

38. ———, *Third memoir on the expansion of certain infinite products*, Proc. London Math. Soc. 26 (1895), 15–32.

39. K. H. Schellbach, *Die Lehre von den Elliptischen Integralen und den Theta-Funktionen*, Reimer, Berlin, 1864.

40. D. Stanton, *Some basic hypergeometric polynomials arising from finite classical groups*, Ph.D. thesis, Univ. of Wisconsin, Madison, 1977.

41. H. P. F. Swinnerton-Dyer, *On l -adic representations and congruences for coefficients of modular forms (II)*, Modular Functions in One Variable V (J.-P. Serre and D. B. Zagier, eds.), Springer-Verlag, Berlin and New York, 1977.

42. J. Wilson, Ph.D. thesis, Univ. of Wisconsin, Madison, 1978.

RICHARD ASKEY

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 1, Number 1, January 1979
©American Mathematical Society 1979
0002-9904/79/0000-0011/\$03.00

Nonlinearity and functional analysis, Melvyn S. Berger, Academic Press, New York, San Francisco, London, 1977, xix + 415 pp., \$24.50.

Linear functional analysis evolved as the natural gathering point for a number of different investigations into the solvability of linear equations which were either in the form of integral equations or in the form of countable systems of linear scalar equations in which the unknown was a sequence of numbers. As the subject developed much broader areas of applicability became evident. These applications, in turn, spawned further abstract development, and the abstract results themselves assumed an intrinsic interest. The basic approach of functional analysis, where one considers functions to be points in a large space of related functions and lets the differential or integral operator act on these points, has also been very successful in treating nonlinear problems.

The division between the linear and the nonlinear theory is, of course, not so sharp. As we know from calculus, a great deal of information about a system of nonlinear equations is obtained from their local linear approximation. Moreover, it is often possible to glean information about a linear problem by considering a related nonlinear problem; this is so strikingly demonstrated in Lomonosov's recent results on the invariant subspace problem for linear operators [5].

Berger's aim is to give a systematic treatment of some of the fundamental abstract nonlinear results and of their application to certain concrete problems in geometry and physics.

The study of nonlinear operators acting on infinite dimensional spaces has an obvious starting point—study the finite-dimensional case, a finite system of scalar equations in a finite number of unknowns. Even at this step we note that a fairly complete description of the solutions would be very difficult; when $p(x)$ is a polynomial in a single variable the study of solutions of $p(x) = 0$ is the subject matter of classical algebraic geometry.

Let n and k be integers, $n \geq 1$, $k + n \geq 1$ and let Θ be a bounded open subset of \mathbf{R}^n . Suppose $f: \Theta \rightarrow \mathbf{R}^{n+k}$ is continuous.

If f is one-to-one then $k \geq 0$. Moreover, assuming f is one-to-one, $f(\Theta)$ is open in \mathbf{R}^{n+k} iff $k = 0$. (F1)

If $k = 0$, Θ contains the origin, and $f(x) \neq \beta x$ when $\beta \leq 0$ and $x \in \partial\Theta$, then there exists $x \in \Theta$ with $f(x) = 0$. (F2)

Clearly (F1) is a direct (local) extension of the linear theory. (F2) is an existence result from which it is easy to see, letting $f = I - h$, that when Θ is convex and $h(\partial\Theta) \subseteq \Theta$, then h has a fixed point. These results, known as the Invariance of Domain Theorem and the Brouwer Fixed Point Theorem are consequences of the topological degree of Brouwer. For many years prior to 1912 generalizations of the notion of winding number for functions of a complex variable had been sought. Brouwer completed this search in 1912 by assigning to each member g of $\mathcal{F} \equiv \{f: \Theta \rightarrow \mathbf{R}^n, f \text{ is continuous, } f(x) \neq 0 \text{ when } x \in \partial\Theta\}$ an integer, which we can denote by $\text{deg}(g, \Theta, 0)$. This integer was invariant under homotopies lying in \mathcal{F} , and its nonvanishing implied g vanished in Θ . It enjoys other properties and is very useful for not only proving existence results, but also for describing the nature of the set of solutions.

Suppose $\Theta = \{x \in \mathbf{R}^n, \|x\| \leq 1\}$, let $S^j = \{x \in \mathbf{R}^{j+1} \mid \|x\| = 1\}$ and define $\hat{f}: S^{n-1} \rightarrow S^{n+k-1}$ by $\hat{f}(x) = f(x)/\|f(x)\|$ for $x \in S^{n-1}$. Then every continuous extension of \hat{f} to Θ with values in \mathbf{R}^{n+k} necessarily has a zero iff \hat{f} is not homotopic, through maps of S^{n-1} into S^{n+k-1} , to a constant map. In particular, if \hat{f} is homotopically nontrivial then $f(x) = 0$ has a solution in Θ . (F3)

The proof of the above is clear. When $k = 0$, $\text{deg}(f, \Theta, 0) \neq 0$ iff \hat{f} is homotopically nontrivial, so no new existence information is gained. When $k > 0$, \hat{f} is always homotopically trivial. When $k < 0$, one has a sufficient, if difficult to check, criteria for existence.

Suppose f is continuously differentiable and at $x_0 \in \Theta$, $df(x_0)$ is nonsingular. Then f is a homeomorphism of a neighborhood of x_0 onto a neighborhood of $f(x_0)$. (F4)

Let f be differentiable, have a Lipschitz continuous derivative, and suppose that at $x_0 \in \Theta$, $df(x_0)$ is nonsingular. Then if $f(x_0)$ is small, the sequence $\{x_k\}$ defined by

$$x_{k+1} = x_k - [df(x_k)]^{-1} f(x_k) \tag{F5}$$

converges to a solution of $f(x) = 0$.

The Inverse Function Theorem, (F4), and Newton's Method, (F5), are distinctly different from the previous two results. Based upon information about f at a point x_0 one can say something about f nearby x_0 . In (F2) and (F3) information about f on $\partial\Theta$ implies the existence of a zero in Θ . Moreover, (F4) and (F5) are not finite-dimensional results in the sense that they are immediately extendable to the general Banach space case. When X and Y are Banach spaces, $U \subseteq X$ is open, $f: U \rightarrow Y$ is differentiable (the derivative of f at x , $df(x)$, is a bounded linear map of X into Y) and $df(x_0)$ being nonsingular means it is a bijection between X and Y , then (F4) and

(F5) hold. The proof is based upon the most venerable of nonlinear techniques—the iterative method which is now well known as the Banach Contraction Principle [3].

Let A be a complete metric space, with $h: A \rightarrow A$ Lipschitz with Lipschitz constant less than 1. Then h has a unique fixed-point x^* , and $\{h^k(x)\}$ converges to x^* for any $x \in A$.

The extension of (F1) and (F2), and, more generally, of topological degree, to general Banach spaces requires an essential modification. From now on let X denote a Banach space equipped with a norm $\| \cdot \|$, and let $B_r = \{x \in X \mid \|x\| < r\}$. If X is infinite dimensional, B_1 is not compact and as a result there always exist continuous maps $g: B_1 \rightarrow B_1$ without fixed-points. This being so, one cannot have a reasonable topological degree defined for *all* continuous maps $g: B_1 \rightarrow X$ with $g(x) \neq 0$ when $x \in \partial B_1$.

In 1922 Birkhoff and Kellogg [2] used the Brouwer Fixed Point Theorem to prove the existence of solutions to certain integral equations. Their method of proof was to discretize the equations, apply the Brouwer Theorem to the resulting finite-dimensional system, and then show that as the discretizations became more refined these approximate fixed-points converged to a solution of the full equation. In 1930 Schauder [8] gave an abstract version of their results. When $D \subseteq X$ a continuous mapping, $g: D \rightarrow X$ is called *compact* if $\overline{g(A)}$ is compact whenever $A \subseteq D$ is bounded. Schauder proved that if D is closed, bounded and convex and $g: D \rightarrow D$ is compact, then g has a fixed-point. The role played by compact perturbations of the identity in the linear theory had long been noted, and in 1934 Leray and Schauder [4] extended Brouwer degree to nonlinear mappings of this type. When $V \subseteq X$ is open and bounded, $h: \overline{V} \rightarrow X$ is compact, $h(x) \neq x$ when $x \in \partial V$, they defined an integer $\text{deg}(I - h, V, 0)$, which they proved had all of the properties of the Brouwer degree provided deformations took place within compact perturbations of the identity. Having a topological degree available permitted investigation of problems by means of the following a priori bound method—if $C: X \rightarrow X$ is compact and linear and $f: X \rightarrow X$ is compact then f has a fixed-point provided that $\{x \in X \mid x = tf(x) + (1 - t)C(x), \text{ for some } t \in [0, 1]\}$ is bounded. This method produced substantial improvement in existence results for nonlinear elliptic partial differential equations, problems which had previously been attacked by local continuation arguments where one starts, say, at the unique solution of $x = tf(x) + (1 - t)C(x)$ at $t = 0$, uses a local analysis to prove existence for small values of t , and by repeating the argument marches across to a solution at $t = 1$.

Of course, for a concrete problem the choice of the spaces in which one sets up the problem will be very important (assuming there is a choice; for some problems, methods of functional analysis have not been shown to be useful).

Let us consider the following problem: Let Ω be a bounded domain in \mathbf{R}^m , with $\partial\Omega$ smooth, let $q: \overline{\Omega} \times \mathbf{R}^{m+1} \rightarrow \mathbf{R}$ be continuous and consider the equation

$$\begin{cases} \Delta u(x) = q(x, u(x), \nabla u(x)), & \text{for } x \in \Omega, \\ u(x) = 0, & \text{for } x \in \partial\Omega. \end{cases} \quad (1)$$

To consider this problem in the context of functional analysis it is clear that one should first find out the spaces where the Laplacian is well behaved. The obvious guess is that one should let $X = \{u \in C^2(\bar{\Omega}), u = 0 \text{ on } \partial\Omega\}$ and $Y = C(\Omega)$, where these spaces are equipped with the norm defined to be the sum of the supremum norms of derivatives up to order 2, and the supremum norm, respectively. It turns out that $\Delta: X \rightarrow Y$ is continuous and one-to-one, but fails to be onto, and, in fact, fails to have closed range. Schauder developed a priori estimates for the Dirichlet problem from which it follows that the following spaces are better for Δ . If $u: \bar{\Omega} \rightarrow \mathbf{R}$ and $\alpha \in (0, 1)$ one defines the α -norm of u , $\|u\|_\alpha$, to be

$$\sup_{x \in \Omega} |u(x)| + \sup_{\substack{x, y \in \Omega \\ x \neq y}} \frac{|u(x) - u(y)|}{|x - y|^\alpha},$$

and let $C^{2,\alpha} = \{u: \bar{\Omega} \rightarrow \mathbf{R}, \text{ derivatives of } u \text{ up to order 2 have finite } \alpha\text{-norm, and } u = 0 \text{ on } \partial\Omega\}$ and $C^{0,\alpha} = \{u: \bar{\Omega} \rightarrow \mathbf{R}, \|u\|_\alpha < \infty\}$. Letting $\|u\|_\alpha$ define the norm on $C^{0,\alpha}$ and letting the norm of u in $C^{2,\alpha}$ be the sum of the α -norms of derivatives less than or equal to order 2, these are Banach spaces and $\Delta: C^{2,\alpha} \rightarrow C^{0,\alpha}$ is a bijection. One can also choose Y to be $L^p(\Omega)$, $p > 1$, and identify a space of functions, X , having generalized derivatives up to order 2, incorporating the null boundary conditions, with $\Delta: X \rightarrow L^p$ a bijection.

In either of the above situations when q satisfies some additional properties one can formulate the above problem as a fixed-point equation for a compact mapping in $C^{0,\alpha}$ or L^p , and one can bring the machinery of topological degree to bear on problem (1).

The above fixed-point theorems and the Leray-Schauder degree have been generalized in many different directions, as has the fundamental method of approximating infinite-dimensional problems by finite-dimensional problems and then verifying the validity of taking limits.

Let us now consider some problems which are related to the classical calculus of variations. Many problems in geometry and physics have a formulation where one has a real valued functional, ψ , defined on a class of functions and one seeks a function in this class where ψ is minimized. The problem of determining geodesics joining two points p and q on a manifold M amounts to minimizing arclength among all paths in M joining p and q . When a body is deformed under the influence of certain forces and one wishes to determine the equilibrium configuration one can associate to each possible configuration an energy, and the configuration which minimizes this energy is the equilibrium.

When the above ψ is differentiable and u_0 is a minimum then by taking directional derivatives of ψ in arbitrary directions at u_0 one obtains necessary conditions for a minima which in many concrete situations are the classical Euler-Lagrange equations. On the other hand, sometimes one has a partial differential equation which one recognizes has a variational principle behind it and then the above ψ can be used as a means of studying the equation. In the situation where the domain of ψ is restricted to a surface one uses Lagrange multipliers to relate the minimization problem to an eigenvalue problem.

Suppose q in equation (1) does not depend on ∇u . Then a classical approach to solving

$$\begin{cases} -\Delta u(x) = q(x, u(x)), & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega, \end{cases} \quad (2)$$

is to minimize $\psi(u) = \int_{\Omega} \left\{ \frac{1}{2} |\nabla u(x)|^2 - g(x, u(x)) \right\} dx$, over a suitable class of functions vanishing on $\partial\Omega$, where $\partial g(x, s)/\partial s = q(x, s)$. In our present context an important aspect of this problem is to determine reasonable conditions under which a continuous function $\psi: H \rightarrow \mathbf{R}$, where H is a Hilbert space, has a minimum. Suppose there exists an $r > 0$ such that $\psi(u) > \psi(0)$ when $\|u\| \geq r$. Then, letting $m = \inf\{\psi(u) | u \in H, \|u\| \leq r\}$, we wish to prove m is attained. Often it is not hard to prove $m > -\infty$, so, supposing this has been done, we may select $\{u_n\}$ with $\|u_n\| \leq r$ and $\psi(u_n) \leq m + 1/n$, for each n . One cannot expect $\{u_n\}$ to have a subsequence which converges in the norm topology. However, assuming H is separable and choosing $\{e_i\}$ to be an orthonormal basis for H , we may use a diagonalization argument to choose a subsequence $\{u_{n_k}\}$ and $u_0 \in H$ with $\|u_0\| \leq r$ such that $\{\langle u_{n_k}, e_i \rangle\} \rightarrow \langle u_0, e_i \rangle$, for each i . If one additionally assumes ψ is *convex* one can show ψ is weakly lower semicontinuous with respect to coordinatewise convergence. So u_0 is a minimum.

This idea of relating convexity and the topology of coordinatewise convergence was known to Hilbert and Beppo Levi, and the method carries over to a general Banach space iff it is reflexive.

When g satisfies appropriate growth conditions the above method can be applied to equation (2) where H is the completion of the C^∞ functions on Ω of compact support in the inner product

$$\langle u, v \rangle = \int_{\Omega} \nabla u(x) \nabla v(x) dx + \int_{\Omega} u(x)v(x) dx.$$

Now the above method establishes the existence of just one solution of equation (2), and classical finite-dimensional results give quite general conditions under which one can establish the existence of a number of critical points of φ on M where $\varphi: M \rightarrow \mathbf{R}$ is C^1 and M is a manifold modelled on \mathbf{R}^n . When $\varphi = (\varphi_1, \dots, \varphi_n)$ and M is open in \mathbf{R}^n one gets solutions of $\nabla \varphi(x) = 0$ and when, say, $M = S^{n-1}$, one gets distinct eigenvalues of $\nabla \varphi$.

Let's be more precise. When $\mathcal{O} \subseteq \mathbf{R}^n$ is open and $\psi: \mathcal{O} \rightarrow \mathbf{R}$ is C^2 , let $C(\psi) = \{x \in \mathcal{O} | \nabla \psi(x) = 0\}$; if $x \in C(\psi)$ implies that the matrix of second partial derivatives of ψ , the Hessian of ψ , at x is nonsingular, ψ is called a Morse function. For such a function, Morse showed that when $x \in C(\psi)$ a local change in coordinates $\xi \rightarrow y(\xi)$ could be made so that $\psi(\xi) - \psi(x) = -\sum_{i=1}^k y_i^2 + \sum_{i=k+1}^n y_i^2$, and $k = k(x)$, defined to be the index of x , is the dimension of the maximal subspace on which the Hessian of ψ at x is negative (this is Morse's Lemma). Even at this stage, once one observes that the local Brouwer degree of $\nabla \psi$ at $x \in C(\psi)$ is $(-1)^{k(x)}$, the properties of the Brouwer degree can be used to obtain information about the number of critical points of ψ of index j . These definitions can be extended to a smooth finite-dimensional manifold M , and when M is compact Morse established

lower bounds for the number of critical points of ψ of index k in terms of the k th Betti number of M (see [7]).

Lyusternik and Shnirelman (see [6]) developed, in the 1920s and 1930s, methods for estimating critical points by certain topological invariants without the above rather strong nongeneracy condition. When W is a topological space, of which V is a subspace, let $\text{cat}_W(V)$ be the least number of closed subsets V_j of V , each contractible in W , such that $V = \cup_j V_j$ (possibly $\text{cat}_W(V) = \infty$). One of the consequences of their theory is that if M is a compact, smooth finite-dimensional manifold, $\psi: M \rightarrow \mathbf{R}$ is C^1 , and there is a subset K of M with $\text{cat}_M(K) \geq m$, then there is a critical point x_m of ψ on M with $\psi(x_m) = c_m$, where

$$c_m = \inf_{K \in \mathcal{Q}_m} \{ \sup \{ \psi(x) \mid x \in K \} \},$$

\mathcal{Q}_m being the collection of closed subsets of M of category greater than m .

The above minimax principle is a direct generalization of the Courant-Weyl minimax principle estimating the eigenvalues of an $n \times n$ symmetric matrix L ; one lets $M = S^{n-1}$, $\psi(x) = \langle L(x), x \rangle$, and replaces \mathcal{Q}_m by the intersections of S^{n-1} with subspaces of dimension greater than m .

The difficulty with applying the category argument is that the category may be hard to compute or the category may be small. That the category may be difficult to compute is intrinsic; one can sometimes determine more easily computable but still useful invariants. The second difficulty may sometimes be circumvented. If $f: S^{n-1} \rightarrow \mathbf{R}$ is C^1 and even, then, by identifying antipodal points, f may be considered to be defined on $(n-1)$ -dimensional projection space, which is known to have category n (in the above eigenvalue problem ψ is even).

The basic idea in both of the above critical point theories is to take $\psi: M \rightarrow \mathbf{R}$, $a \leq b$ and let $M_c = \{x \in M \mid \psi(x) \leq c\}$. If ψ has no critical points in $M_b \setminus M_a$ then the flows $x'(t) = -\nabla \psi(x(t))$, $x(0) = p \in M_b$ will carry M_b to M_a , allowing comparison of topological properties of M_b and M_a , while if $M_b \setminus M_a$ contains critical points and f is a Morse function the difference between M_b and M_a can be described in terms of the indices of the critical points in $M_b \setminus M_a$.

The limitations of local continuation arguments for describing periodic orbits for dynamical systems motivated Poincaré and others to seek more global methods to study such problems, where one reformulated the orbit problem as a description of closed geodesics on a surface. As early as 1917 [1] explicit minimax results were used to obtain the existence of such closed geodesics, the novelty being that these were minimax results on a noncompact space.

When S is a smooth n -dimensional Riemannian manifold and $p, q \in M$, the problem of describing geodesics on M joining p and q is the problem of minimizing arclength on the space P of smooth paths joining p and q on S . By rather ingenious approximation arguments Morse was able to use his finite-dimensional results to describe these geodesics. More recently, such minimization problems have been attacked more directly by formulating the problem as a minimization problem in a manifold modelled on a reflexive

Banach space and then using methods analogous to the gradient flow methods which work in the compact case. Since one is giving up compactness in the domain a price has to be paid. Both the Morse theory and the Lyusternik-Shnirelman theory have extensions to manifolds M modelled on a reflexive Banach space provided the map $\psi: M \rightarrow \mathbf{R}$ is C^1 and satisfies condition (c) of Palais and Smale; if $\{x_n\} \subseteq M$ with $\{\psi(x_n)\}$ bounded and $\{\nabla\psi(x_n)\}$ converging to 0 then $\{x_n\}$ has a convergent subsequence. In practice this amounts to verifying that if $\{x_n\} \subseteq M$ and $\{x_n\}$ converges coordinatewise to x , with $\{\nabla\psi(x_n)\}$ converging to 0, then $\{x_n\}$ converges strongly to x . Such generalizations have proven useful in minimal surface problems; they can also be used in studying equations such as (2), and have been applied to nonlinear eigenvalue problems.

Let's finish this discussion with a method for local analysis, which while simple in concept, is very useful when applicable. Suppose X and Y are Banach spaces, $U \subseteq X$ is open, and $f: U \rightarrow Y$ is C^1 . Moreover suppose one also knows that at $x_0 \in U$, $f(x_0) = 0$ and $df(x_0)$ has an m -dimensional null-space and a range of codimension k . Then, letting Q and P be continuous linear projections of X onto the null-space of $df(x_0)$ and of Y onto the range of $df(x_0)$, respectively, one can use the Implicit Function Theorem to find a neighborhood Θ of x_0 in $Q(x)$ and a function $h: \Theta \rightarrow (I - P)(Y)$ such that x near x_0 is a zero of f iff $P(x) \in \Theta$ is a zero of h . So one has reduced the local description of zeros of f to a system of k real equations in m real unknowns.

This method has been particularly useful in the study of problems involving parametric dependence of solutions, for example, if $h: \mathbf{R} \times \Theta \rightarrow Y$, $(\lambda_0, x_0) \in \mathbf{R} \times \Theta$ and one knows $h(\lambda, x_0) = 0$ for $|\lambda - \lambda_0| < \varepsilon$. One wishes to describe solutions of $h(\lambda, x) = 0$ which are near (λ_0, x_0) but $x \neq x_0$. Such a problem is called a bifurcation problem, and when one applies the above reduction method based on the linearization $d_x f(\lambda_0, x_0)$ and so transfers the parameter to a finite system of equations, the method, known as the Lyapunov-Schmidt procedure, has been widely applied.

As Berger points out in the Introduction, for obvious reasons of space he had to omit many important topics which could naturally be fitted into his overall scheme. For similar reasons the preceding discussion omits a large number of topics covered in the book.

Berger's book is distinguished by the broad variety of problems which have been treated by the abstract results which are developed; there are applications to the determination of periodic solutions of nonlinear ordinary differential equations, the structure of the solutions of von Kármán's equations, nonlinear Dirichlet problems, curvature problems, and more. Moreover, many of the results are from the recent research literature.

The book is divided into three parts. In the first part there is a rather detailed list of the necessary background from linear functional analysis and elliptic partial differential equations (a few brief proofs are included), followed by a discussion of properties of various classes of nonlinear operators. The second part treats the local analysis of a mapping. Infinite dimensional versions of (F4) and (F5) are discussed, as are such classical techniques as the majorant method, asymptotic expansions, and singular perturbations. Bifur-

