CAMBRIDGE

Cambridge University Press
978-0-521-09093-3 - Nonlinear Superposition Operators
Jurgen Appell and Petr P. Zabrejko
Excerpt
More information

Chapter 1

The superposition operator in the space S

In this chapter we study the superposition operator $Fx(s) = f(s, x(s))$ in the complete metric space S of measurable functions over some measure space Ω. First, we consider some classes of functions f which generate a superposition operator F from S into S; a classical example is the class of Carathéodory functions, a more general class that of Shragin functions.

As a matter of fact, there exist functions f, called "monsters", which generate the zero operator $Fx \equiv \theta$, but are not measurable on $\Omega \times \mathbb{R}$, and hence are not Carathéodory functions; this disproves the old-standing Nemytskij conjecture. On the other hand, we show that a function which generates a continuous superposition operator (in measure) is "almost" a Carathéodory function.

We give a necessary and sufficient condition for the function f to generate a bounded superposition operator F in the space S. In particular, this conditions holds always if f is a Carathéodory function. On the other hand, we show that the superposition operator F is "never" compact in the space S, except for the trivial case when F is constant.

Finally, we consider superposition operators which are generated by functions f with special properties (e.g. monotonicity), and characterize the points of discontinuity of such operators.

1.1 The space S

Let Ω be an arbitrary set, $\mathcal{M}$ some σ-algebra of subsets of Ω (which will be called measurable in what follows), and μ a countably additive and σ-finite measure on $\mathcal{M}$. By λ we denote some normalized ("probability") measure on $\mathcal{M}$ which is equivalent to μ (i.e. has the same null sets); one possible choice of λ could be, for instance,

$$\lambda(D) = \int_D n(s) d\mu ,$$

where n is any positive function on Ω with

$$\int_\Omega n(s)d\mu = 1\,.$$

In most examples, we shall deal with either some bounded perfect set Ω with nonempty interior in some finite dimensional space, together with the algebra $\mathcal{M}$ of Borel- or Lebesgue-measurable subsets and the Lebesgue measure μ, or the set of natural numbers, together with the algebra of all subsets and the counting measure. More complicated examples, of course, are also possible: Ω being an arbitrary Lebesgue or Borel subset, and μ the Lebesgue or Borel measure, or Ω being a "nice" subset of a finite dimensional manifold, together with a suitable algebra $\mathcal{M}$ of subsets and some measure μ. Such examples will be considered only in quite exceptional cases. We point out that we do not suppose the σ-algebra $\mathcal{M}$ to be complete with respect to the measure μ.

Recall (Saks' lemma) that the set Ω can be divided, uniquely up to null sets, into two parts Ω_c and Ω_d such that μ is *atomic-free* ("continuous") on Ω_c (i.e. any subset of Ω_c can be divided into two parts of equal measure), and μ is *purely atomic* ("discrete") on Ω_d, i.e. Ω_d is a finite or countable union of atoms of positive measure. In "natural" examples, one of the sets Ω_c or Ω_d is usually empty, and thus one deals with real "function spaces" or "sequence spaces".

As usual, we denote by $S = S(\Omega, \mathcal{M}, \mu)$ the set of all (real or complex-valued) almost everywhere finite μ-measurable functions on Ω; more precisely, S consists of equivalence classes of such functions, where two functions x and y are called *equivalent* if they coincide almost everywhere on Ω. The set S can be equipped with the usual algebraic operations, where the zero element is the function $\theta(s) = 0$ almost everywhere, as well as with the metric $\rho(x, y) = [x - y]$, where

$$[z] = \inf_{0<h<\infty} \{h + \lambda(\{s : s \in \Omega, |z(s)| > h\})\} \tag{1.1}$$

or

$$[z] = \int_\Omega \frac{|z(s)|}{1 + |z(s)|} d\lambda\,. \tag{1.2}$$

With respect to this metric, S becomes a *complete metric space*, and convergence $\rho(x_n, x) \to 0$ is equivalent to *convergence* of x_n *in measure* to x, i.e. $\lambda(\{s : |x_n(s) - x(s)| > h\}) \to 0$ as $n \to \infty$, for any $h > 0$.

It is convenient to introduce also a *partial ordering* in the space S: we write $x \leq y$ $(x, y \in S)$ if $x(s) \leq y(s)$ for almost all $s \in \Omega$. In this way, S

becomes an ordered linear space, i.e. $x \leq y$ implies that $x+z \leq y+z$ for $z \in S$, and that $\lambda x \leq \lambda y$ for $\lambda \geq 0$; moreover, if x_n and y_n are two sequences in S which converge to $x \in S$ and $y \in S$, respectively, then $x_n \leq y_n$ implies that also $x \leq y$. Finally, S is a K-*space* (in the sense of L.V.Kantorovich), which means that any set which is bounded from above (respectively below) admits a least upper bound (respectively greatest lower bound), where these notions are defined as usual.

As in every ordered linear space, one can consider convergence with respect to the above ordering in S. A sequence x_n in S is *order convergent* to $x \in S$ if $\underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = x$, where

$$\underline{\lim}_{n\to\infty} x_n = \sup_k \inf_{m\geq k} x_m\,, \quad \overline{\lim}_{n\to\infty} x_n = \inf_k \sup_{m\geq k} x_m\,.$$

In the space S, this type of convergence coincides with *convergence almost everywhere*. It is well known (Lebesgue's theorem) that convergence almost everywhere implies convergence in measure; the converse is true only if the measure μ is discrete (i.e. $\Omega_c = \emptyset$). Nevertheless (Riesz' theorem), each sequence which is convergent in measure admits a subsequence which converges almost everywhere (to the same limit, of course). We still mention a well known fact about convergence almost everywhere (Jegorov's theorem): if x_n converges almost everywhere to $x \in S$, then x_n converges uniformly outside some set $D \in \mathcal{M}$ of arbitrarily small λ-measure.

In the sequel, we shall denote by χ_D ($D \in \mathcal{M}$) the *characteristic function* of D,

$$\chi_D(s) = \begin{cases} 1 & s \in D\,, \\ 0 & s \notin D\,, \end{cases}$$

and by P_D the *multiplication operator* by χ_D

$$P_D x(s) = \chi_D(s) x(s)\,. \tag{1.3}$$

The functions

$$x(s) = \sum_{j=1}^{m} c_j \chi_{D_j}(s) \quad (D_j \in \mathcal{M};\ j = 1, \cdots, m) \tag{1.4}$$

are usually called *simple functions*. It is not hard to see that the linear space S_0 of simple functions in S is dense in S; this implies, in particular, that the separability of S is equivalent to the separability of the metric space $(\mathcal{M}, d)$, where $d(A,B) = \lambda(A \triangle B)$ is the λ-measure of the "symmetric difference" of

A and B. Moreover (Mikusiński's theorem), every nonnegative function in S is the limit of a monotonically increasing sequence of nonnegative simple functions.

In many situations, the set Ω is a complete metric space, and the algebra $\mathcal{M}$ includes the subalgebra $\mathcal{B}(\Omega)$ of all Borel subsets of Ω. In this case the measure μ is usually supposed to be *regular*, i.e. the following compatibility condition holds between the metric spaces Ω and $(\mathcal{M}, d)$: given $D \in \mathcal{M}$ and $\varepsilon > 0$, there exists a compact subset D_ε of Ω such that $d(D, D_\varepsilon) < \varepsilon$.

We suppose that the reader is familiar with the construction and the basic properties of the (Lebesgue) integral. In what follows, we shall denote by L the set of all (Lebesgue) integrable functions over Ω, equipped with the norm

$$\|x\| = \int_\Omega |x(s)| d\mu(s) . \tag{1.5}$$

If the measure μ under consideration is fixed, we shall write simply ds instead of $d\mu(s)$.

1.2 The superposition operator

Let $f = f(s, u)$ be a function defined on $\Omega \times \mathbb{R}$ (or $\Omega \times \mathbb{C}$), and taking values in $\mathbb{R}$ (respectively $\mathbb{C}$). Given a function $x = x(s)$ on Ω, by applying f we get another function $y = y(s)$ on Ω, defined by $y(s) = f(s, x(s))$. In this way, the function f generates an operator

$$Fx(s) = f(s, x(s)) \tag{1.6}$$

which is usually called *superposition operator* (also outer superposition operator, composition operator, substitution operator, or Nemytskij operator).

In this chapter, and in most other chapters, unless otherwise stated, we shall consider f as function from $\Omega \times \mathbb{R}$ into $\mathbb{R}$.

The superposition operator (1.6) has some remarkable properties. One "algebraic" property which is called the *local determination* of F is described in the following:

Lemma 1.1 *The superposition operator F has the following three (equivalent) properties:*
(a) For $D \subseteq \Omega$,

$$FP_D - P_D F = P_{\Omega \setminus D} F\theta , \tag{1.7}$$

where θ is the almost everywhere zero function.

(*b*) *For* $D \subseteq \Omega$,

$$P_D F P_D x = P_D F x, \ P_{\Omega \setminus D} F P_D x = P_{\Omega \setminus D} F \theta .$$

(*c*) *If two functions* x_1 *and* x_2 *coincide on* $D \subseteq \Omega$, *then the functions* Fx_1 *and* Fx_2 *also coincide on* D.

$\Rightarrow$ The validity of the three conditions follows immediately from the definition of the superposition operator; therefore we shall show only their equivalence (for any operator F). The equivalence of (a) and (b) follows from the equality $FP_D = P_D F P_D + P_{\Omega \setminus D} F P_D$. Now, if (a) holds and x_1 and x_2 coincide on $D \subseteq \Omega$, we have $P_D x_1 = P_D x_2$ and hence, by (1.7), $P_D F x_1 - P_D F x_2 = F P_D x_1 - F P_D x_2 = \theta$, i.e. Fx_1 and Fx_2 coincide on D. Finally, suppose that (c) holds, x is some function on Ω, and $D \subset \Omega$. Then Fx and $FP_D x$ coincide on D, since x and $P_D x$ do so. On the other hand, $P_D x$ and θ coincide on $\Omega \setminus D$, and hence also $FP_D x$ and $F\theta$. This shows that (b) holds. $\Leftarrow$

If, in particular, $F\theta = \theta$ (which means that the function f satisfies

$$f(s, 0) = 0 \tag{1.8}$$

for almost all $s \in \Omega$), all these conditions are equivalent to the fact that F commutes with any of the multiplication operators (1.3), i.e.

$$FP_D = P_D F . \tag{1.9}$$

In this case, the superposition operator is *disjointly additive*; this means that

$$F(x_1 + x_2) = Fx_1 + Fx_2 , \tag{1.10}$$

whenever the functions x_1 and x_2 are *disjoint* (i.e. their supports supp $x_j = \{s : s \in \Omega,\ x_j(s) \neq 0\}$ $(j = 1, 2)$ are disjoint). In fact, from (1.9) we get $F(x_1 + x_2) = F(P_{D_1 \cup D_2} x) = P_{D_1 \cup D_2} F x = P_{D_1} F x + P_{D_2} F x = F P_{D_1} x + F P_{D_2} x = F x_1 + F x_2$, where $x = x_1 + x_2$ and $D_j = \text{supp}\ x_j$ $(j = 1, 2)$. We remark that an analogous partial additivity holds also for countably many functions.

Observe that the condition (1.8) is not really restrictive in many cases, because one can often pass from the superposition operator (1.6) to the superposition operator $\tilde{F}x(s) = \tilde{f}(s, x(s))$ generated by the function

$$\tilde{f}(s, u) = f(s, x_0(s) + u) - f(s, x_0(s)) , \tag{1.11}$$

where x_0 is any fixed function on Ω (for example $x_0 = \theta$).

Cambridge University Press
978-0-521-09093-3 - Nonlinear Superposition Operators
Jurgen Appell and Petr P. Zabrejko
Excerpt
More information

1.3 Sup-measurable functions

Lemma 1.1 implies, in particular, that the superposition operator (1.6) maps equivalent functions on Ω into equivalent ones, i.e. acts actually on classes. Therefore it is natural to ask for *acting conditions* for F in S, i.e. conditions on the function f which guarantee that the corresponding superposition operator F maps all (equivalence classes of) functions in S into such. Surprisingly enough, this problem turns out to be very hard, and a large part of this chapter is actually devoted to the discussion of the known results in this direction. We shall call a function f *superpositionally measurable*, or *sup-measurable*, for short, if the corresponding superposition operator F maps the space S into itself. It is natural to try to characterize sup-measurability by means of (possibly simple) intrinsic properties of f; one basic difficulty in this connection lies in the fact that a sup-measurable function f is by no means uniquely determined by the corresponding operator F, since two functions f_1 and f_2, although generating the same superposition operator, may be "essentially different".

At this point, we introduce special relations between functions of two variables which allow us to formulate most of our results very easily and precisely. Given two functions f_1 and f_2 on $\Omega \times \mathbb{R}$ and some subset Δ of $\Omega \times \mathbb{R}$, we shall write

$$f_1(s,u) \preceq f_2(s,u) \quad ((s,u) \in \Delta) \tag{1.12}$$

if, whenever $x \in S$ has its graph in Δ, we have

$$f_1(s,x(s)) \leq f_2(s,x(s))$$

for almost all $s \in \Omega$ (in case $\Delta = \Omega \times \mathbb{R}$ we drop the condition on the right-hand side of (1.12)). Moreover, the notation

$$f_1(s,u) \simeq f_2(s,u) \quad ((s,u) \in \Delta) \tag{1.13}$$

means that both $f_1(s,u) \preceq f_2(s,u)$ and $f_2(s,u) \preceq f_1(s,u)$, i.e. the operators F_1 and F_2 coincide on the set of all $x \in S$ whose graphs are contained in Δ. In this case we shall call the functions f_1 and f_2 *superpositionally equivalent*, or *sup-equivalent*, for short, on Δ. Note that sup-measurability is then invariant under the equivalence relation (1.13), i.e. if f is sup-measurable then so is every function $\tilde{f}$ which is sup-equivalent to f.

It is very striking that even functions f which are sup-equivalent to the zero function $\tilde{f}(s,u) \equiv 0$ may exhibit a very pathological behaviour; in the literature such functions are called *monsters.* We give now two examples of such monsters which are both constructed on the interval $\Omega = [0,1]$ and essentially build on the validity of the continuum hypothesis.

The first one (which we call the *Russian monster*) was invented by M.A. Krasnosel'skij and A.V. Pokrovskij. Recall first that both the set Ω and the space S (more precisely, a complete representation system of pairwise non-equivalent functions) have the cardinality of the continuum, and therefore their elements ($s_\alpha \in \Omega$ and $x_\beta \in S$, say) can be indexed by the ordinals from 1 to ω, the first uncountable ordinal. Let

$$f(s_\alpha, u) = \begin{cases} 0 & \text{if } u = x_\beta(s_\alpha) \text{ for some } \beta < \alpha, \\ 1 & \text{otherwise.} \end{cases} \tag{1.14}$$

This function is sup-measurable, but not measurable! In fact, for any $x \in S$, the function Fx is different from zero at most on a countable subset of Ω. To see this, fix $x = x_\beta \in S$. By (1.14), we have $f(s, x_\beta(s)) \neq 0$ only for $s = s_\alpha$ with $\alpha \leq \beta$, and these points s form only a finite or countably infinite set (it is here that one uses the continuum hypothesis). This means that $f(s,u) \simeq 0$; in other words, f generates the zero operator, and hence is certainly sup-measurable.

To prove that f is not measurable on the product $\Omega \times \mathbb{R}$, we remark that, for any fixed $s_0 \in \Omega$, the function $f(s_0, \cdot)$ vanishes at most on a countable subset of $\mathbb{R}$. Therefore the set $Q = \{(s,u) : f(s,u) = 1\}$ meets any horizontal line $u = u_0$ in at most countably many points, but contains all vertical lines $s = s_0$ except for at most countably many points. This shows that Q is a non-measurable subset of $\Omega \times \mathbb{R}$ (for example, by Fubini's theorem) and hence f, being the characteristic function of Q, is not measurable either.

The second example (which we call the *Polish monster*) was given by Z. Grande and J. Lipiński and is also quite "exotic". Recall that both the class of all closed subsets of $\Omega \times \mathbb{R}$ of positive measure and the set of all Borel functions on Ω have the cardinality of the continuum, and can therefore again be indexed by the ordinals from 1 to ω (M_α and x_β, say). Further, by transfinite induction one can construct a sequence of points $(s_\alpha, u_\alpha) \in \Omega \times \mathbb{R}$ such that

$$(s_\alpha, u_\alpha) \in M_\alpha \setminus \bigcup_{\beta<\alpha} \Gamma_\beta\,, \quad (s_\alpha \neq s_\beta \text{ for } \beta < \alpha)\,,$$

where M_α is a closed subset of $\Omega \times \mathbb{R}$ of positive measure, and Γ_β is the graph of the Borel function x_β. The possibility of choosing (s_α, u_α) in such a way follows from the fact that the set of all ordinals β "preceding" α is at most countable, the measure of M_α is positive, and the graph Γ_β of a Borel function x_β is a null set in $\Omega \times \mathbb{R}$. The Polish monster

$$f(s,u) = \begin{cases} 1 & \text{if } (s,u) = (s_\alpha, u_\alpha) \text{ for some } \alpha, \\ 0 & \text{otherwise,} \end{cases} \tag{1.15}$$

is then again sup-measurable, but not measurable, as can be seen as follows: if x is any measurable function and x_β is a Borel function which is equivalent to x, the equality $f(s, x_\beta(s)) \neq 0$ holds only for those $s_\alpha \in \Omega$ which satisfy $(s_\alpha, u_\alpha) \in \Gamma_\beta$. But these points form only a countable set, and hence the Polish monster again generates the zero operator, and thus is trivially sup-measurable.

To show that f is non-measurable one can again consider the set $Q = \{(s,u) : f(s,u) = 1\}$; in fact, if Q were measurable, Q would have measure zero (by Fubini's theorem). But in this case one would have $Q \cap M_\alpha = \emptyset$ for some closed set M_α of positive measure, which is impossible since Q contains at least one point of each set M_α.

The above examples show that the sup-measurability of f does not imply its measurability. The converse is not true either; this is easier to see: given an arbitrary non-measurable function z on Ω, just consider

$$f(s,u) = \begin{cases} z(s) & \text{if } u = s, \\ 0 & \text{if } u \neq s; \end{cases} \tag{1.16}$$

this function f is obviously measurable on $\Omega \times \mathbb{R}$, but the corresponding operator F maps the measurable function $x(s) = s$ into the non-measurable function z.

To make the following arguments more transparent, we still introduce some terminology. Suppose that N is some subset of functions of S, and assume that f_1 and f_2 are two sup-measurable functions on $\Omega \times \mathbb{R}$ with the property that the operators F_1 and F_2 take the same values on N. We shall call the set N *thick* if this implies that $f_1(s,u) \simeq f_2(s,u)$, i.e. F_1 and F_2 take the same values on the whole space S. In other words, a function set N is thick if every superposition operator F admits a unique extension from N to S.

A necessary and sufficient condition for "thickness" is provided by the following:

Lemma 1.2 *A subset N of S is thick if and only if, given $x \in S$, one can find a sequence x_n in N and a countable partition $\{D_1, D_2, \cdots\}$ of Ω $(D_n \in \mathcal{M})$ such that $x(s) = x_n(s)$ for almost all $s \in D_n$.*

⇒ The proof is quite simple: in fact, if two superposition operators F_1 and F_2 coincide on N then, by Lemma 1.1, they also coincide on the set N^* of all functions of the form

$$x(s) = \sum_{n=1}^{\infty} P_{D_n} x_n(s) \qquad (x_n \in N),$$

where $\{D_1, D_2, \cdots\}$ is an arbitrary countable partition of Ω (and only on such functions!). But Lemma 1.2 just states that N is thick if and only if $N^* = S$. ⇐

It is also possible to give other simple (sufficient) conditions for the "thickness" of some set of measurable functions. Given $N \subset S$, we call the set $\mathcal{L}u(N)$ of all functions $x \in S$ such that, for each $\varepsilon > 0$, there exists $x_\varepsilon \in N$ with $\lambda(\{s : x(s) \neq x_\varepsilon(s)\}) < \varepsilon$, the *Luzin hull* of N. Obviously, $N \subset S$ is thick if $\mathcal{L}u(N) = S$.

For example, if Ω is a bounded domain in Euclidean space, Luzin's theorem (see Section 6.1) states that $\mathcal{L}u(C) = S$, where C is the set of continuous functions on Ω. It is rather surprising that, if C is replaced by the set C^1 of continuously differentiable functions on Ω, one has $\mathcal{L}u(C^1) \neq S$; in fact, C^1 is not thick in S!

To conclude this section, we shall prove an auxiliary result which will be used several times in the sequel.

Lemma 1.3 *Let a be a nonnegative sup-measurable function on $\Omega \times \mathbb{R}$, and suppose that*

$$\int_\Omega a(s, x(s))ds \leq c < \infty$$

for all $x \in S$. Then there exists a function $\overline{a} \in L$ such that $a(s, u) \preceq \overline{a}(s)$ and

$$\int_\Omega \overline{a}(s)ds \leq c. \qquad (1.17)$$

⇒ By Kantorovich's theorem, the function set $H = \{\arctan Ax : x \in S\}$ admits a least upper bound z_* in S, where A is the superposition operator generated by the function a. Moreover, one can find a sequence x_n in S such that $\sup H = \sup_n \arctan Ax_n$. By induction, we construct a sequence z_n in S putting $z_1 = x_1$ and

$$z_n(s) = \begin{cases} z_{n-1}(s) & \text{if } a(s, z_{n-1}(s)) \geq a(s, x_n(s)), \\ x_n(s) & \text{if } a(s, z_{n-1}(s)) \leq a(s, x_n(s)). \end{cases}$$

Obviously, z_n has the property that arctan Az_n converges monotonically to $\sup H$. By Levi's theorem, the function $\overline{a}(s) = \lim_{n\to\infty} Az_n(s)$ is integrable and satisfies (1.17). $\Leftarrow$

1.4 Carathéodory and Shragin functions

Already in 1918, K. Carathéodory gave the following sufficient condition for sup-measurability: a function $f = f(s,u)$ is sup-measurable if $f(s,\cdot)$ is continuous on $\mathbb{R}$ for almost all $s \in \Omega$, and $f(\cdot,u)$ is measurable on Ω for all $u \in \mathbb{R}$. Such functions are now called *Carathéodory functions* (or functions which satisfy a Carathéodory condition).

To prove the sup-measurability of a Carathéodory function f, note first that the corresponding operator F maps any simple function into a measurable function; in fact, if x has the form (1.4), all functions $f(\cdot,c_j)$ $(j = 1,\cdots,m)$ are measurable, and Lemma 1.1 shows that

$$Fx(s) = \sum_{j=1}^{m} P_{D_j} f(s,c_j)\,,$$

which is clearly a measurable function. Now, if $x \in S$ is arbitrary, and $x_n \in S_0$ are simple functions which converge to x almost everywhere on Ω, we have, by the continuity of $f(s,\cdot)$, that $Fx(s) = \lim_{n\to\infty} Fx_n(s)$ for almost all $s \in \Omega$. By Lebesgue's theorem, Fx is then also measurable.

In the sequel we shall need the following obvious lemma:

Lemma 1.4 *Let f_n be a sequence of sup-measurable functions, and suppose that $f_n(s,u)$ converges to $f(s,u)$ for almost all $s \in \Omega$ and all $u \in \mathbb{R}$. Then f is also sup-measurable.*

$\Rightarrow$ The assertion follows easily from Lebesgue's theorem and the fact that, for any $x \in S$, the sequence $f_n(s,x(s))$ converges almost everywhere to $f(s,x(s))$. $\Leftarrow$

Lemma 1.4 allows us to enlarge the class of sup-measurable functions, by adopting an analogous construction to that leading to the known Baire classes: let B_0 denote the class of all Carathéodory functions, and B_α (α a countable ordinal number) the class of all functions f which admit a representation

$$f(s,u) = \lim_{n\to\infty} f_n(s,u) \qquad (s \in \Omega \setminus D_0\,,\ u \in \mathbb{R})\,, \tag{1.18}$$