Products, Relations and Functions

For a variety of reasons, in this course it will be useful to modify a few of the set-theoretic preliminaries in the first chapter of Munkres. The discussion below explains these differences specifically.

Cartesian Products and Relations

Given two set-theoretic objects a and b, there is a set-theoretic construction which yields an

$$\underline{\textit{ordered pair}}\ (a, b)$$

which has the fundamental property

$$(a, b) = (c, d) \textit{ if and only if } a = c \textit{ and } b = d.$$

Given two sets A and B, the **Cartesian product** $A \times B$ is defined to be the set of all ordered pairs (a, b) where $a \in A$ and $b \in B$.

We shall define a ***binary relation*** from A to B to be a triple

$$\mathcal{R} = (A, R_0, B)$$

where A and B are sets and R_0 is a subset of $A \times B$. This differs from the standard definitions and the one in Munkres in two respects. First of all, the description in Munkres requires $A = B$. Also, we take the sets A and B as specific pieces of extra structure. Following standard notation we say that R_0 is the ***graph*** of $\mathcal{R}$, and for $a \in A$ and $b \in B$ we write

$$a\,\mathcal{R}\,b \textit{ if and only if } (a, b) \in R_0 \text{ and say that } a\,\mathcal{R}\,b \textit{ is true} \text{ or } a \textit{ is } \mathcal{R}-\underline{\textbf{\textit{related}}} \textit{ to } b.$$

If $B = A$ then we often say that $\mathcal{R}$ is a ***binary relation*** on A. Important special cases of binary relations including partial orderings and equivalence relations are discussed in Munkres.

Functions

Formally speaking, a function is a special type of relation. The main difference between our definition and the standard one is that the source set A (formally, the ***domain***) and the target set B (formally, the ***codomain***) are included explicitly as part of the structure. This point is often ignored in discussions of set theory (but an equivalent piece of structure is mentioned on page 16 of Munkres). Given the importance of the concept of function in mathematics, we shall give a complete formal definition.

Definition. A ***function*** is a triple $f = (A, \Gamma, B)$ where A and B are sets and Γ is a subset of $A \times B$ with the following property:

($$) *For each* $a \in A$ *there is a* **unique** *element* $b \in B$ *such that* $(a, b) \in \Gamma$.

The sets A and B are respectively called the ***domain*** and ***codomain*** of f, and Γ is called the ***graph*** of f. Frequently we write $f : A \to B$ to denote a function with domain A and codomain B, and as usual we write $b = f(a)$ if and only if (a, b) lies in the graph of f.

Given a function and a subset of its domain or codomain, we shall often need to work with the image or inverse image of the subset as defined on page 19 or Munkres. As indicated in the course notes, we shall deviate slightly from the notation in Munkres, using $f[S]$ to denote the image of a subset $S \subset A$ under f and $f^{-1}[T]$ to denote the inverse image of a subset $T \subset B$ under f. Many important properties of these constructions are summarized in the exercises on pages 20 and 21 of Munkres. One further property of the same sort is the following identity:

If $f : A \to B$ *and* $g : B \to C$, *then* $g(\, f[S]\,) = (g \circ f)[S]$ *for all* $S \subset A$.

Conceptually, the crucial point in our definition is that ***it keeps track of the codomain.*** This turns out to be crucial for many mathematical constructions (for example, the fundamental group that is discussed in the second half of Munkres' book). In particular, if B is a subset of B^* then Γ can be viewed as the graph of a function $A \to B^*$ as well as the graph of a function $f : A \to B$.

Notation. If $f : A \to B$ is a function whose image lies in $C \subset B$ (*i.e.*, the graph of f is actually a subset of $A \times C$), then the associated function (A, Γ, C) will be denoted by $C \,|\, f$ and called the ***corestriction*** of f to C. The notation is meant to suggest that this is somehow complementary (or "dual") to the notion of restriction defined on page 17 of Munkres; in our notation the latter can be expressed as

$$f|S = \big(S, \Gamma \cap (S \times B), B\big)$$

where $S \subset A$.

Disjoint unions

In many situations it is useful or even necessary to have a set S created from disjoint copies of two given sets A and B. Formally, the ***disjoint sum*** (or ***disjoint union***) is defined to be the set

$$A \sqcup B = A \coprod B = A \times \{1\} \cup B \times \{2\}$$

and the standard ***injection*** mappings $i_A : A \to A \sqcup B$ and $i_B : B \to A \sqcup B$ are defined by

$$i_A(a) = (a, 1) \text{ and } i_B(b) = (b, 2)$$

respectively. By construction, the maps i_A and i_B determine $1-1$ correspondences from A to $i_A[A]$ and from B to $i_B[B]$, the images of A and B are disjoint (because the second coordinates of the ordered pairs are unequal), and their union is all of $A \sqcup B$..

The Natural Numbers

It is nearly impossible to do much mathematical work at all without using the natural numbers (the nonnegative integers) in one way or another. At this point we shall only need enough information about them to work with finite sets and mathematical induction, and it will suffice to view the natural numbers using the axioms for them formulated by G. Peano.

PEANO AXIOMS. The natural numbers are given by a pair $(\mathbf{N}, \sigma)$ consisting of a set $\mathbf{N}$ and a function $\sigma : \mathbf{N} \to \mathbf{N}$ with the following properties [which reflect the nature of sigma as a map taking each natural number m to its "successor" $m + 1$]:

(1) *There is exactly one element (the **zero element**) not in the image of* σ.

(2) *The map sigma is one-to-one.*

(3) *If A is a subset of $\mathbf{N}$ such that*

(i) $0 \in A$,

(ii) *for all* $k \in \mathbf{N}$, $k \in A$ *implies* $\sigma(k) \in A$,

then $A = \mathbf{N}$.

The third axiom is just the ***(weak) principle of finite mathematical induction.***

The Peano axioms are sufficient to yield all the properties of natural numbers that are used in Munkres. Strictly speaking, the proofs of results like Theorem 4.1 and 4.2 require our axioms for the real number system, but we shall not have any need for these results until after the real number system has been introduced. All of the results in

Sections 6 and 8 of Munkres — except those requiring the real numbers explicitly — follow as in the text.

The Axiom of Choice and Logical Consistency

Although set theory provides an effective framework for discussing questions involving finite sets, its initial and most important motivation came from questions about infinite sets. As research on such sets progressed during the late nineteenth and early twentieth century, it eventually became evident that most of the underlying principles involved constructing new sets from old ones and the existence of the set of natural numbers. However, it also became clear that some results depended upon the Axiom of Choice **(AC),** which is an abstract and basically nonconstructive existence statement.

There are many equivalent ways of formulating set-theoretic assumptions that are logically equivalent to the Axiom of Choice. Here is a list of several that appear in Munkres:

Axiom of Choice. *If A is a nonempty set and $\mathcal{P}_{+}(A)$ denotes the set of all nonempty subsets of A, there is a function $f : \mathcal{P}_{+}(A) \rightarrow A$ such that $f(B) \in B$ for all nonempty $B \subset A$.*

"Zermelo's Well-ordering Theorem." *For every nonempty set A, there is a linear ordering such that each nonempty subset B of A has a least element (**i.e.,** there is a **well-ordering** of A).*

Hausdorff Maximal Principle. *Every partially ordered set has a maximal linearly ordered subset.*

"Zorn's Lemma." *If A is a partially ordered set in which linearly ordered subsets have upper bounds, then A has a maximal element.*

For some time there was uncertainty whether the Axiom of Choice should be included in the axioms for set theory. Concern over this point increased with the discovery of apparent paradoxes in set theory around the beginning of the twentieth century. One example of this involves attempts to discuss the "set of all sets." Most of the paradoxes were resolved by a careful foundation of the axioms for set theory, but it was not known if adding **AC** might still lead to a logical contradiction.

The discovery of the so-called ***Banach-Tarski paradox*** in the nineteen twenties illustrated that **AC** had extremely strong consequences that raised questions about whether it should be taken as an axiom for set theory. In its original form, the result of S.

Banach and A. Tarski states that if **AC** is assumed, then it is possible to take a solid ball in 3-dimensional space, cut it up into finitely many pieces and, moving them using only rotation and translation, reassemble the pieces into two balls the same size as the original one **!!** At first glance this may appear to violate the laws of physics, but the sets in question are mathematical rather than physical objects. In particular, there is no meaningful way to define the volumes of the individual pieces, and it is impossible to carry out the construction physically because if one does cut the solid ball into pieces physically (say with a knife or saw), each piece has a specific volume (physically, one can find the volumes by sticking the pieces into a large cylinder with the right amount of water). Even though the Banach-Tarski paradox does not yield a contradiction to the axioms of set theory, it does raise two fundamental questions:

1. If set theory with **AC** yields bizarre conclusions like the existence of the sets described above, is it possible that further pursuits will lead to a contradiction?

2. Is it worthwhile to consider such objects, and if not is it appropriate to have an axiomatic system for set theory that implies the existence of such physically unreal entities?

One way of answering the second question is that **AC** also implies the existence of many things that mathematicians do want for a variety of reasons. For example, one needs **AC** to conclude that every (infinite-dimensional) vector space has a basis. Although some mathematicians think that the subject should only consider objects given by suitably "constructive" methods, the existence results that follow from **AC** are so useful that mathematicians would prefer to include it as part of the axioms if at all possible.

Of course, if **AC** leads to a logical contradiction, then it should not be part of the axioms for set theory, so this brings us back to the first question. Two extemely important and fundamental pieces of research by K. Gödel in the nineteen thirties clarified the role of **AC.** The first of these was his work on the incompleteness properties of axiomatic systems, and the essential conclusion is that mathematics can ***never*** be sure that ***any*** reasonable set of axioms for set theory is logically consistent. His subsequent result showed that **AC** was ***relatively consistent*** with the other axioms for set theory; specifically, if there is a logical contradiction in set theory with the inclusion of **AC,** then there is also a logical contradiction if one does not assume **AC.** If there is an internal contradiction in the axioms for set theory, it must arise either from the assumptions about constructing sets by specifying them in terms of logical statements or from the basic assumption that one can carry out fundamental set theoretic constructions on **N** . Since most mathematicians would prefer to include as many objects as possible in set theory so long as these objects do not lead to a logical contradiction, the effective consequence of relative consistency is that inclusion of **AC** in the axioms for set theory is appropriate.

Subsequent work of P. Cohen in the nineteen sixties completed our current understanding of the role of **AC.** Specifically, he showed that one can construct models for set theory such that **AC** was true for some models and false for others.

The Continuum Hypothesis

Another question about set theory that arose very early in the study of the subject was the ***Continuum Hypothesis :***

(CH) *If* ***A*** *is an infinite subset of the real numbers* **R** *and there is* **no** *one-to-one correspondence between* ***A*** *and the natural numbers* **N***, then there* **is** *a one-to-one correspondence between* ***A*** *and* **R.**

Since there is a one-to-one correspondence between **R** and the set $\mathcal{P}$(**N**) of all subsets of **N**, one can reformulate this as the first case of a more sweeping conjecture known as the ***Generalized Continuum Hypothesis :***

(GCH) *If* ***S*** *is an infinite set and* ***T*** *is a subset of* $\mathcal{P}(S)$, *then* ***either*** (***i***) *there is a one-to-one correspondence between* ***T*** *and a subset of* ***S***, ***or else*** (***ii***) *there is a one-to-one correspondence between* ***T*** *and* $\mathcal{P}(S)$.

G. Cantor originally formulated **CH** in his work establishing set theory, the motivation being that he could not find any subsets whose cardinal numbers were between those of **N** and **R.** As in the case of **AC,** the work of Gödel showed that if a contradiction to the axioms for set theory arose if one assumes **CH** or **GCH,** then one can also obtain a contradiction without such an extra assumption, and the work of P. Cohen shows that one can construct models for set theory such that **CH** was true for some models and false for others. In fact, one can construct models for which the number of cardinalities between those of **N** and **R** can vary to some extent (for example, there might be one or two cardinalities between them). Because of Cohen's work, most mathematicians are ***not*** willing to assume **CH** or **GCH** for the same reason that they are willing to assume **AC:** They would prefer to include as many objects as possible in set theory so long as these objects do not lead to a logical contradiction.