FUNCTION SPACES

TERENCE TAO

1. Function spaces

When working with numbers such as real numbers $x \in \mathbf{R}$ or complex numbers $z \in \mathbf{C}$, there is an unambiguous notion of a *magnitude* $|x|$ or $|z|$ of a number, with which to measure which numbers are large and which are small. One can also use this notion of magnitude to define a distance $|x - y|$ or $|z - w|$ between two real numbers $x, y \in \mathbf{R}$, or between two complex numbers $z, w \in \mathbf{C}$, thus giving a quantitative measure of which pairs of numbers are close and which ones are far apart.

The situation becomes more complicated however when dealing with objects with more degrees of freedom. Consider for instance the problem of determining the "magnitude" of a three-dimensional rectangular box. There are several candidates for such a magnitude: length, width, height, volume, surface area, diameter (i.e. length of the long diagonal), eccentricity, and so forth. Unfortunately, these magnitudes do not give equivalent comparisons: box A may be longer and have more volume than box B, but box B may be wider and have more surface area, and so forth. Because of this, one abandons the idea that there should only be one notion of "magnitude" for boxes, and instead accept that there are instead a multiplicity of such notions, all of which have some utility. Thus for some applications one may wish to distinguish the large volume boxes from the small volume boxes, while in others one may wish to distinguish the eccentric boxes from the round boxes. Of course, there are several relationships between the different notions of magnitude (e.g. the isoperimetric inequality allows one to obtain an upper bound for the volume in terms of the surface area), so the situation is not as disorganized as it may first appear.

Now we turn to *functions* with a fixed domain and range (e.g. functions $f : [-1, 1] \to \mathbf{R}$ from the interval $[-1, 1]$ to the real line $\mathbf{R}$). These objects have infinitely many degrees of freedom, and so it should not be surprising that there are now infinitely many distinct notions of "magnitude", all of which provide a different answer to the question "how large is a given function f?", or to the closely related question "how close together are two functions f, g?". In some cases, certain functions may have infinite magnitude by one such measure, and finite magnitude by another; similarly, a pair of functions may be very close by one measure and very far apart by another. Again, this situation may seem chaotic, but it simply reflects the fact that functions have many distinct characteristics - some are tall, some are broad, some are smooth, some are oscillatory, and so forth - and depending on the application at hand, one may need to give more weight to one of these characteristics than to others. In analysis, this is embodied in the variety of standard *function spaces*, and their associated *norms*, which are available to describe functions both qualitatively and quantitatively. While these spaces and norms are mostly distinct from each other, they are certainly interrelated, for instance through such basic facts of analysis such as approximability by test functions (or in some cases by polynomials), by embeddings such as Sobolev embedding, and by interpolation theorems; we shall discuss these statements in more detail later.

More formally, a *function space* is a class X of functions (with fixed domain and range), together with a *norm*[1] which assigns a non-negative number $\|f\|_X$ to every function f in X; this number is the function space's way of measuring how large a function is. It is common (though not universal) for the class X of functions to consist precisely of those functions for which the definition of the norm $\|f\|_X$ makes sense and is finite; thus the mere fact that a function f has membership in a function space X conveys some qualitative information about that function (e.g. it may imply some regularity, some decay, some boundedness, or some integrability on the function f), while the norm $\|f\|_X$ supplements this qualitative information with a more quantitative measurement of the function (e.g. *how* regular is f? *how much decay* does f have? *by which constant* is f bounded? *what is the integral* of f?). Typically we assume that the function space X and its associated norm $\|\cdot\|_X$ obey a certain number of axioms; for instance, a rather standard set of axioms is that X is a real or complex vector space, that the norm is non-degenerate ($\|f\|_X > 0$ for non-zero f), homogeneous of degree 1, and obeys the triangle inequality $\|f+g\|_X \leq \|f\|_X + \|g\|_X$; furthermore, the space X when viewed using the metric $d(f,g) := \|f-g\|_X$ is a complete metric space. Spaces satisfying all of these axioms are known as *Banach spaces*, and enjoy a number of good properties. A majority (but certainly not all) of the standard function spaces considered in analysis are Banach spaces.

We now present a selected sample of commonly used function spaces. For simplicity we shall consider only spaces of functions from $[-1,1]$ to $\mathbf{R}$.

- **The space $C^0([-1,1])$ of continuous functions**. This is a very familiar space of functions, and one which is regular enough to avoid many of the technical subtleties associated with very rough functions. Given that continuous functions on a compact interval such as $[-1,1]$ are automatically bounded, it is perhaps not surprising that the most natural norm to place on this space is the *supremum norm*[2]
$$\|f\|_{L^\infty([-1,1])} := \sup\{|f(x)| : x \in [-1,1]\};$$

[1] Strictly speaking, this only describes normed function spaces, which are a major and important sub-class of function spaces, but not the only one. For instance one can consider the weaker notion of a *topological function space*, in which there is no precise notion of magnitude to determine which functions are large and which are small, but instead there is just the notion of *convergence*, to determine which sequences (or nets) of functions converge to zero, and which ones do not.

[2] Strictly speaking, the supremum (or least upper bound) should be replaced by the *essential supremum* (or essential least upper bound), i.e. the least $L > 0$ which is an *essential upper bound* in the sense that $|f(x)| \leq L$ for *almost every* x rather than every x, so that we allow a measure zero set of exceptional x. This subtle distinction is only necessary when we generalize beyond the

this is also the norm associated with uniform convergence, thus a sequence $f_1, f_2, \ldots$ converges uniformly to f if and only if $\|f_n - f\|_{L^\infty([-1,1])} \to 0$ as $n \to \infty$. A basic result here is the *Weierstrass M-test*, which asserts that if a sequence $f_1, f_2, \ldots$ is absolutely convergent in $C^0([-1,1])$ (i.e. $\sum_{n=1}^\infty \|f_n\|_\infty < M$ for some finite M) then it is also conditionally convergent in $C^0([-1,1])$ (i.e. there exists $f \in C^0([-1,1])$ such that $\|f - \sum_{n=1}^N f_n\|_{L^\infty([-1,1])} \to 0$ as $N \to \infty$). Actually, the same statement is true for any Banach space (and is in fact one of the defining characteristics of such spaces). This space is also a basic example of a *Banach algebra* - a Banach space which is also closed under multiplication.

- **The space $C^1([-1,1])$ of continuously differentiable functions**. This is a space which has more restrictive membership than $C^0([-1,1])$; functions must not only be continuous but continuously differentiable. (Merely being differentiable is not as useful trait as continuous differentiability; for instance, it is not enough to guarantee the fundamental theorem of calculus $f(b) - f(a) = \int_a^b f'(x)\, dx$). The supremum norm here is no longer natural, because a sequence of continuously differentiable functions (e.g. the partial sums of the Fourier series $\sum_{n=1}^\infty \frac{\sin(nx)}{n^2}$) can converge in this norm to a non-differentiable function. Instead, the natural norm here is the C^1 norm
$$\|f\|_{C^1([-1,1])} := \|f\|_{L^\infty([-1,1])} + \|f'\|_{L^\infty([-1,1])};$$
thus the C^1 norm measures both the size of a function *and* the size of its derivative. (Merely controlling the latter would be unsatisfactory, since it would give constant functions a norm of zero). Thus this is a norm which gives more weight to regularity than the supremum norm. One can similarly define the space $C^2([-1,1])$ of twice continuously differentiable functions, and so forth, all the way up to the space $C^\infty([-1,1])$ of infinitely differentiable functions; there are also fractional versions of these spaces, such as $C^{0,\alpha}([-1,1])$, the space of α-Hölder continuous functions (with $C^{0,1}([-1,1])$ being the space of Lipschitz continuous functions). We will not discuss these variants here.
- **The Lebesgue spaces $L^p([-1,1])$ of p^{th}-power integrable functions**. The supremum norm $\|f\|_\infty$ mentioned earlier gives uniform control on the sizes of $|f(x)|$ for all $x \in [-1,1]$. However, it has the feature of being unstable: if one changes the value of f on a very small set to be very large, this can dramatically increase the supremum norm of f even if f is very small elsewhere. It is thus sometimes more advantageous to work with norms that are less susceptible to such fluctuations, such as the L^p *norms*
$$\|f\|_{L^p([-1,1])} := (\int_{-1}^1 |f(x)|^p\, dx)^{1/p}$$
which we define for $1 \leq p < \infty$ and for any measurable f. (One can also define these norms for $0 < p < 1$ but they are less useful, for instance they no longer obey the triangle inequality). The function space $L^p([-1,1])$ is then the class of measurable functions for which the above norm is finite. The $L^\infty([-1,1])$ norm is the limiting case of the L^p norms as $p \to \infty$ (more

---

case when f is continuous, since for continuous functions the supremum and essential supremum coincide.

precisely, if f is bounded and measurable then $\|f\|_{L^p([-1,1])}$ converges to $\|f\|_{L^\infty([-1,1])}$ as $p \to \infty$). A particularly important norm is the L^2 norm which is the norm which most resembles the familiar Pythagorean norm in Euclidean space; indeed, there is an inner product

$$\langle f, g \rangle := \int_{-1}^{1} f(x)g(x)\ dx$$

which is to the L^2 norm as the Euclidean dot product is to the Pythagorean norm. This space is exceptionally rich in symmetries; there is a wide variety of *unitary transformations* which preserve this space. While the L^∞ norm is concerned solely with the "height" of a function, the L^p norms are instead concerned with a combination of the "height" and "width" of a function. Other such combinations are possible, leading to a wider class of *rearrangement-invariant spaces* which include Lorentz and Orlicz spaces (which we will not discuss here).

- **The Sobolev space $W^{k,p}([-1,1])$ of p^{th}-power integrable functions with k degrees of regularity**. The Lebesgue norms control to some extent the height and width of a function, but say nothing about regularity; there is no reason why a function in L^p should be differentiable or even continuous. To incorporate such information one often turns to the *Sobolev norms* $\|f\|_{W^{k,p}([-1,1])}$, defined for $1 \leq p \leq \infty$ and $k \geq 0$ by[3]

$$\|f\|_{W^{k,p}([-1,1])} := \sum_{j=0}^{k} \|\frac{d^j f}{dx^j}\|_{L^p([-1,1])},$$

and then $W^{k,p}([-1,1])$ is the space of functions for which this norm is finite. Thus, a function lies in $W^{k,p}$ if it and its first k derivatives are p^{th} power integrable. There is one subtlety, which is that we do not require f to be k times differentiable in the classical sense, but merely in the weak sense (the sense of distributions). For instance, the function $f(x) = |x|$ has a weak derivative $f'(x) = \mathrm{sgn}(x)$, which lies (for instance) in $L^\infty([-1,1])$, and thus f lies in $W^{1,\infty}([-1,1])$ (which is the space of *Lipschitz-continuous* functions). We require these generalized differentiable functions in order to ensure that the space $W^{k,p}([-1,1])$ is complete. Sobolev norms are particularly natural and useful in the analytical study of partial differential equations and mathematical physics, for instance the $W^{1,2}$ norm can be interpreted as (the square root of) an "energy" of a function. While we only defined these norms for k a non-negative integer, one can in fact generalize these norms to arbitrary real exponents k by means of fractional differentiation and fractional integration, or via the Fourier transform. One of the basic tools in this theory is the *Sobolev embedding theorem*, which describes which Sobolev norms control which other norms; for instance, a function which is in $W^{1,1}([-1,1])$ (so it is integrable, and its weak derivative is also integrable) will automatically lie in $C^0([-1,1])$. There are also many spaces of a similar flavor to Sobolev spaces (in that they quantify both integrability and regularity), such as Besov spaces, Triebel-Lizorkin spaces, and Hölder spaces, or the space of functions of bounded variation.

[3]This is not the only definition used in the literature; many other formulations exist, but they are all equivalent up to constants, which is usually good enough for most applications in analysis.

There are many ways in which knowledge of the structure of function spaces can assist in the study of functions. For instance, if one has a good basis for the function space, so that every function in the space is a (possibly infinite) linear combination of basis elements, and one has some quantitative estimates on how this linear combination converges to the original function, then this allows one to represent that function efficiently in terms of a number of co-efficients, and also allows one to approximate that function by smoother functions. For instance, one basic result about $L^2([-1,1])$ is the *Plancherel theorem*, which asserts among other things that for every function $f \in L^2([-1,1])$, the *Fourier series* $\sum_{n=-\infty}^{\infty} \hat{f}(n) e^{\pi i n x}$, where $\hat{f}(n) := \frac{1}{2}\int_{-1}^{1} f(x) e^{-\pi i n x}$ are the *Fourier coefficients* of f, is convergent to f in the L^2 sense, or in other words that

$$\|f(x) - \sum_{n=-N}^{N} \hat{f}(n) e^{\pi i n x}\|_{L^2([-1,1])} \to 0 \text{ as } N \to \infty.$$

In particular, this shows that any function in $L^2([-1,1])$ can be approximated to arbitrary accuracy in the L^2 sense by a trigonometric polynomial $\sum_{n=-N}^{N} \hat{f}(n) e^{\pi i n}$ of finite degree. This result can be viewed as a statement about the extent to which the functions $e^{\pi i n}, n \in \mathbf{Z}$ form a basis of $L^2([-1,1])$; in this case, they form a very good basis (they are essentially an orthonormal basis, in fact) and are very useful.

Another very basic fact about function spaces is that certain function spaces embed into others, so that membership in one space automatically conveys membership in other spaces also, and furthermore there is often some inequality that controls one norm by the other. For instance, membership in a high-regularity space such as $C^1([-1,1])$ automatically implies membership in a low regularity space such as $C^0([-1,1])$, whereas membership in a high-integrability space such as $L^\infty([-1,1])$ implies membership in a low integrability space such as $L^1([-1,1])$ (although this statement is false if $[-1,1]$ is replaced by a set of infinite measure, such as the real line $\mathbf{R}$). These inclusions are not reversible; however one does have the *Sobolev embedding theorem*, which allows one to "trade" regularity for integrability, showing that spaces with lots of regularity but low integrability automatically embed into spaces with low regularity but high integrability; a sample estimate of this type is

$$\|f\|_{L^\infty([-1,1])} \leq \|f\|_{W^{1,1}([-1,1])}.$$

Another very useful concept is that of *duality*. Given a function space X, one can define the dual space X^*, which formally is defined as the class of all *continuous linear functionals* on X, or more precisely all maps $\omega : X \to \mathbf{R}$ (or $\omega : X \to \mathbf{C}$, if the function space is complex-valued) which are linear and continuous with respect to the norm of X. While linear functionals are not, strictly speaking, functions, in many cases one can canonically identify linear functionals with functions (or to some generalized concept of function, such as a distribution). For instance, if $1 < p < \infty$ and $\omega : L^p([-1,1]) \to \mathbf{R}$ is a continuous linear functional on $L^p([-1,1])$, then there exists a unique[4] function $g \in L^q([-1,1])$, where $1 < q < \infty$ is the dual

[4]More precisely, unique up to sets of measure zero. When dealing with Lebesgue spaces such as $L^p([-1,1])$, one often declares two functions to be equivalent if they agree except on sets of Lebesgue measure zero.

exponent of p (so $\frac{1}{p} + \frac{1}{q} = 1$), such that

$$\omega(f) := \int_{-1}^{1} f(x)g(x)\ dx \text{ for all } f \in L^p([-1,1]).$$

The integral on the right-hand side turns out to always be absolutely convergent, thanks to *Hölder's inequality.* Because of this fact, we can canonically identify the dual $L^p([-1,1])^*$ of $L^p([-1,1])$ with $L^q([-1,1])$. One can sometimes analyze functions in a function space by instead studying how the linear functionals in the dual space act on those functions; related to this, one can often analyze a continuous linear operator $T : X \to Y$ from one function space to another by first considering the *adjoint operator* $T^* : Y^* \to X^*$, defined for all linear functionals $\omega : Y \to \mathbf{R}$ by letting $T^*\omega : X \to \mathbf{R}$ be the functional $T^*\omega(x) := \omega(Tx)$.

We mention one more important fact about function spaces, which is that certain function spaces X "interpolate" between two other function spaces X_0, X_1. The precise definition of this is technical, but roughly means that every function in X can be decomposed[5] in a number of specific ways as the sum of a function in X_0 and a function in X_1, and conversely any function which lies in both X_0 and X_1 will then lie in X. For instance, given any function $f \in L^2([-1,1])$ and any cutoff parameter $\lambda > 0$, we can partition $f := f_{\leq\lambda} + f_{>\lambda}$, where $f_{\leq\lambda}(x)$ is defined to be equal to $f(x)$ when $|f(x)| \leq \lambda$ and zero otherwise, whereas $f_{>\lambda}(x)$ is equal to $f(x)$ when $|f(x)| > \lambda$ and zero otherwise. One can then verify that $f_{\leq\lambda}$ lies in $L^\infty([-1,1])$ with a norm of at most λ, whereas $f_{>\lambda}$ lies in $L^1([-1,1])$ with a norm of at most $\|f\|^2_{L^2([-1,1])}/\lambda$. Conversely, if a function g lies in $L^\infty([-1,1])$ with norm at most λ and also lies in $L^1([-1,1])$ with norm at most A^2/λ, then one can easily verify that g will also lie in $L^2([-1,1])$ with norm at most A. These two facts basically assert that $L^2([-1,1])$ is an interpolation space between $L^\infty([-1,1])$ and $L^1([-1,1])$; this has a number of consequences, notably the *Marcinkiewicz interpolation theorem*, which among other things asserts that any linear operator which is bounded from $L^\infty([-1,1])$ to $L^\infty([-1,1])$, and also bounded from $L^1([-1,1])$ to $L^1([-1,1])$, is automatically also bounded from $L^2([-1,1])$ to $L^2([-1,1])$. Interpolation methods are remarkably powerful. For instance, they can be used to give a short proof of *Young's inequality*

$$\left(\int_{-\infty}^{\infty} |f * g(x)|^r\ dx\right)^{1/r} \leq \left(\int_{-\infty}^{\infty} |f(x)|^p\ dx\right)^{1/p} \left(\int_{-\infty}^{\infty} |g(x)|^q\ dx\right)^{1/q}$$

whenever $1 \leq p, q, r < \infty$ are such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$, and f and g are measurable functions for which the right-hand side is finite, and $f * g(x) := \int_{-\infty}^{\infty} f(y)g(x-y)\ dy$ is the convolution of f and g. It is more difficult (though not impossible) to prove this inequality without the aid of interpolation theory.

Department of Mathematics, UCLA, Los Angeles CA 90095-1555

E-mail address: `tao@math.ucla.edu`

[5]Strictly speaking, this is what it means for X to be a *real interpolation space* of X_0 and X_1. There is also the notion of being a *complex interpolation space*, which roughly means that every function in X can be embedded in an analytic family of functions, which lie in X_0 for some values of an analytic parameter and lie in X_1 for other values.