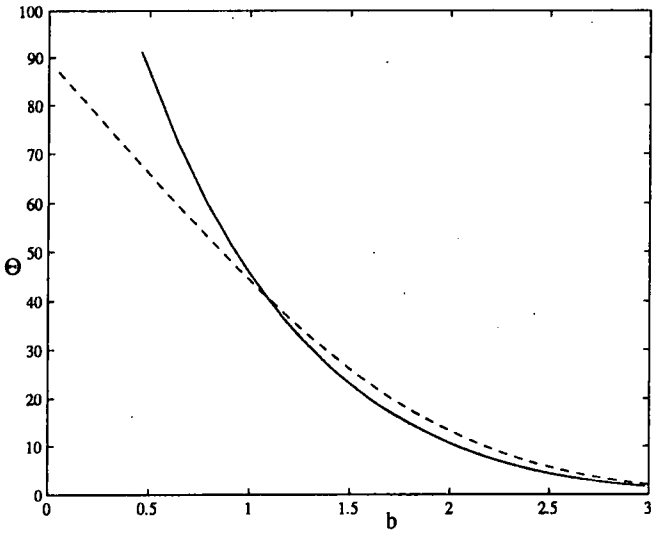

Figure 5.7: Critical vortex scattering: deflection angle Θ versus impact parameter b. The results of the point source approximation (solid curve) should be compared with the results of the numerically implemented geodesic approximation (dashed curve), which is in good agreement with numerical simulations of the full field theory [49].

We evaluate $\vartheta(\infty)$ for a range of r_0 in the interval $[r_d, r_{max}]$, where r_{max} is chosen so that the upper limit of the b range is 3, that is, it is a numerical solution of $r_{max}\sqrt{H(r_{max})} = 6$. Once again, the integrand of (5.130) has a $(r - r_0)^{-\frac{1}{2}}$ singularity. The numerical integration is performed in the same way as described in section 5.5, equation (5.71). Figure 5.7 shows $b(r_0)$ plotted against $\Theta(r_0) = \pi - 2\vartheta(\infty)$, in comparison with Samols' scattering data [49] (see [49] also for a comparison of the geodesic approximation with numerical simulations). The fit is fairly good, and might be improved by adjusting q, a procedure which we eschew on grounds of principle. As one would expect, failure of agreement becomes more marked as b becomes small. Small b collisions probe the small r region of $\mathcal{M}_*$ where the point particle approximation breaks down.

5.8 Conclusion

The essentials of vortex phenomenology in the abelian Higgs model have been known for many years. The qualitative behaviour of vortices depends on a single parameter in the model (μ in our conventions) and splits into three regimes: type I ($\mu < 1$) vortices attract, type II ($\mu > 1$) vortices repel and critical vortices neither attract nor repel. The physical origin of this splitting has also been long known. We may regard μ as the ratio of the Higgs mass to the photon mass (such terminology is, strictly speaking, only appropriate in the quantum, rather than classical, field theory, but its use in this context is widespread) so that the various regimes are seen to arise from the relative ranges of the scalar attraction, transmitted by the Higgs field, and magnetic repulsion, transmitted by the photon field. Using asymptotic properties of the vortex solution, Bettencourt and Rivers [53] found an analytic expression for the intervortex potential at long range. In the present work, this result was rederived from a different viewpoint, a viewpoint which provides a useful physical picture of static vortex interactions. From afar, the vortex looks like a classical point particle carrying both scalar charge q and magnetic dipole moment m in a linear field theory consisting of a real Klein-Gordon field of mass μ and a free vector field of unit mass. The constants q and m were determined numerically for a range of μ^2 values, and it was found that an assumption made about the analogous constants in [53] is ill justified. As an application, the potential was used to formulate a very simple model of type II vortex scattering, the predictions of which were found to be in reasonable agreement with numerical simulations [46], although the model does not do so well as Shah's perturbed moduli space approximation [42].

Velocity dependent corrections to the static interaction were then found by adapting the method of linear retarded potentials originally developed by Manton for the study of BPS monopoles [50]. The difficulty of calculating the induced massive fields due to time varying sources was overcome by means of a formal Fourier transform manoeuvre. The resulting corrections were found to add little of value to the very

simple model of type II scattering already considered, but provided a good quantitative account of critical vortex scattering at modest to large impact parameters. The $\mu = 1$ velocity dependent interaction Lagrangian was reinterpreted in terms of the geodesic problem on a certain manifold $\mathcal{M}$ with metric g, and compared with Samols' metric on the two-vortex moduli space [49]. Here an unfortunate feature was encountered: the natural coordinates for our point source approximation coincide only asymptotically with zeros of the Higgs field, the natural coordinates for the geodesic approximation. This problem has not afflicted previous applications of the method, and so the masses of the linear theory may be responsible. An explicit coordinate transformation was constructed to reduce g to Samols' form, whereupon good asymptotic agreement between the two metrics was found. It is worth pointing out that, compared with other studies of vortex dynamics [46, 49, 42], our calculation required only very lightweight numerical work. The most sophisticated technique was a fourth order Runge-Kutta method to solve an ordinary differential equation, albeit used in a two-parameter shooting algorithm. In fact, if we restrict ourselves to the case of critical coupling, we need only use a one-parameter shooting method. The only other numerical work consisted of the approximate evaluation of integrals on the half line using the trapezium rule. Of course, the price paid for this simplicity was a lengthy and rather messy calculation in section 5.6.2.

A recent development in the study of low energy soliton dynamics is the generation of moduli spaces by computing instanton holonomies [59]. In this approach one explicitly constructs the field configurations from instantons of some Euclidean theory in a space of dimension greater than that of the original soliton theory. The moduli space so generated inherits a potential function (the restriction of the field theoretic potential functional) and a metric induced by the kinetic energy functional. Generically, this metric must be evaluated numerically from first principles, that is, by calculating the $\mathcal{L}^2$ inner products of every (unordered) pair of tangent vectors, at each configuration. This is a very intensive procedure. The computational cost would

be significantly reduced if the need for such numerical work could be contained within a relatively small core region of moduli space where the solitons are close together – if, for example, the asymptotic form of the metric could be found analytically from a point source approximation, a technique employed in [60] in the context of the Skyrme model without pion mass. The present work develops in a simple setting a possible way of doing this when the linearized theory is massive. However, one should note that the unfortunate mismatch of moduli space coordinates encountered in the abelian Higgs model could cause major problems if it occurs generically. We were able to construct a coordinate transformation quite easily, but this was on a two dimensional manifold with rotational symmetry. For higher dimensional moduli spaces (if the field theory is defined on $\mathbb{R}^{3+1}$, or the solitons have orientations or internal degrees of freedom for example) the transformation may be far more complicated. Even so, it would be interesting to find such asymptotic metrics, if only for the aesthetic reason that a formula is always preferable to a collection of numerical data.

Appendix A

Derivation of the velocity dependent interaction Lagrangians

In this appendix we present the detailed derivation of the velocity dependent interaction Lagrangians L_ψ and L_A quoted in section 5.6.2. Recall that we defined the function $\Upsilon : \alpha \mapsto \alpha K_0'(\alpha)$. Beginning with the scalar interaction,

$$L_\psi = \frac{1}{2\pi} S_1 - \frac{1}{4\pi\mu^2} S_2, \tag{A.1}$$

where

$$\begin{aligned} S_1 &= \int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(\mu|\mathbf{x}-\mathbf{x}'|)\rho_1(t,\mathbf{x})\rho_2(t,\mathbf{x}') \\ S_2 &= \int d^2\mathbf{x}\, d^2\mathbf{x}'\, \Upsilon(\mu|\mathbf{x}-\mathbf{x}'|)\dot{\rho}_1(t,\mathbf{x})\dot{\rho}_2(t,\mathbf{x}'). \end{aligned} \tag{A.2}$$

Recall that source 1 is at $\mathbf{y}(t)$ while source 2 is at $-\mathbf{y}(t)$, so from (5.80),

$$\begin{aligned} \rho_1(t,\mathbf{x}) &= q\left(1-\frac{1}{2}|\dot{\mathbf{y}}|^2\right)\delta(\mathbf{x}-\mathbf{y}(t)) \\ \rho_2(t,\mathbf{x}) &= q\left(1-\frac{1}{2}|\dot{\mathbf{y}}|^2\right)\delta(\mathbf{x}+\mathbf{y}(t)). \end{aligned} \tag{A.3}$$

Thus,

$$\begin{aligned} S_1 &= q^2(1_{|}\dot{\mathbf{y}}|^2)\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(\mu|\mathbf{x}-\mathbf{x}'|)\delta(\mathbf{x}-\mathbf{y})\delta(\mathbf{x}'+\mathbf{y})+\cdots \\ &= q^2(1-|\dot{\mathbf{y}}|^2)K_0(2\mu|\mathbf{y}|), \end{aligned} \tag{A.4}$$

while

$$\begin{aligned} S_2 &= q^2\int d^2\mathbf{x}\, d^2\mathbf{x}'\, \Upsilon(\mu|\mathbf{x}-\mathbf{x}'|)(-\dot{\mathbf{y}}\cdot\nabla\delta(\mathbf{x}-\mathbf{y}))(\dot{\mathbf{y}}\cdot\nabla'\delta(\mathbf{x}'+\mathbf{y}))+\cdots \\ &= q^2\int d^2\mathbf{x}\,\dot{\mathbf{y}}\cdot\nabla\delta(\mathbf{x}-\mathbf{y})\int d^2\mathbf{x}'\,\dot{\mathbf{y}}\cdot\nabla'\Upsilon(\mu|\mathbf{x}-\mathbf{x}'|)\,\delta(\mathbf{x}'+\mathbf{y}) \\ &= -q^2\int d^2\mathbf{x}\,\dot{\mathbf{y}}\cot\nabla\delta(\mathbf{x}-\mathbf{y})\,\dot{\mathbf{y}}\cdot\nabla_y\Upsilon(\mu|\mathbf{x}+\mathbf{y}|) \\ &= q^2\int d^2\mathbf{x}\,\delta(\mathbf{x}-\mathbf{y})\,\dot{\mathbf{y}}\cdot\nabla(\dot{\mathbf{y}}\cdot\nabla_y\Upsilon(\mu|\mathbf{x}+\mathbf{y}|)) \\ &= \frac{q^2}{4}(\dot{\mathbf{y}}\cdot\nabla_y)^2\Upsilon(2\mu|\mathbf{y}|). \end{aligned} \tag{A.5}$$

Let $\mathbf{w} := 2\mu\mathbf{y}$. Then

$$\begin{aligned} S_2 &= \frac{q^2}{4}(\dot{\mathbf{w}}\cdot\nabla_w)^2(wK_0'(w)) \\ &= \frac{q^2}{4}\dot{\mathbf{w}}\cdot\nabla_w[\dot{\mathbf{w}}\cdot(\hat{\mathbf{w}}K_0'(w)+\mathbf{w}K_0"(w)] \\ &= \frac{q^2}{4}\dot{\mathbf{w}}\cdot\nabla_w(\dot{\mathbf{w}}\cdot\mathbf{w}\,K_0(w)) \\ &= \frac{q^2}{4}\dot{\mathbf{w}}\cdot(\dot{\mathbf{w}}K_0(w)+\dot{\mathbf{w}}\cdot\mathbf{w}\,K_0'(w)\hat{\mathbf{w}}) \\ &= \mu^2q^2\left[|\dot{\mathbf{y}}|^2K_0(2\mu|\mathbf{y}|)+(\dot{\mathbf{y}}\cdot\hat{\mathbf{y}})^2\Upsilon(2\mu|\mathbf{y}|)\right]. \end{aligned} \tag{A.6}$$

Substituting into (A.1),

$$L_\psi = \frac{q^2}{2\pi}\left[K_0(2\mu|\mathbf{y}|)-\frac{3}{2}|\dot{\mathbf{y}}|^2K_0(2\mu|\mathbf{y}|)-\frac{1}{2}(\dot{\mathbf{y}}\cdot\hat{\mathbf{y}})^2\Upsilon(2\mu|\mathbf{y}|)\right], \tag{A.7}$$

as quoted.

The magnetic interaction Lagrangian, L_A, while considerably more complicated

is evaluated in similar fashion. So,

$$L_A = -\frac{1}{2\pi}(S_3 - S_4) + \frac{1}{4\pi}S_5, \tag{A.8}$$

where

$$\begin{aligned} S_3 &= \int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|) j^0_{(1)}(t,\mathbf{x}) j^0_{(2)}(t,\mathbf{x}') \\ S_4 &= \int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|) \mathbf{j}_{(1)}(t,\mathbf{x}) \cdot \mathbf{j}_{(2)}(t,\mathbf{x}') \\ S_5 &= \int d^2\mathbf{x}\, d^2\mathbf{x}'\, \Upsilon(|\mathbf{x}-\mathbf{x}'|) \frac{\partial j^\mu_{(1)}}{\partial t}(t,\mathbf{x}) \frac{\partial j^{(2)}_\mu}{\partial t}(t,\mathbf{x}'). \end{aligned} \tag{A.9}$$

Source 1 is given explicitly in (5.89), (5.90), while source 2 is obtained from this by mapping $\mathbf{y} \mapsto -\mathbf{y}$ ($\Rightarrow \dot{\mathbf{y}} \mapsto -\dot{\mathbf{y}}$). Substituting these expressions into S_3 and S_4,

$$\begin{aligned} S_3 &= -m^2 \int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)(\hat{\mathbf{k}} \times \dot{\mathbf{y}}) \cdot \nabla \delta(\mathbf{x}-\mathbf{y})\, (\hat{\mathbf{k}} \times \dot{\mathbf{y}}) \cdot \nabla' \delta(\mathbf{x}'+\mathbf{y}) \\ &= \frac{m^2}{4}[(\hat{\mathbf{k}} \times \dot{\mathbf{y}}) \cdot \nabla_y]^2 K_0(2|\mathbf{y}|). \end{aligned} \tag{A.10}$$

$$\begin{aligned} S_4 &= m^2 \int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|). \\ &\quad \Bigg[\overbrace{(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2)\hat{\mathbf{k}} \times \nabla \delta(\mathbf{x}-\mathbf{y})}^{1} - \overbrace{\frac{1}{2}(\hat{\mathbf{k}} \times \dot{\mathbf{y}})\, \dot{\mathbf{y}} \cdot \nabla \delta(\mathbf{x}-\mathbf{y})}^{2} \\ &\qquad - \overbrace{\frac{1}{2}\dot{\mathbf{y}}\, (\hat{\mathbf{k}} \times \dot{\mathbf{y}}) \cdot \nabla \delta(\mathbf{x}-\mathbf{y})}^{3} \Bigg] \\ &\quad \bullet \Bigg[\underbrace{(1 - \frac{1}{2}|\dot{\mathbf{y}}|^2)\hat{\mathbf{k}} \times \nabla' \delta(\mathbf{x}'+\mathbf{y})}_{1} - \underbrace{\frac{1}{2}(\hat{\mathbf{k}} \times \dot{\mathbf{y}})\, \dot{\mathbf{y}} \cdot \nabla' \delta(\mathbf{x}'+\mathbf{y})}_{2} \\ &\qquad - \underbrace{\frac{1}{2}\dot{\mathbf{y}}\, (\hat{\mathbf{k}} \times \dot{\mathbf{y}}) \cdot \nabla' \delta(\mathbf{x}'+\mathbf{y})}_{3} \Bigg] \\ &= m^2 \sum_{i=1}^{3} \sum_{j=1}^{3} S_{ij}, \end{aligned} \tag{A.11}$$

where

$$
\begin{aligned}
S_{11} &= \int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)\left(1-\frac{1}{2}|\dot{\mathbf{y}}|^2\right)^2 (\hat{\mathbf{k}}\times\nabla\delta(\mathbf{x}-\mathbf{y}))\cdot(\hat{\mathbf{k}}\times\nabla'\delta(\mathbf{x}'+\mathbf{y})) \\
&= (1-|\dot{\mathbf{y}}|^2)\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)\nabla\delta(\mathbf{x}-\mathbf{y})\cdot\nabla'\delta(\mathbf{x}'+\mathbf{y})+\cdots \\
&= -\frac{1}{4}(1-|\dot{\mathbf{y}}|^2)\nabla_y^2 K_0(2|\mathbf{y}|), \qquad \text{(A.12)} \\
S_{12} &= -\frac{1}{2}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)(\hat{\mathbf{k}}\times\nabla\delta(\mathbf{x}-\mathbf{y}))\cdot(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\,\dot{\mathbf{y}}\cdot\nabla'\delta(\mathbf{x}'+\mathbf{y})+\cdots \\
&= -\frac{1}{2}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)\,\dot{\mathbf{y}}\cdot\nabla\delta(\mathbf{x}-\mathbf{y})\,\dot{\mathbf{y}}\cdot\nabla'\delta(\mathbf{x}'+\mathbf{y}) \\
&= \frac{1}{8}(\dot{\mathbf{y}}\cdot\nabla_y)^2 K_0(2|\mathbf{y}|), \qquad \text{(A.13)} \\
S_{13} &= -\frac{1}{2}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)(\hat{\mathbf{k}}\times\nabla\delta(\mathbf{x}-\mathbf{y}))\cdot\dot{\mathbf{y}}\,(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\cdot\nabla'\delta(\mathbf{x}'+\mathbf{y})+\cdots \\
&= \frac{1}{2}\int d^2\mathbf{x}\, d^2\mathbf{x}'\, K_0(|\mathbf{x}-\mathbf{x}'|)(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\cdot\nabla\delta(\mathbf{x}-\mathbf{y})\,(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\cdot\nabla'\delta(\mathbf{x}'+\mathbf{y}) \\
&= -\frac{1}{8}[(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\cdot\nabla_y]^2 K_0(2|\mathbf{y}|), \qquad \text{(A.14)}
\end{aligned}
$$

S_{ij} with $i\neq 1$, $j\neq 1$ is of negligible order, and, by construction,

$$S_{ji} = S_{ij}|_{\mathbf{y}\to-\mathbf{y}},$$

which, given the above expressions, implies $S_{12}=S_{21}$, $S_{13}=S_{31}$. Thus,

$$
\begin{aligned}
S_4 &= m^2(S_{11}+2S_{12}+2S_{13}+\cdots) \\
&= \frac{m^2}{4}\left[-(1-|\dot{\mathbf{y}}|^2)\nabla_y^2 K_0(2|\mathbf{y}|)+(\dot{\mathbf{y}}\cdot\nabla_y)K_0(2|\mathbf{y}|)\right. \\
&\qquad \left.-[(\hat{\mathbf{k}}\times\dot{\mathbf{y}})\cdot\nabla_y]^2 K_0(2|\mathbf{y}|)\right]. \qquad \text{(A.15)}
\end{aligned}
$$

Now consider the integral S_5. Recall that $j^0_{(1)} = \dot{\mathbf{y}}\cdot\mathbf{j}_{(1)}+\cdots$, so,

$$
\begin{aligned}
S_5 &= -\int d^2\mathbf{x}\, d^2\mathbf{x}'\,\Upsilon(|\mathbf{x}-\mathbf{x}'|)\frac{\partial\mathbf{j}_{(1)}}{\partial t}(t,\mathbf{x})\cdot\frac{\partial\mathbf{j}_{(2)}}{\partial t}(t,\mathbf{x}')+\cdots \\
&= -m^2\int d^2\mathbf{x}\, d^2\mathbf{x}'\,\Upsilon(|\mathbf{x}-\mathbf{x}'|)\cdot \\
&\qquad (\hat{\mathbf{k}}\times\nabla(\dot{\mathbf{y}}\cdot\nabla_y\delta(\mathbf{x}-\mathbf{y})))\cdot(\hat{\mathbf{k}}\times\nabla'(\dot{\mathbf{y}}\cdot\nabla_y\delta(\mathbf{x}'+\mathbf{y})))
\end{aligned}
$$

$$= \frac{m^2}{16}(\dot{\mathbf{y}} \cdot \nabla_y)^2 \nabla_y^2 \Upsilon(2|\mathbf{y}|). \tag{A.16}$$

After evaluating the differentials in (A.10), (A.15) and (A.16) – an exercise in manipulating Bessel's functions – and substituting S_3, S_4 and S_5 into (A.8), we finally find that

$$\begin{aligned} L_A &= \frac{m^2}{2\pi}\left[-K_0(2|\mathbf{y}|) + |\dot{\mathbf{y}}|^2\left(\frac{2K_0'(2|\mathbf{y}|)}{|\mathbf{y}|} - \frac{3}{2}K_0(2|\mathbf{y}|)\right)\right. \\ &\qquad \left. +(\dot{\mathbf{y}} \cdot \hat{\mathbf{y}})^2\left(4K_0(2|\mathbf{y}|) + \frac{1}{2}\Upsilon(2|\mathbf{y}|) - \frac{4K_0'(2|\mathbf{y}|)}{|\mathbf{y}|}\right)\right]. \end{aligned} \tag{A.17}$$

Bibliography

[1] J.M. Speight and R.S. Ward, "Kink dynamics in a novel discrete sine-Gordon system" *Nonlinearity* **7** (1994) 475.

[2] J.M. Speight, "Kink Casimir energy in a lattice sine-Gordon model" *Phys. Rev.* **D49** (1994) 6914.

[3] J.M. Speight, "Low-energy dynamics of a $\mathbb{C}P^1$ lump on the sphere" *J. Math. Phys.* **36** (1995) 796.

[4] P.W. Higgs, "Spontaneous symmetry breakdown without massless bosons" *Phys. Rev.* **145** (1966) 1156.

[5] T.W.B. Kibble, "Topology of cosmic domains and strings" *J. Phys.* **A9** (1976) 1387.

[6] T.H.R. Skyrme, "A unified field theory of mesons and baryons" *Nucl. Phys.* **31** (1962) 556.

[7] E. Witten, "Current algebra, baryons and quark confinement" *Nucl. Phys.* **B223** (1983) 433.

[8] J.K. Perring and T.H.R. Skyrme, "A model unified field equation" *Nucl. Phys.* **31** (1962) 550.

[9] M. Nakahara, *Geometry, Topology and Physics* (Adam Hilger, Bristol, England, 1990).

[10] N.S. Manton, "A remark on the scattering of BPS monopoles" *Phys. Lett.* **110B** (1982) 54.

[11] P.G. Drazin and R.S. Johnson, *Solitons: an introduction* (Cambridge University Press, Cambridge, England, 1989).

[12] R. Rajaraman, *Solitons and Instantons* (North Holland, Amsterdam, The Netherlands, 1989).

[13] F. Mandl and G. Shaw, *Quantum Field Theory* (John Wiley and Sons, Chichester, England, 1984).

[14] E.B. Bogomol'nyi, "The stability of classical solutions" *Sov. J. Nucl. Phys.* **24** (1976) 449.

[15] J.F. Currie, S.E. Trullinger, A.R. Bishop and J.A. Krumhansl, "Numerical simulation of sine-Gordon soliton dynamics in the presence of perturbations" *Phys. Rev. B* **15** (1977) 5567.

[16] Y. Ishimori and T. Munakata, "Kink dynamics in the discrete sine-Gordon system—A perturbational approach" *J. Phys. Soc. Jpn.* **51** (1982) 3367.

[17] J.A. Combs and S. Yip, "Single-kink dynamics in a one-dimensional atomic chain: A nonlinear atomistic theory and numerical simulation" *Phys. Rev. B* **28** (1983) 6873.

[18] M. Peyrard and M.D. Kruskal, "Kink dynamics in the highly discrete sine-Gordon system" *Physica* **14D** (1984) 88.

[19] R. Boesch, C.R. Willis and M. El-Batanouny, "Spontaneous emission of radiation from a discrete sine-Gordon kink" *Phys. Rev. B* **40** (1989) 2284.

[20] W.J. Zakrzewski, "A modified discrete sine-Gordon model" to appear in *Nonlinearity.*

[21] N.S. Manton, "Unstable manifolds and soliton dynamics" *Phys. Rev. Lett.* **60** (1982) 1916.

[22] R. Boesch and M. Peyrard, "Discreteness effects on a sine-Gordon breather" *Phys. Rev. B* **43** (1991) 8491.

[23] R.S. MacKay and S. Aubry, "Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators" *Nonlinearity* **7** (1994) 1623.

[24] R.S. Ward, "Stable topological Skyrmions on the 2D lattice" to appear in *Lett. Math. Phys.*

[25] G. Plunien, B. Müller and W Greiner, "The Casimir Effect" *Phys. Rep.* **134** (1986) 87.

[26] R. Bellman, *Introduction to Matrix Analysis* (McGraw Hill, New York, USA, 1974).

[27] W. Ledermann and G.E. Reuter, "Spectral Theory for Differential Equations of Simple Birth and Death Processes" *Phil. Trans. Royal Soc. London* **A246** (1953-1954) 321.

[28] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Oxford University Press, Oxford, England, 1965).

[29] R.F. Dashen, B. Hasslacher and A. Neveu, "Particle Spectrum in Model Field Theories from Semiclassical Functional Integral Techniques" *Phys. Rev.* **D11** (1975) 3424.

[30] G.W. Gibbons and N.S. Manton, "Classical and Quantum Dynamics of BPS Monopoles" *Nucl. Phys.* **B274** (1986) 183.

[31] A.M. Din and W.J. Zakrzewski, "General classical solutions in the $\mathbb{C}P^{N-1}$ model" *Nucl. Phys.* **B174** (1980) 397.

[32] R.S. Ward, "Slowly moving lumps in the $\mathbb{C}P^1$ model in $(2+1)$ dimensions" *Phys. Lett.* **158B** (1985) 424.

[33] R.A. Leese, "Low-energy scattering of solitons in the $\mathbb{C}P^1$ model" *Nucl. Phys.* **B344** (1990) 33.

[34] R.A. Leese, M. Peyrard and W.J. Zakrzewski, "Soliton stability in the $O(3)$ σ-model in $(2+1)$ dimensions" *Nonlinearity* **3** (1990) 387.

[35] G. Woo, "Pseudoparticle configurations in two-dimensional ferromagnets" *J. Math. Phys.* **18** (1977) 1264.

[36] A.A. Belavin and A.M. Polyakov, "Metastable states of two-dimensional isotropic ferromagnets" *JETP Lett.* **22** (1975) 245.

[37] H. Schwerdtfeger, *Geometry of Complex Numbers* (Dover, New York, USA, 1979).

[38] R. Penrose and W. Rindler, *Spinors and space-time* Volume 1 (Cambridge University Press, Cambridge, England, 1984).

[39] M.P. Ryan and L.C. Shepley, *Homogeneous Relativistic Cosmologies* (Princeton University Press, Princeton, USA, 1975).

[40] Y. Choquet-Bruhat, C. DeWitt-Morette and M. Dillard-Bleick, *Analysis, Manifolds and Physics* Part I (North-Holland, Amsterdam, The Netherlands, 1982).

[41] B.M.A.G. Piette, B.J. Schroers and W.J. Zakrzewski, "Multisolitons in a two-dimensional Skyrme model" *Zeit. Phys.* **C65** (1995) 165.

[42] P.A. Shah, "Vortex scattering at near-critical coupling" *Nucl. Phys.* **B429** (1994) 259.

[43] D. Stuart, "Dynamics of abelian Higgs vortices in the near Bogomolny regime" *Comm. Math. Phys.* **159** (1994) 51.

[44] H.B. Nielsen and P. Olesen, "Vortex-line models for dual strings" *Nucl. Phys.* **B61** (1973) 45.

[45] E.B. Bogomol'nyi and A.I. Vainshtein, "Stability of strings in abelian gauge theory" *Sov. J. Nucl. Phys.* **23** (1976) 588.

[46] E. Myers, C. Rebbi and R. Strilka, "Study of the interaction and scattering of vortices in the abelian Higgs (or Ginzburg-Landau) model" *Phys. Rev.* **D45** (1992) 1355.

[47] A. Jaffe and C. Taubes, *Vortices and Monopoles* (Birkhäuser, Boston, USA, 1980).

[48] L. Jacobs and C. Rebbi, "Interaction of superconducting vortices" *Phys. Rev.* **B19** (1979) 4486.

[49] T.M. Samols, "Vortex scattering" *Commun. Math. Phys.* **145** (1992) 149.

[50] N.S. Manton, "Monopole interactions at long range" *Phys. Lett.* **154B** (1985) 397 and *Phys. Lett.* **157B** (1985) 475 (errata).

[51] M.F. Atiyah and N.J. Hitchin, "Low energy scattering of non-abelian magnetic monopoles" *Phil. Trans. R. Soc. Lond.* **A315** (1985) 459.

[52] M.F. Atiyah and N.J. Hitchin, *The Geometry and Dynamics of Magnetic Monopoles* (Princeton University Press, Princeton, USA, 1988).

[53] L.M.A. Bettencourt and R.J. Rivers, "Interactions between $U(1)$ cosmic strings: an analytical study" *Phys. Rev.* **D51** (1995) 1842.

[54] B.J. Schroers, "Dynamics of moving and spinning Skyrmions" *Z. Phys.* **C61** (1994) 479.

[55] G.W. Gibbons and P.J. Ruback, "Motion of extreme Reissner-Nordstrom black holes in the low-velocity limit" *Phys. Rev. Lett.* **57** (1986) 1492.

[56] M. Abramowitz and I.A. Stegun (eds), *Pocketbook of Mathematical Functions* (Verlag Harri Deutsch, Frankfurt, Germany 1984).

[57] J.D. Jackson, *Classical Electrodynamics* (John Wiley and Sons, New York, USA, 1975).

[58] E.M. Henley and W. Thirring, *Elementary Quantum Field Theory* (McGraw-Hill, New York, USA, 1962).

[59] M.F. Atiyah and N.S. Manton, "Geometry and kinematics of two Skyrmions" *Commun. Math. Phys.* **152** (1993) 391.

[60] R.A. Leese, N.S. Manton and B.J. Schroers, "Attractive channel Skyrmions and the deuteron" *Nucl. Phys.* **B442** (1995) 228.