A GENERALIZED MORSE THEORY

BY R. S. PALAIS AND S. SMALE

1. Abstract theory. Let M be a C^2-Riemannian manifold without boundary modeled on a separable Hilbert space (see Lang [3]). For $p \in M$ we denote by $\langle \ , \ \rangle_p$ the inner product in the tangent space M_p and we define a function $\| \ \|$ on the tangent bundle $T(M)$ by $\|v\| = \langle v, v \rangle_p^{1/2}$ for $v \in M_p$. Given p and q in the same component of M we define $\rho(p, q) = \mathrm{Inf} \int_0^1 \|\sigma'(t)\| dt$, where the Inf is over all C^1 paths $\sigma \colon [0, 1] \to M$ such that $\sigma(0) = p$ and $\sigma(1) = q$. Just as in the finite dimensional case one shows that ρ is a metric on each component of M which is consistent with the manifold topology. If each component of M is complete in this metric M is called a complete Riemannian manifold and *we assume this in all that follows*. Let $f \colon M \to R$ be a C^2 function. Then df, the differential of f, is a C^1 cross section of the cotangent bundle of M, hence there is a uniquely determined C^1 vector field ∇f on M, the gradient of f, such that $df_p(v) = \langle v, \nabla f(p) \rangle_p$ for $v \in M_p$. We denote by ϕ_t the maximum local one-parameter group generated by $-\nabla f$. A critical point of f is a point where ∇f vanishes; equivalently a stationary point of ϕ_t. At a critical point p of f there is a uniquely determined continuous bilinear form $H(f)_p$ on M_p, the Hessian of f at p, such that $H(f)_p(u, v) = d^2(f \circ \phi^{-1})(d\phi_p(u), d\phi_p(v))$ if ϕ is any chart at p. The supremum of the dimensions of subspaces on which $H(f)_p$ is negative (positive) definite is called the index (coindex) of f at p. $H(f)_p$ is symmetric, hence there is a uniquely determined bounded self-adjoint operator A on M_p such that $H(f)_p(u, v) = \langle Au, v \rangle_p$. The critical point p is called nondegenerate if A has a bounded inverse. In this case p is isolated in the set of critical points.

Let $f^{a,b} = f^{-1}[a, b]$ and $f^b = f^{-\infty,b}$. Morse theory is concerned with relating the structure of the critical point set of f in $f^{a,b}$ with the homology, homotopy, homeomorphism, and diffeomorphism type of the pair (f^b, f^a). We shall be concerned with the Morse theory of pairs (M, f) as above which satisfy at least the following extra condition:

(C) If S is a subset of M on which $|f|$ is bounded but on which $\|\nabla f\|$ is not bounded away from zero, then there is a critical point of f in the closure of S.

Note that if f is proper (which implies that M is finite dimensional) and in particular if M is compact then condition (C) is automatically satisfied. More interesting though is the fact, which we will make

precise in later sections, that a fairly general class of calculus of variations problems can be interpreted as pairs (M, f) satisfying condition (C). The following is thus an existence theorem for minimizing extremals of such problems.

THEOREM 1. *Let (M, f) satisfy (C) and assume f is bounded below on a component M_0 of M. Then f assumes its greatest lower bound on M_0. If $p \in M_0$ then $\phi_t(p)$ is defined for all positive t and $\phi_t(p)$ has a critical point of f as limit point as $t \to \infty$. If all the critical points of f on M_0 are nondegenerate then in fact $\lim_{t \to \infty} \phi_t(p)$ exists and is a critical point of f for each $p \in M_0$. If f is bounded below on all of M then f assumes its greatest lower bound on M provided the critical point set of f has no interior, i.e. provided f is not constant on any nonempty open set, hence in particular if all the critical points of f are nondegenerate.*

REMARK. In any reasonable calculus of variations problem it is a priori evident that the function cannot be anywhere locally constant, hence condition (C) implies the existence of absolute minima.

A real number c is called a critical value of f if $f^{-1}(c)$ contains a critical point of f, otherwise c is a regular value of f. If c is a regular value of f then it is easily seen that f^c is a closed submanifold of M with boundary $f^{-1}(c)$.

THEOREM 2. *Let (M, f) satisfy condition (C) and assume that all critical points of f are nondegenerate. Then*

(1) *For any real numbers $a < b$ there are only finitely many critical points of f in $f^{a,b}$, hence the critical values of f are isolated.*

(2) *Let a and b be regular values of f and suppose that among the critical points of f in $f^{a,b}$ there are r having finite index. Let the indices of these critical points be $d_1, \cdots, d_r$. Then f^b has the homotopy type of f^a with r cells of dimensions $d_1, \cdots, d_r$ attached.*

COROLLARY 1 (MORSE INEQUALITIES). *Let a and b be regular values of f. Let $R_i^{a,b}$ be the ith betti number of (f^b, f^a) with coefficients in a field and let $C_i^{a,b}$ be the number of critical points of index i in $f^{a,b}$. Then*

(1) $$\sum_{i=0}^{m} (-1)^{m-i} R_i^{a,b} \leqq \sum_{i=0}^{m} (-1)^{m-i} C_i^{a,b},$$

(2) $$R_i^{a,b} \leqq C_i^{a,b},$$

(3) $$\sum_{i=0}^{\infty} (-1)^i R_i^{a,b} = \sum_{i=0}^{\infty} (-1)^i C_i^{a,b}.$$

REMARK. Note that if f is bounded below then in the theorem and Corollary 1 we can take $a <$ greatest lower bound of f in which case

$f^a = \emptyset$ so $(f^b, f^a) = f^b$, $R_i^{a,b} = i$th betti number of f^b, and $C_i^{a,b} =$ number of critical points of f in f^b having index i. If we now let $b \to \infty$ through regular values we get

COROLLARY 2. *Assume f is bounded below and let R_i denote the ith betti number of M with coefficients in a field and C_i the number of critical points of f having index i (either or both may be infinite). Then $R_i \leqq C_i$ and if each C_i is finite $i = 0, \cdots, m$ then*

$$\sum_{i=0}^{m} (-1)^{m-i} R_i \leqq \sum_{i=0}^{m} (-1)^{m-i} C_i.$$

Assume the hypothesis of Theorem 2 and let $a < b$ be regular values of f and c the only critical value of f in $[a, b]$. Let $p_1, \cdots, p_r$ be all the critical points of f on $f^{-1}(c)$ and let k_i and l_i be respectively the index and coindex of f at p_i. Then one can show that f^b has the diffeomorphism type of f^a with handles of type $(k_1, l_1), \cdots, (k_r, l_r)$ attached (a handle of type (k, l) is $D^k \times D^l$ where D^r denotes the closed unit disc in a Hilbert space of dimension r, $0 \leqq r \leqq \infty$; what it means to attach a handle to a manifold with boundary we shall not make precise here other than to say it is strictly analogous to the well-known process in the finite dimensional case).

2. **Manifolds of maps.** The prototypes of theorems of this section are due to J. Eells [2].

Let M be a compact differentiable (C^∞) manifold possibly with boundary, ξ a real finite dimensional vector bundle over M with fiber E and $C^r(\xi)$ the vector space of C^r-cross sections of ξ. If we identify an open set $\mathcal{O}$ of M with an open set in a vector space V by a chart and trivialize ξ over $\mathcal{O}$, then every $f \in C^r(\xi)$ defines a C^r-map $\bar{f}: \mathcal{O} \to E$ and for each $p \in \mathcal{O}$ and $m = 0, 1, \cdots, r$ the mth differential for f at p is an element of the space $L_s^m(V, E)$ of symmetric m-linear maps of V into E. The set of $f \in C^r(\xi)$ such that $d^m \bar{f}_p = 0$ $m = 0, 1, \cdots, r$ is a vector subspace $Z_p^r(\xi)$ of $C^r(\xi)$ which does not depend on the chart or trivialization, and the quotient $J^r(\xi)_p = C^r(\xi)/Z_p^r(\xi)$ is a finite dimensional vector space isomorphic to $\oplus_{m=0}^{r} L_s^m(V, E)$. Then $J^r(\xi)_p$ is the fiber at p of a differentiable vector bundle $J^r(\xi)$ over M, the r-jet bundle of ξ, and there is an obvious one-to-one linear map $j_r: C^r(\xi) \to C^0(J^r(\xi))$ called the r-jet extension map. The compact open topology on $C^0(J^r(\xi))$ induces via the injection j_r a topology on $C^r(\xi)$ (the C^r-topology) which can be derived from a complete norm, and henceforth we regard $C^r(\xi)$ as a Banach space whose norm is given only to within equivalence. Let μ be a "smooth" measure on M (i.e. if $\phi: \mathcal{O} \to R^n$ is a chart, then there is a strictly positive C^∞ real valued

function f on $\phi(\mathfrak{O})$ such that for each subset E of $\mathfrak{O}$ $\mu(E) = \int_{\phi(E)} f(x) dx$ where dx is the element of Lebesgue measure in R^n). If we give $J^r(\xi)$ a Riemannian structure then $\langle f, g \rangle = \int \langle j_r(f)(x), j_r(g)(x) \rangle d\mu(x)$ defines a prehilbert space structure on $C^r(\xi)$ which changes to an equivalent one if we change μ or the Riemannian structure on $J^r(\xi)$. The completed space is the Sobolev space $H^r(\xi)$ of "H^r-cross-sections of ξ," a Hilbert space whose inner product is defined up to equivalence. Now let γ be a differentiable fiber bundle over M in the sense of Ehresmann (i.e. γ is a fiber bundle whose total space, fiber, and projection are differentiable and which is differentiably locally trivial) and suppose γ is a "closed sub-bundle of ξ" in the sense that the total space of γ is a closed submanifold of the total space of ξ and the projection of γ is the restriction of the projection of ξ. The space $C^r(\gamma)$ of C^r-cross sections of γ in the C^r topology is then a closed subspace of $C^r(\xi)$. Let $H^r(\gamma)$ denote the closure of $C^r(\gamma)$ in $H^r(\xi)$.

THEOREM 1. $C^r(\gamma)$ *is a closed* C^∞ *submanifold of the Banach space* $C^r(\xi)$. *If* $2r > \dim M$ *then* $H^r(\gamma)$ *is a closed* C^∞-*submanifold of the Hilbert space* $H^r(\xi)$.

The important point is that $C^r(\gamma)$ is always a C^∞-manifold and that if $2r > \dim M$ then $H^r(\gamma)$ is a Hilbert manifold. The next theorem will show that the differentiable structure of $C^r(\gamma)$ is intrinsic, i.e. independent of the embedding of γ as a sub-bundle of a vector bundle. In fact it says that $C^r(\gamma)$ is a covariant functor from differentiable fiber bundles over M to C^∞ manifolds. Similarly it says that if $2r > \dim M$ then $H^r(\gamma)$ is a covariant functor to C^∞-Hilbert manifolds.

THEOREM 2. *Let* ξ' *be a second vector bundle over* M, γ' *a differentiable fiber bundle over* M *that is a closed sub-bundle of* ξ' *and* $\phi: \gamma \to \gamma'$ *a differentiable fiber preserving map. Define* $\bar{\phi}: C^r(\gamma) \to C^r(\gamma')$ *by* $\bar{\phi}(f)$ $= \phi \circ f$. *Then* $\bar{\phi}$ *is a* C^∞-*map and if* $2r > \dim M$ $\bar{\phi}$ *extends to a* C^∞ *map of* $H^r(\gamma) \to H^r(\gamma')$.

REMARK. If we fix an inner product for the Hilbert space $H^r(\xi)$, then this induces a complete Riemannian structure on the closed submanifold $H^r(\gamma)$ $(2r > \dim M)$. However the latter clearly is not intrinsic.

REMARK. Specializing to the important special case of product bundles $\gamma = M \times W$ the above defines a C^∞ manifold structure on the space $C^r(M, W)$ of C^r maps of M into W, and (if $2r > \dim M$) a C^∞ Hilbert manifold $H^r(M, W)$ of H^r maps of M into W. It is well known that $C^r(M, W)$ has the homotopy type of $C^0(M, W)$. The same is true of $H^r(M, W)$ if $2r > \dim M$.

3. **Variational problems.** Maintaining the notation of the preceding section let $F: J^r(\xi) \to R$ be a C^∞ function and let μ be a smooth measure on M. Define $J: C^r(\xi) \to R$ by $J(f) = \int_M F(j_k(f)) d\mu$, $r \geq k$.

THEOREM. *If $r \geq k$ then J is a C^∞ real-valued function on $C^r(\xi)$.*

The calculus of variations is concerned with the study of the critical points of J, usually restricted to some submanifold of $C^r(\gamma)$ defined by certain "boundary conditions." In order to apply the abstract theory of §1 one puts conditions on the integrand F to insure that J extends to a function of at least class C^2 on $H^r(\xi)$. Then if $2r > \dim M$ J restricted to the submanifold of $H^r(\gamma)$ corresponding to the boundary conditions will also be of class C^2. Similarly one imposes conditions on F to insure that condition (C) of §1 will be satisfied on this submanifold. While we are far from having definitive results in this direction we will give in the next two sections theorems which include a great number of classical results and many new results besides.

4. **A generalization of the Morse theory of geodesics.** Let D be a domain with smooth boundary in R^n. Given an n-tuple $\alpha = (\alpha_1, \cdots, \alpha_n)$ of non-negative integers let $|\alpha| = \alpha_1 + \cdots + \alpha_n$. For each such α with $|\alpha| \leq k$ let A_α be a C^∞-map of D into the space of linear transformations of R^p into itself and define a linear map $L = \sum_{|\alpha| \leq r} A_\alpha D^\alpha$ of $C^{|\alpha|}(D, R^p)$ into $C^0(D, R^p)$ by

$$(Lf)(x) = \sum_{|\alpha| \leq k} A_\alpha(x)(\partial^{|\alpha|}/\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n})f(x).$$

Such an L is called a kth order differential operator and its symbol $\sigma_k(L)$ is the map of $D \times (R^n - 0)$ into the space of linear maps of R^p into itself defined by $\sigma_k(L)(x, \xi) = \sum_{|\alpha| = k} \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n} A_\alpha(x)$. If $\sigma_k(L)$ is everywhere nonsingular then L is called elliptic and if $k = 2r$ and $(-1)^r \sigma_k(L)$ is everywhere a positive definite self-adjoint operator then L is called strongly elliptic. The formal adjoint L^* of L is the kth order differential operator $\sum_{|\alpha| \leq k} (-1)^{|\alpha|} A_\alpha^* D^\alpha$ (where $A_\alpha^*(x)$ is the adjoint of $A_\alpha(x)$), so clearly $\sigma_k(L^*) = (-1)^k \sigma_k(L)$. If L' is an mth order differential operator then it is easily seen that $L'L$ is an $(m+k)$th order differential operator and that $\sigma_{m+k}(L'L) = \sigma_m(L') \sigma_k(L)$, hence if L is elliptic of order k then L^*L is strongly elliptic of order $2k$.

We denote the inner product in R^p by $\langle , \rangle$ and for $f, g \in C^0(D, R^p)$ we define $\langle f, g \rangle_0 = \int_D \langle f(x), g(x) \rangle dx$. Let W be a closed submanifold of R^p. Let $f_0 \in C^\infty(D, W)$ and let $C_0^k(D, W)$ be the set of $f \in C^k(D, W)$ such that $D^\alpha f(x) = D^\alpha f_0(x)$ for $|\alpha| < k$ and $x \in \partial D$, and let $H_0^k(D, W)$ denote the closure of $C_0^k(D, W)$ in $H^k(D, R^p)$. Then it follows from

the results of §2 that $C_0^k(D, W)$ is a closed submanifold of $C^k(D, R^p)$ and that if $2k > n$ then $H_0^k(D, W)$ is a closed submanifold of $H^k(D, R^p)$, hence a complete Riemannian manifold. Now let L be a differential operator of order $2k$ and define a C^∞-function $J: C^{2k}(D, R^p) \to R$ by $J(f) = \langle Lf, f \rangle_0$. Then one shows that J restricted to $C_0^{2k}(D, R^p)$ extends to a C^∞-map of $H_0^k(D, W) \to R$ if $2k > n$. We will denote this map by J also.

THEOREM. *Assume that L is a strongly elliptic differential operator of order $2k > n$ such that $Lf = 0$ has no solution $f \in C^\infty(D, R^p)$ satisfying the boundary conditions $D^\alpha f(x) = 0$ for all $|\alpha| < k$ and all $x \in \partial D$. Then $J: H_0^k(D, W) \to R$ satisfies condition (C) of §1. Moreover all the critical points of J on $H_0^k(D, W)$ are in $C_0^\infty(D, W)$.*

If $L = A^*A$ where A is an elliptic kth order differential operator then $J(f)$ is essentially $\|Af\|_0^2 = \langle Af, Af \rangle_0$. Moreover $Lf = 0$ with the boundary conditions of the theorem imply $Af = 0$. A case of special interest is $k = n = 1$, $D = I$ (the unit interval) $A = d/dx$. Then $C_0^1(I, W)$ is all C^1-paths $\sigma: I \to W$ with $\sigma(0)$ and $\sigma(1)$ two given points P and Q of W and $\|A\sigma\|_0^2 = \int_0^1 \|\sigma'(t)\|^2 dt$, the usual action integral. Thus the critical points are the geodesics joining P to Q parameterized proportionately to arc length, and the theorems of §1 in this case yield the usual Morse theory of geodesics [6].

5. A nonlinear Dirichlet problem. As in §3, given the C^∞ integrand $F: J^k(\xi) \to R$, let $J(f) = \int_M F(j_k f) d\mu$ define a C^∞ real function on $C^k(\xi)$. For ("Dirichlet") boundary data of our calculus of variations problem (if boundary M is nonempty), let $f_0 \in C^k(\xi)$, and $C_0^k(\xi)$ be the affine subspace of maps f with $j_{k-1}f = j_{k-1}f_0$ on ∂M. Finally let $H_0^k(\xi)$ be the closure of $C_0^k(\xi)$ in $H^k(\xi)$.

We furthermore give the following conditions on F. There exists a finite local trivialization of $J^k(\xi)$, $U_\alpha \times E \times L_s(V, E) \times \cdots \times L_s^k(V, E) = \{(x, p^0, p^1, \cdots, p^k)\} = \{(x, p)\}$, $\alpha = 1, \cdots, m$, such that on each, the restriction of F, still denoted by F satisfies:

(1) $F(x, p) \leq c_1 \|p\|^2 + c_2,$

(2) $F_{pp}(x, p)(\beta, \beta) \leq c_3 \|\beta\|^2,$ $\beta \in E \times \cdots \times L_s^k(V, E),$

(3) $c_4 \|p^k\|^2 - c_5 \leq \int_M F(x, p) dx,$

(4) $c_6 \|\beta\|^2 \leq F_{p^k p^k}(x, p)(\beta, \beta),$ $\beta \in L_s^k(V, E).$

Here the c_i are constants and the subscripts on F denote partial derivatives. Thus for example

$$F_{p^k p^k}: U_\alpha \times \cdots \times L_s^k(V, E) \to L_s^2(L^k(V, E), R),$$

$$p \in E \times L_s(V, E) \times \cdots \times L_s^k(V, E) \quad \text{and} \quad p^k \in L_s^k(V, E).$$

In the usual case of the (linear) Dirichlet problem, F is derived from a metric on $J^k(\xi)$ in the sense of [3], i.e. is quadratic on each fiber. Then conditions (1) and (2) are automatically satisfied and (3), (4) are a strengthened form of strong ellipticity.

THEOREM. *If F satisfies* (1), (2), *then the real C^∞ function $J = \int_M F(j_k f) du$ on $C^k(\xi)$ can be extended to a C^2 function of $H^k(\xi)$. If furthermore, conditions* (3), (4) *are satisfied, then $J_0: H_0^k(\xi) \to R$, the restriction of this extension to $H_0^k(\xi)$ satisfies* (C) *of §1. Hence J_0 has a minimum on $H_0^k(\xi)$ and if the critical points of J_0 are nondegenerate the Morse theory is valid (cf. §1).*

REMARKS. 1. Conditions (1)–(4) obviously strongly restrict the integrand F, but on the other hand they allow a great deal of nonlinearity and at the same time include the (linear) Dirichlet problem in the general form of Gårding, Browder and Višik, see [8], (when ξ is a product and M a domain of Euclidean space).

2. Since $H_0^k(\xi)$ is contractible, of course no other critical point of J_0 besides the minimum is forced by the homology of $H_0^k(\xi)$. However, the Morse theory implies, for example, that if J_0 has two local nondegenerate minima, then J_0 has some other critical point (such F satisfying our conditions are easy to construct). Morrey (see e.g. [5]) has shown the existence of a minimum (if $k = 1$) under our conditions on F. F. Browder [1] at the same time as this work has proved an existence theorem for partial differential equations related to the above.

3. The critical points of J_0 are weak solutions of the nonlinear strongly elliptic Euler equations. It seems possible that they are C^∞ sections. However the known regularity theory, for example [8], and di Giorgi-Nash cf. [4], [7], implies this only under further drastic assumptions on the problem.

4. Presumably, the theorem of this section extends to sub-bundles η of ξ under similar conditions on F and $2k > \dim M$. Then usually the homology of $H_0^k(\eta)$ will be highly nontrivial and the existence theory will imply much more.

Conversations with several mathematicians have been helpful here, especially those with L. Nirenberg regarding the last section.

REFERENCES

1. F. Browder, *Nonlinear elliptic boundary value problems*, Bull. Amer. Math. Soc. 69 (1963), 862–873.

2. J. Eells, Jr., *On the geometry of function spaces*, International Symposium on Algebraic Topology, pp. 303–308, Universidad Nacional Autónoma de México and UNESCO, Mexico City, 1958.

3. S. Lang, *Introduction to differential manifolds*, Interscience, New York, 1962.

4. C. B. Morrey, *Existence and differentiability theorems for variational problems for multiple integrals*, Partial Differential Equations and Continuum Mechanics, Univ. of Wisconsin Press, Madison, Wis., 1961.

5. ———, *Multiple integral problems in the calculus of variations and related topics*, Ann. Scuola Norm. Sup. Pisa 14 (1960), 1–61.

6. M. Morse, *The calculus of variations in the large*, Amer. Math. Soc. Colloq. Publ. Vol. 18, Amer. Math. Soc., Providence, R. I., 1934.

7. J. Moser, *On the regularity problem for elliptic and parabolic differential equations*, Partial Differential Equations and Continuum Mechanics, Univ. of Wisconsin Press, Madison, Wis., 1961.

8. L. Nirenberg, *On elliptic partial differential equations*, Ann. Scuola Norm Sup. Pisa 13 (1959), 115–162.

BRANDEIS UNIVERSITY AND
COLUMBIA UNIVERSITY