The Three Stages of Morse Theory

John Douglas Moore
Department of Mathematics
University of California
Santa Barbara, CA, USA 93106
e-mail: moore@math.ucsb.edu

During the twentieth century, Morse theory should have developed in three main stages:

1. Critical point theory for finite-dimensional manifolds.

2. Morse theory for ordinary differential equations (geodesics).

3. Morse theory for partial differential equations (minimal surfaces).

(However, event transpired somewhat differently.)

The simplest case of Morse theory for finite-dimensional manifolds is expressed by the so-called mountain pass lemma.

It states: Suppose that all the critical points of a smooth proper function $f : \mathbb{R}^2 \to [0, \infty)$ are nondegenerate. If f has two local minima (lake bottoms), it must also have at least one critical point of index one (mountain pass).

Idea behind the proof of the mountain pass lemma: We consider the gradient of f,

$$\nabla f \quad \text{defined by} \quad \langle \nabla f, v \rangle = df(v),$$

for all tangent vectors v. The vector field $X = -\nabla f$ has a corresponding one-parameter group $\{\phi_t : t \in \mathbb{R}\}$ of local diffeomorphisms of $\mathbb{R}^2$.

One can try to find local minima by the method of steepest descent, that is, by following flow-lines $p \mapsto \phi_t(p)$ for the vector field X.

Suppose now that p and q are distinct local minima for f, and consider

$$\mathcal{F} = \{\gamma : [0,1] \to \mathbb{R}^2 | \gamma(0) = p, \gamma(1) = q\}.$$

Since $\mathbb{R}^2$ is connected there is at least one element $\gamma \in \mathcal{F}$. Applying ϕ_t to this path gives a collection of paths

$$\phi_t \circ \gamma : [0,1] \longrightarrow \mathbb{R}.$$

Take a sequence $t_i \to \infty$, and let $\phi_{t_i} \circ \gamma(s_i)$ be a point in the image $\phi_{t_i} \circ \gamma([0,1])$ at which f assumes its maximum value. We can assume that s_i converges to some point $s_\infty \in (0,1)$. We claim that $r_i = \gamma_i(s_\infty)$ converges to a critical point of index one.

Indeed,

$$X(r_i) = -\nabla f(r_i) \to 0,$$

since otherwise ϕ_t would push the value of f down at a rate of speed bounded below by a positive number, and eventually the value of f would become negative.

Thus we have a sequence $\{r_i\}$ which has the properties:

1. $f(r_i)$ is bounded.

2. $\nabla f(r_i) \to 0$.

Since f is proper, a subsequence of the r_i's converges. The limit r must be a critical point since $\nabla f(r_i) \to 0$. (We leave it to the reader to construct that argument that it must have index one.)

What is needed for this argument to work? Clearly, we can replace properness of f by the following criterion: The function $f : \mathbb{R}^2 \to \mathbb{R}$ satisfies *condition C* if whenever $\{r_i\}$ is a sequence such that

1. $f(r_i)$ is bounded, and

2. $\nabla f(r_i) \to 0$,

then $\{r_i\}$ possesses a convergent subsequence.

The argument can then be carried through in many cases when $\mathbb{R}^2$ is replaced by a complete Riemannian manifold, perhaps even infinite-dimensional.

Finite-dimensional Morse theory in a nutshell: Let M be a complete finite-dimensional Riemannian manifold, $f : M \to [0, \infty)$ a function which satisfies condition C. Then

- For a generic perturbation of f, we can arrange that all the critical points of f are nondegenerate.

- Condition C implies that the number of critical points with $f \leq c$ is finite.

- The critical points form the basis for a chain complex from which one can calculate the homology of M.

(This last point has been emphasized in Witten's work on supersymmetry and Morse theory, but it was implicit also in Milnor's *Lectures on the h-cobordism theorem*.)

From the last fact, one obtains the Morse inequalities, which include the generalized mountain pass lemma: If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a nonnegative proper function whose critical points are nondegenerate, then

$$\begin{aligned}&(\text{number of mountain passes})\\&\qquad\geq (\text{number of lake bottoms}) - 1.\end{aligned}$$

More generally, the Morse inequalities for a generic function $f : M \to [0, \infty)$ satisfying condition C imply that

$$(\text{number of critical points of index } \lambda) \geq \dim H_\lambda(M; F),$$

when F is any field. As described in Milnor's *Morse theory*, these inequalities have many applications to submanifold theory.

Recall the beautiful resolution of the Weyl problem by Pogorelov, Nirenberg, ... Any Riemannian metric of positive curvature on S^2 can be realized by a unique isometric imbedding.

For Riemannian metrics on $\mathbb{R}P^2$, the story is quite different. If $f : \mathbb{R}P^2 \to \mathbb{R}^3$ is any immersion, then a height function $h : \mathbb{R}P^2 \to \mathbb{R}$ (normalized to be nonnegative) must have a critical point of index one, because $H_2(\mathbb{R}P^2; \mathbb{Z}_2) \neq 0$. Such a critical point must be a point of nonpositive curvature. So the standard metric of constant positive curvature on $\mathbb{R}P^2$ cannot be realized by an immersion into $\mathbb{R}^3$.

Question: If M is an n-dimensional manifold with a metric of positive sectional curvatures which admits an isometric immersion into $\mathbb{R}^{2n-1}$, must M be homeomorphic to a sphere?

Yes, if $n = 2$ by above argument and if $n = 3$ by a Morse theory argument in Proc. AMS, vol. 70 (1978), pages 72-74. Morse theory on finite-dimensional manifolds is a natural technique to use on this problem.

The simplest second-order linear ordinary differential equation is

$$\gamma''(t) = 0, \quad \text{with solutions} \quad \gamma(t) = at + b.$$

Here we can assume that γ is vector-valued, that is, that it takes its values in $\mathbb{R}^N$.

The simplest way to make this equation nonlinear is to imagine that

$$\gamma : [0,1] \longrightarrow M \subset \mathbb{R}^N, \quad \gamma(0) = p, \gamma(1) = q,$$

and look for solutions to

$$(\gamma''(t))^{\top} = 0,$$

where $(\cdot)^{\top}$ represents orthogonal projection into the tangent space.

Solutions are called *geodesics* from p to q. The theory of geodesics formed the kernel of Morse's calculus of variations in the large.

Variational formulation: We let

$$\Omega(M; p, q) = \{\gamma : [0, 1] \to M | \gamma(0) = p, \gamma(1) = q\},$$

and define the action function $J : \Omega(M, p, q) \to \mathbb{R}$ by

$$J(\gamma) = \frac{1}{2}\int_0^1 |\gamma'(t)|^2 dt.$$

The Euler-Lagrange equations for J are just the geodesic equations.

Morse studied critical points of J using finite-dimensional approximations to the infinite-dimensional space $\Omega(M; p, q)$. Morse theory of geodesics in a nutshell: Let M be a complete finite-dimensional Riemannian manifold.

- For a generic choice of p and q, all the critical points of J are nondegenerate.

- Condition C implies that the number of critical points with $J(f) \leq c$ is finite.

- The critical points form the basis for a chain complex from which one can calculate the homology of $\Omega(M, p, q)$.

In particular, the last fact yields the Morse inequalities,

$$(\text{number index } \lambda \text{ critical points of } J) \geq \dim H_\lambda(\Omega(M;p,q);F),$$

when F is any field. In his thesis, Serre showed that if M is compact, then $\Omega(M;p,q)$ has non-vanishing H_λ for λ arbitrarily large. This plus the Morse inequalities implied: If M is a compact Riemannian manifold, then any two generic points on M can be connected by infinitely many geodesics.

Conclusion: One can use algebraic topology to prove existence of solutions to ordinary differential equations.

Palais and Smale found a beautiful reformulation of Morse's theory, in which one regards a suitable completion of $\Omega(M; p, q)$ as an infinite-dimensional Hilbert manifold with a complete Riemannian metric, such that

$$J : \Omega(M; p, q) \to \mathbb{R}$$

satisfies condition C.

One can also consider the periodic case, in which

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

is defined by

$$J(\gamma) = \frac{1}{2}\int_{S^1} |\gamma'(t)|^2 dt.$$

The solutions to the Euler-Lagrange equations in this case are periodic geodesics. One can ask: Is it true that any compact manifold with finite fundamental group must contain infinitely many geometrically distinct smooth closed geodesics (Klingenberg)? Note that multiple covers of a single prime geodesics should not be considered to be geometrically distinct. The question has been answered in many cases.

A difficulty in the Morse theory of periodic geodesics is that critical points for

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

can never be nondegenerate in the usual sense, because J is preserved by the group action

$$\phi : \mathrm{Map}(S^1, M) \times S^1 \to \mathrm{Map}(S^1, M),$$

$$\phi(\gamma, s)(t) = \gamma(s + t).$$

Morse theory of periodic geodesics in a nutshell: Let M be a complete finite-dimensional Riemannian manifold.

- For a generic choice of metric on M, all nonconstant geodesics lie on one-dimensional nondegenerate critical submanifolds.

- Condition C implies that the number of such submanifolds with $J(f) \leq c$ is finite.

- The critical submanifolds form the basis for an equivariant chain complex from which one can calculate the homology of Map(S^1, M).

- In particular, one obtains equivariant Morse inequalities.

The simplest second-order linear partial differential equation of elliptic type is

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0,$$

and its solutions are known as *harmonic functions*. Here we can assume that f is vector-valued, that is, that it takes its values in $\mathbb{R}^N$.

The simplest way to make this equation non-linear is to imagine that

$$f : \mathbb{R}^2 \longrightarrow M \subset \mathbb{R}^N,$$

and look for solutions to

$$\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right)^\top = 0,$$

where $(\cdot)^\top$ represents orthogonal projection into the tangent space.

Solutions are called *harmonic maps* from $\mathbb{R}^2$ into M. We can imagine imposing boundary conditions or replacing $\mathbb{R}^2$ by a compact Riemann surface Σ.

We believe the case where the domain is a compact Riemann surface yields the simplest and most beautiful theory. Here is the variational formulation: Consider the energy

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

which is defined by

$$E(f, \omega) = \frac{1}{2} \int_{\Sigma} |df|^2 dA.$$

Here Σ is a Riemann surface of a given genus g and $\mathcal{T}$ is the Teichmüller space of conformal structure on Σ. The norm of df and the area element on Σ are calculated with respect to any Riemannian metric within the conformal equivalence class selected by ω. (It turns out that E does not depend on the choice of metric on Σ, only its conformal equivalence class.)

The critical points for E are not just harmonic, but also conformal. Thus they are in fact *minimal surfaces*.

One would like to develop a Morse theory for E because it should have important applications to minimal surface theory.

Question: Given a choice of genus g, what are the conditions on the topology of a smooth compact manifold M with finite fundamental group which ensure that for generic choice of Riemannian metric on M, there are infinitely many geometrically distinct minimal surfaces of genus g?

Just as in Serre's thesis, a Morse theory for E might well provide an avenue whereby the algebraic topology of $\text{Map}(\Sigma, M)$ can yield information on the solution to nonlinear partial differential equations.

(Bott, 1980, Bulletin AMS)

Marston Morse had developed the abtract setting of the variational theory ... in large part because he hoped to make it applicable to minimal surface theory and other variational problems. Unfortunately, however, a direct extension of the Morse Theory just does not work for variational problems in more than one variable.... In the context of the Palais-Smale theory, one understands this phenomenon in terms of the Sobolev inequalities, which show that the conditions on a map ... to have finite area are far from forcing it to be continuous....

Indeed, the function E determines a natural topology on the space of maps $\text{Map}(\Sigma, M)$, the so-called L_1^2 topology.

When Σ has dimension one, this topology is weakly homotopically equivalent to the compact-open topology familiar to topologists.

When Σ has dimension two, it just **barely** fails to lie within the Sobolev range that would make it homotopy equivalent to the usual space of continuous functions.

(Smale, 1977, Bulletin AMS)

In the theory of Plateau's problem, I had been intrigued by a result of Morse Tompkins and Schiffman in 1939. Their theorem asserted that if a Jordan curve in $\mathbb{R}^3$ spans two stable minimal surfaces, then it spans a third of unstable type. This was suggestive of a Morse theory for Plateau's problem . In the sixties, I tried without success to find such a theory, or to imbed the Morse-Tompkins-Schiffman result in a general conceptual setting. Tromba and Uhlenbeck may now have succeeded in initiating a development of calculus of variations in the larger for more than one independent variable.

What Uhlenbeck (in conjunction with Sacks) had discovered was that when the domain has dimension two—and only dimension two—there is a simple procedure for perturbing the energy function so that the corresponding completion of Map(Σ, M) does in fact lie within Sobolev range.

For each element in Teichmüller space $\mathcal{T}$, we give Σ a canonical metric in its conformal equivalence class. This metric is the constant curvature metric of total curvature one.

Following Sacks and Uhlenbeck, Annals of Math., 1981, we can then define the α-*energy*, for $\alpha > 1$. It is the function

$$E_\alpha : \text{Map}(\Sigma, M) \times \mathcal{T} \to \mathbb{R},$$

given by the formula

$$E_\alpha(f, \omega) = \frac{1}{2} \int_\Sigma [(1 + |df|^2)^\alpha - 1] dA.$$

It is then the case that for fixed choice of ω

$$E_{\alpha,\omega} : \text{Map}(\Sigma, M) \to \mathbb{R}, \qquad E_{\alpha,\omega}(f) = E_\alpha(f, \omega),$$

satisfies condition C in a suitable completion of $\text{Map}(\Sigma, M)$. Note that $E_\alpha \to E$ as $\alpha \to 1$.

There is a full Morse theory for suitable perturbations of $E_{\alpha,\omega}$ (so that all critical points are nondegenerate). However, as $\alpha \to 1$, sequences of α-energy critical points tend to bubble, showing that Morse inequalities cannot hold for E itself. On the other hand, a better understanding of bubbling might provide partial Morse inequalities.

Moreover, the partial Morse theory that does exist when energy is low has had some interesting applications, including the sphere theorem of Micallef and myself: If a compact simply connected Riemannian manifold of dimension at least four has positive curvature on isotropic two-planes, it must be homeomorphic to a sphere.

Recent results that form part of the foundation for a projected partial Morse theory for minimal tori:

- For a generic choice of metric on M, all prime minimal two-sphere, projective planes, tori and Klein bottles lie on nondegenerate critical submanifolds of the dimensions demanded by the group actions.

- If the dimension of the ambient manifold is at least five, there are only finitely many prime minimal two-spheres and projective planes below any given energy level. There are only finitely many minimal tori and Klein bottles with energy below a given bound and conformal structure in a given compact subset of Teichmüller space.

What are the next steps?

The above theorems need to be generalized to arbitrary genus. Once we show that the number of minimal surface configurations for a given genus are finite, we need to understand the boundary maps in the corresponding Morse complex. Although much remains to be done, the prospects for a Morse theory for minimal tori, at least, look bright....