

Supplement to Chapter 5 of Sutherland,

Introduction to Metric and Topological Spaces (Second Edition)

The first section of this chapter (Motivation and definition, pp. 37 – 40) discusses the motivation behind the concept of a metric space from a fairly technical and formal viewpoint. The material in this document takes a somewhat complementary approach, with emphasis on mathematical issues which led to the definition and study of metric and topological spaces.

Origins of point set topology

In ordinary single variable calculus, the basic sets or domains for defining functions are **intervals**, which are sets of real numbers satisfying the following condition:

Intermediate Value Property: *If x and y belong to the set J and z satisfies $x < z < y$, then z also belongs to J .*

If we form the **extended real number system** by adjoining $+\infty$ and $-\infty$, then every interval has a unique lower endpoint and a unique upper endpoint in the extended real number system, and there are three distinct types of intervals in the real line:

Closed intervals: If $a < b$ are real numbers, then $[a, b]$ consists of all real numbers x such that $a \leq x \leq b$.

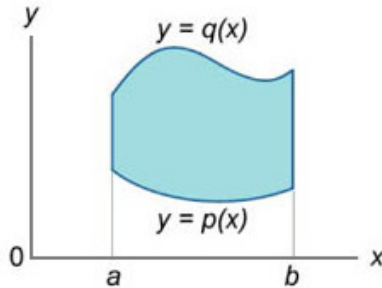
Open intervals: If $a < b$ are extended real numbers, then (a, b) consists of all real numbers x such that $a < x < b$.

Half – open intervals: (*Type 1*) If $a < b$ where a is a real number and b is an extended real number, then $(a, b]$ consists of all real numbers x which satisfy $a < x \leq b$. (*Type 2*) If $a < b$ where a is an extended real number and b is a real number, then $[a, b)$ consists of all real numbers x which satisfy $a \leq x < b$.

Sometimes it is also useful to consider functions defined on unions of pairwise disjoint intervals (for example, the function $1/x$ is defined on the union of $(-\infty, 0)$ and $(0, +\infty)$ and similarly for other quotients of polynomials), but in any case the list of possibilities is still fairly short.

Things are far more complicated in multivariable calculus, particularly if one wants to describe reasonable choices of domains for multiple integration or partial differentiation. Since the main issues are already clear for functions of two variables, we shall consider this special case first. Here are a few obvious possibilities for sets in the coordinate plane on which “nice” functions of two variables can be defined:

1. **Closed rectangles:** Sets defined by pairs of inequalities $a \leq x \leq b$ and $c \leq y \leq d$.
2. **Closed disks:** Sets defined by inequalities $(x - a)^2 + (y - b)^2 \leq r^2$.
3. **Generalizations:** Sets defined by pairs of inequalities $a \leq x \leq b$ and $p(x) \leq y \leq q(x)$ for continuous functions $p(x)$ and $q(x)$ defined on $[a, b]$. These are the types of sets over which one typically integrates functions in multivariable calculus (see the drawing below).



(Source: <http://www.math24.net/iterated-integrals.html>)

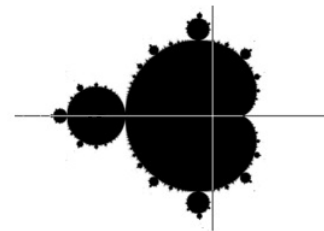
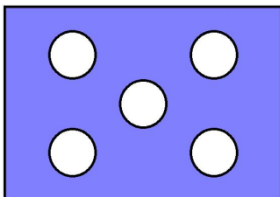
4. **Open analogs:** In the notation of the previous item, sets defined by pairs of **strict** inequalities $a < x < b$ and $p(x) < y < q(x)$; the difference between these and the preceding examples are that the open regions do not include the boundary curves.
5. **Partially open analogs:** In the notation of the previous item, these include sets defined by mixtures of strict and non – strict inequalities, somewhat analogous to half – open intervals; examples of this type include some but not all of the points on the boundary curves.

As noted above, the third class of examples is good for calculating multiple integrals, but they are not necessarily good for partial differentiation. For the latter, it is usually necessary know that the function is defined on a small disk centered at the point at which one is taking derivatives. The fourth class of examples turns out to satisfy this condition.

One way of seeing the added complications in two dimensions is to observe that *there are infinitely many partially open sets between each open example in fourth class and its closed analog in the third class*. Each subset of the boundary curve(s) will yield a distinct partially open subset between the open and closed regions (and there are infinitely many such subsets).

Still further examples. It is easy to draw other potential examples of sets that are reasonable for purposes of multiple integration or partial differentiation. A few are depicted below; in each case it is possible to compute the areas of the closed regions (including boundary points), and if the boundary points are removed the resulting open regions satisfy the previously given condition for taking partial and directional derivatives. The third example is the **Mandelbrot set** which plays a key role in the theory of fractals developed by B. Mandelbrot (1924 – 2010); the drawing below is adapted from the following site:

http://upload.wikimedia.org/wikipedia/commons/thumb/5/56/Mandelset_hires.png/322px-Mandelset_hires.png



In many respects, point set topology is a systematic abstraction of basic methods for analyzing and classifying the types of sets that are most useful for the purposes of multivariable calculus. However, there are several other factors which also contributed significantly to the development of the subject:

1. In single variable calculus, it is often necessary to understand the sets of points where functions are discontinuous or where an infinite series of functions does not converge.
2. In classical Euclidean geometry, at certain points it is necessary to understand simple examples of regions like the two sides of a line in the plane or the interior and exterior regions associated to angles or triangles. Classical Greek geometry only dealt with such notions to a limited extent; coordinate geometry provided more effective ways of describing these and other regions, and point set topology gives abstract, qualitative methods for working with them.
3. In the highly influential setting for geometry proposed by G. F. B. Riemann (1826 – 1866), the fundamental objects (**n – dimensional manifolds** in modern language) are equipped with measurement data which include a notion of distance, and — in analogy with the ordinary 2 – dimensional sphere — a small region around each point is assumed to be equivalent to a small region in n – dimensional Euclidean space. Metric spaces provide a setting in which these concepts can be made mathematically precise.
4. By the end of the 19th century, mathematicians had noticed that several basic ideas and results involving equations in finitely many unknowns have important analogs in the study of certain differential and integral equations. A unified approach to such notions clearly makes everything more efficient, and it also provides a framework for discovering still further similarities between solving equations in finitely many unknowns and solving differential and integral equations.
5. In the study of differential equations arising from various physical problems — most notably celestial mechanics — there are many cases in which one cannot give an explicit analytic formula for the motion of objects in the system (for example, this is true for gravitational systems with three or more objects). The lack of such formulas often makes it difficult to draw qualitative conclusions about the system’s behavior, and often considerations involving metric or topological spaces turn out to be useful in obtaining insights into such qualitative questions.

The existence of a required course on metric topological spaces reflects the usefulness of these concepts for studying the preceding questions and still others which have arisen in the meantime.

The central concepts of this course

This course concentrates on two levels of abstraction and generalization based upon certain properties of coordinate (or Euclidean) n – space \mathbb{R}^n and its subsets. The emphasis at the first level involves the abstract properties the distance between two points, and it is based upon the notion of **metric space** introduced by M. Frechet (1878 – 1973) in 1905. Although such objects satisfy an extremely short list of axioms, the latter suffice to introduce fundamental concepts like open and closed regions, boundary points, continuous functions, infinite sequences, and the convergence of such sequences. Furthermore, in this setting one can prove extremely useful generalizations of many fundamental theorems about continuous functions of a single real variable, including the following:

Extreme value property. *If g is a continuous real valued function defined on a closed interval, then g attains a maximum value and a minimum value at points of that interval.*

Intermediate value property. *Let g be a continuous real valued function defined on an interval, and let p and q be points of that interval such that $g(p) < g(q)$. Then for each y between $g(p)$ and $g(q)$ there is some r between p and q such that $g(r) = y$.*

The second level of abstraction — which for the most part restricts attention to regions, boundaries and continuity — is based upon the notion of **topological space**, which was first considered by F. Hausdorff (1868 – 1942) in 1914 and modified into its current form by K. Kuratowski (1896 – 1980) in 1922. By the early 1950s the precise relationship between metric and topological spaces had been determined, and at that point the general theory of metric and topological spaces was essentially in its current form. Over the past 60 years, the fundamental role of this theory in many parts of mathematics has become even more apparent.

In this course, metric spaces are the main objects of interest, with topological spaces serving as an alternative setting for statements and proofs which do not require the full force of the metric space axioms (however, use of this more generally setting is normally limited to concepts and results for which the added generality make things simpler rather than more complicated).