4

The Mountain-Pass Theorem

1 Critical Points of Minimax Type

Roughly speaking, the basic idea behind the so-called *minimax method* is the following:

> Find a critical value of a functional $\varphi \in C^1(X, \mathbb{R})$ as a *minimax* (or *maximin*) value $c \in \mathbb{R}$ of φ over a suitable class $\mathcal{A}$ of subsets of X:
>
> $$c = \inf_{A \in \mathcal{A}} \sup_{u \in A} \varphi(u) .$$

Example A. Perhaps one of the first examples using a minimax technique is due to E. Fischer (1905) through a well-known minimax characterization of the eigenvalues of a real, symmetric $n \times n$ matrix M (cf. [33], pp. 31 and 47):

$$\lambda_k = \inf_{\{X_{k-1}\}} \sup_{x \perp X_{k-1}, |x|=1} (Mx|x) ,$$

$$\lambda_{-k} = \sup_{\{X_{k-1}\}} \inf_{x \perp X_{k-1}, |x|=1} (Mx|x) .$$

Here, the eigenvalues are numbered so that $\lambda_{-1} \leq \cdots \leq \lambda_{-k} \leq \cdots \leq 0 \leq \cdots \leq \lambda_k \leq \cdots \leq \lambda_1$. Also we are denoting by $(\cdot|\cdot)$ (resp. $|\cdot|$) the usual inner product (resp. norm) in $X = \mathbb{R}^n$, and by $X_j \subset X$ an arbitrary subspace of dimension j. It should be noted that a characterization which is *dual* to the above characterization also holds true, namely:

$$\lambda_k = \sup_{\{X_k\}} \inf_{x \in X_k, |x|=1} (Mx|x) ,$$

$$\lambda_{-k} = \inf_{\{X_k\}} \sup_{x \in X_k, |x|=1} (Mx|x) .$$

Example B. A similar characterization can be obtained for the eigenvalues of a compact, symmetric operator $T : X \longrightarrow X$ on a Hilbert space X. This is part of the so-called *Hilbert-Schmidt theory.*

Example C. A *topological* analogue of such minimax schemes was developed by L. Lusternik and L. Schnirelman from 1925 to 1947. This is known as the (classical) *Lusternik-Schnirelman theory.* It was originally based on the topological notion of *category* Cat (A, X) of a closed subset A of a metric space X. By definition, Cat (A, X) is the smallest number of closed, contractible subsets of X which is needed to cover A (see [53, 54]).

In this context, given a functional $\varphi \in C^1(X, \mathbb{R})$ over, say, a differentiable Riemannian manifold X, the idea is to show that the following values are critical values of φ:

$$c_k = \inf_{A \in \mathcal{A}_k} \sup_{x \in A} \varphi(x) \ , \ \ k = 1, 2, \dots \ ,$$

where $\mathcal{A}_k := \{A \subset X \mid A \text{ is closed, Cat}\,(A, X) \geq k\ \}$. For example, since Cat $(S^n, S^n) = 2$, one obtains, for a given functional $\varphi \in C^1(S^n, \mathbb{R})$, that

$$c_1 \leq c_2 = c_3 = \cdots \ ,$$

and, in this case, $c_1 = \inf \varphi$, $c_2 = \sup \varphi$. Of course this gives us no new information in this case since we know that $\inf \varphi$ and $\sup \varphi$ are attained on the compact manifold S^n and, therefore, are critical values of φ. However, if $\varphi \in C^1(S^n, \mathbb{R})$ is an *even* functional, one obtains more critical values, as shown by the following classical theorem due to Lusternik (1930):

Theorem 1.1. *([53]) Let $\varphi \in C^1(S^n, \mathbb{R})$ be given. If φ is* even, *then it has at least $(n + 1)$ distinct pairs*[1] *of critical points.*

The main idea here is that an *even* functional on S^n can be considered as a functional on the real projective space $\mathbb{RP}^n$ (obtained by identification of the antipodal points in S^n), and the topology of $\mathbb{RP}^n$ is much richer than that of S^n. In fact, it can be shown that Cat $(\mathbb{RP}^n, \mathbb{RP}^n) = n + 1$ (cf. [68]) so that, in this case, one obtains $(n + 1)$ critical values (possibly repeated):

$$c_1 \leq c_2 \leq \cdots \leq c_{n+1}.^2$$

Another way to interpret Lusternik's multiplicity result is to consider it as a consequence of the *symmetry* of the problem (*evenness* of φ, in this

[1] Clearly, since φ is even, its critical points occur in pairs.

[2] Moreover, if $c_j = c_{j+k}$ for some $j, k \geq 1$, it can be shown that the category of the critical set K_c is at least $k + 1$.

case). This question of *multiplicity versus symmetry* will be tackled in a future chapter.

2 The Mountain-Pass Theorem

As already mentioned in the beginning of this chapter, the basic idea behind the minimax method is to *minimaximize* (or *maximinimize*) a given functional φ over a *suitable class* of subsets of X. In particular, such a *suitable class* can be chosen to be *invariant* under the deformation $\eta(t,\cdot)$ given in the deformation theorem 3.2.3.

In this section we will present a first illustration of the minimax method which has proven to be a powerful tool in the attack of many problems on differential equations. It is the celebrated *mountain-pass theorem* of Ambrosetti and Rabinowitz [9]:

Theorem 2.1. *Let X be a Banach space and $\varphi \in C^1(X,\mathbb{R})$ be a functional satisfying the Palais–Smale condition (PS) (or, more weakly, $(BCN)_c$).*[3] *If $e \in X$ and $0 < r < ||e||$ are such that*

$$a =: \max\{\varphi(0), \varphi(e)\} < \inf_{||u||=r} \varphi(u) =: b \ , \tag{2.1}$$

then

$$c = \inf_{\gamma\in\Gamma} \sup_{t\in[0,1]} \varphi(\gamma(t))$$

is a critical value of φ with $c \geq b$. (Here, Γ is the set of paths joining the points 0 and e, that is, $\Gamma = \{\gamma \in C([0,1],X) \mid \gamma(0) = 0,\ \gamma(1) = e\}$.)

Proof: First note that $\gamma([0,1]) \cap \partial B_r$ is *nonempty* for any given $\gamma \in \Gamma$, since $\gamma(0) = 0$, $\gamma(1) = e$ and $0 < r < ||e||$ by assumption. Therefore,

$$\max_{t\in[0,1]} \varphi(\gamma(t)) \geq b = \inf_{\partial B_r} \varphi \ ,$$

so that $c \geq b$.

Let us assume, by negation, that c is not a critical value. Then, by the deformation theorem 3.2.2, there exist $0 < \epsilon < \frac{b-a}{2}$ (recall that $a < b$ by (2.1) and $\eta \in C([0,1] \times X, X)$) such that

$$\eta(t,u) = u \quad if \quad u \notin \varphi^{-1}([c-2\epsilon, c+2\epsilon]) \ , \quad t \in [0,1] \ , \tag{2.2}$$

[3] Recall Remark 3.2.1. One could also use $(Ce)_c$ (cf. [67]).

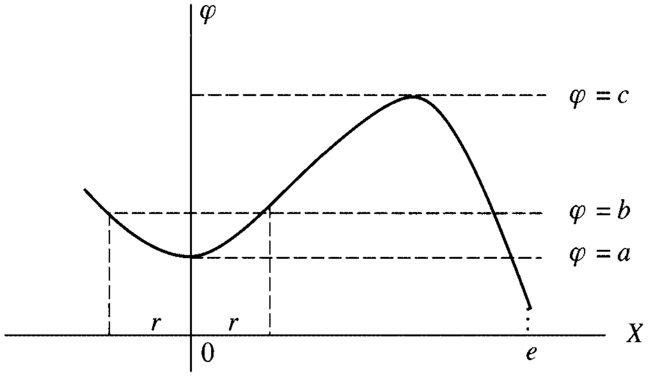

Fig. 4.1.

$$\eta(1, \varphi^{c+\epsilon}) \subset \varphi^{c-\epsilon} \,. \tag{2.3}$$

Now, by definition of c as an infimum over Γ, we can choose $\gamma \in \Gamma$ such that

$$\max_{t \in [0,1]} \varphi(\gamma(t)) \leq c + \epsilon \tag{2.4}$$

and define the path $\widehat{\gamma}(t) = \eta(1, \gamma(t))$. In view of (2.2) and the fact that $2\epsilon < b - a$, it follows that $\widehat{\gamma} \in \Gamma$ (indeed, $\widehat{\gamma}(0) = \eta(1, 0) = 0$ and $\widehat{\gamma}(1) = \eta(1, e) = e$ since $\varphi(0), \varphi(e) \leq a < b - 2\epsilon$). But, then, (2.3) and (2.4) above imply that

$$\max_{t \in [0,1]} \varphi(\widehat{\gamma}(t)) \leq c - \epsilon \,,$$

which contradicts the definition of c. Therefore, c is a critical value of φ. □

Remark 2.1. In the case $u = 0$ is a strict local minimum of φ and $0 \neq e \in X$ is such that $\varphi(e) \leq \varphi(0)$, then Condition 2.1 is clearly satisfied. This situation is common in many application as we shall see next (in this sense, the rough Fig. 4.1 is typical).

3 Two Basic Applications

Application A. Let us show that the following nonlinear Dirichlet problem on a bounded domain $\Omega \subset \mathbb{R}^3$ with smooth boundary possesses a *classical* nontrivial solution:

$$\begin{cases} -\Delta u = u^3 & \text{in } \Omega \\ \quad\; u = 0 & \text{on } \partial\Omega \,. \end{cases} \tag{3.1}$$

To begin with, we observe that since $f(x,u) = u^3$ and $3 < \frac{N+2}{N-2} = 5$, the functional

$$\varphi(u) = \int_\Omega [\frac{1}{2}|\nabla u|^2 - \frac{1}{4}u^4] \, dx$$

is well defined and of class C^1 on the Sobolev space $H_0^1(\Omega)$ by Proposition 2.2.1. The critical points of φ are precisely the weak solutions of (3.1).

Lemma 3.1. *(a) $u = 0$ is a strict local minimum of φ;*
(b) Given $0 \neq v \in H_0^1$ there exists ρ_0 such that $\varphi(\rho_0 v) \leq 0$.

Proof: (a) In view of the Sobolev embedding $H_0^1 \subset L^4$ we have

$$\varphi(u) = \frac{1}{2}||u||^2 - \frac{1}{4}||u||_{L^4}^4 \geq \frac{1}{2}||u||^2 - C||u||^4 ,$$

hence $\varphi(u) > 0 = \varphi(0)$ for all u with $0 < ||u|| \leq r$, for some small $r > 0$.
(b) Letting $\delta = \int_\Omega v^4 dx$ for a given $v \in H_0^1$ with (say) $||v|| = 1$, we have

$$\varphi(\rho v) = \frac{1}{2}\rho^2 - \frac{1}{4}\delta\rho^4 \to -\infty \quad \text{as } \rho \to \infty ,$$

so that the result follows. □

Theorem 3.2. *([9]) Problem* (3.1) *possesses a nontrivial classical solution.*[4]

Proof: We shall use the mountain-pass theorem. Since we already know that $\varphi \in C^1(H_0^1, \mathbb{R})$, we now show that φ satisfies (PS).

Let (u_n) be such that $|\varphi(u_n)| \leq C$, $\varphi'(u_n) \to 0$. Then, for all n sufficiently large, we have

$$|\varphi'(u_n) \cdot u_n| = |\int_\Omega [|\nabla u_n|^2 - u_n^4] \, dx| \leq ||u_n|| ,$$

hence

$$\varphi(u_n) - \frac{1}{4}\varphi'(u_n) \cdot u_n \, dx \leq C + \frac{1}{4}||u_n|| ,$$

that is,

$$\frac{1}{4}||u_n||^2 \leq C + \frac{1}{4}||u_n|| .$$

This implies that $||u_n||$ is bounded, so that we may assume (by passing to a subsequence, if necessary) that $u_n \to \widehat{u}$ weakly in H_0^1. But then, since $\nabla\varphi(u) = u - T(u)$ with T a compact operator (cf. Remark 2.2.1), we obtain

[4] In fact, because of the *eveness* of the corresponding functional φ and its *superquadratic* nature, problem (3.1) has infinitely many solutions, as we shall see later on.

$$u_n = \nabla\varphi(u_n) + T(u_n) \to 0 + T(\widehat{u}) .$$

Therefore, $u_n \to \widehat{u}$ strongly in H_0^1 and we have shown that φ satisfies (PS).

Now, Lemma 3.1 allows us to use Theorem 2.1 (with $e = \rho_0 v$) in order to conclude the existence of a critical point u_0 with $\varphi(u_0) = c \geq b > 0 = \varphi(0)$. Therefore, u_0 is a nontrivial weak solution of (3.1). Moreover, since both $\partial\Omega$ and $f(x,u) = u^3$ are smooth, a *bootstrap* argument shows that u_0 is indeed a classic solution (cf. [2]).

Application B. This next application is a generalization of the previous one. We consider the nonlinear Dirichlet problem (cf. [9])

$$\begin{cases} -\Delta u = f(x,u) & \text{in } \Omega \subset \mathbb{R}^N \\ u = 0 & \text{on } \partial\Omega , \end{cases} \tag{3.2}$$

where $\Omega \subset \mathbb{R}^N$ $(N \geq 2)$ is a bounded smooth domain and, as usual, $f : \overline{\Omega} \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function satisfying the growth condition (f_1) before Proposition 2.2.1 in Chapter 2. Moreover, we shall assume the following conditions:

$$f(x,s) = o(|s|) \quad \text{as } s \to 0, \text{ uniformly in } x. \tag{f_2}$$

There exist $\mu > 2$ and $r > 0$ such that

$$0 < \mu F(x,s) \leq s f(x,s) \quad \text{for } |s| \geq r, \tag{f_3}$$

uniformly in x (where we recall that $F(x,s) = \int_0^s f(x,\tau)d\tau$).

Condition (f_3) is the so-called superquadraticity condition of Ambrosetti and Rabinowitz.

As we know, the fact that f is a Carathéodory function satisfying (f_1) implies (cf. Proposition 2.2.1) that the functional

$$\varphi(u) = \int_\Omega [\frac{1}{2}|\nabla u|^2 - F(x,u)]\, dx \tag{3.3}$$

is well defined and is of class C^1 on the Sobolev space $H_0^1(\Omega)$. Next, we prove an analogue of Lemma 3.1.

Lemma 3.3. *(a) $u = 0$ is a strict local minimum of φ;*
(b) Given $0 \neq v \in H_0^1$ there exists ρ_0 such that $\varphi(\rho_0 v) \leq 0$.

Proof: *(a)* In view of (f_2), given $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that $|f(x,s)| \leq \epsilon|s|$ for all $|s| \leq \delta$, hence

$$|F(x,s)| \leq \frac{1}{2}\epsilon|s|^2 \quad \text{if } |s| \leq \delta. \tag{3.4}$$

Now, since the growth condition (f_1) implies

$$|F(x,s)| \leq A_\epsilon |s|^{\sigma+1} \quad \text{if } |s| \geq \delta = \delta(\epsilon)\ , \tag{3.5}$$

we combine (3.4) and (3.5) to get

$$|F(x,s)| \leq \frac{1}{2}\epsilon|s|^2 + A_\epsilon|s|^{\sigma+1} \quad \forall s \in \mathbb{R},\ \forall x \in \Omega. \tag{3.6}$$

Therefore, using (3.6) we obtain

$$\varphi(u) = \frac{1}{2}||u||^2 - \int_\Omega F(x,u)\ dx \geq \frac{1}{2}||u||^2 - \frac{\epsilon}{2}||u||^2_{L^2} - A_\epsilon||u||^{\sigma+1}_{L^{\sigma+1}}\ ,$$

hence

$$\varphi(u) \geq \frac{1}{2}||u||^2 - \frac{\epsilon}{2\lambda_1}||u||^2 - cA_\epsilon||u||^{\sigma+1} = \frac{1}{2}(1 - \frac{\epsilon}{\lambda_1})||u||^2 - C_\epsilon||u||^{\sigma+1} \tag{3.7}$$

in view of Poincaré's inequality $\lambda_1||u||^2_{L^2} \leq ||u||^2$ and the Sobolev inequality $||u||_{L^{\sigma+1}} \leq c||u||$ (recall that $\sigma + 1 < \frac{2N}{N-2}$). Therefore, since we can take $\epsilon < \lambda_1$ and assume that $\sigma > 1$ in (f_1), the above inequality (3.7) gives $\varphi(u) > 0 = \varphi(0)$ for all u with $0 < ||u|| \leq r$, for some suitably small $r > 0$.

(b) It is easy to see that condition (f_3), together with (f_1), implies that F is *superquadratic* in the sense that there exist constants $c, d > 0$ such that

$$F(x,s) \geq c|s|^\mu - d \quad \forall s \in \mathbb{R},\ \forall x \in \Omega. \tag{3.8}$$

Therefore,

$$\varphi(u) = \frac{1}{2}||u||^2 - \int_\Omega F(x,u)\ dx \leq \frac{1}{2}||u||^2 - c||u||^\mu_{L^\mu} + d|\Omega|\ ,$$

so that, given $v \in H^1_0$ with $||v|| = 1$ and writing $\delta = c||v||^\mu_{L^\mu} > 0$, we obtain

$$\varphi(\rho v) \leq \frac{1}{2}\rho^2 - \delta\rho^\mu + d|\Omega| \longrightarrow -\infty \quad \text{as } \rho \to \infty.$$

In particular, there exists $\rho_0 > 0$ such that $\varphi(\rho_0 v) \leq 0$. □

Remark 3.1. As we have just seen in part (b) of Lemma 3.3, condition (f_3) implies (3.8) with $\mu > 2$ (F is *superquadratic*) and, hence, $\varphi(\rho v) \to -\infty$ as $\rho \to \infty$ for any given $0 \neq v \in H^1_0$. Therefore, the functional φ is *not bounded* *from below.* On the other hand, since $\varphi(u) = \frac{1}{2}||u||^2 - \psi(u)$ where ψ is a weakly continuous functional (recall Example C in Section 2.1 of Chapter 2), then if we let (e_n) denote an orthonormal basis for H_0^1, it follows that $\lim_{n\to\infty} \psi(Re_n) = 0$ for any given $R > 0$, so that $\lim_{n\to\infty} \varphi(Re_n) = \frac{1}{2}R^2$. Since $R > 0$ is arbitrary, we see that φ is also *not bounded from above.*

Theorem 3.4. *([9]) If $f : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function satisfying conditions $(f_1) - (f_3)$, then problem (3.2) possesses a nontrivial weak solution $u \in H_0^1$.*

Proof: As in Theorem 3.2, we start by showing that the functional φ given in (3.3) satisfies the (PS) condition.

Let (u_n) be such that $|\varphi(u_n)| \leq C$, $\varphi'(u_n) \to 0$. Then, for all n sufficiently large, we have

$$|\varphi'(u_n) \cdot u_n| = | \int_\Omega [|\nabla u_n|^2 - f(x, u_n)u_n] \, dx| \leq ||u_n|| \, ,$$

hence

$$\varphi(u_n) - \frac{1}{\mu}\varphi'(u_n) \cdot u_n \, dx \leq C + \frac{1}{\mu}||u_n|| \, ,$$

that is,

$$(\frac{1}{2} - \frac{1}{\mu})||u_n||^2 \leq C + \frac{1}{\mu}||u_n|| \, ,$$

where $(\frac{1}{2} - \frac{1}{\mu}) > 0$, which implies that $||u_n||$ is bounded. The rest of the proof that φ satisfies (PS) is done as in Theorem 3.2. Similarly, Lemma 3.3 and Theorem 2.1 imply the existence of a nontrivial weak solution $u_0 \in H_0^1$ of (3.2). □

Remark 3.2. If $f : \overline{\Omega} \times \mathbb{R} \longrightarrow \mathbb{R}$ is assumed to be locally Lipschitzian, then by a *bootstrap* argument, the weak solution u_0 is a classical solution (see [9]).

Remark 3.3. We point out that the Palais–Smale condition is a compactness condition involving both the functional and the space X in a combined manner. The fact that X is infinite dimensional plays no role in requiring that (PS) (or some other compactness condition) be satisfied in the mountain-pass theorem. Indeed, even in a finite-dimensional space, the geometric conditions alone are not sufficient to guarantee that the level c is a critical level (see Exercise 2 that follows).

4 Exercises

1. Let $\lambda < 0$. Show that the ODE problem

$$\begin{cases} u'' + \lambda u + u^3 = 0 \,, & 0 < t < \pi \\ u'(0) = u'(\pi) = 0 \end{cases}$$

has a solution $u \in C^2[0,\pi]$ which is a mountain-pass critical point of the corresponding functional.

2. Find a polynomial function $p : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ that satisfies the geometric conditions (2.1) of the mountain-pass theorem (so that the minimax value $c \geq b > 0$ does exist), but c is not a critical level of p. (Try to find such a polynomial $p(x,y)$ having $(0,0)$ as a strict local minimum and no other critical point; if giving up, see [20].)

3. Consider the following nonlinear Neumann problem

$$\begin{cases} -\Delta u = f(u) + \rho(x) & \text{in } \Omega \\ \dfrac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega \,, \end{cases} \tag{N}$$

where $\Omega \subset \mathbb{R}^{\mathrm{N}}$ $(N \geq 1)$ is a bounded smooth domain and the continuous functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ (given as p-periodic) and $\rho : \overline{\Omega} \longrightarrow \mathbb{R}$ satisfy the conditions

$$\int_0^p f(s)\, ds = 0 \,, \qquad \int_\Omega \rho(x)\, dx = 0 \,.$$

Recall that, as an application of (the minimum principle) theorem 3.3.1 in Chapter 3 with the Palais–Smale condition replaced by the weaker Brézis–Coron–Nirenberg condition $(BCN)_c$, we proved that (N) had a solution $u_0 \in H^1(\Omega)$ minimizing the corresponding p-periodic functional φ. Clearly, by the periodicity of φ, any translated function $u_k = u_0 + kp$, $k \in \mathbb{Z}$, is also a minimizer of φ. Find another solution for (N) which is different from the u_k's.[5]

4. This is simply a calculus exercise to introduce a function which is *super-linear at infinity* in the sense that

$$\lim_{|s| \to \infty} \frac{f(s)}{s} = +\infty \,,$$

but grows *slower* than any power greater than 1, namely,

$$\lim_{|s| \to \infty} \frac{f(s)}{|s|^\epsilon s} = 0 \qquad \forall \epsilon > 0 \,.$$

[5] The mountain-pass theorem also holds if $b = a$ in (2.1) (cf. [63]).

Indeed, just take $f(s) := F'(s)$, where $F(s) = s^2 ln(1 + s^2)$. You should also check that

$$\lim_{|s|\to\infty} [sf(s) - 2F(s)] = +\infty \ ,$$

which is a condition that is relevant to the next exercise.

5. Consider the Dirichlet problem

$$\begin{cases} -\Delta u = f(x,u) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \ , \end{cases} \qquad (D)$$

where $\Omega \subset \mathbb{R}^N$ is a bounded smooth domain and $f : \overline{\Omega} \times \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, with $f(0) = 0$, $f(x,s) = o(|s|)$ as $s \to 0$ (uniformly for $x \in \Omega$), f satisfying the growth condition (f_1) in Chapter 2 and

$$\liminf_{|s|\to\infty} \frac{f(x,s)}{s} > \lambda_1 \ , \qquad \text{uniformly for } x \in \Omega\,,$$

Moreover, assume that

$$\lim_{|s|\to\infty} [sf(x,s) - 2F(x,s)] = +\infty \ , \qquad \text{uniformly for } x \in \Omega\,,$$ [6]

where, as usual, $F(x,s) = \int_0^s f(x,t)\,dt$. Show that (D) has a nonzero solution. [*Hint:* Use the Fatou lemma to verify that, in view of the above condition, the pertinent functional satisfies the Cerami condition introduced in Exercise 2 of Chapter 3.]

[6] This is a nonquadraticity condition introduced in [31].