

GEOMETRY AND ANALYSIS IN NONLINEAR SIGMA MODELS

D. AUCKLY, L. KAPITANSKI, AND J. M. SPEIGHT

ABSTRACT. The configuration space of a nonlinear sigma model is the space of maps from one manifold to another. This paper reviews the authors' work on nonlinear sigma models with target a homogeneous space. It begins with a description of the components, fundamental group, and cohomology of such configuration spaces, together with the physical interpretations of these results. The topological arguments given generalize to Sobolev maps. The advantages of representing homogeneous-space-valued maps by flat connections are described, with applications to the homotopy theory of Sobolev maps, and minimization problems for the Skyrme and Faddeev functionals. The paper concludes with some speculation about the possibility of using these techniques to define new invariants of manifolds.

CONTENTS

1. Introduction
2. Topology of configuration spaces
 - 2.1. Components of the configuration space
 - 2.2. The fundamental group of the configuration space
 - 2.3. Rational cohomology of configuration spaces
 - 2.4. Other topological results
3. Functionals and analytical techniques
 - 3.1. Local representation of functions by flat connections
 - 3.2. Homotopy theory for finite energy maps
 - 3.3. Minimizing functionals
- References

§1. INTRODUCTION

Physicists have used variational principles to describe physical phenomena for a long time. The classical trajectories of a system are modeled as the stationary points of an action functional. The quantum behavior of the system may be modeled in one of two frameworks: Lagrangian or Hamiltonian. The Lagrangian framework involves path integrals of an exponential of the action functional. This has proved to be a very powerful tool for physical modeling and mathematical inspiration. The Hamiltonian point of view is better understood mathematically, but there is still work to be done along these lines [28].

2000 *Mathematics Subject Classification.* Primary 81T13.

Key words and phrases. Skyrme model, Faddeev model, Yang–Mills equations.

The first author was partially supported by NSF grant DMS-0204651. The second author was partially supported by NSF grant DMS-0436403. The third author was partially supported by EPSRC grant GR/R66982/01.

The first case studied via a variational principle is the motion of a point particle on some configuration manifold. The static configurations are the critical points of a potential energy, and the classical trajectories are described by a nonlinear system of ordinary differential equations. These equations are the Euler–Lagrange equations of the action functional. The second case treats fields such as the electro-magnetic field. A field is a function on a manifold or, more generally, a section of a vector bundle over the manifold. The classical dynamics of a field is described by a generally nonlinear system of partial differential equations (the $U(1)$ Yang–Mills equations for electromagnetic fields, or the $SU(2) \times U(1)$ Yang–Mills equations for electroweak fields). These equations are the Euler–Lagrange equations of a certain functional. Quantum field theory is the quantum version of this second case. Physicists now hedge their bets by speaking of effective field theories [32]. By this they acknowledge that quantum field theories produce very accurate predictions, but are possibly not the correct theories to describe the universe. They argue that in certain limiting conditions one quantum field theory will accurately model physical behavior, whereas in other limiting conditions some other quantum field theory will provide the accurate model. It is reasonable to expect that, under certain limiting conditions, fields will take values close to the critical points of the potential energy. If the potential energy is invariant under a group action, then the set of critical points will be the union of several orbits of the group action. It is well known that an orbit of a smooth group action is diffeomorphic to a coset space G/H . In this situation, the basic objects to be considered are maps from a manifold to some nonlinear space (the critical points of the potential energy). This is the usual origin of nonlinear sigma models.

A nonlinear sigma model is a model with a configuration space consisting of maps from one manifold M , which represents space, to some target manifold N (see [6]). Techniques from the geometry of manifolds can shed light on the behavior of nonlinear sigma models, and it is possible that nonlinear sigma models may prove to be useful tools for the study of the geometry of manifolds.

In this paper, we shall discuss some geometric and analytical techniques that have been useful in the study of two nonlinear sigma models. In §2 we consider the topology of the relevant configuration spaces. In §3, we discuss analytical issues related to these models.

§2. TOPOLOGY OF CONFIGURATION SPACES

Mathematically, the challenges within the study of nonlinear sigma models arise because the target is not a linear or convex space. The topology of the target allows one to consider different types of constraints in an optimization problem. These constraints are an integral feature of the model. For example, it is common to look for a minimizer of a functional in a fixed homotopy class. If the homotopy class is not constrained, it may be that a constant map will minimize the functional.

In fact, many different but related configuration spaces can be considered. We shall concentrate on two general types of configuration spaces: maps into Lie groups and maps into S^2 . Let M be a compact, oriented 3-manifold and G any Lie group. Then the first configuration space we consider is the space of continuous maps $M \rightarrow G$ that send a chosen base point $x_0 \in M$ to the unit element $\mathbf{1} \in G$. This space is denoted by G^M . The second configuration space is $(S^2)^M$.

In this section, we shall give configuration spaces the compact open topology. In practice, some Sobolev topology depending on the energy functional is probably appropriate. The issue of checking the algebraic topology arguments given for classes of Sobolev maps is interesting. We shall address this issue in the next section. The study of the algebraic

topology of these spaces is motivated by the physical interpretation of the nonlinear sigma model. More detailed physical and geometric interpretations of the topological results in this section may be found in [5].

2.1. Components of the configuration space. The Skyrme model has been viewed as a model of several different systems. In one case, it was applied to model nucleons: the protons and neutrons in the center of an atom. The space is assumed to be \mathbb{R}^3 (this is reasonable when we are modeling the nucleus of an atom). The fields are taken to be maps into $\mathrm{Sp}(1)$, the group of unit quaternions. In order to have finite energy, the gradient of such a map must have finite L^2 norm. It is natural, therefore, to impose the boundary condition that the fields approach a constant value as $|\mathbf{x}| \rightarrow \infty$, where \mathbf{x} denotes the position in space. We may choose this boundary value to be 1 without loss of generality. So, we take our configuration space to be $\mathrm{Sp}(1)^{S^3}$. This is the space of base point preserving maps from S^3 to $\mathrm{Sp}(1)$. Notice that topologically $\mathrm{Sp}(1)$ is S^3 . The homotopy classes of maps from S^3 to $\mathrm{Sp}(1)$ are classified by an integer degree. The energy minimizer in the degree 0 class is precisely the constant map, or the vacuum, in physicists' language.

Linearizing the field equations about the vacuum, we obtain a wave equation for a field taking values in $\mathfrak{sp}(1) \cong \mathbb{R}^3$. Traveling wave solutions of this wave equation represent the π^+ , π^- , and π^0 mesons, which transmit the strong nuclear force between nucleons. Upon quantization, the Fourier coefficients of a general solution of the wave equation become the pion creation and annihilation operators of the theory. Particles of these kinds are called the "perturbative quanta" of quantum field theory, and in most theories, they are the only particles present (for example, electrons, positrons, and photons are the perturbative quanta of quantum electrodynamics). For an introductory account of the dynamics of perturbative quanta, see [24]. What makes the Skyrme model particularly interesting is that it possesses extra particle degrees of freedom, beyond its perturbative quanta. The neutron and proton appear in the model as the degree 1 minimizers of the energy functional. More precisely, they are the lowest energy quantum states built on top of a degree 1 minimizer. Deciding which states are protons and which neutrons is a matter of convention, since the theory models only the strong nuclear force, which is insensitive to this distinction (in physicists' language, it is isospin invariant). So the quantum ground state in the degree 1 sector represents a static nucleon. There are Lorentz boosted versions of this state representing a moving nucleon. The higher energy quantum states represent a composite of a nucleon and some number of pions, the more pions, the higher the excitation. There is a fundamental dichotomy between the perturbative quanta (pions) and the topological solitons (nucleons). The number of nucleons is given, as we shall argue shortly, by the degree of the sector on which the quantum wavefunctional has its support. The number of pions is given by the position of the quantum wavefunctional on a ladder of energy eigenstates, analogous to the energy ladder of the simple harmonic oscillator. This scheme is called semiclassical quantization, and is treated in detail for somewhat simpler systems than the Skyrme model in [26].

We return to the physical interpretation of the degree. Quantum states built on degree B minimizers are taken to represent bound states of B nucleons, that is, nuclei of atomic mass B . If the degree 1 minimizers themselves are viewed as nucleons, the particles are smeared out over space. There are several ways to define the center of such a particle. We could take it to be the position of maximum energy density, or maximum particle number density (the latter being the inner product of the volume form on space with the pull-back of the volume form on $\mathrm{Sp}(1)$). One simple and useful choice is to take the particle to be centered where its field differs most from its boundary value. In other words, we imagine a point in space mapped to -1 to be the center

of a particle. The field may be thought of as a smooth bump with support near such a point. The field could also appear as several bump functions representing several particles. There is an orientation in this situation, and the positively oriented bumps may be regarded as particles and the negatively oriented bumps may be regarded as antiparticles. The degree may be computed as the signed count of inverse images of a regular value. Since these images represent the nucleons and antinucleons, we see that the degree B is the net nucleon number of the field. (Actually, the quantum states built on the degree 1 minimizer represent a class of matter particles collectively called baryons, the higher excited states being somewhat exotic. For this reason, the degree of a field is called its baryon number in the Skyrme model literature.) The path components of the configuration space consist of all maps with a fixed degree. Notice that it is possible to create or cancel particle-antiparticle pairs and stay within the same path component. It is desirable to have a functional that will force the trajectories to stay in a fixed path component. This corresponds to a conservation law: there can be no net change in the number of particles. It is exactly these considerations that led physicists to ask about the homotopy classification of maps from one space to another. In the case with Lie group target we have the following result from [1].

Proposition 1. *Let G be a compact Lie group and M a connected, closed 3-manifold. The set of path components of G^M is*

$$G^M / (G^M)_0 \cong H^3(M; \pi_3(G)) \times H^1(M; H_1(G)).$$

In the case of Lie-group-valued maps, the configuration space has the structure of a topological group, so the space of path components is also a group. It would be nice to understand the group structure. The reason the above proposition only describes the set of path components is that in [1] Auckly and Kapitanski only established an exact sequence

$$0 \rightarrow H^3(M; \pi_3(G)) \rightarrow G^M / (G^M)_0 \rightarrow H^1(M; H_1(G)) \rightarrow 0.$$

To understand the group structure on the set of path components, one would have to understand a bit more about this sequence (e.g., does it split?). This is one of the open questions that needs to be addressed. There is a similar result for free maps (see [1]).

The homotopy classification of either based or free maps from a 3-manifold (3-complex in fact) to S^2 was worked out by Pontryagin many years ago, [25].

Theorem 2. *Let M be a closed, connected, oriented 3-manifold, and let μ_{S^2} be a generator of $H^2(S^2; \mathbb{Z}) \cong \mathbb{Z}$. With any based map φ from M to S^2 , one may associate the cohomology class $\varphi^* \mu_{S^2} \in H^2(M; \mathbb{Z})$. Every second cohomology class may be obtained from some map, and any two maps with different cohomology classes lie in distinct path components of $(S^2)^M$. Furthermore, the set of path components corresponding to a cohomology class $\alpha \in H^2(M)$ is in bijective correspondence with $H^3(M) / (2\alpha \sim H^1(M))$.*

The physical description given before these two theorems, interpreting the degree as a count of particles and antiparticles, is the basis of the Pontryagin–Thom construction. Now we shall review this construction, because it is useful for interpreting these theorems in general.

Here we follow the folklore maxim: think with intersection theory and prove with cohomology. The combination of Poincaré duality and the Pontryagin–Thom construction gives a powerful tool for visualizing results in algebraic topology. If W is an n -dimensional oriented homology manifold, the Poincaré duality is the isomorphism $H^k(W) \cong H_{n-k}(W)$. It is tempting to think of the k th cohomology as the dual of the k th homology. This is not far from the truth. Putting these together, we see that every degree k cohomology class corresponds to a unique $(n - k)$ -cycle (codimension

k homology cycle), and the image of the cocycle applied to a k -cycle is the weighted number of intersection points with the corresponding $(n - k)$ -cycle. For field coefficients this is the entire story since there is no torsion and the Ext group vanishes. With other coefficients, this gives the correct answer up to torsion. With any map $W \rightarrow S^k$, the Pontryagin–Thom construction associates a framed codimension k submanifold of W . The associated submanifold is precisely the inverse image of a regular value. This is well defined up to cobordism. Conversely, a framed submanifold produces a map $W \rightarrow S^k$ defined via the exponential map on fibers of a tubular neighborhood of the submanifold and as the constant map outside of the neighborhood.

We first consider the geometric description of the components of G -valued maps for the compact, simply connected Lie groups. Each such group is a product of simple groups. These are listed, along with their center and rational cohomology, in Table 1. The rational cohomology groups are the free unital graded commutative algebras generated by the listed elements. The subscripts are the degrees of the generators. In this case, the $H^1(M; H_1(G))$ -factor in Proposition 1 is trivial. The set of components is isomorphic to \mathbb{Z} when G is simple and M is orientable and connected. In fact, if G has two simple factors and M is connected, closed, and orientable, we have $H^3(M; H_3(G)) \cong \mathbb{Z}^2$; this group represents the number of particles of type one and the number of particles of type two. If M is nonorientable, this group is trivial. If the space were nonorientable, it would not make sense to count the number of particles, because a particle could change into an antiparticle by traversing a nonorientable loop. To interpret the $H^3(M; H_3(G))$ -factor geometrically, we shall use universal coefficient duality to view elements of $H^3(M)$ as functionals on $H_3(M)$. Using Poincaré duality, we view $H_3(G)$ as $H^{\dim G - 3}(G)$. By using this identification, the elements of $H^3(M; H_3(G))$ may be interpreted as functions that associate an integer to a 3-cycle in M and a codimension 3 cycle in G . Any map $u : M \rightarrow G$ generates such a function given by the signed count of the intersection points of the u -image of the 3-cycle and the codimension 3 cycle. There is a different way to say this. Fix a trivialization of the normal bundle to a codimension 3 cycle in G denoted by F (pretend that the normal bundle to F is trivial; it is trivial off a codimension 2 set; see [5]). With a map $u : M \rightarrow G$ (and the fundamental 3-cycle $[M]$ and codimension 3 cycle F), we associate the collection of points $u^{-1}(F)$. Such a point is positively oriented if the push-forward of an oriented frame at the point has the same orientation as the trivialization of the normal bundle at the image. Conversely, with any finite collection of signed points we may associate a based map $u : M \rightarrow G$. Using a positively or negatively oriented frame at each point, we construct a diffeomorphism from the closed tubular neighborhood of each point to the 3-disk of radius π in the space $\mathfrak{sp}(1)$ of purely imaginary quaternions. Via the exponential map $\exp : \mathfrak{sp}(1) \rightarrow \mathrm{Sp}(1)$ given by $\exp(x) = \cos(|x|) + \frac{\sin(|x|)}{|x|}x$, we define a map from the closed tubular neighborhood of the points to $\mathrm{Sp}(1)$. This map may be extended to the whole 3-manifold by sending points in the complement of the neighborhood to -1 . Next, we modify the map by multiplying by -1 , so that the base point will be 1. Finally, we notice that $H_3(G)$ is generated by a homomorphic image of $\mathrm{Sp}(1)$, so we can complete our map into G by composition with this inclusion. We are of course assuming that F is dual to the image of this $\mathrm{Sp}(1)$.

If the group G is not assumed to be simple but is simply connected and compact, it is a product of simple compact groups, so any map will be specified by its components, and the components of a map will be specified by the inverse image of the appropriate codimension 3 cycles as described above. If the group is no longer assumed to be simply connected, we can associate an element of $H^1(M; H_1(G_0))$ with any map $u \in G^M$. This element maps a loop in M to a loop in G_0 by composing with u . If u and v are two maps that generate the same element of $H^1(M; H_1(G_0))$, then $u^{-1}v$ generates the trivial

TABLE 1. Simple groups

group, G	center, $Z(G)$	$H^*(G; \mathbb{Q})$
$A_n = \mathrm{SU}(n+1), n \geq 2$	\mathbb{Z}_{n+1}	$\mathbb{Q}[x_3, x_5, \dots, x_{2n+1}]$
$B_n = \mathrm{Spin}(2n+1), n \geq 3$	\mathbb{Z}_2	$\mathbb{Q}[x_3, x_7, \dots, x_{4n-1}]$
$C_n = \mathrm{Sp}(n), n \geq 1$	\mathbb{Z}_2	$\mathbb{Q}[x_3, x_7, \dots, x_{4n-1}]$
$D_n = \mathrm{Spin}(2n), n \geq 4$	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$ for even n \mathbb{Z}_4 for odd n	$\mathbb{Q}[x_3, x_7, \dots, x_{4n-5}, y_{2n-1}]$
E_6	\mathbb{Z}_3	$\mathbb{Q}[x_3, x_9, x_{11}, x_{15}, x_{17}, x_{23}]$
E_7	\mathbb{Z}_2	$\mathbb{Q}[x_3, x_{11}, x_{15}, x_{19}, x_{23}, x_{27}, x_{35}]$
E_8	0	$\mathbb{Q}[x_3, x_{15}, x_{23}, x_{27}, x_{35}, x_{39}, x_{47}, x_{59}]$
F_4	0	$\mathbb{Q}[x_3, x_{11}, x_{15}, x_{23}]$
G_2	0	$\mathbb{Q}[x_3, x_{11}]$

element, and, thus, lifts to a map to the universal covering group \tilde{G} . Since the universal covering group of a Lie group is a product of a vector space and a compact group, up to homotopy this is a simply connected, compact Lie group. The homotopy class of the lift is specified by the element of $H^3(M; H_3(\tilde{G}))$ as described previously.

Given a map $\varphi : M \rightarrow S^2$, denote by $(S^2)_\varphi^M$ the set of all maps $\psi \in (S^2)^M$ homotopic to φ . The picture of the components of $(S^2)_\varphi^M$ arising from the Pontryagin–Thom construction and Poincaré duality is quite nice. The inverse image of a regular value in S^2 is Poincaré dual to $\varphi^* \mu_{S^2}$. This inverse image is a framed loop in M . Thus, such a map into S^2 can be viewed as modeling a spread-out string in the same way as a map into a Lie group was viewed as modeling a spread-out particle. The relative number of twists in the framing of a second map with the same pull-back is an element of $H^3(M; \mathbb{Z}) / \langle 2\varphi^* \mu_{S^2} \rangle$. We shall identify S^2 with \mathbb{CP}^1 and consider several maps to clarify the situation. Let $\varphi_1, \varphi'_1, \varphi_3 : \mathbb{CP}^1 \times S^1 \rightarrow \mathbb{CP}^1$ be the maps given by

$$\varphi_1([z : w], \lambda) = [z : w], \quad \varphi'_1([z : w], \lambda) = [\lambda z : w], \quad \text{and} \quad \varphi_3([z : w], \lambda) = [z^3 : w^3].$$

We can view $\mathbb{CP}^1 \times S^1$ as $S^2 \times [0, 1]$ (a spherical shell) with the inner and outer ($S^2 \times \{0\}$ and $S^2 \times \{1\}$) spheres identified. Using this convention, we have displayed in Figure 1 the framed 1-manifolds arising as the inverse image of a regular value. These figures also specify a framing. The first vector may be taken perpendicular to the plane of the figure, and the second vector may be obtained from the cross product with the tangent vector.

It may appear that there is a well-defined twist number associated with an S^2 -valued map (count the number of times the framing rotates as it moves around the loop). However, there is a homeomorphism of $\mathbb{CP}^1 \times S^1$ twisting the 2-sphere (such a map is given by $([z : w], \lambda) \mapsto ([\lambda z : w], \lambda)$). This will change the apparent number of twists in a framing, but will not change the relative number of twists. The reason why this relative number of twists is only well defined modulo twice the divisibility of the cohomology class $\varphi^* \mu_{S^2}$ is demonstrated for φ_1 in Figure 2. In physics, this trick (move) is illustrated by Dirac’s spanner (Dirac’s string problem). P. A. M. Dirac used it (with scissors as a spanner) to explain the meaning of the half-spin. One nice reference is [19].

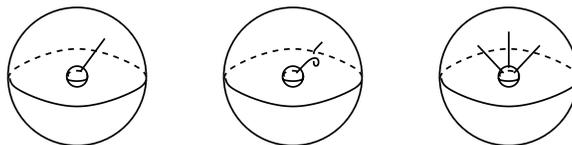


FIGURE 1. Pontryagin–Thom representatives of S^2 -valued maps

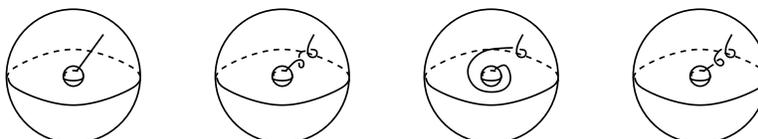


FIGURE 2. Introducing $2d$ twists

2.2. The fundamental group of the configuration space. Typically, the components of the configuration space are labeled by the number of particles in the system; other topological properties of the configuration space may be used to model additional physical properties. Below we shall review the notion of statistics of particles and describe one idea to use the fundamental group of the configuration space to incorporate statistics in a nonlinear model.

Recall the distinction between bosons and fermions: a macroscopic ensemble of identical bosons behaves statistically as if arbitrarily many particles can lie in the same state, while a macroscopic ensemble of identical fermions behaves as if no two particles can lie in the same state. Photons are examples of particles with bosonic statistics, and electrons are examples of particles with fermionic statistics. There are several theoretical models of particle statistics. In quantum mechanics, the wavefunction representing a multiparticle state is symmetric under exchange of any pair of identical bosons, and antisymmetric under exchange of any pair of identical fermions. In conventional perturbative quantum field theory, commuting fields are used to represent bosons and anticommuting fields are used to represent fermions. More precisely, bosons have commuting creation operators and fermions have anticommuting creation operators. However, there are times when fermions may arise within a field theory with purely bosonic fundamental fields. This phenomenon is called emergent fermionicity, and it relies crucially on the topological properties of the underlying configuration space of the model.

Emergent fermionic behavior may sometimes be explained on the basis of symmetry. Symmetries of the classical configuration space often imply symmetries of the space of quantum states. However, the group of quantum symmetries can be different from the group of classical symmetries. It is this difference that may be used to explain emergent fermionicity; see [6, Chapter 7]. A spinning top is a well-known example of this. The classical symmetry group is $SO(3)$, while the quantum symmetry group for some quantizations is $SU(2)$; see [8, 27]. An electron in the field of a magnetic monopole is also a good example; see [6, Chapter 7] and [11, 35].

When the physical space is \mathbb{R}^3 , so that space-time is the usual Minkowski space, the classical rotational symmetries induce quantum symmetries that are representations of the *Spin* group. These irreducible representations are labeled by half integers (one half of the number of boxes in a Young diagram for a representation of $SU(2)$). The integral representations are honest representations of the rotation group, but the fractional ones are not. The spin statistics theorem states that any particle with fractional spin is a fermion, and any particle with integral spin is a boson.

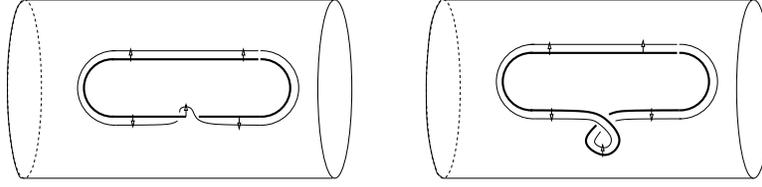


FIGURE 3. The exchange loop

The statistics of a particle may be viewed as a parity. A compound particle made out of an even number of fermions will be a boson, and a compound particle comprised of an odd number of fermions will be a fermion. With this in mind, consider the fundamental group of the space of based maps from S^3 to $\mathrm{Sp}(1)$.

Figures 3 and 4 show some loops in this configuration space. A loop in $\mathrm{Sp}(1)^{S^3}$ is represented by a map from $S^3 \times [0, 1]$ to $\mathrm{Sp}(1)$. Such a map may be described by a framed loop in $S^3 \times [0, 1]$. This framed loop is the inverse image of a frame at a regular value. Conversely, given a framed loop in $S^3 \times [0, 1]$, one can construct a map $S^3 \times [0, 1] \rightarrow \mathrm{Sp}(1)$. The fibers of the closed tubular neighborhood are mapped to $\mathrm{Sp}(1)$ via the exponential map, and the complement of this neighborhood is mapped to the base point. This is exactly analogous to the correspondence between signed points in S^3 and maps into $\mathrm{Sp}(1)$ described in the subsection on the components of Lie-group-valued maps. The horizontal direction in Figures 3 and 4 represents the interval (time) direction in $S^3 \times [0, 1]$. The disks represent the (x, y) -plane, and we suppress the z -direction. Figure 3 shows two copies of a typical loop representing an element in the fundamental group. Only the first vector of the framing is shown in Figure 3. The second vector is obtained by taking the cross product with the tangent vector to the curve in the displayed slice, and the final vector is the z -direction. It is easy to see that the left copy can be deformed into the right copy. Extending the earlier interpretation of inverse images as particle lumps, we describe the left copy as follows: a particle and antiparticle are born (a slice on the far left has no inverse images of -1 and hence no particles or antiparticles, but a slice a bit further to the right has two oppositely oriented inverse images); the particle undergoes a full rotation (the framing twists as the slice moves to the right); the two particles then annihilate (there are no inverse images on a slice on the far right). The right copy may be described as follows: a first particle-antiparticle pair is born; a second pair is born; the two particles exchange positions without rotating (the strand that is second from the bottom in the slice just to the left of the center is third from the bottom in a slice just to the right of the center and the displayed framing vector on each remains pointing down); the first particle and the second antiparticle annihilate; the remaining pair annihilates. Notice that there are two ways a pair of particles can exchange positions; they can face each other or face a fixed reference point. If the particles face each other, each will rotate during the exchange. The exchange described in the second frame is an exchange without rotation. Representing the particles by people in a room, the two people may step sideways forwards/backwards and sideways following diametrically opposite points on a circle and always facing the back of the room. This is the exchange without rotating described in Figure 3. This exchange is nontrivial in $\pi_1(\mathrm{Sp}(1)^{S^3})$. The second way a pair of people may change positions is to walk around a circle at diametrically opposite points always facing the direction that they walk, to end up facing the opposite direction that they started. This second change of position is actually homotopically trivial. Since the framed links in Figure 3 avoid the slices $S^3 \times \{0, 1\}$, they represent a loop based at the constant identity map.

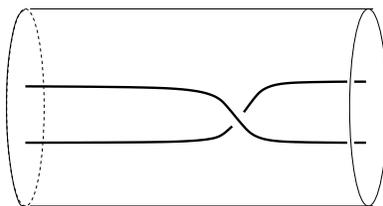


FIGURE 4. The degree 2 exchange loop

It is possible to describe a framing without drawing any normal vectors at all. The first vector may be taken perpendicular to the plane of the figure, the second vector may be obtained from the cross product with the tangent vector, and the third vector may be taken to be the suppressed z -direction. The framing obtained by following this convention is called the blackboard framing. We use the blackboard framing in Figure 4. The Pontryagin–Thom construction may also be used to visualize loops in other components of the configuration space. Figure 4 shows a loop in the degree 2 component of the space of maps from S^3 to $\mathrm{Sp}(1)$.

In fact, the exchange loop that we just described in $\pi_1(\mathrm{Sp}(1)^{S^3})$ has order two. This may be seen as follows. Wrapping twice around the exchange loop is the boundary of a disk in the configuration space. Since $\mathbb{R}P^2$ is the result of identifying the points on the boundary of a disk via a degree 2 map, one expects to find an $\mathbb{R}P^2$ embedded into configuration space. In [30], Sorkin described an embedding $f_{\mathrm{stat}} : \mathbb{R}P^2 \rightarrow \mathrm{Sp}(1)^{S^3}$. We briefly recall Sorkin’s elegant construction. Describe $\mathbb{R}P^2$ as the 2-sphere with antipodal points identified. By addition of particle-antiparticle pairs, we may assume that there are two particles in a representative map. We may place the particles at antipodal points of a sphere, by using frames parallel to the coordinate directions. The loop in configuration space represented by a compound particle making a full rotation is trivial in the fundamental group if the compound particle consists of an even number of basic particles, and is nontrivial if the compound particle consists of an odd number of basic particles. This observation is the basic motivation behind Finkelstein and Rubinstein’s method of incorporating fermionic solitons in the Skyrme model [16].

The fundamental groups of the Lie-group-valued and S^2 -valued configuration spaces were computed in [5]. The results are repeated below.

Theorem 3. *If M is a closed, connected, orientable 3-manifold, and G is any Lie group, then*

$$\pi_1(G^M) \cong \mathbb{Z}_2^s \oplus H^2(M; \pi_3(G)).$$

Here s is the number of symplectic factors in the Lie algebra of G .

Theorem 4. *Let M be closed, connected, and orientable. For any $\varphi \in (S^2)^M$, the fundamental group of $(S^2)_\varphi^M$ is given by*

$$\pi_1((S^2)_\varphi^M) \cong \mathbb{Z}_2 \oplus H^2(M; \mathbb{Z}) \oplus \ker(2\varphi^* \mu_{S^2} \smile).$$

Here $2\varphi^* \mu_{S^2} \smile : H^1(M; \mathbb{Z}) \rightarrow H^3(M; \mathbb{Z})$ is the usual map given by the cup product.

The Pontryagin–Thom construction may also be used to understand the isomorphism

$$\phi : \pi_1(G^M) \cong \mathbb{Z}_2^s \oplus H^2(M; \pi_3(G))$$

asserted in Theorem 3. A loop in $(G^M)_0$ based at the constant map $u(x) = \mathbf{1}$, may be regarded as a based map $\gamma : SM \rightarrow G$. The identifications in the suspension provide a particularly nice way to summarize all of the constraints on γ imposed by the base

points. We shall use the same notation for the map $\gamma : M \times [0, 1] \rightarrow G$ obtained from γ by composition with the natural projection. The group $\pi_3(G)$ is generated by homomorphic images of $\mathrm{Sp}(1)$. Let F be a codimension 3 cycle in G dual to a given generator of $\pi_3(G)$ (as is well known, the normal bundle to a particular such cycle is trivial away from a codimension 2 subset). The inverse image $\gamma^{-1}(F)$ with framing obtained by pulling back the trivialization of $\nu(F)$ may be associated with γ . Conversely, given a framed link in $(M - p_0) \times (0, 1)$, we may construct an element of $\pi_1(G^M)$. By using the framing, each fiber of the closed tubular neighborhood of the link may be identified with the disk of radius π in $\mathfrak{sp}(1)$. As before, -1 times the exponential map may be used to construct a map $\gamma : SM \rightarrow \mathrm{Sp}(1)$ that, after composition with the homomorphism $\mathrm{Sp}(1) \rightarrow G$, represents an element of $\pi_1(G^M)$.

It is now possible to describe the geometric content of the isomorphism in Theorem 3. For a class of loops $[\gamma] \in (G^M)_0$, we describe the isomorphism $\phi(\gamma) = (\phi_1(\gamma), \phi_2(\gamma)) \in \mathbb{Z}_2^s \oplus H^2(M; \pi_3(G))$. We restrict attention to the case of a simply connected G , and make the identifications $\pi_3(G) \cong H_3(G; \mathbb{Z}) \cong H^{\dim G - 3}(G; \mathbb{Z})$. An element of $H^2(M; \pi_3(G))$ may be interpreted as a function that associates an integer to a surface in M , say Σ , and a codimension 3 cycle in G , say F . Set $\phi_2(\gamma)(\Sigma, F) = \#(\Sigma \times [0, 1] \cap \gamma^{-1}(F))$. Note that $\gamma^{-1}(F)$ inherits an orientation from the framing and orientation on M . By using Poincaré duality, this may be said in a different way. The homology class of $\gamma^{-1}(F)$ in $(M - p_0) \times (0, 1)$ projects to an element of $H_1(M)$ dual to the element associated with $\phi_2(\gamma)$. The first component of the isomorphism $(\varphi(\gamma))$ counts the parity of the number of twists in the framing.

Similarly, we can describe the content of Theorem 4. An element of $\pi_1((S^2)_\varphi^M)$ is represented by a map $\gamma : M \times S^1 \rightarrow S^2$. The inverse image of a regular value is a 2-dimensional submanifold, say Σ . This determines an element of $H^1(M; \mathbb{Z})$ as follows. With any 1-cycle in M , say σ , we associate the intersection number of Σ and $\sigma \times S^1$. Since our loop is in the path component of φ , the surface Σ is parallel to the φ -inverse image of a regular value. This implies that our element of $H^1(M; \mathbb{Z})$ is in the kernel of the map $2\varphi^* \mu_{S^2} \smile : H^1(M; \mathbb{Z}) \rightarrow H^3(M; \mathbb{Z})$. Given any element of this kernel, we can define a loop in $(S^2)_\varphi^M$ via what we call the \mathfrak{q} -map. This is the map $\mathfrak{q} : S^2 \times S^1 \rightarrow \mathrm{Sp}(1)$ defined by $\mathfrak{q}(xix^*, \lambda) = x\lambda x^*$. It can be checked that every element of S^2 can be written in the form xix^* , and the expression $x\lambda x^*$ is independent of the choice of x representing the element of S^2 . There is a map $u : M \times S^1 \rightarrow \mathrm{Sp}(1)$ that may be used to change this new loop back into γ . The remaining homotopy invariants of γ are those of u as described previously.

In order to convey some feel for the proofs of these results, we repeat the proof of one of the main steps from [5].

Proposition 5. *The sequence*

$$0 \rightarrow \pi_4(G) \rightarrow \pi_1(G^M) \rightarrow H^2(M; \pi_3(G)) \rightarrow 0$$

splits, and there is a splitting associated with each spin structure on M .

Proof. It suffices to check the result for $G = \mathrm{Sp}(1)$. Since the three-dimensional Spin cobordism group is trivial, every 3-manifold is surgery on a framed link with even self-linking numbers [22]. Such a surgery description induces a Spin structure in M . Let $M = S_L^3$ be such a surgery description, orient the link, and let $\{\mu_j\}_{j=1}^c$ be the positively oriented meridians to the components of the link. These meridians generate $H_1(M) \cong H^2(M; \pi_3(\mathrm{Sp}(1)))$. This last isomorphism is Poincaré duality. We define a splitting:

$$s : H_1(M) \rightarrow \pi_1((\mathrm{Sp}(1))^M); \quad s(\mu_j) = PT(\mu_j \times \{\frac{1}{2}\}, \text{canonical framing}).$$

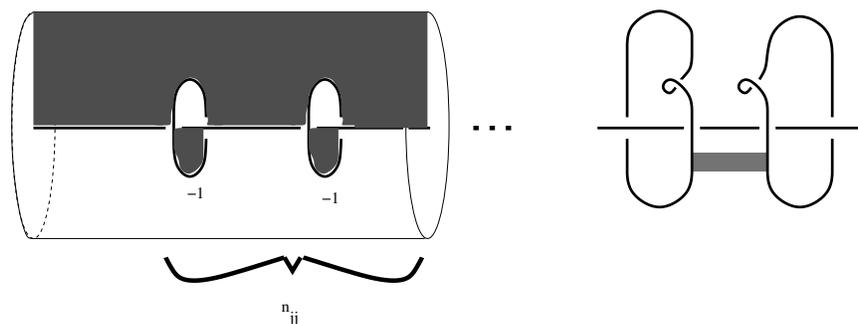


FIGURE 5. The 2-cycle

Here PT represents the Pontryagin–Thom construction, and the canonical framing is constructed as follows. The first vector is chosen to be the 0-framing on μ_j viewed as an unknot in S^3 (i.e., the linking number of μ_j and the curve obtained by applying the exponential map to ϵ times this first vector in S^3 , is zero for sufficiently small ϵ). The second vector is obtained by taking the cross product of the tangent vector with the first vector, and the third vector is just the direction of the interval. Now we shall check that this map respects the relations in $H_1(M)$. Let $Q_L = (n_{jk})$ be the linking matrix, so that $H_1(M) = \langle \mu_j | n_{jk} \mu_j = 0 \rangle$. We are using the summation convention in this description. The 2-cycle representing the relation $n_{jk} \mu_j = 0$ may be constructed from a Seifert surface to the j th component of the link, when this component is viewed as a knot in S^3 . Let Σ_j denote this Seifert surface. The desired 2-cycle is then $\widehat{\Sigma}_j = (\Sigma_j - \mathring{N}(L)) \cup \sigma_j$. Here σ_j is the surface in $S^1 \times D^2$ with n_{jj} meridians depicted on the left in Figure 5.

The boundary of $\widehat{\Sigma}_j$ is exactly the relation $n_{jk} \mu_j = 0$. For $k \neq j$, the framing on each copy of μ_k induced from this surface agrees with the 0-framing. The framing on each copy of μ_j is $-\text{sgn}(n_{jj})$. The surface $\widehat{\Sigma}_j$ may be extended to a surface in $M \times [0, 1] \times [0, 1]$ by adding a collar of the boundary in the direction of the second interval followed by one band for each pair of the μ_j , as depicted on the right of Figure 5. The resulting surface has a canonical framing, and the corresponding homotopy given by the Pontryagin–Thom construction homotopes the loop corresponding to the relation to a loop corresponding to a ± 2 -framed unlink. Such a loop is null-homotopic, as required. \square

It is interesting to notice that spin structures enter into the analysis of the topology of these configuration spaces. This should not be surprising since spin structures were first introduced to model the spin behavior of point particles. The fundamental group of the configuration space may be used to model the spin behavior of solitons. Spin^c structures enter at other points of the story.

2.3. Rational cohomology of configuration spaces. The second cohomology of configuration spaces plays an important role in quantum theory. The full rational cohomology may also appear as topological observables in the theory. To see the role of the second cohomology, recall some of the theory of geometric quantization [28]. In quantum theory one wishes to include the algebra of observables into the collection of operators on some function space. Given a configuration manifold (space of positions), one defines the phase space (collection of positions and momenta) to be the cotangent bundle of the configuration space. This has a canonical symplectic structure. The space of wave functions is chosen to be a set of sections of a complex line bundle over the quotient of phase space by an integrable isotropic distribution of maximal dimension, the so-called

polarization. The most common choice is the vertical polarization, consisting of the tangent spaces to the fibres of the cotangent bundle. In this case, the quotient is naturally identified with the configuration space itself.

There are many different complex line bundles over a space. This is called quantization ambiguity in the physics literature. Line bundles are in one-to-one correspondence with the elements of the second cohomology with integral coefficients. If we choose the vertical polarization, computing the quantization ambiguity requires one to consider the second integral cohomology of configuration space. It is well known that the second integral cohomology group may be determined from the fundamental group and the second rational cohomology group.

The real cohomology ring $H^*((G^M)_0, \mathbb{R})$, including its multiplicative structure, is described in the following theorem [5]. To state the theorem, we shall use a μ -map similar to that used in the Yang–Mills theory,

$$\mu : H_d(M; \mathbb{R}) \otimes H^j(G; \mathbb{R}) \rightarrow H^{j-d}(G^M; \mathbb{R}),$$

and the cohomology ring is generated (as an algebra) by the images of this map. The map μ produces a $(j-d)$ -cocycle in G^M from a d -cycle in M and a j -cocycle in G . At the level of chains, let $e^d : D^d \rightarrow M$ be a d -cell, and let x_j be a closed j -form on G . Given a singular simplex $u : \Delta^{j-d} \rightarrow G^M$, let $\hat{u} : M \times \Delta^{j-d} \rightarrow G$ be the natural map, and write

$$\mu(e^d \otimes x_j)(u) = \int_{D^d \times \Delta^{j-d}} \hat{u}^* x_j.$$

Theorem 6. *Let G be a compact, simply connected, simple Lie group. For a closed, connected, orientable 3-manifold M , the cohomology ring $H^*((G^M)_0; \mathbb{R})$ is the free graded-commutative unital algebra over \mathbb{R} generated by the elements $\mu(\Sigma_j^d \otimes x_k)$, where the Σ_j^d form a basis for $H_d(M; \mathbb{R})$ for $d > 0$ and $k-d > 0$, and the x_k generate $H^*(G; \mathbb{R})$ as a free graded-commutative unital algebra.*

The generators of $H^*(G; \mathbb{R})$ as a graded-commutative unital algebra are shown in Table 1 in Subsection 2.1. The following result of [5] describes the rational cohomology of $(S^2)^M$.

Theorem 7. *Let M be closed, connected, and orientable, let $\varphi : M \rightarrow S^2$, let Σ_j^d form a basis for $H_d(M; \mathbb{R})$ for $d < 3$, and let $\{\alpha_k\}$ form a basis for $\ker(2\varphi^* \mu_{S^2} \smile : H^1(M; \mathbb{Z}) \rightarrow H^3(M; \mathbb{Z}))$. The cohomology ring $H^*((S^2)_\varphi^M; \mathbb{R})$ is the free graded-commutative unital algebra over \mathbb{R} generated by the elements α_k and $\mu(\Sigma_j^d \otimes x)$, where $x \in H^3(\mathrm{Sp}(1); \mathbb{Z})$ is the orientation class. The classes α_k have degree 1, and the $\mu(\Sigma_j^d \otimes x)$ have degree $3-d$.*

It is interesting to compare the configuration space of $\mathrm{Sp}(1)$ connections modulo gauge transformations in the Yang–Mills theory to the configuration space of $\mathrm{Sp}(1)$ -valued maps in the Skyrme model; see Table 2.

In the Yang–Mills theory, there is a family of moduli spaces that may be used to define a family of homology cycles in the configuration space. Evaluating the generators of the cohomology of the configuration space on these homology classes produces a 4-manifold invariant. This invariant is called the Donaldson polynomial. It is a polynomial on the homology of the input 4-manifold.

It is natural to ask if there is a similar family of homology cycles in the Skyrme configuration space. If there is such a family, it would be possible to define a polynomial invariant of 3-manifolds analogous to the Donaldson polynomial.

2.4. Other topological results. Using standard results from algebraic topology, it is possible to generalize the results on the fundamental group and rational cohomology to

TABLE 2. Yang–Mills vs. Skyrme

	Yang–Mills configuration space	Skyrme configuration space
input space	a 4-manifold, X	a 3-manifold, M
homotopy type	$\text{Maps}(X, B\text{SU}(2))$	$\text{Maps}(M, \text{SU}(2))$
integer component labels	second Chern class, c_2	Chern–Simons invariant, $cs = Tc_2$
cohomology generators	$\mu(\Sigma^d \otimes c_2) \in H^{4-d}(\mathcal{C})$	$\mu(\Sigma^d \otimes Tc_2) \in H^{3-d}(\mathcal{C})$

cases where the manifold is not connected, or to arbitrary Lie groups. It is also possible to address configuration spaces of free maps. Details may be found in [5]. The two most subtle results in this direction are listed below.

Theorem 8. *We have*

$$\pi_1(\text{FreeMaps}(M, S^2)_\varphi) \cong \mathbb{Z}_2 \oplus (H^2(M; \mathbb{Z}) / \langle 2\varphi^* \mu_{S^2} \rangle) \oplus \ker(2\varphi^* \mu_{S^2} \smile).$$

Theorem 9. *There is a spectral sequence with $E_2^{p,q} = H^p(S^2; \mathbb{R}) \otimes H^q((S^2)_\varphi^M; \mathbb{R})$ converging to $H^*(\text{FreeMaps}(M, S^2)_\varphi; \mathbb{R})$. The second differential is given by $d_2 \mu(\Sigma^{(2)} \otimes x) = 2\varphi^* \mu_{S^2}[\Sigma] \mu_{S^2}$ with d_2 of any other generator trivial. All higher differentials are trivial as well.*

§3. FUNCTIONALS AND ANALYTICAL TECHNIQUES

There are many different function spaces that may be defined for maps from one manifold to a second manifold. For an analyst, geometry and physics are good guides to the best questions in analysis. Because of the structure of the energy functionals in physics, it is convenient to use the language of Sobolev spaces. For scalar functions on a compact manifold, the Sobolev spaces can be defined by first localizing on charts by means of a partition of unity, and then using the description of the space for functions on \mathbb{R}^n . This generalizes easily to vector-valued functions. To define a Sobolev space of functions taking values in a manifold (say N), one takes an isometric embedding of N into a vector space and considers those elements of the Sobolev space taking values in the vector space that lie in N almost everywhere. We denote by $W^{s,p}(M, N)$ the Sobolev space of maps from M to N that have s derivatives in L^p .

In this paper we consider two energy functionals related to the Skyrme model and the Faddeev model. For maps $u: N \rightarrow G$, Skyrme’s functional is given by the formula

$$(3.1) \quad E(u) = \int_M \frac{1}{2} |u^{-1} du|^2 + \frac{1}{4} |u^{-1} du \wedge u^{-1} du|^2 d\text{vol}_M.$$

In Skyrme’s paper [29] this functional was considered in the case of maps from \mathbb{R}^3 to the group $\text{SU}(2)$. We described the physical interpretation of such maps in the section devoted to topology.

Faddeev’s functional is defined for maps $\psi: M \rightarrow S^2$ by the formula (see [12])

$$(3.2) \quad E(\psi) = \int_M |d\psi|^2 + |d\psi \wedge d\psi|^2 d\text{vol}$$

and is simply related to the Skyrme functional: if we identify the sphere S^2 with an equator of the 3-sphere $S^3 \simeq \text{SU}(2)$, the Faddeev functional will coincide with the restriction of the Skyrme functional to functions with values in S^2 , $\psi(x) = \psi^1(x)\mathbf{i} + \psi^2(x)\mathbf{j} + \psi^3(x)\mathbf{k}$.

For these two models, the finite energy maps are the maps $u \in W^{1,2}(M, G)$, $\psi \in W^{1,2}(M, S^2)$ for which $E(u)$ and $E(\psi)$ are finite. The space of such maps will be denoted by $W_E^{1,2}$.

3.1. Local representation of functions by flat connections. One idea that has turned out to be useful in the study of nonlinear models has been a local representation of functions by flat connections. This first shows up when the codomain is a Lie group. Given a smooth map $u : M \rightarrow G$, one can construct the Lie-algebra-valued 1-form $A = u^{-1}du$. This form satisfies $dA + A \wedge A = 0$ and can be interpreted as a flat connection. Conversely, given a flat connection defined on a cube I^n , one may define a function $u : I^n \rightarrow G$ by solving the parallel transport equation $du(\dot{\gamma}(t)) = u(\gamma(t))A(\dot{\gamma}(t))$ with $u(x_0) = I$. The fact that A is flat implies that the value $u(\gamma(1))$ does not depend on the path connecting x_0 to $\gamma(1)$. Changing the initial value or base point will change the function u by a constant multiple. The same result holds for L^2 connections, [1], but the proof is much more involved. The result obtained in [1] reads as follows.

Theorem 10. *For any \mathfrak{g} -valued 1-form A of class L^2 on I^m such that*

$$(3.3) \quad dA + \frac{1}{2}[A, A] = 0$$

in the sense of distributions, there exists $u \in W^{1,2}(I^m, G)$ such that $u^{-1} \in W^{1,2}(I^m, G)$ and $A = u^{-1}du$. Furthermore, for any two such maps u and v , there exists $g \in G$ so that $u(x) = g \cdot v(x)$ for almost every $x \in I^m$.

The following local representation theorem was proved in [3] for S^2 -valued maps.

Theorem 11. *If φ is in $W^{1,3}(I^3, S^2)$ or $W_E^{1,2}(I^3, S^2)$, then there is $u \in W^{1,3}(I^3, \text{Sp}(1))$ or $u \in W^{1,2}(I^3, \text{Sp}(1))$, respectively, so that $\varphi = u^*iu$. In either case $\text{Re}((u^*du)^3) \in L^1$. Furthermore, for any two such maps u and v there is a map λ in $W^{1,3}(I^3, S^1)$ or $W^{1,2}(I^3, S^1)$, respectively, so that $v = \lambda u$.*

The space $W_E^{1,2}(I^3, S^2)$ was defined to analyze the Faddeev model. We shall recall its definition in a subsection below. The Hopf map $\text{Sp}(1) \rightarrow S^2$ given by $u \mapsto u^*iu$ is a convenient way to express the quotient projection $\text{Sp}(1) \rightarrow \text{Sp}(1)/S^1$. More generally, if H is a closed subgroup of a Lie group G , then the functions from I^n to G/H may be factored through G .

For simply connected, compact homogeneous spaces, this was done by Koshkin in [23]. Then, as before, a G -valued map can be specified by a flat connection. This is important for one class of sigma models in physics called coset models, where the target is a homogeneous space G/H .

3.2. Homotopy theory for finite energy maps. Our first application of these local representation theorems will be to the study of homotopy theory in classes of maps with finite energy. This question arises naturally, because one common feature among many sigma models is the restriction to a fixed homotopy class. The homotopy classes divide the particles into different sectors, and numerical invariants distinguishing the classes (sectors) bear some definite physical meaning (depending on the model). Good physical models describe interesting physics and have energy functionals that respect the geometry of the configuration space.

There has been a fair amount of work done relating homotopy theory to Sobolev spaces $W^{s,p}(M, N)$; see [10] for a survey of recent results. A paraphrase of the result of White [33] gives a good idea of what to expect: The homotopy classes of maps with one derivative in L^p restricted to the k -skeleton of a manifold agree with the homotopy classes of smooth maps provided $k < p$. For example, the second cohomology of a space is

completely determined by the homotopy classes of maps from the 3-skeleton of the space into $K(\mathbb{Z}, 2)$ (or $\mathbb{C}P^N$ for N sufficiently large). Thus, one would expect the pull-back of a second cohomology class under a $W^{1,3}$ map to be well defined. In general, the third cohomology of a space depends on the 4-skeleton; of course, if the space is a 3-manifold, then the 3-skeleton will suffice and the pull-back will be well defined for $W^{1,3}$ maps. Similarly, one would expect the Hopf invariant to be well defined for $W^{1,3}$ maps.

It is useful to keep several explicit maps in mind when considering the homotopy theory of maps in generalized function spaces. First, consider the map $\eta_1 : D^3 \rightarrow S^2$ from the 3-disk to the 2-sphere given by $\eta_1(x) = x/|x|$. Simple integration suffices to verify that η_1 is in $W^{1,p}$ for $p < 3$, but is not in $W_E^{1,2}$ or $W^{1,3}$. Notice that the integral of the η_1 -pull-back of the normalized area form on S^2 integrates to 1 on the boundary of D^3 . Thus, on second cohomology the pull-back may not be reasonably defined for such maps. Even more regularity is required to define the pull-back on the third cohomology. The composition of projection of S^3 to D^3 with this η_1 map may be patched into any smooth map from a 3-manifold to obtain a similar example. Similarly, the function $\eta_2 : D^3 \rightarrow S^1$ given by $\eta_2(x) = \cos(\ln|x|)\mathbf{i} + \sin(\ln|x|)\mathbf{j}$ is in $W_E^{1,p}$ for $p < 3$ but not in $W^{1,3}$, and the function $\eta_3 : D^3 \rightarrow S^1$ given by $\eta_3(x) = \cos(\ln|\ln|x/e||)\mathbf{i} + \sin(\ln|\ln|x/e||)\mathbf{j}$ is in $W^{1,3}$ but is not continuous. These last two functions may also be patched into maps from an arbitrary 3-manifold. Furthermore, they may be composed with maps into any nontrivial compact Lie group or into S^2 .

Our applications of the local representation theorems to homotopy problems are similar to the definition of Čech cohomology. As our first example, consider the space of flat connections modulo gauge equivalence. It is well known that the path components are labeled by the holonomy representations.

It is possible to define the holonomy for distributionally flat connections of class L_{loc}^2 . The sample functions listed above show that the local developing maps for such a connection can look like Swiss cheese. Here is a sketch of the definition. For a sufficiently fine cover of a manifold by cubes, the nerve of the cover will be homotopy equivalent to the manifold. (The nerve is a complex with one vertex for each cube, an edge for each nonempty intersection, a triangle for each nonempty triple intersection, etc.) The edge path group of a complex is an analog of the fundamental group of a topological space. It is isomorphic to the fundamental group of the topological realization of the complex. Given a distributionally flat connection of class L_{loc}^2 , say A , we can find local developing maps u_p such that $A = u_p^{-1}du_p$ on each cube by the local representation theorem. Furthermore, each edge of the nerve may be labeled by the group element relating the two local developing maps. This gives a representation of the edge path group. The actual definition includes the transition functions of the bundle supporting the connection, and it is only valid for central bundles. It is an open question to generalize the definition of holonomy to arbitrary flat connections; see [1].

Using the notion of holonomy, it is possible to generalize the local representation theorem to a global representation theorem.

Lemma 12. *Two distributionally flat connections of class L_{loc}^2 on a central bundle are gauge equivalent if and only if they have the same holonomy.*

The homotopy invariant of a map from M to G living in $H^1(M; H_1(G_0))$ may be interpreted as the holonomy of a flat connection, and therefore, is well defined for $W^{1,2}$ maps.

There are similar applications of the local representation theorem for S^2 -valued maps. Given a smooth map $\varphi : M \rightarrow S^2$, one may construct the pull-back of the orientation class $\mu_{S^2} \in H^2(S^2; \mathbb{Z})$. This class $\varphi^*\mu_{S^2} \in H^2(M; \mathbb{Z})$ is also the obstruction to lifting

the map φ to a map $u : M \rightarrow \mathrm{Sp}(1)$ such that $\varphi = u\mathbf{i}u^*$. Using this obstruction, we can prove the following global representation theorem; see [2, 3].

Proposition 13. *If φ and ψ are two maps in $W^{1,3}(M, S^2)$ or $W_E^{1,2}(M, S^2)$, then there is $w \in W^{1,3}(M, \mathrm{Sp}(1))$ or $w \in W^{1,2}(M, \mathrm{Sp}(1))$, respectively, with $\int_M \mathrm{Re}((w^*dw)^3) < \infty$ and $\psi = w\varphi w^*$ if and only if $\varphi^*\mu_{S^2} = \psi^*\mu_{S^2}$. If w_1 and w_2 are two such maps, then there is λ in $W^{1,3}(M, S^1)$ or in $W^{1,2}(M, S^1)$, respectively, with $w_1 = \mathbf{q}w_2$.*

The proof of this lifting theorem is quite technical and involves ideas of the Čech theory. It is similar to the proof of local gauge slices in gauge theory. The geometry suggests a natural system of differential equations that may be solved to give the lift.

Using de Rham theory, some homotopy invariants may be represented as integrals of differential forms. When an invariant is integral, this is the first thing anyone would try. Invariants living in the top cohomology are usually represented in this way. Sometimes a homotopy invariant takes values in a group with torsion. There are two ways we address torsion. We use the Čech picture as we described above, or we lift the map to a different space and realize the torsion invariant of the original map as an integral invariant of the lift. Torsion is introduced in this second picture if there is more than one possible lift. In any such applications of de Rham theory, there is a new issue that arises. The de Rham cohomology is cohomology with real coefficients, and the integral is the cap product. Thus, *a priori* such integrals are only known to be real numbers. It is possible to use the local representation theorems to prove that these integrals take integral values, even for Sobolev maps. Some sample theorems from [3] are listed below.

Theorem 14. *Let G be a compact, simply connected Lie group, let M be a closed oriented 3-manifold, let Θ be a smooth form representing an integral class in $H^3(G)$, and let u be a map in $W^{1,3}(M, G)$ or in $W_E^{1,2}(M, G)$. Then the number $\int_M u^*\Theta$ is an integer.*

It follows that the $H^3(M; H_3(\tilde{G}))$ invariants of homotopy classes of maps from M to G are integers for Sobolev maps and general compact Lie groups. By Lemma 12, any distributionally flat connection of class L_{loc}^2 is gauge equivalent to a smooth connection. The Chern–Simons invariants of these two connections are related by the degree of the gauge transformation, so we have the following corollary.

Corollary 15. *If A is a finite energy distributionally flat connection of class L_{loc}^2 on a central bundle, then there is a smooth connection on the same bundle with the same holonomy and Chern–Simons invariant.*

This integrality proposition combines with the lifting theorem for S^2 -valued maps to give a version of the secondary invariant for Sobolev S^2 -valued maps. We need a second integrality result to know that the degree of the lift changes by multiples of the right amount when the lift is changed. This is the content of the next result.

Proposition 16. *If ψ and λ are in $W^{1,3}$, or if ψ is in $W_E^{1,2}$ and λ is in $W^{1,2}$, then the expression $-\frac{1}{16\pi^2i} \int_M \psi d\psi \wedge d\psi \wedge \lambda^*d\lambda$ is an integer.*

Here is the formal definition of the secondary homotopy invariant for Sobolev S^2 -valued maps.

Definition 17. Given $\varphi, \psi : M \rightarrow S^1$ in $W^{1,3}$ or in $W_E^{1,2}$ such that $\varphi^*\mu_{S^2} = \psi^*\mu_{S^2}$, we define

$$\Upsilon(\varphi, \psi) = -\frac{1}{12\pi^2} \int_M \mathrm{Re}((w^*dw)^3) \in \mathbb{Z}_{2m_\psi},$$

where w is the map given by Proposition 13 and m_ψ is the divisibility of the class $\psi^*\mu_{S^2}$.

From Theorem 14, it follows that Υ is an integer. To see that Υ is well defined, notice that

$$-\frac{1}{12\pi^2} \int_M \operatorname{Re}((wq(\psi, \lambda)^* dw)^3) = -\frac{1}{8\pi^2 i} \int_M \psi d\psi \wedge d\psi \wedge \lambda^* d\lambda - \frac{1}{12\pi^2} \int_M \operatorname{Re}((w^* dw)^3).$$

See [2] for a proof of this. These invariants serve to generalize Pontryagin's theorem to Sobolev finite energy maps, because two smooth S^2 -valued maps φ and ψ are homotopic if and only if $\varphi^* \mu_{S^2} = \psi^* \mu_{S^2}$ and $\Upsilon(\varphi, \psi) = 0$.

Our expression for Υ generalizes an integral formula for the Hopf invariant; see [9]. Indeed, when $\varphi^* \mu_{S^2}$ is torsion (in particular if $M = S^3$), the pull-back $\varphi^* \omega_{S^2}$ of the volume form on S^2 is exact, and there is a 1-form, θ^φ such that $d\theta^\varphi = \varphi^* \omega_{S^2}$. Then the Hopf invariant of φ is defined to be $\int_M \theta^\varphi d\theta^\varphi$. The following proposition relates the Hopf invariant and Υ .

Proposition 18. *Let $\varphi : M \rightarrow S^2$ in $W^{1,3}$ or in $W_E^{1,2}$ be such that $R\varphi^* \mu_{S^2} = 0$ for some nonzero integer. For the minimal positive R , we have*

$$\operatorname{Hopf}(\varphi) = \frac{1}{R^2} \Upsilon(\varphi_R, \mathbf{i}),$$

where φ_R is the composition of φ with the map $z \mapsto z^R$ of S^2 to itself.

Thus, we see that our generalization of Pontryagin's theorem specializes to an integrality result for the Hopf invariant of Sobolev maps.

3.3. Minimizing functionals. One motivation for our study of the topology of spaces of maps and extensions to Sobolev maps was to consider minimization problems for the Skyrme and Faddeev functionals. In physical lingo, the minimizers of the energy functional (in every sector) are called the ground states. Their geometric structure and stability are of great importance. The reader may find interesting pictures and discussion of purported ground states for the original Skyrme model in [17] and [7] and for the Faddeev model in [14] and [20].

In this subsection we present our results establishing the existence of minimizers and compactness of the set of minimizers for the Skyrme and Faddeev functionals. Along with the energy functionals (3.1) and (3.2) for maps, we consider functionals for connections. In [29], Skyrme noticed and used the fact that maps could be represented by flat connections, $a = u^{-1} du$. This leads to a second version of the Skyrme functional,

$$(3.4) \quad E[a] = \int_M \frac{1}{2} |a|^2 + \frac{1}{16} |[a, a]|^2 d\operatorname{vol}_M.$$

The Faddeev functional can also be written in terms of flat connections; see [2]. For any smooth reference map $\varphi : M \rightarrow S^2$, any ψ homotopic to φ can be represented as $\psi = u\varphi u^*$ with $u : M \rightarrow \operatorname{Sp}(1)$. Plugging this expression into the energy functional, we obtain

$$(3.5) \quad E(\psi) = E_\varphi[a] := \int_M |D_a \varphi|^2 + |D_a \varphi \wedge D_a \varphi|^2,$$

where $D_a \varphi = d\varphi + [a, \varphi]$. As was observed in [21, 1, 2], the variational problems for maps and for connections are, in general, different. The way we approach these variational problems requires several steps; see [1, 2]. First, we describe the homotopy classes for smooth maps and analytically define expressions to label different classes. Second, we show that those expressions for labels of the components extend to finite energy Sobolev maps (or connections). Third, we show that minimizers exist in every sector of finite energy maps specified by given values of the label. We also prove in [3] that the values

these labels take on finite energy configurations agree with the possible values on smooth configurations.

There is an open question remaining, related to the components of the spaces $W_E^{1,2}$: the set of finite energy maps (or connections) with a fixed label is not known to be path connected. For this reason, we call such sets sectors rather than components. We expect that the sectors are the components of the spaces of finite energy maps when a suitable topology is used.

The existence of minimizers is established by the direct method of the calculus of variations; for any given sequence of maps in a fixed sector with energy approaching the infimum, there is a subsequence converging to a minimizer in the same sector. In particular, any sequence of minimizers in a fixed sector contains a convergent subsequence, so that the set of minimizers of the Skyrme or Faddeev functional in a fixed sector is compact in $W_E^{1,2}$.

Future work will address the existence of minimizers of a generalized Skyrme functional for maps into general homogenous spaces. We also plan to study the regularity of the limiting maps. It is common for minimizers of a functional in a class of functions to be more regular than generic members of the class. We expect to see applications of the local representation results for functions to the regularity of minimizers.

One may speculate that the minimizers of the Skyrme functional may form a cycle in the configuration space of maps from M to G . Given such a cycle, one could evaluate cohomology classes from the configuration space to obtain topological invariants. We believe that this is slightly naïve, because the Skyrme functional lacks an important property shared by the functionals that do serve to define invariants of manifolds. We recall that the sufficiently regular minimizers of functionals similar to the Skyrme functional satisfy second order Euler–Lagrange equations. When a functional is topologically saturated, the minimizers are characterized as the solutions of a first order system of equations. For example, the anti-self-dual condition in Yang–Mills theory, the pseudo-holomorphic condition in Gromov–Witten theory, or the Seiberg–Witten equations in Seiberg–Witten theory are all examples of first order systems characterizing minima. One can write out what the first order equations corresponding to the Skyrme functional should be. However, these equations are overdetermined and, essentially, have no solutions. The equations are analogous to what one would have by considering the Seiberg–Witten equations for a fixed connection and attempting to solve for the spin field. We believe that a suitable modification of the Skyrme functional incorporating new fields (perhaps gauge bosons?) will be topologically saturated and may lead to new smooth invariants of manifolds.

REFERENCES

- [1] D. Auckly and L. Kapitanski, *Holonomy and Skyrme’s model*, Comm. Math. Phys. **240** (2003), 97–122. MR2004981 (2005g:58033)
- [2] ———, *Analysis of S^2 -valued maps and Faddeev’s model*, Comm. Math. Phys. **256** (2005), 611–620. MR2161273 (2006e:58019)
- [3] ———, *The Pontryagin–Hopf invariants for Sobolev maps* (submitted).
- [4] ———, *Integrality of homotopy invariants in the Skyrme model* (in preparation).
- [5] D. Auckly and J. M. Speight, *Fermionic quantization and configuration spaces for the Skyrme and Faddeev–Hopf models*, Comm. Math. Phys. **263** (2006), no. 1, 173–216. MR2207327
- [6] A. Balachandran, G. Marmo, B. Skagerstam, and A. Stern, *Classical topology and quantum states*, World Sci. Publ. Co., Inc., River Edge, NJ, 1991. MR1254876 (95b:81001)
- [7] R. A. Battye and P. M. Sutcliffe, *Skyrmions, fullerenes and rational maps*, Rev. Math. Phys. **14** (2002), 29–85. MR1877915 (2003b:81227)
- [8] F. Bopp and Z. Haag, *Über die Möglichkeit von Spinmodellen*, Z. Naturforschung **5a** (1950), 644–653. MR0043309 (13:239e)

- [9] R. Bott and L. W. Tu, *Differential forms in algebraic topology*, Grad. Texts in Math., vol. 82, Springer-Verlag, New York–Berlin, 1982. MR0658304 (83i:57016)
- [10] H. Brezis, *The interplay between analysis and topology in some nonlinear PDE problems*, Bull. Amer. Math. Soc. (N.S.) **40** (2003), 179–201. MR1962295 (2004a:35038)
- [11] P. Dirac, *The theory of magnetic poles*, Phys. Rev. (2) **74** (1948), 817–830. MR0027713 (10:345a)
- [12] L. D. Faddeev, *Quantization of solitons*, Preprint IAS print-75-QS70 (1975).
- [13] ———, *Knotted solitons and their physical applications*, Topological Methods in the Physical Sciences (London, 2000), R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **359** (2001), 1399–1403. MR1853628 (2002g:81091)
- [14] L. D. Faddeev and A. J. Niemi, *Stable knot-like structures in classical field theory*, Nature **387** (1997), 58–61.
- [15] H. Federer, *A study of function spaces by spectral sequences*, Trans. Amer. Math. Soc. **82** (1956), 340–361. MR0079265 (18:59b)
- [16] D. Finkelstein and J. Rubinstein, *Connection between spin, statistics, and kinks*, J. Math. Phys. **9** (1968), 1762–1779. MR0234695 (38:3011)
- [17] T. Gisiger and M. B. Paranjape, *Recent mathematical developments in the Skyrme model*, Phys. Rep. **306** (1998), no. 3, 109–211. MR1659576 (99h:81129)
- [18] D. Giulini, *On the possibility of spinorial quantization in the Skyrme model*, Modern Phys. Lett. A **8** (1993), 1917–1924. MR1228374 (94k:81166)
- [19] V. L. Hansen, <http://www.mat.dtu.dk/people/V.L.Hansen/string.html>
- [20] J. Hietarinta and P. Salo, *Ground state in the Faddeev-Skyrme model*, Phys. Rev. D **62** (2000), 081701(R).
- [21] L. Kapitanski, *On Skyrme’s model*, Nonlinear Problems in Mathematical Physics and Related Topics, II, Int. Math. Ser. (N. Y.), vol. 2, Kluwer/Plenum, New York, 2002, pp. 229–241. MR1971999
- [22] R. C. Kirby, *The topology of 4-manifolds*, Lecture Notes in Math., vol. 1374, Springer-Verlag, Berlin, 1989. MR1001966 (90j:57012)
- [23] S. Koshkin, *Homogeneous spaces and Faddeev-Skyrme models*, Dissertation, Kansas State Univ., 2006.
- [24] F. Mandl and G. Shaw, *Quantum field theory*, Wiley, 1991.
- [25] L. Pontryagin, *A classification of mappings of the three-dimensional complex into the two-dimensional sphere*, Mat. Sb. **9** (51) (1941), no. 2, 331–363. (English) MR0004780 (3:60g)
- [26] R. Rajaraman, *Solitons and instantons*, North-Holland, Amsterdam, 1982. MR0719693 (85i:81002)
- [27] L. Schulman, *A path integral for spin*, Phys. Rev. (2) **176** (1968), 1558–1569. MR0237149 (38:5440)
- [28] D. J. Simms and N. M. J. Woodhouse, *Lectures in geometric quantization*, Lecture Notes in Phys., vol. 53, Springer-Verlag, Berlin–New York, 1976. MR0672639 (58:32470)
- [29] T. H. R. Skyrme, *A unified field theory of mesons and baryons*, Nuclear Phys. **31** (1962), 556–569. MR0138394 (25:1841)
- [30] R. Sorkin, *A general relation between kink-exchange and kink-rotation*, Comm. Math. Phys. **115** (1988), 421–434. MR0931669 (89e:81173)
- [31] A. F. Vakulenko and L. V. Kapitanskiĭ, *Stability of solitons in S^2 of a nonlinear σ -model*, Dokl. Akad. Nauk SSSR **246** (1979), no. 4, 840–842. (Russian) MR0543541 (80i:81045)
- [32] S. Weinberg, *The quantum theory of fields*. Vol. 1, Cambridge Univ. Press, Cambridge, 1996. MR1410064 (98a:81107a)
- [33] B. White, *Homotopy classes in Sobolev spaces and the existence of energy minimizing maps*, Acta Math. **160** (1988), 1–17. MR0926523 (89a:58031)
- [34] E. Witten, *Current algebra, baryons, and quark confinement*, Nuclear Phys. B **233** (1983), 433–444. MR0717916 (86i:81104)
- [35] F. Zaccaria, E. Sudarshan, J. Nilsson, N. Mukunda, G. Marmo, and A. Balachandran, *Universal unfolding of Hamiltonian systems: From symplectic structure to fiber bundles*, Phys. Rev. D (3) **27** (1983), 2327–2340. MR0704119 (85e:58052)

DEPARTMENT OF MATHEMATICS, KANSAS STATE UNIVERSITY, MANHATTAN, KANSAS 66506

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MIAMI, CORAL GABELS, FLORIDA 33124

SCHOOL OF MATHEMATICS, UNIVERSITY OF LEEDS, LEEDS LS2 9JT, ENGLAND

Received 16/JUN/2005

Translated by THE AUTHORS