
Preface

This book is devoted to giving the foundations for a partial Morse theory of minimal surfaces in Riemannian manifolds. It is based upon lecture notes for graduate courses on “Topics in Differential Geometry”, given at the University of California, Santa Barbara, during the fall quarter of 2009 and again in the spring of 2014, but it also includes several topics not treated in these courses.

It might be helpful to start with a description of the goal of our presentation. Morse theory is concerned with the relationship between the critical points of a smooth function on a manifold and the topology of that manifold. It might have developed in three main stages.

The first stage in our fictional history would have been finite-dimensional Morse theory, which relates critical points of proper nonnegative functions on finite-dimensional manifolds to the homology of these manifolds via the Morse inequalities. The foundations were laid in Marston Morse’s first landmark article on what is now known as Morse theory [Mor25], but Morse quickly turned his attention to problems from the calculus of variations, which ultimately became part of the infinite-dimensional theory. Nevertheless, finite-dimensional Morse theory became one of the primary tools for studying the topology of finite-dimensional manifolds and had many successes, including the celebrated h -cobordism theorem of Smale [Mil65], crucial for the classification of manifolds in high dimensions. Modern expositions of finite-dimensional Morse theory often construct a chain complex from the free abelian group generated by the critical points of a “generic” function, the boundary being defined by orbits of the gradient flow which connect the critical points. The homology of this chain complex, called the

Morse-Witten complex, is isomorphic to the usual integer homology of the manifold.

What might have been the second stage, the Morse theory of geodesics, formed the core of what Morse [Mor34] called “the calculus of variations in the large”. Motivated to some extent by earlier work on celestial mechanics by Poincaré and Birkhoff, Morse studied the calculus of variations for the length function or action function on the space of paths joining two points in a Riemannian manifold (or the space of closed paths in a Riemannian manifold), the critical points of these functions being geodesics. His idea was to approximate the infinite-dimensional space of paths by a finite-dimensional manifold of very high dimension and then apply finite-dimensional Morse theory to this approximation. This approach is explained in Milnor’s classical book on Morse theory [Mil63] and produced many striking results in the theory of geodesics in Riemannian geometry, such as the theorem of Serre [Ser51] that any two points on a compact Riemannian manifold can be joined by infinitely many geodesics. It also provided the first proof of the Bott periodicity theorem from homotopy theory. A Morse theory of closed geodesics, representing periodic motion in certain mechanical systems, was also constructed. One might regard the Morse theory of geodesics as an application of topology to the study of ordinary differential equations, in particular, to those equations which like the equation for geodesics arise from classical mechanics.

Palais and Smale were able to provide an elegant reformulation of the Morse theory of geodesics in the language of infinite-dimensional Hilbert manifolds [PS64]. They showed that the action function on the infinite-dimensional manifold of paths satisfies “Condition C”, a condition replacing “proper” in the finite-dimensional theory, and they showed that Condition C is sufficient for the development of Morse theory in infinite dimensions. This became a standard approach to the Morse theory of geodesics during the last several decades of the twentieth century, and we will describe it in some detail later.

One might regard the third stage of Morse theory as encompassing many strands, but our viewpoint is to focus on techniques for applying Morse theory to nonlinear elliptic partial differential equations coming from the calculus of variations in which the domain is a two-dimensional compact surface. Morse himself hoped to apply the ideas of his theory to what is arguably the central case—the partial differential equations for minimal surfaces—and he focused on the case in which the ambient space is Euclidean space. The first steps in this direction were taken by Morse and Tompkins, as well as Shiffman, who established the theorem that if a simple closed curve in Euclidean space \mathbb{R}^3 bounds two stable minimal disks, it bounds a

third, which is not stable. This provided a version of the so-called “mountain pass lemma” for minimal disks in Euclidean space. The results of Morse, Tompkins, and Shiffman suggested that Morse inequalities should hold for minimal surfaces in Euclidean space with boundary constrained to lie on a given Jordan curve and, indeed, such inequalities were later established under appropriate hypotheses (as explained, for example, in [JS90]).

But the most natural extension of the Morse theory of closed geodesics in Riemannian manifolds to partial differential equations would be a Morse theory of closed two-dimensional minimal surfaces in a general curved ambient Riemannian manifold M , instead of the ambient Euclidean space used in the classical theory of minimal surfaces with boundary. The generalization to ambient Riemannian manifolds with arbitrary curvature introduces complexity and requires new techniques. Unfortunately, if Σ is a connected compact surface, it becomes awkward to extend the finite-dimensional approximation procedure—so effective in the theory of geodesics—to the space $\text{Map}(\Sigma, M)$ of mappings from Σ to M . One might hope for a better approach based upon the theory of infinite-dimensional manifolds, as developed by Palais and Smale, but a serious problem is encountered: the standard energy function

$$E : \text{Map}(\Sigma, M) \longrightarrow \mathbb{R},$$

used in the theory of harmonic maps and parametrized minimal surfaces, fails to satisfy the compactness condition, the Condition C which Palais and Smale had used as a foundation for their theory, when $\text{Map}(\Sigma, M)$ is completed with respect to a norm strong enough to lie within the space of continuous functions.

To get around this difficulty, Sacks and Uhlenbeck introduced the α -energy [SU81], [SU82] in which $\alpha > 1$ is a parameter, a perturbation of the usual energy which is defined on the completion of $\text{Map}(\Sigma, M)$ with respect to a Banach space norm which is both weak enough to satisfy Condition C and strong enough for the α -energy to be a C^2 function on a Banach space having the same homotopy type as the space of continuous maps from Σ to M . We can express the fact that such a completion exists by saying that the α -energy lies within “Sobolev range”. The α -energy approaches the usual energy as the parameter α in the perturbation goes to one, and we can therefore say that the usual energy on maps from compact surfaces is “on the border of Sobolev range”, approachable by Morse theory via approximation. Using this approximation via the α -energy, Sacks and Uhlenbeck were able to establish many striking results in the theory of minimal surfaces in Riemannian manifolds, including the fact that any compact simply connected Riemannian manifold contains at least one nonconstant minimal two-sphere, which parallels the classical theorem of Fet and Lyusternik stating that any

compact Riemannian manifold contains at least one smooth closed geodesic. But they also discovered a serious obstruction: The phenomenon of “bubbling” as $\alpha \rightarrow 1$ prevents full Morse inequalities from holding for compact parametrized minimal surfaces in $\text{Map}(\Sigma, M)$ in complete generality.

A somewhat different approach to existence of parametrized minimal surfaces in Riemannian manifolds was developed at about the same time by Schoen and Yau [SY79], using Morrey’s solution to the Plateau problem for minimal disks bounded by a Jordan curve in a Riemannian manifold and arguments based upon a “replacement procedure”. Their approach provided striking theorems relating positive scalar curvature to the topology of three-manifolds, including a step toward the first proof of the positive mass theorem of general relativity. The Schoen-Yau replacement procedure can also be used to obtain many of the existence results of Sacks and Uhlenbeck, and, indeed, an alternate treatment of many of their theorems is provided by Jost [Jos91]. Yet other proofs of these theorems were developed using heat flow (see Struwe [Str90] or Hang and Lin [HL03]).

An even more general approach to minimal surfaces is provided by geometric measure theory [Mor08], which constructs minimal varieties of arbitrary dimension and codimension via generalized surfaces such as integral currents and varifolds, and then attempts to show that the resulting generalized surfaces are regular. But regularity cannot always succeed because according to results of René Thom, not every homology class is represented by a smooth submanifold [Tho54] and much work has focused on the special dimensions in which regularity can be established. In this book we restrict our attention to minimal surfaces of dimension two for which a parametrized theory often provides additional information not directly accessible via geometric measure theory, such as the genus of the surface, and this additional information is crucial for many of the applications we will describe. Moreover, the Sacks-Uhlenbeck perturbation is available in the mapping context, and it provides a suggestive link with classical Morse theory, and perhaps the clearest insight into bubbling, the phenomenon observed as the perturbation is turned off, which turns out to be the main technical difficulty in establishing Morse inequalities for (two-dimensional) closed minimal surfaces in Riemannian manifolds. Moreover, analogs of bubbling appear in the study of other nonlinear partial differential equations, such as the Yang-Mills equation on four-dimensional manifolds, suggesting links between theories that can sometimes be exploited.

Indeed, the theory of two-dimensional minimal surfaces can be placed in a broader context—that of nonlinear partial differential equations which lie

on the border of Sobolev range and are conformally invariant or closely related to conformally invariant equations. These equations include the Yang-Mills equations on four-dimensional manifolds from the standard model for particle physics, the anti-self-dual equations used so effectively by Donaldson, the Seiberg-Witten equations, and Gromov's equations for pseudoholomorphic curves. It is useful to develop a common technology for studying these equations and a common schema for setting up theories: First one needs to develop a transversality theory using Smale's generalization of Sard's theorem from finite-dimensional differential topology. This generally shows that in generic situations, solutions to the nonlinear partial differential equation form a finite-dimensional submanifold of an infinite-dimensional function space. The tangent space to this submanifold is studied via the linearization of the nonlinear partial differential equation at a given solution, and often the dimension of the tangent space is obtained by application of the Atiyah-Singer index theorem (which reduces to the Riemann-Roch theorem in the case of parametrized minimal surfaces). Next one develops a suitable compactness theorem, which in Donaldson's theory provides solutions localized near points, in analogy with bubbling. Finally, one uses topological and geometric methods to derive important geometric conclusions (for example, existence of minimal surfaces under various topological conditions) or to construct differential topological invariants of manifolds (in the Seiberg-Witten or Donaldson theories). The reader can refer to [DK90] for a definitive treatment of Donaldson theory, which might be regarded as a model for the other theories.

Although bubbling implies that the most obvious extension of the Morse inequalities to minimal surfaces in Riemannian manifolds cannot hold in general, it also suggests a framework for analyzing how the Morse inequalities fail and to what extent a remnant of the Morse inequalities might still be retained. Of course, other difficulties also need to be controlled. When constructing a minimax critical point, one must allow for variations in the conformal structure on the surface, thought of as an element of Teichmüller space or moduli space, and a sequence of harmonic maps may degenerate as the conformal structure moves to the boundary of moduli space. Moreover, branched coverings of a given minimal surface count as new critical points within the space of functions, although they are not geometrically distinct from the covered surface. We call these the "sources of noncompactness" for the two-variable energy

$$E : \text{Map}(\Sigma, M) \times \mathcal{T} \longrightarrow \mathbb{R},$$

in which \mathcal{T} represents Teichmüller space. Note that this energy is invariant under various groups, an action of the group of complex automorphisms of

the domain Σ and an action of the mapping class group on Teichmüller space, suggesting use of equivariant Morse theory as described by Bott [Bot82].

One might suspect that a procedure for constructing minimal surfaces that might fail in several different ways is too flawed to be of much use. However, we argue that a different perspective is more productive—since the minimax procedure for a given homology class does not always yield nontrivial minimal surfaces, one should divide homology classes into various categories depending upon which of the possible difficulties can arise. This approach appears ambitious at first, but then one realizes that the topology of $\text{Map}(\Sigma, M)$ has a rich internal structure that can compensate for the many ways in which minimax sequences might fail to converge.

Our purpose here is to provide the foundation for a theory which might explain when sequences of α -energy critical points for minimax constraints converge to minimal surfaces without bubbling as $\alpha \rightarrow 1$, a theory which should ultimately have important applications similar to those found in the Morse theory of geodesics.

It has long been known that the extension of Morse theory to infinite-dimensional manifolds is not really necessary for the study of geodesics. Bott expressed it this way in his 1982 survey on Morse theory [Bot82]: “I know of no aspect of the geodesic question where [the infinite-dimensional approach] is essential; however it clearly has some aesthetic advantages, and points the way for situations where finite-dimensional approximations are not possible....” On the other hand, finite-dimensional approximations suitable for studying the α -energy when $\alpha > 1$ appear to break down as $\alpha \rightarrow 1$. This suggests that for any partial Morse theory of minimal surfaces, in contrast to closed geodesics, calculus on infinite-dimensional manifolds should play an essential conceptual and simplifying role.

Our presentation in this book therefore starts with an extension of finite-dimensional calculus to calculus on infinite-dimensional manifolds. We assume the reader is familiar with the basics of finite-dimensional differential geometry, including geodesics, curvature, and the tubular neighborhood theorem. We also assume some familiarity with basic complex analysis, the foundations of Banach and Hilbert space theory, and the willingness to accept results from the linear theory of elliptic partial differential operators, in particular, the theory of Fredholm operators on Sobolev spaces. All of these topics are treated very well in highly accessible sources, to which we can refer at the appropriate time.

The first chapter gives an introduction to global analysis on infinite-dimensional manifolds of maps. The second chapter studies the theory of geodesics on Riemannian manifolds using the action integral

$$J : \text{Map}(S^1, M) \rightarrow \mathbb{R}$$

and owes much to the beautiful work of Bott, Gromoll, Klingenberg, and Meyer. This leads to the question of determining the topological invariants of the space $\text{Map}(S^1, M)$ and motivates the exposition in Chapter 3 of Sullivan's theory of minimal models, which provides an efficient means of calculating rational cohomology of $\text{Map}(S^1, M)$. Sullivan and Vigué-Poirrier were able to use this theory to provide a striking extension of an earlier theorem of Gromoll and Meyer: Most compact smooth manifolds have infinitely many geometrically distinct smooth closed geodesics for arbitrary choice of Riemannian metric. More generally, Sullivan's theory provides an algorithm for calculating the rational cohomology of $\text{Map}(\Sigma, M)$, when Σ is a compact manifold of arbitrary dimension and M is a nilpotent manifold.

In Chapter 4, we turn to the theory of minimal surfaces in Riemannian manifolds and discuss well-known theorems of Sacks, Uhlenbeck, Schoen, and Yau which use minimal surfaces to elucidate the topology of three-dimensional manifolds. We also describe Uhlenbeck's Morse theory for the α -energy and its perturbations, which provide examples of energy functions for which complete Morse inequalities hold, relating critical points to rational cohomology of $\text{Map}(\Sigma, M)$, when Σ is a compact smooth surface.

Just as a generic choice of proper function $f : M \rightarrow \mathbb{R}$ on a smooth manifold M is necessary for Morse inequalities on finite-dimensional manifolds, so a generic choice of Riemannian metric on a compact manifold M is needed to simplify the relationships between the topology of M and the types of minimal surfaces within M . Chapter 5 gives a proof of the Bumpy Metric Theorem from [Moo06], which describes the generic behavior: It states that for generic choice of Riemannian metric on a smooth manifold M of dimension at least three, closed prime minimal surfaces $f : \Sigma \rightarrow M$ have no branch points and are as nondegenerate as allowed by their invariance under the group of complex automorphisms of Σ . This is complemented by a transversality theory which in accordance with Whitney's theorems implies that for generic metrics minimal surfaces $f : \Sigma \rightarrow M$ are imbeddings when M has dimension at least five and immersions with transverse crossings when M has dimension four.

We believe that the Bumpy Metric Theorem will be useful in establishing partial Morse inequalities for parametrized minimal surfaces of given genus when the dimension of the ambient manifold is sufficiently large. Such partial Morse inequalities require additional techniques not treated in this book, and we hope to return to this topic in a later publication. We intend to apply the Bumpy Metric Theorem to area-minimizing minimal surfaces in four-manifolds with generic metrics in a subsequent article [Moox], using the twistor degrees studied by Eells and Salamon [ES64].

We have left out many topics from the theory of minimal surfaces which are treated very well in existing sources. Readers unfamiliar with the rich theory of minimal surfaces in Euclidean space \mathbb{R}^3 should seek other sources, such as the excellent books of Osserman [O69] and Colding and Minnicozzi [CM11].

I would like to thank the students in my classes, and in particular Casey Blacker, Robert Ream, and Changliang Wang, for questions and remarks that helped correct minor errors and smooth out the exposition in several places. I am also grateful to anonymous readers of earlier versions of this book for their helpful suggestions and to Xianzhe Dai for his encouragement.

Doug Moore

Santa Barbara, CA, June 2017