Chicago Lectures in Physics

Robert M. Wald, Editor
Hellmut Fritzsche
Riccardo Levi-Setti
Roland Winston

Mathematical Physics

Robert Geroch

The University of Chicago Press
Chicago and London

Robert Geroch is professor in the Departments of Physics and Mathematics at the University of Chicago.

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

94 93 92 91 90 89 88 87 86 85   5 4 3 2 1

Library of Congress Cataloging in Publication Data

/fRGeroch, Robert.

Mathematical physics.

(Chicago lectures in physics)
1. Mathematical physics. I. Title. II. Series.
QC20.G47   1985   530.1'5   85-17764
ISBN 0-226-28861-7
ISBN 0-226-28862-5 (pbk.)

Contents

1

Introduction

One sometimes hears expressed the view that some sort of uncertainty principle operates in the interaction between mathematics and physics: the greater the mathematical care used to formulate a concept, the less the physical insight to be gained from that formulation. It is not difficult to imagine how such a viewpoint could come to be popular. It is often the case that the essential physical ideas of a discussion are smothered by mathematics through excessive definitions, concern over irrelevant generality, etc. Nonetheless, one can make a case that mathematics as mathematics, if used thoughtfully, is almost always useful—and occasionally essential—to progress in theoretical physics.

What one often tries to do in mathematics is to isolate some given structure for concentrated, individual study: what constructions, what results, what definitions, what relationships are available in the presence of a certain mathematical structure—and only that structure? But this is exactly the sort of thing that can be useful in physics, for, in a given physical application, some particular mathematical structure becomes available naturally, namely, that which arises from the physics of the problem. Thus mathematics can serve to provide a framework within which one deals only with quantities of physical significance, ignoring other, irrelevant things. One becomes able to focus on the physics. The idea is to isolate mathematical structures, one at a time, to learn what they are and what they can do. Such a body of knowledge, once established, can then be called upon whenever it makes contact with the physics.

An everyday example of this point is the idea of a derivative. One could imagine physicists who do not understand, as mathematics, the notion of a derivative and the properties of derivatives. Such physicists could still formulate physical laws, for example, by speaking of the "rate of change of . . . with . . ." They could use their physical intuition to obtain, as needed in various applications, particular properties of these "rates of change." It would be more convenient, however, to isolate the notion "derivative" once and for all, without direct reference to later physical applications of this concept. One learns what a derivative is and what its properties are: the geometrical significance of a derivative, the rule for taking the derivative of a product, etc. This established body of knowledge then comes into play automatically when the physics requires the use of derivatives. Having mastered the abstract concept "rate of change" all by itself, the mind is freed for the important, that is, the physical, issues.

The only problem is that it takes a certain amount of effort to learn mathematics. Fortunately, two circumstances here intervene. First, the mathematics one needs for theoretical physics can often be mastered simply by making a sufficient effort. This activity is quite different from, and far more straightforward than, the originality and creativity needed in physics itself. Second, it seems to be the case in practice that the mathematics one needs in physics is not of a highly sophisticated sort. One hardly ever uses elaborate theorems or long strings of definitions. Rather, what one almost always uses, in various areas of mathematics, is the five or six basic definitions, some examples to give the definitions life, a few lemmas to relate various definitions to each other, and a couple of constructions. In short, what one needs from mathematics is a general idea of what areas of mathematics are available and, in each area, enough of the flavor of what is going on to feel comfortable. This broad and largely shallow coverage should in my view be the stuff of "mathematical physics."

There is, of course, a second, more familiar role of mathematics in physics: that of solving specific physical problems which have already been formulated mathematically. This role encompasses such topics as special functions and solutions of differential equations. This second role has come to dominate the first in the traditional undergraduate and graduate curricula. My purpose, in part, is to argue for redressing the balance.

We shall here take a brief walking tour through various areas of mathematics, providing, where appropriate and available, examples in which this mathematics provides a framework for the formulation of physical ideas.

By way of general organization, chapters 2–24 deal with things algebraic and chapters 25–42 with things topological. In chapters 43–50 we discuss some special topics: structures which combine algebra and topology, Lebesque integrals, Hilbert spaces. Lest the impression be left that no difficult mathematics can ever be useful in physics, we provide, in chapters 51–56, a counterexample: the spectral theorem. Strictly speaking, the only prerequisites are a little elementary set theory, algebra, and, in a few places, some elementary calculus. Yet some informal contact with such objects as groups, vector spaces, and topological spaces would be most helpful.

The following texts are recommended for additional reading: A. H. Wallace, *Algebraic Topology* (Elmsford, NY: Pergamon, 1963), and C. Goffman and G. Pedrick, *First Course in Functional Analysis* (Englewood Cliffs, NJ: Prentice-Hall, 1965). Two examples of more advanced texts, to which the present text might be regarded as an introduction, are: M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (New York: Academic, 1972), and Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (Amsterdam: North-Holland, 1982).

2

Categories

In each area of mathematics (e.g., groups, topological spaces) there are available many definitions and constructions. It turns out, however, that there are a number of notions (e.g., that of a product) that occur naturally in various areas of mathematics, with only slight changes from one area to another. It is convenient to take advantage of this observation. Category theory can be described as that branch of mathematics in which one studies certain definitions in a broader context—without reference to the particular area to which the definition might be applied. It is the "mathematics of mathematics." Although this subject takes a little getting used to, it is, in my opinion, worth the effort. It provides a systematic framework that can help one to remember definitions in various areas of mathematics, to understand what many constructions mean and how they can be used, and even to invent useful definitions when needed. We here summarize a few facts from category theory.

A ***category*** consists of three things—i) a class O (whose elements will be called ***objects***), ii) a set $\mathrm{Mor}(A,B)$ (whose elements will be called ***morphisms*** from A to B), where A and B are any two[1] objects, and iii) a rule which assigns, given any objects A, B, and C and any morphism φ from A to B and morphism ψ from B to C, a morphism, written $\psi \circ \varphi$, from A to C (this $\psi \circ \varphi$ will be called the ***composition*** of φ with ψ)—subject to the following two conditions:

1. Composition is associative. If A, B, C, and D are any four objects, and φ, ψ, and λ are morphisms from A to B, from B to C, and from C to D, respectively, then

$$(\lambda \circ \psi) \circ \varphi = \lambda \circ (\psi \circ \varphi) \quad .$$

(Note that each side of this equation is a morphism from A to D.)

2. Identities exist. For each object A, there is a morphism i_A from A to A (called the ***identity*** morphism on A) with the following property: if φ is any morphism from A to B, then

$$\varphi \circ i_A = \varphi \quad ;$$

if μ is any morphism from C to A, then

1. Here and hereafter, "two elements" means "two elements in a specific order," or, more formally, an "ordered pair."

$$i_A \circ \mu = \mu \quad .$$

That is the definition of a category. It all seems rather abstract. In order to see what is really going on with this definition—why it is what it is—one has to look at a few examples. We shall have abundant opportunity to do this: almost every mathematical structure we look at will turn out to be an example of a category. In order to fix ideas for the present, we consider just one example (the simplest, and probably the best).

To give an example of a category, one must say what the objects are, what the morphisms are, what composition of morphisms is—and one must verify that conditions 1 and 2 above are indeed satisfied. Let the objects be ordinary sets. For two objects (now, sets) A and B, let Mor(A,B) be the set of all mappings from the set A to the set B. (Recall that a mapping from set A to set B is a rule that assigns, to each element of A, some element of B.) Finally, let composition of morphisms, in this example, be ordinary composition of mappings. (That is, if φ is a mapping from set A to set B and ψ is a mapping from set B to set C, then $\psi \circ \varphi$ is the mapping from set A to set C which sends the element a of A to the element $\psi(\varphi(a))$ of C.) We now have the objects, the morphisms, and the composition law. We must check that conditions 1 and 2 are satisfied. Condition 1 is indeed satisfied in this case: it is precisely the statement that composition of mappings on sets is associative. Condition 2 is also satisfied: for any set A, let i_A be the identity mapping (i.e., for each element a of A, $i_A(a) = a$) from A to A. Thus we have here an example of a category. It is called the ***category of sets***.

This example is in some sense typical. It is helpful to think of the objects as being "really sets" (perhaps, as in later examples, with additional structure) and of the morphisms as "really mappings" (which, in these later examples, will be "structure preserving"). With this mental picture, it is easy to remember the definition of a category—and to follow the constructions we shall shortly introduce on categories.

This example suggests the introduction of the following notation for categories. We shall write $A \xrightarrow{\varphi} B$ to mean "A and B are objects, and φ is a morphism from A to B."

We now wish to give a few examples of how one carries over notions from categories in general to specific categories.

Let φ be a morphism from A to B. This φ is said to be a ***monomorphism*** if the following property is satisfied: given any object X and any two morphisms, α and α', from X to A such that $\varphi \circ \alpha = \varphi \circ \alpha'$, it follows that $\alpha = \alpha'$ (figure 1). This φ is said to be an ***epimorphism*** if the following property is satisfied: given any object X and any two morphisms, β and β', from B to X such that $\beta \circ \varphi = \beta' \circ \varphi$, it follows that $\beta = \beta'$ (figure 2). (That is, monomorphisms are the things that can be "canceled out of morphism equations on the left"; epimorphisms can be "canceled out of morphism equations

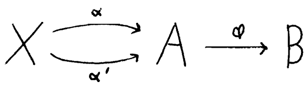

Figure 1

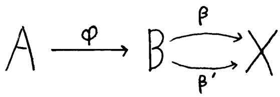

Figure 2

on the right.")

As usual, one makes sense out of these definitions by appealing to our example, the category of sets.

THEOREM 1. *In the category of sets, a morphism is a monomorphism if and only if it is one-to-one.*

(Recall that a mapping from set A to set B is said to be one-to-one if no two distinct elements of A are mapped to the same element of B.)

Proof. Let φ be a mapping from set A to set B, which is one-to-one. We show that this φ is a monomorphism. Let X be any set, and let α and α' be mappings from X to A such that $\varphi \circ \alpha = \varphi \circ \alpha'$. We must show that $\alpha = \alpha'$. If α and α' were different, they would differ on some element of X; that is, there would be an x in X such that $\alpha(x)$ would be different from $\alpha'(x)$. Then, since φ is one-to-one, we would have $\varphi(\alpha(x))$ different from $\varphi(\alpha'(x))$. But this contradicts $\varphi \circ \alpha = \varphi \circ \alpha'$. Hence φ is a monomorphism.

Let φ be a mapping from set A to set B which is a monomorphism. We show that this φ is one-to-one. Let a and a' be elements of A such that $\varphi(a) = \varphi(a')$. We must show that $a = a'$. Let X be the set having only one element, x. Let α be the mapping from X to A with $\alpha(x) = a$, and let α' be the mapping from X to A with $\alpha'(x) = a'$. Then, since $\varphi(a) = \varphi(a')$, $\varphi \circ \alpha(x) = \varphi \circ \alpha'(x)$. That is, $\varphi \circ \alpha = \varphi \circ \alpha'$. But φ is supposed to be a monomorphism; hence $\alpha = \alpha'$. In particular, we must have $\alpha(x) = \alpha'(x)$; that is, we must have $a = a'$. Hence, φ is one-to-one. □

THEOREM 2. *In the category of sets, a morphism is an epimorphism if and only if it is onto.*

(Recall that a mapping from set A to set B is said to be onto if every element of B is the image, under the mapping, of some element of A.)

Proof. Let φ be a mapping from set A to set B, which is onto. We show that this φ is an epimorphism. Let X be any set, and let β and β' be mappings from B to X such that $\beta \circ \varphi = \beta' \circ \varphi$. We must show that $\beta = \beta'$. If β and β' were different, they would differ on some element of B; that is, there would be a b in B such that $\beta(b)$ would be different from $\beta'(b)$. But, since φ is onto, there is an a in A such that $\varphi = b$. Hence $\beta \circ \varphi(a)$ would be different from $\beta' \circ \varphi(a)$. This contradicts $\beta \circ \varphi = \beta' \circ \varphi$. Hence φ is an epimorphism.

Let φ be a mapping from set A to set B, which is an epimorphism. We show that this φ is onto. Suppose, on the contrary, that there were some element b of B which was not the image, under φ, of any element of A. Let X be the set having just two elements, x and y. Let β be the mapping from B to X which sends b to x and the rest of B to y. Let β' be the mapping from B to X which sends all of B to y. Then, since φ sends no element of A to b, we have $\beta \circ \varphi = \beta' \circ \varphi$ (for both of these mappings from A to X send all of A to y). Since φ was assumed an epimorphism, we must have $\beta = \beta'$. But, by construction, β does not equal β'. We have a contradiction. Hence φ is onto. □

It should be clear that there is no real content to these proofs: all one has to do to obtain a proof is keep from getting confused. One should think of a monomorphism as a fancy way of saying "one-to-one" and of an epimorphism as a fancy way of saying "onto." Why does one bother to invent fancy words and fancy ways of saying these simple things? The point is that the definition of, for example, monomorphism is different in an important way from the definition of one-to-one. The latter refers directly to the sets themselves (i.e., to the elements of the sets and what happens to those elements under mappings). The former, however, refers only to objects, morphisms, and composition of morphisms. That is, "monomorphism" is a statement one can make about a category, while "one-to-one" is a statement one can only make about sets. As we shall see, there are lots of categories other than the category of sets: in every such category, the notion of a monomorphism will be available. Theorems 1 and 2 are good examples of the sort of activity that takes place in category theory. One takes a notion (e.g., one-to-one or onto) that refers directly to the detailed nature of the objects to which it is applied and finds a "categorical version" of that notion, a version that refers only to the things that go into making a category (objects, morphisms, composition). One thus acquires the ability to carry over this same notion to many different areas of mathematics.

Recall that any composition of one-to-one mappings on sets is itself one-to-one, and similarly for onto. This observation about sets and mappings suggests a theorem in category theory.

THEOREM 3. *Let* $A \xrightarrow{\varphi} B \xrightarrow{\psi} C$. *Then* $\psi \circ \varphi$ *is a monomorphism if both* φ *and* ψ *are monomorphisms;* $\psi \circ \varphi$ *is an epimorphism if both* φ *and* ψ *are epimorphisms.*

Proof. Let φ and ψ be monomorphisms. We show that $\psi \circ \varphi$ is a monomorphism. Let X be any object, and α and α' morphisms from X to A

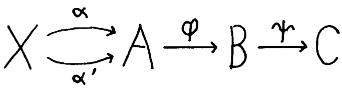

Figure 3

such that $(\psi \circ \varphi) \circ \alpha = (\psi \circ \varphi) \circ \alpha'$ (figure 3). We must show that $\alpha = \alpha'$. By condition 1 for a category, $\psi \circ (\varphi \circ \alpha) = \psi \circ (\varphi \circ \alpha')$. Since ψ is a monomorphism, we have $\varphi \circ \alpha = \varphi \circ \alpha'$. But now, since φ is a monomorphism, we have $\alpha = \alpha'$. Hence $\psi \circ \varphi$ is a monomorphism. Similarly for epimorphism. □

Note that theorem 3 is actually easier for categories in general than it is for the special case of sets. This phenomenon is by no means rare.

Everyone knows what a subset is. We formulate this notion categorically. A *subobject* of an object A is an object A' along with a monomorphism $A' \to A$. Since, in the category of sets, monomorphisms are just one-to-one mappings, it is clear that, in the category of sets, subobjects are just subsets.

There is, in set theory, the notion of a "one-to-one correspondence" between sets. We formulate categorically. A morphism φ from A to B is said to be an *isomorphism* if there is a morphism φ' from B to A such that

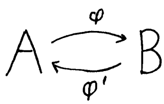

Figure 4

$\varphi' \circ \varphi = i_A$ and $\varphi \circ \varphi' = i_B$ (figure 4). (In words, a morphism is an isomorphism if it has "an inverse, which works on both the left and the right.") It is clear that, in the category of sets, an isomorphism is just a one-to-one, onto mapping. In fact, the statement of what an isomorphism is, when applied to the case of sets, is just what one really means by a "correspondence" between sets.

Example. A set is said to be *countable* if there exists an isomorphism between that set and the set of positive integers. Thus the set of real numbers is not countable.

We emphasize that these notions—monomorphism, epimorphism, isomorphism, subobject—while of some importance in themselves, are important primarily as examples of the point of view of category theory. We shall shortly give two additional, somewhat richer, examples. We first need a little terminology. By a *diagram* we mean any collection of objects along with a collection of morphisms between various of those objects. (E.g., a diagram is what

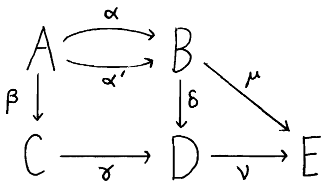

Figure 5

is pictured.) A diagram is said to *commute* if it has the following property: given any two objects in the diagram, and any two morphisms between those objects, obtained by composition of the morphisms in the diagram, those two morphisms are equal. Thus the statement that the diagram of figure 5 commutes is the statement that $\alpha = \alpha'$, $\gamma \circ \beta = \delta \circ \alpha$ (both morphisms from A to D), and $\nu \circ \delta = \mu$ (both morphisms from B to E). It turns out that many statements in category theory can be formulated as the statement that an appropriate diagram commutes. For example, one could (with no gain in clarity) define a monomorphism as follows: $A \xrightarrow{\varphi} B$ is a monomorphism if, whenever the first diagram of figure 6 commutes, so does the second.

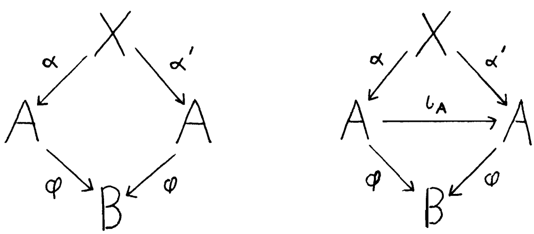

Figure 6

We now proceed to the final two examples of definitions in category theory. Let A and B be objects. A *product* of A and B is an object C, together with morphisms from C to A and from C to B, such that the following property holds: if C' is any object, and α' and β' any morphisms from C' to A and from C' to B, respectively, there is a unique morphism γ from C' to C such that the diagram of figure 7 commutes. A *direct sum* of A and

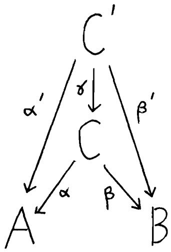

Figure 7

B is an object C, together with morphisms from A to C and from B to C, such that the following property holds: if C' is any object, and α' and β' any morphisms from A to C' and from B to C', respectively, there is a unique morphism γ from C to C' such that the diagram of figure 8 commutes.

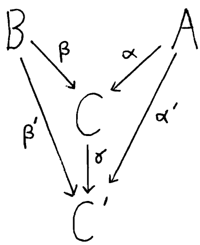

Figure 8

Note that a direct product consists not only of an object but also of a certain pair of morphisms, and similarly for a direct sum. We emphasize that there is no guarantee that, given two objects, either a direct product or a direct sum will exist: in fact, there are examples in which one or both do not exist. However, most of the categories one commonly deals with have the property that any two objects in the category do have both a direct product and a direct sum. Finally, we remark that, although we have here defined the product and sum of only two objects, the definition has an obvious extension to an arbitrary collection of objects (including infinite collections). The corresponding diagrams for the direct product and the direct sum are shown in figure 9. Nothing of consequence happens in the passage from the case of two objects to the case of an arbitrary collection of objects: proofs, for example, can be repeated almost word for word. We treat the case of only two objects because it makes arguments less confusing and because it is this case which is normally needed in practice.

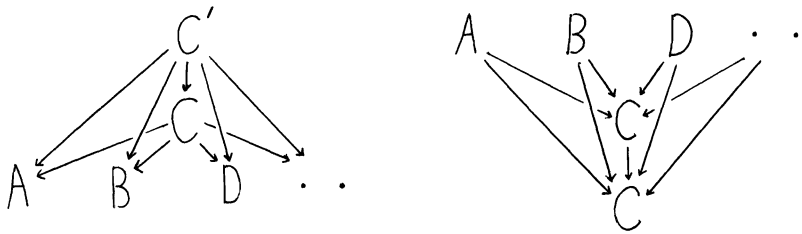

Figure 9

We now find out what these general definitions mean in the specific context of the category of sets. Recall a construction from set theory. The *Cartesian product,* $A \times B$, of sets A and B is the set of all pairs (a,b) with a an element of A and b an element of B. There is a natural mapping, α, from the Cartesian product $A \times B$ to A, given by $\alpha(a,b) = a$. (That is, α is the mapping which "ignores the second entry" of an element, (a,b), of the Cartesian product.) Similarly, $\beta(a,b) = b$ is a mapping from $A \times B$ to B.

THEOREM 4. *In the category of sets, $(A \times B, \alpha, \beta)$ is a direct product of sets A and B, where $A \times B$ is the Cartesian product, and α and β are the mappings above.*

Proof. Let C' be any set, and α' and β' any mappings from C' to A and from C' to B, respectively. We define a mapping γ from C' to $A \times B$ as follows: for c' in C', set $\gamma(c') = (\alpha'(c), \beta'(c))$. (Note that the right side is indeed an element of $A \times B$.) We must show that this γ makes the diagram of figure 10 commute (i.e., that $\alpha \circ \gamma = \alpha'$ and $\beta \circ \gamma = \beta'$), and that this γ is the only mapping from C' to $A \times B$ which makes it commute. To see that $\alpha \circ \gamma = \alpha'$, apply the mapping on the left to a typical element, c', of C': $\alpha(\gamma(c')) = \alpha(\alpha'(c'), \beta'(c')) = \alpha'(c')$. Hence $\alpha \circ \gamma = \alpha'$ and, similarly, $\beta \circ \gamma = \beta'$. Thus the diagram of figure 10 indeed commutes. All that remains is to show that this γ is the only mapping which does the job. Let $\bar{\gamma}$ be another mapping from C' to $A \times B$ which gives a commuting diagram. Fix an element c' of C', and consider the element $\bar{\gamma}(c')$ of $A \times B$. Since $\alpha \circ \bar{\gamma} = \alpha'$, we must have $\alpha(\bar{\gamma}(c')) = \alpha'(c')$; since $\beta \circ \bar{\gamma} = \beta'$, we must have $\beta(\bar{\gamma}(c')) = \beta'(c')$. But the only element x of $A \times B$ such that $\alpha(x) = \alpha'(c')$ and $\beta(x) = \beta'(c')$ is the element $(\alpha'(c'), \beta'(c'))$. Hence $\bar{\gamma}(c') = (\alpha'(c'), \beta'(c'))$. In other words, $\bar{\gamma} = \gamma$, establishing uniqueness of the mapping. □

To paraphrase this proof, "γ decides where (in $A \times B$) it will send an element c' of C' as follows: it sees where (in A) α' sends c' and where (in B) β' sends c' and puts these two into a pair to get an element of $A \times B$."

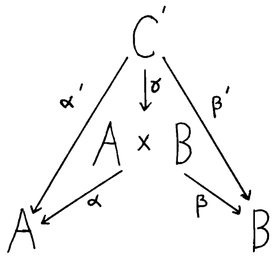

Figure 10

In order to obtain the analogous result for the direct sum, we must again recall a construction from set theory. The *disjoint union* of two sets, A_1 and A_2, is the set, written $A_1 \cup_d A_2$, consisting of all pairs (x,n), where n is either the number "1" or the number "2," and where x is an element of A_n. (It is convenient here to call the sets A_1 and A_2 rather than A and B.) This definition requires a little explanation. The elements of $A_1 \cup_d A_2$ of the form $(x,1)$ form a "copy of A_1," for the x in $(x,1)$ is allowed to range over A_1. Similarly, the elements of the form $(x,2)$ form a copy of A_2. But the elements of the form $(x,1)$ together with the elements of the form $(x,2)$ exhaust the elements of $A_1 \cup_d A_2$. Thus $A_1 \cup_d A_2$ is "the union of a copy of A_1 with a copy of A_2." Why this use of copies? Why not just take the union of A_1 and A_2? This would be fine if A_1 and A_2 were disjoint. If, however, A_1 and A_2 have some elements in common, then these elements would be "included only once" in $A_1 \cup A_2$, whereas they are "included twice" (once as an element of A_1 and once as an element of A_2) in $A_1 \cup_d A_2$. Denote by α_1 the mapping from A_1 to $A_1 \cup_d A_2$ which sends a typical element, a_1, of A_1 to the element $(a_1,1)$ of $A_1 \cup_d A_2$; similarly for α_2.

THEOREM 5. *In the category of sets, $(A_1 \cup_d A_2, \alpha_1, \alpha_2)$ is a direct sum of the sets A_1 and A_2, where $A_1 \cup_d A_2$ is the disjoint union, and α_1 and α_2 are the mappings above.*

Proof. Let C' be any set, and α_1' and α_2' any mappings from A_1 to C' and from A_2 to C', respectively. We define a mapping γ from $A_1 \cup_d A_2$ to C' as follows: $\gamma(x,1) = \alpha_1'(x)$, $\gamma(x,2) = \alpha_2'(x)$. (Note that, in the first equation, x must be an element of A_1 while, in the second, it must be an element of A_2.) We must show that this γ is the only mapping which makes the diagram of figure 11 commute (i.e., that this γ, and only this one, satisfies $\gamma \circ \alpha_1 = \alpha_1'$ and $\gamma \circ \alpha_2 = \alpha_2'$). To see that $\gamma \circ \alpha_1 = \alpha_1'$, apply the mapping on the left to a typical element, x, of A_1: $\gamma(\alpha_1(x)) = \gamma(x,1) = \alpha_1'(x)$. Hence $\gamma \circ \alpha_1 = \alpha_1'$, and, similarly, $\gamma \circ \alpha_2 = \alpha_2'$. Thus the diagram of

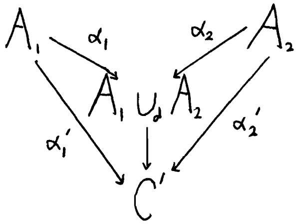

Figure 11

figure 11 indeed commutes. To show that this γ is the only mapping which gives a commuting diagram, let $\bar{\gamma}$ be another mapping. Then, since $\bar{\gamma} \circ \alpha_1 = \alpha_1{}'$, we have, applying this to an arbitrary element x of A_1, $\bar{\gamma}(\alpha_1(x)) = \alpha_1{}'(x)$. But $\alpha_1(x) = (x,1)$, so $\bar{\gamma}(x,1) = \alpha_1{}'(x)$, and, similarly, $\bar{\gamma}(x,2) = \alpha_2{}'(x)$. But these last two equations are precisely the statement that $\bar{\gamma} = \gamma$. □

Theorems 4 and 5 show, in particular, that direct products and direct sums always exist in the category of sets. (It is interesting to note that the definition of a direct sum, in particular, is perhaps more straightforward than that of the disjoint union.) One expects from these theorems that, even in other categories, direct products will be "product-like," and direct sums "union-like." This will turn out in some sense to be the case. However, for objects in more complicated categories, one cannot just go taking naive "products" and "unions" and expect to get, as a result, things having the structure appropriate for objects in that category. These categorical definitions will force us within each category to do the "right" thing and will ensure, in particular, that we always end up with an object in that category. Whereas there is nothing very subtle in monomorphisms and epimorphisms (they are always the obvious things), direct products and direct sums can, as we shall see, be very clever in combining two objects, and making all the structure work out right, to get new objects.

One of the things that makes direct products and direct sums interesting is that they are unique in a certain sense. This is not completely obvious even for sets. One could, offhand, imagine working very hard to find, for two sets A and B, a set C and mappings from C to A and from C to B such that the definition of a direct product is satisfied but such that this C is essentially different from $A \times B$. In fact, direct products and direct sums are, in an appropriate sense, unique, and, in fact, this is true in any category. The "appropriate sense" is the following:

THEOREM 6. *Let A and B be objects, and let (C,α,β) and (C',α',β') be two direct products of these objects. Then there is one and only one isomorphism from C' to C for which the diagram of figure 12 commutes. Similarly for the direct sum.*

(That is, not only are C and C' isomorphic as objects, but there is a unique isomorphism between them which preserves their "relationship with A and B.")

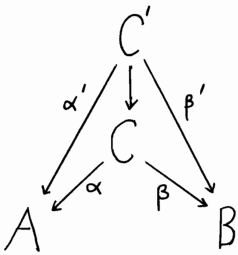

Figure 12

Proof. Since C is a product, there is a (unique) morphism γ such that the diagram of figure 12 commutes. Since C' is a product, there is a (unique) morphism γ' such that the diagram of figure 13 commutes. The first sentence implies $\alpha \circ \gamma = \alpha'$; the second implies $\alpha' \circ \gamma' = \alpha$. Hence

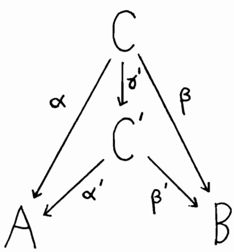

Figure 13

$\alpha \circ (\gamma \circ \gamma') = (\alpha \circ \gamma) \circ \gamma' = \alpha' \circ \gamma' = \alpha$. Similarly, $\beta \circ (\gamma \circ \gamma') = \beta$. Now consider the diagram of figure 14. (Note the two C's.) The result just derived is precisely the statement that this diagram commutes if μ is replaced by $\gamma \circ \gamma'$. Clearly, this diagram commutes if μ is replaced by i_C. But C is a direct product, so there can be only one μ which makes this diagram commute. Hence $\gamma \circ \gamma' = i_C$. Similarly, $\gamma' \circ \gamma = i_{C'}$. Thus γ is an isomorphism. Similarly for the direct sum. ☐

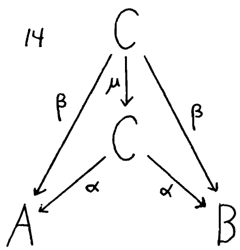

Figure 14

This completes our brief survey of category theory. We emphasize again that theorems 3 and 6 are about category theory while theorems 1, 2, 4, and 5 are about the category of sets. There are, of course, many more definitions and theorems in category theory. We shall introduce a few of these, when needed, later.

Exercise 1. Prove that every isomorphism is both a monomorphism and an epimorphism. (The converse is false, for example, in the category of Hausdorff topological spaces.)

Exercise 2. Let $A \overset{\varphi}{\to} B \overset{\psi}{\to} C$. Prove that, if $\psi \circ \varphi$ is a monomorphism, so is φ; that, if $\psi \circ \varphi$ is an epimorphism, so is ψ. Find examples, in the category of sets, to show that the converses are false.

Exercise 3. Let the objects consist of pairs, (A,A'), where A is a set and A' is a subset of A. Let a morphism from (A,A') to (B,B') consist of a mapping φ from set A to set B such that, whenever a' is in A', $\varphi(a')$ is in B'. Let composition of morphisms be composition of mappings. Prove that this is a category. Discuss monomorphisms, epimorphisms, direct products, and direct sums.

Exercise 4. Let the objects be sets with exactly 17 elements, the morphisms mappings of such sets, and composition composition. Verify that this is a category. Prove that in this category no two objects have either a direct product or a direct sum.

Exercise 5. Prove that the φ' in the definition of an isomorphism is unique.

Exercise 6. Fix a set A. For A' and A'' subsets of A, write $A' \leq A''$ if A' is a subset of A''. It is immediately evident that i) $A' \leq A'$, ii) $A' \leq A'' \leq A'''$ implies $A' \leq A'''$, and iii) $A' \leq A''$ and $A'' \leq A'$ imply $A' = A''$. Define a similar "$\leq$" for subobjects of a fixed object in an arbitrary category and prove that these three properties again hold.

Exercise 7. In the category of sets, the two morphisms (α and β) in a direct product are monomorphisms and the two morphisms in a direct sum are epimorphisms. Is this true in every category?

Exercise 8. In the category of sets, there is a "natural isomorphism," given three sets A, B, and C, from the set $(A \times B) \times C$ to the set $A \times (B \times C)$, where $\times$ is a Cartesian product. This observation suggests a theorem in category theory. State it and prove it. Similarly for disjoint union.

Exercise 9. For purposes of this exercise, call an object A in a category atomic if it has the following property: given any object B, there is one and only one morphism, φ_B, from B to A. Prove that any two atomic objects are isomorphic. Prove that any morphism from an atomic object to another object is a monomorphism. Prove that, given any objects A and B, with A atomic, (B, φ_B, i_B) is a direct product of A and B. What are the atomic objects in the category of sets? Do the "arrows reversed" version of all this.

Exercise 10. In the category of sets, $A \times (B \cup_d C)$ is naturally isomorphic with $(A \times B) \cup_d (A \times C)$. Is this a special case of some theorem in category theory?

Exercise 11. Fix two categories. Introduce a new category that can be thought of as the "product" of these. (Hint: choose, for the objects in this category, pairs consisting of one object from each of the given categories.)

3

The Category of Groups

A *group* consists of two things—i) a set G, and ii) a rule which assigns, given two elements (in a specific order) g and g' of G, an element (normally written gg' and called the product of g with g') of G—subject to the following three conditions:

1. The product is associative. For any three elements, g, g', and g'', of G,

$$g(g'g'') = (gg')g'' \ .$$

2. An identity exists. There is an element of G (called the *identity*, and normally written e) with the following property: for any element g of G,

$$eg = ge = g \ .$$

3. Inverses exist. Given any element g of G, there is an element of G (normally written g^{-1} and called the *inverse* of g) such that

$$gg^{-1} = g^{-1}g = e \ .$$

It is immediate that the identity is unique [proof: if e' were another, then $ee' = e'$, since e is an identity, while $ee' = e$, since e' is an identity, whence $e = e'$] and that inverses are unique [proof: if h and h' are both inverses of g, then $h' = h'e = h'(gh) = (h'g)h = eh = h$] and that the inverse of e is e [proof: $ee = ee = e$].

To give an example of a group, one must, of course, say what the set is and what the product rule is, and then one must verify that conditions 1, 2, and 3 above are satisfied.

Example. Denote by Z the collection of all integers (positive or negative, including zero). The rule "associate with any two integers their sum" assigns to any two integers another. This is a group—condition 1 is associativity of addition of integers; the integer zero will serve as an identity for condition 2; the integer $-n$ will serve as an inverse for the integer n—called the *additive group of integers.*

Example. The collection of all real numbers, R, with the rule, again, being addition is a group, called the *additive group of reals.*

Example. Let S be any set. Denote by G the collection of all one-to-one, onto mappings from S to S. Associate with any two elements, μ and ν, of G their composition, $\nu \circ \mu$, another element of G. This G, together with this product rule, is a group: condition 1 follows from associativity of composition of mappings; a suitable identity for condition 2 is the identity mapping from S to S; condition 3 follows from the fact that one-to-one, onto mappings have inverses. This group is called the ***permutation group*** on the set S.

Example. Let A be any object in any category, and denote by G the collection of all isomorphisms from A to A. Choose, for the product rule, composition of isomorphisms (noting that the composition of two isomorphisms is an isomorphism). The result is a group: conditions 1 and 2 for a group are immediate from conditions 1 and 2, respectively, for a category; condition 3 for a group is immediate from the definition of an isomorphism (in a category).

Most of the groups which occur naturally in physics seem to arise by an application of this last example in some special case. To describe a physical situation, one introduces some "space, with an appropriate structure." One makes these "spaces with this structure" into a category, so the particular physical situation is described by some object A in that category. Then the corresponding group, as in the last example, becomes the group of "structure-preserving transformations on A," that is, the "symmetry group." The last example above is, if you like, a reason why groups are important in physics.

A group G is said to be ***abelian*** if, for any two elements g and g' of G, we have

$$gg' = g'g \ .$$

(Thus the additive group of integers and the additive group of reals are abelian. The permutation group on a set S is abelian when and only when S has no more than two elements.) One almost always uses the following special notation for abelian groups: instead of gg', one writes $g + g'$ (and calls it the sum); instead of e, one writes 0; instead of g^{-1}, one writes $-g$; instead of gh^{-1}, one writes $g - h$.

Let G and H be groups. A ***homomorphism*** from G to H is a mapping φ from the set G to the set H such that, for any two elements, g and g', of G,

$$\varphi(gg') = \varphi(g)\varphi(g') \ .$$

(That is, "given two elements of G, one gets the same element of H whether one i) first takes the product of these elements in G and sends the result, by φ, to H or ii) first sends these elements to H, by φ, and there takes the product.") It is immediate that, for φ a homomorphism, $\varphi(e) = e$ (the e on the left is in G; the e on the right in H) [proof: $\varphi(e) = \varphi(e)\varphi(e)[\varphi(e)]^{-1} = \varphi(ee)[\varphi(e)]^{-1} = \varphi(e)[\varphi(e)]^{-1} = e$] and $\varphi(g^{-1}) = [\varphi(g)]^{-1}$ for any g in G [proof: $\varphi(g^{-1}) = \varphi(g^{-1})\varphi(g)[\varphi(g)]^{-1} = \varphi(g^{-1}g)[\varphi(g)]^{-1} = \varphi(e)[\varphi(g)]^{-1} = e[\varphi(g)]^{-1} = [\varphi(g)]^{-1}$]. (If either of these properties did not follow from the definition of a homomorphism, one would change the definition so that they would follow.) Thus a homomorphism from one group to another is a mapping which "preserves, in the strongest sense, all the structure available." Note also that the composition of two homomorphisms is another: if $G \xrightarrow{\varphi} H$ and $H \xrightarrow{\psi} K$ are homomorphisms of groups, then $G \xrightarrow{\psi \circ \varphi} K$ is also a homomorphism [proof: for g and g' elements of G, $\psi \circ \varphi(gg') = \psi[\varphi(gg')] = \psi[\varphi(g)\varphi(g')] = \psi[\varphi(g)]\psi[\varphi(g')] = \psi \circ \varphi(g)\psi \circ \varphi(g')$].

Let the objects be groups, the morphisms homomorphisms from groups to groups, and the composition composition of homomorphisms. We thus obtain a category—composition of morphisms (here, homomorphisms) is associative because composition of mappings is associative; the identity morphism from group G to itself is the identity mapping (obviously a homomorphism) from G to G. This is called the ***category of groups.*** Replacing "group" everywhere above by "abelian group," we obtain the ***category of abelian groups.*** Note the way that "additional structure" (in this case a product structure) is incorporated in the passage from sets to groups, in the passage from mappings to homomorphisms, and in the passage from the category of sets to the category of groups.

Categorical definitions now become applicable, in particular, to groups. For monomorphisms, the explicit meaning is the same for groups as for sets.

THEOREM 7. *In the category of groups, monomorphisms are one-to-one homomorphisms.*

Proof. Let $G \xrightarrow{\varphi} H$ be a one-to-one homomorphism, and consider $K \overset{\alpha}{\underset{\alpha'}{\rightrightarrows}} G \xrightarrow{\varphi} H$ with $\varphi \circ \alpha = \varphi \circ \alpha'$. If, for some k in K, we had $\alpha(k) \neq \alpha'(k)$, then, since φ is one-to-one, we would have $\varphi \circ \alpha(k) \neq \varphi \circ \alpha'(k)$, contradicting $\varphi \circ \alpha = \varphi \circ \alpha'$. Hence φ is a monomorphism. Suppose, conversely, that φ is a monomorphism. Let g and g' be elements of G with $\varphi(g) = \varphi(g')$. Then $\varphi(x) = e$, where $x = g^{-1}g'$. Let α be the homomorphism from Z, the additive group of integers, to G which sends every integer to e (in G). Let α' be the homomorphism which sends the positive integer n to $xx \cdots x$ (n times), the integer 0 to e, and the negative integer $-n$ to $(x^{-1}) \cdots (x^{-1})$ (n times). (It is easily checked that this α' is indeed a homomorphism.) Since $\varphi(x) = e$, we have $\varphi \circ \alpha = \varphi \circ \alpha'$ (for each side is the homomorphism which maps Z to the identity of H). But φ is a monomorphism, whence $\alpha = \alpha'$. Hence $e = \alpha(1) = \alpha'(1) = x = g^{-1}g'$. That is, $g = g'$. Hence φ is one-to-one. □

Note that the first half of this proof is identical to that for sets, while, for the second half, one replaces the "simplest set" (one with just one element) by the "simplest group" (additive group of integers). (It is also true that the epimorphisms in the category of groups are onto homomorphisms, but the proof is easier with a little of the technology of group theory and will be postponed.)

We also have, from category theory: a ***subgroup*** of a group G is a group H and monomorphism $H \to G$. This and other categorical notions will be discussed, for groups, later.

Both the statement and the proof of the following result are useful in group theory.

THEOREM 8. *Every group is a subgroup of the permutation group on some set.*

Proof. Let G be any group, and denote by Perm(G) the group of all one-to-one, onto mappings from the set G to itself. Fix an element g of G, and denote by φ_g the mapping from the set G to itself given by $\varphi_g(x) = gx$ (x in G). This mapping is one-to-one (for $\varphi_g(x) = \varphi_g(x')$ implies $gx = gx'$ implies $x = x'$) and onto (for $\varphi_g(g^{-1}x) = x$). That is, φ_g is an element of Perm(G). Thus we have obtained, for each element of G, an element of Perm(G); that is, we have a mapping φ from G to Perm(G). We claim that this φ is a homomorphism from the group G to the group Perm(G). Indeed, for fixed elements g and g' in G, and for x in G, we have $\varphi_g \circ \varphi_{g'}(x) = \varphi_g(\varphi_{g'}(x)) = \varphi_g(g'x) = gg'x = \varphi_{gg'}(x)$. That is, $\varphi_g \circ \varphi_{g'} = \varphi_{gg'}$, which is precisely the statement that φ is a homomorphism. Finally, we claim that this φ is in fact a monomorphism. By theorem 7 it suffices to show that φ is one-to-one, that is, that $\varphi_g = \varphi_{g'}$ implies $g = g'$. But, if $\varphi_g = \varphi_{g'}$, we have $g = \varphi_g(e) = \varphi_{g'}(e) = g'$. This completes the proof. □

The proof seems a bit complicated because there are so many mappings around. The idea is that each element of G defines, via "left multiplication by that element," a permutation on the set G itself, and this observation represents G as a subgroup of the permutation group of the set G. The statement of the theorem is of interest because it gives one a sense of control over groups. There is not much to sets; every set gives rise to a group (of permutations on that set); every group is a subgroup of one of these. It is not exactly an algorithm for writing down explicitly all groups, but one can feel more secure knowing that nothing happens in general groups much trickier than what happens already in the permutation groups. Finally, the proof itself is a useful construction: it is often convenient to regard a group G as a subgroup of the group of permutations on the set G.

Looking around for subgroups of permutation groups is one way to obtain groups. We conclude this section with another. Let S be any set. A *free group* on the set S is a group G together with a mapping α from the set S to the set G such that, given any other group G' and mapping α' from S to G', there is a unique homomorphism μ from group G to group G' such that the diagram of figure 15 commutes. This definition requires a remark. It is not (at least, as we have stated it) a categorical definition. The α, for example, maps S (a set) to G (which begins, at least, as a group): this α does not know what category it should be a morphism in. On the other hand, it certainly has the flavor of a categorical definition. Such definitions are (for an obvious reason) called universal definitions.

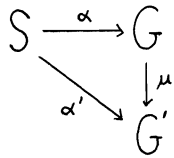

Figure 15

We show existence and uniqueness.

THEOREM 9. *Let (G,α) and (G',α') be free groups on the set S. Then there is a unique isomorphism (in the category of groups) from G to G' such that the diagram of figure 16 commutes.*

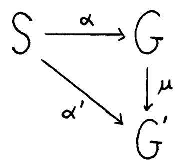

Figure 16

Proof. Since G is a free group on S, there is a (unique) μ such that the diagram of figure 16 commutes; since G' is free, there is a (unique) μ' such that the diagram of figure 17 commutes. Hence $(\mu' \circ \mu) \circ \alpha = \alpha$. Hence

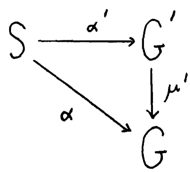

Figure 17

the diagram of figure 18 commutes both with ν replaced by $\mu' \circ \mu$ and with ν replaced by i_G. Since G is a free group, we have therefore $\mu' \circ \mu = i_G$. Similarly, $\mu \circ \mu' = i_{G'}$. □

The proof is essentially identical to that of theorem 6.

THEOREM 10. *For any set S, there exists a free group on S.*

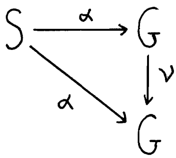

Figure 18

Proof. Denote by G the collection of all finite, ordered sequences having the following properties: i) each entry in the sequence is either an element of the set S or such an element with a prime attached, and ii) no two consecutive entries consist of the same element of S, once primed and once unprimed. (For example, a typical element of G is $ab'b'adca'cccd'$, where a, b, c, and d are elements of S.) We now define a product rule on this set G. Given two elements of G (i.e., sequences), their product is the sequence obtained by writing these sequences one after another and then removing from the result any pair of consecutive entries which violate ii) above. (For example, the product of abc' and $b'd'ca$ is $abc'b'd'ca$; the product of $ab'c'$ and cbd' is ad'.) We claim that this set with this product rule is a group. (The product is obviously associative. The identity is the sequence with no entries. The inverse of a sequence is that sequence obtained by reversing the order of entries, and interchanging "primed" and "unprimed.") Denote by α the following mapping from the set S to the set G: for each element a of S, $\alpha(a)$ is the element "a" of G, that is, the sequence having this one entry.

We now claim that this (G,α) is a free group on S. Let G' be any group, and α' any mapping from the set S to the set G'. Let μ be the mapping from G to G' which sends, for example, the sequence $ab'b'c'd$ to $\alpha'(a)[\alpha'(b)]^{-1}[\alpha'(b)]^{-1}[\alpha'(c)]^{-1}\alpha'(d)$ in G'. This mapping is clearly (because of our product rule for sequences) a homomorphism from the group G to the group G'. Furthermore, this μ makes the diagram commute (for, given any element a of S, $\alpha(a)$ is the sequence (element of G) with the single entry "a," while μ of this sequence is, by definition of μ, $\alpha'(a)$; that is, $\mu \circ \alpha = \alpha'$). We now claim, finally, that this μ is the only homomorphism from G to G' which makes the diagram commute. Let $\tilde{\mu}$ be another. Then $\tilde{\mu}$ and μ must certainly agree on sequences (elements of G) having but a single entry without a prime, for example, on the sequence "a," for, by commutativity, $\tilde{\mu}(a)$ must be $\alpha'(a)$. Since the sequence "a'" is the inverse (in G) of the sequence "a" and since homomorphisms take inverses to inverses, we must also have $\tilde{\mu}(a') = [\tilde{\mu}(a)]^{-1}$. That is, $\tilde{\mu}$ and μ must also agree on sequences with but a single primed entry. But homomorphisms take products to products (definition of homomorphism), while every element of G can be written as a product of sequences each having a single entry. Hence, since $\tilde{\mu}$ and μ agree on sequences having but a single entry, they agree on all sequences. That is, $\tilde{\mu} = \mu$. □

Intuitively, "since S is only a set, you cannot take inverses or products therein. So one enlarges S, introducing formal inverses (primed entries) and formal products (sequences), to obtain a group, the free group on S": a free group on S is what results if you "force S to become a group." Note that the definition of a free group on S is much simpler than the actual construction of such a group (in the proof of theorem 10). The uniqueness theorem (theorem 9) makes it clear that all one ever needs to use about the free group is the universal property of its definition. As we shall see, there are a number of constructions which begin with a free group. We remark, finally, that many other categories have similar "free objects."

Example. The free group on a set with just one element is isomorphic with the additive group of integers.

If, in the definition of a free group, one everywhere replaces "group" by "abelian group," the result is called a ***free abelian group*** on the set S.

Exercise 12. Show that, in any group G, $(gg')^{-1} = g'^{-1}g^{-1}$.

Exercise 13. Why is the set of real numbers, with the product rule multiplication of numbers, not a group? Show that the set of all positive reals, with product multiplication, is a group and is isomorphic to the additive group of reals.

Exercise 14. Prove that, for any set S, there exists a free abelian group on S. Prove that the free abelian group on S is isomorphic with the free group on S when and only when S has fewer than two elements.

Exercise 15. Fix a set S. Let the objects be elements of S, the morphisms from s to s' elements of Perm(S) which send s to s', and composition product in the group Perm(S). Is this a category?

Exercise 16. Fix a group G. Find a natural homomorphism from the free group on the set G to the group G.

Exercise 17. Let S and S' be sets, and G and G' corresponding free groups on these sets. Construct, for each mapping from the set S to the set S', a "corresponding" (in a sense to be defined) homomorphism from the group G to the group G'.

Exercise 18. For which groups G is the homomorphism from the group G to the permutation group on the set G (in the proof of theorem 8) an isomorphism?

Exercise 19. Is every group a subgroup of a free group?

Exercise 20. Prove that the only group which is both a permutation group and a free group is the group with but a single element.

Exercise 21. Can there exist a homomorphism from an abelian group to a group which is not abelian? from a nonabelian group to an abelian group? Same for monomorphism and isomorphism.

Exercise 22. Define a "free set." What is the theme which relates the proof of theorem 1 to the proof of theorem 7?

Exercise 23. Find the atomic objects (in the sense of exercise 9) in the category of groups.

Exercise 24. Find, for each positive integer n, a group having exactly n elements.

4

Subgroups

Recall that a subgroup of a group G is a monomorphism $H \xrightarrow{\tau} G$. But (theorem 7) monomorphisms, in the category of groups, are just one-to-one homomorphisms. Denote by $\tau[H]$ the subset of G consisting of elements of G which can be written in the form $\tau(h)$ with h in H. Then τ is a one-to-one, onto mapping from the set H to the set $\tau[H]$. Next, note that, if this subset $\tau[H]$ contains g $(= \tau(h))$ and g' $(= \tau(h'))$, then it also contains gg' $(= \tau(hh'))$ and g^{-1} $(= \tau(h^{-1}))$. (It follows that this subset also contains $e = gg^{-1}$.) Thus, if we impose on this subset $\tau[H]$ of G the product rule of G, this $\tau[H]$ itself becomes a group and τ becomes an isomorphism from the group H to the group $\tau[H]$. Clearly, we could just as well have defined a subgroup of G as a subset of G having the properties that the product (with respect to G) of any two elements of the subset is again in the subset, and the inverse of any element of the subset is again in the subset. This is what one normally calls a subgroup.

It is convenient to introduce the following notation for manipulating subsets of a group G. For A and B subsets (not necessarily subgroups) of a group G, denote by AB the subset of G consisting of all elements of G which can be written in the form ab, with a in A and b in B, and by A^{-1} the subset consisting of all elements of G which can be written in the form a^{-1} with a in A. (For example, $(AB)C = A(BC)$.) In this notation, a subset $\underline{G}$ of group G is a subgroup provided $\underline{G}\underline{G} = \underline{G}^{-1} = \underline{G}$.

Example. For G any group, the subset consisting of the identity element alone is a subgroup; the subset consisting of G itself is a subgroup. The additive group of integers is a subgroup of the additive group of reals.

Example. Let S' be a subset of the set S. The collection of all μ in Perm(S) (the permutation group on the set S) such that $\mu(s') = s'$ for each s' in S' is a subgroup of Perm(S). The collection of all μ in Perm(S) such that $\mu(s')$ is in S' for each s' in S' is another subgroup of Perm(S).

We introduce the following terminology. When we say "let X_λ (λ in Λ) be ...s," we mean that Λ is a set, and for each element λ of this set, X_λ is a ... Thus, if Λ were the set consisting only of "1" and "2," we would have two ...s, X_1 and X_2. The terminology is useful because often it does not make any difference whether one has two ...s, a finite number, a countably infinite number (e.g., X_1, X_2, . . .), or even more; one is able to give an argument without committing oneself as to how many ...s are involved.

One of the most useful properties of subgroups is this.

THEOREM 11. *Let G_λ (λ in Λ) be subgroups of group G. Then $\bigcap_\Lambda G_\lambda$ is a subgroup of G.*

Proof. Let g and g' be in $\bigcap_\Lambda G_\lambda$. Then g and g' are in each of the G_λ. But each G_λ is a subgroup, so gg' and g^{-1} are in each G_λ. Hence gg' and g^{-1} are both in $\bigcap_\Lambda G_\lambda$. That is, $\bigcap_\Lambda G_\lambda$ is a subgroup of G. □

As an example of theorem 11, consider the following. Let A be any subset of group G. The intersection of all subgroups of G which contain A itself contains A and is, by theorem 11, also a subgroup of G. This is called the subgroup (of G) *generated* by the subset A. It is clearly the smallest subgroup of G containing A, in the sense that any other subgroup of G containing A also contains the subgroup generated by A. It is also clear that the subgroup of G generated by A consists precisely of the elements of G which can be written as a product of elements of A and their inverses. Thus, for example, A is a subgroup of G if and only if the subgroup generated by A is just A itself. (Not much work is saved here by the use of theorem 11, but the savings increase quickly as the objects become more complicated.)

Example. Regard a set S as a subset of the free group on S. Then the subgroup of this free group generated by S is precisely the free group itself.

Example. Let F be a finite set, and consider the elements of the permutation group on F which interchange two elements of F, leaving the rest invariant. This subset of Perm(F) generates Perm(F).

The remarks above permit one to obtain subgroups of a group G: pick any subset of G, and find the subgroup it generates. There is also a way of getting subgroups from subgroups. Let ~~G~~ H be any subgroup of G, and let g be any element (fixed once and for all) of G. Consider the subset $K = gHg^{-1}$ of G (i.e., K is the collection of all elements of G which can be written in the form ghg^{-1} with h in H). We claim that this subset K of G is in fact a subgroup of G. [Proof: Let ghg^{-1} and $gh'g^{-1}$ be elements of K. Then $(ghg^{-1})(gh'g^{-1}) = g(hh')g^{-1}$ is in K, for, since H is a subgroup, hh' is in H. Similarly, $(ghg^{-1}) = gh^{-1}g^{-1}$ is in K.] Thus a subgroup H of G, together with any particular choice, g, of an element of G, gives rise to some subgroup K. This subgroup K is in general different from the subgroup H, although it can be the same. (For example, if G is abelian, then always $K = H$.) On the other hand, regarding H and K as groups in their own right, they are always isomorphic. Indeed, consider $H \underset{\tau}{\overset{\sigma}{\rightleftarrows}} K$, where σ is the mapping from H to K which sends the element h of H to the element ghg^{-1} of K, and τ is the mapping from K to H which sends the element k of K to the element $g^{-1}kg$ of H. It is easily checked that these mappings are both homomorphisms of groups, and that their compositions are the identity homomorphisms on H and K. Thus this construction yields, given one subgroup of a group, various other subgroups, each of which is a "copy" of the original subgroup.

Example. Let S be any set, and let $G = \text{Perm}(S)$, the permutation group on the set S. Let T be any subset of S, and let H be the subgroup of G consisting of permutations μ which leave T pointwise invariant (i.e., which are such that $\mu(t) = t$ for every t in T). Let g be any fixed element of G, that is, any permutation on the set S. Then gHg^{-1} is the subgroup of G consisting of those permutations which leave $g[T]$ pointwise fixed, where $g[T]$ denotes the collection of all elements of S of the form $g(t)$ for t in T.

An important tool in the study of the structure of subgroups is the notion of a coset. Let H be a subgroup of group G. A subset A of G is called a *left coset* (of H in G) if $A = gH$ for some g in G, and a *right coset* if $A = Hg$ for some g in G. (Note that the "g" such that, e.g., $A = gH$ may not be unique.) For example, H itself is both a left coset and a right coset of H in G, for $H = eH = He$. The basic properties of cosets are given in the following.

THEOREM 12. *Let H be a subgroup of group G. Then each element of G is contained in one and only one left coset of H in G. Furthermore, given any two left cosets of H in G, A and A', there is a one-to-one, onto mapping from the set A to the set A'.*

Proof. Each element g of G is certainly in some left coset, namely gH, for $g = ge$, and e is in H. This left coset is furthermore unique, for if g is in xH, (x in G), then $g = xh$ for some h in H, whence $gH = (xh)H = x(hH) = xH$, where, in the last step, we have used $hH = H$ (which follows from the fact that H is a subgroup). Finally, let $A = aH$ and $A' = a'H$ be left cosets. Then left multiplication by $a'a^{-1}$ certainly maps A to A', while left multiplication by $a(a')^{-1}$ maps A' to A. But, since $(a'a^{-1})(aa'^{-1}) = (aa'^{-1})(a'a^{-1}) = e$, the compositions of these mappings yield the identity on A and the identity on A'. We thus have one-to-one, onto mappings from A to A'. □

Thus the left cosets of H in G are all "copies" of each other, and they "cover G without overlapping." The situation is similar, of course, for right cosets (although a left coset and a right coset can overlap without coinciding). There are a large number of assorted facts about cosets, all of which are easier to prove when needed than to remember. For example: the element g of G is in the left coset A of H in G when and only when $A = gH$; the left cosets gH and $g'H$ coincide when and only when $g^{-1}g'$ is in H; the only coset of H in G which is also a subgroup is H $(= eH)$ itself.

Example. Let G be a group with n (a finite number) elements, and let H be a subgroup with p elements. Then every coset of H in G has p elements. If there are q left cosets of H in G, then (since these cosets cover G without overlapping) $n = pq$. In particular, p divides n. Thus the only subgroups of a group with, for example, 17 elements are the group itself and the subgroup containing only the identity (since the only integers which divide 17 are 1 and 17). In particular, if G is a group with 17 elements and $g \neq e$ is an element of G, then the subgroup of G generated by g is G itself.

The power of cosets is well illustrated by this example. The claims of the example seem subtle if one has not heard of cosets. Yet facts about cosets are easy to prove, and such facts about groups are easy to prove using cosets.

Example. Let S be a set, G the permutation group on S, T a subset of S, and H the subgroup of G consisting of elements of G which leave T pointwise invariant. Then, for any g in G, gH is the left coset consisting of permutations which, restricted to T, have the same action as g.

Let H be a subgroup of group G, and denote by L the collection of all left cosets of H in G. Then, since each element of G is contained in a unique left coset (theorem 12), we have a mapping from the set G to the set L. This mapping is always onto, and is one-to-one when and only when H consists only of the identity of G.

Fix an element g of G. Then, for A any left coset of H in G, gA is another. Thus this g defines a mapping from L to L. It is easily verified that this mapping is one-to-one and onto. Thus the mapping from L to L defined by g in G is an element of Perm(L), the permutation group on the set L. But this is true for each g in G; hence we have a mapping κ from the set G to the set Perm(L). Since $g(g'A) = (gg')A$ (A a left coset), this mapping is in fact a homomorphism of groups, $G \xrightarrow{\kappa} \text{Perm}(L)$.

Thus there are at least two modes of interaction between the group G and the set L of left cosets of H in G.

Exercise 25. Prove that a nonempty subset A of a group G is a subgroup if and only if, for any a and a' in A, $a^{-1}a'$ is in A. (In practice this is often the easiest criterion to test whether a subset of a group is a subgroup.)

Exercise 26. Let G be a group, and consider the collection of all subsets of G with product (for A and A' subsets) AA'. Do we thus make the subsets of G into a group?

Exercise 27. Find all subgroups of the additive group of integers.

Exercise 28. Let p be a prime number. Find all groups having exactly p elements.

Exercise 29. Is any union of subgroups of a group a subgroup?

Exercise 30. Let S be a finite set. Let A denote the elements μ of Perm(S) having the property that there are distinct elements s, s', and s'' of S with $\mu(s) = s'$, $\mu(s') = s''$, $\mu(s'') = s$, and with μ the identity on all other elements of S. Show that the subgroup of Perm(S) generated by A is not Perm(S) itself.

Exercise 31. Do there exist groups G and G' such that there is a monomorphism from G to G' and a monomorphism from G' to G, but such that there is no isomorphism from G to G'?

Exercise 32. Let H be a subgroup of group G. Let A and A' be subsets of G, each of which is the intersection of a left coset of H in G with a right coset of H in G. Does there exist a one-to-one, onto mapping from A to A'?

Exercise 33. A group is said to be finitely generated if it contains a finite subset that generates the entire group. Prove that, if $G \xrightarrow{\varphi} H$ is an onto homomorphism of groups, with G finitely generated, then H is finitely generated. Is every subgroup of a finitely generated group finitely generated?

5

Normal Subgroups

Subgroups of a group G are "group-like subsets": the definition refers only to the internal structure of the subset itself (and the product in that subset induced by G). There is also the notion of a subset which is "not only internally group-like, but also particularly well situated in the whole group G." The appropriate notion is that of a normal subgroup. By far the most important class of subgroups is the class of normal ones.

A subgroup N of group G is said to be a *normal subgroup* if

$$gNg^{-1} = N$$

for each element g of G (i.e., if, for any g in G and n in N, gng^{-1} is in N).

Thus the normal subgroups are those for which the technique (following theorem 11) for getting subgroups from subgroups yields nothing new. There is still another characterization of normal subgroups: a subgroup N of G is normal if and only if the left cosets of N in G are precisely the right cosets of N in G. [Proof: The left coset of N containing g is gN, the right coset Ng. If N is normal, then $gN = (gNg^{-1})g = Ng$. If, on the other hand, $gN = Ng$ for all g, then $gNg^{-1} = N$ for all g, so N is normal.]

It is often possible to guess with reasonable accuracy whether a given subgroup of a given group is normal. Normal subgroups usually "sit naturally in G, without making preferred choices that G itself does not make."

Example. The subgroups in the second example of chapter 4 are not normal unless $S' = S$ or S' is empty.

Example. Let G be the permutation group on set S. The subgroup of G consisting of all one-to-one, onto mappings from S to S which leave invariant all but a finite number of elements of S is normal.

Example. Every subgroup of every abelian group is normal.

Note that (by the same proof, essentially, as for theorem 11) any intersection of normal subgroups of group G is a normal subgroup of G. Thus, for A any subset of G, the intersection of all normal subgroups of G containing A is a normal subgroup, the normal subgroup *generated* by A. (Since G is always a normal subgroup of G, there is at least one normal subgroup in this intersection.) Even if A itself is a subgroup, the normal subgroup generated by A may be larger than A (and, of course, is larger if and only if A is not normal).

Example. Let G be any group, and let C be the subset consisting of elements of G which can be written in the form $gg'g^{-1}g'^{-1}$ with g and g' in G.

The subgroup of G generated by the subset C is called the *commutator subgroup* of G. The commutator subgroup is, in fact, a normal subgroup of G, so this commutator subgroup is also the normal subgroup generated by C.

The crucial fact which makes normal subgroups interesting is the following.

THEOREM 13. *Let N be a normal subgroup of group G, and denote by G/N the collection of cosets (left and right are the same) of N in G. Then, for A and A' elements of G/N, AA' is also a coset of N, and, furthermore, this product structure makes G/N a group.*

Proof. Let $A = aN$ and $A' = a'N$ be cosets of N in G. Then $AA' = aNa'N = a(a'Na'^{-1})a'N = aa'N(a'^{-1}a')N = aa'NeN = (aa')N$, so AA' is also a coset of N in G. This product is associative, since $(AA')A'' = A(A'A'')$. The identity is the element of G/N which is the coset N. The inverse of $A = aN$ is $a^{-1}N$. □

The group G/N is called the *quotient group* of G by (the normal subgroup) N. Thus, if N is the normal subgroup consisting of the identity alone, G/N is isomorphic with G; if N is the normal subgroup G of G, G/N is the group whose only element is the identity. In general, G/N is not a subgroup of G.

Intuitively, G/N is the group which results if "the elements of G which lie in N are forced to equal the identity element, with this forcing extended, in a consistent way, over G."

Example. For any group G, the quotient of G by the commutator subgroup of G is abelian. (The elements of C, the elements of the form $gg'g^{-1}g'^{-1}$, are the things whose "not being e is the reason the group is not abelian." When we "force these elements to be the identity," there results an abelian group.)

Three notions—that of a free group, that of a generated subgroup, and that of a quotient group—are often used together. One has some set S and a collection A of formal products of elements of S. One wishes to "make S into a group, but such that these formal products, in this group, reduce to the identity." One proceeds by first constructing the free group on S, then regarding the set A of formal products as a subset of this free group F, then taking the normal subgroup of F generated by A, and finally taking the quotient of F by this normal subgroup.

Example. Let G be any group, and let F be the free group on the set G. Denote by A the subset of F consisting of all formal products $g \cdots g'$ ($g, \ldots, g'$ in G) such that the real product in G (i.e., the element $g \cdots g'$ of G) is e. Let N be the normal subgroup of F generated by A. Then F/N is isomorphic with G. Thus every group is a quotient group of a free group. (Compare, theorem 8.)

Let N be a normal subgroup of group G. Let φ denote the mapping from the set G to the set G/N which sends g to gN. Then, since $gNg'N = gg'N$, this φ is a homomorphism. Thus, for example, we have the following sequence of homomorphisms: $N \xrightarrow{\mu} G \xrightarrow{\varphi} G/N \xrightarrow{\nu} \mathrm{Perm}(G/N)$, where the first is the monomorphism which represents N as a subgroup of G, and the last the homomorphism of theorem 8.

Exercise 34. Prove that a group is abelian if and only if its commutator subgroup consists only of the identity.

Exercise 35. Guess whether the subgroup of exercise 30 is normal, and verify.

Exercise 36. Regarding the additive group of integers as a subgroup of the additive group of reals, find the quotient group.

Exercise 37. Show that any subgroup of a group having just two left cosets is normal.

Exercise 38. Let G be a group, and N a normal subgroup such that G/N is abelian. Show that N contains the commutator subgroup of G.

Exercise 39. Let G be a group. The *center* of G is the collection Z of all elements z of G such that $zg = gz$ for every g in G. Show that Z is a normal subgroup of G.

Exercise 40. Let H and H' be subgroups of group G. Prove that i) HH' is not necessarily a subgroup of G, ii) if one of H or H' is a normal subgroup, then HH' is a subgroup of G, and iii) if both H and H' are normal, then HH' is a normal subgroup of G.

Exercise 41. Does there exist a universal definition of a normal subgroup?

Exercise 42. Find a nontrivial normal subgroup of the free group on a set with, say, three elements.

Exercise 43. Prove that, in the category of groups, $G \xrightarrow{\varphi} H$ is an epimorphism if and only if φ is onto. (Hints: The proof that φ onto implies φ an epimorphism is the same for groups as for sets. For the converse, consider $G \xrightarrow{\varphi} H \overset{\alpha}{\underset{\alpha'}{\rightrightarrows}} X$, where X is the permutation group on the set H. Let α be the natural homomorphism from H to $\mathrm{Perm}(H)$. Let α' be defined by $\alpha'(h) = x\alpha(h)x^{-1}$, where x is an element of X. One must now choose $x \neq e$ so that $\alpha \circ \varphi = \alpha' \circ \varphi$. Denote by Y the subgroup of X consisting of elements of the form $\alpha \circ \varphi(g)$ for g in G. Choose x so that it "rearranges the cosets of Y in X.")

6

Homomorphisms

Let $G \xrightarrow{\varphi} H$ be a homomorphism of groups. Since subgroups are in some sense easier to think about than homomorphisms, one would like to analyze the structure of this homomorphism in terms of certain subgroups.

The *kernel* of φ, $\mathrm{Ker}(\varphi)$, is the collection of all elements g of G such that $\varphi(g) = e$. The *image* of φ, $\mathrm{Im}(\varphi)$, is the collection of all elements h of H such that $\varphi(g) = h$ for some g in G.

The kernel is a measure of "how one-to-one the homomorphism is: the larger the kernel, the less the one-to-one–ness." (More precisely, $\varphi(g) = \varphi(g')$ if and only if $g^{-1}g'$ is in $\mathrm{Ker}(\varphi)$.) Similarly, "the larger $\mathrm{Im}(\varphi)$, the more nearly onto φ is." The basic properties of these subsets are the following.

THEOREM 14. *Let $G \xrightarrow{\varphi} H$ be a homomorphism of groups. Then* $\mathrm{Ker}(\varphi)$ *is a normal subgroup of G, and* $\mathrm{Im}(\varphi)$ *a subgroup of H.*

Proof. If g and g' are in $\mathrm{Ker}(\varphi)$, then $\varphi(gg') = \varphi(g)\varphi(g') = e$ and $\varphi(g^{-1}) = [\varphi(g)]^{-1} = e$. So $\mathrm{Ker}(\varphi)$ is a subgroup of G. Furthermore, for x in G and g in $\mathrm{Ker}(\varphi)$, $\varphi(xgx^{-1}) = \varphi(x)\varphi(g)[\varphi(x)]^{-1} = \varphi(x)e[\varphi(x)]^{-1} = e$. So $\mathrm{Ker}(\varphi)$ is a normal subgroup. For h $(= \varphi(g))$ and h' $(= \varphi(g'))$ in $\mathrm{Im}(\varphi)$, $hh' = \varphi(gg')$ and $h^{-1} = \varphi(g^{-1})$, whence $\mathrm{Im}(\varphi)$ is a subgroup of H. □

Since $\mathrm{Ker}(\varphi)$ is a normal subgroup of G, we can form the group $G/\mathrm{Ker}(\varphi)$. Let μ be the homomorphism from G to $G/\mathrm{Ker}(\varphi)$ (so μ assigns to each element of G the coset of $\mathrm{Ker}(\varphi)$ in G in which that element lies). This μ is onto (since every coset of $\mathrm{Ker}(\varphi)$ in G contains some element), and, furthermore, $\mathrm{Ker}(\mu) = \mathrm{Ker}(\varphi)$ (since the elements of $\mathrm{Ker}(\mu)$ are precisely the elements of the identity coset of $\mathrm{Ker}(\varphi)$ in G, i.e., the elements of the normal subgroup $\mathrm{Ker}(\varphi)$ of G). Next, consider two elements, g and g', of G that lie in the same coset of $\mathrm{Ker}(\varphi)$ in G; that is, let $g' = gk$ for k in $\mathrm{Ker}(\varphi)$. Then $\varphi(g') = \varphi(gk) = \varphi(g)\varphi(k) = \varphi(g)$. Thus φ takes an entire coset of $\mathrm{Ker}(\varphi)$ in G to a single element of H. We have therefore a mapping $\tilde{\varphi}$ from the set $G/\mathrm{Ker}(\varphi)$ to the set H, which is easily checked to be a homomorphism. This homomorphism is one-to-one (for, for x in G, $\tilde{\varphi}(x\,\mathrm{Ker}(\varphi)) = \varphi(x)$, whence $\tilde{\varphi}(x\,\mathrm{Ker}(\varphi)) = \tilde{\varphi}(y\,\mathrm{Ker}(\varphi))$ implies $\varphi(x) = \varphi(y)$ implies $x^{-1}y$ in $\mathrm{Ker}(\varphi)$ implies $x\,\mathrm{Ker}(\varphi) = y\,\mathrm{Ker}(\varphi)$), while $\mathrm{Im}(\tilde{\varphi}) = \mathrm{Im}(\varphi)$. Finally, let ν be the monomorphism from $\mathrm{Im}(\varphi)$ to H which represents $\mathrm{Im}(\varphi)$ as a subgroup of H.

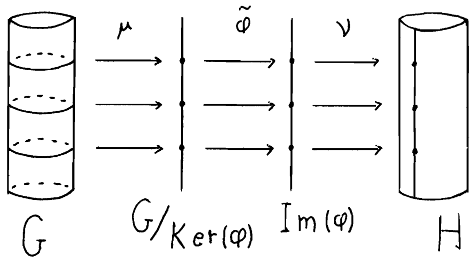

Figure 19

We thus have the sequence of homomorphisms of figure 19, where μ is onto, $\tilde{\varphi}$ is an isomorphism, and ν is one-to-one. Furthermore, $\varphi = \nu \circ \tilde{\varphi} \circ \mu$. Thus we have a decomposition of φ into simpler homomorphisms. The point of this decomposition is that the first two groups, G and $G/\mathrm{Ker}(\varphi)$, refer to G while the last two, $\mathrm{Im}(\varphi)$ and H, refer to H. The only link between G and H is via the isomorphism $\tilde{\varphi}$. In this sense, the structure of the homomorphism φ is carried by the structure of the subgroups $\mathrm{Ker}(\varphi)$ and $\mathrm{Im}(\varphi)$.

We conclude with a few assorted remarks about homomorphisms. Let G and H be abelian groups, and consider $\mathrm{Mor}(G,H)$, the set of (homo)morphisms from G to H. Given two such, α and β, we define a new mapping, $\mu = \alpha + \beta$, from G to H by $\mu(g) = \alpha(g) + \beta(g)$ (for g in G; the sum on the right is in the abelian group H). It is immediate that, since α and β are homomorphisms, so is μ. Hence we have a rule which assigns, given two elements of $\mathrm{Mor}(G,H)$, another element of $\mathrm{Mor}(G,H)$. We claim that $\mathrm{Mor}(G,H)$ thus becomes an abelian group. [Proof: Associativity follows from the fact that, for α, β, and γ in $\mathrm{Mor}(G,H)$, $((\alpha + \beta) + \gamma)(g) = \alpha(g) + \beta(g) + \gamma(g) = (\alpha + (\beta + \gamma))(g)$. The identity is the element of $\mathrm{Mor}(G,H)$ which sends all of G to e in H. For α in $\mathrm{Mor}(G,H)$, its inverse is the element of $\mathrm{Mor}(G,H)$ which sends g in G to $-\alpha(g)$. The group is abelian since $(\alpha + \beta)(g) = \alpha(g) + \beta(g) = (\beta + \alpha)(g)$.] Thus, in the category of abelian groups, for any two objects G and H $\mathrm{Mor}(G,H)$ also has the structure of an object in the category! It turns out that this important property holds in a large number of categories (but not, however, in the category of groups).

Let G be any group, and fix an element g of G. Then the mapping from the set G to G which sends x in G to gxg^{-1} is clearly an isomorphism from group G to G: call it φ_g. Then $\varphi_g \circ \varphi_{g'}$ sends x to $g(g'xg'^{-1})g^{-1} = (gg')x(gg')^{-1}$, whence $\varphi_g \circ \varphi_{g'} = \varphi_{gg'}$. In other words, we have a homomorphism from G to the group of isomorphisms from G to G. The elements of the image of this homomorphism are called *inner isomorphisms* on G. Thus, for example, if G is abelian, the only inner isomorphism on G is the identity isomorphism. (One can now easily understand the structure of the technique following theorem 11 for obtaining subgroups from subgroups. This is why normal subgroups are often called invariant subgroups.)

Exercise 44. Prove that the kernel of the homomorphism above from G to the group of isomorphisms on G is precisely the center of G.

Exercise 45. Find the kernel and image of the homomorphism of exercise 16.

Exercise 46. Prove that $G \xrightarrow{\varphi} H$ is one-to-one if and only if $\mathrm{Ker}(\varphi)$ consists only of e, and onto if and only if $\mathrm{Im}(\varphi) = H$.

Exercise 47. Show that, in the category of groups, $\mathrm{Mor}(G,H)$ does not in general have the structure of a group.

Exercise 48. Say what can be said about the kernel and image of the composition of two homomorphisms in terms of the kernel and image of the original homomorphisms.

Exercise 49. Let G and H be groups, N a normal subgroup of G and K a subgroup of H, such that G/N is isomorphic with K. Find $G \xrightarrow{\varphi} H$ with $\mathrm{Ker}(\varphi) = N$ and $\mathrm{Im}(\varphi) = K$.

Exercise 50. For G a set and H an abelian group, endow the set of mappings from G to H with the structure of an abelian group. Now let G also be an abelian group, and consider the subset of this collection of mappings consisting of the homomorphisms from G to H. Show that this is a subgroup. Is it normal?

Exercise 51. Let $G \xrightarrow{\varphi} H$ be a homomorphism, and fix h in H. Show that $G \xrightarrow{\varphi'} H$ given by $\varphi'(g) = h\varphi(g)h^{-1}$ (g in G) is also a homomorphism. Relate the kernel and image of φ' to those of φ.

7

Direct Products and Sums of Groups

We have the category of groups and the category of abelian groups. We know what it means in any category to say that an object is a direct product or a direct sum of others. We now ask what these general notions mean in our specific categories.

Let G and H be groups. We define a new group. The set is the Cartesian product of sets $G \times H$, that is, the set of all pairs (g,h) with g in G and h in H. We introduce, on this set, the rule $(g,h)(g',h') = (gg',hh')$, which assigns, to two elements of the set, another. That is, we multiply "component-wise." It is immediate that, since G and H are groups, this product structure makes the set $G \times H$ into a group. (The identity is (e,e); the inverse of (g,h) is (g^{-1},h^{-1}).) We shall also denote this group $G \times H$. Let α be the mapping from $G \times H$ to G with action $\alpha(g,h) = g$, and β the mapping from $G \times H$ to H with action $\beta(g,h) = h$. These mappings are homomorphisms of groups (e.g., $\alpha[(g,h)(g',h')] = \alpha(gg',hh') = gg' = \alpha(g,h)\alpha(g',h')$). We have

THEOREM 15. *Let G and H be groups (resp. abelian groups). Then $(G \times H,\alpha,\beta)$ is a direct product of G and H in the category of groups (resp. abelian groups).*

Proof. Let K be any group, and α' and β' any homomorphisms in the diagram of figure 20. Then, for k in K, $\mu(k)$ must, in order that the diagram

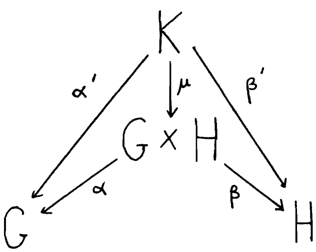

Figure 20

commute, be such that $\alpha(\mu(k)) = \alpha'(k)$ and $\beta(\mu(k)) = \beta'(k)$; hence, for commutativity, we must have $\mu(k) = (\alpha'(k),\beta'(k))$. But this μ is a homomorphism of groups, for $\mu(kk') = (\alpha'(kk'),\beta'(kk')) = (\alpha'(k)\alpha'(k'),\beta'(k)\beta'(k'))$

$= (\alpha'(k),\beta'(k))(\alpha'(k'),\beta'(k')) = \mu(k)\mu(k')$. The proof is the same in the abelian case, noting that, for G and H abelian, so is $G \times H$. □

This proof is nearly identical to that for sets. Thus direct products exist in our categories: one imposes the "obvious group structure on the Cartesian product."

The structure of the direct product can be seen in more detail using the diagram of figure 21. There, γ is the homomorphism from group G to group

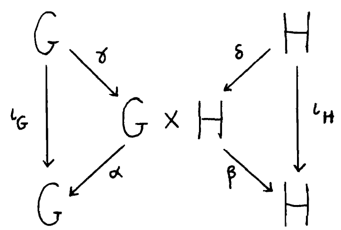

Figure 21

$G \times H$ with action $\gamma(g) = (g,e)$ (where e, of course, is the identity of H), and δ is the homomorphism from H to $G \times H$ with action $\delta(h) = (e,h)$. Note that this diagram is commutative (e.g., $\alpha \circ \gamma(g) = \alpha(\gamma(g)) = \alpha(g,e) = g = i_G(g)$). The various homomorphisms available in the presence of a direct product are easily kept in mind by remembering that there is a commutative "butterfly diagram." Note that γ and δ are monomorphisms (i.e., Ker(γ) consists only of the identity in G, and Ker(δ) only of the identity in H), that α and β are epimorphisms (i.e., Im(α) $= G$ and Im(β) $= H$), and that the remaining kernels and images are related by Im(δ) $=$ Ker(α) and Im(γ) $=$ Ker(β).

Direct sums also exist in both the category of groups and the category of abelian groups. This statement has somewhat more content than the corresponding statement for direct products. Recall that, in the category of sets, direct products are "product-like," and direct sums "union-like." To take a direct product of groups, one simply takes a direct product of sets and imposes on the resulting set an appropriate group structure. One might imagine therefore that direct sums of groups could be obtained in a similar way: first take the direct sum (disjoint union) of sets and look around for a suitable group structure on the resulting set. This fails (e.g., Should the identity element, in the disjoint union of G and H, be selected from G or from H? Where should the product of an element from G with an element from H lie?): there is no natural way to impose, on the disjoint union of sets G and H, a product so that this set becomes a group. Categorical definitions, however, are clever in such matters.

THEOREM 16. *Let G and H be abelian groups. Then $(G \times H, \gamma, \delta)$ is a direct sum of G and H in the category of abelian groups.*

Proof. Let K be any abelian group, and γ' and δ' any homomorphisms in the diagram of figure 22. Then, in order that the diagram commute, we

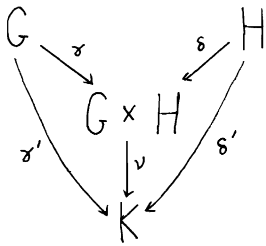

Figure 22

must have, for g in G, $\nu(g,0) = \nu(\gamma(g)) = \gamma'(g)$ and, for h in H, $\nu(0,h) = \delta'(h)$. Now, in order that ν be a homomorphism of groups, we must have $\nu(g,h) = \nu[(g,0) + (0,h)] = \nu(g,0) + \nu(0,h) = \gamma'(g) + \delta'(h)$. But this ν, so defined, is indeed a homomorphism, for $\nu[(g,h) + (g',h')] = \nu(g + g', h + h') = \gamma'(g + g') + \delta'(h + h') = \gamma'(g) + \gamma'(g') + \delta'(h) + \delta'(h') = \nu(g,h) + \nu(g',h')$. Thus one and only one homomorphism ν makes the diagram commute, so we have a direct sum. □

Thus, in the category of abelian groups, the direct product and direct sum are essentially the same thing (for a finite number of groups; they differ for an infinite number). In the category of groups, $G \times H$ is not a direct sum. The construction of the appropriate group is slightly more technical than of $G \times H$, seems never to arise in practice, and will be left to the exercises.

Exercise 52. Try to repeat the proof of theorem 16 in the category of groups, and find out what goes wrong.

Exercise 53. Display explicitly a direct sum in the category of groups.

Exercise 54. Find the direct product and direct sum of an infinite collection of abelian groups (in the category of abelian groups).

Exercise 55. Regard (G,γ) as a subgroup of $G \times H$. Show that this is a normal subgroup and that the quotient group is isomorphic to H. (That is, $(G \times H)/G = H$.)

Exercise 56. Let G, H, and K be groups, and γ and δ monomorphisms, and α and β epimorphisms such that the diagram of figure 23 commutes, and let

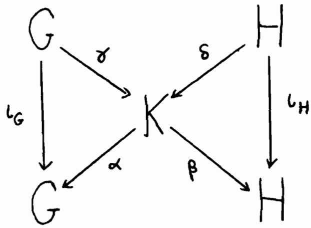

Figure 23

Ker(α) = Im(δ) and Ker(β) = Im(γ). State and prove: there is a natural isomorphism from K to $G \times H$.

Exercise 57. Let N be a normal subgroup of group G. Is there an isomorphism from G to $N \times (G/N)$?

Relations

It is convenient at this point to introduce a few facts about relations. This subject has very much the flavor of a "service subject"; that is, it is useful in certain arguments in mathematics, but rarely directly in applications.

Let S be a set. A ***relation*** on S is a subset of $S \times S$ (Cartesian product). Instead of "(s,s') (where s and s' are elements of S) is in this subset," one writes sRs' and says "s is in the relation R to s'." Just plain relations are not very interesting; but relations satisfying certain additional conditions are.

Let S be a set. A relation (normally written, in this case, "$\approx$" instead of "R") on S is said to be an ***equivalence relation*** if the following three conditions are satisfied:

1. For any s, s', and s'' in S with $s \approx s'$ and $s' \approx s''$, we have $s \approx s''$ (transitive).
2. For any s in S, $s \approx s$ (reflexive).
3. For any s and s' in S with $s \approx s'$, we have $s' \approx s$ (symmetric).

These are properties one would intuitively associate with the word "equivalent."

Example. Let G be a group, and H a subgroup of G (here regarded as a subset). For g and g' in G, write $g \approx g'$ if $g^{-1}g'$ is in H. This is an equivalence relation on the set G. (For $g^{-1}g'$ in H and $g'^{-1}g''$ in H, $g^{-1}g''$ $(= (g^{-1}g')(g'^{-1}g''))$ is in H, since H is a subgroup; for any g in G, $g^{-1}g$ $(= e)$ is in H; for $g^{-1}g'$ in H, $g'^{-1}g$ $(= (g^{-1}g')^{-1})$ is in H.)

Let S be a set, and consider any collection of equivalence relations on S. Their intersection (i.e., regard each equivalence relation as a subset of $S \times S$, and intersect these subsets to obtain a new subset of $S \times S$, i.e., to obtain a new relation) is, as is easily checked, also an equivalence relation. Now let R be any (not necessarily equivalence) relation on S. The intersection of all equivalence relations containing R (in the sense that, regarded as subsets of $S \times S$, they contain the subset of $S \times S$ defined by R) is thus an equivalence relation containing R and is clearly the smallest one (i.e., any equivalence relation containing R is contained in this intersection). This is called the equivalence relation ***generated*** by R. Many of the equivalence relations one introduces are obtained in this way.

Let "$\approx$" be an equivalence relation on the set S. An ***equivalence class*** (of this relation) is a subset T of S with the following property: for each t in T, $t \approx s$ if and only if s is also in T. That is, for an equivalence class T, "any two elements of T are equivalent and no element of T is equivalent to an element of S not in T."

There is essentially only one theorem in the subject of equivalence relations. The introduction of an equivalence relation is invariably followed immediately by application of this theorem.

THEOREM 17. *Let "$\approx$" be an equivalence relation on set S. Then every element of S is contained in one and only one equivalence class.*

Proof. Let s be in S, and denote by T the collection of all elements t of S such that $s \approx t$. Then T is an equivalence class (for, if t and t' are in T, so $s \approx t$ and $s \approx t'$, then by conditions 1 and 3 above, $t \approx t'$, while, for t in T and s' not in T, we could not have $t \approx s'$, for, since $s \approx t$, we would then have $s \approx s'$). Furthermore, since $s \approx s$, s itself is in T. Finally, if U is another equivalence class containing s, then u is in U if and only if $s \approx u$ (since U is an equivalence class), whence u is in U if and only if u is in T, whence $U = T$. □

Example. The equivalence classes for the equivalence relation given earlier in this chapter are the left cosets of H in G.

Theorem 17 means that "the equivalence classes cover S without overlapping."

There is a second special kind of relation which is of interest. A relation (normally written, in this case, "$\leq$") on a set S is called a *partial ordering* if the following three conditions are satisfied:

1. For any s, s', and s'' in S with $s \leq s'$ and $s' \leq s''$, we have $s \leq s''$.
2. For any s in S, $s \leq s$.
3. For any s and s' in S with $s \leq s'$ and $s' \leq s$, we have $s = s'$.

These are the properties one associates with "less than or equal to."

Example. Let P be any set, and consider any collection of subsets of P. For two such subsets, U and V, write $U \leq V$ if U is a subset of V. This is a partial ordering. One says that this collection of subsets is *ordered by inclusion.* (Many useful partial orderings arise in this way.)

Example. Let S be the set consisting of all pairs, (t,x), of real numbers, each in the open interval (0,1). Write $(t,x) \leq (t',x')$ if $t \leq t'$ and $(x - x')^2 \leq (t - t')^2$ (figure 24). (Geometrically, $(t,x) \leq (t',x')$ if (t',x') "lies on or above" the lines with slope 45° through (t,x) in the t-x plane.) This is a partial ordering on S.

It is convenient, unfortunately, to introduce three definitions. Let "$\leq$" be a partial ordering on the set S, and let A be a subset of S. This A is said to be ***bounded above*** if there is an element s of S such that $a \leq s$ for every a in A. This A is said to be ***totally ordered*** if, for any a and a' in A, either $a \leq a'$ or $a' \leq a$ (or, of course, both, if $a = a'$). To get an intuitive feeling for these definitions, consider the four diagrams of figure 25 (referring to the

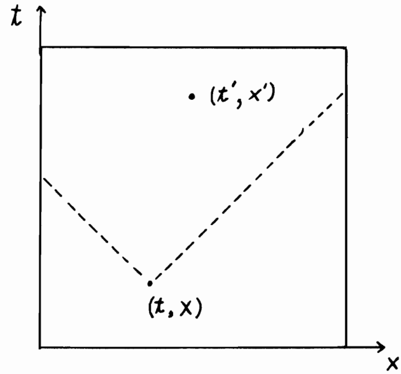

Figure 24

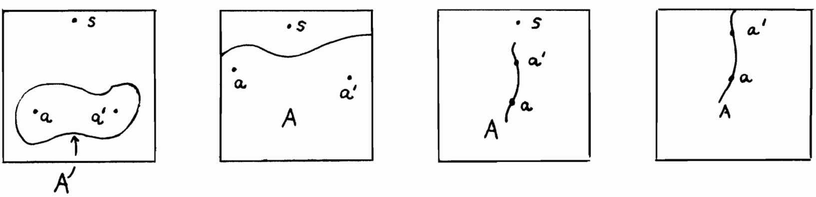

Figure 25

previous example). In the first, A is bounded above but not totally ordered (the s does the job for bounded above; neither $a \leq a'$ nor $a' \leq a$ for totally ordered); in the second, A is neither bounded above nor totally ordered (the s in the figure does not work for bounded above, for it is not true that $a \leq s$); in the third, A is both totally ordered and bounded above (use s for bounded above; since the slope of the curve is never less than 45°, we have, e.g., $a \leq a'$); in the fourth, A is not bounded above but is totally ordered. Finally, an element s of S is said to be a *maximal element* if there is no s' in S different from s itself such that $s \leq s'$. In the previous example, there are no maximal elements (for one can always "move up a little in the open square to obtain s'"). If, however, we had chosen for S the collection of pairs (t,x) with $0 < x < 1$ and $0 < t \leq 1$ (i.e., if we had "included the top edge of the square"), then any (t,x) with $t = 1$ would be a maximal element.

Now consider the following.

Statement. Let S be a set with a partial ordering, and suppose that S has the property that every totally ordered subset of S is bounded above. Then S possesses a maximal element.

It is not difficult to convince oneself that this statement seems "true." Pick s_1 in S. If this s_1 is not a maximal element, there is an s_2, different from s_1, with $s_1 \leq s_2$. If s_2 is not maximal, there is an s_3, different from s_2, with $s_2 \leq s_3$, etc. Either we eventually obtain a maximal element (in which case we are done), or we obtain $s_1 \leq s_2 \leq s_3 \leq \cdots$. In this latter case, $\{s_1, s_2, \cdots\}$ is a totally ordered subset of S, whence it is bounded above, say by s_1'. If s_1' is not a maximal element, there is s_2' with $s_1' \leq s_2'$, etc. (We "go right through infinity.") If we find no maximal element in s_1', s_2', . . . , we obtain, again, a totally ordered set, which must be bounded above, say by s_1'', etc. Suppose we are unlucky and go through an infinite sequence of such sequences without finding a maximal element. Then we obtain totally ordered s_m^n $(m,n = 1,2, \ldots)$, which must be bounded above, say by t_1, etc. We "continue in this way. At every stage we have a totally ordered subset, which must be bounded above, so we add in this upper bound to obtain a new totally ordered subset. We can go through infinity, an infinity of infinities, etc. We never get stuck; we can always go on. We should obtain eventually a maximal element."

The discussion above is not a proof—it simply gives the flavor of what the statement above asserts. Is there a real proof? The answer is remarkable: there is neither a proof (unless one assumes essentially the statement) nor a counterexample. The point is that sets themselves must be regarded, not as having been somehow "handed down from above," but rather as just symbols which may be related in a certain way (e.g., by "is an element of"), with certain properties to be satisfied by this "is an element of." (A set theory is a mathematical structure rather like the notion of a group.) Suppose we introduce the relation "is an element of" and postulate various properties which reflect things we expect to be satisfied for sets. What happens is that the statement above does not follow from these properties (in the sense that the denial of the statement, together with the properties, forms a consistent system), while neither does the denial of the statement follow from these properties. (It is like writing in group theory "statement: $gg' = g'g''$. Either the statement or its denial is consistent with the conditions for a group.)

What one normally does is include our statement, called *Zorn's lemma,* as an additional axiom on one's set theory. (The result, it has been shown, is internally consistent.) One normally uses Zorn's lemma when one wants to do what was described intuitively just above ("keep going, on and on"). Although Zorn's lemma is useful in mathematics, I know of no example in which one's stance regarding Zorn's lemma has a direct impact on one's mode of description of physical phenomena. Such an example might be very interesting.

Exercise 58. Introduce the category of partially ordered sets, and discuss therein direct product and direct sums (both of which exist).

Exercise 59. Let R be a relation on the set S. Prove that, in the equivalence relation generated by R, $s \approx s'$ if and only if there is a finite sequence $(s, s_1, s_2, \ldots, s_n, s')$ such that either sRs_1 or s_1Rs, either s_1Rs_2 or $s_2Rs_1, \ldots,$ and either s_nRs' or $s'Rs_n$.

Exercise 60. Fix a category. For any two objects, A and B, therein, write $A \approx B$ if there exists an isomorphism from A to B. Prove that "$\approx$" is an equivalence relation. What are the equivalence classes?

Exercise 61. Is it true that, for any set with a partial ordering having a maximal element, every totally ordered subset is bounded above?

Exercise 62. Let S be a set with partial ordering "$\leq$." Denote by T the collection of all totally ordered subsets of S, ordered by inclusion, so T is also a partially ordered set. Show that the hypothesis of Zorn's lemma is satisfied by T. What are the maximal elements of T like?

9

The Category of Vector Spaces

The next structure we shall discuss is that of vector spaces.

A *real vector space* consists of three things—i) a set V (whose elements are called *vectors*), ii) a rule which assigns, given vectors v and v', a third vector (written $v + v'$ and called their *sum*), and iii) a rule which assigns, given a vector v and a real number a, a vector (written av)—subject to the following conditions:

1. The set V is an abelian group under addition. (The identity is written "0" (not to be confused with the number "0"); the inverse of v, $-v$.)

2. For a and a' real numbers, v and v' vectors,

$$(a + a')v = av + a'v$$

(the "+" on the left is addition of numbers, that on the right addition of vectors) and

$$a(v + v') = av + av'$$

(both "+" signs are addition of vectors).

3. For a and a' real numbers, and v a vector,

$$a(a'v) = (aa')v$$

(aa' is ordinary multiplication of numbers).

4. For any vector v,

$$1v = v \; ,$$

where 1 is that number.

Replacing "real" above everywhere by "complex," we obtain the definition of a *complex vector space*. (There are more general vector spaces, in which "real numbers" and "complex numbers" are replaced by other things, but they do not seem very useful for applications.) Most arguments are identical for real vector spaces and complex vector spaces. We shall use the generic word "number." It can be replaced, consistently throughout any argument, by either "real number" or "complex number."

Thus a vector space is a set on which one can "add" and "scale by numbers," satisfying all the properties one would expect of these operations. A number of elementary facts follow immediately from the definition, for example, $0v = 0$ (the "0" on the left is the number, the "0" on the right the vector; proof: $0v = 0v + 0v - 0v = (0 + 0)v - 0v = 0v - 0v = 0$), $a0 = 0$ (in both instances "0" is a vector; same proof), and $(-a)v = -(av)$ (the "$-$" on the left denotes the negative of the number, that on the right the inverse of the vector; proof: $(-a)v = (-a)v + (av) - (av) = (-a + a)v - (av) = 0v - (av) = -(av)$).

Example. Let S be any set, and denote by V the collection of all mappings from S to the set of numbers, so V is the collection of all "number-valued functions" on S. Given two such functions, v and v', we define their sum, $v + v'$, as the function with action $(v + v')(s) = v(s) + v'(s)$ (s in S); given a function v and a number a, we define a new function, av, by $(av)(s) = a[v(s)]$. This set V, with these operations, is clearly a vector space; that is, it satisfies conditions 1, 2, 3, and 4 above.

Let V and W be vector spaces (both real or both complex). A mapping φ from the set V to the set W is called a *linear mapping* if, for any v and v' in V, and any number a, we have $\varphi(v + av') = \varphi(v) + a\varphi(v')$. It is immediate that, for φ a linear mapping, $\varphi(v + v') = \varphi(v) + \varphi(v')$ (so φ is also a homomorphism from the abelian group V to the abelian group W) and $\varphi(av) = a\varphi(v)$. [Proofs: $\varphi(v + v') = \varphi(v + 1v') = \varphi(v) + 1\varphi(v') = \varphi(v) + \varphi(v')$; $\varphi(av) = \varphi(0 + av) = \varphi(0) + a\varphi(v) = a\varphi(v)$.] Linear mappings are "structure preserving." Note that we do not define the notion of a linear mapping from a real vector space to a complex vector space, or vice versa. Note, finally, that the composition of two linear mappings (defined by composing the mappings of sets) is also a linear mapping.

Let the objects be real vector spaces, the morphisms linear mappings from real vector spaces to real vector spaces, and composition composition. We obtain the *category of real vector spaces* and, similarly, the *category of complex vector spaces.* The monomorphisms and epimorphisms in these categories are one-to-one linear mappings and onto linear mappings, respectively. (See exercises.)

Example. Let S and S' be sets, and let V and V' be the corresponding vector spaces in the previous example. Let $S \xrightarrow{\psi} S'$ be any mapping from set S to set S'. We define a mapping from the set V' to the set V. (Note this reversal of order.) For v' in V', let $\varphi(v')$ be the element of V (i.e., the function on S) whose value at s in S is $v'(\varphi(s))$, a number. (That is, $\varphi(v')$ assigns, to s in S, the number that v' assigns to the element $\psi(s)$ of S'.) It is easily checked that this mapping from set V' to set V is a linear mapping from vector space V' to vector space V.

A good way to get hold of some objects is to look at the free ones. Let S be any set. A *free vector space* on S is a vector space V, together with a mapping from the set S to the set V, such that, if W is any other vector space, and β any mapping from set S to set W, there is a unique linear mapping from vector space V to vector space W such that the diagram of figure 26 commutes.

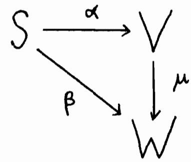

Figure 26

THEOREM 18. *For any set S, there exists a free vector space on S.*

Proof. Let S be any set, and denote by V the collection of all mappings v from the set S to the set of numbers such that $v(s) = 0$ for all but a finite collection of s in S. Setting $(v + v')(s) = v(s) + v'(s)$ and $(av)(s) = a[v(s)]$ (for v and v' in V, a a number, and s in S), we obtain (noting that $v + v'$ and av are always in V) a vector space V. Let α be the following mapping from set S to set V: for s in S, $\alpha(s)$ is the element of V (i.e., the function on S) given by $[\alpha(s)](s') = 1$ if $s' = s$, and 0 otherwise. (Note that this statement defines, for each fixed s in S, a function on S by giving the value of this function for each s' in S.) We claim that this (V,α) is a free vector space on S. Let W be a vector space, and β a mapping from set S to set W. We must find, and show the uniqueness of, a linear mapping from V to W such that figure 26 commutes. First, note that, in order that the diagram commute, μ must have the following action on the elements of V of the form $\alpha(s)$: $\mu[\alpha(s)] = \beta(s)$. Next, note that, since elements of V consist of functions on S which vanish on all but a finite subset of S, every element v of V can be written in the form $v = a_1\alpha(s_1) + \cdots + a_n\alpha(s_n)$. Then, note that, in order that μ be a linear mapping, we must have, for this v, $\mu(v) = a_1\mu[\alpha(s_1)] + \cdots + a_n\mu[\alpha(s_n)] = a_1\beta(s_1) + \cdots + a_n\beta(s_n)$. Finally, note that this μ, so defined, is indeed a linear mapping from V to W. □

Intuitively, V is "the vector space of finite formal sums, each term of which is a formal product of a number and an element of S. Since we cannot add or multiply by numbers in S (it is only a set), we appropriately enlarge S to allow this to be done, obtaining the free vector space on S." Note that the construction in this proof differs slightly (but significantly) from that of the first example of this chapter. Exactly as with groups, the free vector space on any set is unique in the appropriate sense.

Free vector spaces, as we shall see, play an important role in this subject (more important than, say, free groups do in the subject of groups). It is worthwhile therefore to study the structure of free vector spaces in a bit more detail. We first need some definitions. Let V be any vector space, and K any subset of V. An element v of V is said to be a ***linear combination*** of elements of K if v can be written in the form $a_1k_1 + \cdots + a_nk_n$, with $a_1, \ldots, a_n$

numbers and $k_1, \ldots, k_n$ elements of K. The subset K of V is said to *span* V if every element of V is a linear combination of elements of K. Thus, for example, V spans V; if K spans V and K is a subset of K' (a subset of V), then K' spans V. A subset K of V is said to be *linearly independent* if, whenever $k_1, \ldots, k_n$ are distinct elements of K and $a_1, \ldots, a_n$ are numbers, with $a_1k_1 + \cdots + a_nk_n = 0$, we necessarily have $a_1 = a_2 = \ldots = a_n = 0$. Thus, for example, K is never linearly independent if 0 is in K (for $1 \cdot 0 = 0$); if K is linearly independent and K' is a subset of K, then K' is linearly independent. Finally, a subset K of V is said to be a *basis* for V if K spans V and also K is linearly independent. (Note that, in these definitions, one only considers "finite sums." That is because, in a just plain vector space, one does not know what an "infinite sum" means. This notion requires some further structure, e.g., a topology.)

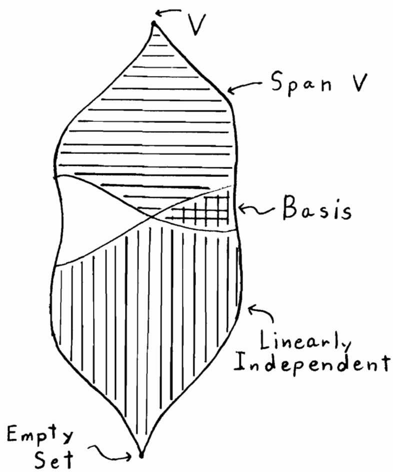

Figure 27

We illustrate these definitions schematically in figure 27. Each point of the figure represents a subset of vector space V. We order by inclusion, so the collection of all subsets of V is partially ordered. "Increase in size" of the subset is represented by "moving upward" in the figure. (Thus V itself, the largest subset of V, is at the top, while the empty subset—the smallest—is at the bottom.) The hatched regions represent the subsets which are linearly independent and the subsets which span V. The intersection represents the subsets that are a basis for V.

There are a number of assorted facts connecting these definitions, of which we give just two examples. If K is a linearly independent subset of V, and v is an element of V which is not a linear combination of elements of K, then K', the subset of V consisting of K with v included, is also linearly independent. [Proof: Suppose K' were not linearly independent, so $a_1k_1' + \cdots + a_nk_n' = 0$ with $k_1', \ldots, k_n'$ distinct elements of K', and $a_1, \ldots, a_n$ numbers, not all zero. Then v had better be included in this linear combination—and with nonzero numerical factor—for otherwise we would violate linear independence of K. Hence the linear combination above must really be of the form $bv + b_1k_1 + \cdots + b_{n-1}k_{n-1} = 0$, with $b \neq 0$, and $k_1, \ldots, k_{n-1}$ in K. But now we have $v = (-b_1/b)k_1 + \cdots + (-b_{n-1}/b)k_{n-1}$. This violates the assumption that v is not a linear combination of elements of K.] If K spans V, and k in K is a linear combination of elements of K', the set consisting of K with k excluded, then K' also spans V. [Proof: Since K spans V, and v in V can be written as a linear combination, $v = a_1k_1 + \cdots + a_nk_n$, of elements of K. If k does not appear in this linear combination, then here is an expression for v as a linear combination of elements of K', and we are done. If k does appear, on the other hand, write $k = a_1k_1' + \cdots + b_mk_m'$ with $k_1', \ldots, k_m'$ in K'. Substituting this expression for k in the above expression for v, we eliminate k in that linear combination and obtain an expression for v as a linear combination of elements of K'. Hence K' spans V.]

We now return to the free vector space constructed in the proof of theorem 18. Recall that, there, V was the vector space of all number-valued functions on the set S which vanish on all but a finite number of points of S. Also, α was the mapping from S to V which associates, with s in S, the function whose value is 1 on s and 0 elsewhere. Denote by K the subset of V consisting of all elements of V of the form $\alpha(s)$, s in S, that is, the subset of V consisting of all functions whose value is 1 at one point of S and 0 elsewhere. Since the only linear combination of such functions which is the zero function is a linear combination with all numerical coefficients zero, K is linearly independent in V. Furthermore, since V consists of functions on S which vanish on all but a finite number of elements of S, every element of V is a finite linear combination of elements of K. That is, K spans V. What we have just shown is that K is a basis for V. It is equally clear, conversely, that, if V is any vector space and K is a basis for V, then V is isomorphic with the free vector space on the set K. Thus we have the following conclusion: a vector space V is isomorphic to a free vector space if and only if V possesses a basis.

We now give the main structure theorem in vector spaces: every vector space has a basis. The idea of the proof is the following. Begin with any old linearly independent subset of vector space V (e.g., the subset with just one element, that nonzero). If this subset is not a basis (i.e., if it does not span V), there is a v in V which cannot be written as a linear combination of elements of the subset. We can therefore include this v in the subset and still have a linearly independent subset. If it is still not a basis, we can add another element, etc. We "go on and on, adding more and more elements, until we finally do get a basis." That is, we use Zorn's lemma.

THEOREM 19. *Every vector space V is isomorphic to a free vector space on some set.*

Proof. By the remarks above, it suffices to prove that V possesses a basis. Denote by **L** the collection of all linearly independent subsets of V, ordered by inclusion. Let L_λ (λ in Λ) be any totally ordered subset of **L**. (That is, each L_λ is a linearly independent subset of V, and any two L_λ's have the property that one is a subset of the other.) Then $\bigcup_\Lambda L_\lambda$ is also linearly independent (for any finite collection of elements of $\bigcup_\Lambda L_\lambda$ are all elements of some L_λ, so no linear combination of these could vanish without violating linear independence of that L_λ). This $\bigcup_\Lambda L_\lambda$ is clearly an upper bound (in the partially ordered set **L**) of L_λ (λ in Λ). Thus the hypothesis of Zorn's lemma is satisfied by **L**—so there exists a maximal element K of **L**. (That is, K is a linearly independent subset of V and is such that, if any other elements of V are included in K, the resulting subset is no longer linearly independent.) We claim that this K is a basis for V (i.e., that K spans V). Suppose not. Then there would be a v in V which could not be expressed as a linear combination of elements of K. But, were this the case, we could include v in K and still have a linearly independent subset of V, violating maximality of K. Hence this K is in fact a basis for V. ☐

The proof is a bit tricky because the essential idea ("keep on adding vectors to a linearly independent set, preserving linear independence, until you get a basis") tends to get lost in all the groundwork at the beginning which prepares the way for Zorn's lemma. (If you think there is no content to Zorn's lemma, try to find explicitly a basis for the vector space of the first example of this chapter.)

Theorem 19 is the reason why vector spaces are so simple. There is nothing much to sets, every set gives rise to a vector space (the free one on that set), and every vector space is isomorphic to one of these. The situation is much nicer than even that for groups.

The mopping-up is completed by

THEOREM 20. *Let V be a free vector space on S, and V' a free vector space on S'. If V and V' are isomorphic (as vector spaces), then S and S' are isomorphic (as sets).*

Proof. From the remarks above, it suffices to prove that, if V is any vector space, and K and L are two bases for V, then there is a one-to-one, onto mapping from set K to set L. Denote by **B** the collection of all triples, (P,Q,φ), where P is a subset of K, Q is a subset of L, and φ is a one-to-one onto mapping from P to Q, which are such that the following condition is satisfied: $(L - Q) \cup P$ (i.e., the set of elements of V which are either in L and not Q, or in P) is a basis for V. Introduce the following partial ordering on **B**: $(P,Q,\varphi) \leq (P',Q',\varphi')$ if P is a subset of P', Q is a subset of Q', and $\varphi = \varphi'$ on P (i.e., wherever both are defined). Any totally ordered subset of **B**, $(P_\lambda, Q_\lambda, \varphi_\lambda)$ (λ in Λ), is bounded above by (P,Q,φ), where $P = \bigcup_\Lambda P_\lambda$, $Q = \bigcup_\Lambda Q_\lambda$, and where φ is defined as follows: for p in P (so p is in one of the P_λ), $\varphi(p) = \varphi_\lambda(p)$. By Zorn's lemma there is a maximal element, (P,Q,φ), of **B**. We must show that $P = K$ and $Q = L$ (for then φ is a one-to-one, onto mapping from K to L). Suppose, for example, that P did not equal all of K: then there would be an element k in K but not in P. Since $(L - Q) \cup P$ is a basis, this k can be written as a linear combination of elements from this set. This linear combination cannot contain only elements from P, for that would violate linear independence of K. Hence at least one element, $\mathbf{l}$, of $L - Q$ must be included in the linear combination. But now (P',Q',φ'), where P' consists of P with k included, Q' consists of Q with $\mathbf{l}$ included, and φ' is the mapping from P' to Q' which agrees with φ on P and which sends k to $\mathbf{l}$, is also an element of **B**. Furthermore, $(P,Q,\varphi) \leq (P',Q',\varphi')$. This violates maximality of (P,Q,φ). Thus $P = K$ and, similarly, $Q = L$. □

The proof of theorem 20 is rather like that of theorem 19, but with even more technical details around to obscure the idea of the proof. The idea of the proof is this (figure 28). We have a basis K and a basis L for V, and we wish to obtain a one-to-one, onto mapping from K to L. We want to "replace the basis vectors in L by those in K, one at a time, until L itself has disappeared and been replaced by K." A (P,Q,φ) represents a "stage in this replacement

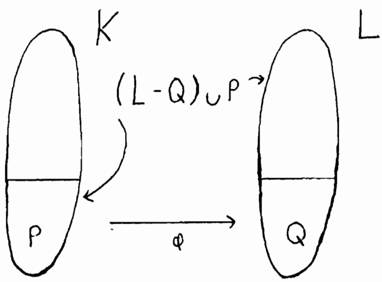

Figure 28

process, namely the stage in which we can replace all the vectors in the part Q of L by the vectors in the part P of K and still get a basis (the φ just ensures that we put back into L the same number of vectors that we take out)." The partial ordering represents "how far along we are in the replacement process." We "keep replacing vectors in L by those in K, going on and on until we cannot go further." The result is the maximal (P,Q,φ). This maximal "stage of replacement" must be such that "we have replaced all of L by K, for if it were not, we could always switch one more vector over from K to L." This is just a fancy version of the familiar row-reduction procedure for matrices.

The notion of the dimension of a vector space arises from theorems 19 and 20. Let V be any vector space. Then V is isomorphic to the free vector space on some set S (and any other set for which this is true is isomorphic to this S). If S has a finite number n of elements, V is said to have ***dimension n***. If S has an infinite number of elements, V is said to be ***infinite-dimensional.*** (Of course, one could further subdivide the infinite-dimensional vector spaces according to the cardinality of S.)

Exercise 63. Prove that, in the category of vector spaces, a linear mapping is one-to-one if and only if it is a monomorphism. (Hint: The proof that one-to-one implies monomorphism is easy. For the converse, consider $X \underset{\alpha'}{\overset{\alpha}{\rightrightarrows}} V \xrightarrow{\varphi} W$, where X is a free vector space on a set with one element.)

Exercise 64. Let V be a vector space, K a subset of V which spans V, and K' a subset of K which is linearly independent. Prove that there is a basis for V which contains K' and is contained in K.

Exercise 65. Give an alternative proof of theorem 19 in which one "begins with a subset which spans V, and keeps removing elements."

Exercise 66. Let V be a vector space. Prove that a subset K of V is a basis if and only if it has the following property: if any element of V not in K is included in K, the result is not linearly independent, while if any element of K is removed from K, the result does not span V.

Exercise 67. Let V and V' be free vector spaces on sets S and S', respectively. Prove that there is a monomorphism from V to V' (category of vector spaces) if and only if there is a monomorphism from S to S' (category of sets). Similarly for epimorphism. What does this mean for finite-dimensional vector spaces?

Exercise 68. Let V and V' be free vector spaces on sets S and S', respectively. Show that any mapping from S to S' induces a natural linear mapping from V to V'. (Compare the second example of this chapter. What goes wrong if one tries to use that method to obtain, from $S \xrightarrow{\mu} S'$, a linear mapping from V' to V?)

Exercise 69. Can the underlying abelian groups of two vector spaces be isomorphic (as groups), although the vector spaces themselves are not isomorphic? Is there a simple characterization of which abelian groups are underlying groups of vector spaces?

10

Subspaces

Let V be a vector space. A *subspace* of V is a vector space W together with a monomorphism $W \xrightarrow{\varphi} V$. Denote by W' the collection of all elements of V of the form $\varphi(w)$ with w in W. Then 0 $(= \varphi(0))$ is in W', and, for w_1' $(= \varphi(w_1))$ and w_2' $(= \varphi(w_2))$ in W', so is $w_1' + aw_2'$ $(= \varphi(w_1 + aw_2))$. Thus, under addition and multiplication by numbers (as given from V), W' itself is a vector space, and, clearly, φ is an isomorphism from W to W'. Thus we could just as well have defined a subspace as follows: for V a vector space, a subspace of V is a subset W of V such that i) 0 is in W, and ii) whenever w and w' are in W, and a is a number, $w + aw'$ is in W. Thus a subspace of a vector space is a subset such that "one remains in the subset under all vector-space operations on elements of that subset."

Note that any intersection of subspaces of a vector space is again a subspace. [Proof: Let W_λ (λ in Λ) be subspaces, and set $W = \bigcap_{\Lambda} W_\lambda$. Then 0 is in W, since it is in each W_λ. Furthermore, for w and w' in W, so is $w + aw'$, for w and w' must be in each W_λ, whence $w + aw'$ must be in each W_λ.] Let K be any subset of vector space V. The intersection of all subspaces of V containing K is a subspace of V called the subspace *generated* by K. (Note that this is precisely the subspace containing zero and all linear combinations of elements of K.) Thus, for example: a subset of V spans V if and only if V is in fact the subspace of V generated by this subset.

Example. A subset K of V is a basis for V if and only if K has the following property: the subspace generated by K is V itself, and the subspace generated by any K', consisting of K with one or more elements removed, is not V.

We introduce the following notation for subsets of a vector space (similar to that for groups). For K and L subsets of vector space V, we write $K + L$ for all elements of V which can be written in the form $k + l$ with k in K and l in L. We write aK (a a number) for the collection of all elements of V which can be written in the form ak with k in K. For subsets having only one element, we represent the subset by writing that element, for example, $k + L$. Thus, for W a subspace of V, we have $W + W = W$ and, for $a \neq 0$, $aW = W$.

Note that a subspace of a vector space is also a subgroup of the corresponding abelian group.

Let W be a subspace of vector space V. A *coset* of W in V is a subset of V which can be written in the form $v + W$ for some v in V. Note that every element of V is in some coset of W in V (namely, v is in $v + W$ for $v = v + 0$) and, in fact, in a unique one (for, if v were in $v' + W$, we would have $v = v' + w$, with w in W, whence $v' + W = v - w + W = v + W$). The cosets of W in V "cover V without overlapping." (Note that we do not have to discuss left and right cosets; they would turn out to be the same, because addition in V is commutative.)

Let W be a subspace of vector space V. We denote the collection of all cosets of W in V by V/W. We now define a few operations on this set V/W. For $v + W$ and $v' + W$ cosets, we define their sum by $(v + v') + W$ (a result which, clearly, depends only on the cosets themselves and not on how each is represented as "$v + W$"). For $v + W$ a coset and a a number, we define a new coset by $av + W$ (which, again, depends only on the coset itself). Thus V/W is now a set on which we can "add and multiply by numbers." We now claim that, with these operations, V/W is in fact a vector space, that is, that conditions 1–4 in the definition of a vector space are indeed satisfied. (Each property follows immediately from the corresponding property in V.)

This V/W is called the *quotient space* of V by (the subspace) W.

We now define a mapping φ from the set V to the set V/W as follows: for v in V, $\varphi(v)$ is the coset $v + W$. This φ is, in fact, a linear mapping from vector space V to vector space V/W. [Proof: $\varphi(v + av') = v + av' + W$, which is precisely the coset $v + W$ plus a times the coset $v' + W$, which is precisely $\varphi(v) + a\varphi(v')$.]

We have already seen that vector spaces are simpler than groups because all vector spaces are free, while not all groups are. Now, vector spaces are again simpler, because all subspaces have quotients, while all subgroups (i.e., the ones which are not normal) do not.

Not only does any subspace of a vector space have a quotient space, but the quotient space can, in a certain sense, be "put back into the vector space." We now explain what this means. Let V be a vector space. Two subspaces, U and W, of V will be said to be *complementary* if i) every element v of V can be written in the form $v = u + w$ with u in U and w in W, and ii) whenever $u + w = 0$, with u in U and w in W, we have $u = 0$ and $w = 0$. (The first condition ensures that the union of U and W spans V; the second states that the intersection of U and W is the subspace consisting only of 0. In this sense, U "complements" W.) This definition has rather the flavor of that of a basis.

Let W be a fixed subspace of vector space V. Let U be another subspace of V such that U and W are complementary. We wish to show that U represents "V/W inserted into V." First note that, for any u in U, $u + W$ is certainly a coset of W in V. Now consider any coset of W in V, for example, $v + W$ (v in V). Write $v = u + w$, with u in U and w in W. Then $v + W =$

$u + w + W = u + W$. Thus every coset of W in V, that is, every element of V/W, can be written in the form $u + W$ with u in U. Finally, note that, if $u + W = u' + W$ (u and u' in U), so $u - u' = w$ (w in W), then, by condition ii) above, we have $u - u' = 0$, whence $u = u'$. Thus every coset of W in V can be written uniquely in the form $u + W$, with u in U. We have therefore an isomorphism from the set U to the set V/W. But the mapping from, for example, U to V/W (which sends u in U to $u + W$) is clearly a linear mapping of vector spaces. We have proven the following.

THEOREM 21. *Let U and W be complementary subspaces of vector space V. Then every coset of W in V can be written in one and only one way in the form $u + W$ with u in U, and this correspondence between U and V/W is an isomorphism of vector spaces.*

In this sense, then, U "represents V/W in V."

What we have not yet shown, however, is that such complementary subspaces exist. We do this now.

THEOREM 22. *Let W be a subspace of vector space V. Then there exists a subspace U of V such that U and W are complementary.*

Proof. Denote by $\mathbf{U}$ the collection of all subspaces U of V such that $u + w = 0$ (u in U and w in W) implies $u = w = 0$. Order by inclusion, so $\mathbf{U}$ is partially ordered. If U_λ (λ in Λ) is any totally ordered subset of $\mathbf{U}$, then it is bounded above, by $\bigcup_\Lambda U_\lambda$. (Note that $\bigcup_\Lambda U_\lambda$ is indeed in $\mathbf{U}$, for it is a subspace, and, if $u + w = 0$, with w in W and u in $\bigcup_\Lambda U_\lambda$, then u is in some U_λ, whence $u = 0$.) By Zorn's lemma, there is a maximal element, U, of $\mathbf{U}$. We claim that U and W are complementary, that is, that every v in V can be written $v = u + w$ with u in U and w in W. Suppose not, for some v in V. Let U' be the subspace of V generated by U and v, so every element of U' is of the form $u + av$ (u in U). Then U' is also in $\mathbf{U}$, for $u + av + w = 0$ implies (since v cannot be the sum of an element from U and one from w) $a = 0$, $u = 0$, and $w = 0$. But $U \leq U'$, contradicting maximality of U. Hence U and W are complementary. □

It is the usual Zorn lemma thing. We consider subspaces which satisfy half the properties for complementarity and maximize to get the other half of the properties satisfied too.

We illustrate the situation with the following example.

Example. Let V be a three-dimensional vector space. Let x, y, and z be three elements of V which form a basis, so every element v of V can be written $v = ax + by + cz$ (uniquely), with a, b, and c numbers. Let W be the subspace of V consisting of all elements of V of the form ax, so W is the subspace generated by the element x of V. The cosets of W in V are subsets of the form $v + W$. Thus two elements of V, $ax + by + cz$ and $a'x + b'y + c'z$, are in the same coset if and only if $b' = b$ and $c' = c$. In figure 29, the

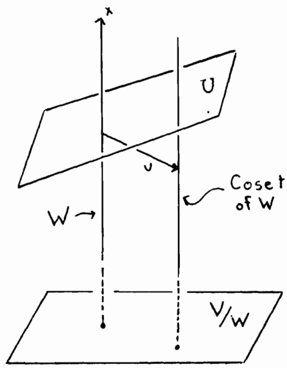

Figure 29

cosets are "vertical straight lines." Then V/W is the "space of all vertical straight lines," that is, the horizontal plane in figure 29. The mapping from V to V/W is the mapping which projects each vector in V downward to the plane at the bottom. A subspace U complementary to W is, for example, the plane in the top half of the figure. (The most general such U is that generated by $y + ax$ and $z + bx$ for fixed numbers a and b.) The isomorphism from V/W to U is "vertical motion upward from V/W until you reach U." The coset of W in V shown in the figure is $u + W$.

Exercise 70. Prove that, in the category of vector spaces, a linear mapping is onto if and only if it is an epimorphism. (Hint: The proof that onto implies epimorphism is easy. For the converse, consider $V \xrightarrow{\varphi} W \overset{\beta}{\underset{\beta'}{\rightrightarrows}} X$, and choose $X = W$ and $\beta = i_W$. Let $P = \varphi[V]$, and let Q be a complementary subspace. Let β' agree with β on P, and let $\beta'(q) = 0$ for q in Q.)

Exercise 71. Let W be a subspace of vector space V. Prove that any basis for W is a subset of some basis for V. Obtain, from a basis for V so obtained, a basis for V/W.

Exercise 72. Let U and U' be complementary subspaces of V, and let S, S', and T be sets the free vector spaces over which are isomorphic to U, U', and V, respectively. Prove that T is isomorphic to $S \cup_d S'$ (disjoint union).

Exercise 73. Let W be a fixed subspace of vector space V. Let U be a complementary subspace, and let ψ be a linear mapping from U to W. Let U' consist of all elements of V which can be written in the form $u + \psi(u)$, with u in U. Prove that U' is a subspace and is, in fact, a complementary subspace to W. (Thus complementary subspaces are not, except in degenerate cases, unique.) Prove, furthermore, that every complementary subspace to W can be obtained, from U, by this construction.

Exercise 74. Let K, L, and M be subsets of vector space V, and let $K + L = M$. Does therefore $K = M - L$?

Exercise 75. Let V be a vector space, and U, U', U'' subspaces of V with U contained in U' and U' contained in U''. Consider: $(U''/U)/(U'/U) = U''/U'$. State the theorem and prove it.

Exercise 76. Find all vector spaces whose only subspaces are the one containing only 0 and the entire vector space.

Exercise 77. Find an example of a vector space V which has a subspace (other than V itself) that is isomorphic to V.

11

Linear Mappings; Direct Products and Sums

The analysis of linear mappings of vector spaces is essentially the same as the analysis of homomorphisms of groups.

Let $V \xrightarrow{\varphi} W$ be a linear mapping of vector spaces. We denote by Ker(φ), the *kernel* of φ, the collection of all v in V with $\varphi(v) = 0$, and by Im(φ), the *image* of φ, the collection of all w in W such that $w = \varphi(v)$ for some v in V. Then Ker(φ) is a subspace of V [proof: $\varphi(0) = 0$, and, for v and v' in Ker(φ), we have $\varphi(v + av') = \varphi(v) + a\varphi(v') = 0$, so $v + av'$ is in Ker(φ)] and Im(φ) is a subspace of W [proof: $\varphi(0) = 0$, and, for w ($= \varphi(v)$) and w' ($= \varphi(v')$) in W, so is $w + aw'$ ($= \varphi(v + av')$)].

Consider $V \xrightarrow{\alpha} V/\text{Ker}(\varphi) \xrightarrow{\tilde{\varphi}} \text{Im}(\varphi) \xrightarrow{\beta} W$, where α is the mapping that associates, with each v in V, the coset v + Ker of Ker(φ) in V; $\tilde{\varphi}$ is the mapping that associates, with the coset v + Ker(φ) of Ker(φ) in V, the element $\varphi(v)$ of Im(φ) (noting that this $\varphi(v)$ is independent of how the coset is written as v + Ker(φ)); and β is the mapping that represents Im(φ) as a subset of W. These are all linear mappings; α is an epimorphism, $\tilde{\varphi}$ an isomorphism, and β a monomorphism.

Now let V and W be any two vector spaces (both real or both complex). We construct a new vector space (written $V \oplus W$). Let the set $V \oplus W$ consist of pairs, (v,w), with v in V and w in W. We define addition of such pairs by $(v,w) + (v',w') = (v + v', w + w')$. We define multiplication of such pairs by numbers by $a(v,w) = (av,aw)$ (a a number). We claim that, with these operations, the set $V \oplus W$ becomes a vector space. (Each of properties 1–4 is immediate from the corresponding property for V and W.) Now consider the "butterfly diagram" shown in figure 30. Here, α is the linear mapping from vector space V to vector space $V \oplus W$ with action $\alpha(v) = (v,0)$, and β is the mapping with action $\beta(w) = (0,w)$. Furthermore, α' is the linear mapping from vector space $V \oplus W$ to vector space V with action $\alpha'(v,w) = v$, and β the mapping with action $\beta(v,w) = w$. We claim, first, that this diagram commutes (for, e.g., for v in V, $\alpha' \circ \alpha(v) = \alpha'(v,0) = v = i_V(v)$). Note also that α and β are monomorphisms, that α' and β' are epimorphisms, and that Im(α) = Ker(β') and Im(β) = Ker(α'). Finally, V and W (regarded as subspaces, via the monomorphisms α and β, of $V \oplus W$) are complementary.

One can also describe the structure of this $V \oplus W$ using bases. Let K and L be bases for V and W, respectively. Then we claim that the collection

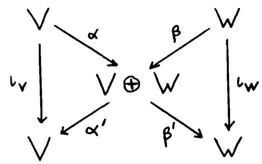

Figure 30

of all elements of $V \oplus W$ of the form $(k,0)$ (k in K), together with those of the form $(0,l)$ (l in L), forms a basis for $V \oplus W$. [Proof: Since K spans V and L spans W, any element of $V \oplus W$, $(v,w) = (v,0) + (0,w)$, can be written as a linear combination of elements of the form $(k,0)$ and $(0,l)$. If some such linear combination vanished, e.g., if $(a_1k_1 + \cdots + a_nk_n, b_1l_1 + \cdots + b_ml_m)$ vanished, then we would have $a_1k_1 + \cdots + a_nk_n = 0$ and $b_1l_1 + \cdots + b_ml_m = 0$, whence, by linear independence of K in V, $a_1 = \cdots = a_n = 0$, and, by linear independence of L in W, $b_1 = \cdots = b_m = 0$.] Thus a basis for $V \oplus W$ is isomorphic (as a set) to the disjoint union of a basis for V and a basis for W. (In this sense, at least, $V \oplus W$ is more "union-like" than "product-like.")

We have

THEOREM 23. *Let V and W be vector spaces. Then $(V \oplus W,\alpha,\beta)$ is a direct sum of V and W, while $(V \oplus W,\alpha',\beta')$ is a direct product of V and W.*

Proof. Consider the diagram of figure 31. Since $(v,0) = \alpha(v)$, we must have, in order that the diagram commute, $\gamma(v,0) = \mu(v)$, and, similarly, $\gamma(0,w) = \nu(w)$. In order that γ be a linear mapping, we must therefore have $\gamma(v,w) = \gamma[(v,0) + (0,w)] = \mu(v) + \nu(w)$. But this γ, so defined, is indeed a linear mapping. Hence $V \oplus W$ is a direct sum.

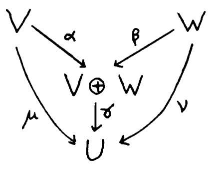

Figure 31

In the diagram of figure 32, we must have, in order that the diagram commute, that $\gamma(u)$ (u in U) is that element of $V \oplus W$ with $\alpha'[\gamma(u)] = \sigma(u)$ and $\beta'[\gamma(u)] = \tau(u)$. But there is only one such element of $V \oplus W$; hence we must have $\gamma(u) = (\sigma(u),\tau(u))$. But this γ, so defined, is indeed a linear mapping. Hence $V \oplus W$ is a direct product. □

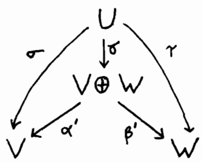

Figure 32

This $V \oplus W$ is normally called the direct sum of V and W. For vector spaces, as for abelian groups, the direct product and direct sum are the same thing (for a finite number of vector spaces; they actually differ slightly for an infinite number).

Exercise 78. Let V be a vector space. Find a necessary and sufficient condition on vector space W such that there exists an isomorphism from V to $V \oplus W$.

Exercise 79. Given linear mappings from V to V' and from W to W', find a corresponding linear mapping from $V \oplus W$ to $V' \oplus W'$. When is the latter a monomorphism? When an epimorphism?

Exercise 80. Let U and V be vector spaces. Make $\mathrm{Mor}(U, V)$ into a vector space. Now let W be another vector space, and fix a linear mapping κ from V to W. Then "composition with κ" is a mapping from $\mathrm{Mor}(U, V)$ to $\mathrm{Mor}(U, W)$. Prove that this is a linear mapping of vector spaces. What is its image and kernel?

Exercise 81. Construct a direct product and a direct sum for an arbitrary collection of vector spaces.

Exercise 82. Describe a subspace of $V \oplus W$ completely in terms of subspaces, linear mappings, etc., but involving only V and W separately.

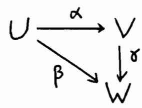

Figure 33

Exercise 83. In the diagram of figure 33, let α and β be given linear mappings. Prove that there exists a linear mapping γ which makes the diagram commute if and only if $\mathrm{Ker}(\alpha)$ is a subset of $\mathrm{Ker}(\beta)$. When is there a unique γ which makes the diagram commute?

12

From Real to Complex Vector Spaces and Back

It is by no means rare in applications that one wishes to obtain, from a given real vector space, a complex one, or vice versa. There are at least three techniques available for making such transitions. We here describe them.

Let V be a fixed complex vector space. We obtain a real vector space. Let $\tilde{V}$ be a set, and $V \xrightarrow{\varphi} \tilde{V}$ an isomorphism of sets. (Instead of writing $\varphi(v)$ all the time, we shall write $\tilde{v}$, so a tilde denotes membership in $\tilde{V}$.) We now define some operations on this set $\tilde{V}$. For $\tilde{v}$ and $\tilde{v}'$ in $\tilde{V}$, we define the left side of $\tilde{v} + \tilde{v}' = \widetilde{(v + v')}$ by the right side. (That is, to "add" two elements of the set $\tilde{V}$, you carry them, via the isomorphism, back to V, there add them (in the vector-space structure of V), and then carry the result, via the isomorphism, back to $\tilde{V}$.) Next, for $\tilde{v}$ in $\tilde{V}$, and a a real number, we define the left side of $a\tilde{v} = \widetilde{(av)}$ by the right side. (Carry $\tilde{v}$ back to V, there multiply by the real number a (we even know how to multiply by complex numbers in V), and carry the result back to $\tilde{V}$.) We claim that this set $\tilde{V}$, with these operations, is a real vector space. [Proof: Properties 1–4 hold in V. Applying a tilde to each, we obtain properties 1–4 in $\tilde{V}$.]

Thus we obtain, from a complex vector space, a real one. All we have done is to "forget the possibility of multiplying by complex numbers in V, using instead only multiplication by real numbers." Not only is this $\tilde{V}$, so constructed, a real vector space; it has some additional structure. We introduce a linear mapping ι from the real vector space $\tilde{V}$ to $\tilde{V}$, defined as follows: for $\tilde{v}$ in $\tilde{V}$, $\iota(\tilde{v}) = \widetilde{(iv)}$. (Here i is that complex number; iv is well defined, because V is just a complex vector space.) This ι is clearly a mapping from $\tilde{V}$ to $\tilde{V}$—and is clearly a linear mapping. Furthermore, we have $\iota \circ \iota = -i_{\tilde{V}}$, where $-i_{\tilde{V}}$ is the linear mapping from $\tilde{V}$ to $\tilde{V}$ which sends $\tilde{v}$ in $\tilde{V}$ to $-\tilde{v}$. ("Since $\tilde{V}$ is a real vector space, you cannot multiply by i therein. But the fact that this $\tilde{V}$ originally arose from a complex vector space is nonetheless reflected within this $\tilde{V}$ by the mapping ι.")

The second construction is just the reversal of the first. Let $\tilde{V}$ be a real vector space, and let ι be a linear mapping from $\tilde{V}$ to $\tilde{V}$ satisfying $\iota \circ \iota = -i_{\tilde{V}}$. (It follows that this ι is an isomorphism from $\tilde{V}$ to $\tilde{V}$, for $\iota \circ (\iota \circ \iota \circ \iota) = (\iota \circ \iota \circ \iota) \circ \iota = i_{\tilde{V}}$.) We define a certain complex vector space. Let V be a set, and $V \xrightarrow{\varphi} \tilde{V}$ an isomorphism of sets. As above, we use attachment and removal of a tilde to denote passage, via this isomorphism, from set V to set

$\tilde{V}$. We define addition within this V as follows: $\widetilde{(v + v')} = \tilde{v} + \tilde{v}'$. (That is, $v + v'$ is defined as that element of the set V which, when carried over to the real vector space $\tilde{V}$, is the element $\tilde{v} + \tilde{v}'$ of $\tilde{V}$.) We next define, within this set V, multiplication by complex numbers. Let $(a + ib)$ be a complex number (a and b real numbers), and let v be an element of V. The definition is $[\widetilde{(a + ib)v}] = a\tilde{v} + b\iota(\tilde{v})$. (The instructions: For v in V, $(a + ib)v$ is that element of V which, when carried to the real vector space $\tilde{V}$, is the sum in $\tilde{V}$ of the element $a\tilde{v}$ of $\tilde{V}$ and the element $b\iota(\tilde{v})$ of $\tilde{V}$. Note that this is all well defined and that, e.g., $[\widetilde{(a + ib)v}] = (a + ib)\tilde{v}$ is not, for we do not know how to multiply, in $\tilde{V}$, by complex numbers.) We now claim that this set V, with these operations, is indeed a complex vector space. (All properties are immediate except possibly property 3, which, in turn, is immediate from $\iota \circ \iota = -i_{\tilde{V}}$.)

Thus a complex vector space defines a real vector space (on which there is a certain linear mapping ι), while a real vector space with such an ι defines a complex vector space.

The final construction yields a complex vector space from a real one, even though there is not given, on the latter, any preferred ι. Let W be a real vector space, and consider the real vector space $W \oplus W$. We define a linear mapping $W \oplus W \xrightarrow{\iota} W \oplus W$ by $\iota(w,w') = (-w',w)$ (w and w' in W, so (w,w') is in $W \oplus W$). We have, for this ι, $\iota \circ \iota(w,w') = \iota[(-w',w)] = (-w,-w') = i_{W \oplus W}(w,w')$. That is, we have $\iota \circ \iota = -i_{W \oplus W}$. Thus, by the preceding construction, we obtain a complex vector space. (This is the construction in which one allows oneself to "naively multiply vectors in a real vector space by complex numbers." Write $(w,w') = (w,0) + (0,w') = (w,0) + \iota(w',0)$. Think of $(w,0)$ as the "real part" of (w,w') and of $(0,w')$ as the "imaginary part," so ι is "multiplication by i.")

Of these three constructions, the first is used rarely (because, I suppose, it has so little content), and the second and third occasionally. We conclude with the following remark: there is no such thing, in a complex vector space, as a "real vector" or as the "complex-conjugate of a vector." These notions are simply not part of the structure that goes into a complex vector space. One needs, for such notions, a preferred isomorphism from the complex vector space to a complex vector space of the form $W \oplus W$ (W a real vector space) as obtained in the last construction above.

Exercise 84. Let $\tilde{V}$ be a real vector space. Prove that there exists a linear mapping ι from $\tilde{V}$ to $\tilde{V}$ satisfying $\iota \circ \iota = -i_{\tilde{V}}$ if and only if either i) $\tilde{V}$ is infinite-dimensional, or ii) $\tilde{V}$ has (finite) even dimension.

Exercise 85. Let $\tilde{V}$ be a real vector space, with linear mapping ι from $\tilde{V}$ to $\tilde{V}$ satisfying $\iota \circ \iota = -i_{\tilde{V}}$. Let $\tilde{W}$ be a subspace of $\tilde{V}$. Find a simple necessary and sufficient condition that the corresponding subset W of the corresponding complex vector space V be a subspace.

Exercise 86. Let $\tilde{V}$ and $\tilde{V}'$ be real vector spaces, with $\tilde{V} \xrightarrow{\iota} \tilde{V}$ and $\tilde{V}' \xrightarrow{\iota'} \tilde{V}'$ satisfying $\iota \circ \iota = -i_{\tilde{V}}$ and $\iota' \circ \iota' = -i_{\tilde{V}'}$. Let $\tilde{V} \xrightarrow{\alpha} \tilde{V}'$. Find a necessary and sufficient condition that the corresponding mapping from the complex vector space V to the complex vector space V' be linear.

Exercise 87. In each of the three constructions of this section, one obtains one vector space from another. Given in each case a basis for the original vector space, find a set isomorphic to a basis for the constructed vector space.

Exercise 88. Let $\tilde{V}$ and $\tilde{V}'$ be real vector spaces, with $\tilde{V} \xrightarrow{\iota} \tilde{V}$ and $\tilde{V}' \xrightarrow{\iota'} \tilde{V}'$ satisfying $\iota \circ \iota = -i_{\tilde{V}}$ and $\iota' \circ \iota' = -i_{\tilde{V}'}$. Introduce a similar linear mapping from $\tilde{V} \oplus \tilde{V}'$ to $\tilde{V} \oplus \tilde{V}'$. Show that the corresponding complex vector space is isomorphic to the complex vector space $V \oplus V'$.

13

Duals

One of the very useful constructions in working with vector spaces is that of taking the dual. We here describe the situation. It is convenient (because we shall so frequently refer to the numbers) to carry out the discussion in the real case (denoting the collection of real numbers by **R**), noting at the end that nothing whatever changes in the complex case.

Fix a (real) vector space V. A mapping f from the set V to the set **R** of real numbers (i.e., a real-valued function on the set V) is said to be linear if $f(av + v') = af(v) + f(v')$ (v and v' in V, a in **R**). (Thus f is a linear mapping from the vector space V to the vector space **R**.) Denote by V^* the collection of all such linear mappings. We define, on this set V^*, addition and multiplication by real numbers as follows:

$$(f + f')(v) = f(v) + f'(v)$$

and

$$(af)(v) = a[f(v)]$$

(the left side in each case defines an element of V^* by specifying the number it associates with v in V). With these operations, this set V^* becomes a real vector space. This vector space V^* is called the *dual* of the vector space V.

The following is a geometrical picture of an element of V^*. Let f be a nonzero element of V^*, and consider Ker(f) (i.e., the collection of all v in V such that $f(v) = 0$). This Ker(f) is, of course, a subspace of V. We claim, furthermore, that the quotient space, $V/\text{Ker}(f)$, is one-dimensional (for there is an isomorphism from $V/\text{Ker}(f)$ to Im(f), while Im(f) = **R**). Next, note that, if f and f' are nonzero elements of V^* and are proportional (i.e., if $f' = af$), then Ker(f) = Ker(f') (for, evidently, $f(v) = 0$ if and only if $f'(v) = 0$). There is a sort of converse to this last statement: if W is any subspace of V, with V/W one-dimensional, then there is an f in V^*, unique up to a factor, with Ker(f) = W. [Proof: Let U be a (necessarily one-dimensional) complementary subspace to W. Fix $u \neq 0$ in U, so this u is a basis for U. Then, for v in V, write v (uniquely) as $v = au + w$ (a in **R**, w in W), and set $f(v) = a$. Then Ker(f) = W, and this f is unique up to a factor.] Thus subspaces W of V with V/W one-dimensional represent "elements of V^* up to a nonzero factor."

Since every vector space has a basis, the following example is, in a sense, generic. It will be necessary to refer to it again later.

Example. Let V be a vector space, and K (a subset of V) a basis for V. Fix any real-valued function $\hat{f}$ on the set K. We define, using this $\hat{f}$, an element f of V^*. For any v in V, write $v = a_1k_1 + \cdots + a_nk_n$ ($a_1, \ldots, a_n$ in $\mathbf{R}$, $k_1, \ldots, k_n$ in K), and set $f(v) = a_1\hat{f}(k_1) + \cdots + a_n\hat{f}(k_n)$. (That is, one writes v as a linear combination of elements of the basis K, evaluates the function $\hat{f}$ on each of these basis vectors to get a real number, and takes the corresponding linear combination of real numbers to get $f(v)$.) Thus every real-valued function on the set K defines an element of V^*. A converse is also true: every element of V^* arises in this way. Indeed, let f be in V^* (so f associates a real number with each element of V and, in particular, with each element of K). Then $\hat{f}(k) = f(k)$, for each k in K, defines a real-valued function $\hat{f}$ on the set K. Evidently, $f(a_1k_1 + \cdots + a_nk_n) = a_1\hat{f}(k_1) + \cdots + a_n\hat{f}(k_n)$. Thus this function $\hat{f}$ on the set K indeed defines our original element f of V^*. That is, V^* is precisely the vector space of real-valued functions on the set K.

The discussion above is misleading in one important respect. Recall that we can regard V as the vector space of all real-valued functions on the set K having the property that they vanish for all but a finite number of elements of K. But V^* is the vector space of all real-valued functions on the set K. Hence we may regard V as a subspace of V^*. Unfortunately, this representation of V as a subspace of V^* depends on the choice of basis: if we choose a different basis, a fixed element of V will define a quite different element of V^*. That is, although there always exists a monomorphism from V to V^*, there is no "natural" (i.e., basis-independent) one.

Now let V and W be (real) vector spaces, and $V \xrightarrow{\varphi} W$ a linear mapping. We define a linear mapping $W^* \xrightarrow{\varphi^*} V^*$ as follows: for f in W^* (i.e., f a real-valued function on W),

$$\varphi^*(f) = f \circ \varphi$$

(the right side is the real-valued function on V obtained by first sending v in V to W using φ, and there (in W) evaluating f). This φ^* is clearly a linear mapping of vector spaces. (It is often called the adjoint of φ.) Thus, not only does every vector space give rise to another vector space (its dual), but every linear mapping between vector spaces gives rise to an "order reversed" linear mapping on the corresponding dual spaces.

Now let $U \xrightarrow{\alpha} V \xrightarrow{\beta} W$ be linear mappings of vector spaces. We claim that

$$(\beta \circ \alpha)^* = \alpha^* \circ \beta^* .$$

(Note first that this is meaningful, for $\beta \circ \alpha$ maps U to W, whence $(\beta \circ alpha)^*$ maps W^* to U^*, while β^* maps W^* to V^* and α^* maps V^* to U^*, whence $\alpha^* \circ \beta^*$ also maps W^* to U^*.) [Proof: $\alpha^* \circ \beta^*(f) = \alpha^*[\beta^*(f)] = \alpha^*[f \circ \beta] = f \circ \beta \circ \alpha = (\beta \circ \alpha)^*(f)$, for f in W^*.] A number of other properties are satisfied by this $*$-operation on linear mappings. Thus, if $V \xrightarrow{\varphi} W$ is a monomorphism, then φ^* is an epimorphism. [Proof: Let f be in V^*; we must find h in W^* so $\varphi^*(h) = f$. Since φ is a monomorphism, V is a subspace of W. Choose a complementary subspace U. Then, for $w = v + u$ in W, set $h(w) = f(v)$. This h is certainly in W^* and clearly satisfies $\varphi^*(h) = f$.] Similarly, if φ is an epimorphism, then φ^* is a monomorphism, and if φ is an isomorphism, so is φ^*.

This operation of taking the dual is often applied twice. Let V be a vector space. Then V^*, the dual of V, is also a vector space, so we can, in turn, take its dual, V^{**}. (Note that, e.g., if $V \xrightarrow{\varphi} W$ is a monomorphism, then $V^{**} \xrightarrow{\varphi^{**}} W^{**}$ is also a monomorphism, etc.) Of course, elements of V^{**} are real-valued functions on the collection of all linear, real-valued functions on V. There is an important relation between V and V^{**}: there is a linear mapping $V \xrightarrow{\mu} V^{**}$ of vector spaces. To say what this μ is, we must give a rule which associates, with each v in V, an element of V^{**}; that is, a rule which associates, with v, a real-valued function on linear functions on V. The rule is this: for v in V, $\mu(v)(f) = f(v)$, for each f in V^*. This formula indeed defines an element, $\mu(v)$, of V^{**}, for it associates a real number (namely, $f(v)$) with each element f of V^*, that is, for it defines a function, $\mu(v)$ on V^*. Thus, for each v in V, $\mu(v)$ is an element of V^{**}, so μ is a (clearly linear) mapping from V to V^{**}. This construction is a bit complicated; we repeat it in different words. Our V^* is a certain collection of real-valued functions on V. For fixed v in V, "evaluation at v" associates a number with each such function, that is, defines a function on V^*, that is, defines an element of V^{**}. Repeating for each v in V, we obtain a mapping from V to V^{**}.

We next claim that this linear mapping μ is in fact a monomorphism. What we must show is that if v in V has the property $\mu(v) = 0$, then $v = 0$. That is, we must show that, if v in V has the property that $f(v) = 0$ for every f in V^*, then $v = 0$. But this is true: for $v \neq 0$ in V, let U be a complementary subspace to the subspace generated by v. Then every element of V can be written in the form $av + u$ (a in $\mathbf{R}$, u in U): set $f(av + u) = a$. Then f is in V^*, and $f(v) = 1 \neq 0$. Thus we have a monomorphism from V to V^{**}. (Note that, since we have made no mention of bases lately, this monomorphism is basis-independent.) In other words, the original vector space V is in fact a subspace of V^{**}.

It is not too easy, offhand, to think of an element of V^{**} not in $\mathrm{Im}(\mu)$. One is thus led to ask when the monomorphism μ is in fact an isomorphism. The answer is pretty.

THEOREM 24. *Let V be a vector space, and $V \xrightarrow{\mu} V^{**}$ the natural monomorphism. Then μ is an isomorphism if and only if V is finite-dimensional.*

Proof. Let V have finite dimension n. Then, choosing a basis K for V (so K has exactly n elements), both V and V^* can (by the previous example) be regarded as the vector space of real-valued functions on the set K. (Since K has a finite number of elements, every such function vanishes outside of a finite set.) Hence $\dim(V) = \dim(V^*)$ and, similarly, $\dim(V^*) = \dim(V^{**})$. So $\dim(V) = \dim(V^{**})$. Thus $V \xrightarrow{\mu} V^{**}$ is a monomorphism from a finite-dimensional vector space to another of the same dimension, whence μ must be an isomorphism.

Let V be infinite-dimensional, and K a basis for V. We shall find an element of V^{**} not of the form $\mu(v)$. Regard V^* as the vector space of all real-valued functions on the set K, and denote by B the subspace of V^* consisting of bounded functions (noting that, since K has an infinite number of elements, there are nonbounded functions on K, so $B \neq V^*$). Let W be a complementary subspace to B, and fix, once and for all, an element $w \neq 0$ of W. Let U be a complementary subspace, in W, of the subspace generated by w. Finally, let X be the subspace of V^* generated by U and B. Then X is a complementary subspace (in V^*) to that generated by w. Now consider the function κ on V^* which associates, with the element $aw + x$ (a in $\mathbf{R}$, x in X) of V^*, the number a. This κ is certainly an element of V^{**}. Furthermore, (since B is a subspace of $X = \mathrm{Ker}(\kappa)$) $\kappa(f) = 0$ for any element of V^* (i.e., any function on κ) which is bounded. We claim, finally, that this element κ of V^{**} is not of the form $\mu(v)$ for any v in V, for, for any $v \neq 0$ in V, there is a bounded function on K (e.g., the function which represents v itself) which $\phi(v)$ fails to annihilate, while κ annihilates all bounded functions on K. □

The complication in the proof is the presence of so many subspaces. The idea of the proof is this: if we can only find an element κ of V^{**} which annihilates all the elements of V^* (i.e., all the functions on K) which are bounded, then we are done, for no element of V^{**} of the form $\mu(v)$ has this property. What functions, then, will κ annihilate? It must annihilate a subspace X of V^*, with V^*/X dimension one, and with X including all the founded functions. All that talk about complementary subspaces at the beginning is the construction of an appropriate X. You begin with the subspace B of bounded functions, pick one element w of a complementary subspace, and "join the rest of that complementary subspace (namely, U) with B to get our X." This X then leads immediately to κ, an element of V^{**} which does not come, via μ, from any element of V.

Theorem 24 is perhaps the basic reason why finite-dimensional vector spaces are simpler than infinite-dimensional ones. With infinite-dimensional vector spaces, one "keeps getting bigger things on taking duals," while, in the finite-dimensional case, the situation remains more manageable. (We remark that, when a sufficiently nice topology is available on V (e.g., for a Hilbert space), the definition of the dual must be modified to take into account this additional topological structure, with the result that we can have $V = V^{**}$ even in the infinite-dimensional case.)

It is interesting that intricate arguments such as the one above should be available in vector spaces when, since every vector space is free, the subject looks at the beginning as though it will be straightforward.

This entire chapter can be repeated, replacing everywhere "real" by "complex."

Exercise 89. State and prove: $(V \oplus W)^* = V^* \oplus W^*$.

Exercise 90. Let V be a vector space. We have seen that a one-dimensional subspace of V^* is completely and uniquely determined by a subspace W of V with V/W one-dimensional. Show that a two-dimensional subspace of V^* is similarly determined by a subspace W of V with V/W two-dimensional.

Exercise 91. Show explicitly that the linear mapping from V to V^* (in the example above) in fact depends on choice of basis.

Exercise 92. Let V be a vector space. For f in V^*, denote by $U(f)$ the collection of all v in V such that $f(v) = 1$. Show that $U(f) = U(f')$ implies $f = f'$. Characterize the subsets of V which are of the form $U(f)$ for some f in V^*. Thus one could as well have defined V^* as the collection of all subsets of V which satisfy this characterization. Now figure out, geometrically, how to "add" such subsets, and "multiply by numbers," to reflect the operations available in V^*. Describe various things about the dual geometrically, for example, the action of μ.

Exercise 93. Is there a reasonable notion of a "dual" for abelian groups?

Exercise 94. Fix a vector space X. Define the X-dual of a vector space V, V^X, to be the vector space of linear mappings from V to X. Show that a linear mapping φ from V to W gives rise to a linear mapping, φ^X, from W^X to V^X. Does φ a monomorphism imply φ^X an epimorphism? Does $(\varphi \circ \psi)^X =$

$\psi^X \circ \varphi^X$? Is there a monomorphism from V to V^{XX}? When is there an isomorphism?

Exercise 95. Let V be a real vector space. Consider the mapping from $V \oplus V^*$ to $\mathbf{R}$ which associates, with (v,f) in $V \oplus V^*$, the number $f(v)$. Why is this not an element of $(V \oplus V^*)^*$?

14

Multilinear Mappings; Tensor Products

In this chapter we shall introduce two further constructions which yield vector spaces from vector spaces.

Let $V_1, \ldots, V_n, V$ be $(n+1)$ vector spaces (all real or all complex). Let υ be a mapping from the set $V_1 \times V_2 \times \cdots \times V_n$ (Cartesian product of sets) to the set V. Recall that this Cartesian product is just the set of all n-tuples, $(v_1, \ldots, v_n)$, with v_1 in $V_1, \ldots, v_n$ in V_n. Hence υ assigns to each such n-tuple an element, $\upsilon(v_1, \ldots, v_n)$, of V. In other words, υ is just a vector-valued (in V) function of n vector variables (in $V_1, \ldots, V_n$). Such an υ is said to be *multilinear* if it satisfies the following n conditions

$$\upsilon(v_1 + av_1', v_2, \ldots, v_n) = \upsilon(v_1, \ldots, v_n) + a\upsilon(v_1', \ldots, v_n) \ ,$$

$$\upsilon(v_1, v_2 + av_2', \ldots, v_n) = \upsilon(v_1, v_2, \ldots, v_n) + a\upsilon(v_1, v_2', \ldots, v_n) \ ,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\upsilon(v_1, \ldots, v_{n-1}, v_n + av_n') = \upsilon(v_1, \ldots, v_n) + a\upsilon(v_1, \ldots, v_n')$$

for any number a, and for any v_1 and v_1' in $V_1, \ldots, v_n$ and v_n' in V_n. (Thus a multilinear mapping is "linear in each of its variables separately, keeping the others fixed.") For $n = 1$, multilinear mappings are called just plain linear; for $n = 2$, bilinear; for $n = 3$, trilinear, etc. We denote the collection of all such multilinear mappings by $\mathrm{Lin}(V_1, \ldots, V_n; V)$.

Example. $\mathrm{Lin}(V; W)$ is just $\mathrm{Mor}(V, W)$, the collection of all morphisms (in the category of vector spaces, i.e., linear mappings) from vector space V to vector space W.

Example. Let V be a real vector space. Then $\mathrm{Lin}(V;\mathbf{R})$ (where $\mathbf{R}$ is the vector space of real numbers) is V^*, the dual of V. Consider the rule which associates, with (v,f) (v in V and f in V^*), the number $f(v)$. This is an element of $\mathrm{Lin}(V, V^*;\mathbf{R})$.

Example. Consider the rule which associates, with the element (v,w) of $V \times W$, the element (also written) (v,w) of the direct sum, $V \oplus W$. This is not an element of $\mathrm{Lin}(V, W; V \oplus W)$, for the mapping is not bilinear. (For example, (av,w) does not always equal $a(v,w)$ in $V \oplus W$, for the latter is (av,aw).)

We next introduce structure on the set $\mathrm{Lin}(V_1, \ldots, V_n; V)$ so that it, too, becomes a vector space. For υ and υ' in $\mathrm{Lin}(V_1, \ldots, V_n; V)$, define a new element, $\upsilon + \upsilon'$, of $\mathrm{Lin}(V_1, \ldots, V_n; V)$ by its action:

$$(v + v')(v_1, \ldots, v_n) = v(v_1, \ldots, v_n) + v'(v_1, \ldots, v_n) \ .$$

For v in $\text{Lin}(V_1, \ldots, V_n; V)$ and a a number, define a new element, av, of $\text{Lin}(V_1, \ldots, V_n; V)$ by its action:

$$(av)(v_1, \ldots, v_n) = a[v(v_1, \ldots, v_n)] \ .$$

With these operations, the set $\text{Lin}(V_1, \ldots, V_n; V)$ clearly becomes a vector space.

Thus V^* is a vector space, as we have seen. In the category of vector spaces, $\text{Mor}(V, W)$ has the structure of an object in the category.

Since the "Lins" are vector spaces, we can combine them just like other vector spaces.

Example. Fix vector spaces V and W. With each pair (φ, v), with $V \xrightarrow{\varphi} W$ and v in V, associate the element $\varphi(v)$ of W. This is an element of $\text{Lin}(\text{Lin}(V, W), V; W)$.

Example. Fix vector spaces V and W. Associate, with $V \xrightarrow{\varphi} W$, $W^* \xrightarrow{\varphi^*} V^*$. This is an element of $\text{Lin}(\text{Lin}(V, W); \text{Lin}(W^*, V^*))$.

Example. Composition of linear mappings is an element, for example, of $\text{Lin}(\text{Lin}(U, V), \text{Lin}(V, W); \text{Lin}(U, W))$.

Example. Fix vector spaces U, V, W, and fix $U \oplus V \xrightarrow{\varphi} W$. Then, restricting the action of φ to elements of $U \oplus V$ of the form $(u,0)$, we obtain a linear mapping from U to W; similarly, a linear mapping from V to W. Conversely, given $U \xrightarrow{\alpha} W$ and $V \xrightarrow{\beta} W$, the mapping from $U \oplus V$ to W which sends (u,v) to $\alpha(u) + \beta(v)$ is linear. Thus $\text{Lin}(U \oplus V; W) = \text{Lin}(U; W) \oplus \text{Lin}(V; W)$. Similarly, $\text{Lin}(U; V \oplus W) = \text{Lin}(U; V) \oplus \text{Lin}(U; W)$.

The complexities here seem to be getting out of hand. We shall return to this point shortly.

The final method we shall discuss for constructing vector spaces from vector spaces is the tensor product. Let V and W be vector spaces (both real or both complex). A ***tensor product*** of V and W consists of a vector space T, together with a bilinear mapping ς from $V \times W$ to T, such that, if T' is any vector space, and ς' a bilinear mapping from $V \times W$ to T', there is a unique linear mapping γ from vector space T to vector space T' such that the diagram of figure 34 commutes. (Although we have, to simplify the discussion, given the definition of the tensor product of only two vector spaces, there is an obvious generalization to any finite number (and, in fact, an only slightly less obvious generalization to an arbitrary number).) Tensor products are unique in the following sense: if (T, ς) and (T', ς') are two tensor products of V and W, the unique linear mapping γ in the diagram is in fact an isomorphism of vector spaces. (The proof is identical to that (theorem 9) for free groups.)

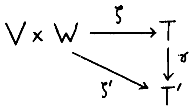

Figure 34

It remains to be shown, however, that tensor products even exist. We now show this. Let V and W be vector spaces. Let X denote the free vector space on the set $V \times W$, and let α be the natural mapping from the set $V \times W$ to the set X (i.e., the mapping which is part of the definition of a free vector space; chapter 9). Thus, for each (v,w) in $V \times W$, there is an element, $\alpha(v,w)$, of X. Every element of X is a linear combination of elements of the form $\alpha(v,w)$, and the only such linear combination which vanishes is the one with zero coefficients. We first note that this α is not a bilinear mapping from $V \times W$ to X, for, for example, $\alpha(v + v',w) - \alpha(v,w) - \alpha(v',w)$ is not the zero element of X, but rather the nonzero element which is that linear combination of the three basis vectors $\alpha(v + v',w)$, $\alpha(v,w)$, and $\alpha(v',w)$. We wish to "force α to be bilinear." To this end, we denote by Y the subspace of X generated by the collection of all elements of X of the forms

$$\alpha(v + av',w) - \alpha(v,w) - a\alpha(v',w) \quad ,$$

$$\alpha(v,w + aw') - \alpha(v,w) - a\alpha(v,w')$$

for all v and v' in V, w and w' in W, and a a number. The elements of Y are precisely "the things which would have to vanish if α were to be bilinear." Denote by β the natural linear mapping $X \xrightarrow{\beta} X/Y$ from X to the quotient space, X/Y. Then $\mathrm{Ker}(\beta) = Y$. We now claim that $\varsigma = \beta \circ \alpha$ (so $V \times W \xrightarrow{\varsigma} X/Y$) is in fact bilinear. Indeed, $\varsigma(v + av',w) - \varsigma(v,w) - a\varsigma(v',w) = \beta[\alpha(v + av',w) - \alpha(v,w) - a\alpha(v',w)]$ (since β is linear). But the expression $\alpha(v + av',w) - \alpha(v,w) - a\alpha(v',w)$ is in Y (that is how we constructed Y) and hence is in $\mathrm{Ker}(\beta)$. Thus β, applied to this element of X, gives zero. We have shown $\varsigma(v + av',w) - \varsigma(v,w) - a\varsigma(v',w) = 0$, and similarly reversing the roles of V and W. That is, we have shown that ς is bilinear. Thus we have the diagram of figure 35.

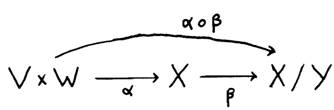

Figure 35

We now claim that this $(X/Y,\varsigma)$ is a tensor product of V and W. Let T be any vector space, and $V \times W \xrightarrow{\xi} T$ any bilinear mapping. Then, since X is the free vector space on the set $V \times W$, there is a unique linear mapping μ from X to T such that the diagram of figure 36 commutes. (This is just the

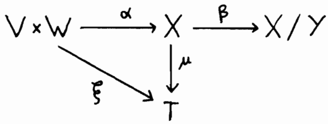

Figure 36

definition of a free vector space—we do not, here, even need bilinearity of ξ.) Now consider any element of X which happens to lie in the subspace Y, for example, the element $\alpha(v + av',w) - \alpha(v,w) - a\alpha(v',w)$. Then, applying μ to this element, we obtain $\mu[\alpha(v + av',w) - \alpha(v,w) - a\alpha(v',w)] = \mu \circ \alpha(v + av',w) - \mu \circ \alpha(v,w) - a\mu \circ \alpha(v',w) = \xi(v + av',w) - \xi(v,w) - a\xi(v',w) = 0$, where, in the first step, we have used linearity of μ, in the second, the fact that the diagram of figure 36 commutes, and, in the third, the fact that ξ is bilinear. That is, every element of X which happens to lie in the subspace Y is annihilated by μ. In other words, μ takes all the elements of a single coset of Y in X to the same element of T. That is, there is a mapping γ from X/Y (the space of cosets of Y in X) to T such that the diagram of figure 37 commutes. This mapping γ is linear (because of the definition of

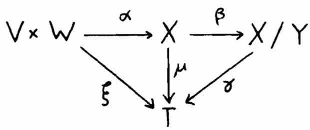

Figure 37

vector-space operations on X/Y) and unique (for, given any element of X/Y, e.g., the coset $x + Y$, we must have $\gamma(x + Y) = \mu(x)$ in order that the diagram commute). Thus every bilinear $V \times W \xrightarrow{\xi} T$ gives rise to a unique μ and γ such that the diagram of figure 37 commutes. It follows that every bilinear $V \times W \xrightarrow{\xi} T$ gives rise to a unique γ such that the diagram of figure 38 commutes, for every γ determines also a μ (namely, $\mu = \gamma \circ \beta$). Thus we have demonstrated precisely the defining property of a tensor product.

It is conventional to write $V \otimes W$ instead of X/Y. We have proven

THEOREM 25. *Let V and W be vector spaces (both real or both complex). Then $(V \otimes W, \varsigma)$ is a tensor product of V and W.*

The following feature of the proof above should be noted. It is, on the one hand, explicit (i.e., we actually construct a tensor product) and, on the other, somewhat abstract (for the substance of the proof is to push arrows around, using, e.g., the universal definition of a free vector space). One can, if one has the inclination, often use the defining (often universal) properties of various things in proofs rather than the more explicit construction.

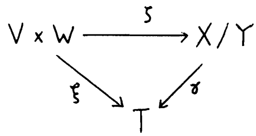

Figure 38

The structure of this tensor product can be seen still more concretely. For (v,w) in $V \times W$, we write, instead of $\varsigma(v,w)$ (for the corresponding element of $V \otimes W$), $v \otimes w$. Then bilinearity of ς becomes $(v + av') \otimes w = v \otimes w + av' \otimes w$ and $v \otimes (w + aw') = v \otimes w + av \otimes w'$. Furthermore, "any linear combination of these $v \otimes w$ (in $V \otimes W$) is equal to some other linear combination of these $v \otimes w$ only when the two linear combinations can be made identical by repeated use of the two bilinearity formulae above." (This is just the intuitive statement of the definition of the subspace Y in the construction above.) Finally, every element of $V \otimes W$ is some (finite) linear combination of various $v \otimes w$ (that is how free vector spaces work). To see in more detail what is going on, we choose bases: let K be a basis for V, and L a basis for W. Then, for k in K and l in L, $k \otimes l$ is certainly an element of $V \otimes W$. Furthermore, every element of $V \otimes W$ of the form $v \otimes w$ is a linear combination of those of the form $k \otimes l$ (for v must be a linear combination of k's, and w a linear combination of l's, so we can expand out $v \otimes w$ as a linear combination of $k \otimes l$'s using bilinearity of "$\otimes$"). Thus the collection of elements of $V \otimes W$ of the form $k \otimes l$ spans $V \otimes W$. Note, furthermore, that this collection of elements of $V \otimes W$ is linearly independent (for no linear combination of distinct $k \otimes l$'s, with nonzero coefficients, could be zero in $V \otimes W$, i.e., could be reduced to zero using bilinearity of "$\otimes$," for the k are linearly independent in V, and the l in W). Thus the subset of $V \otimes W$ consisting of elements of the form $k \otimes l$ (k in K, l in L) is a basis for $V \otimes W$. Note that this subset is isomorphic to the set $K \times L$ (Cartesian product). (Thus one could have defined the tensor product of V and W as the free vector space on $K \times L$, where K is a basis for V, and L a basis for W.)

One can think of the direct sum as being "product-like on the spaces" (e.g., the direct sum of a one-dimensional vector space ("a line") and a one-dimensional vector space is a two-dimensional vector space ("a plane")) and "sum-like on the bases" (for dimensions add under direct sum). From this viewpoint, the tensor product is "ultra-product-like: even product-like on the bases" (for dimensions multiply under tensor product).

There are a number of properties of the tensor product, of which we give a few examples. Let V and W be vector spaces, and fix an element, (v,w), of $V \times W$. We define, using this (v,w), an element of $\mathrm{Lin}(V^*; W)$. Given f in V^*, so f is a linear function on V, associate with it the element $[f(v)]w$ of W (i.e., the element of W which is the product of the number $f(v)$ with the vector w). Thus we have here a mapping ξ from $V \times W$ to $\mathrm{Lin}(V^*; W)$. This mapping is clearly bilinear. Hence (by definition of tensor product) there is a unique linear mapping γ from $V \otimes W$ to $\mathrm{Lin}(V^*; W)$ such that the diagram of

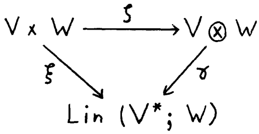

Figure 39

figure 39 commutes. It is easily checked that this $V \otimes W \xrightarrow{\gamma} \mathrm{Lin}(V^*; W)$ is in fact a monomorphism. (For finite dimensions it is actually an isomorphism, but not in general for infinite dimensions.) Thus $V \otimes W$ can be regarded as a subspace of $\mathrm{Lin}(V^*; W)$—similarly, as a subspace of $\mathrm{Lin}(W^*; V)$. Now again fix (v,w) in $V \times W$. For (f,g) in $V^* \times W^*$ (so f is a linear function on V, g a linear function on W), consider the number $[f(v)][g(w)]$ (product of numbers). This is a mapping (for fixed (v,w)) from $V^* \times W^*$ to $\mathbf{R}$ (or $\mathbf{C}$, in the complex case)—a mapping which is certainly bilinear. Hence this is an element (for fixed (v,w)) of $\mathrm{Lin}(V^*, W^*; R)$. We thus have a (bilinear) mapping from $V \times W$ to $\mathrm{Lin}(V^*, W^*; R)$, and so, by the tensor product definition, we have a linear mapping from $V \otimes W$ to $\mathrm{Lin}(V^*, W^*; R)$. This, too, is a monomorphism, so we may also regard $V \otimes W$ as a subspace of $\mathrm{Lin}(V^*, W^*; R)$ (and, if you like, as a subspace of $(V^* \otimes W^*)^*$).

As a final example, let V, U, and W be vector spaces, and consider $V \otimes (U \oplus W)$. A typical element of $U \oplus W$ is (u,w) (u in U, w in W). A typical element of V is v. Given a (u,w) and a v, we obtain an element, $v \otimes (u,w)$, of $V \otimes (U \oplus W)$. Every element of $V \otimes (U \oplus W)$ can be written as a linear combination of elements of this form. Suppose we now associate, with $v \otimes (u,w)$, the element $(v \otimes u, v \otimes w)$ of $(V \otimes U) \oplus (V \otimes W)$. (That is, $v \otimes u$ is in $V \otimes U$, while $v \otimes w$ is in $V \otimes W$, so $(v \otimes u, v \otimes w)$ is in $(V \otimes U) \oplus (V \otimes W)$.) This is a linear mapping from $V \otimes (U \oplus W)$ to $(V \otimes U) \oplus (V \otimes W)$, which is easily checked to be an isomorphism. Hence we can write $V \otimes (U \oplus W) = (V \otimes U) \oplus (V \otimes W)$. ("Multiplication distributes over addition.")

What can one do with vector spaces to get vector spaces? One can take the direct sum (which adds dimensions), the tensor product (which multiplies dimensions), multilinear mappings (which normally increase dimension even faster), and duals (which are a special case of multilinear mappings). (One can also take quotient spaces, but let's ignore that for the present.) Suppose we begin with just one vector space V. Then, using just this vector space, we can construct an enormous collection of vector spaces, for example, $\mathrm{Lin}(V^* \otimes \mathrm{Lin}(V \oplus V^*, V; V^*), V, V^* \otimes \mathrm{Lin}(V, V^*; V \oplus V); V^* \oplus V^*) \otimes \mathrm{Lin}(V; V^*)$. Some of these vector spaces will have "natural preferred elements"; some will be subspaces of others; some will be "naturally isomorphic" to others.

Problem. Organize this situation so that one has an overview of what is going on—when preferred elements, preferred monomorphisms or isomorphisms, etc., exist. Then invent a notation in which moving around from vector space to vector space (e.g., via natural monomorphisms), introduction of preferred elements, etc., can be done effortlessly.

Such an organization is in fact available in the finite-dimensional case; we shall introduce it later. One would like, however, to cover also the infinite-dimensional case in order to have facility with the infinite-dimensional vector spaces (e.g., Hilbert spaces) which arise in applications. Ideally, one would also like to choose a notation such that, when the vector spaces have additional structure (so that, e.g., more things become isomorphic), these additional structural features are incorporated easily into the notation (e.g., by "ignore all primes for topological vector spaces"). Finally, we remark that one can forget about the direct sum, since, because $\mathrm{Lin}(V \oplus W; U) = \mathrm{Lin}(V; U) \oplus \mathrm{Lin}(W; U)$, $\mathrm{Lin}(U; V \oplus W) = \mathrm{Lin}(U; V) \oplus \mathrm{Lin}(U; W)$, and $V \otimes (U \oplus W) = V \otimes U \oplus V \otimes W$, every vector space one can construct using our three techniques can be written as a direct sum of vector spaces constructed without using direct sum.

Exercise 96. State and prove: $U \otimes (V \otimes W) = (U \otimes V) \otimes W$.

Exercise 97. Just prior to the "problem" above, there appears a complicated expression for a vector space constructed from V. Write out the isomorphic vector space which is a direct sum of vector spaces, each constructed from V without using direct sum.

Exercise 98. Construct a natural isomorphism from $\mathrm{Lin}(V_1, \ldots, V_n; V)$ to $\mathrm{Lin}(V_1 \otimes \cdots \otimes V_n; V)$. (Thus tensor products allow one to consider only spaces of linear mappings rather than of multilinear mappings.)

Exercise 99. Consider the mapping from $V \otimes V^*$ to $\mathrm{Lin}(V; V)$ which sends $v \otimes f$ (v in V, f in V^*) to the linear mapping from V to V which sends v' in V to $[f(v')]v$. Prove that this is a linear mapping and that it is a monomorphism. Prove that it is an isomorphism if and only if V is finite-dimensional.

Exercise 100. Which of $(V \otimes W)^*$, $V^* \otimes W^*$ can be regarded as a subspace of the other?

Exercise 101. Find a preferred element of $(V \otimes V^*)^*$. (In matrix algebra, this element is called the trace.)

Exercise 102. Prove that every vector space is isomorphic to some vector space of the form $V \otimes W$.

Exercise 103. Is there any natural way in which V can be regarded as a subspace of $V \otimes W$?

Exercise 104. Consider, on the collection of all vector spaces, the rule which associates, with any two, their tensor product. Does this collection thus become a group?

15

Example: Minkowski Vector Space

We now give an example of the application of some of these notions to the description of a physical situation.

Let V be a fixed real, four-dimensional vector space. Let g be an element of Lin($V,V;\mathbf{R}$), so, for vectors v and v', $g(v,v')$ is a real number. Let this g by symmetric in the following sense: for any vectors v and v', $g(v,v') = g(v',v)$. Call a vector v (in V) timelike if $g(v,v)$ is negative, null if $g(v,v)$ is zero, and spacelike if $g(v,v)$ is positive. Let us suppose, finally, that the following (signature) condition is satisfied: there exist timelike vectors, and, for any timelike vector t and nonzero vector v, with $g(t,v) = 0$, v is necessarily spacelike. Such a (V,g) will be called Minkowski vector space.

We now describe the physical setup that a Minkowski vector space is intended to represent. By an event, we shall mean an occurrence in the physical world having extension in neither space nor time. Thus the snapping of one's fingers or the explosion of a firecracker would represent an event. Now fix, once and for all, a particular reference event 0. The vector space V represents the collection of all "nearby events" to the event 0. The event 0 itself is represented by the zero vector in V. (Think of v in V as being the "displacement vector" from 0 to the event described by v.) The multilinear mapping g describes, within the mathematics, the results of the following thought experiment (figure 40). Fix an event, described by vector v, and let the event 0 be the snapping of fingers by some observer. Now let our observer send out a light signal (a pulse of light) at just the right time, and in just the right direction, so that the light signal arrives at the event v when this event occurs. Let another light signal be sent, at the occurrence of v, in just the right direction so that it returns to our observer. Denote by p and p' the events "the sending out of the light signal" and "the reception of the return signal" (by our observer), respectively. Denote by t (a real number) the elapsed time, according to our observer, between the event p (when our observer sent out the light) and the event 0; similarly, denote by t' the elapsed time between event 0 and event p'. Then $g(v,v)$ is supposed to be the real number tt' (product of numbers).

In order to get a feeling for what this $g(v,v)$ means physically, we consider some special cases. Suppose $t = t'$. Then one would say "it took light t s (before 0) to get to the event, and t' ($= t$) s after 0 to get back, so the event v occurred at the same time as 0." Furthermore, one would regard the product tt' as the "square of the distance (in units of light-seconds) of the

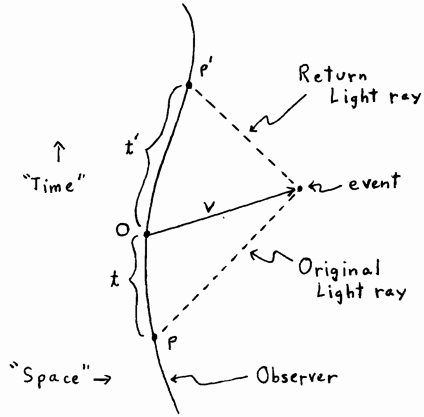

Figure 40

event v from the event 0." If t' were a little larger than t, then one would say that "the event v occurred shortly after the event 0." If t were zero (so $g(v,v) = 0$), that would correspond to the situation in which the light signal,

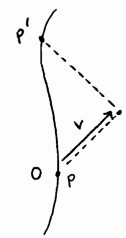

Figure 41

so timed to reach v, would have to be sent out at event 0 itself (figure 41). One would say that "the elapsed time between 0 and v, and the spatial displacement of v from 0, have been so adjusted that light just makes it from 0 to v." As t continues getting smaller, that is, goes negative, one regards "the event v as occurring later and later in time." The limiting case $t = -t'$ (figure 42) corresponds to the physical situation "the return light signal arrives essentially the moment the original signal is sent out," that is, to the situation "the event v occurs in the immediate presence of the observer, but later than 0 by an elapsed time t'." In this last case, $g(v,v) = tt'$ is negative.

To summarize, our observer regards event v as having occurred "a distance $(1/2)(t + t')$ from 0 and a time $(1/2)(t' - t)$ later than 0." But $g(v,v) = tt' = [(1/2)(t' + t)]^2 - [(1/2)(t' - t)]^2$, the difference between the square of

Figure 42

the "spatial distance" and the "elapsed time." Thus timelike vectors are those v whose corresponding events have an "elapsed time" (in seconds) from 0 which exceeds the "spatial distance" (in light-seconds) from 0. Vice versa for spacelike vectors.

So that is what $g(v,v)$ means physically. What does $g(v,v')$ (with $v \neq v'$) mean? Here is where multilinearity and symmetry come in. We have $g(v+v',v+v') - g(v-v',v-v') = [g(v,v) + g(v,v') + g(v',v) + g(v',v')] - [g(v,v) - g(v,v') - g(v',v) + g(v',v')] = 4g(v,v')$. The quantity on the left has already been interpreted physically, while $g(v,v')$ appears on the right.

We can now reintroduce our observer into the formalism. Denote by v the event which occurs in the immediate presence of the observer but 1 s later than 0. Then, for this v, we have $t = -1$ and $t' = 1$, so $g(v,v) = -1$. This v "describes the observer within the mathematics." Consider, for example, another event, $\tilde{v}$, with times $\tilde{t}$ and $\tilde{t}'$. Let's evaluate $g(v,\tilde{v})$. The event $v + \tilde{v}$ has times $-1 + \tilde{t}$ and $1 + \tilde{t}'$, while the event $v - \tilde{v}$ has times $-1 - \tilde{t}$ and $1 - \tilde{t}'$. Hence $g(v+\tilde{v},v+\tilde{v}) - g(v-\tilde{v},v-\tilde{v}) = [(-1+\tilde{t})(1+\tilde{t}')] - [(-1-\tilde{t})(1-\tilde{t}')] = 2(\tilde{t} - \tilde{t}')$. Thus $g(\tilde{v},v) = -(1/2)(\tilde{t}' - \tilde{t})$, so $g(v,\tilde{v})$ is minus the quantity we interpreted as the "elapsed time between the event $\tilde{v}$ and the basic event 0." In particular, $g(v,\tilde{v}) = 0$ means that "our observer thinks that the event $\tilde{v}$ has a purely spatial (no temporal) displacement from 0." Thus the results of various physical observations of our observer can be expressed by replacing our observer with a certain vector v and writing expressions (to represent those observations) involving this v.

The signature condition, it should now be clear, merely represents the incorporation into the mathematics of the physical observations described in the preceding paragraph.

It is intended that the discussion above make the point that all the structure of (V,g) has physical meaning, that anything one says within this structure can be interpreted as a physical statement, that any theorem about this structure is a physical prediction, etc. There are no "irrelevant things" around. We emphasize, however, that a Minkowski vector space is to be nothing more and nothing less than its mathematical definition above. One must prove things within that framework, providing physical interpretations where they are interesting or useful. (The idea is to use one's physical insight in the interaction between the mathematics and the physics, i.e., in the definition of a Minkowski vector space, and the following discussion.) In particular, the use of the two "times," t and t', just gives the flavor of what is happening physically and can now be dispensed with.

We give a few examples of this viewpoint.

1. Let t be a timelike vector, with $g(t,t) = -1$ (this vector t not to be confused with the number t above), and denote by $t^{\perp}$ the collection of all vectors v with $g(t,v) = 0$. Then this $t^{\perp}$ is a subspace of V and is complementary to the subspace generated by t. [Proof: Since $g(t,v + av') = g(t,v) + ag(t,v')$, and since the right side vanishes if v and v' are in $t^{\perp}$, this $t^{\perp}$ is certainly a subspace. For any v in V, we have $v = [v + tg(t,v)] - tg(t,v)$. But the first vector is in $t^{\perp}$ (for $g(t,v + tg(t,v)) = g(t,v) + g(t,t)g(t,v) = g(t,v) - g(t,v) = 0$), and the second is in the subspace generated by t. Hence every vector is a sum of vectors, one from each subspace. Finally, let $at + s = 0$, where a is a number and s is in $t^{\perp}$. Then $0 = g(t,at + s) = ag(t,t) + g(t,s) = -a$. Thus $a = 0$, whence $s = 0$. This establishes that the subspaces are complementary.]

Physically, think of t as describing an observer. Then vectors v in $t^{\perp}$ represent, as we have seen, "pure spatial displacements" (according to this observer), while vectors in the subspace generated by t represent "pure temporal displacements." The statement above thus asserts, physically, that "every displacement can be decomposed uniquely into its spatial part and its temporal part, according to any observer." One can now understand physically what addition of vectors means: one "adds the spatial parts, and adds the temporal parts, separately." Finally, let $v = at + s$, with s in $t^{\perp}$. Then $g(v,v) = g(at + s, at + s) = a^2 g(t,t) + 2ag(t,s) + g(s,s) = g(s,s) - a^2$. This little calculation reflects, within the formalism, the formula in the fifth paragraph of this chapter. It also shows that $g(s,s)$ should be interpreted as the "square of the magnitude of the spatial displacement represented by s."

2. For t and t' timelike vectors, write $t \approx t'$ if $g(t,t')$ is negative. Then this is an equivalence relation on the collection of all timelike vectors and there are precisely two equivalence classes. [Proof: For any timelike vector t, $t \approx t$, for $g(t,t)$ is negative. If $t \approx t'$, then $t' \approx t$, for $g(t,t') = g(t',t)$. Finally, let $t \approx t'$ and $t \approx t''$, so $g(t,t')$ and $g(t,t'')$ are both negative. Then, since $g(t,t' - at'') = g(t,t') - ag(t,t'')$, there is a positive number a such that $g(t,t' - at'')$ vanishes. But $g(t' - at'', t' - at'') = g(t',t') - 2ag(t',t'') + a^2 g(t'',t'')$. If $g(t',t'')$ were nonnegative, this right side would be negative, whence $t' - at''$ would be a timelike vector which (by construction) would satisfy $g(t,t' - at'') = 0$. This would violate the signature condition. Hence $\approx$ is an equivalence relation. There are just two

equivalence classes since, for $g(t,t')$ and $g(t,t'')$ positive, $g(t',t'')$ is negative: two vectors not in the equivalence class of t are in the same equivalence class.]

This result presents us with the possibilities for the distinction between past and future. Pick an equivalence class, and call its elements "future-directed timelike vectors," those in the other equivalence class "past-directed." Note that the mathematics does not tell us which class to pick; it only tells us what the possibilities are. This is a physical distinction which must, here, be incorporated into the formalism.

3. Let t and t' be timelike vectors in the same equivalence class, and let n be a nonzero null vector. Then the numbers $g(t,n)$ and $g(t',n)$ are both positive or both negative. [Proof: If one were positive and the other negative, there would be a positive number a with $g(t + at',n) = 0$. But $g(t + at',t + at') = g(t,t) + 2ag(t,t') + a^2g(t',t')$ is negative, since t and t' are in the same equivalence class. Thus $t + at'$ is timelike, and n nonzero and null, with $g(t + at',n) = 0$, violating the signature condition.]

This result allows us to extend the notion "future- and past-directed" to null vectors: call null n future-directed if $g(t,n)$ is negative for every future-directed timelike t, and past-directed otherwise. Physically, future-directed null vectors represent events which "can be reached from 0 by a light signal from 0," and past-directed events "from which 0 can be reached by a light signal." One expects physically that a future-past distinction should be available for null vectors—and indeed it is. Figure 43 is useful for keeping these

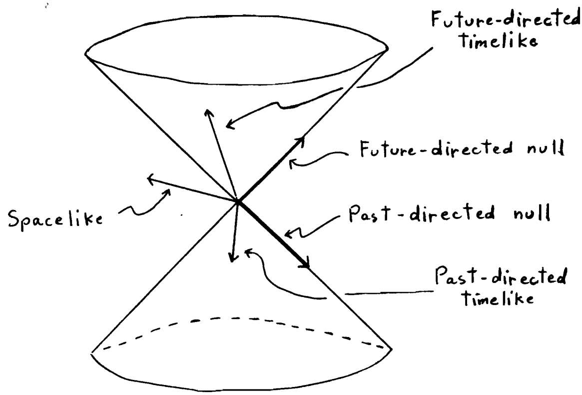

Figure 43

properties straight, suggesting new properties, and suggesting proofs. One thinks of "increasing time going upward in the figure." Null vectors lie on the cones, timelike inside, and spacelike outside. (Warning: One dimension is, of necessity, suppressed in the figure, so it can on rare occasions be misleading.)

4. Let n be a nonzero null vector, and v some vector with $g(n,v) = 0$. Then either v is spacelike or v is a numerical multiple of n (hence null). [Proof: Certainly v cannot be timelike, for that would violate the signature condition. If v were spacelike, we would be done. So let v be null. Then, for every real number a, $v - an$ is null (since v and n are null, and $g(n,v) = 0$). Choose any timelike t, and consider $g(t,v - an) = g(t,v) - ag(t,n)$. Since $g(t,n)$ cannot vanish (signature condition), there is an a such that this expression, $g(t,v - an)$, vanishes. Such a null $v - an$ would violate the signature condition unless $v - an = 0$, whence v is a numerical multiple of n.]

Thus, for a null n, $n^{\perp}$ is not a complementary subspace to that generated by n.

5. The sum of two future-directed timelike vectors is future-directed timelike; the sum of two future-directed null vectors (neither a numerical multiple of the other) is future-directed timelike. [Proof: For t and t' future-directed timelike (so $g(t,t')$ is negative), $g(t + t',t + t') = g(t,t) + 2g(t,t') + g(t',t')$ is negative, so the sum is timelike. Since $g(t,t + t') = g(t,t) + g(t,t')$ is negative, the sum is future-directed timelike. Let n and n' be future-directed null, neither a numerical multiple of the other. Fix a future-directed timelike t, and consider $g(t,n - an') = g(t,n) - ag(t,n')$. Since $g(t,n)$ and $g(t,n')$ are both negative, some positive a makes the right side above vanish. By the signature condition, this $n - an'$ must be spacelike, whence $g(n - an',n - an') = -2ag(n,n')$ must be positive. Hence $g(n,n')$ must be negative. Hence $g(n + n',n + n') = 2g(n,n')$ is negative, whence $n + n'$ is timelike. It is future-directed, since $g(t,n + n') = g(t,n) + g(t,n')$ is negative.]

These examples are intended to illustrate the point that there are numerous geometrical properties of Minkowski vector space and that all are easily derived from the definitions. (In practice, of course, a hundred or so such properties become incorporated into one's repertoire, to be used, without derivation, when needed.)

We next give some examples of more physical calculations within a Minkowski vector space.

Fix, once and for all, an observer, represented by a future-directed timelike vector t with $g(t,t) = -1$. Now consider a second observer, who is "moving by, but passes our original observer just at the event 0." This observer could be represented by a future-directed timelike t', with $g(t',t') = -1$. By property 1 above, we can write $t' = at + s$, with s in $t^{\perp}$. This t' represents "an event occurring in the presence of the second observer and 1 sec later (according to the second observer) than 0." According to the fundamental observer (figure 44), this event has a "spatial displacement" s from 0 and a "temporal displacement" a. Hence the vector $a^{-1}s$ would be interpreted by the fundamental observer as the "velocity of the second observer as that second observer passes by": write $v = a^{-1}s$. Then $g(v,v)$ is to be interpreted

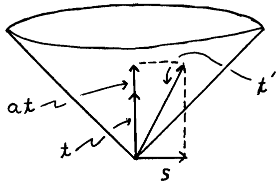

Figure 44

as the "speed squared of the second observer, as seen by the fundamental one." (We shall write v^2 for $g(v,v)$.) Let's evaluate this v^2. We have $v^2 = g(a^{-1}s, a^{-1}s) = a^{-2}g(s,s)$. But $-1 = g(t',t') = g(at + s, at + s) = -a^2 + g(s,s)$, while $g(t,t') = g(t, at + s) = ag(t,t) + g(t,s) = -a$. Thus $v^2 = [g(t,t')^2 - 1]/g(t,t')^2$. Inverting, $g(t,t') = -(1 - v^2)^{-1/2}$. These formulae show, in particular that "the speed with which the t-observer sees the t'-observer go by is the same as the speed with which the t'-observer sees the t-observer go by." Note, for example, that the familiar time-dilation formula, $a = (1 - v^2)^{-1/2}$, is immediate.

Let us now consider the situation in which our fixed observer (described by vector t) sees two objects go by. Let future-directed τ and τ' describe these objects, with $g(\tau,\tau) = g(\tau',\tau') = -1$. Let us suppose, further, that the objects are seen to "go in the same direction" (by our observer), that is, that t, τ, and τ' are linearly dependent. We first decompose these vectors according to our observer: $\tau = \alpha t + \sigma$ and $\tau' = \alpha' t + \sigma'$, with σ and σ' in $t^\perp$. Then $g(\tau,\tau) = g(\tau',\tau') = -1$ imply $g(\sigma,\sigma) = \alpha^2 - 1$ and $g(\sigma',\sigma') = \alpha'^2 - 1$. Now consider $g(\tau,\tau') = g(\alpha t + \sigma, \alpha' t + \sigma') = -\alpha\alpha' + g(\sigma,\sigma')$. But σ and σ' are in $t^\perp$, while t, σ, and σ' are linearly dependent: hence σ' is a multiple of σ. Therefore $g(\sigma,\sigma') = -[g(\sigma,\sigma)g(\sigma',\sigma')]^{1/2} = -[(\alpha^2 - 1)(\alpha'^2 - 1)]^{1/2}$. Thus we have $g(\tau,\tau') = -\alpha\alpha' - [(\alpha^2 - 1)(\alpha'^2 - 1)]^{1/2}$. This formula provides a relationship between velocities. We have $\alpha = (1 - v^2)^{-1/2}$, where v is the speed that our observer sees the τ-object go by, and, similarly, $\alpha' = (1 - v'^2)^{-1/2}$. Substituting, $g(\tau,\tau') = -(1 + vv')[(1 - v^2)(1 - v'^2)]^{-1/2}$. But $g(\tau,\tau') = -(1 - V^2)^{-1/2}$, where V is the relative speed of the two objects. Hence $V = (v + v')/(1 + vv')$, the familiar formula for addition of velocities.

Now again fix our fundamental observer (described by t), and consider a particle going by. This particle is described by a certain future-directed timelike vector p, where, setting $p = m\tau$, τ represents the "1 s later event of an observer following the particle," so $g(\tau,\tau) = -1$, and m the rest mass of the particle. Thus $g(p,p) = -m^2$. Decompose as usual: $p = Et + P$, where E is a number (the energy of the particle, as seen by our observer) and P is in $t^\perp$ (its "spatial momentum"). Thus $E = -g(t,p) = -mg(t,\tau) = m(1 - v^2)^{-1/2}$, where v is the speed of the particle as seen by our observer. Furthermore, $m^2 =$

$-g(p,p) = -g(Et + P, Et + P) = E^2 - g(P,P)$. That is, $E^2 = m^2 + g(P,P)$, the usual energy-momentum formula.

Finally, consider a light ray going by. We would expect it to be described by a future-directed null vector n. As usual, decompose according to our fundamental observer: $n = \omega t + \kappa$, where ω is a number (the angular frequency of the light) and κ is a vector in $t^{\perp}$ (the wavenumber of the light beam). We have $0 = g(n,n) = g(\omega t + \kappa, \omega t + \kappa) = -\omega^2 + g(\kappa,\kappa)$, so $g(\kappa,\kappa) = \omega^2$, the usual formula relating frequency and wavelength for a light ray. Now consider a second observer, described by $t' = at + s$, with s in $t^{\perp}$. This second observer looks at the same light beam and says that its frequency is $\omega' = -g(t',n)$ (for, above, $\omega = -g(t,n)$). Thus $\omega' = -g(t',n) = -g(at + s, \omega t + \kappa) = a\omega - g(s,\kappa)$. This is the Doppler shift formula. For example, if $g(s,\kappa) = 0$ (physically, the fundamental observer sees the second observer and the light ray going off in orthogonal directions), then $\omega' = a\omega = \omega(1 - v^2)^{-1/2}$, where v is the relative speeds of the two observers. This is the formula for the transverse Doppler shift.

Other standard formulae are obtained with similar ease.

We emphasize the following points in connection with the discussion above.

1. One uses very little mathematics, and certainly nothing sophisticated. One does, however, need an overview of vector spaces, multilinear mappings, etc., and the flavor of what can be done therein.

2. One takes care at the beginning to be sure that "the mathematics is appropriate for the physics," that is, that everything in the mathematics has physical meaning and that all of the physics one wishes to talk about is describable in terms of the mathematics. In particular, one tries to avoid structural features (e.g., a basis for the vector space) which have no physical significance.

3. One allows physical things to be described by the mathematics which naturally describes them. For example, one does not regard a vector or a multilinear mapping as "less real," "less physical," or "less explicit" than a number or a function of one variable.

4. To answer a physical question, one first translates that question into various objects (including only objects which are relevant to the question) in the mathematics, with various properties describing the physical setup. Then one manipulates these objects within the mathematics and translates results back into physical terms. (These "translations" ultimately become automatic.)

16

Example: The Lorentz Group

Let V,g be a Minkowski vector space. A Lorentz transformation on V,g is an isomorphism φ from vector space V to V which satisfies the following condition: for any v and v' in V, $g(\varphi(v),\varphi(v')) = g(v,v')$. Thus a Lorentz transformation "preserves all the structure of the Minkowski vector space." In particular, if φ is a Lorentz transformation and v is timelike (resp., null, spacelike), then $\varphi(v)$ is also timelike (resp., null, spacelike). Denote by L the collection of all Lorentz transformations on V,g. Let φ and ψ be in L, and consider $\psi \circ \varphi$, a linear mapping from V to V. This $\psi \circ \varphi$ is certainly an isomorphism from V to V (for composition of isomorphisms is an isomorphism) and is, in fact, a Lorentz transformation, for $g(\psi \circ \varphi(v),\psi \circ \varphi(v')) = g(\psi[\varphi(v)],\psi[\varphi(v')]) = g(\varphi(v),\varphi(v')) = g(v,v')$. With this product rule on the set L, it becomes a group. (The identity is the identity isomorphism on V; the inverse of the Lorentz transformation φ is the $V \xrightarrow{\lambda} V$ such that $\varphi \circ \lambda = \lambda \circ \varphi = i_V$ (the existence of which is guaranteed by the fact that φ must be an isomorphism from V to V). One easily checks that this λ is in fact a Lorentz transformation.) This L is called the Lorentz group (on V,g).

Thus L is a subgroup of the group of all isomorphisms from vector space V to V (but not, e.g., a subspace of Lin(V;V)).

We next give some explicit examples of Lorentz transformations. Let t be a future-directed timelike vector, with $g(t,t) = -1$. Denote by R_t the collection of all Lorentz transformations φ such that $\varphi(t) = t$. (Physically, these represent "spatial rotations as seen by the observer described by t.") Then, for φ in R_t, and s in $t^\perp$, we have that $\varphi(s)$ is also in $t^\perp$, for $g(\varphi(s),t) = g(\varphi(s),\varphi(t)) = g(s,t) = 0$. Next, note that the inverse of an element of R_t is in R_t and that the composition of two elements of R_t is in R_t. That is, R_t is a subgroup of L (called the subgroup of spatial rotations defined by t). Consider, in particular, the element σ of R_t with the following action: for v in V, write $v = at + s$, with a a real number and s in $t^\perp$; then set $\sigma(v) = at - s$. (This is indeed a Lorentz transformation: $g(\sigma(v),\sigma(v')) = g(\sigma(at + s),\sigma(a't + s')) = g(at - s, a't - s') = -aa' + g(s,s') = g(at + s, a't + s')$.) This σ is called the spatial reflection defined by t. Similarly, the Lorentz transformation τ given by $\tau(at + s) = -at + s$ (a a real number, s in $t^\perp$) is called the temporal reflection defined by t. (Note that τ is not in the subgroup R_t.) Of course, for $t \neq t'$, the spatial and temporal reflections defined by t will not be the same as those defined by t' (although the product of the two for t, the mapping which sends v in V to $-v$, is the same as the corresponding product for t').

To obtain other Lorentz transformations, we proceed as follows. Let P and Q be complementary subspaces of V, each two-dimensional, and such that, for p in P and q in Q, $g(p,q) = 0$. Suppose, furthermore, that P contains both timelike and spacelike vectors, while every nonzero vector in Q is spacelike. (It is easy to construct such. Choose timelike t and spacelike s, and let P be the subspace generated by the subset consisting of t and s. Then let Q be the subspace of all q with $g(t,q) = g(s,q) = 0$.) A Lorentz transformation φ such that $\varphi(p) = p$ for each p in P (whence $\varphi(q)$ is in Q for each q in Q) is called a rotation in Q, or a rotation about the axis P. (Note that, in four dimensions, rotations have a two-dimensional axis.) Similarly, a Lorentz transformation φ such that $\varphi(q) = q$ for each q in Q (whence $\varphi(p)$ is in P for each p in P) is called a boost in P, or a boost about the axis Q. To interpret these physically, choose future-directed timelike t in P, with $g(t,t) = -1$. Find spacelike s in P with $g(t,s) = 0$. Then a rotation in Q, according to this observer, is "a spatial rotation about the spatial axis s." A boost in P, according to this observer, "gives every object a certain increment of velocity in the s-direction."

Write R_P (resp., R_Q) for the collection of all rotations (resp., boosts) about the axis P (resp., Q). (To construct an element of R_P, choose q and q' in Q with $g(q,q) = g(q',q') = 1$, and $g(q,q') = 0$. Then, for $v = aq + a'q' + p$, with p in P, set $\varphi(v) = (a \cos b + a' \sin b)q + (-a \sin b + a' \cos b)q' + p$, where b is any real number. Similarly for R_Q.) The subgroup of L generated by the union of R_P and R_Q is that consisting of φ in L such that $\varphi(p)$ is in P and $\varphi(q)$ in Q, for any p in P and q in Q. But note that every such φ can be written uniquely in the form $\varphi = \alpha \circ \beta$ $(= \beta \circ \alpha)$ with α in R_P and β in R_Q. Thus the subgroup generated by the union of R_P and R_Q is isomorphic to the direct product of groups R_P and R_Q. (Of course, each of R_P and R_Q is itself a subgroup of the Lorentz group L.)

Denote by K the group with just two elements, e (the identity) and k (so $ee = kk = e$, $ek = ke = k$). We define a homomorphism μ from the Lorentz group L to K. For φ a Lorentz transformation, let $\mu(\varphi)$ be e if φ takes future-directed timelike vectors to future-directed timelike vectors, and k if φ takes future-directed to past-directed. (Note that a Lorentz transformation necessarily takes either all future-directed timelike vectors to future-directed, or all to past-directed.) Thus, for φ the temporal reflection defined by t, $\mu(\varphi) = k$, while, for φ the spatial reflection defined by t, $\mu(\varphi) = e$. Thus this homomorphism μ is in fact an epimorphism of groups. The kernel of μ, Ker(μ), is a normal subgroup of L (the Lorentz transformations which do not reverse "time sense"). Since μ is an epimorphism, there are precisely two cosets of Ker(μ) in L. Fix τ, the temporal reflection defined by t. Then every element of the other coset (i.e., besides Ker(μ) itself) can be written uniquely in the form $\tau \circ \alpha$, with α in Ker(μ). Note, however, that there is no "preferred element" of this coset.

An analogous construction is available for spatial reflections. Denote by X the subspace of the 256-dimensional vector space Lin(V, V, V, V;**R**) consisting of ϵ in this Lin(V, V, V, V;**R**) which "reverse sign under interchange of any two vectors," that is, which satisfy $\epsilon(v,v',v'',v''') = -\epsilon(v,v''',v'',v')$, and similarly for all other interchanges. This X is in fact a one-dimensional vector space, for, fixing a basis v_1, v_2, v_3, v_4 for V, ϵ in X is completely and uniquely determined by the value of the number $\epsilon(v_1,v_2,v_3,v_4)$. We define a homomorphism γ from the Lorentz group L to the group G of all isomorphisms on X. For φ a Lorentz transformation, associate with it the isomorphism on X which sends ϵ in X to the multilinear mapping $\epsilon(\varphi(v),\varphi(v'),\varphi(v''),\varphi(v'''))$ from $V \times V \times V \times V$ to **R**. Next, note that, since X is one-dimensional, the only isomorphisms on X are those which send x in X to ax, with a a nonzero real number. Hence we have a homomorphism δ from G (the group of isomorphisms on X) to K, where δ sends the isomorphism "x goes to ax" to e if a is positive, and to k if a is negative.

Thus we have $L \xrightarrow{\gamma} G \xrightarrow{\delta} K$, homomorphisms of groups. Set $\nu = \delta \circ \gamma$, a homomorphism from the Lorentz group L to the group K (with just two elements). Clearly, if φ is either a temporal or a spatial reflection, then $\nu(\varphi) = k$. For φ the Lorentz transformation which sends v in V to $-v$, we have $\nu(\varphi) = e$. Intuitively, "ν sends Lorentz transformation φ to e if neither a spatial reflection nor a temporal reflection is involved, or if both a spatial reflection and a temporal reflection are involved, and ν sends φ to k otherwise."

The subgroup $\underline{L}$ of L consisting of all Lorentz transformations φ such that $\mu(\varphi) = e$ and $\nu(\varphi) = e$ is called the proper Lorentz (sub)group. Clearly, $\underline{L}$ (since it is an intersection of two normal subgroups) is a normal subgroup of L. Elements of $\underline{L}$ are "those Lorentz transformations which reverse neither spatial nor temporal sense." There are precisely four cosets of $\underline{L}$ in L. (The cosets can be thought of as "Lorentz transformations which reverse spatial and not temporal sense," "Lorentz transformations which reverse temporal and not spatial sense," "Lorentz transformations which reverse both," and "Lorentz transformations which reverse neither.")

17

Functors

The notion of a category can, as we have seen, help one organize one's thinking in mathematics. There is a second, closely related notion—that of a functor. As is the case with categories, there is essentially only one use we shall make of functors: to call them functors when they arise. (We have postponed until this point the discussion of functors in order to accumulate first a few examples.)

Let **C** and **C**′ be categories (so **C** consists of objects, morphisms, and a composition rule, subject to associativity of composition and the existence of identities, and similarly for **C**′). A *covariant functor* **F** from category **C** to category **C**′ consists of two things—i) a rule which associates, with each object A of category **C**, an object, written $\mathbf{F}(A)$, of category **C**′, and ii) a rule which associates, with each morphism φ from object A to object B in category **C**, a morphism, written $\mathbf{F}(\varphi)$, from object $\mathbf{F}(A)$ to object $\mathbf{F}(B)$ in category **C**′—subject to the following two conditions:

1. Composition is preserved. For $A \xrightarrow{\varphi} B \xrightarrow{\psi} C$ a diagram in category **C**, we have

$$\mathbf{F}(\psi \circ \varphi) = \mathbf{F}(\psi) \circ \mathbf{F}(\varphi) \quad .$$

(Note that this is meaningful, for each side of the equation is a morphism, in the category **C**′, from object $\mathbf{F}(A)$ to object $\mathbf{F}(C)$ of that category.)

2. Identities are preserved. For A any object in category **C**, we have

$$\mathbf{F}(i_A) = i_{\mathbf{F}(A)} \quad .$$

(Note that this is meaningful, for each side is a morphism, in **C**′, from object $\mathbf{F}(A)$ to object $\mathbf{F}(A)$ therein.) Thus a functor is a "structure-preserving mapping from one category to another."

As one might expect, the key to the definition is examples.

Example. Let **C** be the category of groups, and **C**′ the category of sets. For G a group (an object of category **C**), let $\mathbf{F}(G)$ be its underlying set (an object of category **C**′). For $G \xrightarrow{\varphi} H$ a homomorphism of groups (a morphism of category **C**), let $\mathbf{F}(G) \xrightarrow{\mathbf{F}\varphi} \mathbf{F}(H)$ be the corresponding mapping from the set G to the set H (a morphism of category **C**′). The first property above holds by definition of composition of homomorphisms (compose the corresponding mappings of sets). The second property above holds since the identity homomorphism from a group to itself is just the identity mapping from the corresponding underlying set to itself. Thus we have a covariant functor from the category of groups to the category of sets. (This functor "just ignores the group structure.")

Example. Let **C** be the category of vector spaces, and **C**′ the category of abelian groups. For V a vector space, let $\mathbf{F}(V)$ be its underlying abelian group. For $V \xrightarrow{\varphi} W$ a linear mapping of vector spaces, let $\mathbf{F}(V) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(W)$ be the corresponding homomorphism of abelian groups. (We have seen that this is indeed a homomorphism in chapter 9.) This is a covariant functor from the category of vector spaces to the category of abelian groups. ("Forget how to multiply vectors by numbers, but do not forget how to add vectors.")

In each case, the covariant functor "merely ignores some (or all) of the structure which is available, so the functor is from a category to a category whose structure is less rich." Such functors are often called "forgetful functors." Similarly, there is the forgetful functor from the category of vector spaces to the category of sets. The first construction of chapter 12 is of a forgetful functor from the category of complex vector spaces to the category of real vector spaces.

We have already encountered examples of more subtle functors.

Example. Let **C** be the category of sets, and **C**′ the category of groups. For S a set, let $\mathbf{F}(S)$ be the free group on set S. (Thus this **F** associates an object in **C** with each object in **C**′.) Next, let $S \xrightarrow{\varphi} S'$ be a mapping of sets, and consider the diagram of figure 45. Here, μ is the mapping (part of the

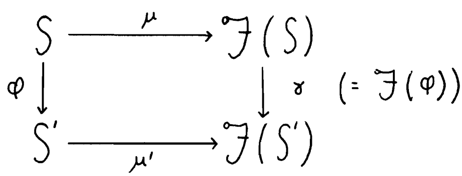

Figure 45

definition of a free group) from the set S to the underlying set of the free group $\mathbf{F}(S)$, and similarly for μ'. By the definition of a free group, there is a unique homomorphism γ of groups which makes the diagram commute. In this way, with $S \xrightarrow{\varphi} S'$ (morphism in category **C**), we associate $\mathbf{F}(S) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(S')$ (set $\mathbf{F}(\varphi) = \gamma$; morphism in category **C**′). Properties 1 and 2 are easily checked. Thus we obtain a covariant functor from the category of sets to the category of groups.

Example. Let **C** be the category of real vector spaces, and **C**′ the category of complex vector spaces. For W a real vector space, let $\mathbf{F}(W)$ be the complex vector space constructed in chapter 12. For $V \xrightarrow{\varphi} W$ a linear mapping of real vector spaces, let $V \oplus V \xrightarrow{\tilde{\varphi}} W \oplus W$ be given by $\tilde{\varphi}(v,v') = (\varphi(v),\varphi(v'))$. This mapping is in fact a linear mapping, $\mathbf{F}(V) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(W)$, of the corresponding complex vector spaces. Properties 1 and 2 are immediate. Hence we have a covariant functor from the category of real vector spaces to the category of complex vector spaces.

One normally thinks of a functor, as in these last two examples, as being a "construction" (or, in the case of a forgetful functor, as a "destruction"). An intermediate example is the covariant functor from the category of abelian groups to the category of groups which associates, with abelian group G, the group G. (An abelian group is nonetheless a group.) "Take the free vector space" is a covariant functor from the category of sets to the category of vector spaces.

We now consider an example of something which is nearly, but not quite, a covariant functor. Let $\mathbf{C}$ and $\mathbf{C}'$ each be the category of (say, real) vector spaces. For V a real vector space, let $\mathbf{F}(V)$ be the real vector space V^* (the dual of V, i.e., Lin $(V;\mathbf{R})$). Now let $V \xrightarrow{\varphi} W$ be a linear mapping of real vector spaces. If we are going to obtain a covariant functor, we shall have to say what $\mathbf{F}(\varphi)$ is supposed to be: that is, we must specify a certain linear mapping $\mathbf{F}(\varphi)$ from V^* $(= \mathbf{F}(V))$ to W^* $(= \mathbf{F}(W))$. Unfortunately, there does not seem to be any natural linear mapping of this type available. What is available is the adjoint of φ. But $W^* \xrightarrow{\varphi^*} V^*$, that is, the adjoint, although it is indeed a linear mapping on the dual spaces, is not a candidate for $\mathbf{F}(\varphi)$, for it maps W^* to V^* rather than V^* to W^*. Nonetheless, the adjoint does satisfy a property analogous to property 1 for a functor, for (chapter 13) $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$. This example suggests the following definition.

A *contravariant functor* $\mathbf{F}$ from category $\mathbf{C}$ to category $\mathbf{C}'$ consists of two things—i) a rule which associates, with each object A of category $\mathbf{C}$, an object, written $\mathbf{F}(A)$, of category $\mathbf{C}'$, and ii) a rule which associates, with each morphism φ from object A to object B in category $\mathbf{C}$, a morphism, written $\mathbf{F}(\varphi)$, from object $\mathbf{F}(B)$ to object $\mathbf{F}(A)$ in category $\mathbf{C}'$—subject to the following two conditions:

1. Composition is preserved. For $A \xrightarrow{\varphi} B \xrightarrow{\psi} C$ a diagram in category $\mathbf{C}$, we have

$$\mathbf{F}(\psi \circ \varphi) = \mathbf{F}(\varphi) \circ \mathbf{F}(\psi) \ .$$

2. Identities are preserved. For A any object in category $\mathbf{C}$, we have

$$\mathbf{F}(i_A) = i_{\mathbf{F}(A)} \ .$$

Note that there are only two minor differences between this definition and that of a covariant functor. First, from $A \xrightarrow{\varphi} B$, we obtain $\mathbf{F}(B) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(A)$

(rather than $\mathbf{F}(A) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(B)$, as for a covariant functor). Second, property 1 is $\mathbf{F}(\psi \circ \varphi) = \mathbf{F}(\varphi) \circ \mathbf{F}(\psi)$ (rather than $\mathbf{F}(\psi \circ \varphi) = \mathbf{F}(\psi) \circ \mathbf{F}(\varphi)$, as for a covariant functor). The second change is actually made necessary by the first. Consider $A \xrightarrow{\varphi} B \xrightarrow{\psi} C$ in category $\mathbf{C}$. Then, for $\mathbf{F}$ a contravariant functor, we have $\mathbf{F}(C) \xrightarrow{\mathbf{F}(\psi)} \mathbf{F}(B) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(A)$. Thus $\mathbf{F}(\varphi) \circ \mathbf{F}(\psi)$ and $\mathbf{F}(\psi \circ \varphi)$ are both morphisms (in $\mathbf{C}'$) from $\mathbf{F}(C)$ to $\mathbf{F}(A)$, while $\mathbf{F}(\psi) \circ \mathbf{F}(\varphi)$ is not even defined. It is clear from the discussion above that the construction of taking the dual is a contravariant functor (now using $\mathbf{F}(\varphi) = \varphi^*$) from the category of real (resp., complex) vector spaces to the same category.

Example. Let $\mathbf{C}$ be the category of sets, and $\mathbf{C}'$ the category of (say) real vector spaces. For S any set, let $\mathbf{F}(S)$ be the vector space of all real-valued functions on the set S. (See the first example in chapter 9.) For $S \xrightarrow{\varphi} S'$ a mapping of sets, let $\mathbf{F}(S') \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(S)$ be the linear mapping of vector spaces given in the second example in chapter 9. This is a contravariant functor from the category of sets to the category of real vector spaces.

Note, for example, that functors take isomorphisms to isomorphisms, for, if $A \xrightarrow{\varphi} B$ is an isomorphism, so the top diagram of figure 46 commutes, then so does the bottom diagram, where the solid arrowheads are to be used for contravariant functors, and the regular arrowheads for covariant functors. In general, functors do not take monomorphisms or epimorphisms to themselves (although this often turns out to be the case for categories of interest).

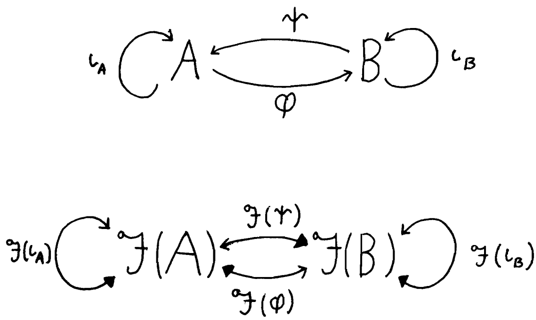

Figure 46

It is sometimes convenient to adopt the following picture when thinking about functors: imagine that there is some sort of "supercategory" whose "objects" are ordinary categories and whose "morphisms" are functors. This picture, for example, suggests the following definition. Let $\mathbf{C}$, $\mathbf{C}'$, and $\mathbf{C}''$ be three categories, and let $\mathbf{F}$ be a functor from category $\mathbf{C}$ to category $\mathbf{C}'$, and $\mathbf{G}$ a functor from category $\mathbf{C}'$ to category $\mathbf{C}''$. We define a new functor, written $\mathbf{G} \circ \mathbf{F}$, from category $\mathbf{C}$ to category $\mathbf{C}''$. For A an object of category $\mathbf{C}$, let $\mathbf{G} \circ \mathbf{F}(A) = \mathbf{G}[\mathbf{F}(A)]$ (i.e., $\mathbf{F}(A)$ is an object of category $\mathbf{C}'$, so $\mathbf{G}[\mathbf{F}(A)]$ is an object of category $\mathbf{C}''$). For $A \xrightarrow{\varphi} B$ a morphism of category $\mathbf{C}$, let $\mathbf{G} \circ \mathbf{F}(\varphi) = \mathbf{G}[\mathbf{F}(\varphi)]$. Thus, if $\mathbf{F}$ were covariant, and $\mathbf{G}$ contravariant, we would have $\mathbf{F}(A) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(B)$, and hence $\mathbf{G} \circ \mathbf{F}(B) \xrightarrow{\mathbf{G} \circ \mathbf{F}(\varphi)} \mathbf{G} \circ \mathbf{F}(A)$. That is, in this case, $\mathbf{G} \circ \mathbf{F}$ would be contravariant. Clearly, $\mathbf{G} \circ \mathbf{F}$ is covariant if both $\mathbf{F}$ and $\mathbf{G}$ are covariant, or if both are contravariant, while $\mathbf{G} \circ \mathbf{F}$ is contravariant if one of $\mathbf{F}$,$\mathbf{G}$ is covariant and the other contravariant.

Example. Let $\mathbf{C}$ be the category of sets, $\mathbf{C}'$ the category of vector spaces, and $\mathbf{C}''$ the category of vector spaces. Let $\mathbf{F}$ be the covariant functor "take the free vector space" from $\mathbf{C}$ to $\mathbf{C}'$. Let $\mathbf{G}$ be the contravariant functor "take the dual" from $\mathbf{C}'$ to $\mathbf{C}''$. Then $\mathbf{G} \circ \mathbf{F}$ is the contravariant functor from the category of sets to the category of vector spaces given in the example above.

This $\mathbf{G} \circ \mathbf{F}$ is called the *composition* of functors $\mathbf{F}$ *and* $\mathbf{G}$.

We conclude with a final example of how the notion of a functor can be used as a tool to organize ideas. Fix a covariant functor $\mathbf{F}$ from category $\mathbf{C}$ to category $\mathbf{C}'$. Let A' be an object in category $\mathbf{C}'$. A *free object* on A' (via $\mathbf{F}$) consists of an object B in category $\mathbf{C}$ together with a morphism α' from A' to $\mathbf{F}(B)$ (in category $\mathbf{C}'$) such that the following property is satisfied: given any object C in category $\mathbf{C}$ together with morphism β' from A' to $\mathbf{F}(C)$ (in category $\mathbf{C}'$), there is a unique morphism γ from B to C (in category $\mathbf{C}$) such that the left diagram of figure 47 commutes.

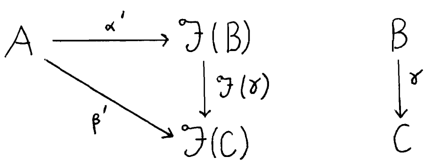

Figure 47

Like all such definitions, this one sounds more complicated than it is (perhaps because definitions are conventionally single sentences). Pictorially, we have figure 48.

Example. Let $\mathbf{C}$ be the category of groups, $\mathbf{C}'$ the category of sets, and $\mathbf{F}$ the forgetful functor. Let S be a set. A free group on S consists of a group G, together with a mapping α' from set S to set G, such that the following property is satisfied: given any group H, together with mapping β' from set S to set H, there is a unique homomorphism γ of groups such that the diagram of figure 49 commutes. This will be recognized, on the one hand, as a special case of the free object definition above and, on the other, as precisely our

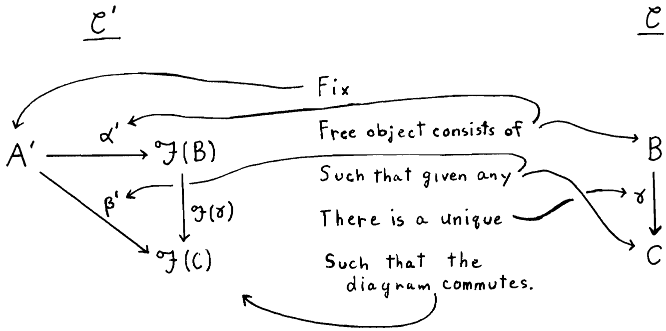

Figure 48

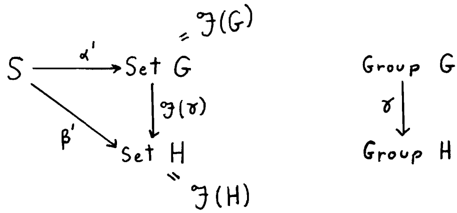

Figure 49

earlier definition of a free group.

Of course, free objects may not exist given a covariant functor. The definition is normally used for forgetful functors. Thus the definition of a free vector space is just that of a free object via the forgetful functor from the category of vector spaces to the category of sets. What is nice about the above definition is that one can use it via functors which "are forgetful, but not completely forgetful," that is, for forgetful functors which are not all the way to the category of sets. Thus one is able in a systematic way to "expand the structure of things which already have some structure, preserving all the old structure and 'freeing the new.' "

Exercise 105. Is composition of functors associative?

Exercise 106. Define the identity functor from a category to that same category. What do you suppose is meant by equivalent categories?

Exercise 107. Associate with each vector space a basis for that vector space (regarded as a set). Does this lead to a functor from the category of vector spaces to the category of sets?

Exercise 108. Compose the forgetful functor from the category of vector spaces to the category of abelian groups with the forgetful functor from the category of abelian groups to the category of sets. What is the result?

Exercise 109. State and prove: free objects are unique.

Exercise 110. Let $\mathbf{F}$ be a covariant functor from category $\mathbf{C}$ to category $\mathbf{C}'$. Suppose that every object A' of $\mathbf{C}'$ possesses a free object via $\mathbf{F}$. Construct the corresponding "free object" functor from the category $\mathbf{C}'$ to category $\mathbf{C}$, and prove that it is indeed a (covariant) functor.

Exercise 111. Does there exist some notion of "free objects" but via contravariant rather than covariant functors?

Exercise 112. Investigate the existence of free objects via various covariant functors which are not forgetful (e.g., via the functor from the category of sets to the category of groups: "free group"). Are there any interesting such examples for which free objects exist?

Exercise 113. Try (and fail) to prove that every covariant functor takes a monomorphism to a monomorphism.

Exercise 114. Is there any sense in which functors take direct sums to direct sums, or direct products to direct products?

Exercise 115. For S a set, let $\mathbf{F}(S)$ be the collection of all subsets of S. For $S \xrightarrow{\varphi} S'$ a mapping of sets, let $\mathbf{F}(S') \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(S)$ be the mapping which sends subset A' of S' to the subset of S consisting of all s in S with $\varphi(s)$ in A'. Prove that this is a contravariant functor from the category of sets to the category of sets.

18

The Category of Associative Algebras

We continue, in our study of algebraic categories, in the direction of "increasing structure."

A *real associative algebra* consists of two things—i) a real vector space V, and ii) a rule that assigns, given vectors v and v', an element (written vv' and called the product of v and v') of V—subject to the following two conditions:

1. The product is linear in each factor. For v, v', and v'' vectors, and a a number, we have

$$(v + av')v'' = vv'' + av'v''$$

and

$$v(v' + av'') = vv' + avv'' \quad .$$

2. The product is associative. For v, v', and v'' vectors, we have

$$(vv')v'' = v(v'v'') \quad .$$

Replacing "real" everywhere above by "complex," we obtain the definition of a *complex associative algebra.* Note that the first condition above is precisely the statement that this "product structure" is an element of $\text{Lin}(V, V; V)$. This same condition implies that $0v = v0 = 0$, where "0" is the zero element of the vector space. (It follows immediately that the only associative algebra which is also a group under the product operation is that based on a zero-dimensional vector space, for $0v = 00$, together with the assumption that the product operation satisfies the conditions for a group, implies $v = 0$.) Note that three operations are actually available in an associative algebra: sum of vectors, product of vectors, and product of vector with number. Finally, note that it is not necessarily true in an associative algebra that $vv' = v'v$.

Example. Let S be any set, and denote by V the vector space of all real-valued functions on the set S (so addition of vectors is addition of functions; multiplication of vectors by real numbers is multiplication of functions by real numbers). For v and v' in V (so v and v' are functions on S), let vv' be the product function (i.e., the function on S whose value at s in S is the number $v(s)v'(s)$, product of numbers). Properties 1 and 2 are immediate: we have a real associative algebra.

Example. Let G be any group. Denote by V the free vector space on the set G, and let α denote the corresponding mapping from set G to set V. Thus every element of V is of the form $a_1\alpha(g_1) + \cdots + a_n\alpha(g_n)$, where $a_1, \ldots, a_n$ are numbers, and $g_1, \ldots, g_n$ are elements of the group G. Consider that linear product in V which, for elements of V of the form $\alpha(g)$, is $\alpha(g)\alpha(g') = \alpha(gg')$ (g and g' in G). (Knowledge of the action of the product on this basis for V, together with linearity, determines the product completely. Thus, e.g., $(a_1\alpha(g_1) + a_2\alpha(g_2))(a_3\alpha(g_3) + a_4\alpha(g_4)) = (a_1a_3)\alpha(g_1g_3) + (a_1a_4)\alpha(g_1g_4) + (a_2a_3)\alpha(g_2g_3) + (a_2a_4)\alpha(g_2g_4)$.) Associativity of the product is immediate from associativity in the group G. Thus we obtain an associative algebra, called the *group algebra* of G.

Example. Let W be any vector space, and let V denote the vector space Lin $(W; W)$. We introduce a product in V: for φ and ψ in V (so each is a linear mapping from W to W), let the "product" be $\psi \circ \varphi$, the composition of linear mappings. Then property 1 is immediate from the definition of linear combinations of elements of Lin$(W; W)$, while property 2 is the statement that composition of mappings is associative. Thus Lin$(W; W)$ has the structure of an associative algebra.

Now let V and W both be associative algebras (both real or both complex). A mapping φ from set V to set W is called a *homomorphism* (of associative algebras) if it is "structure preserving," that is, if φ is a linear mapping of vector spaces and if, furthermore, $\varphi(vv') = \varphi(v)\varphi(v')$ for any v and v' in V. Note that the composition of two homomorphisms is a homomorphism. Let the objects be real associative algebras, the morphisms be homomorphisms of real associative algebras, and the composition be composition of homomorphisms. We obtain the *category of real associative algebras* and, similarly, the *category of complex associative algebras.*

We have available now a way to get a quick picture of what structure is involved in a new category: look for the forgetful functors. From the category of associative algebras, there are forgetful functors to the category of vector spaces ("forget how to take products of vectors"), to the category of abelian groups ("forget both how to take products of vectors and how to take products of vectors by numbers"), and to the category of sets ("forget everything"). There is also a forgetful functor from the category of complex associative algebras to the category of real associative algebras ("forget how to multiply vectors by the complex number 'i'").

Perhaps the most useful free construction in this category is that via the forgetful functor from the category of associative algebras to the category of vector spaces. We now introduce this construction. Let W be a vector space. A *free associative algebra* on W (via the forgetful functor) consists of an associative algebra V, together with a linear mapping α from vector space W to the underlying vector space of V, such that the following property is satisfied: given any associative algebra U, together with linear mapping β from vector space W to vector space U, there is a unique homomorphism $V \xrightarrow{\gamma} U$ of associative algebras such that the diagram of figure 50 commutes. What we

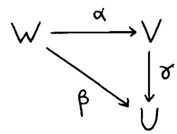

Figure 50

propose to do is to construct explicitly this associative algebra V.

Now W is just a vector space: we do not know how to take products of vectors therein. We wish to "free a product structure on W, preserving its intrinsic vector-space structure." Let w and w' be vectors in W. The "product" (that we wish to define) of w and w' could hardly be an element again of W (for that is not the way free constructions work). What vector space, then, should it be an element of? A perfect candidate is available: $W \otimes W$, the tensor product. Now, for the "product" of w and w', we can choose the element $w \otimes w'$ of $W \otimes W$. The problem now is that this "product" takes us out of our original vector space W. That is, we have gained a product structure (within W) but lost our vector-space structure, for we do not know how, for example, to add an element w of W to an element $w' \otimes w''$ of $W \otimes W$. To "restore a vector-space structure," we take a direct sum; that is, we consider $W \oplus (W \otimes W)$. Then w in W can be represented as the element $(w,0)$ of $W \oplus (W \otimes W)$, while $w' \otimes w''$ in $W \otimes W$ can be represented as the element $(0,w' \otimes w'')$. But now we can add these elements (in $W \oplus (W \otimes W)$), to obtain the element $(w,w' \otimes w'')$ of this direct sum.

Thus we decide to consider $W \oplus (W \otimes W)$. This is a vector space. Furthermore, we take, for the product of elements $(w,0)$ and $(w',0)$ of this vector space, the element $(0,w \otimes w')$ of this vector space. Unfortunately, we do not yet have an associative algebra, for we do not know how to take the "product" of any two elements of this $W \oplus (W \otimes W)$, for example, of the two elements $(w,0)$ and $(0,w' \otimes w'')$, or of the two elements $(0,w \otimes w')$ and $(0,w'' \otimes w''')$. In what vector space would, for example, the "product of $(w,0)$ and $(0,w' \otimes w'')$" like to be? Perhaps the best vector space would be $W \otimes W \otimes W$, so the product could then be $w \otimes w' \otimes w''$ therein. Similarly, the "product" of $(0,w \otimes w')$ and $(0,w'' \otimes w''')$ would like to be $w \otimes w' \otimes w'' \otimes w'''$ in $W \otimes W \otimes W \otimes W$. To now "restore a vector-space structure," we again use the direct sum; that is, we consider $W \oplus (W \otimes W) \oplus (W \otimes W \otimes W) \oplus (W \otimes W \otimes W \otimes W)$. Then, for example, the product of $(w,0,0,0)$ and $(0,w' \otimes w'',0,0)$ in this vector space is to be the element $(0,0w \otimes w' \otimes w'',0)$. We still, however, do not have an associative algebra, for some products (e.g., of $(0,0,w \otimes w' \otimes w'',0)$ and $(0,w''' \otimes w'''',0,0)$) now want to be in "still larger tensor products" (in this case, in $W \otimes W \otimes W \otimes W \otimes W$). What is needed (to get an associative algebra) is an infinite direct sum of tensor products. That is, we must consider the vector space $W \oplus (W \otimes W) \oplus (W \otimes W \otimes W) \oplus \cdots$ (Recall that, for $W_1, W_2, \cdots$ vector spaces, $W_1 \oplus W_2 \oplus \cdots$ is the vector space of all $(w_1, w_2, w_3, \cdots)$ with w_i in W_i $(i = 1,2, \cdots)$, and with all but a finite number of the w_i zero.)

The discussion above, of course, is merely intended to motivate the following construction. Let W be a vector space, and let V be the vector space $V = W \oplus (W \otimes W) \oplus (W \otimes W \otimes W) \oplus \cdots$ We define a product in this V. For $(0, \ldots, 0, w_1 \otimes \cdots \otimes w_n, 0 \cdots)$ and $(0, \ldots, 0, w_1' \otimes \cdots \otimes w_m', 0, \cdots)$ in V (so $w_1, \ldots, w_n, w_1', \ldots, w_m'$ are all in W), let their product be $(0, \ldots, 0, w_1 \otimes \cdots \otimes w_n \otimes w_1' \otimes \cdots \otimes w_m', 0, \cdots)$ in V. Extend this product to all of V by linearity. Then this V becomes an associative algebra. Now let α be the following linear mapping from vector space W to vector space V: for w in W, let $\alpha(w)$ be the element $(w,0,0, \cdots)$ of V. We claim that this (V,α) is in fact a free associative algebra on W. Indeed, let U be any associative algebra, and $W \xrightarrow{\beta} U$ a linear mapping of vector spaces. We must find, and show the uniqueness of, a homomorphism $V \xrightarrow{\gamma} U$ of associative algebras which makes the diagram of figure 50 commute. For any element of V of the form $(w,0,0, \cdots)$ (i.e., of the form $\alpha(w)$), we must choose $\gamma(w,0, \cdots) = \beta(w)$ in order that the diagram commute. But, in order that γ be a homomorphism, we must have $\gamma(0, w \otimes w', 0, \cdots) = \gamma[(w,0, \cdots)(w',0, \cdots)] = \gamma(w,0, \cdots)\gamma(w',0, \cdots) = \beta(w)\beta(w')$. Similarly, we must have $\gamma(0,0,w \otimes w' \otimes w'',0, \cdots) = \beta(w)\beta(w')\beta(w'')$, etc. But every element of V is a linear combination of elements of this form (i.e., of elements with only one nonzero entry, and that a tensor product of elements of W). Hence γ, if it is to be a homomorphism (and, in particular, a linear mapping) is completely and uniquely determined by its action on these elements. Thus (V,α) is indeed a free associative algebra.

The construction above has the characteristic feature of free constructions: it uses "brute force" to make available the operations desired. One then "expands the size of what one has, minimally, to get the appropriate structure." The result is always a free object because, in this example, "the action of γ on elements of V which come directly from W (i.e., elements of the form $(w,0, \cdots)$) is determined by commutativity of the diagram, while the action on other elements of V is determined (since every element of V can be obtained, using the operations in V, from elements which come directly from W) by the requirement that γ be a homomorphism."

Example. Let W be a one-dimensional vector space: let w in W be a basis for W. Then $w \otimes w$ is a basis for $W \otimes W$, $w \otimes w \otimes w$ is a basis for $W \otimes W \otimes W$, etc. (Each of $W \otimes W$, etc., is one-dimensional.) Then the general element of V is $(a_1 w, a_2 w \otimes w, \ldots, a_n w \otimes \cdots \otimes w, 0, 0, \cdots)$, where $a_1, \ldots, a_n$ are numbers. It is convenient to denote this element not as above, but rather as follows: $a_1 w + a_2 w^2 + \cdots + a_n w^n$. (This is simply a change to a more suggestive notation.) Thus V, as a vector space, is just "the vector space of all polynomials in the 'variable w' having no 'constant term.'" Note, furthermore, that the product structure in V is just the usual rule for multiplication of polynomials. Thus the free associative algebra on a one-dimensional vector space is the "associative algebra of polynomials in one variable, having no constant term." The statement that this V is a free associative algebra becomes, in the present example, the following: "given an associative algebra U, if you know where in U w is to be sent, then you also know where in U every polynomial in w must be sent, provided you require that the resulting mapping from the algebra V of polynomials to U be a homomorphism of associative algebras."

Fix an associative algebra V. A *subalgebra* of V consists of a nonempty subset W of V such that i) for any w and w' in the subset W, and any number a, $w + aw'$ is also in W, and ii) for any w and w' in W, ww' is also in W. Thus a subalgebra is a subset "in which one remains under all the operations available, if applied to elements of the subset." (This definition is, of course, a special case of that of a subobject.) Note that a subalgebra of V is itself an associative algebra, with the operations of V. Of course, a subalgebra W of associative algebra V is, as a vector space, a subspace of vector space V.

Fix an associative algebra V. An *ideal* of V consists of a nonempty subset W of V such that i) for any w and w' in the subset W, and any number a, $w + aw'$ is also in W, and ii) for any w in W and v in V, vw and wv are both in W. Note that an ideal is not only a subalgebra; it is more: let W be a subspace of vector space V—then, in order that W be a subalgebra, wv (w in W) must also be in W provided v is in W, while, in order that W be an ideal, wv and vw must be in W for all v in V.

Example. Let V be the real associative algebra of all real-valued functions on an infinite set S. Denote by W the collection of all bounded functions. Then W is a subalgebra of V (for linear combinations and products of bounded functions are bounded), but W is not an ideal of V (for the product of a bounded function and an arbitrary function need not be bounded). Let S' be a subset of S, and denote by U the collection of all functions on S which vanish on the subset S'. Then U is an ideal of V (for not only do linear combinations and products of functions which vanish on S' vanish on S', but, furthermore, the product of a function which vanishes on S' with an arbitrary function on S is a function which vanishes on S').

Next note that any intersection of subalgebras of an associative algebra is a subalgebra and that any intersection of ideals of an associative algebra is an ideal. Let K be any subset of associative algebra V. The intersection of all subalgebras containing K is called the subalgebra *generated by* K; the intersection of all ideals containing K is called the ideal *generated by* K. Of course, the subalgebra generated by K is the collection of all linear combinations (in V) of products of elements of K, while the ideal generated by K is the collection of all linear combinations (in V) of products of elements of V, where each product contains at least one factor from the subset K.

Example. Let V be the real associative algebra of all real-valued functions on set S. Let K be the subset of V consisting of all functions whose values are always between the numbers 4 and 5. Then the subalgebra generated by K is the subalgebra of bounded functions, while the ideal generated by K is V itself.

We now describe the sense in which the following statement is true: "Subalgebras are to subgroups as ideals are to normal subgroups." Fix an associative algebra V, and let W be a subalgebra of V. Then, in particular, W is a subspace of vector space V: hence we may take the quotient space, V/W. For W a just plain subalgebra, that is all we can get—a just plain vector space V/W. Now suppose that this W is not only a subalgebra, but actually an ideal. Consider two elements of V/W, that is, two cosets of W in vector space V: $v + W$ and $v' + W$. We wish to define the product of these cosets to be the coset $vv' + W$. We must, however, check that this is well defined. Our cosets could as well have been represented as $(v + w) + W$ and $(v' + w') + W$, with w and w' in the subspace W. Had we written them this way, we would have for the product $(v + w)(v' + w') + W = vv' + vw' + wv' + ww' + W$. But, since W is now assumed to be an ideal, vw', wv', and ww' are all in W. Hence the above becomes $(v + w)(v' + w') + W = vv' + W$. That is, our product is indeed independent of how the cosets are represented. Thus, on the vector space V/W, we have defined a product operation, an operation which is easily checked to satisfy the two conditions for an associative algebra. That is, V/W, where W is an ideal in V, has the structure of an associative algebra. This V/W is called the ***quotient algebra*** of the associative algebra V by the ideal W.

Example. Let S be a set, V the associative algebra of functions on S, and W the ideal consisting of functions which vanish on some fixed subset S' of S. Then V/W ("functions on S modulo addition of a function which vanishes on S'") is isomorphic (as an associative algebra) to the associative algebra of functions on the set S'. ("It does not make any difference what values a function takes off S', for one can always add to any function, while remaining within the same coset of W, any function which vanishes on S'.")

Exercise 116. Let $V \xrightarrow{\varphi} W$ be a homomorphism of associative algebras. Define $\mathrm{Im}(\varphi)$, and show that it is a subalgebra of W; define $\mathrm{Ker}(\varphi)$, and show that it is an ideal of V. Give the homomorphisms in $V \xrightarrow{\alpha} V/\mathrm{Ker}(\varphi) \xrightarrow{\beta} \mathrm{Im}(\varphi) \xrightarrow{\gamma} W$.

Exercise 117. Prove that every associative algebra is a quotient algebra of a free associative algebra.

Exercise 118. A semigroup is a set on which there is given an associative product. Introduce the category of semigroups, and a forgetful functor from the category of associative algebras to the category of semigroups. Prove that a group algebra is a free object via this forgetful functor.

Exercise 119. Find all associative algebras based on a two-dimensional vector space. (Meaning: Find a collection of associative algebras on two-dimensional vector spaces such that no two are isomorphic and such that any associative algebra on a two-dimensional vector space is isomorphic to one in the collection.)

Exercise 120. Find the direct product and direct sum of two associative algebras (both of which exist).

Exercise 121. Let W be a subalgebra of associative algebra V. Define a complementary subalgebra. Does one always exist? Does one always exist if W is an ideal?

Exercise 122. Find all finite-dimensional free associative algebras.

Exercise 123. A unit of an associative algebra V is an element e of V such that $ev = ve = v$ for every v in V. Find an example of an associative algebra which has no unit. State and prove: every associative algebra is a subalgebra of a "minimal" associative algebra with unit.

Exercise 124. Let W be a vector space, and let V be the free associative algebra on W. Let $\underline{W}$ be the associative algebra of functions on the set W^*. Consider the linear mapping from vector space V to vector space $\underline{W}$ which sends, for example, $(0,0,w \otimes w' \otimes w'',0,\cdots)$ in V to the function on W^* whose value at φ (in W^*) is $\varphi(w)\varphi(w')\varphi(w'')$. Show that this is a homomorphism of associative algebras. What is its image and kernel?

19

The Category of Lie Algebras

A *real* (resp., *complex*) *Lie algebra* consists of two things—i) a real (resp., complex) vector space V, and ii) a rule which assigns, given vectors v and v', an element (written $[v,v']$, and called the bracket of v and v') of V—subject to the following three conditions:

1. The bracket is linear in each factor. For v, v', and v'' vectors, and a a number, we have

$$[v + av',v''] = [v,v''] + a[v',v'']$$

and

$$[v,v' + av''] = [v,v'] + a[v,v''] \ .$$

2. The bracket is antisymmetric. For v and v' any vectors, we have

$$[v,v'] = -[v',v] \ .$$

3. The *Jacobi relation* is satisfied. For any vectors v, v', and v'', we have

$$[v,[v',v'']] + [v',[v'',v]] + [v'',[v,v']] = 0 \ .$$

Note that the three terms in the Jacobi relation are obtained by cyclic permutations of v, v', and v''. It is immediate from antisymmetry of the bracket that $[v,v] = 0$ for any vector v. Setting any two of v, v', v'' equal to each other, the Jacobi relation becomes an identity. Note also that a Lie algebra does not differ all that much from an associative algebra. More generally, a vector space V, together with an element of $\mathrm{Lin}(V,V;V)$, is called an algebra. Both Lie algebras and associative algebras are special cases of algebra in which the particular element of $\mathrm{Lin}(V,V;V)$ satisfies certain additional conditions. Thus one thinks of conditions 2 and 3 above as analogous to the associativity condition for an associative algebra.

Example. Let V be any vector space, and let κ be any fixed element of V^*. For v and v' in V, set $[v,v'] = \kappa(v)v' - \kappa(v')v$. That this bracket satisfies conditions 1 and 2 above is immediate. We check the third condition: $[v,[v',v'']] = [v,\kappa(v')v'' - \kappa(v'')v'] = \kappa(v')(\kappa(v)v'' - \kappa(v'')v) - \kappa(v'')(\kappa(v)v' - \kappa(v')v) = \kappa(v)\kappa(v')v'' - \kappa(v)\kappa(v'')v'$. Similarly for $[v',[v'',v]]$ and $[v'',[v,v']]$. Adding these three expressions, one obtains the Jacobi relation. Hence we have a Lie algebra.

Let V and W be Lie algebras (both real or both complex). A mapping φ from set V to set W is called a ***homomorphism*** (of Lie algebras) if φ is a linear mapping of vector spaces and if, furthermore, $\varphi([v,v'])=[\varphi(v),\varphi(v')]$ for any v and v' in V. Note that the composition of two homomorphisms is a homomorphism. Letting the objects be real (resp., complex) Lie algebras, the morphisms homomorphisms of real (resp., complex) Lie algebras, and the composition composition of homomorphisms, one obtains the *category of real* (resp., *complex*) *Lie algebras.* As was the case with associative algebras, we have forgetful functors from the category of Lie algebras to the category of vector spaces, the category of abelian groups, and the category of sets.

The discussion of subalgebras and ideals for associative algebras can be repeated, without change, for Lie algebras. A *subalgebra* of Lie algebra V consists of a nonempty subset W of V such that i) for any w and w' in the subset W, and any number a, $w+aw'$ is also in W, and ii) for any w and w' in W, $[w,w']$ is also in W. An *ideal* of Lie algebra V consists of a nonempty subset W of V such that i) for any w and w' in the subset W, and any number a, $w+aw'$ is also in W, and ii) for any w in W and v in V, $[w,v]$ (and hence also $[v,w]$) is in W. Any intersection of subalgebras (of a Lie algebra) is a subalgebra; any intersection of ideals is an ideal. Let V be a Lie algebra, and K any subset of V. The intersection of all subalgebras (resp., ideals) containing K is called the subalgebra (resp., ideal) *generated by* K. Let W be a fixed subalgebra of Lie algebra V. Then V/W (quotient space of vector spaces) is a vector space. Suppose now that W is in fact an ideal of V. Then the bracket operation $[v+W,v'+W]=[v,v']+W$ on the vector space V/W makes this vector space into a Lie algebra called the *quotient algebra* of the Lie algebra V by the ideal W.

Example. In the example just above, let W be the subset of Lie algebra V consisting of all elements v of V with $\kappa(v)=0$. Then W is certainly a subspace of vector space V. For w in W and v in V, we have $[w,v]=\kappa(w)v-\kappa(v)w=-\kappa(v)w$. Thus W is an ideal of V. The quotient algebra, V/W, is one-dimensional (as a vector space). Since, given any two elements of the quotient algebra, one is a numerical multiple of the other, the bracket of these two elements (in the Lie algebra V/W) must be zero. Thus, in the quotient algebra V/W, the bracket of any two elements vanishes.

There is an important functor from the category of associative algebras to the category of Lie algebras. Let V be any associative algebra (with product written vv'). Given any two vectors, v and v', define the left side of

$$[v,v']=vv'-v'v$$

by the right side. Thus we have defined a "bracket operation" on the associative algebra V. This bracket operation is clearly linear in each factor (because the product in the associative algebra V is). Furthermore, $[v,v']=-[v',v]$. We claim, finally, that this bracket operation satisfies the Jacobi relation. Indeed, we have $[v,[v',v'']]=[v,v'v''-v''v']=vv'v''-vv''v'-v'v''v+v''v'v$. (No parentheses are needed in these products, since V is an associative algebra.) Evaluating $[v',[v'',v]]$ and $[v'',[v,v']]$ similarly, and adding, one easily checks (all the terms cancel) that the Jacobi relation is satisfied. Thus the vector space V, with this bracket operation, is a Lie algebra. Now, for V any associative algebra, let $\mathbf{F}(V)$ be this Lie algebra just constructed. Next, note that, for $V \xrightarrow{\varphi} W$ a homomorphism of associative algebras, the mapping φ from set V to set W is also a homomorphism (which we can write $\mathbf{F}(\varphi)$) from Lie algebra $\mathbf{F}(V)$ to Lie algebra $\mathbf{F}(W)$. The two properties for a functor are immediate. Thus we have obtained a covariant functor from the category of associative algebras to the category of Lie algebras.

We shall regard the functor above as forgetful. The reason is this. Let V be an associative algebra. Then, for v and v' in V, we have the identity $vv' = (1/2)(vv' + v'v) + (1/2)(vv' - v'v)$. The first term on the right is a product in the vector space V which is linear in each factor, and which remains the same under interchange of v and v'. The second term (which we wrote as $(1/2)[v,v']$ above) is also a product in the vector space V which is linear in each factor, but which reverses sign under interchange of v and v'. Thus we "decompose the product in the associative algebra V into two products, one symmetric and one antisymmetric." The covariant functor $\mathbf{F}$ "forgets the symmetric part of the (associative) product of V." In this sense, $\mathbf{F}$ is a forgetful functor.

The following situation arises occasionally. One has a certain Lie algebra V. One wishes to obtain the free associative algebra on V via the forgetful functor above. We now carry out this construction of a free object. The result is conventionally called the universal enveloping algebra of V (rather than "the free associative algebra on Lie algebra V via the forgetful functor from associative to Lie algebras"). A *universal enveloping algebra* of Lie algebra V is an associative algebra W, together with a linear mapping φ from vector space V to vector space W satisfying $\varphi([v,v']) = \varphi(v)\varphi(v') - \varphi(v')\varphi(v)$, such that the following condition is satisfied: given any associative algebra U, together with a linear mapping ψ from vector space V to vector space U satisfying $\psi([v,v']) = \psi(v)\psi(v') - \psi(v')\psi(v)$, there is a unique homomorphism

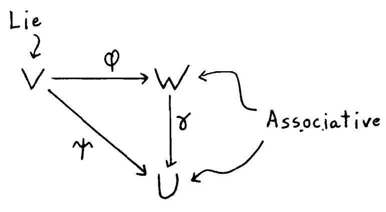

Figure 51

$W \xrightarrow{\gamma} U$ of associative algebras such that the diagram of figure 51 commutes. Note that this definition is a special case of the general definition of a free

object. (Let V be a Lie algebra, and W an associative algebra. Then a "homomorphism from Lie algebra V to Lie algebra $\mathbf{F}(W)$, where $\mathbf{F}$ is the forgetful functor," is precisely the same thing as a "linear mapping φ from vector space V to vector space W satisfying $\varphi([v,v']) = \varphi(v)\varphi(v') - \varphi(v')\varphi(v)$." The former, which appears in the general definition of a free object, is replaced by the latter in the definition above of a universal enveloping algebra.)

Fix once and for all a Lie algebra V. We construct its universal enveloping algebra. The first step is to consider the free associative algebra on vector space V, $K = V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$ Let α be the linear mapping from vector space V to associative algebra K which sends v in V to $(v,0,0,\cdots)$ in K. Thus the first step is to "ignore completely the Lie algebra structure of V, but not its vector-space structure, and 'free' on this vector space an associative product." Of course, this (K,α) is not even a good candidate for the universal enveloping algebra of V, for we do not have $\alpha([v,v']) = \alpha(v)\alpha(v') - \alpha(v')\alpha(v)$. In fact, for v and v' in V, $\alpha([v,v'])$ is the element $([v,v'],0,\cdots)$ of K, while $\alpha(v)\alpha(v') - \alpha(v')\alpha(v)$ is the element $(0,v \otimes v' - v' \otimes v,0,\cdots)$ of K: these are not even close to being equal. To "make them be equal," we proceed as follows. Denote by I the ideal of associative algebra K generated by elements of K of the form $\alpha([v,v']) - \alpha(v)\alpha(v') + \alpha(v')\alpha(v)$, with v and v' in V. (Thus the ideal I represents the "elements of K we wish to make equal to zero.") Let W be the quotient algebra, $W = K/I$, and let β be the homomorphism $K \xrightarrow{\beta} W$ (of associative algebras) which sends k in K to the coset $k + I$ of I in K. Finally, set $\varphi = \beta \circ \alpha$, so φ is a linear mapping from vector space V to vector space W. We now claim: $\varphi([v,v']) = \varphi(v)\varphi(v') - \varphi(v')\varphi(v)$. Indeed, for v and v' in V, we have $\varphi([v,v']) - \varphi(v)\varphi(v') + \varphi(v')\varphi(v) = \beta[\alpha([v,v']) - \alpha(v)\alpha(v') + \alpha(v')\alpha(v)]$. But this last quantity in square brackets is in the ideal I of K (that is the way we defined I), whence β of that quantity is zero. Thus we have shown $\varphi([v,v']) = \varphi(v)\varphi(v') - \varphi(v')\varphi(v)$. (This is not very surprising. One generates an ideal by "what one wants to make be zero," takes a quotient, and "what one wants to make be zero indeed becomes zero.")

We began with a Lie algebra V. We have now obtained a certain associative algebra W and a mapping φ from vector space V to vector space W satisfying $\varphi([v,v']) = \varphi(v)\varphi(v') - \varphi(v')\varphi(v)$. We claim that this (W,φ) is a universal enveloping algebra of V. Indeed, let U be an associative algebra, and $V \xrightarrow{\psi} U$ a linear mapping of vector spaces satisfying $\psi([v,v']) = \psi(v)\psi(v') - \psi(v')\psi(v)$. Then, since K is the free associative algebra on vector space V, there is a unique homomorphism $K \xrightarrow{\mu} U$ such that the diagram of figure 52 commutes. Given any element, for example, $\alpha([v,v']) - \alpha(v)\alpha(v') + \alpha(v')\alpha(v)$, of the ideal I of K, μ of this element is zero, for $\mu[\alpha([v,v']) - \alpha(v)\alpha(v') + \alpha(v')\alpha(v)] = \psi([v,v']) - \psi(v)\psi(v') + \psi(v')\psi(v) = 0$, where, in

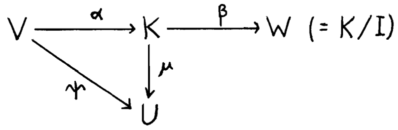

Figure 52

the first step, we have used commutativity of the diagram and, in the second, we have used the property above of ψ. In other words, μ takes entire cosets of I in K to single elements of U. Hence there is a homomorphism γ of associative algebras such that the diagram of figure 53 commutes. Noting that this γ is unique, we have that (W,φ) is indeed the universal enveloping algebra of the Lie algebra V, as claimed. (Note that the structure of this construction is almost identical to that of the tensor product of vector spaces.)

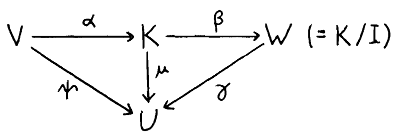

Figure 53

Example. Let V be the two-dimensional Lie algebra in which $[v,v'] = 0$ for every v and v' in V. Let x and y (elements of V) be a basis for V. Then $x \otimes x$, $x \otimes y$, $y \otimes x$, $y \otimes y$ is a basis for $V \otimes V$, etc. Thus a typical element of K is $(x + 2y,\ x \otimes x - x \otimes y,\ 3x \otimes x \otimes y + y \otimes x \otimes y,\ 0, \cdots)$. We agree to write this element (for example) in the following simplified form: $x + 2y + xx - xy + 3xxy + yxy$. Thus K is the associative algebra of "polynomials in the variables x and y having no constant term, where the order of the x and y in each term of the polynomial is relevant." Since all brackets in V vanish, any two such elements of K which "differ only in the order of the x and y in each term" lie in the same coset of I. Thus, for example, the above element of K is in the same coset of I as the element $x + 2y + xx - yx + 3xyx + yyx$ of K, for the difference (in K) between these two elements can be written $-(xy - yx) + 3x(xy - yx) + y(xy - yx)$. Since "$xy - yx$" is in the ideal I, so must be this expression. Thus "cosets ignore order," whence the quotient algebra, $W = K/I$, is the associative algebra of "all polynomials in variables x and y having no constant term, and where the order of the x and y in each term is irrelevant." This W is the universal enveloping algebra of the Lie algebra V.

We conclude our discussion of Lie algebras with a definition. Let V be a Lie algebra, and W its universal enveloping algebra. Denote by C the subset

of W consisting of all w in W with $ww' - w'w = 0$ for all w' in W. (A similar C could, of course, be defined for any associative algebra W.) Clearly, linear combinations and products of elements of C are in C: hence C is a subalgebra of W (not in general an ideal). This C is called the *Casimir subalgebra* (of the universal enveloping algebra of the Lie algebra V). This Casimir subalgebra, as we shall see, plays an important role in the theory of representations.

Example. In the example above, Casimir subalgebra of the universal enveloping algebra of Lie algebra V is the entire universal enveloping algebra (since products in W commute).

Exercise 125. Find all Lie algebras whose underlying vector space is two-dimensional.

Exercise 126. Suppose that, in the definition of a Lie algebra, one had changed one of the signs in the Jacobi relation. Prove from this "modified Jacobi relation" that $[v,[v',v'']] = 0$ for any v, v', v''.

Exercise 127. Construct direct sums and direct products of Lie algebras.

Exercise 128. Is it true that, given any Lie algebra V, there exists an associative algebra W, together with an isomorphism from Lie algebra $\mathbf{F}(W)$ to Lie algebra V, where $\mathbf{F}$ is the forgetful functor?

Exercise 129. Define the kernel and image of a homomorphism of Lie algebras, and verify that the former is an ideal and the latter a subalgebra.

Exercise 130. Let V be any vector space, and let K be the vector space $\mathrm{Lin}(V;V) \oplus V$. Introducing, on this K, the bracket $[(\varphi,v),(\varphi',v')] = (\varphi \circ \varphi' - \varphi' \circ \varphi, \varphi(v') - \varphi'(v))$, verify that K becomes a Lie algebra. (Check that the product $(\varphi,v)(\varphi',v') = (\varphi \circ \varphi', \varphi(v))$ does not, however, make K an associative algebra.) Next, show that the subspace of K consisting of elements of the form $(\varphi,0)$ is a subalgebra and that the subspace consisting of elements of the form $(0,v)$ is an ideal. Find, for the latter, the quotient algebra.

Exercise 131. Find an example of an associative algebra which is not the universal enveloping algebra of any Lie algebra.

Exercise 132. Find associative algebras V and V' which are not isomorphic, but which are such that $\mathbf{F}(V)$ and $\mathbf{F}(V')$ are isomorphic, where $\mathbf{F}$ is the forgetful functor to the category of Lie algebras.

Exercise 133. Let V be a Lie algebra, W its universal enveloping algebra, and $\mathbf{F}(W)$ the corresponding Lie algebra via the forgetful functor. Find a natural homomorphism $V \xrightarrow{\varphi} \mathbf{F}(W)$ of Lie algebras. Prove that this φ is a monomorphism. Find a necessary and sufficient condition (on V) that it be an isomorphism.

Exercise 134. Let V be a two-dimensional vector space, and let x and y be two elements of V that form a basis. Introduce, on V, the bracket $[ax + by, cx + dy] = (ad - bc)x$. Prove that V becomes a Lie algebra. Find its universal enveloping algebra and Casimir subalgebra.

Example: The Algebra of Observables

We have some physical system whose properties we wish to study. We adopt the point of view that the description of this system is to be carried out by means of certain observations on the system. We wish to obtain a mathematical formalism in terms of which these observations can be organized.

Let V be a two-dimensional, complex vector space, and let P and Q be two elements of V which form a basis. Consider the associative algebra $O = \mathbf{C} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$, where $\mathbf{C}$ denotes the vector space of complex numbers. (That is, except for the inclusion of $\mathbf{C}$, this O is the free associative algebra on the vector space V.) Thus an element of O is "a polynomial in P and Q, in which the order ot the P and Q in each term is relevant," for example, $4i + 2P - iQPQQ + PPQQP$.

This setup is intended to represent the following physical situation. We think of Q as representing a certain "observation of configuration" that can be made on the system, and of P as an "observation of momentum." We imagine, furthermore, that we know, given two observations which can be made on our system, how to construct physically a new observation which can be regarded as "a linear combination of the two observations" and another which can be regarded as "the product of the two observations." (We can imagine such a thing if we like. I do not know, however, how actually to give a physical prescription for so combining actual physical measuring instruments to obtain, in each case, the new measuring instrument. In other words, that the set of observables on a physical system seems to have the structure of an associative algebra is apparently a rather subtle feature of our description of physical systems.) Thus we form an associative algebra out of the basic elements, P and Q, an algebra whose elements are to represent "observables of the physical system." (In order to simplify this discussion, we have suppressed one feature, which should, for correctness, be noted. Denote by "*" the mapping from O to O satisfying the following properties: i) for A and B in O, and a a complex number, $(A + aB)^* = A^* + aB^*$, ii) for A and B in O, $(AB)^* = B^*A^*$, and iii) $Q^* = Q$ and $P^* = P$. Note that these properties define the mapping "*" completely and uniquely, e.g., $(4i + 2P - iQPQQ + PPQQP)^* = -4i + 2P + iQQPQ + PQQPP$. In fact, the observables are only those elements A of O satisfying $A^* = A$.)

We next note that there are, in fact, some additional physical things about the space of observables that we have not yet incorporated into O. Let us suppose, first, that our system is subject to the laws of classical mechanics. But in classical mechanics, "observations can be made on a physical system in such a way that the disturbance of the system by the observation is negligible." In other words, the order in which observations (and, in particular, our fundamental observations, P and Q) are made on the system, in classical mechanics, is irrelevant. Thus, if our system is to be classically described, we would expect to have $PQ = QP$. Suppose, on the other hand, that our system is subject to the laws of quantum mechanics. Then "an observation has a non-negligible effect on the system observed," a fact which is to be reflected, within the formalism of quantum mechanics, by the canonical commutation relation, $PQ - QP = \hbar/i$, on the fundamental observables, where $\hbar$ is a certain real number called Planck's constant. Of course, within the associative algebra O, $PQ - QP$ is just $PQ - QP$: it is neither zero nor $\hbar/i$. It is convenient to treat the classical and quantum cases simultaneously: we wish to impose, within our algebra O, the additional condition $PQ - QP = r/i$, where r is an (as yet unspecified) real number.

Of course, the way to "impose additional conditions" on an algebra is to take a quotient algebra by an appropriate ideal. Thus, denote by I_r the ideal of O generated by the element $PQ - QP - r/i$ of O. Then the most general element of I_r is a linear combination of elements of O of the form $A(PQ - QP - r/i)B$, with A and B in O. Set $O_r = O/I_r$, the quotient algebra of the associative algebra O by the ideal I_r. Thus an element of O_r is "a polynomial in P and Q in which order of factors is relevant, but in which P and Q can be interchanged, provided an additional term r/i is included in the polynomial." Thus the element of O_r represented as PPQ is the same as the element $PQP + Pr/i$ (for both lie in the same coset of I_r in O).

The situation is now the following. The elements of O represent "potential observables of the system—potential in the sense that they have not yet been told whether the system is to be treated classically or quantum mechanically; only after they are provided with this additional information will they become actual observables." The quotient algebra O_0 represents "the associative algebra of classical observables," and $O_\hbar$ "the associative algebra of quantum observables." A given element A of O "defines the classical observable (element of O_0) given by the coset $A + I_0$" and "defines the quantum observable (element of $O_\hbar$) given by the coset $A + I_\hbar$." The only relationship between the classical and quantum observables is that which arises from the fact that each is a quotient algebra of the associative algebra O (figure 54).

Note in particular that, since I_0 is a different ideal in O than $I_\hbar$, each coset of I_0 in O intersects many cosets of $I_\hbar$ and vice versa. In this sense, "with each classical observable there are associated many quantum observables (namely, those whose cosets intersect the coset representing the classical observable)" and vice versa. Nonetheless, it is possible, at least in this example, to invent a correspondence between classical and quantum observables

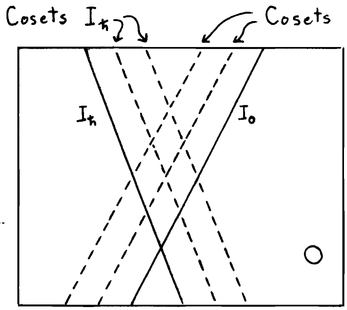

Figure 54

(although this prescription fails for systems with more degrees of freedom). Denote by K the subspace of vector space O generated by polynomials which consist of the sum of a given number of P and Q written in all orders (e.g., $PPQQ + PQPQ + PQQP + QPQP + QPPQ + QQPP$). Then this K, it can be checked, is complementary, as a subspace of vector space O, to both the subspace I_0 and the subspace $I_\hbar$. Thus the quotient spaces $O_0 = O/I_0$ and $O_\hbar = O/I_\hbar$ are each isomorphic to vector space K, whence we have an isomorphism between vector space O_0 and vector space $O_\hbar$. Thus the space of classical observables is isomorphic, as a vector space, to the space of quantum observables. This is not, however, an isomorphism of associative algebras (e.g., because the product in O_0 is commutative, while that in $O_\hbar$ is not).

The mathematical formalism makes it easy to decide what is true and what is not about these classical and quantum observables, and how the observables in the two cases are related to each other.

21

Example: Fock Vector Space

In this chapter we describe the mathematical formalism within which one treats quantum systems of many identical, noninteracting particles.

The states of a quantum system form a complex vector space (actually, a Hilbert space, which is a complex vector space with certain additional structure. We shall here be concerned only with the vector-space structure on the space of states. We introduce Hilbert spaces later.) The operation of taking linear combinations of vectors in this vector space corresponds physically to that of taking superpositions of quantum states.

We first wish to decide how the vector spaces of quantum states for two systems are to be combined when the systems are combined in certain ways. Consider two systems, the states of the first described by complex vector space V_1, and the states of the second by V_2. Suppose we "regard these two separate systems as one," that is, we consider a new system having "these two systems as components." (Note that we are not here turning on an interaction between the systems; rather, we are simply looking at the two systems from a different viewpoint.) Then, we claim, the vector space of the combined system's states is just $V_1 \otimes V_2$, the tensor product. Thus, if the first system is in state v_1 (element of V_1), and the second system in state v_2, then the combined system would be in state $v_1 \otimes v_2$. Note, however, that it is not true in general that every element of $V_1 \otimes V_2$ is of the form $v_1 \otimes v_2$. Thus it is possible to have a state of the combined system (e.g., $v_1 \otimes v_2 + v_1' \otimes v_2'$) for which the individual components are not in definite states. We think of the state above as "a superposition of the states in which the first system is in v_1 and the second in v_2, with the state in which the first system is in v_1' and the second in v_2'." If the combined system were known to be in this state, and if we somehow determined the state of the first system and found it to be v_1, then the second system would have to be in v_2.

There is a second way to combine systems, but it is difficult to describe this combination procedure in the same generality as that above. Suppose, for example, that each of the two systems consists of a potential in which a particle moves. For the combined system, we wish to take the two potentials, but introduce only a single particle to move about in these potentials. Then the complex vector space of states of this combined system would be $V_1 \oplus V_2$, the direct sum. Thus the most general state of the combined system is of the form (v_1, v_2), with v_1 in V_1 and v_2 in V_2. We regard this state as "the superposition of the state, $(v_1,0)$, in which the particle is certainly in the first potential, with the state $(0,v_2)$, in which the particle is certainly in the second potential." For the state (v_1,v_2), we think of v_1 as "the amplitude that the particle has in the first potential" and of v_2 as "the amplitude for the second potential." In such a state, of course, the particle is not definitely in either potential.

We now return to the system of interest, a system of many identical, noninteracting particles (e.g., photons). Suppose that we are given a certain complex vector space V which represents the "space of one-particle states." (We shall discuss later how one obtains this V.) Then, by the remarks above, one should take, for "the space of two-particle states," $V \otimes V$, and, "for the space of three-particle states," $V \otimes V \otimes V$, etc. For the "space of zero-particle states (i.e., vacuum states)," one takes $\mathbf{C}$, the vector space of complex numbers. Now, our system need not, in general, be in a state corresponding to a definite number of particles: it could, for example, be in a superposition of states having different numbers of particles. These remarks suggest that we take, for the space of states of our system, the complex vector space $F = \mathbf{C} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$ a direct sum. (Note that we are here interested in this F only as a vector space, not as an associative algebra.)

What we must do next is "correct" this F to take account of the fact that our particles are to be indistinguishable. Denote by σ the linear mapping from $V \otimes V$ to $V \otimes V$ which sends $v \otimes v'$ to $(1/2)(v \otimes v' + v' \otimes v)$, where the action of σ is extended to all of $V \otimes V$ (i.e., to linear combinations of elements of the form $v \otimes v'$) by linearity. Similarly, denote by σ (this use of the same letter for different mappings does not lead to confusion) the linear mapping from $V \otimes V \otimes V$ to $V \otimes V \otimes V$ which sends $v \otimes v' \otimes v''$ to $(1/6)(v \otimes v' \otimes v'' + v \otimes v'' \otimes v' + v' \otimes v'' \otimes v + v' \otimes v \otimes v'' + v'' \otimes v \otimes v' + v'' \otimes v' \otimes v)$. (Note that one simply writes the vectors in all orders.) Similarly for $V \otimes V \otimes V \otimes V$, etc. Application of this σ to an element of $V \otimes \cdots \otimes V$ is called symmetrization of the element of the tensor product. An element w of $V \otimes \cdots \otimes V$ such that $\sigma(w) = w$ is said to be symmetric. Since, on $V \otimes \cdots \otimes V$, we have $\sigma \circ \sigma = \sigma$, it follows that, for any w in $V \otimes \cdots \otimes V$, $\sigma(w)$ is symmetric. Of course, the symmetric elements form a subspace of $V \otimes \cdots \otimes V$. Similarly, denote by τ the linear mapping from $V \otimes V$ to $V \otimes V$ which sends $v \otimes v'$ to $(1/2)(v \otimes v' - v' \otimes v)$; by τ the linear mapping from $V \otimes V \otimes V$ to $V \otimes V \otimes V$ which sends $v \otimes v' \otimes v''$ to $(1/6)(v \otimes v' \otimes v'' + v' \otimes v'' \otimes v + v'' \otimes v \otimes v' - v'' \otimes v' \otimes v - v' \otimes v \otimes v'' - v \otimes v'' \otimes v')$, etc. (The formula is the same as for σ, except that one attaches a minus sign to odd permutations of the vectors.) Application of this τ to an element of $V \otimes \cdots \otimes V$ is called antisymmetrization of the element of the tensor product. An element w of $V \otimes \cdots \otimes V$ such that $\tau(w) = w$ is said to be antisymmetric. We have $\tau \circ \tau = \tau$. Antisymmetrization yields an antisymmetric element of the tensor product;

the antisymmetric elements of $V \otimes \cdots \otimes V$ form a subspace.

The discussion above permits one to construct elements of the tensor product $V \otimes \cdots \otimes V$ which "treat each V the same." One might expect, therefore, that such elements might be appropriate for the description of identical particles. The question is, Should one use the symmetric elements, the antisymmetric elements, or perhaps elements satisfying some other type of symmetry? It turns out that the appropriate description of systems of identical particles, for particles that actually occur in nature, requires either the use of symmetric elements or the use of antisymmetric elements (depending on the type of particle being considered), and that no "other types of symmetries" apparently are necessary. Denote by F_σ the subspace of vector space F consisting of elements of the direct sum each entry of which is a symmetric element of the corresponding tensor product, and by F_τ the subspace of F consisting of elements each entry of which is an antisymmetric element of $V \otimes \cdots \otimes V$. This F_σ is called the symmetric Fock space (on vector space V), F_τ the antisymmetric Fock space. Particles (e.g., pi-mesons, photons, gravitons) whose space of (many-particle) states is the symmetric Fock space F_σ are called bosons; particles (e.g., neutrinos, electrons) whose space of states is the antisymmetric Fock space F_τ are called fermions. Of course, whether a particle is to be a boson or a fermion must be decided by experiment.

We give one example of the difference between bosons and fermions. Let v be an element of V (a "one-particle state"), and consider the element $(0,0,v \otimes v,0,\cdots)$ of F (a "two-particle state, with both particles in state v"). Then the corresponding element of the antisymmetric Fock space is $(0,0,\tau(v \otimes v),0,\cdots) = (0,0,(1/2)(v \otimes v - v \otimes v),0,\cdots) = (0,0,0,0,\cdots)$. (The corresponding element of the symmetric Fock space is, of course, just $(0,0,v \otimes v,0,\cdots)$ itself.) This state of affairs is described by saying that "you cannot have a two-particle state, for fermions, with both particles in the same state." This is the Pauli exclusion principle.

We have now obtained the space of quantum states for our system of many identical particles (at least, once it has been decided whether those particles are bosons or fermions). We now wish to introduce a few linear mappings on these spaces of states.

Fix a complex vector space V, and let $F = \mathbf{C} \oplus V \oplus (V \otimes V) \oplus \cdots$ Thus a typical element of F is $f = (v_0,v_1,v_2,\cdots)$, where v_0 is a complex number, v_1 is an element of V, v_2 is an element of $V \otimes V$, etc. Denote by N the linear mapping from complex vector space F to F which sends this f to

$$N(f) = (0,1v_1,2v_2,3v_3,\cdots) \ .$$

Clearly, if f is in F_σ (i.e., in the symmetric Fock space on V), then so is $N(f)$; if f is in F_τ, then so is $N(f)$. Thus we obtain a linear mapping (which we also denote by N) from F_σ to F_σ, and another linear mapping N from F_τ to F_τ. These N are called the number (of particles) operators on the Fock spaces.

For f in N_σ, for example, we have $N(f) = 5f$ if and only if $f = (0,0,0,0,0,v_5,0, \cdots)$, that is, if and only if "f represents a state having precisely five particles."

Another class of linear mappings are those which "create or annihilate particles." Fix an element v of V. We introduce, using this v, a certain linear mapping C_v from F to F as follows: for $f = (v_0,v_1,v_2, \cdots)$ in F, set

$$C_v(f) = (0,\sqrt{1}v_0v,\sqrt{2}v \otimes v_1,\sqrt{3}v \otimes v_2, \cdots) \ .$$

Note that this is well defined; for example, v_2 is an element of $V \otimes V$, so $v \otimes v_2$ is an element of $V \otimes V \otimes V$, and this $v \otimes v_2$ indeed appears in the "$V \otimes V \otimes V$ entry" of $C_v(f)$. Of course, this C_v is a linear mapping (which, however, depends on the choice of v) from F to F. (The reason for the introduction of the numerical factors, $\sqrt{1},\sqrt{2}, \ldots$, in the definition of C_v depends on a Hilbert space structure on F, a structure we shall only introduce at a later stage.) We shall think of the action of this C_v as representing "creation of an additional particle in state v." This interpretation is reasonable: for $f = (v_0,0, \cdots)$ ("a vacuum state"), $C_v(f) = (0,v_0v,0, \cdots)$ ("a one-particle state, with the particle in state a multiple of v"); for $f = (0,v_1,0, \cdots)$ ("one particle, in state v_1"), $C_v(f) = (0,0,\sqrt{2}v \otimes v_1,0, \cdots)$ ("a two-particle state, with one particle in state v, the other in state v_1"), etc. There is, in fact, a little formula which reflects this interpretation of C_v (it "creates a particle"), namely $N \circ C_v - C_v \circ N = C_v$. (We check that this formula holds, e.g., on the state $f = (v_0,v_1,v_2,0, \cdots)$. We have $C_v(f) = (0,v_0v,\sqrt{2}v \otimes v_1,\sqrt{3}v \otimes v_2,0, \cdots)$, whence $N \circ C_v(f) = (0,v_0v,2\sqrt{2}v \otimes v_1,3\sqrt{3}v \otimes v_2,0, \cdots)$. But $N(f) = (0,v_1,2v_2,0, \cdots)$, whence $C_v \circ N(f) = (0,0,\sqrt{2}v \otimes v_1,2\sqrt{3}v \otimes v_2,0, \cdots)$. Thus $(N \circ C_v - C_v \circ N)(f) = (0,v_0v,\sqrt{2}v \otimes v_1,\sqrt{3}v \otimes v_2,0, \cdots)$. But this is precisely $C_v(f)$.) Note that our formula can also be written in the form

$$N \circ C_v = C_v \circ (N + I_F) \ ,$$

where I_F is the identity mapping on F. In this form, it says "counting the number of particles after applying C_v is the same as counting the number of particles before applying C_v and adding one."

Similarly, fix an element φ of V^*, the dual of V. We introduce, using this φ, a certain linear mapping A_φ from F to F. The rule is this: for $f = (a,0, \cdots)$ (a in $\mathbf{C}$), $A_\varphi(f) = (0,0, \cdots)$; for $f = (0,v,0, \cdots)$ (v in V), $A_\varphi(f) = (\sqrt{1}\varphi(v),0, \cdots)$; for $f = (0,0,v \otimes v',0, \cdots)$, $A_\varphi(f) = (0,\sqrt{2}\varphi(v)v',0, \cdots)$; for $f = (0,0,0,v \otimes v' \otimes v'',0, \cdots)$, $A_\varphi(f) = (0,0,\sqrt{3}\varphi(v)v' \otimes v'',0, \cdots)$; etc. (That is, "one lets the element φ of the dual of V seize the first element of a tensor product, to produce a complex number, which then is to multiply the remaining vectors in a tensor product.") Since this A_φ takes "a one-particle state to a vacuum state, a two-particle state to a one-particle state, etc.," we think of the action of A_φ as "annihilation of a particle (via the element φ of V^*)." This remark can be expressed symbolically by the formula

$$(N + I_F) \circ A_\varphi = A_\varphi \circ N \ .$$

Note that the information one needs to "know how to create a particle" is a knowledge of v, "the state the created particle is to be in," while the information one needs to "know how to annihilate a particle" is a knowledge of φ, "an element of the dual of V, which tells how much vacuum (namely, $\varphi(v)$) is to be produced from one-particle state v."

A few minor modifications are necessary to take into account the fact that our particles are actually identical. First, note that, for f in F_σ (resp., in F_τ), $A_\varphi(f)$ is also in F_σ (resp., in F_τ). Thus, for each element φ of V^*, we have a linear mapping, A_φ, from symmetric Fock space F_σ to F_σ, and a linear mapping (also written A_φ) from antisymmetric Fock space F_τ to F_τ. Things are not quite so simple for the creation operators. Fix v in V. Then, for example, for the element $f = (0,v_1,0, \cdots)$ of F_σ, $C_v(f) = (0,0,\sqrt{2}v \otimes v_1,0, \cdots)$ is not in general in F_σ. That is, creation of a particle (since it always "puts the created particle in the first entry of the tensor product") in general destroys symmetry or antisymmetry. To correct this, we simply symmetrize or antisymmetrize after application of C_v. Thus, for $f = (v_0, v_1, v_2, \cdots)$ in F_σ, we now set $C_v(f) = (0, v_0 v, \sqrt{2}\sigma(v \otimes v_1), \sqrt{3}\sigma(v \otimes v_2), \cdots)$, so $C_v(f)$ is again an element of F_σ. We thus obtain a linear mapping C_v from F_σ to F_σ. Similarly, for $f = (v_0, v_1, v_2, \cdots)$ in F_τ, we now set $C_v(f) = (0, v_0 v, \sqrt{2}\tau(v \otimes v_1), \sqrt{3}\tau(v \otimes v_2), \cdots)$. We thus obtain a linear mapping C_v from F_τ to F_τ. (This use of the same symbol for different mappings does not lead to confusion in practice, because one deals with just one type of particle at a time.)

To summarize, given the vector space V of "one-particle states for a boson particle," one can construct the Fock space F_σ of "many-particle states of these particles" together with linear mappings N ("number of particles"), C_v ("creation of a particle in state v"), and A_φ ("annihilation of a particle via φ of V^*") from F_σ to F_σ. Similarly for fermions.

We now claim that, in the boson case, these creation and annihilation operators satisfy the following commutation relations:

$$C_v \circ C_{v'} - C_{v'} \circ C_v = 0 \ ,$$

$$A_\varphi \circ C_v - C_v \circ A_\varphi = [\varphi(v)] I_{F_\sigma} \ ,$$

$$A_\varphi \circ A_{\varphi'} - A_{\varphi'} \circ A_\varphi = 0 \ .$$

The verification is straightforward: one applies both sides to a typical element f of F_σ and checks that the two sides agree. As an example, we check the second equation, applied to the element $f = (0,v_1,0, \cdots)$ of F_σ. We have

$C_v(f) = (0,0,(\sqrt{2}/2)(v \otimes v_1 + v_1 \otimes v),0, \cdots)$, whence $A_\varphi \circ C_v(f) = (0,\varphi(v)v_1 + \varphi(v_1)v,0, \cdots)$. But $A_\varphi(f) = (\varphi(v_1),0, \cdots)$, whence $C_v \circ A_\varphi(f) = (0,\varphi(v_1)v,0, \cdots)$. Subtracting, we have $(A_\varphi \circ C_v - C_v \circ A_\varphi)(f) = (0,\varphi(v)v_1,0, \cdots)$. But this is just $\varphi(v)(0,v_1,0, \cdots) = \varphi(v)f = [\varphi(v)]I_F(f)$. Similarly for other elements of F_σ and for the other equations just above. The equations above will be recognized as the standard commutation relations for the creation and annihilation operators for bosons.

In the fermion case, one has instead the following (called anticommutation) relations:

$$C_v \circ C_{v'} + C_{v'} \circ C_v = 0 \ ,$$

$$A_\varphi \circ C_v + C_v \circ A_\varphi = [\varphi(v)]I_F \ ,$$

$$A_\varphi \circ A_{\varphi'} + A_{\varphi'} \circ A_\varphi = 0 \ .$$

(Of course, nothing stops one from working out the left sides of the commutation expressions in the fermion case. What happens is that one obtains a complicated, and uninteresting, formula. It is only when one uses plus signs, as above, that one obtains a simple expression in the fermion case.) These are the standard anticommutation relations for creation and annihilation operators for fermions. Note, incidentally, that, setting $v = v'$ in the first formula, we obtain $C_v \circ C_v = 0$. "You get zero if you try to create two fermions in the same state." This, again, is essentially the Pauli exclusion principle.

We remark that the only subtle part of all this is at the beginning when one decides what mathematics to use, what spaces things are going to be in, etc. Everything is quite simple after that. In particular, one gets along quite well without any bases for one's vector spaces.

22

Representations: General Theory

In this chapter we introduce a general framework for the theory of representations. This framework is useful, it turns out, for a number of reasons: i) it provides a simple, broad, and easily remembered context into which the details of the subject fit; ii) the framework actually includes a number of other ideas that one does not normally associate with representation theory; iii) a significant fraction of the definitions and properties of representations in many common special cases fit easily into the general theory. Representations on vector spaces—the most important special case for physical applications—will be discussed in more detail in the next chapter.

The idea of a representation is to "represent an object as a collection of morphisms." In this way, one "makes the elements of the object more concrete—instead of being just abstract elements, they actually do something, namely act as morphisms." The setup is the following. Let **C** and **C**$'$ be two categories. We suppose that we are given the following two things: i) a forgetful functor **F** from the category **C** to the category of sets, and ii) a rule which assigns, to each object P' in category **C**$'$, an object Z in category **C** and an isomorphism (in the category of sets) from set **F**(Z) to the set Mor(P',P'). The situation is much simpler than the above suggests. There is an obvious forgetful functor from every category we shall consider to the category of sets, and we shall always use this one for item i). The purpose of i) is to allow us to speak of "elements" of objects in category **C**. Item ii) is just a fancy way of saying "we are to have available a procedure for introducing, on the set Mor(P',P'), structure so that it becomes an object in category **C**." We have already seen several examples of categories **C**$'$ such that each Mor(P',P') is more than just a set—such that additional structure appears naturally on Mor(P',P') so that it actually becomes an object in some category. We shall simply incorporate i) and ii) into our terminology as follows: we shall allow ourselves to speak of elements of objects in category **C**, and we shall simply regard Mor(P',P') (for P' an object in **C**$'$) as an object in category **C**.

For **C** and **C**$'$ as above, a *representation* of object A in **C** consists of an object P' in **C**$'$ together with a morphism $A \xrightarrow{\psi} \text{Mor}(P',P')$ (in category **C**). Thus, for each element a of object A, we must have a certain morphism, which we write ψ_a, from object P' to itself. This "rule ψ which assigns to each element a of A a morphism from P' to P' cannot in general be arbitrary, but must be such that ψ itself is a morphism in category **C**." Thus "the object A is represented as a bunch of morphisms (from P' to P'), with this representation (by which a in A goes to morphism ψ_a from P' to P') reflecting the structure within the object A." Some examples will make this clearer.

Example. Let **C** be the category of sets, and **C**$'$ the category of Lie algebras. Fix a set S. A representation of S (in the category of Lie algebras) consists of a Lie algebra P', together with a rule which assigns, to each element s of S, a homomorphism ψ_s from Lie algebra P' to Lie algebra P'.

Example. Let **C** be the category of associative algebras, and **C**$'$ the category of vector spaces. Fix an associative algebra V. A representation of V (in the category of vector spaces) consists of a vector space W, together with a rule which assigns, to each element v of associative algebra V, a linear mapping ψ_v from vector space W to W, subject to the following properties: $\psi_{(av+v')} = a\psi_v + \psi_{v'}$ (a a number), and $\psi_{vv'} = \psi_v \circ \psi_{v'}$. Note that, for W any vector space, Mor(W,W) has the structure of an associative algebra. Thus the two properties above are precisely the statement that $V \xrightarrow{\psi}$ Mor(W,W) be a homomorphism of associative algebras. The associative algebra V is "represented as linear mappings on vector space W."

Example. Let G be a group. A representation of G (in the category of vector spaces) consists of a vector space W, together with a rule which assigns, to each element g of G, a linear mapping ψ_g from vector space W to W, such that $\psi_{gg'} = \psi_g \circ \psi_{g'}$ and $\psi_e = i_W$. This is also a special case of the definition of a representation (exercise 137).

As one might expect, there are a number of ways to combine and manipulate representations. We give a few examples. Fix categories **C** and **C**$'$ as above. Let A be a fixed object of category **C**, and consider two representations of A, $A \xrightarrow{\psi}$ Mor(P',P') and $A \xrightarrow{\varphi}$ Mor(Q',Q'). Thus, for each a in A, we have $P' \xrightarrow{\psi_a} P'$ and $Q' \xrightarrow{\varphi_a} Q'$. Using the definition of the direct product and direct sum in the diagrams of figure 55, we obtain a morphism γ_a from the direct product of P' and Q' to itself, and a morphism δ_a from the direct sum of P' and Q' to itself. It is normally the case that $A \xrightarrow{\gamma}$ Mor(Prod,Prod) and $A \xrightarrow{\delta}$ Mor(Sum,Sum) are both also representations of A. The former is called the ***direct product representation,*** the latter the ***direct sum representation*** (of the representations $A \xrightarrow{\psi}$ Mor(P',P') and $A \xrightarrow{\varphi}$ Mor(Q',Q')). Thus, from two representations, one normally obtains, by taking direct products or direct sums, two others.

Again, let A be a fixed object in **C**, and let $A \xrightarrow{\psi}$ Mor(P',P') be a representation of A. Let $K' \xrightarrow{\tau} P'$ be a subobject of P', so τ is a monomor-

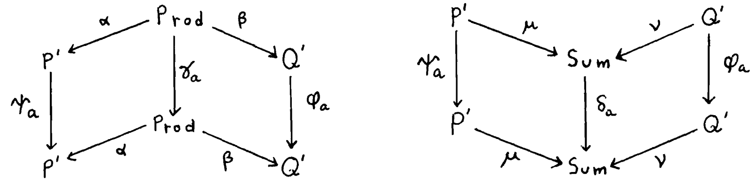

Figure 55

phism. We wish to obtain a representation of A on K'. However, this will not always be possible, because, intuitively, "it may happen that, for some a in A, ψ_a takes elements of the subobject K' out of K'." We proceed therefore as follows. The subobject K' is said to be an *invariant subobject* if, for every a in A, there is a λ_a such that the diagram of figure 56 commutes.

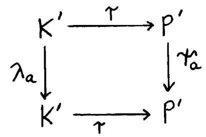

Figure 56

(Note that this λ_a, if it exists, is unique, for the τ on the bottom line is a monomorphism.) Thus, for K' an invariant subobject, we obtain, for each a in A, $K' \xrightarrow{\lambda_a} K'$. This is normally a representation of A on K'. Thus a representation $A \xrightarrow{\psi} \mathrm{Mor}(P',P')$, and an invariant subobject, $K' \xrightarrow{\tau} P'$, normally leads to another representation (called a *subrepresentation*), $A \xrightarrow{\lambda} \mathrm{Mor}(K',K')$.

Example. Let $A \xrightarrow{\psi} \mathrm{Mor}(P',P')$ and $A \xrightarrow{\varphi} \mathrm{Mor}(Q',Q')$ be representations. Consider their direct sum representation, $A \xrightarrow{\delta} \mathrm{Mor(Sum,Sum)}$. The morphism $P' \xrightarrow{\mu} \mathrm{Sum}$ is normally a monomorphism, whence P' is a subobject of Sum. This subobject is easily checked to be invariant. Hence $A \xrightarrow{\psi} \mathrm{Mor}(P',P')$ is a subrepresentation of $A \xrightarrow{\delta} \mathrm{Mor(Sum,Sum)}$.

Example. Let S be any set, and let V be the vector space of all real-valued functions on set S. For s in S, let ψ_s be the linear mapping from V to V that sends v in V (so v is a function on S) to the function whose value at s is $v(s)$ and whose value elsewhere is zero. This $S \xrightarrow{\psi} \mathrm{Mor}(V,V)$ is a representation of S (in the category of real vector spaces). Let W be the subspace of V

consisting of functions on S that vanish for all but a finite number of elements of S. This is an invariant subobject (for, for each w in W, $\psi_s(w)$ is in the subspace W). Thus we have a subrepresentation, $S \xrightarrow{\varphi} \text{Mor}(W, W)$.

In the constructions above, we were concerned with "manipulations on the object on which the representation acts." We next consider "manipulations of the object represented."

Let A, again, be an object in category $\mathbf{C}$, and let $A \xrightarrow{\psi} \text{Mor}(P', P')$ be a representation of A. Let B be any other object in $\mathbf{C}$, and let $B \xrightarrow{\kappa} A$ be a morphism. Then, of course, $B \xrightarrow{\psi \circ \kappa} \text{Mor}(P', P')$ is also a representation of B. Thus a morphism from B to A, together with a representation of A, leads immediately to a representation of B. In particular, B might be a subobject of A.

Next, let $A \xrightarrow{\psi} \text{Mor}(P', P')$ and $B \xrightarrow{\varphi} \text{Mor}(P', P')$ be representations of A and B (both on P'). Write $A \oplus B$ for the direct sum of A and B. Then, by definition of the direct sum, there is a unique morphism ς in the diagram of figure 57 such that the diagram commutes. This $A \oplus B \xrightarrow{\varsigma} \text{Mor}(P', P')$ is,

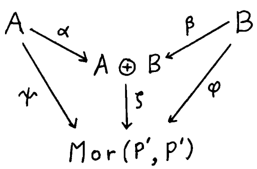

Figure 57

of course, a representation of $A \oplus B$. Thus, from a representation of A and a representation of B (on the same P'), we get a representation of $A \oplus B$.

Example. Let S and T be disjoint sets, and let $S \xrightarrow{\psi} \text{Mor}(V, V)$ and $T \xrightarrow{\varphi} \text{Mor}(V, V)$ be representations of S and T (say, in the category of complex vector spaces). Then we have a representation $S \cup_d T \xrightarrow{\varsigma} \text{Mor}(V, V)$, where $\varsigma_s = \psi_s$ for s in S and $\varsigma_t = \varphi_t$ for t in T. Now regard S as a subset of $S \cup_d T$. Then we recover from $S \cup_d T \xrightarrow{\varsigma} \text{Mor}(V, V)$, by the first construction above, the original representation of S, $S \xrightarrow{\psi} \text{Mor}(V, V)$.

Exercise 135. Fix categories $\mathbf{C}$ and $\mathbf{C}'$ as above, and fix an object A of category $\mathbf{C}$. Let the objects consist of representations, $A \xrightarrow{\psi} \text{Mor}(P', P')$, of A in category $\mathbf{C}$. Let the morphisms be as follows: for $A \xrightarrow{\psi} \text{Mor}(P', P')$ and

$A \xrightarrow{\varphi} \mathrm{Mor}(Q', Q')$ objects, a morphism is a morphism $P' \xrightarrow{\tau} Q'$ (in category $\mathbf{C}'$) such that $\tau \circ \psi_a = \varphi_a \circ \tau$ for each element a of A. Prove that this is a category (the category of representations of A in $\mathbf{C}'$). Show that the constructions of this chapter are just those of direct product, direct sum, and subobjects in this category.

Exercise 136. Let $\mathbf{C}$ and $\mathbf{C}'$ both be the category of sets. Prove that a representation of S, $S \xrightarrow{\psi} \mathrm{Mor}(T, T)$, is the same thing as a mapping of sets, $S \times T \rightarrow T$.

Exercise 137. A semigroup with unit is a set on which there is given a product that is associative and has an identity. Define morphisms, and obtain the category of semigroups with unit. Show that, for any object P' in category $\mathbf{C}'$, $\mathrm{Mor}(P', P')$ has the structure of a semigroup with unit (the "product" is composition of morphisms). Thus we know what it means to talk about a representation of a semigroup with unit in any category. Note that there is a forgetful functor from the category of groups to the category of semigroups with unit. Now complete the third example in this chapter.

Exercise 138. A representation $A \xrightarrow{\psi} \mathrm{Mor}(P', P')$ is said to be faithful if ψ is a monomorphism. Are subrepresentations of a faithful representation faithful?

Exercise 139. Let V be a vector space, and $V \xrightarrow{\sigma} U$ the free associative algebra on V. For v in V, let ψ_v be the mapping from U to U which sends u in U to $\sigma(v)u$. Is this a representation of V on U?

Representations on Vector Spaces

It turns out that the most important representations for applications are those in which the category **C**′ is that of (usually complex) vector spaces. (The reason, I guess, is that the space of states of a quantum system is a complex vector space.) In fact, the term "representation" is often taken to mean a representation on a vector space. We here discuss some properties of representations on vector spaces.

What sort of objects will have representations on vector spaces? It all depends, as we have seen in the previous chapter, on what sort of structure one can discover on the set Mor(V, V), for V any vector space. But we have already pretty well decided that issue: Mor(V, V) has the structure of an associative algebra. (The vector-space structure of Mor(V, V) reflects the fact that any linear combination of linear mappings from V to V is again a linear mapping from V to V; the product structure on Mor(V, V) is composition of linear mappings.) Thus, in particular, any category one can obtain via a forgetful functor from the category of associative algebras will be a category whose objects can have representations on vector spaces. (This is, of course, because Mor(V, V) can be regarded as an object in any such category.) It turns out that, among all the possibilities, only three are of practical interest: groups (see exercise 137), Lie algebras, and associative algebras. Thus, specializing from the previous chapter, we have:

1. Let G be a group. A representation of G (on a vector space) consists of a vector space V, together with a rule which assigns, to each g in G, a linear mapping $V \xrightarrow{\psi_g} V$ such that $\psi_{gg'} = \psi_g \circ \psi_{g'}$ and $\psi_e = i_V$.

2. Let U be a Lie algebra. A representation of U (on a vector space) consists of a vector space V, together with a rule which assigns, to each u in U, a linear mapping $V \xrightarrow{\psi_u} V$ such that $\psi_{(au+u')} = a\psi_u + \psi_{u'}$ and $\psi_{[u,u']} = \psi_u \circ \psi_{u'} - \psi_{u'} \circ \psi_u$.

3. Let T be an associative algebra. A representation of T (on a vector space) consists of a vector space V, together with a rule which assigns, to each t in T, a linear mapping $V \xrightarrow{\psi_t} V$ such that $\psi_{(at+t')} = a\psi_t + \psi_{t'}$ and $\psi_{tt'} = \psi_t \circ \psi_{t'}$.

It is immediate that, for a representation of a group, each ψ_g is invertible (for $\psi_g \circ \psi_{g^{-1}} = \psi_{gg^{-1}} = \psi_e = i_V$). For a representation of a Lie or associative algebra, this is not in general true. (In fact, we have immediately in this case that ψ_0 is the zero linear mapping.)

Consider two representations (say, of a group G), $G \xrightarrow{\psi} \text{Mor}(V,V)$ and $G \xrightarrow{\varphi} \text{Mor}(W,W)$. Consider the direct sum of vector spaces, $V \oplus W$. With each g in G, associate the linear mapping from $V \oplus W$ to $V \oplus W$ which sends (v,w) to $(\psi_g(v),\varphi_g(w))$. This is clearly a representation of G on $V \oplus W$, a particular case of the direct sum of general representations. (Similarly, of course, for representations of associative or Lie algebras.) Since direct products, in the category of vector spaces, are the same as direct sums, that yields nothing new. However, there is, in the case of vector spaces, an additional way to combine representations: using the tensor product. Consider again our two representations of G, $G \xrightarrow{\psi} \text{Mor}(V,V)$ and $G \xrightarrow{\varphi} \text{Mor}(W,W)$. Consider the tensor product, $V \otimes W$, of vector spaces. For g in G, let κ_g be the linear mapping from $V \otimes W$ to $V \otimes W$ which sends $v \otimes w$ to $\psi_g(v) \otimes \varphi_g(w)$. (Of course, these instructions determine the linear mapping completely, since every element of $V \otimes W$ is a linear combination of elements of the form $v \otimes w$.) Clearly, we have a representation, $G \xrightarrow{\kappa} \text{Mor}(V \otimes W, V \otimes W)$. This representation is called the ***tensor product representation*** of our two representations. (Similarly for representations of Lie or associative algebras.) Finally, as a special case of the previous chapter, a subspace W of vector space V is an invariant subspace of the representation $G \xrightarrow{\psi} \text{Mor}(V,V)$ (say, of a group) if, for each g in G and w in W, $\psi_g(w)$ is also in W. For W an invariant subspace, we obtain a subrepresentation, $G \xrightarrow{\mu} \text{Mor}(W,W)$, of G on W (since each ψ_g maps W to W, whence each ψ_g can be regarded as an element of $\text{Mor}(W,W)$).

Example. Let G be a group, V a vector space, and $G \xrightarrow{\psi} \text{Mor}(V,V)$ a representation. Consider the tensor product of this representation with itself, $G \xrightarrow{\kappa} \text{Mor}(V \otimes V, V \otimes V)$. Denote by K the subspace of $V \otimes V$ generated by elements of the form $v \otimes v' + v' \otimes v$, and by L the subspace generated by elements of the form $v \otimes v' - v' \otimes v$. Then, evidently, K and L are complementary subspaces of $V \otimes V$. Next, note that K is an invariant subspace of the representation $G \xrightarrow{\kappa} \text{Mor}(V \otimes V, V \otimes V)$, for $\kappa_g(v \otimes v' + v' \otimes v) = \psi_g(v) \otimes \psi_g(v') + \psi_g(v') \otimes \psi_g(v)$ is also in K. Similarly, L is an invariant subspace. Thus we have subrepresentations $G \xrightarrow{\mu} \text{Mor}(K,K)$ and $G \xrightarrow{\nu} \text{Mor}(L,L)$. Since K and L are complementary in $V \otimes V$, there is a natural isomorphism of vector spaces from $V \otimes V$ to $K \oplus L$. Clearly, therefore, the direct sum representation, $G \xrightarrow{\tau} \text{Mor}(K \oplus L, K \oplus L)$ is essentially the same as the representation $G \xrightarrow{\kappa} \text{Mor}(V \otimes V, V \otimes V)$.

Example. Let $G \xrightarrow{\psi} \text{Mor}(V, V)$ be a representation. For each g in G, let $V^* \xrightarrow{\gamma} V^*$ be given by $\gamma_g = [\psi_{g^{-1}}]^*$ (adjoint mapping on the dual). Then this is also a representation, for $\gamma_g \circ \gamma_{g'} = [\psi_{g^{-1}}]^* \circ [\psi_{g'^{-1}}]^* = [\psi_{g'^{-1}} \circ \psi_{g^{-1}}]^* = [\psi_{g'^{-1}g^{-1}}]^* = [\psi_{(gg')^{-1}}]^* = \gamma_{gg'}$. (Note that, in this example, we actually require that G be a group, for we need a g^{-1}.) Take the tensor product of the two, $G \xrightarrow{\kappa} \text{Mor}(V \otimes V^*, V \otimes V^*)$. Now denote by λ the linear mapping from $V \otimes V^*$ to the numbers that sends $v \otimes \varphi$ (so φ is in V^*) to the number $\varphi(v)$. Then Ker (λ) is an invariant subspace of the representation $G \xrightarrow{\kappa} \text{Mor}(V \otimes V^*, V \otimes V^*)$.

There are a number of ways to get from one to another of groups, Lie algebras, and associative algebras, many of which lead to ways to get from one representation to another.

Let G be a group, and let U be the group algebra of G (chapter 18). Thus, with each g in G there is associated an element, $\alpha(g)$, of U, and the most general element of U is a linear combination of such elements: $a_1\alpha(g_1) + \cdots + a_n\alpha(g_n)$, with $a_1, \ldots, a_n$ numbers. The product in U is $\alpha(g)\alpha(g') = \alpha(gg')$, extended to all of U by linearity. Now let $G \xrightarrow{\psi} \text{Mor}(V, V)$ be a representation of G. We introduce a representation of the associative algebra U. For $u = a_1\alpha(g_1) + \cdots + a_n\alpha(g_n)$ in U, set $\gamma_u = a_1\psi_{g_1} + \cdots + a_n\psi_{g_n}$, so γ_u is a certain linear mapping from V to V. This is clearly a representation of U. One often uses essentially this idea to obtain a representation of a given group G. Let G be given, and let U be the group algebra of G. For g in G, let φ_g be the linear mapping from U to U which sends $a_1\alpha(g_1) + \cdots + a_n\alpha(g_n)$ to $a_1\alpha(gg_1) + \cdots + a_n\alpha(gg_n)$. One thus obtains a representation $G \rightarrow \text{Mor}(U, U)$ (called the regular representation of G).

Next, let U be any associative algebra, and denote by $\mathbf{F}(U)$ the corresponding Lie algebra (via the forgetful functor). Consider a representation $U \xrightarrow{\psi} \text{Mor}(V, V)$ of U. Then, applying the forgetful functor to both of the associative algebras U and $\text{Mor}(V, V)$, we have a morphism $\mathbf{F}(U) \xrightarrow{\varphi = \mathbf{F}(\psi)} \mathbf{F}(\text{Mor}(V, V))$. Thus we obtain a representation of the Lie algebra $\mathbf{F}(U)$. In this representation, for example, $\varphi_{[u, u']} = \psi_{(uu' - u'u)} = \psi_u \circ \psi_{u'} - \psi_{u'} \circ \psi_u$.

Finally, let L be a Lie algebra, and U its universal enveloping algebra. Consider a representation of L, $L \xrightarrow{\psi} \text{Mor}(V, V)$. Then, by definition of the universal enveloping algebra, there is a unique homomorphism τ of associative algebras such that the diagram of figure 58 commutes. Thus we obtain $U \xrightarrow{\tau} \text{Mor}(V, V)$, a representation of the universal enveloping algebra of L. This little construction is essentially the reason why one introduces the universal enveloping algebra. "Mor(V, V) really wants to be an associative algebra, but

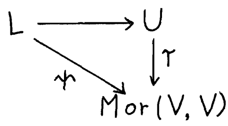

Figure 58

we can force it to be a Lie algebra by applying the forgetful functor. Then, we have the notion of a representation of a Lie algebra. By 'freeing' an associative structure on the Lie algebra L, one obtains its universal enveloping algebra, the thing that really should have been represented on V in the first place."

Example. Let L be a three-dimensional real vector space, with basis x, y, and z. Define a bracket on L by $[x,y] = -[y,x] = z$, $[y,z] = -[z,y] = x$, $[z,x] = -[x,z] = y$, extending to all of L by linearity. Then, noting that the Jacobi relation is satisfied, we have on L the structure of a Lie algebra. The free associative algebra on vector space L is the algebra of "all polynomials in x,y,z, having no constant term, where order of factors in each term is relevant." The universal enveloping algebra, U, of L is the algebra of "all polynomials in x,y,z, having no constant term, where the order of factors is relevant, but in which one may substitute $xy - yx$ for z, $yz - zy$ for x, and $zx - xz$ for y." (Thus, in U, $xxyz = xyxz + xzz$.) Now suppose that we have a representation, $L \xrightarrow{\psi} \mathrm{Mor}(V,V)$, of L. Then, in particular, since x, y, and z are elements of L, we have ψ_x, ψ_y, and ψ_z, linear mappings from V to V. Since ψ is a representation of our Lie algebra, $\psi_x \circ \psi_y - \psi_y \circ \psi_x = \psi_z$, $\psi_y \circ \psi_z - \psi_z \circ \psi_y = \psi_x$, and $\psi_z \circ \psi_x - \psi_x \circ \psi_z = \psi_y$. Clearly, this representation is uniquely determined by ψ_x, ψ_y, and ψ_z. The corresponding representation of the universal enveloping algebra U is this. For, for example, $u = xzyx + 2yz$ in U, set $\gamma_u = \psi_x \circ \psi_z \circ \psi_y \circ \psi_x + 2\psi_y \circ \psi_z$. This actually is a representation of U. For example, $xxyz = xyxz + xzz$ in U (since $xy - yx = z$ in U), and, indeed, $\psi_x \circ \psi_x \circ \psi_y \circ \psi_z = \psi_x \circ \psi_y \circ \psi_x \circ \psi_z + \psi_x \circ \psi_z \circ \psi_z$ (since $\psi_x \circ \psi_y - \psi_y \circ \psi_x = \psi_z$). Note what has happened: "since our original representation was of the Lie algebra L, the bracket structure of L was reflected in the ψ. This bracket structure is also incorporated into the construction of the universal enveloping algebra U. Everything works out just right so that 'polynomials which are the same as elements of U lead to the same element of $\mathrm{Mor}(V,V)$.' "

Note that any representation (of a group, Lie algebra, associative algebra, etc.) on a vector space V always has two invariant subspaces: the subspace of V consisting only of the zero element, and V itself. A representation in which these two are the only invariant subspaces is said to be ***irreducible***. **In many situations, one can reduce the study of all representations to the study of** irreducible representations by the following technique. Let $G \xrightarrow{\psi} \mathrm{Mor}(V, V)$ be a representation (say, of a group). If this representation is not irreducible, there is some invariant subspace W (neither V nor the zero subspace). Suppose one managed to find a subspace U of V which was complementary to W and which was also invariant. Then we would have representations of G on W and U, and our original representation would be a direct sum of these (since V is naturally isomorphic to $W \oplus U$). If, for example, the representation on W is not irreducible, then there is an invariant subspace. If, again, one can find a complementary subspace, one can repeat the process. With any luck, one will end up writing the original representation as a direct sum of irreducible ones ("keep decomposing in this way until no further decomposition is possible, i.e., one hopes, until one has irreducible representations"). Unfortunately, it is not true in general that one can always find a complementary invariant subspace in order to continue the "decomposition process" above. (Although certain situations are known for which a complementary invariant subspace can always be found—e.g., when G has only a finite number of elements—there is no simple criterion, as far as I know, in the general case.) Fortunately, it turns out that many of the representations one meets in practice are in fact a direct sum of irreducible ones (actually, most are already irreducible). Thus irreducible representations are interesting.

One of the crucial facts about irreducible representations is this.

THEOREM 26 *(Schur). Let* $S \xrightarrow{\psi} \mathrm{Mor}(V, V)$ *be an irreducible representation (of anything) on a finite-dimensional, complex vector space* V. *Let* $V \xrightarrow{\alpha} V$ *be a linear mapping such that* $\alpha \circ \psi_s = \psi_s \circ \alpha$ *for every* s *in* S. *Then* $\alpha = ai_V$ *for some number* a.

Proof. Since V is finite-dimensional and complex, there exists an eigenvector of α, that is, there exists a nonzero $\underline{v}$ in V and number a such that $\alpha(\underline{v}) = a\underline{v}$. (This follows because $\det(\alpha - aI) = 0$ must have complex root a.) Denote by W the subspace of V consisting of all v in V such that $\alpha(v) = av$ (fixing this particular number a). Then W is an invariant subspace: for w in W (so $\alpha(w) = aw$), we have $\alpha(\psi_s(w)) = \psi_s(\alpha(w)) = \psi_s(aw) = a[\psi_s(w)]$, whence $\psi_s(w)$ is in W. (We used $\alpha \circ \psi_s = \psi_s \circ \alpha$ in the first step.) Since our representation is irreducible (and since W cannot be the zero subspace, for it contains $\underline{v} \neq 0$), we must have $W = V$. That is, $\alpha(v) = av$ for every v in V, or, what is the same thing, $\alpha = ai_V$. $\square$

The conclusion of theorem 26 holds in many cases more general than that of the hypothesis (in particular, for certain infinite-dimensional V).

We give an example of a use of theorem 26. Let L be a Lie algebra, and U its universal enveloping algebra. Consider an irreducible representation of

L, $L \xrightarrow{\psi} \mathrm{Mor}(V,V)$, on a finite-dimensional, complex V. Then we obtain a representation of U, $U \xrightarrow{\gamma} \mathrm{Mor}(V,V)$. Now let $\underline{u}$ be an element of the Casimir subalgebra of U (chapter 19). Then, for each u in U, we have $\gamma_{\underline{u}} \circ \gamma_u = \gamma_u \circ \gamma_{\underline{u}}$ (since $\underline{u}u = u\underline{u}$ in U). By theorem 26 there is a number a such that $\gamma_{\underline{u}} = ai_V$. Thus, in this representation, $\underline{u}$ is represented as a numerical multiple of the identity linear mapping. Thus one can classify the irreducible representations of L (at least, complex ones on a finite-dimensional vector space) by the "numerical values" taken by the elements of the Casimir subalgebra. The attractive feature of this situation is that the universal enveloping algebra U and the Casimir subalgebra are already known once one knows L—they do not refer directly to the representation being considered. (In physics, "spin" and "mass" arise in just this way.)

Example. In the prior example, consider the element $\underline{u} = xx + yy + zz$ of U. Then this $\underline{u}$ is in the Casimir subalgebra, for, for example, $\underline{u} - \underline{u}x = x(xx + ' + zz) - (xx + ' + zz)x = xyy + xzz - yyx - zzx = yxy + zy + zxz - yz - yyx - zzx = yyx + yz + zy + zzx - zy - yz - yyx - zzx = 0$. Thus, given any irreducible representation of L on a finite-dimensional, complex vector space V, we have $\gamma_{\underline{u}} = \psi_x \circ \psi_x + \psi_y \circ \psi_y + \psi_z \circ \psi_z$ a multiple of the identity.

Exercise 140. Let U be an associative algebra, and write $\underline{U}$ for vector space U. Consider $U \xrightarrow{\varphi} \mathrm{Mor}(\underline{U},\underline{U})$ with the following action: for u in U, φ_u sends $\underline{u}$ in $\underline{U}$ to $u\underline{u}$. Prove that this is a representation. Prove that every ideal of U, regarded as a subspace of $\underline{U}$, is an invariant subspace.

Exercise 141. Let G be the additive group of reals, V a two-dimensional vector space with basis x,y. For g in G, let ψ_g send x to x and y to $y + gx$ (g is just a real number). Prove that the subspace generated by x is invariant and that there is no complementary invariant subspace.

Exercise 142. Let G be the additive group of reals, and V a real two-dimensional vector space with basis x,y. For g in G, let ψ_g send x to $x \cos g + y \sin g$ and y to $y \cos g - x \sin g$. Prove that this is an irreducible representation. Find $V \xrightarrow{\alpha} V$ such that $\alpha \circ \psi_g = \psi_g \circ \alpha$, but with α not a multiple of the identity.

Exercise 143. Are direct sums, direct products, and subrepresentations of irreducible representations irreducible?

Exercise 144. Let V be a vector space, and regard $\mathrm{Mor}(V,V)$ as an associative algebra. Then the identity mapping from $\mathrm{Mor}(V,V)$ to $\mathrm{Mor}(V,V)$ is a representation of $\mathrm{Mor}(V,V)$ on V. Prove that this representation is irreducible.

Exercise 145. Let S be any set, V the free vector space on S, and G the permutation group of set S. Then, for g in G, let ψ_g be the mapping from V to V which sends v (a function on S) to the function $v(g(s))$. Prove that this is a representation of G on V. For which S is this representation irreducible?

Exercise 146. Let $G \xrightarrow{\psi} \mathrm{Mor}(V,V)$ be an irreducible representation, and fix nonzero v in V. Show that the subset of V consisting of elements of the form $\psi_g(v)$ (g in G) spans V.

Exercise 147. Prove that, if a finite-dimensional, complex V is the vector space for an irreducible representation of an abelian group, then V is one-dimensional or zero-dimensional.

24

The Algebraic Categories: Summary

We have now completed our study of things "whose only structure is purely algebraic." There are many other categories which would also fall under this description—for example, semigroups, lattices, rings, integral domains, modules, fields—but which we have omitted because they are less important for applications.

The forgetful functors provide an easy way to get an overall picture of the nine algebraic categories we have discussed. In figure 59, the solid arrows

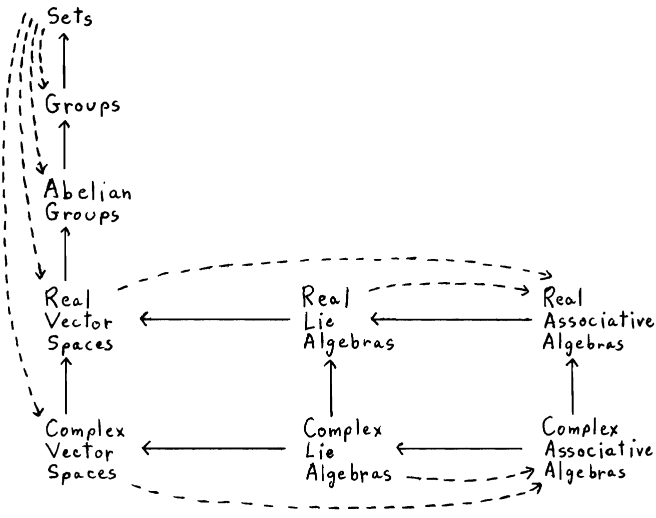

Figure 59

represent the action of the forgetful functors while the dashed arrows represent the free constructions we have introduced. (Other forgetful functors are obtained by composition.)

Other assorted properties of these categories are in the table below.

Category	Quotient Set for	Quotient Object for	Ker(φ)	Im(φ)	Mor(A,B)	Remark
Abelian Groups	Subgroup	Subgroup	Subgroup	Subgroup	Abelian Group	Sum = Product
Groups	Subgroup	Normal Subgroup	Normal Subgroup	Subgroup	Set	
Vector Spaces	Subspace	Subspace	Subspace	Subspace	Assoc. Algebra	Tensor Product
Lie Algebras	Sub-algebra	Ideal	Ideal	Sub-algebra	Assoc. Algebra	
Assoc. Algebras	Sub-algebra	Ideal	Ideal	Sub-algebra	Assoc. Algebra	

25

Subsets and Mappings

It is convenient at this point to introduce some definitions and notation, and to recall a few facts about sets, subsets, and mappings. It is, in my opinion, a waste of time to try to memorize all the various properties listed below. Rather, one should try to become skillful at guessing with reasonable accuracy what is true, and to become adept at quickly finding the (always easy) proofs and at effortlessly manipulating the symbols.

Fix a set S. We denote by $\emptyset$, the ***empty set,*** the subset of S having no elements. For A and B subsets of S, we write $A \subset B$ or $B \supset A$ if every element of A is also an element of B. In this case, one says that A is a ***subset*** of B and that B is a ***superset*** of A. (Note that $A = B$ is not excluded.) Thus $\emptyset \subset B$ for all B. For A a subset of S, denote by A^c, called the ***complement*** of A (in S), the set consisting of all elements of S which are not in A. Then $(A^c)^c = A$ for any subset A of S, while $\emptyset^c = S$ and $S^c = \emptyset$. Furthermore, $A \subset B$ if and only if $A^c \supset B^c$. For A and B subsets of S, one writes $A - B$ for the set of all elements of S which are elements of A and not elements of B. Thus $A - B = A \cap B^c$, while $A - A = \emptyset$ and $A - \emptyset = A$. Finally, let A_λ (λ in Λ) be a collection of subsets of S. Then $(\bigcup_\Lambda A_\lambda)^c = \bigcap_\Lambda A_\lambda^c$ [proof: s (in S) is in $(\bigcup_\Lambda A_\lambda)^c$ if and only if s is not in $\bigcup_\Lambda A_\lambda$ if and only if s is in no A_λ if and only if s is in every A_λ^c if and only if s is in $\bigcap_\Lambda A_\lambda^c$], and, similarly, $(\bigcap_\Lambda A_\lambda)^c = \bigcup_\Lambda A_\lambda^c$.

Now let S and T be two sets, and let $S \xrightarrow{\varphi} T$ be a mapping of sets. For A a subset of S, one writes $\varphi[A]$, called the ***image*** of A (by φ), for the set of all t in T such that $\varphi(a) = t$ for some a in A. For B a subset of T, one writes $\varphi^{-1}[B]$, called the ***inverse image*** of B (by φ), for the set of all s in S with $\varphi(s)$ in B. Note that $\varphi[A]$ is a subset of T, while $\varphi^{-1}[B]$ is a subset of S. Thus $\varphi^{1-}[T] = S$, while it is false in general that $\varphi[S] = T$. This notation can be confusing: for example, it is false in general that $\varphi[\varphi^{-1}[B]] = B$ or that $\varphi^{-1}[\varphi[A]] = A$. (Let S have just two elements, s and s', and T have just two elements, t and t'. Let $\varphi(s) = \varphi(s') = t$. Let A be the subset of S consisting of s'. Then $\varphi^{-1}[\varphi[A]] = S \neq A$. Let B be the subset of T consisting of t'. Then $\varphi[\varphi^{-1}[B]] = \emptyset \neq B$.)

The behavior of complements under inverse images is this: for B a subset of T, we have $\varphi^{-1}[B^c] = (\varphi^{-1}[B])^c$. [Proof: s (in S) is in $\varphi^{-1}[B^c]$ if and only if $\varphi(s)$ is in B^c if and only if $\varphi(s)$ is not in B if and only if s is not in $\varphi^{-1}[B]$ if and only if s is in $(\varphi^{-1}[B])^c$.] Complements do not behave so well under images: for example, it is false in general that $\varphi[A^c] = (\varphi[A])^c$. (For example, let S have just three elements, s, s', and s'', and let T have just three elements, t, t', and t''. Let $\varphi(s) = \varphi(s') = t'$ and $\varphi(s'') = t''$. Let A be the subset of S consisting of s' and s''. Then $\varphi[A^c]$ consists only of t', while $(\varphi[A])^c$ consists only of t.)

Next, we consider the behavior of unions and intersections under images and inverse images. Let B_λ (λ in Λ) be a collection of subsets of T. Then $\varphi^{-1}[\bigcap_\Lambda B_\lambda] = \bigcap_\Lambda \varphi^{-1}[B_\lambda]$. [Proof: s (in S) is in $\varphi^{-1}[\bigcap_\Lambda B_\lambda]$ if and only if $\varphi(s)$ is in $\bigcap_\Lambda B_\lambda$ if and only if $\varphi(s)$ is in every B_λ if and only if s is in every $\varphi^{-1}[B_\lambda]$ if and only if s is in $\bigcap_\Lambda \varphi^{-1}[B_\lambda]$.] Similarly, $\varphi^{-1}[\bigcup_\Lambda B_\lambda] = \bigcup_\Lambda \varphi^{-1}[B_\lambda]$. To consider the behavior under images, let A_λ (λ in Λ) be a collection of subsets of S. Then $\varphi[\bigcup_\Lambda A_\lambda] = \bigcup_\Lambda \varphi[A_\lambda]$. [Proof: t (in T) is in $\varphi[\bigcup_\Lambda A_\lambda]$ if and only if $t = \varphi(s)$ for some s in $\bigcup_\Lambda A_\lambda$ if and only if $t = \varphi(s)$ for some s in some A_λ if and only if t is in $\varphi[A_\lambda]$ for some A_λ if and only if t is in $\bigcup_\Lambda \varphi[A_\lambda]$.] On the other hand, it is false in general that $\varphi[\bigcap_\Lambda A_\lambda] = \bigcap_\Lambda \varphi[A_\lambda]$. (For example, let S have elements s and s', and T elements t and t'. Let $\varphi(s) = \varphi(s') = t$. Let A consist of s, and A' consist of s'. Then $\varphi[A \cap A'] = \emptyset$, while $\varphi[A] \cap \varphi[A']$ consists of t.)

Finally, let $S \xrightarrow{\varphi} T \xrightarrow{\psi} U$ be mappings of sets. Then, for A a subset of S, $(\psi \circ \varphi)[A] = \psi[\varphi[A]]$ and, for C a subset of U, $(\psi \circ \varphi)^{-1}[C] = \varphi^{-1}[\psi^{-1}[C]]$.

Exercise 148. Does $(A - A')^c = A'^c - A^c$?

Exercise 149. Does $\varphi^{-1}[B - B'] = \varphi^{-1}[B] - \varphi^{-1}[B']$? Does $\varphi[A - A'] = \varphi[A] - \varphi[A']$?

Exercise 150. Does $A \subset A'$ imply $\varphi[A] \subset \varphi[A']$? Does $B \subset B'$ imply $\varphi^{-1}[B] \subset \varphi^{-1}[B']$?

Exercise 151. Let $S \xrightarrow{\varphi} T$ be a mapping of sets. Prove that φ is a monomorphism if and only if $\varphi^{-1}[\varphi[A]] = A$ for every subset A of S. Prove that φ is an epimorphism if and only if $\varphi[\varphi^{-1}[B]] = B$ for every subset B of T.

Topological Spaces

We now begin the study of topological spaces. As we shall see, the viewpoint is somewhat different in topology than it was in the study of groups, vector spaces, etc. If one had to characterize this difference in terms of a single feature, perhaps it would be this: whereas in the algebraic categories one is concerned principally with elements (of sets), one is in topology more concerned with subsets. Nonetheless, a number of the ideas used in the algebraic categories have topological versions.

A *topological space* consists of two things—i) a set X, and ii) a collection of subsets of X (subsets in this collection are called *open* (sub)*sets*)—subject to the following three conditions:

1. The subsets $\emptyset$ and X of X are both open.
2. For O_λ (λ in Λ) any collection of open sets, $\bigcup_\Lambda O_\lambda$ is open.
3. For O and O' open sets, $O \cap O'$ is open.

It is immediate from condition 3 that the intersection of any finite number of open sets is open (while, by condition 2, the union of any number of open sets is open). Elements of the set X are normally called *points.* Given a set X, a collection of subsets of X satisfying conditions 1, 2, and 3 is called a *topology* on X.

Example. Let S be any set. Consider the collection of all subsets of S. Then (since $\emptyset$ and S are both subsets of S, since any union of subsets of S is a subset of S, since any intersection of two subsets of S is a subset of S) we have a topology on S. This is called the *discrete topology.*

Example. Let S be any set. Consider the collection of subsets of S which includes only these two: $\emptyset$ and S. Condition 1 is clearly satisfied. Since any union involving only $\emptyset$ and S is either $\emptyset$ or S, condition 2 is satisfied. Since $\emptyset \cap \emptyset = \emptyset$, $\emptyset \cap S = \emptyset$, and $S \cap S = S$, condition 3 is satisfied. Thus we have a topology on S, called the *indiscrete topology.*

Example. Let the set be the set $\mathbf{R}$ of all real numbers. Let the collection of subsets of $\mathbf{R}$ consist of the following: we call a subset O of $\mathbf{R}$ open if, for each point r (a real number) of O, there exists a positive number ϵ such that all r' with $|r' - r| < \epsilon$ are also in O. (Intuitively, a set O is open if O includes all points "sufficiently close" to any point of O.) We verify the three conditions above. Clearly, both $\emptyset$ (because it contains no r) and $\mathbf{R}$ (because it contains every r') satisfy the condition above for openness. Let O_λ (λ in Λ) be any collection of open sets: we show that $\bigcup_\Lambda O_\lambda$ is open. Let r be a point of

$\bigcup_{\Lambda} O_\lambda$, so r is in some O_λ. Then, since that O_λ is open, there is an $\epsilon > 0$ such that every r' with $|r' - r| < \epsilon$ is in that O_λ. Hence every such r' is in $\bigcup_{\Lambda} O_\lambda$. That is, $\bigcup_{\Lambda} O_\lambda$ is open. Finally, let O_1 and O_2 be open: we show that $O_1 \cap O_2$ is open. Let r be in $O_1 \cap O_2$, so r is in both O_1 and O_2. Since O_1 is open, there is an ϵ_1 such that all r' with $|r' - r| < \epsilon_1$ are in O_1; since O_2 is open, there is an ϵ_2 such that all r' with $|r' - r| < \epsilon_2$ are in O_2. Let ϵ be the smaller of ϵ be the smaller of ϵ_1 and ϵ_2. Then all r' with $|r' - r| < \epsilon$ are in both O_1 and O_2, whence all such r' are in $O_1 \cap O_2$. That is, $O_1 \cap O_2$ is open. We have verified the three conditions above. This topological space is called the *real line.*

For a and b real numbers with $a < b$, write (a,b) for the set of all numbers r with $a < r < b$, $(a,b]$ for the set of all r with $a < r \leq b$, $[a,b)$ for the set of all r with $a \leq r < b$, and $[a,b]$ for the set of all r with $a \leq r \leq b$. We now claim that each subset (a,b) of $\mathbf{R}$ is open in the topological space above. (That is why (a,b) is called an "open interval.") Indeed, let r be a

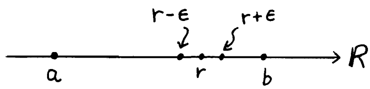

Figure 60

point of (a,b). Clearly (figure 60), there is a positive number ϵ such that $r + \epsilon$ and $r - \epsilon$ are both in (a,b). Then all r' with $|r' - r| < \epsilon$ (i.e., all r' in $(r - \epsilon, r + \epsilon)$) are in (a,b). Hence (a,b) is open. On the other hand, for example, $[a,b)$ is not open: choose $r = a$ and note that no $\epsilon > 0$ does the job.

Let, in the last example above, $O_1 = (0,2/1)$, $O_2 = (0,3/2)$, $O_3 = (0,4/3)$, $\cdots$ Then each of O_1, O_2, $\cdots$ is open. But $\bigcap O_i$ (the intersection of an infinite number of open sets) $= (0,1]$ is not open.

There is another way to obtain topological spaces, using the idea of the real line. We first need a definition. A *metric space* consists of a set X, together with a mapping of sets, $X \times X \xrightarrow{d} \mathbf{R}$ (so, for x and x' in X, $d(x,x')$ is a real number, which we think of as the "distance between x and x'"), such that i) for any x and x', $d(x,x') \geq 0$, with equality if and only if $x = x'$, ii) for any x and x', $d(x,x') = d(x',x)$, and iii) for any x, x', and x'',

$$d(x,x') + d(x',x'') \geq d(x,x'') \ .$$

(Condition i) says that "distances are non-negative"; condition ii) says that "the distance from x to x' is the same as the distance from x' to x"; condition iii) says that "it is generally shorter to go directly from x to x'' than to go by way of x'.") For example, R is a metric space with, for r and r' numbers, $d(r,r') = |r - r'|$.

Example. Let (X,d) be a metric space. Call a subset O of X open if, for any x in O, there is a positive number ϵ such that all x' with $d(x',x) < \epsilon$ are also in O. Then this collection of subsets of X is a topology on X. [Proof: The argument of the previous example goes through word for word, replacing $| - |$ by $d(\, , \,)$.]

Example. Let $X = R \times R$, Cartesian product of sets. Let $d(\, , \,)$ be the usual Euclidean distance on the plane X. We thus have a metric space. By the example above, we have a topological space. This topological space (a convenient one for drawing pictures) is called the *topological plane.*

We have now accumulated enough examples to discuss in general terms what a topological space is. Think of a metric space as a set on which one knows a "distance of any point from any other," that is, on which one has a notion of "how close" one point is to another. A topological space has just a little less structure. While one does not know "how close" (numerically) one point is to another, one still has a notion of "sufficiently close." This notion is embodied in the open sets. One should think of an open set as a set having the property that "any point sufficiently close to a point of the open set is also in the open set." In other words, one thinks of an open set as having the property that "given any point of the open set, sufficiently small variations of location of that point do not take one out of the open set." The three conditions for a topological space can now be seen as reflecting this intuitive picture of an open set. We now introduce a number of definitions and remarks which are intended to amplify the discussion above.

Let X be a topological space. A subset C of X is said to be *closed* if C^c (complement) is open. (Intuitively, "sufficiently small variations of any point not in a closed set yield points also not in that closed set.") For example, in the real line, each $[a,b]$ ("closed interval") is closed. Caution: Some subsets of X can be neither open nor closed, for example, the subset $[a,b)$ of the real line. The basic properties of closed (sub)sets follow from those of open sets by taking complements: i) $\emptyset$ and X are closed. (We have $\emptyset^c = X$ and $X^c = \emptyset$.) ii) For C_λ (λ in Λ) closed sets, $\bigcap_\Lambda C_\lambda$ is also closed. (We have $(\bigcap_\Lambda C_\lambda)^c = \bigcup_\Lambda C_\lambda^c$. Since each C_λ is closed, each C^c is open. Hence $\bigcup_\Lambda C_\lambda^c$ is open. That is, $(\bigcap_\Lambda C_\lambda)^c$ is open, whence $\bigcap_\Lambda C_\lambda$ is closed.) iii) For C and C' closed, $C \cup C'$ is closed. (Since $(C \cup C')^c = C^c \cap C'^c$, and since C^c and C'^c are open, $(C \cup C')^c$ is open. Hence $C \cup C'$ is closed.) In short, the three properties of closed sets are just the three properties of open sets, but with union and intersection interchanged.

Example. The integers form a closed subset of the real line. Any subset of a set with the discrete topology is closed.

Let X be a topological space, and let x be a point of X. A subset N of X is said to be a *neighborhood* of x if N is a superset of some open set O containing x. Note that a neighborhood need not be either open or closed. Thus, on the real line, $(-1,2)$, $[-1,2)$, and $[-1,2]$ are all neighborhoods of the point 0. Every open set is a neighborhood of each of its points. There are two fundamental properties of neighborhoods: i) Any superset of a neighborhood of x is a neighborhood of x. [Proof: Let N be a neighborhood of x, so $N \supset O$ with O open and containing x. Then, for N' a superset of N, $N' \supset O$, whence N' is a neighborhood of x.] ii) If N and N' are neighborhoods of x, so is $N \cap N'$. [Proof: We have $N \supset O$ and $N' \supset O$, with each of O and O' open, and each containing x. Then $O \cap O'$ is open and contains x, while $N \cap N' \supset O \cap O'$.] The following example provides an intuitive picture of a neighborhood.

Example. Let (X,d) be a metric space, and let x be a point of X. Fix a positive number ϵ, and let N denote the subset of X consisting of all x' with $d(x',x) < \epsilon$. Then N is a neighborhood of x. [Proof: Since $d(x,x) = 0$, N contains x. We show that N is open. Let $\underline{x}$ be a point of N, so $d(\underline{x},x) = a$ is less than ϵ. Choose $\underline{\epsilon} = \epsilon - a > 0$. Then, for $d(\underline{x}',\underline{x}) < \underline{\epsilon}$, we have $d(\underline{x}',x) \leq d(\underline{x}',\underline{x}) + d(\underline{x},x) < a + \underline{\epsilon} = a + (\epsilon - a) = \epsilon$, whence $\underline{x}'$ is in N. Thus N is open.]

This example suggests that one think of a neighborhood of x as representing "all points which can be reached from x under a sufficiently small, but otherwise arbitrary, variation of position." Our earlier intuitive discussion of open sets can now be sharpened as follows.

THEOREM 27. *Let X be a topological space. A subset A of X is open if and only if, for each point x of A, some neighborhood of x is a subset of A. A subset A of X is closed if and only if, for each point x not in A, some neighborhood of x does not intersect A.*

Proof. Suppose that A is open. Then A (a subset of A) is a neighborhood of each point of A. Suppose, conversely, that, for each x in A, there is a neighborhood N_x of x with $A \supset N_x$. Then, for each x in A, there is an open set O_x containing x and with $A \supset O_x$. Set $B = \bigcup O_x$, where the union is over all x in A. Then B, as a union of open sets, is open. Since $A \supset O_x$ for each x, $A \supset B$. But each x in A is in some O_x and hence is in B ; thus $A \subset B$. We have shown $A = B$, whence A is open. The last statement is immediate from its predecessor, noting that A is closed if and only if A_c is open. □

Note that there is very little real content to theorem 27: open sets lead to neighborhoods which can then be used to characterize open sets. The only point to this theorem is that "neighborhood" somehow has more intuitive content than "open set."

Let X be a topological space, and let A be any subset of X. Denote by **Int(A) the union of all open subsets of A, so Int(A) is itself an open subset of A. This Int(A) (which should perhaps be called the open set generated by A)**

is called the *interior* of A. The interior of A is, of course, the largest open subset of A in the following sense: for any open O with $A \supset O$, $\text{Int}(A) \supset O$. Similarly, denote by $\text{Cl}(A)$, the *closure* of A, the intersection of all closed supersets of A, so $\text{Cl}(A)$ is itself a closed superset of A. This $\text{Cl}(A)$ is the smallest closed superset of A: for any closed C with $A \subset C$, $\text{Cl}(A) \subset C$. Finally, set $\text{Bndy}(A) = \text{Cl}(A - \text{Int}(A)$, the *boundary* of A.

Example. Let $A = [a,b)$, a subset of the real line. Then $\text{Int}(A) = (a,b)$ [proof: this (a,b) is certainly an open subset of A ; the only larger subset of A is A itself: but A is not open] and $\text{Cl}(A) = [a,b]$ [proof: this $[a,b]$ is certainly a closed superset of A ; the only smaller superset of A is A itself: but A is not closed]. Hence $\text{Bndy}(A)$ is the subset consisting of just two points, a and b.

Example. For A any subset of a set with the discrete topology, $\text{Int}(A) = \text{Cl}(A) = A$, whence $\text{Bndy}(A) = \emptyset$. For A (say, neither $\emptyset$ nor X) a subset of X with the indiscrete topology, $\text{Int}(A) = \emptyset$ and $\text{Cl}(A) = X$, whence $\text{Bndy}(A) = X$.

Example. Let A be the set of integers in the real line. Then $\text{Int}(A) = \emptyset$ while $\text{Cl}(A) = A$, whence $\text{Bndy}(A) = A$.

Example. Let A be the subset of the real line consisting of all the rational numbers. Then, since the only open set containing only rationals is the empty set, $\text{Int}(A) = \emptyset$ while, since the only open set containing no rationals is the empty set (whence the only closed set containing all the rationals is $\mathbf{R}$), $\text{Cl}(A) = \mathbf{R}$. Thus $\text{Bndy}(A) = \mathbf{R}$. (Every real number is a boundary point of the rationals.)

As usual, one can get a quick intuitive picture of something by characterizing it in terms of neighborhoods.

THEOREM 28. *Let A be a subset of topological space X. Let x be a point of X. Then i) x is a point of* $\text{Int}(A)$ *if and only if some neighborhood of x is a subset of A, ii) x is a point of* $\text{Cl}(A)$ *if and only if every neighborhood of x intersects A, and iii) x is a point of* $\text{Bndy}(A)$ *if and only if every neighborhood of x contains both points in A and points not in A.*

Proof:

i) For x a point of $\text{Int}(A)$, $\text{Int}(A)$ itself is both a neighborhood of x and a subset of A. Next, let point x have a neighborhood N with $N \subset A$. Then there is an open set O containing x and with $O \subset N$. Thus O is an open subset of A containing x, whence $\text{Int}\ (A)$ (the union of all open subsets of A) also contains x.

ii) Let x be a point of $\text{Cl}(A)$. Then every closed superset of A contains x, whence (taking complements) every open set not intersecting A fails to contain x, whence every open set containing x intersects A, whence every neighborhood of x intersects A. Next, let x be a point with the property that every neighborhood of x intersects A. Then every closed superset C of A contains x (for, if not, C^c would be a neighborhood of x not intersecting A). But $\text{Cl}(A)$ is the intersection of all such closed supersets, whence x is a point of Cl(A).

iii) The point x is in Bndy(A) if and only if x is in Cl(A) and not in Int(A), that is (by i) and ii)), if and only if every neighborhood of x intersects A but is not a subset of A. □

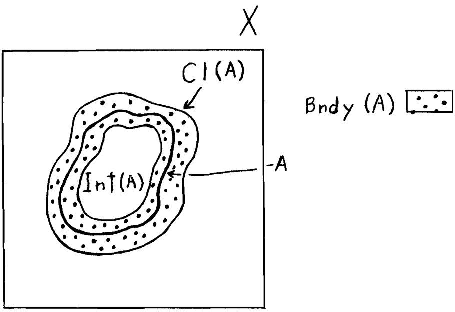

Figure 61

The setup regarding the interior, closure, and boundary of a set A is illustrated in figure 61. We regard theorem 28 as a means of strengthening one's intuition about these notions. These concepts, in turn, can be regarded as a means of strengthening one's intuition about open sets. We have

THEOREM 29. *Let A be a subset of topological space X. Then A is open if and only if no point of* Bndy(A) *is a point of A, while A is closed if and only if every point of* Bndy(A) *is a point of A.*

Proof. It is immediate (e.g., from figure 61) that, if no point of Bndy(A) is a point of A, then A = Int(A) is open, and that, if every point of Bndy(A) is a point of A, then A = Cl(A) is closed. Conversely, if A is open, then, since Int(A) is the union of all open subsets of A, we have A = Int(A), whence no point of Bndy(A) is a point of A. If A is closed, then, since Cl(A) is the intersection of all closed supersets of A, we have A = Cl(A), whence every point of Bndy(A) is a point of A. □

Thus theorem 29 provides still another characterization of open sets: they are sets which include none of their boundary points. (E.g., (a,b) is open in the real line, while neither $(a,b]$ nor $[a,b]$ is open.) Once again, this is all circular: the open sets lead to the notion of a boundary, which can, in turn, be used to characterize the open sets.

Everything we have done since the definition of a topological space has been intended as an amplification of that definition, a picture of how one thinks of an open set. What we now wish to do is obtain an overview of the possibilities for introducing a topology on a given set X.

Fix a set X. We denote by $\mathbf{P}(X)$ the collection of all subsets of X. Thus, for example, a point of $\mathbf{P}(X)$ is a subset of X while a subset of $\mathbf{P}(X)$ is a collection of subsets of X. Next, note that a topology on X is also just a certain collection of subsets of X (namely, a collection of subsets of X satisfying the three conditions defining a topological space). Thus a topology on X is just a certain subset of $\mathbf{P}(X)$. What is the advantage of thinking of topologies on X in these terms? It is that subsets of $\mathbf{P}(X)$, and hence topologies on X, are ordered by inclusion. Thus, for $\mathbf{T}$ and $\mathbf{T}'$ topologies on X (regarded as subsets of $\mathbf{P}(X)$), we write $\mathbf{T} \leq \mathbf{T}'$ if every open set with respect to topology $\mathbf{T}$ is also an open set with respect to topology $\mathbf{T}'$. ("The larger the topology, the more open sets it has.") For $\mathbf{T} \leq \mathbf{T}'$, one says that topology $\mathbf{T}'$ is *finer* than $\mathbf{T}$ and that $\mathbf{T}$ is *coarser* than $\mathbf{T}'$. Now, we can draw a little picture of the partially ordered set of all topologies on set X (figure 62). One "moves to

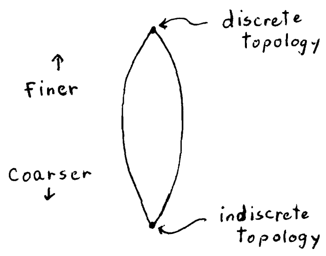

Figure 62

finer and finer topologies (those with more and more open sets)" as one moves upward in the picture. At the top is the finest of all topologies, the discrete topology (with respect to which every set is open). At the bottom is the coarsest topology, the indiscrete topology (in which there are as few open sets as possible consistent with having a topology, namely $\emptyset$ and X). Obviously, the interesting topologies on X (those which convey some genuine information) are those around the center of the figure. (Think of X as a rock pile and of the rocks as open sets. Then finer topologies correspond to smaller rocks. In the discrete topology, every point is an open set—the finest topology—while, in the indiscrete topology, there is just one large rock—X itself.)

We give one example of the use of this organization of the collection of topologies on X. Let $\mathbf{T}_\lambda$ (λ in Λ) be a collection of topologies on X (so, for each λ, $\mathbf{T}_\lambda$ is a certain subset of $\mathbf{P}(X)$). Consider $\mathbf{T} = \bigcap_\Lambda \mathbf{T}_\lambda$, the intersection of subsets of $\mathbf{P}(X)$. We now claim: this $\mathbf{T}$ is in fact also a topology on X. We already know that $\mathbf{T}$ is a subset of $\mathbf{P}(X)$ (i.e., a collection of subsets of X—namely, the collection of all subsets of X which are open in each of the topologies $\mathbf{T}_\lambda$): what we must show is that the three conditions for a topology are satisfied. We do this: (1) Since $\emptyset$ (a subset of X and hence a point of $\mathbf{P}(X)$) is in each $\mathbf{T}_\lambda$, $\emptyset$ is in $\mathbf{T}$. Since X is in each $\mathbf{T}_\lambda$, X is in $\mathbf{T}$. (2) Let O_μ (μ in $\boldsymbol{\mu}$) each be in $\mathbf{T}$. Then each O_μ is open in each topology $\mathbf{T}_\lambda$. Hence (since each $\mathbf{T}_\lambda$ is a topology), $\bigcup_{\mu} O_\mu$ is open in each topology $\mathbf{T}_\lambda$. Hence $\bigcup_{\mu} O_\mu$ is in $\mathbf{T}$. (3) Let O and O' each be in $\mathbf{T}$. Then both O and O' are open in each topology $\mathbf{T}_\lambda$. Hence (since each $\mathbf{T}_\lambda$ is a topology) $O \cap O'$ is open in each topology $\mathbf{T}_\lambda$. Hence $O \cap O'$ is in $\mathbf{T}$. Thus $\mathbf{T}$ is indeed a topology on X: any intersection of topologies on X is again a topology on X (figure 63). We can, for example, characterize this $\mathbf{T}$ as follows: it is the finest topology on X which is coarser than each topology $\mathbf{T}_\lambda$.

Let $\mathbf{A}$ be any collection of subsets of X. Consider topologies on X with respect to which every subset in the collection $\mathbf{A}$ is open. (There certainly is one such topology: the discrete one.) The intersection of all these topologies is, as we have seen above, a topology on X. This is called the topology *generated by* the collection $\mathbf{A}$ of subsets of X. Clearly, the topology generated by $\mathbf{A}$ is the coarsest topology with respect to which every subset (of X) in $\mathbf{A}$ is open (figure 64). One "requires that the subsets in $\mathbf{A}$ be open, and includes whatever other open sets are (absolutely) necessary in order to obtain a topology."

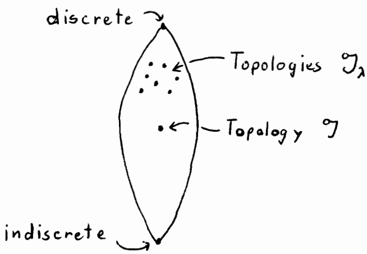

Figure 63

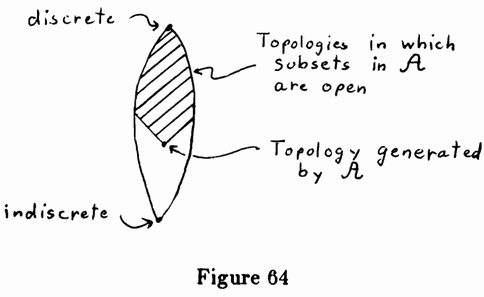

Figure 64

Example. Let $X = \mathbf{R}$, the set of real numbers. Let $\mathbf{A}$ be the following collection of subsets of $\mathbf{R}$: the subsets of the form (a,b). Now, this $\mathbf{A}$ is not already a topology on $\mathbf{R}$: it is false, for example, that the union of two such "open intervals" is another. We claim that the topology generated by $\mathbf{A}$ is the usual topology, $\mathbf{T}$, of the real line. First, note that every element of $\mathbf{A}$ is indeed open in the topology $\mathbf{T}$. We must show that the topology $\mathbf{T}$ is the coarsest one which has this property. Let $\mathbf{T}'$ be another topology on the set $\mathbf{R}$ with respect to which each subset (a,b) is open. Then all unions of subsets of this form must be open with respect to $\mathbf{T}'$. But every open subset with respect to $\mathbf{T}$ (the usual topology) is a union of subsets of this form: hence every set open with respect to $\mathbf{T}$ is also open with respect to $\mathbf{T}'$. That is, $\mathbf{T}$ is coarser than $\mathbf{T}'$.

Example. Let X be any set, and let $\mathbf{A}$ be the collection of all subsets of X having but a single element. Then the topology generated by $\mathbf{A}$ is the discrete topology on X.

Finally, we introduce an example of an additional "niceness" condition on topological spaces. Let X be a topological space. This X is said to be *Hausdorff* if the following condition is satisfied: given any two distinct points x and x' of X, there are neighborhoods N of x and N' of x' which do not intersect: $N \cap N' = \emptyset$.

Example. Any X with the discrete topology is Hausdorff (for, given any $x \neq x'$ in X, x itself is a neighborhood of x, and x' is a neighborhood of x' (since x and x' are open sets)).

Example. If X has two or more elements and has the indiscrete topology, then X is not Hausdorff (for, for x in X, the only neighborhood of x is X itself).

Example. The real line $\mathbf{R}$ is Hausdorff. Given distinct numbers a and b, then, for sufficiently small positive ϵ, $(a - \epsilon, a + \epsilon)$ is a neighborhood of a, and $(b - \epsilon, b + \epsilon)$ a neighborhood of b, with these neighborhoods not intersecting (figure 65).

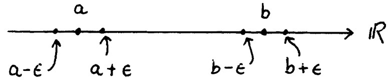

Figure 65

Clearly, any topology on X finer than a Hausdorff topology is Hausdorff. (Intuitively, "finer topologies give one more open sets, hence more neighborhoods, hence an even better chance of finding two neighborhoods, of given x and x', respectively, which do not intersect.") This situation is illustrated in figure 66.

Finally, we give an example of a consequence of Hausdorffness.

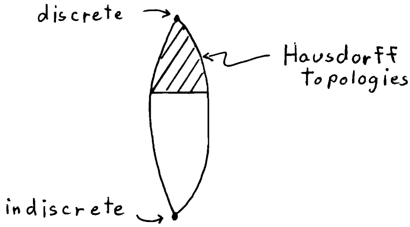

Figure 66

THEOREM 30. *Let X be a Hausdorff topological space. Then each point x of X (regarded as a subset of X) is closed.*

Proof. Let $x' \neq x$. Then, since X is Hausdorff, there are neighborhoods N of x and N' of x' which do not intersect. Hence, in particular, x is not a point of N'. By theorem 28 (part ii) x' is not in $\mathrm{Cl}(x)$. Since this is true for every $x' \neq x$, we have $\mathrm{Cl}(x) = x$. Hence the point x, regarded as a subset of X, is closed. ☐

Example. The real line is Hausdorff. Every point of the real line is a closed subset.

Example. For X a set with two or more elements with the indiscrete topology, X is not Hausdorff. Indeed, for any x in X, $\mathrm{Cl}(x) = X$, so the subset consisting of only x is not closed.

Essentially every topology which arises in applications is Hausdorff. As a general rule, one's intuition is fairly reliable for Hausdorff topological spaces—and somewhat less so for non-Hausdorff.

Exercise 152. Let A_λ (λ in Λ) be subsets of topological space X. Prove: $\mathrm{Cl}(\bigcap A_\lambda) \subset \bigcap \mathrm{Cl}(A_\lambda)$, $\mathrm{Cl}(\bigcup A_\lambda) \supset \bigcup \mathrm{Cl}(A_\lambda)$, $\mathrm{Int}(\bigcap A_\lambda) \subset \bigcap \mathrm{Int}(A_\lambda)$, $\mathrm{Int}(\bigcup A_\lambda) \supset \bigcup \mathrm{Int}(A_\lambda)$.

Exercise 153. Let A be a subset of topological space X. Prove that $\mathrm{Cl}(A^c) = (\mathrm{Int}(A))^c$. Prove that $\mathrm{Bndy}(A^c) = \mathrm{Bndy}(A)$. Find an example to show that this is false: $\mathrm{Int}(\mathrm{Cl}(A)) = \mathrm{Int}(A)$.

Exercise 154. Let A be a subset of topological space X. Why are these definitions pointless? Let $\cdots$ of A be the union of all open supersets of A. Let $\cdots$ of A be the union of all closed subsets of A.

Exercise 155. Prove that every subset of the real line is an intersection of open sets. (It is easier than it looks.)

Exercise 156. Prove that, for any subset A of topological space X, Bndy(A) is closed.

Exercise 157. Find an example of a topological space X such that X is not Hausdorff, but such that every point of X is a closed subset.

Exercise 158. Prove that, given any metric space, (X,d), the corresponding topological space X is Hausdorff.

Exercise 159. Prove: given any neighborhood N of point x of topological space X, there exists a neighborhood N' of x such that N is a neighborhood of every point of N'.

Exercise 160. Prove that a Hausdorff topological space having the property that every open set is also closed is discrete.

Exercise 161. Define a neighborhood space as a set X, together with a rule which assigns, to each point x of X, a collection off supersets of x (to be called neighborhoods of x) subject to the conditions $\cdots$ Now define open subsets of X by the statement of theorem 27. Figure out what $\cdots$ should be in order that i) these open sets define a topology on X, and ii) the neighborhoods, as defined by this topology, are precisely the neighborhoods given in the original neighborhood space.

Exercise 162. Let X be a set on which there is given a collection of topologies. Show that their union is not in general a topology. Prove that nonetheless there exists a finest topology which is coarser than all the topologies in the collection.

Exercise 163. Let **A** be a collection of subsets of set X. State explicitly which subsets of X are open in the topology generated by **A**.

Exercise 164. Let (X,d) be a metric space. Define, for x and x' in X, $D(x,x') = d(x,x')/(1 + d(x,x'))$. Show that (X,D) is also a metric space. Prove that these two metrics define the same topology on X.

Exercise 165. Prove that the only subsets of the real line which are both open and closed are the empty subset and **R** itself.

Exercise 166. Let V be a real vector space. Call a subset O of V open if, for any point v of O and any v' in V, there is a number $\epsilon > 0$ such that $v + av'$ is in O whenever $|a| < \epsilon$. Prove that this is a topology on V. Is it Hausdorff?

Exercise 167. Prove that every open subset of the real line is a union of open intervals.

Exercise 168. Let X be a set. Let the open subsets of X be either those whose complement is all of X or whose complement has only a finite number of points. Prove that this is a topology on X.

27

Continuous Mappings

The primary tool for the study of the structure of topological spaces and their relationship with each other is the notion of a continuous mapping.

Let X and Y be topological spaces, and let $X \xrightarrow{\varphi} Y$ be a mapping of sets. This φ is said to be a *continuous mapping* if, for each open O in Y, $\varphi^{-1}[O]$ is open in X. That is, a mapping is continuous if inverse images of open sets are open. We shall be concerned for the remainder of this chapter with understanding this definition.

Example. Let X and Y be topological spaces, with X discrete. Then any mapping of sets, $X \xrightarrow{\varphi} Y$, is continuous. Indeed, for any open O in Y, $\varphi^{-1}[O]$ is a subset of X and hence, since X has the discrete topology, is open in X.

Example. Let X and Y be topological spaces, with Y indiscrete. Then any mapping of sets, $X \xrightarrow{\varphi} Y$, is continuous. Indeed, since Y has the indiscrete topology, the open subsets of Y are just $\emptyset$ and Y. But $\varphi^{-1}[\emptyset] = \emptyset$ and $\varphi^{-1}[Y] = X$, while, since X is a topological space, $\emptyset$ and X are open subsets of X.

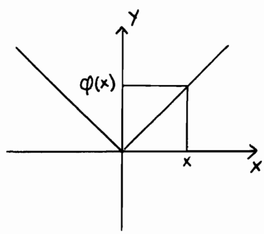

Figure 67

Example. Let X and Y each be the real line, and let $X \xrightarrow{\varphi} Y$ be given by $\varphi(x) = |x|$ (absolute value of the real number x). This mapping is continuous (figure 67). First, note that, for (a,b) any open interval (in Y), $\varphi^{-1}[(a,b)]$ is open in X. (If a and b are both positive, $\varphi^{-1}[(a,b)] = (a,b) \cup (-a,-b)$. If a is negative and b positive, with $|a| \leq |b|$, then $\varphi^{-1}[(a,b)] = (-b,b)$. Similarly for other possibilities.) Continuity follows from the fact that the open subsets of Y include the unions of open intervals.

Example. Let X and Y each be the real line, and let $X \xrightarrow{\varphi} Y$ be given by $\varphi(x) = 1$ for $x \geq 0$ and $= -1$ for $x < 0$ (figure 68). This φ is not continuous.

Indeed, (1/2,3/2) is an open subset of Y, but $\varphi^{-1}[(1/2,3/2)]$ = the collection of all $x \geq 0$, which is not an open subset of X.

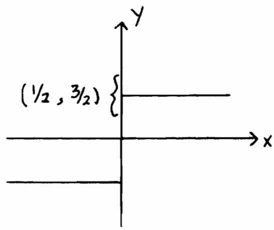

Figure 68

To get an intuitive picture of what is involved in the notion of continuity, we express it in terms of neighborhoods.

THEOREM 31. *Let X and Y be topological spaces, and let $X \xrightarrow{\varphi} Y$ be a mapping of sets. Then φ is continuous if and only if the following condition is satisfied: given any point x of X and any neighborhood M of $\varphi(x)$, there is a neighborhood N of x such that $\varphi[N] \subset M$.* [See figure 69.]

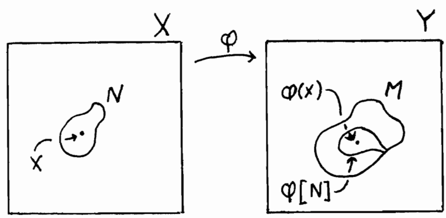

Figure 69

Proof. Suppose first that φ is continuous. Let x be a point of X, and M a neighborhood of $\varphi(x)$. Then there is an open subset O of M with $\varphi(x)$ a point of O. By continuity of φ, $N = \varphi^{-1}[O]$ is an open set in X. But x is a point of this N (since $\varphi(x)$ is in O), so N is a neighborhood of x. But $\varphi[N] \subset o \subset M$. Suppose, conversely, that the condition of the theorem is satisfied. Let O be any open subset of Y: we must show that $\varphi^{-1}[O]$ is open in X. For any point x of $\varphi^{-1}[O]$, $\varphi(x)$ is in O, whence O is a neighborhood of $\varphi(x)$. By hypothesis, there is a neighborhood N of x with $\varphi[N] \subset O$. That is, $N \subset \varphi^{-1}[O]$. That is, $\varphi^{-1}[O]$ contains a neighborhood of each of its points, whence (by theorem 27) $\varphi^{-1}[O]$ is open. □

Intuitively, theorem 31 says the following. Let $X \xrightarrow{\varphi} Y$. Suppose that, given any point x of X and given "how close you want to be to $\varphi(x)$," I can tell you "how close you must be to x in order that φ send you that close to $\varphi(x)$." Then φ is continuous. Suppose, for example, that, for some x in X, "small

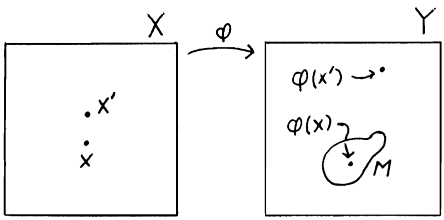

Figure 70

variations in x result in large jumps in the location of $\varphi(x)$." This is the signal that φ is not continuous. Choose the neighborhood M of $\varphi(x)$ (figure 70) so that these "large jumps in the location of $\varphi(x)$ take one out of M." Then one will never find a neighborhood N of x with $\varphi[N] \subset M$, for N, if it is to be a neighborhood, must allow "small variations in x," while these "nearby points to x will be sent, by φ, out of M." The beautiful thing is that all this content is expressed so neatly and cleanly by the definition of continuity.

Example. Let X and Y each be the real line, and let $X \xrightarrow{\varphi} Y$ be a mapping of sets. Then φ is continuous if and only if, given any number x (point of X) and any $\delta > 0$ (which will shortly define a neighborhood of $\varphi(x)$), there is an $\epsilon > 0$ (which will shortly define a neighborhood of x) such that, whenever $|x - x'| < \epsilon$, we have $|\varphi(x) - \varphi(x')| < \delta$. (Thus M is all y in Y with $|\varphi(x) - y| < \delta$, while N is all x' in X with $|x - x'| < \epsilon$. The phrase "$\cdots$ whenever $|x - x'| < \epsilon$, we have $|\varphi(x) - \varphi(x')| < \delta$ " above is another way of waying $\varphi[N] \subset M$.) Note, on the one hand, that, by theorem 31, the statement above is true and, on the other hand, that this statement is just the usual characterization of continuity of functions of a real variable.

Since $\varphi^{-1}[O^c] = (\varphi^{-1}[O])^c$, it is immediate that, for a continuous mapping, the inverse image of each closed set is closed. On the other hand, it is false in general that, for a continuous mapping, the image of an open set is open or that the image of a closed set is closed. (For example, let $X \xrightarrow{\varphi} X$ be the identity mapping of sets, and let the X on the left have the discrete topology, the X on the right the indiscrete topology.)

Next, let X, Y, and Z be topological spaces, and let $X \xrightarrow{\varphi} Y \xrightarrow{\psi} Z$ be continuous mappings. Then $\psi \circ \varphi$ (composition of mappings) is continuous. Proof: Let O be open in Z. Then $(\psi \circ \varphi)^{-1}[O] = \varphi^{-1}[\psi^{-1}[O]]$. Since ψ is continuous, $\psi^{-1}[O]$ is open in Y; then, since φ is continuous, $\varphi^{-1}[\psi^{-1}[O]]$ is open in

X. Thus composition of continuous mappings yields a continuous mapping.

We have been discussing whether a mapping of topological spaces is continuous. The following type of question is also often of interest: if the mapping is already given, what are the possibilities for choosing topologies on the sets involved so that this mapping is continuous?

Let X be a set, Y a topological space, and $X \xrightarrow{\varphi} Y$ a mapping of sets. We wish to find a topology on X such that this mapping is continuous. (Of course, the discrete topology on X would do the job, but that is not a very interesting choice.) We proceed as follows. Consider the collection of all subsets of X of the form $\varphi^{-1}[O]$, for O open in Y. Then: (1) The empty set $\emptyset$ and X itself are in this collection (for $\emptyset = \varphi^{-1}[\emptyset]$ and $X = \varphi^{-1}[Y]$). (2) Any union of sets in this collection is a set in this collection. (Let $\varphi^{-1}[O_\lambda]$ be sets in our collection, so each O_λ is open in Y. Then $\bigcup \varphi^{-1}[O_\lambda] = \varphi^{-1}[\bigcup O_\lambda]$, whence, since $\bigcup O_\lambda$ is open in Y, $\bigcup \varphi^{-1}[O_\lambda]$ is in our collection.) (3) The intersection of two sets in our collection is a set in our collection. (For O and O' open in Y, $\varphi^{-1}[O] \cap \varphi^{-1}[O'] = \varphi^{-1}[O \cap O']$.) That is, the subsets of X of the form $\varphi^{-1}[O]$ form a topology on X. We call this the topology on X *induced* by $X \xrightarrow{\varphi} Y$. Now let X and Y both be given topological spaces, and $X \xrightarrow{\varphi} Y$ a mapping of sets. Then, if φ is continuous, every subset of X of the form $\varphi^{-1}[O]$ (O open in Y) must be open in X. That is, every open set in X in the induced topology must be open in the given topology on X. It is equally clear that, if every open set in the induced topology on X is open in the given topology on X, then φ is continuous. Thus: $X \xrightarrow{\varphi} Y$ is continuous if and only if the topology on X is finer than the induced topology on X. Another way of saying the same thing: the induced topology on X is the coarsest for which $X \xrightarrow{\varphi} Y$ is continuous. The situation is illustrated in figure 71.

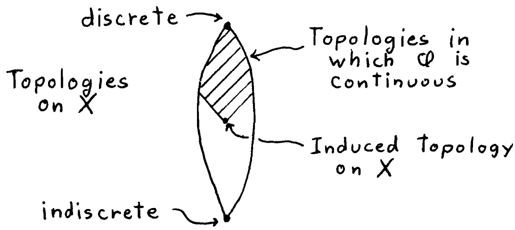

Figure 71

We give one example of the use of the discussion above. Let Y be a topological space, and let A be a subset of Y. Let $A \xrightarrow{\varphi} Y$ be he natural monomorphism which inserts A into Y. Then we have, on A, the induced topology, with respect to which φ is continuous. In more detail, the open subsets of A in this topology are subsets of the form $A \cap O$, with O open in Y (figure 72). This A, with this (induced) topology, is called a ***subspace*** of Y.

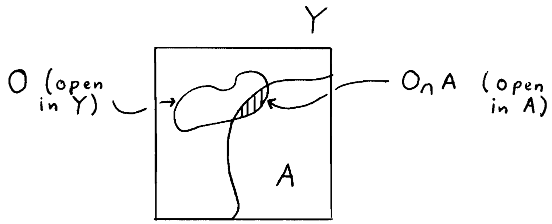

Figure 72

We next consider the reverse of the situation above. Let X be a topological space, and Y a set, and let $X \xrightarrow{\varphi} Y$ be a mapping of sets. Consider the collection of all subsets O of Y for which $\varphi^{-1}[O]$ is open in X. We claim that this collection of subsets of Y defines a topology on Y. [Proof: (1) $\varphi^{-1}[\emptyset] = \emptyset$, $\varphi^{-1}[Y] = X$. (2) For O_λ (λ in Λ) in this collection, so each $\varphi^{-1}[O_\lambda]$ is open in X, $\cup\, \varphi^{-1}[O_\lambda] = \varphi^{-1}[\cup\, O_\lambda]$ is open in X, whence $\cup\, O_\lambda$ is in the collection. (3) For O and O' in this collection, so is $O \cap O'$, since $\varphi^{-1}[O \cap O'] = \varphi^{-1}[O] \cap \varphi^{-1}[O']$.] We call this the topology on Y ***induced*** by $X \xrightarrow{\varphi} Y$. The induced topology on Y is the finest for which φ is continuous, so, for X and Y topological spaces, $X \xrightarrow{\varphi} Y$ is continuous if and only if the topology of Y is coarser than the induced topology on Y (figure 73). (No confusion results from these two meanings for the term "induced topology.")

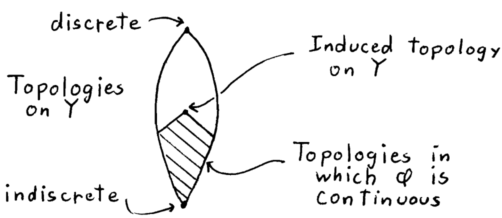

Figure 73

Finally, we give an example of a use of the induced topology above. Let X be a topological space, and suppose that we are given an equivalence relation, "$\approx$," on the set X. Denote by Y the collection of all equivalence classes. Let $X \xrightarrow{\varphi} Y$ be the mapping which sends each point of X to its equivalence class. Then we have, on Y, the induced topology, with respect to which φ is continuous. In more detail, this topology is the following. Let O be a subset of Y, so each point of O is a subset of X (an equivalence class). Then $\varphi^{-1}[O]$ is just the union (in X) of all the equivalence classes represented by points of O. Thus, in the induced topology, this O is open in Y whenever this union of equivalence classes is an open subset of X. In figure 74, the equivalence relation on X is "two points of X are equivalent if they lie on the same horizontal line." Then the equivalence classes are the horizontal lines. Hence Y (the set of all horizontal lines) can be represented as a vertical line. A subset of Y is open in the induced topology if and only if the union of the corresponding collection of horizontal lines is open. In general, the topology on Y, obtained as above, is called the ***quotient topology*** (on the set of equivalence classes of the equivalence relation "$\approx$" on the topological space X).

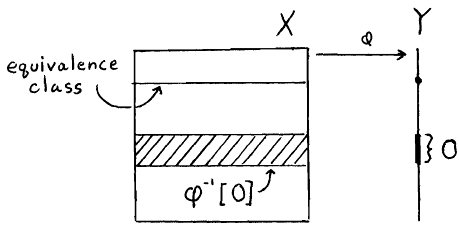

Figure 74

Exercise 169. Let (X,d) be a metric space, and fix a point $\underline{x}$ of X. Let $X \xrightarrow{\varphi} \mathbf{R}$ be the mapping with $\varphi(x) = d(x,\underline{x})$. Prove that φ is continuous.

Exercise 170. Let $R \xrightarrow{\varphi} Y$ be continuous, where Y has the discrete topology. Show that, for any numbers r and r', $\varphi(r) = \varphi(r')$.

Exercise 171. Let X and Y be sets. Find all mappings of sets, $X \xrightarrow{\varphi} Y$, which are continuous for all topologies on X and Y.

Exercise 172. Let X be a set, Y a topological space, and $X \xrightarrow{\varphi_\lambda} Y$ (λ in Λ) a collection of mappings from X to Y. Construct the coarsest topology on X such that all these mappings are continuous.

Exercise 173. Is it true that, for X and Y topological spaces and $X \xrightarrow{\varphi} Y$ a mapping of sets, φ is continuous if and only if inverse images of closed sets are closed?

Exercise 174. Let Y be the topological plane, so a point of Y is a pair, (y_1,y_2), of real numbers. Let A be the subset of Y consisting of all points for which $(y_1)^2 + (y_2)^2 = 1$. (The subspace A, as a topological space, is called the *topological one-sphere.*) Find explicitly the open sets in A. Let $\mathbf{R} \xrightarrow{\varphi} A$ be given by $\varphi(r) = (\cos r, \sin r)$. Show that this φ is continuous.

Exercise 175. Let V be a real vector space with the topology of exercise 166. Show that every linear mapping, $V \xrightarrow{\varphi} V$ is continuous.

Exercise 176. Let V be a real vector space, W a subspace of V, and V/W the quotient space. Then, giving V the topology of exercise 166, $V \xrightarrow{\varphi} V/W$ induces a topology on the vector space V/W. Show that this topology is also that of exercise 166.

Exercise 177. Let X be a topological space. Call an equivalence relation on X Hausdorff if the induced quotient topology on the set of equivalence classes is Hausdorff. Show that the intersection of all Hausdorff equivalence relations on X is a Hausdorff equivalence relation. (Thus one "makes equivalent as few points of X as necessary to get a Hausdorff quotient space.") Give this equivalence relation explicitly.

Exercise 178. Let $X \xrightarrow{\varphi} Y$ be a continuous mapping of topological spaces, and let A be a subset of Y. Is it true that $\varphi^{-1}\,[\mathrm{Cl}(A)] = \mathrm{Cl}(\varphi^{-1}[A])$? that $\varphi^{-1}[\mathrm{Int}(A)] = \mathrm{Int}(\varphi^{-1}[A])$?

Exercise 179. Let $X \xrightarrow{\varphi} Y$ be a continuous mapping of topological spaces. Let L be the subspace $\varphi[X]$ of Y, and let K be the quotient space of X by the equivalence relation $x \approx x'$ if $\varphi(x) = \varphi(x')$ on X. Show that the natural mappings $X \xrightarrow{\alpha} K \xrightarrow{\beta} L \xrightarrow{\gamma} Y$ are all continuous and that $\varphi = \gamma \circ \beta \circ \alpha$.

The Category of Topological Spaces

Let the objects be topological spaces, the morphisms continuous mappings of topological spaces, and composition composition of continuous mappings (noting that the composition of two such is another). We thus obtain a category, the *category of topological spaces.* (Similarly, e.g., the *category of Hausdorff topological spaces.*) In this chapter, we shall specialize various categorical notions to the category of topological spaces.

There is apparently only one forgetful functor from the category of topological spaces—that to the category of sets (i.e., for X a topological space, let $\mathbf{F}(X)$ be the set X). The corresponding free construction would be this. Let S be a set. A free topological space on S consists of a topological space X, together with a mapping $S \xrightarrow{\mu} X$ of sets, such that the following condition is satisfied: given any topological space Y and mapping $S \xrightarrow{\nu} Y$ of sets, there is a unique continuous mapping $X \xrightarrow{\gamma} Y$ of topological spaces such that the diagram of figure 75 commutes. We now claim: (X,μ) is a free topological

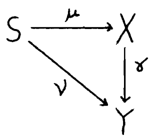

Figure 75

space on S provided $S \xrightarrow{\mu} X$ is an isomorphism of sets (i.e., provided "set X is just a copy of set S"), with X having the discrete topology. [Proof: Let Y be a topological space, and $S \xrightarrow{\nu} Y$ a mapping of sets. Then, since $S \xrightarrow{\mu} X$ is an isomorphism of sets, there is certainly a unique mapping of sets, $X \xrightarrow{\gamma} Y$, such that the diagram of figure 75 commutes. But, since X has the discrete topology, this γ is necessarily a continuous mapping of topological spaces.] Thus nothing essentially new results from this construction. Instead of "taking the free topological space on set S," one might just as well "endow the set S with the discrete topology" (since X is just a copy of set S anyway).

Next, let X and Y be topological spaces. An ***isomorphism*** from X to Y is **a continuous mapping $X \xrightarrow{\varphi} Y$ of topological spaces for which there exists a continuous mapping $Y \xrightarrow{\psi} X$ with $\psi \circ \varphi = i_X$ and $\varphi \circ \psi = i_Y$. We now claim:**

$X \xrightarrow{\varphi} Y$ is an isomorphism if and only if i) φ is one-to-one and onto, and ii) images and inverse images of open sets, under φ, are also open sets. [Proof: Clearly, these conditions imply (choosing for ψ the inverse of φ, the existence of which is guaranteed by condition i)) that φ is an isomorphism. Suppose, conversely, that φ is an isomorphism. Then φ must be an isomorphism of sets, i.e., φ must be one-to-one and onto. By continuity of φ, inverse images of open sets must be open. Finally, for O open in X, $\varphi[O] = \psi^{-1}[O]$ must be open in Y, by continuity of ψ.] Thus an isomorphism, in the category of topological spaces, is "a correspondence between set X and set Y which takes open sets in X to open sets in Y, and vice versa." It "makes X and Y identical as topological spaces." (An isomorphism, in this category, is normally called a homeomorphism.)

It is easily checked that monomorphisms, in the category of topological spaces, are just one-to-one continuous mappings. Thus, for Y a topological spaces, a subobject (in this category) consists of a topological space A together with a one-to-one continuous mapping $A \xrightarrow{\varphi} Y$. Since φ is one-to-one, we may regard A as a subset of Y. Then, in order that φ be continuous, the topology on A must be finer than the topology induced by $A \xrightarrow{\varphi} Y$. Thus, for Y a topological space, any subset of Y, given a topology finer than the induced topology on that subset, yields a subobject. In particular, there are in general many topologies on such a subset (e.g., the induced topology itself, the discrete topology, or any in between) which make it a subobject. (Note that no such feature arises, e.g., for subgroups.) It is conventional to reserve the term "subspace" for the case when the subset is given the induced topology.

These preliminaries out of the way, we now turn to the interesting part: direct products and direct sums.

Let X and Y be topological spaces. Then (from category theory) a *direct product* of X and Y consists of a topological space Z, together with continuous mappings $Z \xrightarrow{\alpha} X$ and $Z \xrightarrow{\beta} Y$, such that the following condition is satisfied: given any topological space W, and any continuous mappings $W \xrightarrow{\mu} X$ and $W \xrightarrow{\nu} Y$, there is a unique continuous mapping $W \xrightarrow{\gamma} Z$ such that the diagram of figure 76 commutes. In fact, direct products indeed exist in this category. Choose, for the set Z, $Z = X \times Y$, the Cartesian product of sets. (As regards underlying sets, topology is very "set-theory-like," much more so than is, say, group theory.) Thus a point of Z is an ordered pair, (x,y), with x a point of X and y a point of Y, and $\alpha(x,y) = x$ and $\beta(x,y) = y$. For the topology of Z, we choose the coarsest one for which the mappings $Z \xrightarrow{\alpha} X$ and $Z \xrightarrow{\beta} Y$ are continuous. (Motivation: We are trying to get a direct product here, so we want to choose a topology on Z so that γ is "as likely as possible to be continuous."

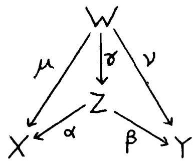

Figure 76

Hence one wants to place on Z the coarsest topology possible. We cannot, however, go too far in this direction—e.g., choose the indiscrete topology on Z—for we might thereby fail to have α and β continuous. So we strike the best possible compromise.) In more detail, the topology on Z is that generated by the subsets of Z of the form $\alpha^{-1}[O]$, with O open in X, and $\beta^{-1}[U]$, with U open in Y. (Note that, for O open in X, $\alpha^{-1}[O]$ consists of all points (x,y) of Z with x in O.) In still more detail, the open sets in Z are unions of sets of the form $\alpha^{-1}[O] \cap \beta^{-1}[U]$, with O open in X and U open in Y. (Note that $\alpha^{-1}[O] \cap \beta^{-1}[U]$ consists of points (x,y) of Z with x in O and y in U.) This is clearly a topology on Z and is clearly the coarsest one with respect to which both $Z \xrightarrow{\alpha} X$ and $Z \xrightarrow{\beta} Y$ are continuous. We now claim that this (Z,α,β) is a direct product of X and Y. Proof: Let W, μ, and ν be as above. Then, since Z is the Cartesian product of sets X and Y, there is certainly a unique mapping of sets, $W \xrightarrow{\gamma} Z$, such that the diagram commutes. We have only to prove that this γ is continuous. First, note that, for open sets in Z of the form $\alpha^{-1}[O] \cap \beta^{-1}[U]$ (O open in X and U open in Y), we have $\gamma^{-1}[\alpha^{-1}[O] \cap \beta^{-1}[U]] = \gamma^{-1}[\alpha^{-1}[O]] \cap \gamma^{-1}[\beta^{-1}[U]] = (\alpha \circ \gamma)^{-1}[O] \cap (\beta \circ \gamma)^{-1}[U] = \mu^{-1}[O] \cap \nu^{-1}[U]$ which (since μ and ν are continuous and since the intersection of two open sets is open) is open in W. Thus inverse images, by γ, of open subsets of Z of this form are open. But every open set in Z is a union of open sets of this form: hence γ is continuous. Thus (Z,α,β) is indeed a direct product of X and Y.

Example. The direct product of two topological spaces with the discrete topology is the Cartesian product with the discrete topology. The direct product of two topological spaces with the indiscrete topology is the Cartesian product with the indiscrete topology.

Example. Let X and Y each be the real line. Then the set Z of the topological space that is the direct product of X and Y is the collection of all pairs of real numbers. Given an open set O in X, for example, the open interval (a,b), and an open set U in Y, for example, the open interval (c,d), the open set $\alpha^{-1}[O] \cap \beta^{-1}[U]$ in Z is the collection of all pairs (x,y) with $a < x < b$ and $c < y < d$, that is, an "open square" in the plane Z. (Note that the notation "(,)" has two different meanings in the sentence above.) (See figure 77.) The most general open set in Z is a union of such "open squares." Clearly,

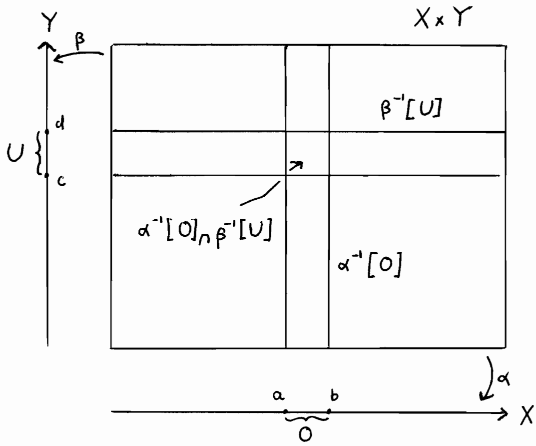

Figure 77

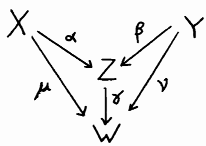

Figure 78

this topological space Z is just the topological plane.

Once again, let X and Y be topological spaces. Then (again from category theory), a *direct sum* of X and Y consists of a topological space Z, together with continuous mappings $X \xrightarrow{\alpha} Z$ and $Y \xrightarrow{\beta} Z$, such that the following condition is satisfied: given any topological space W, and any continuous mappings $X \xrightarrow{\mu} W$ and $Y \xrightarrow{\nu} W$, there is a unique continuous mapping $Z \xrightarrow{\gamma} W$ such that the diagram of figure 78 commutes. Direct sums also exist in this category. Choose, for the set Z, $Z = X \cup_d Y$, the disjoint union of sets, and for $X \xrightarrow{\alpha} Z$ and $Y \xrightarrow{\beta} Z$ the usual mappings of sets. For the topology of Z, we choose the finest one for which $X \xrightarrow{\alpha} Z$ and $Y \xrightarrow{\beta} Z$ are continuous (for "we want to make it as likely as possible that γ will be continuous, whence we want the topology of Z to be as fine as possible, while we cannot afford to make it so fine that either α or β fail to be continuous.") Thus a subset U of Z will be taken to be open if $\alpha^{-1}[U]$ is open in X and $\beta^{-1}[U]$ is open in Y. This is clearly a topology on Z, and clearly the finest one for which α and β are continuous. (Note that, in this case, we do not have to consider the topology generated by something.) We now claim that this (Z,α,β) is a direct sum of X

and Y. Proof: Let W, μ, and ν be as above. Then, since Z is the disjoint union of sets X and Y, there is certainly a unique mapping of sets, $Z \xrightarrow{\gamma} W$, which makes the diagram commute. We must show that this γ is continuous. Let O be open in W: we must show that $\gamma^{-1}[O]$ is open in Z. That is, we must show that $\alpha^{-1}[\gamma^{-1}[O]]$ is open in X, and $\beta^{-1}[\gamma^{-1}[O]]$ is open in Y. But $\alpha^{-1}[\gamma^{-1}[O]] = (\gamma \circ \alpha)^{-1}[O] = \mu^{-1}[O]$ is indeed open in X, by continuity of μ, and similarly for $\beta^{-1}[\gamma^{-1}[O]]$. Thus γ is continuous, whence this Z is indeed a direct sum of X and Y.

The situation is illustrated in figure 79. Since the set Z is the disjoint union of sets X and Y, Z consists of "a copy of X placed beside a copy of Y." A typical open set O in Z is a subset whose "overlap with the copy of X is open in X and whose overlap with the copy of Y is open in Y." (This admission of open sets in Z which intersect both "the copy of X and the copy of Y" is necessary because, if Z is to be a topological space, unions of open sets must be open.)

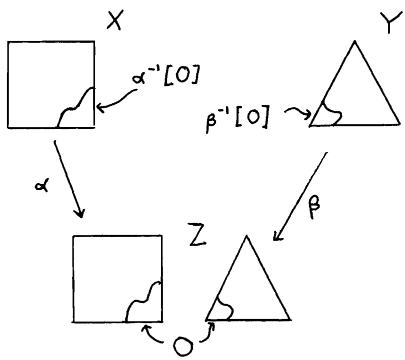

Figure 79

Example. The direct sum of two topological spaces with the discrete topology is the disjoint union with the discrete topology. The direct sum of two topological spaces with the indiscrete topology is the disjoint union, with just four open sets: the empty set, the disjoint union itself, the "copy of X," and the "copy of Y."

Example. The subspace $(0,1) \cup [2,3]$ of the real line is isomorphic to the direct sum of the subspace $(0,1)$ with the subspace $[2,3]$. It is false, however, that the subspace $(0,1) \cup [1,2]$ of the real line is isomorphic to the direct sum of the subspace $(0,1)$ with the subspace $[1,2]$.

Exercise **180. Prove that any subspace of a Hausdorff topological space is Hausdorff.**

Exercise 181. Find an isomorphism from the subspace (0,1) of the real line to the real line. Prove that there is no such isomorphism from the subspace [0,1].

Exercise 182. Give a universal definition which leads to the introduction of the indiscrete topology on a set.

Exercise 183. Prove that every topological space is a direct product of two topological spaces.

Exercise 184. Find an isomorphism from X, the subspace of the topological plane consisting of all points except the origin, to Y, the direct product of the topological one-sphere and the real line.

Exercise 185. Prove that both the direct product and the direct sum of two Hausdorff topological spaces is Hausdorff.

Exercise 186. Prove that the continuous mappings α and β in the definition of a direct product take open sets to open sets. Find an example in which they do not take closed sets to closed sets.

Exercise 187. Prove that the continuous mappings α and β in the definition of a direct sum take open sets to open sets and closed sets to closed sets.

Exercise 188. Prove that, in the category of topological spaces, epimorphisms are onto continuous mappings. Prove that, in the category of Hausdorff topological spaces, a continuous mapping $X \xrightarrow{\varphi} Y$ with $\mathrm{Cl}[\varphi[X]] = Y$ is an epimorphism.

Exercise 189. Let X be the rational numbers (as a subspace of the real line), Y the real line, and $X \xrightarrow{\varphi} Y$ the mapping with $\varphi(x) = x$. Prove that φ is a continuous mapping, a monomorphism, and an epimorphism—and that φ is not an isomorphism.

Exercise 190. Prove that, if the topologies of X and Y are replaced by finer topologies, then the topology of the direct product becomes finer. Similarly for the direct sum.

Exercise 191. Let A and B be disjoint subsets of topological space X. Find a simple necessary and sufficient condition that the natural mapping from the subspace $A \cup B$ of X to the direct sum of subspace A with subspace B be an isomorphism.

Exercise 192. Let X and Y be topological spaces. Define an equivalence relation on the direct product: $(x,y) \approx (x',y')$ if $x = x'$. Find an isomorphism from the quotient space to Y.

29

Nets

We introduce in this chapter the notion of a net. One can regard this notion as analogous to that of a boundary or to that of a neighborhood in the following sense: each contributes a slightly different point of view to the concept of an open set.

Fix a topological space X. A *sequence* in X is a mapping, $Z^+ \xrightarrow{\kappa} X$, from the set Z^+ of positive integers to the set X. Instead of writing, for n a positive integer, $\kappa(n)$ for the corresponding point of X, we may write x_n. Thus a sequence in X consists of x_1, x_2, x_3, . . . , a "numbered list of points of X." The intuitive notion (see figure 80) of such a sequence's "approaching a point

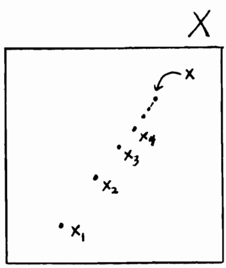

Figure 80

x of X" certainly has a topological flavor. Thus one might imagine introducing such a notion in topology and proving theorems about it. It turns out, however, to be easier to deal with things somewhat more general than sequences: these "things" will be called nets. The idea is to deny the positive integers the special role they play with regard to sequences.

A *directed set* is a nonempty partially ordered set Δ having the following property: given any elements δ and δ' of Δ, there is an element δ'' of Δ with $\delta \leq \delta''$ and $\delta' \leq \delta''$.

Example. The positive integers Z^+ (on which "$\leq$" is given its usual meaning) is a directed set: given any two positive integers, there exists a positive integer greater than or equal to both.

Example. The set of real numbers (with the usual "$\leq$") is a directed set.

Example. Let X be a topological space, and x a point of X. Denote by Δ the collection of all neighborhoods of x. For δ and δ' neighborhoods of x, write $\delta \leq \delta'$ if $\delta' \subset \delta$. (Note that "larger elements of the directed set Δ are smaller neighborhoods.") This Δ is a directed set. Indeed, for δ and δ' elements of Δ (i.e., neighborhoods of x), $\delta'' = \delta \cap \delta'$ is also a neighborhood of x, while $\delta \leq \delta''$ and $\delta' \leq \delta''$. This Δ will be called the directed set of neighborhoods of x.

Example. The partially ordered set of figure 24 (see chapter 8) is not a directed set.

Fix a topological space X. A ***net*** in X consists of a directed set Δ together with a mapping $\Delta \xrightarrow{\kappa} X$ from set Δ to set X. Instead of writing, for δ in Δ, $\kappa(\delta)$ for the corresponding point of X, we shall write x_δ. (We shall usually refer to a net thus: "$\cdots$ net x_δ (δ in Δ) $\cdots$") Thus, for example, every sequence is a net (where the directed set is the directed set of positive integers). The advantage of nets over sequences is that, since the former "allow larger and more complicated directed sets, nets are able to approach points in more sophisticated ways."

Let X be a topological space, and let x_δ (δ in Δ) be a net in X. A point x of X is said to be a *limit point* of this net if the following condition is satisfied: given any neighborhood N of x, there is an element δ of Δ such that all $x_{\delta'}$ with $\delta \leq \delta'$ are points of N (figure 81). One says, when this is the case, that the net *converges* (to x). (Intuitively, a net converges to x in X if "the points of the net eventually get into, and remain in, every neighborhood of x.")

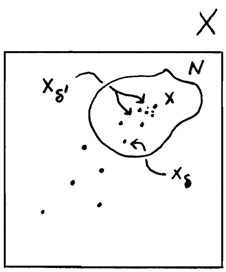

Figure 81

Example. Let X be any topological space, x any point of X, and Δ any directed set. Then the net with $x_\delta = x$ for every δ converges to x.

Example. Let X be the real line, and consider the sequence with $x_1 = 1/2$, $x_2 = 1/4$, $x_3 = 1/8, \cdots$ This net converges to the point 0 (for, given any neighborhood of 0, there is a positive integer such that all x_n with n greater than or equal to this integer are in the given neighborhood).

Example. Let X be the real line, and consider the sequence with $x_1 = 0$, $x_2 = 1$, $x_3 = 0$, $x_4 = 1, \cdots$ This net does not have any limit points. Indeed, given any real number r, choose a neighborhood of r which excludes either the point 0 or the point 1 (always possible). Then, since "every other x_n is outside this neighborhood," there is no integer such that every x_n with n greater than or equal to this integer is a point of our neighborhood.

Example. For X indiscrete, every point of X is a limit point of every net in X.

This last example, in particular, suggests

THEOREM 32. *Let X be a Hausdorff topological space. Then any net in X, if it has any limit point, has a unique one.*

Proof. Let x_δ (δ in Δ) be a net in X, and let this net have two different limit points, x and x'. We obtain a contradiction. Since X is Hausdorff, there are neighborhoods N of x and N' of x', with $N \cap N' = \emptyset$. Since the net converges to x, there is a δ with $x_{\delta''}$ in N whenever $\delta \leq \delta''$. Since the net converges to x', there is a δ' with $x_{\delta''}$ in N' whenever $\delta' \leq \delta''$. Since Δ is directed, there is a δ'' with $\delta \leq \delta''$ and $\delta' \leq \delta''$. Hence, for this δ'', we have $x_{\delta''}$ a point of N and $x_{\delta''}$ a point of N', contradicting $N \cap N' = \emptyset$. ☐

We now describe the "net point of view" toward open sets and continuous mappings.

THEOREM 33. *Let A be a subset of topological space X. Then* Cl(A) *consists precisely of the limit points of nets in A (i.e., nets each of whose x_δ is a point of A).*

Proof. Let x be a limit point of net x_δ (δ in Δ) in A. Then, given any neighborhood N of x, there is a δ with (in particular) x_δ a point of N. But this x_δ is also a point of A, whence this neighborhood N intersects A. Since every neighborhood of x intersects A, x is a point of Cl(A). Now suppose, conversely, that x is a point of Cl(A), so every neighborhood of x intersects A. We must find a net in A converging to x. Let Δ be the directed set of neighborhoods of x. For δ in Δ (so δ is a neighborhood of x), choose x_δ a point of $\delta \cap A$. This is a net in A. Furthermore, it converges to x, for, given any neighborhood N ($= \delta$) of x, all $x_{\delta'}$ with $\delta \leq \delta'$ (i.e., with $\delta' \subset N$) are points of N. ☐

The only tricky part of the proof is in the last half, in which one uses the directed set of neighborhoods of x to construct a net. It is crucial here that one have available the notion of a net (rather than just that of a sequence). In fact, theorem 33 is false if "net" is replaced by "sequence" (for "sequences are not long enough, in general, to probe out all points of Cl(A) while remaining in A").

It is immediate from theorem 33 that a subset C of topological space X is closed if and only if C contains every limit point of every net in C.

(For example, [0,1] is a closed subset of the real line.) Thus a subset O of X is open provided no point of O is a limit point of a net none of those x are points of O. (Nets "cannot converge to a point of open O without actually entering O.")

Next, we characterize continuous mappings.

THEOREM 34. *Let X and Y be topological spaces, and $X \xrightarrow{\varphi} Y$ a mapping of sets. Then φ is continuous if and only if the following property is satisfied: given any point x of X, and any net x_δ (δ in Δ) in X that converges to x, the net $\varphi(x_\delta)$ (δ in Δ) in Y converges to $\varphi(x)$.*

Proof. Suppose first that φ is continuous. Let x_δ (δ in Δ) be a net in X that converges to x. By continuity of φ there exists, for any neighborhood M of $\varphi(x)$ in Y, a neighborhood N of x with $\varphi[N] \subset M$. By convergence of the net to x, there exists a δ with $x_{\delta'}$ in N whenever $\delta \leq \delta'$. Hence $\varphi(x_\delta)$ is a point of M whenever $\delta \leq \delta'$. This is so for every neighborhood M of $\varphi(x)$, so the net $\varphi(x_\delta)$ (δ in Δ) converges to $\varphi(x)$ in Y. Now suppose, conversely, that the property of the theorem is satisfied. We show that the assumption that φ is not continuous leads to a contradiction. By theorem 31 (assuming φ not continuous), there is a point x of X and a neighborhood M of $\varphi(x)$ such that for no neighborhood N of x is $\varphi[N] \subset M$. Let Δ be the directed set of neighborhoods of x. For δ in Δ, choose x_δ a point of neighborhood δ with $\varphi(x_\delta)$ not a point of M (possible, since $\varphi[\delta]$ is not a subset of M). Then this net in X converges to x, while, since no $\varphi(x_\delta)$ is a point of neighborhood M in Y, the net $\varphi(x_\delta)$ in Y does not converge to $\varphi(x)$. This contradicts our supposition that the property of the theorem is satisfied. Hence φ must be continuous. ☐

Theorem 34 gives perhaps the best intuitive picture of a continuous mapping: it "preserves limit points, whence it does not permit the image of a point to move too wildly with motion of the point itself."

We conclude this chapter with one more definition. For x_δ (δ in Δ) a net in topological space X, point x of X is said to be an *accumulation point* (of this net) if, for any neighborhood N of x and any δ in Δ, there exists a δ' with $\delta \leq \delta'$ and with $x_{\delta'}$ in N. (Intuitively, x is an accumulation point of a net if the net "continually reenters" every neighborhood of x, i.e., if the net "never gets out of and remains thereafter out of any neighborhood of x.") Clearly, every limit point of a net is an accumulation point of that net, although the converse is false: the numbers 0 and 1 are both accumulation points of the net on the reals consisting of a sequence which alternates between 0 and 1.

Exercise 193. State and prove the converse of theorem 32.

Exercise 194. Prove that every net x_δ (δ in Δ) on a topological space X with Δ finite converges.

Exercise 195. Let x be a point of topological space X. Let Δ be the directed set of neighborhoods of x, and let x_δ be a point of δ for each δ. Prove that this net converges to x.

Exercise 196. Let z_δ (δ in Δ) be a net in Z, the direct product of X and Y (with $Z \xrightarrow{\alpha} X$ and $Z \xrightarrow{\beta} Y$ the corresponding continuous mappings). Prove that this net converges if and only if the nets $\alpha(z_\delta)$ in X and $\beta(z_\delta)$ in Y converge.

Exercise 197. Characterize the nets, on a discrete topological space, that converge.

Exercise 198. Prove that, for any topological space X, there exists a net in X having every point of X as an accumulation point. (Hints: Every set can be given a partial ordering so that it is a directed set. One might as well give X the discrete topology.)

Exercise 199. Find an example of a continuous $X \xrightarrow{\varphi} Y$, and a net x_δ (δ in Δ) in X which does not converge to x but with $\varphi(x_\delta)$ converging to $\varphi(x)$.

Exercise 200. Prove that every net in the subspace $[0,1]$ of the real line has an accumulation point.

Exercise 201. Let $X = \mathbf{R}$, the set of real numbers. Let the open sets in X consist of the empty set together with sets whose complements are countable. Prove that this is a topology on X. Prove that the element 0 of X is in the closure of the subset $(0,1)$ of X. Show that no sequence in $(0,1)$ converges, in this topology, to 0. Find a net in $(0,1)$ that converges to 0.

Compactness

We now introduce the notion of compactness in topology.

It is convenient to have available the following definition. Let S be a set, A a subset of S, and A_λ (λ in Λ) some collection of subsets of S. Then the A_λ are said to *cover* A if $A \subset \bigcup_{\Lambda} A_\lambda$, that is, if every point of A is a point of some A_λ. In particular, the collection A_λ (λ in Λ) of subsets of S covers S itself provided $S = \bigcup_{\Lambda} A_\lambda$.

Let X be a topological space. This X is said to be *compact* if it has the following property: given any collection O_λ (λ in Λ) of open sets that cover X, some finite number of these O_λ also cover X.

Example. Let X be any infinite set with the discrete topology. Then X is not compact. Consider the collection of all subsets of X that contain just one point of X. These are open sets (since X is discrete), and they cover X (since every point of X is a point of one of them). But, since X is infinite, no finite number of these cover X.

Example. Let X have the indiscrete topology. Then X is compact. Consider any collection O_λ (λ in Λ) of open sets that cover X. Then, since the only open sets in X are X and $\emptyset$, at least one of these O_λ must be X itself. This single O_λ itself covers X.

Example. Let X be the subspace of the real line consisting of the open interval $(0,1)$. Then X is not compact. Let $O_1 = (1/4,3/4)$, $O_2 = (1/8,7/8)$, $O_3 = (1/16,15/16)$, . . . , open subsets of X. These clearly cover X. But no finite number cover X (for, given any finite number of O_1, O_2, . . . , there is some small $\epsilon > 0$ (so ϵ is in $(0,1)$) which is in none of that finite number of open sets).

Example. Let X be the subspace of the real line consisting of the closed interval $[0,1]$. Then X is compact. (First, note that the technique of the previous example will not work here. The O_1, O_2, $\cdots$ of that example do not cover $[0,1]$, for the points 0 and 1 are in none of these open sets. Suppose that we add a couple of open sets to obtain a collection which does cover this X. Thus we might add $[0,1/16)$ and $(15/16,1]$. (Note that these are actually open sets in X. Since X is the subspace $[0,1]$ of the real line, the open sets in X are the intersections of $[0,1]$ with open sets in the real line.) Now we have a collection of open sets in X that cover X, but a finite number will also work: in this example, $(1/32,31/32)$, $[0,1/16)$, and $(15/16,1]$.) To prove that this X is compact, let O_λ (λ in Λ) be any collection of open sets that cover X. We suppose that no finite number cover X and obtain a contradiction. Since no finite number cover $X = [0,1]$, it must be true that either no finite number cover $[0,1/2]$ or no finite number cover $[1/2,1]$ (or both, in which case, pick one). Suppose that no finite number cover $[1/2,1]$. Then either no finite number cover $[1/2,3/4]$ or no finite number cover $[3/4,1]$: say, no finite number cover $[1/2,3/4]$. Then either no finite number cover $[1/2,5/8]$ or no finite number cover $[5/8,3/4]$, etc. We "keep dividing the half that no finite number cover in half, and asking which half no finite number cover." Repeating this process, we obtain successively smaller closed intervals, which must converge on some real number a in $[0,1]$. But our original collection O_λ of open sets cover $X = [0,1]$, whence one of them, say $O_{\underline{\lambda}}$, must contain a. But since the closed intervals in the construction above converge on a, one of them must be a subset of this $O_{\underline{\lambda}}$. But this closed interval does have the property that a finite number of the O_λ (λ in Λ) cover it, namely the single $O_{\underline{\lambda}}$ itself. This contradicts our construction of the closed intervals above. Hence X is compact.

Nets often provide an easy way to get a feeling of what is involved in something. The situation in the present case, as suggested by the previous two example, is this:

THEOREM 35. *Let X be a topological space. Then X is compact if and only if every net in X has an accumulation point.*

Proof. Suppose first that X is compact, and let x_δ (δ in Δ) be a net in X. We show that the assumption that this net has no accumulation points leads to a contradiction. For $\underline{x}$ any point of X, then, since $\underline{x}$ is not an accumulation point of the net, there is an open neighborhood $O_{\underline{x}}$ of $\underline{x}$ and a $\delta_{\underline{x}}$ in Δ such that no x_δ is in $O_{\underline{x}}$ with $\delta_{\underline{x}} \leq \delta$. These open $O_{\underline{x}}$ (as $\underline{x}$ ranges over X) certainly cover X. Since X is compact, a finite number also cover X: denote by A this finite collection of points $\underline{x}$ such that $\bigcup_{\gamma \text{ in } A} O_{\underline{x}} = X$. Since Δ is a directed set, there exists an element δ of Δ with $\delta_{\underline{x}} \leq \delta$ for all $\underline{x}$ in A. The corresponding point x_δ is therefore in no $O_{\underline{x}}$ ($\underline{x}$ in A). This contradicts the fact that the $O_{\underline{x}}$ ($\underline{x}$ in A) cover X.

Next, suppose that X is not compact. We construct a net on X that has no accumulation point. Since X is not compact, there exists a collection O_λ (λ in Λ) of open sets that cover X and that are such that no finite number cover X. Denote by Δ the directed set of all finite subsets of Λ, ordered by inclusion. For δ in Δ (so δ is a finite subset of Λ), set $K_\delta = \bigcup_{\lambda \text{ in } \delta} O_\lambda$, a union of a finite number of the O_λ. Since no finite number of the O_λ cover X, no K_δ is all of X: hence we may choose, for each δ in Δ, a point x_δ of $X - K_\delta$. We have thus constructed a net in X. We claim that it has no accumulation points. Let x be a point of X. (We want to show that this x is not an accumulation point of our net.) Since the O_λ cover X, there is some $O_{\underline{\lambda}}$ containing the point x. Denote by $\underline{\delta}$ the finite subset of Λ consisting of this single element $\underline{\lambda}$. Then, for $\underline{\delta} \leq \delta$ (i.e., for δ a finite subset of Λ containing the element $\underline{\lambda}$), we have $O_{\underline{\lambda}} \subset K_\delta$ (since $K_\delta = \bigcup_{\lambda \text{ in } \delta} O_\lambda$ and since $\underline{\lambda}$ is among the λ in the union). The corresponding element x_δ of our net is not in K_δ (by construction), and so this x_δ is also not in $O_{\underline{\lambda}}$. That is, we have found, for each x in X, a neighborhood $O_{\underline{\lambda}}$ of x and a $\underline{\delta}$ such that no x_δ is in $O_{\underline{\lambda}}$ for $\underline{\delta} \leq \delta$. In other words, we have shown that our net x_δ (δ in Δ) has no accumulation points. ☐

This difficult proof requires some explanation. For the first part, one supposes that X is compact and that some net in X has no accumulation point. Then every point of X has an open neighborhood such that "the net gets out of that neighborhood and remains out of that neighborhood" (for otherwise that point would be an accumulation point). But we have one of these open neighborhoods for each point of X. Thus we have a collection of open sets that cover X. By compactness, a finite number will serve to cover X. Now we have a finite number of open sets that cover X and such that, for each one, "the net gets out of it and remains out of it." But there are only a finite number of these open sets: how can the net "get out of and remain out of all of them"? Where is the net supposed to go? This is the contradiction for the first half of the proof. For the second half, one supposes that X is not compact—so one has some collection of open sets that cover X such that no finite number will do. Thus, given any finite number of these open sets, one can find a point of X outside all of them. These "points outside" are made into a net, where the directed set Δ is the set of all "finite numbers of open sets." This construction makes the net unable to have an accumulation point. Suppose a point x of X wanted to be an accumulation point. Since one of our open sets, $O_{\underline{\lambda}}$, contains x, all of the points of the net associated with a finite number of the open sets which include this $O_{\underline{\lambda}}$ must be outside this $O_{\underline{\lambda}}$. That is, the net "gets out of and remains out of this $O_{\underline{\lambda}}$." Thus x (for every point x of X) cannot be an accumulation point.

Now let X be a set, and consider the partially ordered set of all topologies on X. The coarser the topology on X, the fewer the number of open sets on X, the more difficult it will be to find a collection of open sets that cover X such that no finite number cover X, the more likely it will be that X is compact in this topology. More precisely: if some topology on X makes X compact, then any coarser topology on X also makes X compact. The situation is illustrated in figure 82.

Next, let X be a topological space, and let A be a subset of X. The subset A is said to be ***compact*** if the subspace A (a topological space based on set A) is compact as a topological space. Since the open sets in A are the intersections of A with the open sets in X, we have the following: the subset A is compact provided the following condition is satisfied: given any collection O_λ

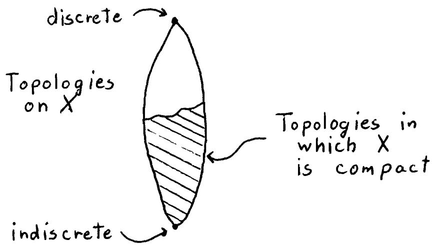

Figure 82

(λ in Λ) of open sets in X that cover A, a finite number of these also cover A.

One can think of a compact subset of X in the following terms. A subset A of X is closed if "all boundary points of A made available by X are contained in A." For compactness of A, we require somewhat more—that "all points which could conceivably be made available as boundary points (whether or not those points are actually made available by X) for A are included in A." Thus, for A to be closed, we require that "given any net in A, and if it has an accumulation point (in X), then that accumulation point be in fact a point of A," while compactness requires that every net in A actually have an accumulation point (in A). Thus, in particular, "whether or not A is closed depends on what X is, while whether or not A is compact depends only on the subspace A itself." Take the case $A = X$. The subset X of X is always closed (in X) (for, intuitively, X certainly contains all boundary points made available by X), while X may not be compact (since there may be other points which "could conceivably be made available to this X as boundary points").

The following two results illustrate these remarks.

THEOREM 36. *A closed subset C of a compact topological space X is compact.*

Proof. Let O_λ (λ in Λ) be any collection of open sets in X that cover C. Then, including the open set C^c to this collection, we obtain a collection of open sets in X that cover X. Since X is compact, a finite number cover X. Hence, in particular, this finite number cover C. If C^c happens to be in this finite number, we still have, on deleting it, a finite number of the O_λ that cover C (since C^c "does not contribute to covering C"). That is, given any collection of open sets in X that cover C, a finite number cover C; so C is compact. □

In intuitive terms, "since X is compact, every boundary point which could conceivably be made available to X is included in X," while, since C is closed in X, "every boundary point which X makes available to C is included in C."

Hence "every point which could conceivably be made available to C as a boundary point must be included in C," whence C is also compact. The second result is:

THEOREM 37. *Let C be a compact subset of Hausdorff topological space X. Then C is closed.*

Proof. Fix a point x of X that is not a point of C. Given any point c of C, then, since X is Hausdorff, there is an open neighborhood O_c of c and a neighborhood N_c of x, with $O_c \cap N_c = \emptyset$. The collection of all these O_c (as c ranges over C) clearly cover C. Since C is compact, a finite number cover C: denote by A this finite number of c such that $\bigcup_{c \text{ in } A} O_c \subset C$ (figure 83). Then $\bigcap_{c \text{ in } A} N_c$ (the intersection of a finite number of neighborhoods of x) is certainly a neighborhood of x. Since each N_c does not intersect the corresponding O_c and since the O_c for c in A cover C, this $\bigcap_{c \text{ in } A} N_c$ does not intersect C. Thus we have found, for each point x of X that is not a point of C, a neighborhood of x that does not intersect C. Hence (theorem 27) C is closed. □

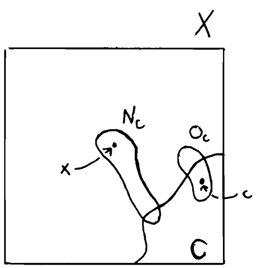

Figure 83

Intuitively, since C is compact, it must include "every point which could conceivably be made available to it as a boundary point," and so C must certainly include "every boundary point actually made available by X." Thus C must be closed.

As an example of these theorems, we find all the compact subsets of the real line. A subset of the real line is said to be bounded if it is a subset of some closed interval. Suppose first that C is a compact subset of the real line. Then C must certainly be bounded, for, were it not, one could find a sequence in C ("going off to infinity") having no accumulation point. (The existence of such a sequence would, by theorem 33, violate compactness of C.) Furthermore, this C must be closed (by theorem 37, since C is a compact subset of a Hausdorff topological space, namely the real line). Thus every compact subset of the real line is closed and bounded. Now suppose, conversely, that C is a closed and bounded subset of the real line. Then, since C is bounded, C is a subset of some closed interval. But, as we have seen, every closed interval in the real line is compact. Thus this C is a closed subset of a compact space, whence (by theorem 36) C must itself be compact. We have shown: a subset of the real line is compact if and only if it is closed and bounded. Finally, note that the closed and bounded subsets of the real line are precisely those which satisfy our intuitive picture of compactness.

Perhaps the most useful fact involving compactness is the following:

THEOREM 38. *Let* $X \xrightarrow{\varphi} Y$ *be a continuous mapping of topological spaces. Then, for* C *any compact subset of* X, $\varphi[C]$ *is a compact subset of* Y.

Proof. Let O_λ (λ in Λ) be a collection of open sets in Y that cover $\varphi[C]$. Then $\varphi^{-1}[O_\lambda]$ is a collection of open sets (since φ is continuous) in X that cover C. Since C is compact, a finite number of these $\varphi^{-1}[O_\lambda]$ cover C. The corresponding O_λ then cover $\varphi[C]$. Since, for every collection of open sets in Y that cover $\varphi[C]$, a finite number cover $\varphi[C]$, this $\varphi[C]$ is compact. □

Example. Let X be a compact topological space. Let $X \xrightarrow{\varphi} \mathbf{R}$ be a continuous mapping to the real line, so φ is a real-valued function on X. By theorem 38, $\varphi[X]$ is a compact subset of the real line, that is, a closed bounded subset of $\mathbf{R}$. But $\varphi[X]$ is just the range of the function φ, that is, the set of all values assumed by this function. Since $\varphi[X]$ is closed and bounded in $\mathbf{R}$, this set contains a maximum real number. Thus every continuous, real-valued function on a compact topological space assumes a maximum value.

Exercise 202. Prove that the direct product and direct sum of two compact topological spaces are compact.

Exercise 203. When is the topology generated by a collection of subsets of set X compact?

Exercise 204. Find all compact subsets of the topological plane.

Exercise 205. Let X be a topological space. Prove that the union of a finite number of compact subsets of X is compact and that the intersection of an arbitrary collection of compact subsets of X is compact. (Thus, if X itself is compact, the complements of compact subsets of X define a topology on X. When is this the original topology on X?)

Exercise 206. Let X be compact, Y Hausdorff, and $X \xrightarrow{\varphi} Y$ continuous. Prove that images of closed sets, by φ, are closed.

Exercise 207. Let X be a compact, Hausdorff topological space. Prove that, for any topology on X finer than, and different from, the given topology, X is no longer compact and that, for any topology on X coarser than, and different from, the given topology, X is no longer Hausdorff. (Compare "linearly independent" and "span.")

Exercise 208. Find, on the set of integers, a compact, Hausdorff topology.

Exercise 209. A topological space X is said to be locally compact if every point of X possesses a compact neighborhood. Prove that the real line is locally compact and that the rationals (as a subspace of the real line) is not.

Exercise 210. Let X be a locally compact topological space. Prove that there exists a compact topological space Y, and point y of Y, such that $Y - y$ is isomorphic to X. (Hint: One already knows what the set Y is. What are the open sets to be?)

31

The Compact-Open Topology

We have seen several examples of categories in which the set of morphisms from one object to another has more structure than just that of a set. For example, in the category of vector spaces, Mor(V, W) has the structure of a vector space. It is natural to ask what happens in this regard in the category of topological spaces.

Let X and Y be topological spaces, so Mor(X, Y) is the set of all continuous mappings from X to Y. We introduce a topology on this set. For C a compact subset of X, and O an open subset of Y, denote by $K(C, O)$ the collection of all continuous mappings $X \xrightarrow{\varphi} Y$ such that $\varphi[C] \subset O$. Thus, for each C and O, $K(C, O)$ is a subset of Mor(X, Y). The *compact-open topology* on Mor(X, Y) is that topology on this set generated by the subsets of the form $K(C, O)$.

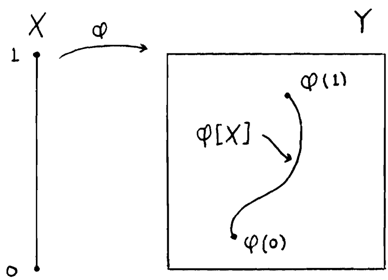

Figure 84

Example. Let X be the subspace [0,1] of the real line, and let Y be the topological plane. Thus a continuous mapping $X \xrightarrow{\varphi} Y$ is a "curve in Y, with an endpoint on each end." (See figure 84.) Then Mor(X, Y) is the set of all such curves. Fix one curve, $X \xrightarrow{\underline{\varphi}} Y$. We describe some open neighborhoods of this $\underline{\varphi}$, that is, some open sets in Mor(X, Y) (in the compact-open topology) containing the point $\underline{\varphi}$ of Mor(X, Y). Let C be the compact subset [0,1/10] of X = [0,1], and let O be some open subset of Y = topological plane with $\underline{\varphi}[C] \subset O$. Then the collection of all such curves $X \xrightarrow{\varphi} Y$ with $\varphi[C] \subset O$ is a neighborhood of our curve $\underline{\varphi}$. For this neighborhood, one places no restriction on "where the curve goes for x in X outside [0,1/10]." (See figure 85.) Thus there are many curves in this neighborhood $K(C, O)$, some quite different from our original curve $\underline{\varphi}$. In other words, this is a "rather large" neighborhood of the point $\underline{\varphi}$ of Mor(X, Y). To get "smaller neighborhoods," we take an intersection of several (a finite number!) of these. Let $C_1 = [0,1/10]$, $C_2 = [1/10,2/10], \ldots, C_{10} = [9/10,1]$, compact subsets of X. Choose $O_1, \ldots, O_{10}$ open subsets of Y with $\underline{\varphi}[C_1] \subset O_1, \ldots, \underline{\varphi}[C_{10}] \subset O_{10}$. (To get a good small neighborhood of $\underline{\varphi}$, make these O small.) Then $K(C_1, O_1) \cap \cdots \cap K(C_{10}, O_{10})$, the intersection of 10 neighborhoods of $\underline{\varphi}$, is a neighborhood of $\underline{\varphi}$. A curve $X \xrightarrow{\varphi} Y$ is in this neighborhood provided $\varphi[C_1] \subset O_1, \ldots,$ and $\varphi[C_{10}] \subset O_{10}$. Thus a curve φ, in order that it be in this neighborhood, must "not differ very much from $\underline{\varphi}$ along the whole of [0,1]." The compact-open topology on Mor(X, Y) reproduces in this example one's intuitive notion of "nearby curves."

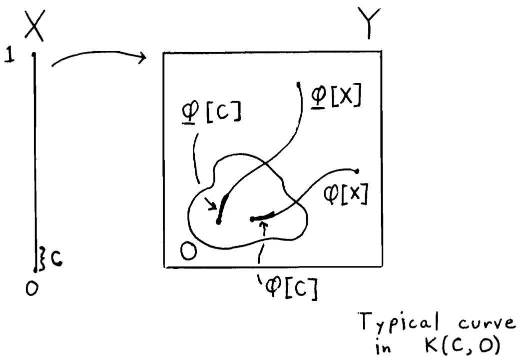

Figure 85

Example. Let X and Y each be the real line. Then a continuous mapping $X \xrightarrow{\varphi} Y$ is just a continuous real-valued function of one real variable. We can represent such a function by its graph as in figure 86. Then Mor(X, Y) is the set of all such continuous functions. Fix one continuous $X \xrightarrow{\underline{\varphi}} Y$. Then a typical neighborhood of this $\underline{\varphi}$, in the compact-open topology on Mor(X, Y), would be the following. Fix some compact subsets, $C_1, \ldots, C_5$, of X, and some open subsets, $O_1, \ldots, O_5$, of Y, with $\underline{\varphi}[C_1] \subset O_1, \ldots, \underline{\varphi}[C_5] \subset O_5$. Then the corresponding neighborhood consists of all continuous $X \xrightarrow{\varphi} Y$ with $\varphi[C_1] \subset O_1, \ldots, \varphi[C_5] \subset O_5$ as shown in figure 87. Note that a function φ can be in this neighborhood and yet "do anything it wants (provided only that it be continuous)" outside $C_1, \ldots, C_5$. In particular, since these C are compact, this neighborhood "does not restrict functions for large x."

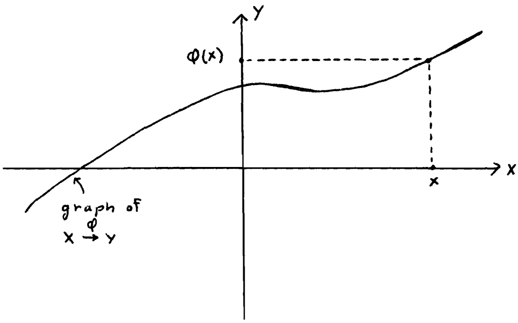

Figure 86

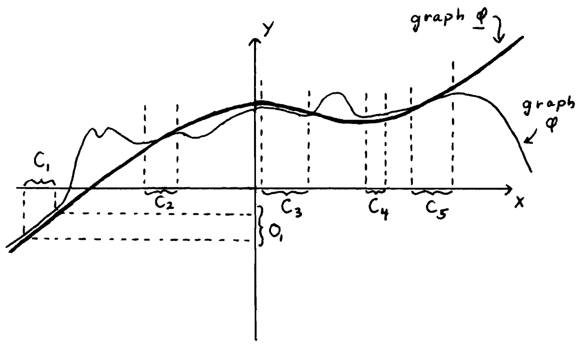

Figure 87

In this example the compact-open topology on Mor(X, Y) perhaps does not completely reproduce one's intuitive picture of "nearby functions." The compact-open topology seems just a bit too coarse (because "one cannot get neighborhoods as small as one might like"). We introduce another topology on the set Mor(X, Y), for this example, which is finer than the compact-open topology. Given any continuous $X \xrightarrow{\underline{\varphi}} Y$, and any number $\epsilon > 0$, denote by $L(\underline{\varphi}, \epsilon)$ the collection of all curves φ with $|\varphi(x) - \underline{\varphi}(x)| < \epsilon$ for all x in X (figure 88). Consider the topology on Mor(X, Y) for which a subset of Mor(X, Y) is open provided that, for any $\underline{\varphi}$ in that subset, there exists an $\epsilon > 0$ such that all φ in $L(\underline{\varphi}, \epsilon)$ are also in that subset. In this topology, a neighborhood of function $\underline{\varphi}$ consists, for example, of "all functions which are everywhere within ϵ of the function $\underline{\varphi}$." Clearly, this topology is finer than the compact-open topology and perhaps closer to one's intuitive picture of "nearby functions." (Note that this last topology uses essentially the fact

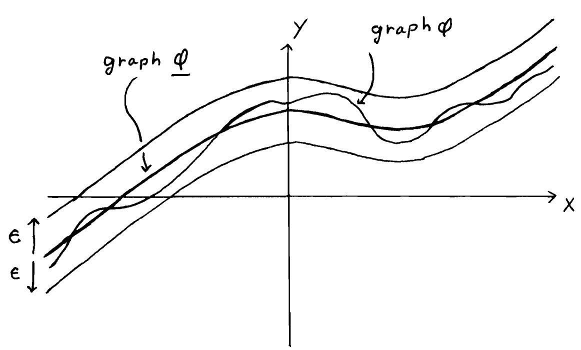

Figure 88

that Y is not any old topological space, but rather the real line.)

There are still finer topologies, for this example, on the set Mor(X, Y). Consider this one: given any open set O in $X \times Y$ (direct product), consider the collection of all functions $X \xrightarrow{\varphi} Y$ whose graphs lie in O (figure 89). These are the open sets for a topology on Mor(X, Y). In this topology, a neighborhood of $X \xrightarrow{\underline{\varphi}} Y$ consists of functions φ which "are permitted to differ from $\underline{\varphi}$ by a little bit for each x, where this 'little bit' must vary continuously, but otherwise arbitrarily, with x." (The "continuously varying 'little bit' is described by the open set O in the topological plane $X \times Y$.") This topology differs from the previous one in that, there, the "little bit" was required to remain constant—at ϵ; it could not "get smaller and smaller, for example, for large x." Clearly, the present topology is finer than its predecessor.

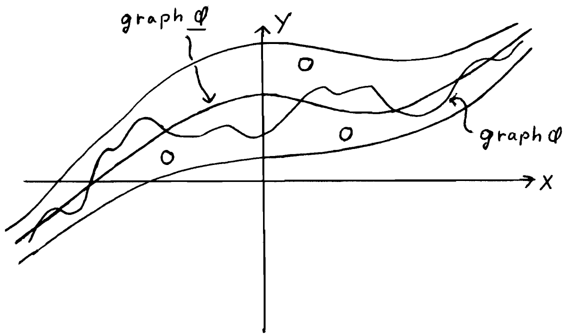

Figure 89

Of course, one could invent other topologies in this and in other situations. The topology one chooses in practice depends on what one wants the topology to do. If one had to make a general statement about the compact-open topology, perhaps it would be this: when X is compact, the compact-open topology on Mor(X, Y) is usually the "right" one; when X is not compact, this topology is often too coarse.

Exercise 211. Let X and Y each be the real line, and consider the set Mor(X, Y) of the last example above. Let $R \xrightarrow{\psi} \text{Mor}(X, Y)$ be the mapping which sends real number a to the element φ of Mor(X, Y) with $\varphi(x) = a$ for all x in X (i.e., each real number goes to that constant function). Prove that this ψ is continuous when Mor(X, Y) is given either of the first two topologies above and is not continuous for the third topology.

Exercise 212. Let X and Y be topological spaces, and let $X \times \text{Mor}(X, Y) \xrightarrow{\psi} Y$ be given by the mapping which sends (x,φ) to $\varphi(x)$ in Y. Assigning Mor(X, Y) the compact-open topology, when is ψ continuous?

Exercise 213. Let X, Y, and Z be topological spaces, and let $\text{Mor}(X, Y) \times \text{Mor}(Y,Z) \xrightarrow{\psi} \text{Mor}(X,Z)$ be composition. Using the compact-open topology for each Mor(,), when is ψ continuous?

Exercise 214. Prove that, if X and Y are both compact, so is Mor(X, Y). Prove that, if Y is Hausdorff, so is Mor(X, Y).

Exercise 215. Let X and Y be topological spaces. For $X \xrightarrow{\varphi} Y$ continuous, let K_φ be the subset of $X \times Y$ consisting of all pairs (x,y) with $y = \varphi(x)$ (the "graph" of φ). For each open set O in $X \times Y$, consider the subset of Mor(X, Y) consisting of all $X \xrightarrow{\varphi} Y$ with $K_\varphi \subset O$. Show that these subsets are the open sets for a topology on Mor(X, Y). Show that this topology is finer than the compact-open topology on Mor(X, Y). Prove that, when X is compact, the two topologies coincide.

Exercise 216. Let X and Y be topological spaces. For x a point of x and O open in Y, let $K(x,O)$ denote the subset of Mor(X, Y) consisting of $X \xrightarrow{\varphi} Y$ with $\varphi(x)$ a point of O. The topology on Mor(X, Y) generated by these subsets is called the point-open topology. Prove that the point-open topology is coarser than the compact-open topology. When do they coincide?

Connectedness

One has an intuitive notion of what it means to say, of a topological space X, that "X is not connected; it consists of several pieces which do not touch each other." We now wish to introduce a definition in topology which represents this idea. To get a notion of what is involved, consider the following situation. Let X and Y be topological spaces, and let Z be their direct sum, with $X \xrightarrow{\alpha} Z$ and $Y \xrightarrow{\beta} Z$ the corresponding continuous mappings. Then one would certainly think of $\alpha[X]$ (the "copy of X in Z") and $\beta[Y]$ as "separate pieces of Z, pieces which cause Z not to be connected." We look for some characteristic property of the subset $\alpha[X]$ of Z which might be used to formulate a notion of connectedness. There is indeed one: the subset $\alpha[X]$ of Z is both open and closed in Z.

The remarks above are motivation. Let X be a topological space. Then X is said to be *connected* if the only subsets of X which are both open and closed are X itself and the empty set (each of these always being an open and closed subset of X). (Note that, by theorem 29, a subset A of X is both open and closed if and only if the boundary of A is empty. Thus X is connected if and only if the only subsets of X having empty boundary are X and $\emptyset$.)

Example. Let X be a set having two or more points, with the discrete topology. Then X is not connected. Indeed, each point x of X is both open and closed.

Example. Let X be a set with the indiscrete topology. Then X is connected, since the only open and closed subsets of X (in fact, the only open subsets of X) are X and $\emptyset$.

Example. Letting X be the real line, X is connected. Indeed, let A be any nonempty subset of the real line which is both open and closed. Choose a number a in A. If A does not include all numbers greater than a, there exists a number b such that A is a superset of $[a,b)$ and such that b is the largest number with this property. Since A is closed, b must be a point of A. But now, since A is open, A must contain a neighborhood of b, which contradicts our assertion that b must be the largest number having the property above. Hence A includes all real numbers greater than a and, similarly, all real numbers less than a. Thus $A = \mathbf{R}$. Since the only nonempty subset of R which is both open and closed is R itself, R is connected.

Example. Let X be the subspace of the real line consisting of the rational numbers. Then X is not connected. Let A be the subset of X consisting of all rational numbers a with $-\pi < a < +\pi$ (where π is that irrational number). Then A is an open subset of X (for it is the intersection of X with the open subset $(-\pi,+\pi)$ of the real line), and A is also a closed subset of X (for it is the intersection of X with the closed subset $[-\pi,+\pi]$ of the real line).

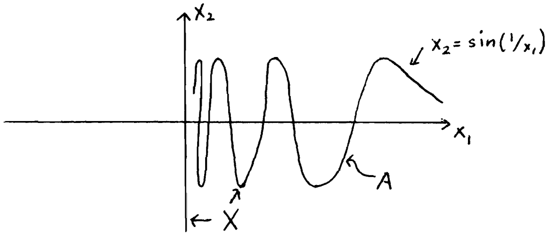

Figure 90

Example. Let X be the subspace of the topological plane consisting of all points (x_1,x_2) of that plane with either i) $x_1 > 0$ and $x_2 = \sin(1/x_1)$, or ii) $x_1 = 0$. (See figure 90.) This topological space X is in fact connected. The best candidate for a subset of X both open and closed is the following: let A be the subset of X consisting of all points (x_1,x_2) with $x_1 > 0$ and $x_2 = \sin(1/x_1)$. Then A is indeed an open subset of X, for A is the intersection of X with the open subset O of the topological plane consisting of all (x_1,x_2) with $x_1 > 0$. However, this A is not a closed subset of X: there is no closed subset C of the topological plane whose intersection with X is A. (For example, the closed C consisting of (x_1,x_2) with $x_1 \geq 0$ will not do, for, for this C, $C \cap X = X$.)

Let X be a set. It is immediate from the definition that, given some topology on X such that X is connected, any coarser topology on X also makes X connected. (See figure 91.)

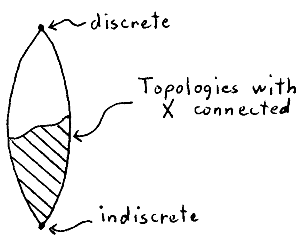

Figure 91

Let X be a topological space, and let A be a subset of X. This subset A is said to be ***connected*** if the subspace A (a topological space based on set A) is connected. Recall, however, that the open sets in the subspace A are the intersections of A with open sets in X, while the closed sets in the subspace A

are the intersections of A with closed sets in X. Clearly, then, subset A of X is connected if and only if the following property is satisfied: given any open O in X and closed C in X, with $O \cap A = C \cap A$, we have either $O \cap A = A$ or $O \cap A = \emptyset$. Thus, for example, the rationals do not form a connected subset of the real line, while the last example above gives a connected subset of the topological plane. Any subset of a topological space containing either just one, or no, points is connected.

We give two examples of properties that are satisfied by this notion of connectedness.

THEOREM 39. *Let X be a topological space, and A a connected subset of X. Then $\mathrm{Cl}(A)$ is also connected.*

Proof. Let O be an open subset of X, and C a closed subset of X, with $O \cap \mathrm{Cl}(A) = C \cap \mathrm{Cl}(A)$. We must show that either $O \cap \mathrm{Cl}(A) = \mathrm{Cl}(A)$ or $O \cap \mathrm{Cl}(A) = \emptyset$. We have, in particular, $O \cap A = C \cap A$, whence, since A is connected, either $O \cap A = A$ or $O \cap A = \emptyset$. If $O \cap A = A$, then $C \cap A = A$, that is, A is a subset of C, whence (by definition of closure) $\mathrm{Cl}(A)$ is a subset of C, that is, $C \cap \mathrm{Cl}(A) = \mathrm{Cl}(A)$, whence $O \cap \mathrm{Cl}(A) = \mathrm{Cl}(A)$. If $O \cap A = \emptyset$, that is, A is a subset of O^c, then (by definition of closure) $\mathrm{Cl}(A)$ is a subset of O^c, that is, $O \cap \mathrm{Cl}(A) = \emptyset$. Thus either $O \cap \mathrm{Cl}(A) = \mathrm{Cl}(A)$ or $O \cap \mathrm{Cl}(A) = \emptyset$, whence $\mathrm{Cl}(A)$ is connected. □

The statement of theorem 39 is in agreement with one's intuitive picture of connectedness. It says that "if you attach, to a connected set, its boundary, the result is again a connected set."

THEOREM 40. *Let X be a topological space. Let A_λ (λ in $\Lambda \neq \emptyset$) be a collection of connected subsets of X, any two of which intersect. Then their union, $\bigcup_\Lambda A_\lambda$, is also connected.*

Proof. Let O be an open subset of X, and C a closed subset of X, with $O \cap (\bigcup_\Lambda A_\lambda)$. We must show that $O \cap (\bigcup_\Lambda A_\lambda)$ is either $\bigcup_\Lambda A_\lambda$ or $\emptyset$. Fix one of these sets, $A_{\underline{\lambda}}$. Then, since this $A_{\underline{\lambda}}$ is connected, we have either $O \cap A_{\underline{\lambda}} = A_{\underline{\lambda}}$ or $O \cap A_{\underline{\lambda}} = \emptyset$, say $O \cap A_{\underline{\lambda}} = A_{\underline{\lambda}}$. Then, since each A_λ intersects $A_{\underline{\lambda}}$ and hence O, and since each A_λ, being connected, satisfies $O \cap A_\lambda = A_\lambda$ or $O \cap A_\lambda = \emptyset$, we must have $O \cap A_\lambda = A_\lambda$ for each λ. That is, we must have $O \cap (\bigcup_\Lambda A_\lambda) = \bigcup_\Lambda A_\lambda$. Similarly, the supposition $O \cap A_{\underline{\lambda}} = \emptyset$ leads to $O \cap (\bigcup_\Lambda A_\lambda) = \emptyset$. □

Intuitively, since each A_λ is connected, that is, since each "cannot be divided into separate pieces which do not touch, and since all the A_λ overlap with each other," the union of the A_λ is also connected. (See figure 92.)

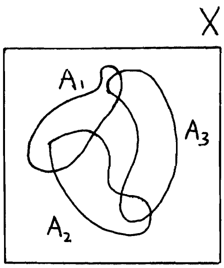

Figure 92

We next introduce an important tool for analyzing the connectivity properties of a topological space. Let X be a topological space. A subset C of X is called a *connected component* of X if i) C is a connected subset of X, and ii) no superset of C (except C itself) is connected. Thus the connected components of X are "maximal connected subsets."

Example. For X a connected topological space, the only connected component of X is X itself. For X discrete, the connected components of X are just the points of X.

Example. If X and Y are connected topological spaces, and Z is their direct sum (with $X \xrightarrow{\alpha} Z$ and $Y \xrightarrow{\beta} Z$), then there are two connected components of Z, namely $\alpha[X]$ and $\beta[Y]$.

Example. The connected components of the rationals (regarded as a subspace of the real line) are precisely the points of this topological space (exercise 217).

It is immediate from theorem 39 that connected components are always closed subsets. However, we have as yet no guarantee that, given a topological space, it even has any connected components. The following theorem not only settles this question, but shows that "the connected components of X cover X without overlapping."

THEOREM 41. *Let X be a topological space. Then each point of X is a point of one and only one connected component of X.*

Proof. Let x be a point of X, and let C denote the union of all connected subsets of X which contain x. Then (since any two sets in this union intersect, namely at x) we have from theorem 40 that C is connected. Any connected superset of C containing x would have been included in the union which defined C and hence must be C itself. Thus C is a connected component containing x. If C' were another, then, by theorem 40, $C \cup C'$ would be connected, whence, since C and C' are connected components, we must have $C \cup C' = C$ and $C \cup C' = C'$. That is, we must have $C = C'$. Hence there is a unique connected component containing x. ☐

Thus the connected components of X provide a "subdivision of X into disjoint connected pieces." Note, incidentally, that every connected subset of X is a subset of some connected component of X. [Proof: Let A be a connected component of X containing x. Then, by theorem 40, $A \cup C$ is connected, whence, since C is a connected component, $A \cup C = C$, i.e., A is a subset of C.]

Finally, we consider the behavior of connected sets under continuous mappings. The main result, analogous to theorem 38 in the compact case, is this:

THEOREM 42. *Let* $X \xrightarrow{\varphi} Y$ *be a continuous mapping of topological spaces, and let* A *be a connected subset of* X. *Then* $\varphi[A]$ *is a connected subset of* Y.

Proof. Let O be open, and C closed, in Y, and let $O \cap \varphi[A] = C \cap \varphi[A]$. Then, by continuity of φ, $\varphi^{-1}[O]$ is open and $\varphi^{-1}[C]$ is closed, in X, and $\varphi^{-1}[O] \cap A = \varphi^{-1}[C] \cap A$. Since A is connected, we have either $\varphi^{-1}[O] \cap A = A$ (in which case $O \cap \varphi[A] = \varphi[A]$ or $\varphi^{-1}[O] \cap A = \emptyset$ (in which case $O \cap \varphi[A] = \emptyset$). Thus $\varphi[A]$ is connected in Y. $\square$

Intuitively, since, for a continuous φ, "$\varphi(x)$ cannot change a great deal with small changes in x," it follows that, for A connected, $\varphi[A]$ cannot consist of several pieces which do not touch. For example, it is immediate from theorem 42 that, for Y discrete and $X \xrightarrow{\varphi} Y$ continuous, each connected component of X is taken, by φ, to a single point of Y. As another simple application of theorem 42, we have

Example. Let X be a topological space, and suppose that X has the following property: for any two points, x and x', of X, there is a continuous mapping $\mathbf{R} \xrightarrow{\varphi} X$ with $\varphi(0) = x$ and $\varphi(1) = x'$ (i.e., "there is a curve in X which passes through both x and x'"). This is illustrated in figure 93. Then (as one would expect) it follows that X is connected. Proof: Fix a point x of X, and let C be the connected component of X containing x. Since the real line $\mathbf{R}$ is connected, for each continuous mapping $\mathbf{R} \xrightarrow{\varphi} X$, $\varphi[\mathbf{R}]$ is connected in X. Hence, for each such curve with $\varphi(0) = x$, $\varphi[\mathbf{R}]$ is a subset of C. But, for every point x' of X there is a curve with $\varphi(0) = x$ and $\varphi(1) = x'$; hence every point x' of X is a point of C. That is, $C = X$, whence X is connected.

Exercise 217. A topological space X is said to be totally disconnected if every connected component of X consists of but a single point. Show that the rationals (a subspace of the real line) is totally disconnected.

Exercise 218. Prove that no subset of a Hausdorff topological space, with the subset consisting of exactly two points, can be connected.

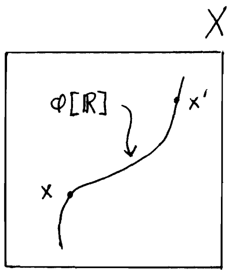

Figure 93

Exercise 219. Prove the following generalization of theorem 39. If A is a connected subset of topological space X and if B is a superset of A and a subset of Cl(A), then B is connected.

Exercise 220. Find an example (e.g., in the topological plane) of two connected sets whose intersection is not connected.

Exercise 221. Find an example of sets A and B, neither of which are connected and which do not intersect each other, such that $A \cup B$ is connected. Find a connected set whose interior is not connected.

Exercise 222. Find an example to show that the converse of the result of the last example above is false.

Exercise 223. Let X be a topological space having only a finite number of connected components. Prove that each connected component of X is open.

Exercise 224. Let A and B be connected subsets of topological space X such that A intersects Cl(B). Prove that $A \cup B$ is connected. Find an example to show that it is not enough to assume that Cl(A) intersects Cl(B).

Exercise 225. Prove the following generalization of theorem 40. Let A_λ (λ in $\Lambda \neq \emptyset$) be a collection of connected subsets of topological space X, and let the equivalence relation on this collection of sets generated by $A_\lambda \approx A_{\lambda'}$ if $A_\lambda \cap A_{\lambda'} \neq \emptyset$ have just one equivalence class. Then $\bigcup_\Lambda A_\lambda$ is connected.

Exercise 226. Let X be a topological space. Write $x \approx x'$ if there exists a connected subset of X containing both x and x'. Prove that this is an equivalence relation. What are the equivalence classes? Prove that the quotient space by this equivalence relation is totally disconnected.

33

Example: Dynamical Systems

A topology on a set endows that set with "a notion of closeness of its points." Such a structure is often what is needed for physical applications, for the following reason. We examine physical systems—we learn about their structure—by making observations. But any physical observation, apparently, is necessarily accompanied by a certain error. Thus, from a physical viewpoint, we are concerned not so much with "exactly what the situation is" as with "what the situation is to within the errors of our observations." In other words, "sufficiently close is all that is physically relevant anyway, where 'what is sufficient' depends on how good our measuring instruments are." One expects that the notion of a topology will be appropriate for the description of these physical ideas. As an example of an application of topology, we now consider the description of (e.g., classical) dynamical systems.

We imagine some physical system that we wish to study. It is perhaps difficult to say exactly what one means by "a physical system," but we imagine it as consisting of some mechanism which sits in a box on a table, unaffected by "uncontrollable external influences." We shall, however, allow ourselves to influence the system (e.g., by hitting it with a stick) so that we may manipulate the system to study its properties. We introduce the notion of "the state of the system," where we think of the state of the system as a complete description of what every part of the system is like at a given instant of time. (For example, the state of a harmonic oscillator is specified by giving its position and momentum.) Thus, at each instant of time, the system is in a certain state, and all we can ever hope to know about the system, at a given instant, is what state it is in. By an extended series of manipulations on the system, we "discover all the states which are available to it," that is, we introduce a set Γ whose points represent the states of our system. Thus our mathematical model of the system so far consists simply of a certain set Γ.

We next decide that, more or less, we know what it means physically to say that two states of the system (i.e., two points of Γ) are "nearby." Roughly speaking, two states are "nearby" if "the system does not have to change all that much in passing from the first state to the second." We wish to incorporate this physical idea as mathematical structure on the set Γ. The notion of a topology seems to serve this purpose well. Thus we suppose that the set of states, Γ, of our system is in fact a topological space, where the topology on Γ reflects "physical closeness" of states. This topological space Γ is usually called the phase space of the system.

We have now completed the description of the kinematics of our system—what possibilities are available to it. The following remarks are intended to illustrate further what is going on.

How would one describe an observable of the system within this framework? We think of an observable as an instrument with a dial such that, when the instrument is brought into contact with the system, the dial reads a certain real number. Thus an observable assigns a real number to each state of the system, that is, it is a mapping of sets, $\Gamma \xrightarrow{\varphi} \mathbf{R}$. (That an observable assign a unique number to each state of the system is, if you like, part of what we mean by a "state.") We now wish to claim that it is only the continuous mappings (from topological space Γ to the real line) that one should regard as observables. There are at least two points of view one could take toward this statement. On the one hand, a discontinuous mapping, if taken as an observable, could be regarded as "unphysical," or at least as "unrepeatable," for even arbitrarily small changes in the state of the system (which should make no difference from a physical viewpoint) could result in large changes in the value of the observable. On the other hand, one could regard this continuity as part of what we mean physically by the topology which is to be selected for Γ. (If there are things that we want to call observables and if they are not continuous in a given topology on Γ, then we should choose a finer topology on Γ so that they become continuous.) In any case, the physical idea of an observable is incorporated into the mathematics as a continuous, real-valued function on Γ.

Is it reasonable to assume that Γ is Hausdorff? If not, we would have distinct points x and x' of Γ such that every neighborhood of x intersects every neighborhood of x'. That is, we would have "two distinct states, x and x', such that, if you tell me how close you want to be to x and how close you want to be to x', I can find a state which is within the given tolerances of both x and x'." This contradicts one's intuitive picture of what one wants to mean by "distinct states." It therefore seems reasonable, physically, to suppose that Γ is Hausdorff.

Consider next two systems, with phase spaces Γ_1 and Γ_2. We now "consider these two separate systems as one," that is, we put them side by side in a large box without allowing them to interact. What should be the phase space of the combined system? To specify a state of the combined system, one would have to give the state of each of the two subsystems, that is, one would have to give a point of Γ_1 and a point of Γ_2. We therefore choose, for the set Γ of states of the combined system, the Cartesian product of sets, $\Gamma_1 \times \Gamma_2$. We must now select a suitable topology on this set Γ. Physically, "two states, (x_1,x_2) and $(x_1{}',x_2{}')$, of the combined system may be regarded as nearby if x_1 and $x_1{}'$ (states of the system with phase space Γ_1) are close, and also x_2 and x_2' are close." We incorporate this idea into the mathematics by placing on Γ the product topology (for that is, intuitively, just what it is). Thus we choose, for the phase space Γ of the combined system $\Gamma_1 \times \Gamma_2$, the direct product of topological spaces.

We next wish to describe the dynamics of our system—how it actually evolves with time. It is observed, in the physical world, that prediction is normally possible for physical systems: if you tell me the state of the system now and how long you are going to wait (and if you do not interfere with the system during its evolution), then I can tell you what state the system will be in at the end of that time. This dynamical information would be described by a mapping of sets, $\Gamma \times \mathbf{R} \xrightarrow{\psi} \Gamma$, where there appears on the left the Cartesian product of sets: for x a point of Γ, and t a real number, $\psi(x,t)$ is to represent the state in which the system will be after elapsed time t if it was initially in state x. (Thus, e.g., $\psi(x,0) = x$.) This mapping ψ is, of course, to be determined physically by merely watching the system as it evolves.

It seems normally to be the case for actual physical systems that this $\Gamma \times \mathbf{R} \xrightarrow{\psi} \Gamma$ is in fact a continuous mapping of topological spaces (where there now appears on the left the direct product of topological spaces). Let us see what this means physically. For fixed initial state $\underline{x}$, continuity of $\psi(\underline{x},t)$ in t means that "the system does not, during the course of its evolution, change its state suddenly with time." For fixed time $\underline{t}$, continuity of $\psi(x,\underline{t})$ in x means that "the state the system will be in a time $\underline{t}$ from now does not change discontinuously with the present state of the system." In fact, continuity of $\Gamma \times \mathbf{R} \xrightarrow{\psi} \Gamma$ seems to be necessary in order that one be able to make accurate predictions. One can specify the initial state of the system, x, and the elapsed time, t, only to within a certain error. Were it not true that one could, by making these errors small, make as small as one wished the error in the predicted final state of the system, $\psi(x,t)$, one's "prediction" would be subject to large and essentially uncontrollable errors. This would hardly be a prediction at all. Thus the dynamics of the system is described by a continuous mapping $\Gamma \times \mathbf{R} \xrightarrow{\psi} \Gamma$ of topological spaces.

We give two examples of consequences of this formulation of the dynamics of our system.

Suppose that the phase space Γ of our system were compact. It then follows that the system is almost-periodic, in the following sense: there exists a state x such that, given any neighborhood N of x and any time t, the system can begin in N and return to N after some elapsed time greater than t. (That is, "the system continually returns, as closely as you wish, to the state x.") Proof: Fix point $\underline{x}$, and consider $\psi(\underline{x},t)$ a mapping from $\mathbf{R}$ to Γ. Since $\mathbf{R}$ is a directed set, this is a net in Γ. Since Γ is assumed compact, this net has some accumulation point, x. The statement that this x is an accumulation point of our net is precisely the statement of almost-periodicity above.

What is the significance of Γ having more than one connected component? Fix an initial state $\underline{x}$, and regard $\psi(\underline{x},t)$ (the evolution of the system from this initial state) as a mapping (necessarily continuous) from **R** to Γ. Since the real line **R** is connected, the image of **R** under this mapping (the set of "states the system visits as it evolves from $\underline{x}$") is necessarily connected. Hence all of these states are in the same connected component of Γ as is $\underline{x}$. That is, the system remains throughout the course of its evolution in a single connected component. (It is perhaps not unreasonable, therefore, to require that Γ have a single connected component, i.e., that Γ be connected.)

Finally, we show how the topology on the phase space can be used to formulate the notion of stability for a system. A point $\underline{x}$ of Γ will be called a stationary state if $\psi(\underline{x},t) = \underline{x}$ for all t. Thus the system, initially in a stationary state, remains in that state for all time as it evolves. Stability of a stationary state refers to the behavior of the evolution from nearby initial states (figure 94). Does the system in the course of such an evolution remain near the stationary state $\underline{x}$, or does it eventually wander away?

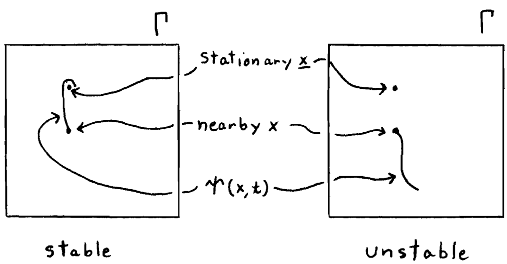

Figure 94

We may call a stationary state $\underline{x}$ weakly stable if it satisfies the following property: given any neighborhood N' of $\underline{x}$, there exists a neighborhood N of $\underline{x}$, with N a subset of N', such that, for any x in N, $\psi(x,t)$ is in N for all t. In other words, weak stability of $\underline{x}$ requires that there exist "arbitrarily small neighborhoods of $\underline{x}$ which the system, once in, always remains in." This is illustrated in figure 95. The stationary state of a harmonic oscillator, for example, is weakly stable in this sense.

We may call a stationary state $\underline{x}$ strongly stable if it satisfies the following property: there exists a neighborhood N of $\underline{x}$ such that, given any neighborhood M of $\underline{x}$, there is a number t_0 with $\psi(x,t)$ in M for all x in N and $t > t_0$. Thus strong stability of $\underline{x}$ requires that, "if you start off within N, you eventually get, and remain thereafter, as close as you like to $\underline{x}$." Note that a harmonic oscillator is not strongly stable at its stationary state but that, for example, a damped harmonic oscillator is.

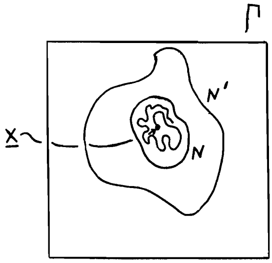

Figure 95

Of course, one could invent, using only the topology of phase space, other notions of stability. More generally, other intuitive properties of the system, provided only that they involve no more structure than that of "closeness," should be expressible in topological terms. Still more generally, if one wishes to speak of things other than just "closeness of states," then one introduces other, appropriate, structure on the set Γ.

It should be emphasized, finally, that topology does not "tell one anything about the way nature behaves." Rather, topology is just a few definitions, a few constructions, and a few theorems which often happen, for some reason, to provide a convenient and appropriate framework for the description of the way nature actually does behave. It is just like the notion of a derivative: derivatives just exist as mathematics, but often happen to find application to physics. In neither case does the mathematics serve, in any sense of which I am aware, as a substitute for the physics.

34

Homotopy

It is often the case that a complete and detailed examination of the internal structure of some given topological space can become quite a complicated business. One would like to have available simpler ways to look at a space. It is natural, therefore, to search for various characterizations of a topological space that refer not so much to its detailed structure as to its structure as a whole. We have already seen one example of such a characterization: that of connectedness (which refers to "the possibilities for dividing the space into several pieces"). We now wish to study another global property of topological spaces, this one referring to "the multiple-connectedness of the space, that is, the presence of holes or of interconnections between various regions." Consider, for example, the ***punctured plane,*** the topological space, as illustrated in figure 96, consisting of the subspace of the topological plane given by $(x^1)^2 + (x^2)^2 > 1$. Clearly, the "presence of a hole" represents an intuitive property of this space not possessed, for example, by the topological plane itself. This is the type of structure we now wish to examine.

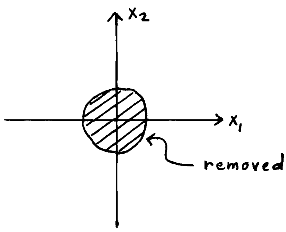

Figure 96

In fact, there are two different approaches to this intuitive idea of multiple-connectedness—that via homotopy and that via homology. These two approaches are very similar not only in terms of the structure they describe, but also technically. Nonetheless, we shall consider them both. Homotopy is perhaps conceptually somewhat simpler, and leads to an important notion with a wide variety of applications—that of a "continuous deformation." Homology, on the other hand, is very often what is actually used in practice, for example, for the description of regions of integration. We shall discuss homotopy in the present chapter and homology in the next.

Fix, once and for all, a topological space X and a point x of X. Denote by I the subspace of the real line consisting of the closed interval [0,1]. By a *loop* (based at x) we mean a continuous mapping $I \xrightarrow{\omega} X$ with $\omega(0) = \omega(1) = x$. Thus, for each real number t in $[0,1] = I$, $\omega(t)$ is a point of X: we have a

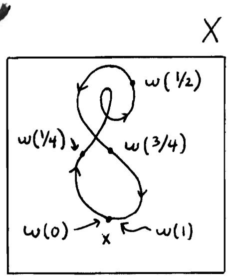

Figure 97

"curve in X which begins at x and ends at x." (Note, however, that by a loop we mean the continuous mapping ω, not just the subset $\omega[I]$ of X.) A loop is illustrated in figure 97. The loop given by $\omega(t) = x$ for all t in I (note that this ω is necessarily continuous) will be called the *zero loop*. Next, denote by $\Omega(X,x)$, the *loop space* of X (based at x), the collection of all such loops. Thus, for example, if X is discrete, then (since I is connected, since the image of a connected set under a continuous mapping is connected, since no connected subset of X has more than one point) the only loop based at x is the zero loop, whence $\Omega(X,x)$ has just this one element.

The loop space is not quite the right thing to describe the "multiple-connectedness structure" we have in mind. The problem is that it is far too large. For example, the topological plane is "not very multiply-connected," although it has an enormous loop space. What we wish to do is "regard two loops as essentially the same if one can be continuously deformed into the other" and then consider only "loops up to such continuous deformations."

Let x be a point of topological space X, and let ω and ω' be two loops based at x. We say that ω and ω' are *homotopic* if there exists a continuous mapping $I \times I \xrightarrow{\psi} X$ (where we shall write, for s in I and t in I, $\psi_s(t)$ instead of $\psi(s,t)$) satisfying the following conditions: i) $\psi_s(0) = x$ and $\psi_s(1) = x$ for all s in I, ii) $\psi_0(t) = \omega(t)$ for all t in I, and iii) $\psi_1(t) = \omega'(t)$ for all t in I. Let us see what this definition means geometrically. For each fixed value of s, ψ_s is a continuous mapping from I to X (namely, the mapping which sends t in I to $\psi_s(t)$). Condition i) just guarantees that this mapping is a loop. Thus each ψ_s is a loop, so we have a one-parameter family (parameterized by s in I) of loops. Condition ii) requires that "the first loop in this family," ψ_0, is the given loop ω, while condition iii) requires that the "last loop," ψ_1, is the given loop ω'. Finally, continuity of $I \times I \xrightarrow{\psi} X$ guarantees that "this family of loops ψ_s varies continuously with s." In other words, loops ω and ω' are homotopic if theie exists a "continuous deformation of ω to ω'" (figure 98). We shall call the corresponding continuous mapping $I \times I \xrightarrow{\psi} X$ a *homotopy* from ω *to* ω'.

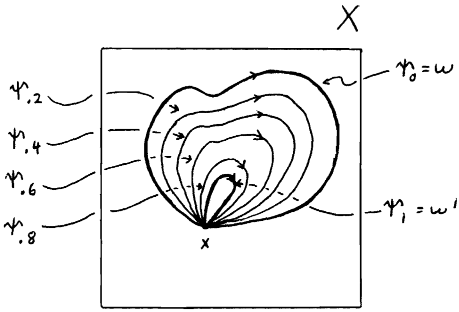

Figure 98

Example. Let X be the topological plane, and let x be the origin, that is, the point (0,0) of x. We claim that every loop based at x is homotopic to the zero loop. Indeed, let ω be a loop. Then, for each t in I, $\omega(t)$ is a point of X, whence there are real-valued functions $x_1(t)$ and $x_2(t)$ with $\omega(t) = (x_1(t), x_2(t))$ and with $x_1(0) = x_2(0) = x_1(1) = x_2(1) = 0$. Next, set $\psi_s(t) = ((1-s)x_1(t), (1-s)x_2(t))$, for s and t in I. This $I \times I \xrightarrow{\psi} X$ is clearly continuous. We have $\psi_s(0) = ((1-s)x_1(0), (1-s)x_2(0)) = (0,0)$, and similarly $\psi_s(1) = (0,0)$, whence condition i) above is satisfied. Furthermore, $\psi_0(t) = (x_1(t), x_2(t)) = \omega(t)$, while $\psi_1(t) = (0,0)$. Thus this ψ is a homotopy from the loop ω to the zero loop. (Geometrically, we "uniformly shrink the loop ω, radially, to the zero loop.")

Example. Let x be a point of topological space X, and let $I \xrightarrow{\omega} X$ be a loop based at x. Choose any continuous function $I \xrightarrow{f} I$ with $f(0) = 0$ and $f(1) = 1$. We claim, first, that $\omega' = \omega \circ f$ (so, for t in I, $\omega'(t)$ is the point $\omega(f(t))$ of X) is also a loop. [Proof: This $I \xrightarrow{\omega'} X$, as a composition of continuous mappings, is continuous. Furthermore, $\omega'(0) = \omega(f(0)) = \omega(0) = x$, and $\omega'(1) = x$.] We next claim that the loops ω and ω' are in fact homotopic. Let $I \times I \xrightarrow{\psi} X$ be given by $\psi_s(t) = \omega((1-s)t + sf(t))$ (where "+" denotes addition in the additive group of reals). First, note that ψ is continuous (for the mapping $I \times I \xrightarrow{\kappa} I$ which sends the point (s,t) of $I \times I$ to $\kappa(s,t) = (1-s)t + sf(t)$ is continuous, and $\psi = \omega \circ \kappa$). Furthermore, $\psi_s(0) = \omega((1-s)0 + sf(0)) = \omega(0) = x$, and $\psi_s(1) = x$, so each ψ_s is a loop. Finally, $\psi_0(t) = \omega((1-0)t + 0f(t)) = \omega(t)$, and $\psi_1(t) = \omega((1-1)t + 1f(t)) = \omega(f(t)) = \omega'(t)$. Thus ψ is a homotopy from the loop ω to the loop ω'. To see what this means geometrically, first note that $\omega[I] = \omega'[I]$, that is, our two loops "pass over precisely the same points of X." The only difference between the two loops, then, is "the way in which the points $\omega[I]$ of X are parameterized by t." We shall say that loop ω' is obtained from loop ω by *reparameterization*. This reparameterization is described, of course, by the function $I \xrightarrow{f} I$. Since $f(0) = 0$ and $f(1) = 1$, the two loops "start out the same and end up the same." It is just that "one loop can get ahead of, or fall behind, the other, with the passage of t, as they traverse the same actual path on X." The homotopy is a "continuous adjustment from the parameterization via ω to the parameterization via ω'." Thus we have shown that one loop obtained from another by reparameterization is homotopic to it. The following two points should be noted. First, we need not require that $I \xrightarrow{f} I$ be monotonic (i.e., ω' can "backtrack a little bit along the path of ω before going ahead"). Second, the homotopy argument above does not work if we impose on f the alternative boundary conditions $f(0) = 1$ and $f(1) = 0$.

Example. Let X be the punctured plane, and let x be the point given by (2,0). Let $I \xrightarrow{\omega} X$ be the loop given by $\omega(t) = (\cos(2\pi t), \sin(2\pi t))$. (The loop "goes around the hole," as shown in figure 99.) It is true, intuitively clear, but rather difficult actually to prove that this loop is not homotopic to the zero loop.

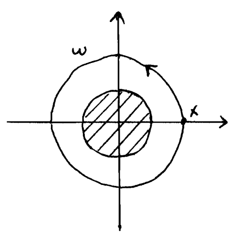

Figure 99

Let x be a point of topological space X. For ω and ω' loops based at x, write $\omega \approx \omega'$ if ω is homotopic to ω'. One expects intuitively that this relation "$\approx$" on $\Omega(X,x)$ will be an equivalence relation, and this in fact turns out to be the case. Proof: (1) Every loop is homotopic to itself, for, for ω a loop, $\psi_s(t) = \omega(t)$ is clearly a homotopy from ω to ω. (2) Let loops ω and ω' be homotopic, with homotopy ψ (so $\psi_0 = \omega$ and $\psi_1 = \omega'$). Then $\tilde{\psi}$ given by $\tilde{\psi}_s(t) = \psi_{(1-s)}(t)$ is clearly a homotopy from ω' to ω. Thus $\omega \approx \omega'$ implies

$\omega' \approx \omega$. (3) Let ω, ω', and ω'' be loops, and let $\hat{\psi}$ be a homotopy from ω to ω', and $\hat{\psi}'$ a homotopy from ω' to ω''. We obtain a homotopy from ω to ω''. Let $I \times I \xrightarrow{\psi} X$ be given by $\psi_s(t) = \{\hat{\psi}_{2s}(t)$ for $0 \leq s \leq 1/2$, and $\hat{\psi}'_{2s-1}(t)$ for $1/2 < s \leq 1\}$. Since $\hat{\psi}_1 = \hat{\psi}'_0$ $(= \omega')$, this mapping ψ is continuous. Clearly, each ψ_s is a loop. Finally, $\psi_0 = \hat{\psi}_0 = \omega$, and $\psi_1 = \hat{\psi}'_1 = \omega''$, so ψ is a homotopy from ω to ω'' (figure 100). (Intuitively, we first "follow the continuous deformation from ω to ω' and then follow the continuous deformation from ω' to ω'', obtaining a continuous deformation from ω to ω''. The $(2s)$ arise because we are only allowed to have s in the homotopy ψ go from 0 to 1, while s already goes from 0 to 1 in $\hat{\psi}$ and from 0 to 1 in $\hat{\psi}'$. In order to complete the whole deformation ψ by $s = 1$, therefore, we must first do the deformation $\hat{\psi}$ twice as quickly (in terms of s), followed by the deformation $\hat{\psi}'$ done twice as quickly.")

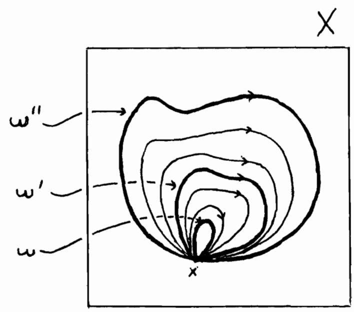

Figure 100

Thus, "is homotopic to" is an equivalence relation. We denote the set of equivalence classes by $\pi_1(X,x)$. Thus an element of $\pi_1(X,x)$ is a collection of loops all based at x, where any two loops in this collection are homotopic, and any loop homotopic to one of these is included in the collection. We call a loop in the collection a representative of the corresponding element of $\pi_1(X,x)$.

Example. Since, in the topological plane X, any loop (say, based at the origin x) is homotopic to the zero loop, $\pi_1(X,x)$ in this case has a single element. The zero loop is a representative of this element.

Example. It is intuitively clear that, for the punctured plane, two loops are homotopic if and only if they "go around the hole the same number of times and in the same direction." Thus one expects (and it is in fact true) that $\pi_1(X,x)$ in this case is isomorphic to the set of integers.

In some sense, we are now done. This $\pi_1(X,x)$ is a fairly simple set (e.g., usually simpler than the topological space X) which seems to characterize the "multiple-connectivity properties of X." The remarkable thing is that we can actually go one step further: we can endow the set $\pi_1(X,x)$ with the structure of a group.

Let x be a point of topological space X, and let ω and ω' be loops based at x. Consider the loop $I \xrightarrow{\omega''} X$ given by $\omega''(t) = \{\omega(2t)$ if $0 \leq t \leq 1/2$, and $\omega'(2t-1)$ for $1/2 < t \leq 1\}$. (This ω'' is continuous, since $\omega(1) = \omega'(0)$ ($= x$), while $\omega''(0) = \omega(0) = x$, and $\omega''(1) = \omega'(1) = x$.) We shall write this

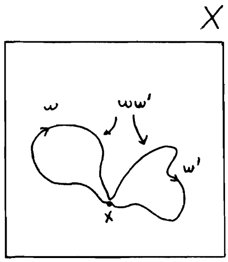

Figure 101

loop ω'' as $\omega\omega'$ (figure 101). (Note that this is not a composition of mappings. In fact, the composition of ω and ω' is not even defined.) Intuitively, we "first follow ω and then ω', going twice as fast all the time in order to finish by $t = 1$."

This "product structure" on the set $\Omega(X,x)$ has almost no nice properties: it is not associative, there is no identity, etc. Niceness comes only when we pass to $\pi_1(X,x)$. This "passage," in turn, is accomplished by the following fact: if loops ω_1 and ω_1' are homotopic, and ω_2 and ω_2' are homotopic, then $\omega_1\omega_2$ and $\omega_1'\omega_2'$ are homotopic. Indeed, if $\hat{\psi}$ is a homotopy from ω_1 to ω_1', and $\hat{\psi}'$ is a homotopy from ω_2 to ω_2', then ψ, given by $\psi_s(t) = \{\hat{\psi}_s(2t)$ for $0 \leq t \leq 1/2$, and $\hat{\psi}'_s(2t-1)$ for $1/2 < t \leq 1\}$, is a homotopy from $\omega_1\omega_2$ to $\omega_1'\omega_2'$. Thus we may define the product of two elements of $\pi_1(X,x)$ by taking the product (in $\Omega(X,x)$) of their representatives and finding the equivalence class (element of $\pi_1(X,x)$) in which it lies. Independence of choice of representatives is guaranteed by the result just proven. The whole point of this construction is the following:

THEOREM 43. *Let x be a point of topological space X. Then the set $\pi_1(X,x)$, with the product structure above, is a group.*

Proof:

i) Consider three elements of $\pi_1(X,x)$, with representatives ω, ω', and ω''. Then $(\omega\omega')\omega''$ is the loop given by $\{\omega(4t)$ for $0 \leq t \leq 1/4$, $\omega'(4t-1)$ for $1/4 < t \leq 1/2$, and $\omega''(2t-1)$ for $1/2 < t \leq 1\}$, while $\omega(\omega'\omega'')$ is the loop given by $\{\omega(2t)$ for $0 \leq t \leq 1/2$, $\omega'(4t-2)$ for $1/2 < t \leq 3/4$, and

$\omega''(4t-3)$ for $3/4 < t \leq 1\}$. The former is clearly a reparameterization of the latter, whence $(\omega\omega')\omega''$ and $\omega(\omega'\omega'')$ are homotopic. Thus the product in $\pi_1(X,x)$ is associative.

ii) The equivalence class containing the zero loop is an identity in $\pi_1(X,x)$. Denote by ω_0 the zero loop, and consider any element of $\pi_1(X,x)$, with representative ω. Then $\omega_0\omega$ is the loop given by $\{x$ if $0 \leq t \leq 1/2$, and $\omega(2t-1)$ for $1/2 < t \leq 1\}$. But this loop is a reparameterization of the loop ω itself, so $\omega_0\omega$ is homotopic to ω. Similarly, $\omega\omega_0$ is homotopic to ω.

iii) Consider an element of $\pi_1(X,x)$, with representative ω. Denote by ω' the loop with $\omega'(t) = \omega(1-t)$. Then $\omega\omega'$ is the loop given by $\{\omega(2t)$ for $0 \leq t \leq 1/2$, and $\omega(2-2t)$ for $1/2 < t \leq 1\}$. But this loop is homotopic to the zero loop, for example, by choosing $\psi_s(t) = \{\omega((1-s)2t)$ for $0 \leq t \leq 1/2$, and $\omega((1-s)(2-2t))$ for $1/2 < t \leq 1\}$. Similarly, $\omega'\omega$ is also homotopic to the zero loop. Thus inverses exist in $\pi_1(X,x)$ under this product structure. □

Like most proofs, this one is much simpler conceptually than in detail. For the first part, both $(\omega\omega')\omega''$ and $\omega(\omega'\omega'')$ are loops in which one "first follows along ω, then along ω', and then along ω''. The only difference is that the rates are different, because you have to follow each loop in a product twice as fast. The rate correction is a reparameterization, which does not count, up to homotopy." For the second part, the zero loop is "a representative of the identity, because $\omega_0\omega$ is the loop in which you sit at x for the first half of your allotted t-interval, and then cover ω twice as fast, which differs from ω itself only up to a reparameterization." For the third part, "ω' is just the loop for which you follow ω, but in the other direction. Thus $\omega\omega'$ is the loop in which you first follow along ω (twice as fast), and then turn around and follow ω in the opposite direction. But this is homotopic to the zero loop, the homotopy being that for which one only follows ω for a little while (less and less as s gets nearer to one) before turning back to retrace ω to x," as in figure 102.

The group $\pi_1(X,x)$ is called the *first homotopy group* of X (based at x). ("First" because there is a second, third, etc., although we shall not discuss them.) A topological space X is said to be *simply connected* if, for every point x of X, $\pi_1(X,x)$ is the group with a single element (the identity). (Thus X is simply connected if and only if any two loops based at the same point are homotopic.)

Example. The first homotopy group of the topological plane has only a single element, so the topological plane is simply connected. The first homotopy group of the punctured plane is the additive group of integers.

Example. Every discrete topological space, and every indiscrete one, is simply connected.

Example. Let X be the topological space consisting of the subspace of the topological plane that results from the removal of the points (0,1) and

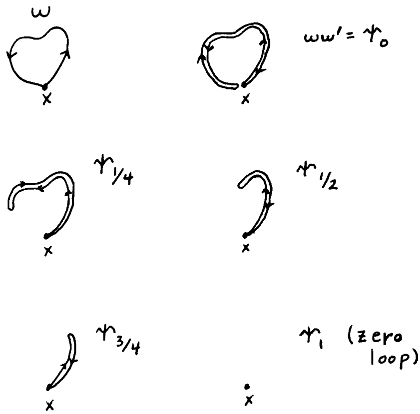

Figure 102

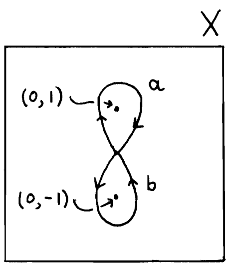

Figure 103

(0,–1), as in figure 103. Let x be the origin. Denote by a and b the two loops shown, and by a^{-1} and b^{-1} these loops, traversed in the opposite direction. It is clear intuitively that any loop based at x is homotopic to some product of the form $aaba^{-1}bab^{-1}b^{-1}$. Thus one expects that the first homotopy group of this space will be the free group on a set with two elements. This is in fact true. (Note, in particular, that homotopy groups need not be abelian.)

There is one final issue we wish to discuss: what is the dependence of $\pi_1(X,x)$ on the base point x? Let x and x' be points of the topological space X. A *curve* from x to x' is a continuous mapping $I \xrightarrow{\gamma} X$ with $\gamma(0) = x$ and $\gamma(1) = x'$ (so a loop is a curve from x to x). Consider first the case when there exists no curve from x to x'. Then there is no necessary relation between $\pi_1(X,x)$ and $\pi_1(X,x')$; for example, X could be the direct sum of two topological spaces, with x in one piece and x' in the other. The case when there is a curve γ from x to x' is more interesting. We define in this case a mapping κ_γ from $\Omega(X,x')$ to $\Omega(X,x)$ as follows: κ sends loop ω based at x' to the loop (based at x) given by $\{\gamma(3t)$ for $0 \le t \le 1/3$, $\omega(3t-1)$ for $1/3 < t \le 2/3$, and $\gamma(3-3t)$ for $2/3 < t \le 1\}$. (That is, "we follow γ from x to x',

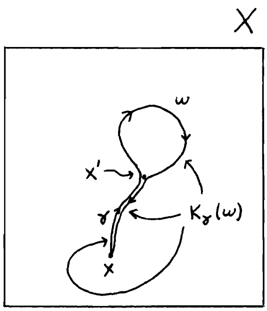

Figure 104

then the loop ω, then retrace γ back from x' to x, to obtain a loop at x.") (See figure 104.) We now wish to make two claims about this mapping κ_γ. First, if ω and ω' are loops based at x' and if these loops are homotopic, then so are the loops $\kappa_\gamma(\omega)$ and $\kappa_\gamma(\omega')$ (both based at x). [Proof: If ψ is a homotopy from ω to ω', then $\tilde{\psi}$, given by $\tilde{\psi}_s(t) = \{\gamma(3t)$ for $0 \le t \le 1/3$, $\psi_s(3t-1)$ for $1/3 < t \le 2/3$, and $\gamma(3-3t)$ for $2/3 < t \le 1\}$, is a homotopy from $\kappa_\gamma(\omega)$ to $\kappa_\gamma(\omega')$.] Second, if ω and ω' are loops based at x', then $\kappa_\gamma(\omega\omega')$ is homotopic to $\kappa_\gamma(\omega)\kappa_\gamma(\omega)$. [Outline of proof: For the loop $\kappa_\gamma(\omega\omega')$ at x, you first follow γ from x to x', then go around ω, then go around ω', and then retrace γ back to x. For the loop $\kappa_\gamma(\omega)\kappa_\gamma(\omega')$ at x, you first follow γ from x to x', then go around ω, then go back to x along γ, then return to x' along γ, then go around ω', and finally again return to x along γ. To show that these loops are homotopic, first "homotopy away" from the second loop the extra visit to x between going around ω and going around ω'. Then note that the resulting loop is a reparameterization of the first loop.]

The first observation above is just the statement that $\Omega(X,x') \xrightarrow{\kappa_\gamma} \Omega(X,x)$ sends all the points of an entire equivalence class of $\Omega(X,x')$ into the same equivalence class in $\Omega(X,x)$. Thus we obtain a corresponding mapping, $\pi_1(X,x') \xrightarrow{\varphi_\gamma} \pi_1(X,x)$ between the sets of equivalence classes (i.e., given an element of $\pi_1(X,x')$, of which ω is a representative, φ_γ sends that element to the element of $\pi_1(X,x)$ of which $\kappa_\gamma(\omega)$ is a representative). The second observation above is now precisely the statement that φ_γ is in fact a homomorphism of groups. To summarize, we have shown so far that a curve from x to x' leads to a homomorphism φ_γ from $\pi_1(X,x')$ to $\pi_1(X,x)$.

We next claim that this homomorphism φ_γ is in fact an isomorphism of groups. Indeed, denote by γ' the curve from x' to x given by $\gamma'(t) = \gamma(1 - t)$. Then, for ω any loop based at x', $\kappa_\gamma(\omega)$ is a loop based at x, whence $\kappa_\gamma(\kappa_\gamma(\omega))$ is again a loop based at x'. In fact, this loop $\kappa_{\gamma'}(\kappa_\gamma(\omega))$ is homotopic to ω. [Outline of proof: For the loop $\kappa_{\gamma'}(\kappa_\gamma(\omega))$ at x', you first go from x' to x via γ, then go back to x' via γ, then go around ω, then go again to x and back to x' along γ. To see that this loop at x' is homotopic to ω, just "homotopy away" the two visits to x.] Expressed in terms of our homotopy groups, the statement that $\kappa_{\gamma'}(\kappa_\gamma(\omega))$ is homotopic to ω is just the statement that $\varphi_{\gamma'} \circ \varphi_\gamma$ is the identity homomorphism on $\pi_1(X,x')$. (Note that this makes sense, for $\pi_1(X,x') \xrightarrow{\varphi_\gamma} \pi_1(X,x)$ and $\pi_1(X,x) \xrightarrow{\varphi_{\gamma'}} \pi_1(X,x')$, so $\pi_1(X,x') \xrightarrow{\varphi_{\gamma'} \circ \varphi_\gamma} \pi_1(X,x')$.) Similarly, $\varphi_\gamma \circ \varphi_{\gamma'}$ is the identity homomorphism on $\pi_1(X,x)$. Thus φ_γ is an isomorphism of groups.

To summarize, there need be no relation between $\pi_1(X,x)$ and $\pi_1(X,x')$ when there exists no curve in X from x to x'. When such a curve does exist, these groups are necessarily isomorphic. Thus, for a topological space X having the property that any two of its points can be joined by a curve (e.g., the topological plane), we may speak of the first homotopy group of X (since the base point makes no difference).

Exercise 227. Let X be a topological space. For x and x' points of X, write $x \approx x'$ if there exists a curve from x to x'. Prove that this is an equivalence relation. (The equivalence classes are called curve-connected components of X.) Prove that $\pi_1(X,x)$ depends only on the curve-connected component of X containing x in the following sense: if there exists an isomorphism (of topological spaces) between the curve-connected component of X containing x and the curve-connected component of Y containing y, then there exists an isomorphism (of groups) between $\pi_1(X,x)$ and $\pi_1(Y,y)$.

Exercise 228. The category of pointed topological spaces is that whose objects are pairs (X,x), where X is a topological space and x is a point of X, and whose morphisms (from (X,x) to (Y,y)) consist of a continuous mapping $X \xrightarrow{\varphi} Y$ of topological spaces with $\varphi(x) = y$. Show that "take the first homotopy group" is a covariant functor from the category of pointed topological spaces to the category of groups. Why not just from the category of topological spaces?

Exercise 229. Analyze the dependence of the isomorphism φ_γ above on γ. In particular, let γ and γ' be two curves from x to x', and let $\mathbf{q}$ be the element of $\pi_1(X,x')$ having as representative $\{(2 - 2t)$ for $0 \leq t \leq 1/2$, and $\gamma'(2t - 1)$

for $1/2 < t \leq 1\}$. Prove that, for p any element of $\pi_1(X,x')$, $\varphi_\gamma^{-1} \circ \varphi_\gamma(p) = \mathbf{q}^{-1}p\mathbf{q}$.

Exercise 230. The two-sphere is the subspace of $\mathbf{R} \times \mathbf{R} \times \mathbf{R}$ given by $(x_1)^2 + (x_2)^2 + (x_3)^2 = 1$. Prove that the two-sphere is simply connected.

Exercise 231. A topological space X is said to be contractible (to point x_0 of X) if there is a continuous mapping $I \times X \xrightarrow{\psi} X$ with $\psi(0,x) = x$ and $\psi(1,x) = x_0$ for all x. Show that the topological plane, for example, is contractible. Prove that every contractible topological space is connected and simply connected. Find an example to show that the converse is false.

Exercise 232. A topological space is said to be locally simply connected if each of its points has a neighborhood which, as a subspace, is simply connected. Find an example of a topological space which is not locally simply connected.

Exercise 233. Let X be a topological space, x a point of X, Y a topological space, and y a point of Y. Prove that $\pi_1(X \times Y,(x,y))$ is isomorphic to $\pi_1(X,x) \times \pi_1(Y,y)$ (where $\times$ is direct product in the appropriate category).

Exercise 234. Is any given group some $\pi_1(X,x)$?

Exercise 235. Is $\Omega(X,x') \xrightarrow{\kappa_\gamma} \Omega(X,x)$ an isomorphism of sets?

Exercise 236. Find two topological spaces that are not isomorphic but whose first homotopy groups are.

Exercise 237. Assigning to the loop spaces the compact-open topology, is $\Omega(X,x') \xrightarrow{\kappa_\gamma} \Omega(X,x)$ continuous?

35

Homology

We now consider the second approach to the "multiple-connectedness structure" of a topological space.

Fix a non-negative integer n. We denote by K_n the subspace of $\mathbf{R} \times \mathbf{R} \times \cdots \times \mathbf{R}$ (the direct product of the real line with itself $(n + 1)$ times) consisting of $(a_0, a_1, \ldots, a_n)$ with $a_0 \geq 0$, $a_1 \geq 0, \ldots, a_n \geq 0$, and with $a_0 + a_1 + \cdots + a_n = 1$. The first few are illustrated in figure 105. Thus K_0 is the subspace of $\mathbf{R}$, the real line, consisting of the single point, (1). K_1 is the subspace of $\mathbf{R} \times \mathbf{R}$, the topological plane, consisting of pairs of real numbers, (a_0, a_1), with each non-negative and with $a_0 + a_1 = 1$. Thus K_1 is a "line segment (with endpoints)" in the topological plane. Similarly, K_2 is the subspace of $\mathbf{R} \times \mathbf{R} \times \mathbf{R}$ consisting of (a_0, a_1, a_2), with each non-negative and with $a_0 + a_1 + a_2 = 1$. That is, K_2 is "a triangle in Euclidean 3-space." K_3 is a "tetrahedron," and so on for higher dimensions. Note that each of K_0, K_1, $K_2, \cdots$ is an explicit, given topological space, fixed once and for all.

The idea is to use K_0 (a "point"), K_1 (a "line"), K_2 (a "triangle"), etc., to construct various geometrical figures within a topological space X which can then, in turn, be used to describe the structure of that space. The first step is to "insert the K_n into the space X." To this end, we define an *n-simplex* (in X) as a continuous mapping $K_n \xrightarrow{\sigma} X$. Thus a 0-simplex in X is a mapping from the "one-point space" K_0 to X (i.e., it is essentially a point of X). A 2-simplex is a continuous mapping from the "triangle" K_2 to X. We can think of a 2-simplex in X as a sort of "curvilinear triangle which sits in X" (figure 106). (It should be emphasized, however, that an n-simplex consists of a continuous mapping, $K_n \xrightarrow{\sigma} X$, and not just the image, $\sigma[K_n]$, of that mapping. Thus, even if $K_2 \xrightarrow{\sigma} X$ takes all of K_2 to a single point of X, this is nonetheless a 2-simplex in X.) Note, incidentally, that a loop in X is also a 1-simplex in X.

We regard the n-simplices as the "building blocks" which are to be used to construct more complicated figures in X. Thus we need some way to "string together a lot of n-simplices." One could imagine, for example, proceeding by the "prefabrication method": one could first "hook together" a lot of K_n (e.g., in some Euclidean space) and then map the entire thing, all at once (continuously), into X. It turns out, however, to be more convenient to use "on-site construction." Thus what we next wish to do is "consider a lot

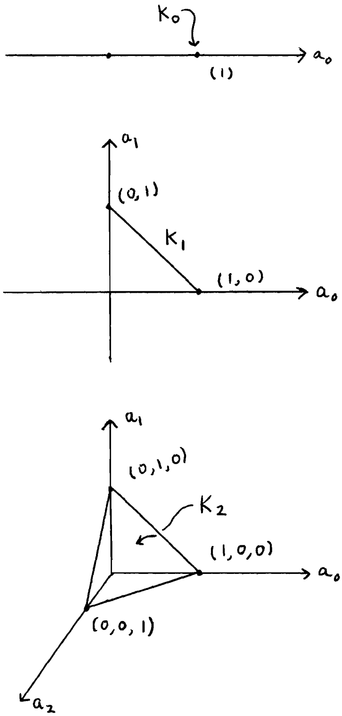

Figure 105

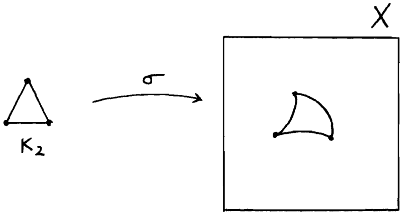

Figure 106

of n-simplices in X, all together."

Fix a topological space X and a non-negative integer n. Denote by $C_n(X)$ the free abelian group on the set of all n-simplices in X. Thus a typical element of $C_n(X)$ is "a formal linear combination of n-simplices in X, where the coefficients are (positive or negative) integers." For example, $7\sigma - 3\sigma' + 2\sigma''$ is a typical element of $C_2(X)$, where σ, σ', and σ'' are 2-simplices in X.

Consider $C_0(X)$. Since a 0-simplex in X is essentially a point of X, an element of $C_0(X)$ is "a finite list of points of X, each assigned a multiplicity (some integer)." Similarly, one can picture an element of, for example, $C_2(X)$ by drawing in X the "curvilinear triangles" (which represent the 2-simplices which appear in this element of $C_2(X)$) and assigning to each a "multiplicity integer." (One can think of, e.g., "a curvilinear triangle of multiplicity three" as three identical curvilinear triangles all on top of each other. We allow negative multiplicities to make it easier to "remove curvilinear triangles" when necessary.) Note that each $C_n(X)$ is always an abelian group and that, except for rather trivial X, $C_n(X)$ is enormous in size.

Example. Let X be discrete. Then (since K_n is connected, since the image of a connected topological space under a continuous mapping is connected, since no connected subset of X has more than one point) every n-simplex, $K_n \xrightarrow{\sigma} X$, takes all of K_n to a single point of X. Thus, for example, if X contains exactly five points, then each $C_n(X)$ is the direct sum of the additive group of integers with itself five times.

We think of the elements of $C_n(X)$ as representing the desired "geometrical figures in X." Two things have yet to be done. First, since an n-simplex is simply a continuous mapping (into X), we have not yet incorporated a notion of "the simplices in an element of $C_n(X)$ all joining together properly to make reasonable geometrical figures." What we are really interested in is elements of $C_n(X)$ which represent "n-dimensional surfaces in X." Second, we have to figure out some way to cut down the size of $C_n(X)$ to obtain a simple description of X. In the previous chapter, "homotopic" served this purpose: we need an analogous notion here. It turns out that a single concept—the notion of a "boundary"—will serve both of these needs.

Fix a positive integer n, and consider K_n. We now introduce $(n + 1)$ mappings, $f_0, f_1, \cdots, f_n$, from K_{n-1} to K_n as follows: $K_{n-1} \xrightarrow{f_0} K_n$ is the mapping which sends $(a_0, \ldots, a_{n-1})$ (a point of K_{n-1}) to the point $(0, a_0, \ldots, a_{n-1})$ of K_n; $K_{n-1} \xrightarrow{f_1} K_n$ sends the point $(a_0, \ldots, a_{n-1})$ of K_{n-1} to the point $(a_0, 0, a_1, \ldots, a_{n-1})$ of K_n; $\ldots$; $K_{n-1} \xrightarrow{f_n} K_n$ sends the point $(a_0, \ldots, a_{n-1})$ of K_{n-1} to the point $(a_0, \ldots, a_{n-1}, 0)$ of K_n. (That is, given a point, $(a_0, \ldots, a_{n-1})$ of K_{n-1}, f_i sends this point to the point of K_n which results from "adding an extra zero, in the ith place.") Note that these mappings are well defined, for, if $(a_0, \ldots, a_{n-1})$ is n non-negative numbers with sum 1, then $(a_0, \ldots, 0, \ldots, a_{n-1})$ is $(n + 1)$ non-negative numbers with sum 1. Note also that each of the mappings $f_0, f_1, \ldots, f_n$ is continuous. There is a simple geometrical interpretation for these mappings, as illustrated in figure 107. Consider K_2, the "triangle." Then $K_1 \xrightarrow{f_0} K_2$ sends (a_0, a_1) to $(0, a_0, a_1)$, f_1 sends (a_0, a_1) to $(a_0, 0, a_1)$, and f_2 sends (a_0, a_1) to $a_0, a_1, 0)$. But the points of K_2 of the

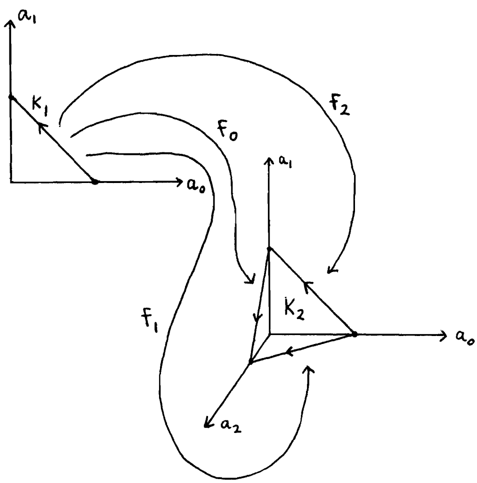

Figure 107

form $(0, a_0, a_1)$ form one "edge" of the triangle and, similarly, $(a_0, 0, a_1)$ and $(a_0, a_1, 0)$ represent the other two edges. Thus the three mappings $K_1 \xrightarrow{f_0} K_2$, $K_1 \xrightarrow{f_1} K_2$, and $K_1 \xrightarrow{f_2} K_2$ describe the "three edges of the triangle K_2." Similarly, $K_0 \xrightarrow{f_0} K_1$ and $K_0 \xrightarrow{f_1} K_1$ describe the "two endpoints of the line segment K_1"; the four mappings from K_2 to K_3 describe the "four faces of the tetrahedron K_3"; etc. (Note, incidentally, that, since, for any topological space Y, $\text{Bndy}(Y) = \emptyset$, the images of the f do not really form $\text{Bndy}(K_n)$ in this strict sense.)

We now know how to describe the "boundary of a K_n." We now wish to "carry this idea over to X," that is, to define the boundary of an n-simplex. Let $K_n \xrightarrow{\sigma} X$ be a fixed n-simplex in X. Then, since $K_{n-1} \xrightarrow{f_0} K_n$ and $K_n \xrightarrow{\sigma} X$ are continuous, so is $K_{n-1} \xrightarrow{\sigma \circ f_0} X$. Thus $\sigma \circ f_0$ is an $(n-1)$-simplex in X, and similarly for $\sigma \circ f_1, \ldots, \sigma \circ f_n$. That is, from a single n-simplex σ, we obtain a total of $(n+1)$ $(n-1)$-simplices. For example, the three 1-simplices that arise in this way (figure 108) from a 2-simplex σ represent the three "curved edges of the curvilinear triangle represented by σ." But the whole idea of going to the free abelian groups, the C_n, was to be able to talk about a number of simplices at the same time. Thus, for σ an n-simplex, we define $\partial(\sigma)$ as the element of $C_{n-1}(X)$ given by $\partial(\sigma) = \sigma \circ f_0 - \sigma \circ f_1 + \sigma \circ f_2 - \sigma \circ f_3 + \cdots \pm \sigma \circ f_n$. (For σ a 0-simplex, we set $\partial(\sigma) = 0$.) Thus, for σ a 2-simplex, $\partial(\sigma)$ is the element of $C_1(X)$ obtained by "adding (in $C_1(X)$) the

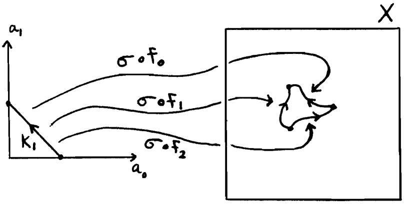

Figure 108

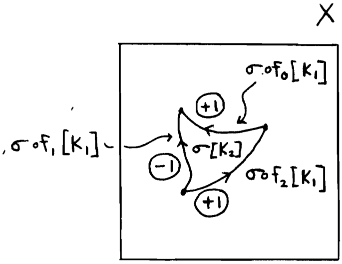

Figure 109

edges of the curvilinear triangle σ." The reason for the alternating signs of the multiplicities in the formula above can be seen in figure 109. A 2-simplex σ, together with $\partial(\sigma)$, is shown. Note that the second "edge of the triangle" gets "mapped in backward" (i.e., the three arrows do not all go around the triangle in the same direction). This is corrected by the negative multiplicity. To take a second example, let σ be a 1-simplex in X as in figure 110. Then $\partial(\sigma)$ is the difference of two 0-simplices (the "endpoints of the curve σ"). The one with plus multiplicity is the one that rises from the point (0,1) of K_1; negative multiplicity, from the point (1,0). In this case, the multiplicities reflect the "starting point" and "ending point" of the curve.

Thus, for σ an n-simplex, $\partial(\sigma)$ is a linear combination (with integer coefficients) of $(n - 1)$-simplices, that is, an element of $C_{n-1}(X)$. We next wish to extend this action of ∂ from n-simplices to all linear combinations of n-simplices, that is, to define an action of ∂ on $C_n(X)$. Given an element of $C_n(X)$ (e.g., $7\sigma - 3\sigma' + 2\sigma''$, where σ, σ', and σ'' are n-simplices), we define ∂ of that element, for example, by the formula $\partial(7\sigma - 3\sigma' + 2\sigma'') = 7\partial(\sigma) - 3\partial(\sigma') + 2\partial(\sigma'')$. (Note that the sums on the left are in $C_n(X)$; those on the right are in $C_{n-1}(X)$.) (More precisely, the action of ∂ on $C_n(X)$ is defined by requiring that, on a single n-simplex, it be the ∂ above and that this ∂ satisfy $\partial(\alpha + \beta) = \partial(\alpha) + \partial(\beta)$ for any α and β in $C_n(X)$.) Thus, for any element α of $C_n(X)$, $\partial(\alpha)$ is an element of $C_{n-1}(X)$, so ∂ is a mapping, $C_n(X) \xrightarrow{\partial}$

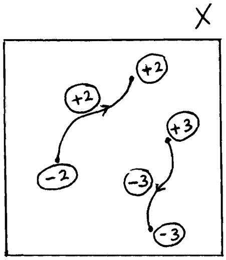

Figure 110

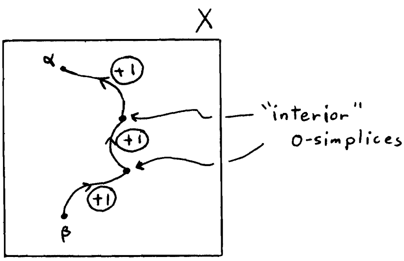

Figure 111

$C_{n-1}(X)$. It is obvious from the definition that this mapping ∂ is in fact a homomorphism of (abelian) groups. For α in $C_n(X)$, the element $\partial(\alpha)$ of $C_{n-1}(X)$ is called the ***boundary*** of α. Thus the boundary of an n-simplex (or of any linear combination, with integer coefficients, of n-simplices) is a linear combination of $(n-1)$-simplices (as one would expect intuitively).

A few intuitive examples will make the setup clearer. Consider the element of $C_1(X)$ illustrated in figure 111. (Our element of $C_1(X)$ is the linear combination $2\sigma - 3\sigma'$ of 1-simplices.) The boundary of this element is a certain linear combination of four 0-simplices with multiplicities -2, +2, +3, and -3. The whole point of the multiplicities, it can now be seen, is that they "allow interior boundaries (i.e., common boundaries of two simplices that fit together)" to cancel. Thus the boundary of the element of $C_1(X)$ pictured at the right is just $\alpha - \beta$, where α and β are the 0-simplices shown. The two "interior 0-simplices of the curve" appear twice in the boundary, once with multiplicity +1 and once with -1, and hence cancel (in the group $C_0(X)$). Similarly, the boundary of the element of $C_2(X)$ shown in figure 112 is a combination of four 1-simplices, the "interior simplex appearing twice in the boundary with opposite multiplicities, and hence canceling out." A similar phenomenon occurs in higher dimensions. It is precisely this cancellation

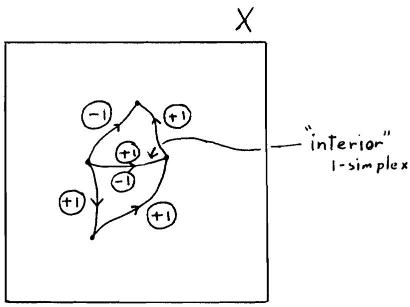

Figure 112

property which makes what we define as the boundary have the correct geometrical properties.

To summarize, we first introduce a family, K_0, K_1, $\cdots$ of "figures in Euclidean space." Continuous mappings from these "figures" to topological space X define 0-simplices, 1-simplices, . . . , in X. By taking "formal linear combinations, with integer coefficients" of n-simplices, we obtain an abelian group $C_n(X)$. Next, we introduce, in a natural way, the notion of the boundary of an n-simplex (where this boundary is an element of $C_{n-1}(X)$) and then extend this notion, by linearity, to $C_n(X)$. Thus we obtain, finally,

$$\cdots \xrightarrow{\partial} C_n(X) \xrightarrow{\partial} C_{n-1}(X) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_2(X) \xrightarrow{\partial} C_1(X) \xrightarrow{\partial} C_0(X) \xrightarrow{\partial} 0 \quad .$$

It is only these groups and these mappings which will now be needed to obtain the homological description of topological spaces. (The fact that the same symbol, ∂, denotes a number of homomorphisms of groups does not lead to confusion.)

We next obtain an important property of the homomorphisms ∂. We have $C_n(X) \xrightarrow{\partial} C_{n-1}(X)$, and hence $C_n(X) \xrightarrow{\partial\circ\partial} C_{n-2}(X)$. We now claim that this homomorphism $\partial \circ \partial$ is in fact the zero homomorphism, that is, $\partial \circ \partial(\alpha) = 0$ for every α in $C_n(X)$. (That is, "the boundary of the boundary is zero.") It clearly suffices to show that $\partial \circ \partial$ gives zero on any n-simplex. Fix $n \geq 2$, and let i and j be integers with $0 \leq i < j \leq n$. Both $f_i \circ f_{j-1}$ and $f_j \circ f_i$ are mappings from K_{n-2} to K_n. We claim that they are in fact the same mapping: each sends the point $(a_0, \ldots, a_{n-2})$ of K_{n-2} to the point $(a_0, \ldots, 0, \ldots, 0, \ldots, a_{n-2})$ of K_n, where the zeros occur in the ith and jth places. (Indeed, f_i sends $(a_0, \ldots, a_{n-2})$ to $(a_0, \ldots, 0, \ldots, a_{n-2})$, with 0 in the ith place, while f_j sends this point of K_{n-1} to $(a_0, \ldots, 0, \ldots, 0, \ldots, a_{n-2})$, with zeros in the ith and jth places. Similarly, f_{j-1} puts in a zero in the $(j-1)$st place of $(a_0, \ldots, a_{n-2})$, while f_i puts in a zero in the ith place. But, since $i < j$, this last putting in a zero in the ith place shifts the original zero (which was in the $(j-1)$st place) to the jth place.) Next, let σ be an n-simplex in X. Then $\partial \circ \partial(\sigma) = \sum_{j=0}^{n} \sum_{i=0}^{n-1} (-1)^{i+j}\sigma \circ f_i \circ f_j$. But since, for $0 \leq i < j \leq n$, $f_i \circ f_{j-1} = f_j \circ f_i$, all the terms in this last sum cancel. Thus the boundary of the boundary of any element of $C_n(X)$ vanishes. This property is much easier to see explicitly for low dimensions, for example, by finding the boundary of the element of $C_1(X)$ in figure 109.

We now wish to restrict consideration to a certain class of "particularly interesting geometrical figures in X." An element α of $C_n(X)$ will be called an *n-cycle* if $\partial(\alpha) = 0$, that is, if "all the terms in the linear combination of $(n-1)$-simplices which constitute the boundary of α cancel." Intuitively, one thinks of an element of $C_n(X)$ as "little pieces of n-surfaces sitting in X" and of an n-cycle (a particular kind of element of $C_n(X)$) as a "closed n-surface in X, one that connects up with itself along all edges." For example, figure 113

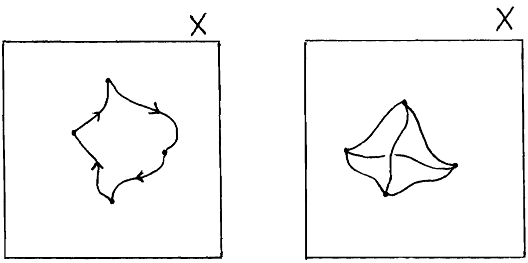

Figure 113

represents a typical 1-cycle and a typical 2-cycle in X. Note, in particular, that every loop in X (regarded as an element of $C_1(X)$) is a 1-cycle. Since $\partial = 0$ on $C_0(X)$, every element of $C_0(X)$ is a 0-cycle. Next, note that, since $C_n(X) \xrightarrow{\partial} C_{n-1}(X)$ is a homomorphism, we have $\partial(\alpha + \beta) = \partial(\alpha) + \partial(\beta)$ and $\partial(-\alpha) = -\partial(\alpha)$ for any elements α and β of $C_n(X)$. In particular, the sum of two n-cycles is necessarily an n-cycle and the (additive) inverse (in the group $C_n(X)$) of an n-cycle is an n-cycle. (Thus a 1-cycle could consist, for example, of "several separate closed curves.") Thus the collection, $Z_n(X)$, of all n-cycles forms a subgroup of the group $C_n(X)$. What we have said in this paragraph, with motivation removed, is: let $Z_n(X)$ denote the kernel of $C_n(X) \xrightarrow{\partial} C_{n-1}(X)$, and call its elements cycles.

We have now, finally, obtained the "geometrical figures in X" that we wish to study, namely the n-cycles. Only one further difficulty remains: there are, for any reasonable X, too many n-cycles to make $Z_n(X)$ provide a reasonable description of X. One can think of the n-cycles in homology as playing a role somewhat analogous to that of loops in homotopy. What we now want is something in homology analogous to "is homotopic to" in homotopy. In order to get an idea of what one should try, we ask, "How does one know that there tend to be so many n-cycles?" For one thing, if we take any element γ of $C_{n+1}(X)$, then $\partial(\gamma)$ is necessarily an n-cycle (for its boundary, $\partial(\partial(\gamma))$, is zero since $\partial \circ \partial = 0$). This observation suggests the following definition. Two n-cycles, α and β, will be called *homologous* if there is an element γ of $C_{n+1}(X)$ with $\alpha - \beta = \partial(\gamma)$. (Note that there is no point in requiring that γ be an $(n+1)$-cycle, for then we would have $\partial(\gamma) = 0$, and hence $\alpha = \beta$.)

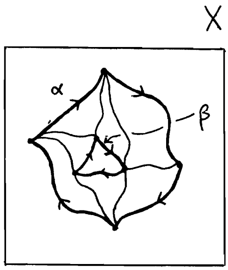

Figure 114

This notion "is homologous to" has very much the intuitive flavor of "is homotopic to." For example, the two 1-cycles represented in figure 114 are homologous with each other. The "triangulation of the region between α and β" suggests how one would construct an element γ of $C_2(X)$ with $\alpha - \beta = \partial(\gamma)$. Consider next the situation illustrated in figure 115. The 1-cycle

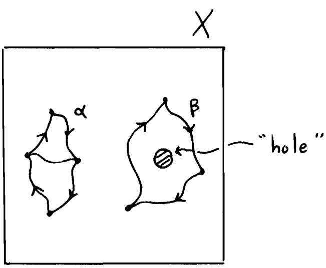

Figure 115

denoted α is homologous to the zero 1-cycle (i.e., $\alpha = \partial(\gamma)$ for some γ), for example, by the "triangulation" suggested by the figure. (There is little point to writing everything out explicitly. How one would actually write out γ in detail is perhaps clearer from the figure than it would be from explicit formulae.) On the other hand, the 1-cycle β in this figure is not homologous to the zero 1-cycle, for "the hole in X prevents one from finding any γ in $C_2(X)$ with

$\beta = \partial(\gamma)$." The difference (at least formally—a difference we shall refine in the next chapter) between homologous 1-cycles and homotopic loops is that the latter require the existence of a "continuous deformation" from one loop to the other, while the former require the existence of "some 2-surface in X whose boundary is the difference of the 1-cycles." This intuitive state of affairs continues into higher dimensions. Thus two 2-cycles are homologous, intuitively, iff "there is a 3-surface in X of which the two 2-cycles form a boundary." It is easy, for 0-cycles, to describe the situation more precisely.

Example. Let X be a topological space, and let X have the property that between any two points of X there passes a curve. Let σ and σ' be 0-simplices (and hence 0-cycles) on X, and let x and x' be the corresponding image-points of these mappings. Then, evidently, $\sigma - \sigma' = \partial(\gamma)$, where γ is the element of $C_1(X)$ which is the curve from x to x'. Thus any two 0-simplices are homologous. Now consider two 0-cycles, for example, $2\sigma + \sigma'$ and $-\sigma'' + 3\sigma'''$. Since any two 0-simplices are homologous, it is clear that these two will be homologous if and only if the "sum of the multiplicities" (i.e., $2 + 1 = 3$ and $-1 + 3 = 2$) is the same for the two (so these two 0-cycles are not homologous).

Thus we now have the notion of a "boundaryless surface" (an n-cycle), the set $Z_n(X)$ of these n-cycles, and the relation "is homologous to" on these n-cycles. We next claim that this is in fact an equivalence relation on $Z_n(X)$. [Proof: (1) Each n-cycle α is homologous to itself, for $\alpha - \alpha = \partial(0)$. (2) If α is homologous to β (so $\alpha - \beta = \partial(\mu)$ for some μ in $C_{n+1}(X)$), then β is homologous to α for $\beta - \alpha = \partial(-\mu)$. (3) If α is homologous to β, and β to γ (so $\alpha - \beta = \partial(\mu)$ and $\beta - \gamma = \partial(\nu)$), then α is homologous to γ for $\alpha - \gamma = \partial(\mu + \nu)$.] We denote the collection of equivalence classes by $H_n(X)$. Thus an element of $H_n(X)$ is "an n-cycle (i.e., a boundaryless thing) modulo the boundary of something of one higher dimension." (This $H_n(X)$ is somewhat analogous to the set of equivalence classes of loops under "is homotopic to.") Next, note that, if α is homologous to α', and β is homologous to β', then $\alpha + \beta$ is homologous to $\alpha' + \beta'$. Thus we can define addition in $H_n(X)$: given two elements of $H_n(X)$ (i.e., two equivalence classes in the set $Z_n(X)$ of n-cycles), pick an n-cycle from each equivalence class, add them, and see what equivalence class (element of $H_n(X)$) the sum lies in. By the observation above, this rule is independent of the original particular choices of n-cycles. Thus $H_n(X)$ has the structure of an abelian group and is called the nth ***homology group*** of the topological space X. (In a more algebraic setting, the above could be stated as follows. Since $\partial \circ \partial = 0$, the image of the first ∂ in $C_{n+1}(X) \xrightarrow{\partial} C_n(X) \xrightarrow{\partial} C_{n-1}(X)$ is a subgroup of the kernel, $Z_n(X)$, of the second ∂. The quotient group is $H_n(X)$.)

To summarize, we begin with n-simplices in X and then "string them together, with integral multiplicities" to obtain the abelian group $C_n(X)$.

Next, one constructs the "boundary homomorphism," $C_n(X) \xrightarrow{\partial} C_{n-1}(X)$. One next obtains two subgroups of $C_n(X)$: $Z_n(X)$, the subgroup of n-cycles, consists of elements of $C_n(X)$ with zero boundary; and (say) $B_n(X)$, the subgroup of $C_n(X)$, consists of elements which are boundaries of elements of $C_{n+1}(X)$. Since $\partial \circ \partial = 0$, $B_n(X)$ is a normal subgroup of $Z_n(X)$. Set $H_n(X) = Z_n(X)/B_n(X)$, the nth homology group.

Example. Let X be a topological space any two points of which can be joined by a curve. Then, as we saw above, any two 0-simplices in X are homologous. Fix a 0-simplex $\underline{\sigma}$ in X. Then any 0-cycle is homologous to some multiple of $\underline{\sigma}$, for example, $2\sigma + 9\sigma' - \sigma''$ is homologous to $10\underline{\sigma}$. Thus $H_0(X)$ in this example is the additive group of integers.

Example. Let X be the topological space with a single point. Then each $C_n(X)$ is the additive group of integers (since there is only one n-simplex for each n, so only the multiplicity counts). From the definition of ∂, it is immediate that $C_{n+1}(X) \xrightarrow{\partial} C_n(X)$ is the zero mapping for n even, and the identity mapping for n odd. Hence $H_0(X) = Z$, with all other $H_n(X)$ the zero group.

Example. The homology groups of the real line are the same as those of the previous example. We sketch a proof. It suffices to show that every n-cycle for $n \geq 1$ is homologous to the zero n-cycle. Introduce, as in the last chapter, a "uniform contraction" of the real line to its origin, and use this to construct, for any n-cycle α, a μ with $\alpha = \partial(\mu)$. A similar argument shows that the homology groups of the topological plane are the same as those of the real line.

Exercise 238. Prove that the homology groups of the punctured plane are the following: $H_0(X)$ and $H_1(X)$ are the additive group of integers, all others zero.

Exercise 239. Let X be a topological space with seven path-connected components. Prove that $H_0(X)$ is the direct sum of the additive group of integers with itself seven times.

Exercise 240. Let X and Y be topological spaces. Prove that, for each n, $H_n(X + Y)$ is isomorphic to $H_n(X) + H_n(Y)$, where each "+" denotes the direct sum in the appropriate category.

Exercise 241. Find an example of a topological space which has a nonzero nth homology group for every n.

Exercise 242. Find the homology groups of an indiscrete topological space.

Exercise 243. Investigate the behavior of homology groups under direct products of topological spaces.

Exercise 244. Find an example of a topological space X such that $H_1(X)$ is not isomorphic to $\pi_1(X,x)$.

Homology: Relation to Homotopy

It is clear intuitively that the first homology group of a topological space should be closely related to its first homotopy group (relative to some base point). In each case, one deals essentially with "equivalence classes of curves," the only difference being that the equivalence relation, for homotopy, is "is continuously deformable into" and, for homology, is "together form the boundary of a 2-surface." We briefly discuss this connection in this chapter.

Recall, from chapter 34, that I denotes the subspace $[0,1]$ of the real line and, from chapter 35, that K_1 denotes the subspace of the topological plane consisting of (a_0,a_1) with $a_0 \geq 0$, $a_1 \geq 0$, and $a_0 + a_1 = 1$. Let $K_1 \xrightarrow{\tau} I$ be the mapping which sends (a_0,a_1) to the number (point of I) a_1. This τ is, of course, an isomorphism of topological spaces.

Next, let X be a topological space, and x a point of X. Let ω be a point of $\Omega(X,x)$ (i.e., a loop based at x, i.e., a continuous mapping $I \xrightarrow{\omega} X$ with $\omega(0) = \omega(1) = x$). Then $\omega \circ \tau$ is a continuous mapping $K_1 \xrightarrow{\omega\circ\tau} X$, that is, a 1-simplex in X, that is (since $C_1(X)$ is the abelian group of all linear combinations, with integral coefficients, of 1-simplices), an element of $C_1(X)$. Thus a loop (based at x) defines an element of $C_1(X)$. This mapping of sets, $\Omega(X,x) \xrightarrow{\varsigma} C_1(X)$, will be used to relate homotopy to homology groups.

We next claim that, for any loop ω, $\varsigma(\omega)$ is in fact a 1-cycle (i.e., $\partial(\varsigma(\omega)) = 0$). [Proof: For ω a loop, the 1-simplex $\omega \circ \tau$ sends the point $(0,1)$ of K_1 to the point x of X, and the point $(1,0)$ of K_1 to the point x of X. Since $\partial(\omega \circ \tau)$ is the difference of these two 0-simplices and since they are the same 0-simplex, $\partial(\omega \circ \tau) = 0$.] (This property is, of course, clear geometrically: "a loop closes on itself.") Thus $\Omega(X,x) \xrightarrow{\varsigma} C_1(X)$ sends loops to 1-cycles, whence (since $Z_1(X)$ is the subgroup of $C_1(X)$ consisting of just the 1-cycles) we have $\Omega(X,x) \xrightarrow{\varsigma} Z_1(X)$.

The next step is to compare the two equivalence relations: "is homotopic to" and "is homologous to." We claim: if loops ω and ω' are homotopic, then the corresponding 1-cycles, $\varsigma(\omega)$ and $\varsigma(\omega')$, are homologous. Proof: Let $I \times I \xrightarrow{\psi} X$ be a homotopy from ω to ω', so $\psi_0 = \omega$, $\psi_1 = \omega'$, and $\psi_s(0) = \psi_s(1) = x$ for all s. Let $K_2 \xrightarrow{\rho} I \times I$ be the continuous mapping that sends (a_0,a_1,a_2) to the point $(1 - a_2,a_1)$ of $I \times I$, and $K_2 \xrightarrow{\rho'} I \times I$ the continuous mapping that sends (a_0, a_1, a_2) to $(a_0, 1 - a_1)$ (figure 116). Then each of $\psi \circ \rho$ and $\psi \circ \rho'$ is a continuous mapping from K_2 to X, that is, a 2-simplex in X. Let α be the element of $C_2(X)$ given by $\alpha = \psi \circ \rho - \psi \circ \rho'$. Then $\partial(\alpha) = \varsigma(\omega') - \varsigma(\omega)$, that is, $\varsigma(\omega)$ and $\varsigma(\omega')$ are homologous. (Intuitively, we "triangulate the square $I \times I$," so the homotopy $I \times I \xrightarrow{\psi} X$ leads to two 2-simplices in X. The triangulation is so chosen that the boundary of the difference of these two 2-simplices in X is just $\varsigma(\omega') - \varsigma(\omega)$. Thus, since $\varsigma(\omega') - \varsigma(\omega)$ is the boundary of this element of $C_2(X)$, $\varsigma(\omega)$ and $\varsigma(\omega')$ are homologous. Still more intuitively, "the continuous deformation from one loop to the other must be over a certain 2-surface, so the difference of the loops must be the boundary of this surface, so the loops must be homologous.")

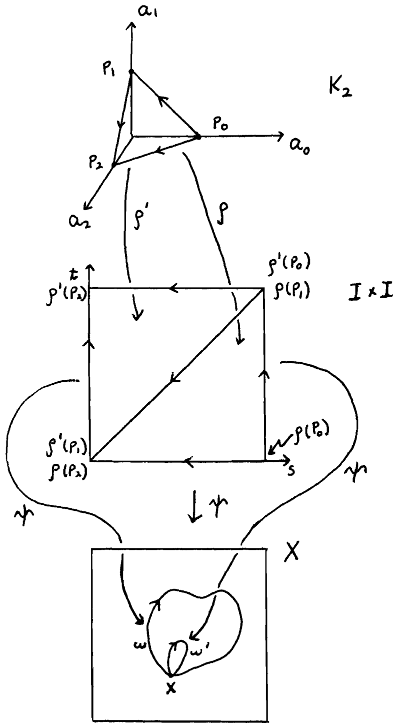

Figure 116

Thus every loop defines a 1-cycle, and homotopic loops define homologous 1-cycles. Recall, however, that $\pi_1(X,x)$ is the set of equivalence classes of $\Omega(X,x)$ under the equivalence relation "is homotopic to," while $H_1(X)$ is the set of equivalence classes of $Z_1(X)$ under the equivalence relation "is homologous to." Thus, by the remark above, $\Omega(X,x) \xrightarrow{\varsigma} Z_1(X)$ induces a mapping

$\pi_1(X,x) \xrightarrow{\varphi} H_1(X)$ (namely, given an element of $\pi_1(X,x)$, choose a loop in that equivalence class, apply ς to it, and find the equivalence class (element of $H_1(X)$) in which the resulting 1-cycle lies).

We now have a mapping of sets, $\pi_1(X,x) \xrightarrow{\varphi} H_1(X)$. But each set is actually a group. Our final claim: this φ is in fact a homomorphism of groups. To prove this, it suffices to show that, for any two loops ω and ω', the 1-cycle $\varsigma(\omega\omega')$ is homologous to the 1-cycle $\varsigma(\omega) + \varsigma(\omega')$. To this end, let $K_2 \xrightarrow{\sigma} X$ be the 2-simplex in X given by $\sigma(a_0,a_1,a_2) = \{\omega(a_0 - a_2)$ for $a_2 \leq a_0$, and $\omega(a_2 - a_0)$ for $a_0 < a_2\}$. (The edge $(0,a_1,a_2)$ is sent to the loop ω', the edge $(a_0,0,a_2)$ to $\omega\omega'$, and the edge $(a_0,a_1,0)$ to ω.) Thus $\partial(\omega) = \varsigma(\omega) + \varsigma(\omega') - \varsigma(\omega\omega')$, so $\varsigma(\omega) + \varsigma(\omega')$ is homologous to $\varsigma(\omega\omega')$.

Thus we obtain, finally, a homomorphism of groups, $\pi_1(X,x) \xrightarrow{\varphi} H_1(X)$. Note that this need not be an isomorphism, for example, because $\pi_1(X,x)$ need not be abelian, while $H_1(X)$ is always abelian. In fact, one can see this feature more geometrically: there exist (for certain X) loops ω and ω' that are not homotopic, but whose corresponding 1-cycles are in fact homologous. An intuitive example of such (the "doughnut with two holes") is shown in figure 117. The loop ω is not homotopic to ω' (the zero loop) because "you cannot continuously deform ω over the hole." Yet the corresponding 1-cycles together bound a 2-surface (the left half of the figure).

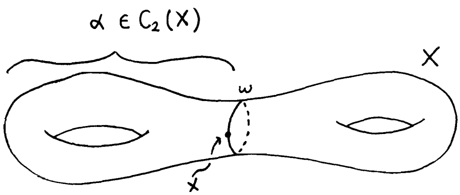

Figure 117

Exercise 245. Show that, for the real line, the topological plane, and the punctured plane, $\pi_1(X,x) \xrightarrow{\varphi} H_1(X)$ is an isomorphism.

Exercise 246. Why is there no way in general to obtain a homomorphism from $H_1(X)$ to $\pi_1(X,x)$?

Exercise 247. Prove that the commutator subgroup, $[\pi,\pi]$, of $\pi_1(X,x)$ is a subgroup of the kernel of φ. Thus, obtain a homomorphism from $\pi_1(X,x)/[\pi,\pi]$ (quotient group) to $H_1(X)$. Prove that, if X is curve-connected, this is an isomorphism. Why is "curve-connected" needed?

37

The Homology Functors

Fix a non-negative integer n. Then, as one would expect, "take the nth homology group" is a covariant functor from the category of topological spaces to the category of abelian groups. We now verify this assertion.

Recall (from chapter 17) that a covariant functor from one category to another is a rule which associates, with each object of the first category, an object of the second, and with each morphism of the first category a morphism of the second, such that the identity morphism and composition of morphisms are preserved. Fix a non-negative integer n. Then, for X any topological space, let $\mathbf{F}(X)$ be the nth homology group of X, $H_n(X)$. This is the first of the two "rules" needed to specify our functor.

Next, let $X \xrightarrow{\varphi} Y$ be a continuous mapping of topological spaces. We must find a homomorphism of groups, $H_n(X) \xrightarrow{\mathbf{F}(\varphi)} H_n(Y)$. Consider first an n-simplex in X, $K_n \xrightarrow{\sigma} X$. Then $K_n \xrightarrow{\varphi \circ \sigma} Y$ is continuous, whence $\varphi \circ \sigma$ is an n-simplex in Y. More generally, for $\alpha = 3\sigma - 7\sigma' + \sigma''$ an element of $C_n(X)$ (so σ, σ', and σ'' are n-simplices in X), set $\tilde{\varphi}(\alpha) = 3\varphi \circ \sigma - 7\varphi \circ \sigma' + \varphi \circ \sigma''$, so $\tilde{\varphi}(\alpha)$, as a linear combination, with integral coefficients, of n-simplices in Y, is an element of $C_n(Y)$. Thus we obtain a mapping $C_n(X) \xrightarrow{\tilde{\varphi}} C_n(Y)$, a mapping which is clearly a homomorphism of abelian groups. We next wish to figure out how these homomorphisms $\tilde{\varphi}$ interact with the boundary homomorphisms ∂. Consider the diagram of figure 118. We claim that

$$\begin{array}{ccccccccc} \cdots \xrightarrow{\partial} & C_{n+1}(X) & \xrightarrow{\partial} & C_n(X) & \xrightarrow{\partial} & C_{n-1}(X) & \xrightarrow{\partial} \cdots \\ & \downarrow \tilde{\varphi} & & \downarrow \tilde{\varphi} & & \downarrow \tilde{\varphi} & \\ \cdots \xrightarrow{\partial} & C_{n+1}(Y) & \xrightarrow{\partial} & C_n(Y) & \xrightarrow{\partial} & C_{n-1}(Y) & \xrightarrow{\partial} \cdots \end{array}$$

Figure 118

the interaction between $\tilde{\varphi}$ and ∂ is the following: this diagram commutes. Indeed, let $K_n \xrightarrow{\sigma} X$ be an n-simplex in X. Then $\tilde{\varphi}(\sigma) = \varphi \circ \sigma$, whence $\partial(\tilde{\varphi}(\sigma)) = \sum_{i=0}^{n} (-1)^i(\varphi \circ \sigma) \circ f_i$. On the other hand, $\partial(\sigma) = \sum_{i=0}^{n} (-1)^i \sigma \circ f_i$, whence $\tilde{\varphi}(\partial(\sigma)) = \sum_{i=0}^{n} (-1)^i \varphi \circ (\sigma \circ f_i)$. Thus $\partial \circ \tilde{\varphi} = \tilde{\varphi} \circ \partial$ when applied to any n-simplex in X, and hence (since everything is a homomorphism, since

$C_n(X)$ is the free abelian group on the set of n-simplices in X) $\partial \circ \tilde{\varphi} = \tilde{\varphi} \circ \partial$ when applied to any element of $C_n(X)$. Thus the diagram above indeed commutes. (This is not unexpected geometrically. Think of an n-simplex in X as an "n-dimensional tetrahedron sitting in X." Then φ "sends this n-dimensional tetrahedron sitting in X over to an n-dimensional tetrahedron sitting in Y, via $X \xrightarrow{\varphi} Y$." The statement of commutativity of the diagram above is, essentially, "you get the same result whether you first take the boundary of this tetrahedron in X and send that boundary, via φ, over to Y, or first send the whole tetrahedron over to Y via φ and there take its boundary.")

Commutativity of the diagram having been established, all that remains is algebra. We first claim that $C_n(X) \xrightarrow{\tilde{\varphi}} C_n(Y)$ takes n-cycles in X to n-cycles in Y. In fact, if α is any n-cycle in X (so $\partial(\alpha) = 0$), then $\tilde{\varphi}(\alpha)$ is an n-cycle in Y, for $\partial(\tilde{\varphi}(\alpha)) = \tilde{\varphi}(\partial(\alpha)) = \tilde{\varphi}(0) = 0$, where we used commutativity in the first step. Thus (recalling that Z_n is the subgroup of C_n consisting of n-cycles) we have $Z_n(X) \xrightarrow{\tilde{\varphi}} Z_n(Y)$. Now let α be an n-cycle in X which is in fact a boundary, that is, let $\alpha = \partial(\mu)$, with μ in $C_{n+1}(X)$. Then the n-cycle $\tilde{\varphi}(\alpha)$ in Y is also a boundary (of something in $C_{n+1}(Y)$), for $\tilde{\varphi}(\alpha) = \tilde{\varphi}(\partial(\mu)) = \partial(\tilde{\varphi}(\mu))$, where we used commutativity in the last step. Thus (recalling that B_n is the subgroup of Z_n consisting of n-cycles which are boundaries of elements of C_{n+1}) we have that $Z_n(X) \xrightarrow{\tilde{\varphi}} Z_n(Y)$ takes elements of the subgroup $B_n(X)$ of $Z_n(X)$ to the subgroup $B_n(Y)$ of $Z_n(Y)$. But $H_n = Z_n/B_n$, the quotient group. We define the homomorphism $H_n(X) \xrightarrow{\mathbf{F}(\varphi)} H_n(Y)$ as follows: given an element of $H_n(X)$ (i.e., a coset of $B_n(X)$ in $Z_n(X)$), choose an element of $Z_n(X)$ in that coset, apply $\tilde{\varphi}$ to it to get an element of $Z_n(Y)$, and find the coset of $B_n(Y)$ (i.e., the element of $H_n(Y)$) in which that element lies. This rule is independent of the choice of element of $Z_n(X)$, since $Z_n(X) \xrightarrow{\tilde{\varphi}} Z_n(Y)$ takes $B_n(X)$ to $B_n(Y)$.

We are now nearly done. We have, for each topological space X, an abelian group $\mathbf{F}(X)$ $(= H_n(X))$ and, for each continuous mapping $X \xrightarrow{\varphi} Y$, a homomorphism $\mathbf{F}(X) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(Y)$, namely that constructed above. We must verify that the two conditions for a covariant functor are satisfied. It is obvious that, for $X \xrightarrow{i_X} X$ the identity (continuous) mapping, $\mathbf{F}(X) \xrightarrow{\mathbf{F}(i_X)} \mathbf{F}(X)$ is the identity homomorphism. So let $X \xrightarrow{\varphi} Y \xrightarrow{\psi} Z$ be continuous mappings of topological spaces. We must show that $\mathbf{F}(\psi \circ \varphi) = \mathbf{F}(\psi) \circ \mathbf{F}(\varphi)$ (each side being a homomorphism from $\mathbf{F}(X)$ to $\mathbf{F}(Z)$). Let $K_n \xrightarrow{\sigma} X$ be an n-simplex in X. Then the n-simplices $\psi \circ (\varphi \circ \sigma)$ and $(\psi \circ \varphi) \circ \sigma$ in Z are the same (since composition of continuous mappings is associative). Hence $\widetilde{(\psi \circ \varphi)} = \tilde{\psi} \circ \tilde{\varphi}$ (each side a homomorphism from $C_n(X)$ to $C_n(Z)$. But this fact, together with the definition of $H_n(X) \xrightarrow{F(\varphi)} H_n(Y)$, implies immediately that $\mathbf{F}(\psi \circ \varphi) = \mathbf{F}(\psi) \circ \mathbf{F}(\varphi)$. Thus we have a functor.

Note that we obtain not just a single covariant functor from the category of topological spaces to the category of abelian groups, but an infinite number (i.e., one for each non-negative integer n). These are called the *homology functors.* Note also that the homology functors are "nontrivial" in the following sense: they are certainly not forgetful, and neither are they free (for we have no forgetful functor from the category of abelian groups to the category of topological spaces). Note finally that the statement that the homology functors are functors has some real content to it: we actually had to work, and prove things. In other words, the idea of a functor here made a contribution, in that it suggested what one should try to prove.

Finally, we remark that this chapter provides an example of how one works with homology. One first deals with the simplices, using properties of continuous mappings. Then, using properties of free abelian groups, one extends what one knows to the C_n. The next step is to see how this structure interacts with the boundary homomorphisms (here using both properties of continuous mappings and properties of abelian groups). Finally, since the C_n and the boundary homomorphisms together define the homology groups, the H_n, one proceeds (purely algebraically) to say what one can about the homology groups.

Exercise 248. Is $\mathbf{F}(\varphi)$ necessarily a monomorphism (resp., epimorphism) if $X \xrightarrow{\varphi} Y$ is?

Exercise 249. Let X be a topological space. The nth cohomology group of X is the group of all homomorphisms from $H_n(X)$ to Z, the additive group of integers. Construct the (contravariant) cohomology functors.

Uniform Spaces

Let X be a set. We may regard a topology on this set as providing one with the following structure: the notion of what are the neighborhoods of each point of X. On the other hand, a metric (chapter 26) on X provides somewhat more structure: not only can one speak of the neighborhoods of each point of X, but also one can assign to these neighborhoods a "numerical size." (For example, a neighborhood of "radius 7" about point x consists of all x' with $d(x,x') < 7$.) In particular, one can, in a metric space, "compare sizes of different neighborhoods." The structure we now wish to consider can be regarded as an intermediate one between these two extremes. We shall have "neighborhoods, together with a certain ability to compare them with regard to size, but with no precise numerical measure of that size."

It is convenient first to introduce a little notation. Let X be a set, and consider $X \times X$, the Cartesian product of X with itself. Denote by D the subset of $X \times X$ consisting of all elements of the form (x,x). This D is called the *diagonal* of $X \times X$. For A any subset of $X \times X$, we write A^{-1} for the subset of $X \times X$ consisting of all pairs (x,x') with (x',x) in A. Finally, for A and B subsets of $X \times X$, we write AB for the subset of $X \times X$ consisting of all pairs (x,x') for which there exists an $\underline{x}$ with $(x,\underline{x})$ in A and $(\underline{x},x')$ in B. Thus, for example, $D^{-1} = D$, $AD = DA = A$ and $(A^{-1})^{-1} = A$ for every A, and $(AB)C = A(BC)$ for every A, B, and C. (We do not have a group, however, for, e.g., it is false in general that $A^{-1}A = D$.)

A *uniform space* consists of two things—i) a set X, and ii) a collection of subsets of $X \times X$ (subsets in this collection will be called *entourages*)—subject to the following four conditions:

1. The intersection of all the entourages is precisely the diagonal D.
2. For any entourage A, A^{-1} is also an entourage.
3. For any entourage A, there exists an entourage B with $BB \subset A$.
4. The intersection of two entourages is another, and any superset of an entourage is another.

A number of properties of the entourages are immediate from this definition. Condition 1 implies that every entourage is a superset of the diagonal. Hence, for A and B entourages, AB is a superset of both A and B (for $AD = A$ and $DB = B$). By conditions 2 and 4, for A any entourage, $A \cap A^{-1} = B$ is another. This B is symmetric, in the sense that $B^{-1} = B$ and, by the remark above, is a subset of A. Thus every entourage is a superset of a symmetric entourage. Repeated application of condition 3 establishes that, for A

any entourage, there is an entourage B with $BB \cdots B$ (n times) a subset of A. Finally, condition 4 guarantees that the intersection of any finite number of entourages is an entourage.

There is one important example which both motivates this definition and makes sense out of it.

Example. Let X, $d(\ ,\ )$ be a metric space. For each positive number ϵ, denote by K_ϵ the subset of $X \times X$ consisting of all (s,s') with $d(x,x') < \epsilon$. Let the entourages be the subsets A of $X \times X$ for which there exists an ϵ with A a superset of K_ϵ. We claim that this set X, with these entourages, is a uniform space. (1) Since, for every x, $d(x,x) = 0$, the diagonal is a subset of every entourage. For $x \neq x'$, $d(x,x')$ is positive, whence there is some K_ϵ (i.e., some entourage) not containing (x,x'). Thus the intersection of all the entourages is the diagonal. (2) Since, for any x and x', $d(x,x') = d(x',x)$, we have $K_\epsilon^{-1} = K_\epsilon$ for each ϵ. Hence, for A any entourage (so $A \supset K_\epsilon$), we have $A^{-1} \supset K_\epsilon$, whence A^{-1} is an entourage. (3) Let A be any entourage, so $A \supset K_\epsilon$. By the triangle inequality, $K_{\epsilon/2}K_{\epsilon/2} \subset K_\epsilon$. But $K_{\epsilon/2}$ is an entourage. (4) Let A and A' be entourages (so $A \supset K_\epsilon$ and $A' \supset K_{\epsilon'}$). Denote by $\underline{\epsilon}$ the smaller of ϵ and ϵ'. Then $A \cap A' \supset K_{\underline{\epsilon}}$, whence $A \cap A'$ is an entourage. It is obvious that any superset of an entourage is an entourage. Thus every metric space yields, by the construction above, a uniform space.

This example suggests that one think of a uniform space as "a metric space modified so that the real numbers lose their special role." Thus one interprets the four conditions in the definition of a uniform space as follows. The first condition reflects the condition for a metric space that $d(x,x')$ is non-negative and vanishes when and only when $x = x'$. The second condition reflects symmetry of the metric, $d(x,x') = d(x',x)$. The third condition for a uniform space is a sort of truncated version of the triangle inequality for metric spaces. The first part of the fourth condition reflects the ordering on the set of real numbers. Finally, the last part of the fourth condition ensures that no "additional structure has been surreptitiously included in the uniform space by cutting down the number of entourages."

Some further special cases of the example above will make the situation clearer. It is convenient, first, to have available the following definition. Given a set X, a collection of subsets of $X \times X$ satisfying the four conditions above is called a *uniformity* on X.

Example. Let $X = \mathbf{R}$, the set of real numbers, and, for r and r' numbers, set $d(r,r') = |r - r'|$ (the usual metric on the reals). Then $X \times X$ can be represented as the plane as shown in figure 119, and the diagonal D is the line indicated. For each positive ϵ, K_ϵ is a "strip centered on D." The entourages are the supersets of these K_ϵ. This uniform space is called the ***uniform space of reals.***

Example. Let $X = \mathbf{R}$, and set

$$\tilde{d}(r,r') = |r - r'|/(1 + |r - r'|) .$$

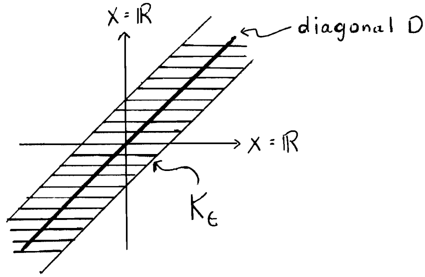

Figure 119

Then, noting that this is a metric, we have a metric space. (Note that this metric space is essentially different from the one above, for here "no two points are a distance greater than 1 apart.") Each $\tilde{K}_{\tilde{\epsilon}}$ in this case is just one of the K_ϵ above (namely for $\epsilon = \tilde{\epsilon}/(1 - \tilde{\epsilon})$). Thus "the K_ϵ in the two cases are essentially the same, but are labeled in a different way by the number ϵ." We obtain, therefore, the same uniformity in this example as in the previous one.

Example. Let $X = \mathbf{R}$, and set

$$\hat{d}(r,r') = |(2r + \sin r) - (2r' + \sin r')| \ ,$$

noting that we thus obtain a metric space. A typical $\hat{K}_\epsilon$ is shown in figure 120 (the "waviness" is caused by the sin's in $\hat{d}(\ ,\ )$). Note first that each of

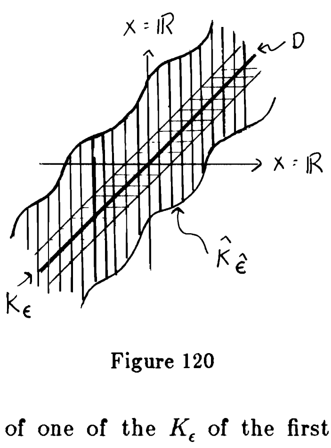

Figure 120

these $\hat{K}_\epsilon$ is a superset of one of the K_ϵ of the first example. Thus every entourage in the present uniformity on X is an entourage in the uniformity above. Similarly, each $\hat{K}_\epsilon$ is a subset of a K_ϵ. Thus every entourage in the uniform space of reals is an entourage in the present uniformity. Thus the metrics $d(\ ,\ )$ and $\hat{d}(\ ,\ )$ define the same uniformity on X.

Example. Let $X = \mathbf{R}$, and set

$$\breve{d}(r,r') = |\tan^{-1}r - \tan^{-1}r'|$$

(where, by "$\tan^{-1}$," we mean the value between $-\pi/2$ and $\pi/2$). This is a metric space. A typical $\breve{K}_\epsilon$ is shown in figure 121. Note that each such $\breve{K}_\epsilon$—and hence each entourage in this uniformity—is a superset of some K_ϵ, and hence an entourage in the uniform space of reals. The converse, however, is false. A K_ϵ (i.e., an entourage) in the uniform space of reals is not a superset of one of these $\breve{K}_\epsilon$. Thus none of these entourages in the uniform space of reals is an entourage in the present uniformity. We have here a uniformity different from that of the uniform space of reals.

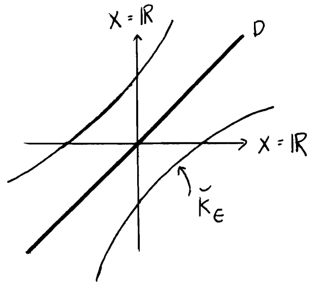

Figure 121

It is clear from these examples that the uniform space obtained from a metric space "ignores certain of the finer numerical details of the metric but does retain the asymptotic, gross shapes of the K_ϵ."

We next consider the relation between uniform spaces and topological spaces. Let X be a uniform space. We call a subset O of X open if, for each point x of O, there is an entourage A such that, whenever (x,x') is in A, x' is in O. We claim that these open sets define a topology on the set X. [Proof: (1) The empty set (which contains no x) and X itself (which contains every x') are certainly open. (2) Let O_λ (λ in Λ) be open, and let x be a point of $\bigcup_\Lambda O_\lambda$. Then x is a point of some O_λ, whence there is an entourage A with x' in O_λ whenever (x,x') is in A. Thus x' is in $\bigcup_\Lambda O_\lambda$ whenever (x,x') is in A, whence $\bigcup_\Lambda O_\lambda$ is open. (3) Let O and $\tilde{O}$ be open, and let A and $\tilde{A}$ be corresponding entourages. Then $A \cap \tilde{A}$ is an entourage. But, for x in $O \cap \tilde{O}$, and (x,x') in $A \cap \tilde{A}$, x' is in both O and $\tilde{O}$, whence x' is in $O \cap \tilde{O}$. Thus $O \cap \tilde{O}$ is open.] Thus a uniformity on a set X yields, by the construction above, a topology on the set X. We can now go back and reinterpret the entourages in our original uniform space topologically. In fact, we have

THEOREM 44. *Let X be a uniform space, x a point of X, and A an entourage. Denote by N the collection of all points x' of X with (x,x') in A. Then N is a neighborhood (in the topology above) of x.*

Proof. Choose entourage B_1 with $B_1B_1 \subset A$, then entourage B_2 with $B_2B_2 \subset B_1$, then entourage B_3 with $B_3B_3 \subset B_2$, etc. Then $B_1 \subset A$, and $B_1B_2 \subset B_1B_1 \subset A$, and $B_1B_2B_3 \subset B_1B_2B_2 \subset B_1B_1 \subset A$, etc. Thus, since each $B_1B_2 \cdots B_n$ is a subset of A, $B = B_1 \cup B_1B_2 \cup B_1B_2B_3 \cup \cdots$ is a subset of A. Denote by O the subset of X consisting of all x' with (x,x') in B. Note that x is a point of O and that (since $B \subset A$) O is a subset of N. The proof that N is a neighborhood is completed by showing that this O is open. Let x' be any point of O (i.e., let (x,x') be a point of B, say (x,x') is in $B_1B_2 \cdots B_n$). Then, for any x'' with (x',x'') a point of B_{n+1}, (x,x'') is a point of $B_1B_2 \cdots B_nB_{n+1}$, and therefore a point of B. That is, for any x'' with (x',x'') a point of B_{n+1}, we have that x'' is a point of O. That is, O is open. □

Thus the "topological viewpoint" toward uniform spaces is this. We regard the entourages of uniform space X as obtained as follows. For each point (x,x) of the diagonal, one chooses a neighborhood of x and considers the

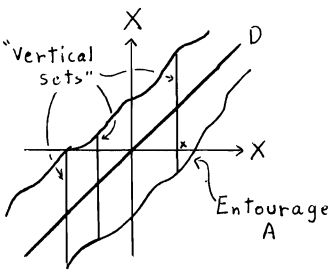

Figure 122

"vertical set" (figure 122) consisting of all (x,x') with x' in this neighborhood. Repeating for each point of the diagonal and taking the union of these "vertical sets," we obtain a candidate for an entourage. Condition 4 for a uniform space just reflects the basic properties of neighborhoods (finite intersections and supersets of neighborhoods are neighborhoods), while condition 1 guarantees that one can find "small neighborhoods." However, not every candidate, constructed as above, will in general be an entourage. Condition 2 requires, roughly speaking, that "the neighborhood chosen for each point x vary continuously with x." Finally, condition 3 requires that "into every neighborhood one can fit neighborhoods about half as large." Since an entourage represents a "system of neighborhoods, one attached to each point of x, varying continuously with x," the entourages give more information than just the neighborhoods. That is, a uniformity gives more information than a topology. On the other hand, a metric gives even more information, namely that of "the numerical size of a neighborhood."

Note that all four uniformities on the set **R** of real numbers given in the earlier examples yield the same topology on **R**, namely that of the real line. Thus different uniformities can yield the same topology. Note also that, for X, $d(\ ,\ )$ a metric space, if we first take the corresponding uniformity and then the corresponding topology, we obtain the same topology as would be obtained directly from the metric space (by the construction of chapter 26). Finally, note that every topology that arises from a uniformity is Hausdorff. [Proof: Let X be a uniform space, and let x and x' be distinct points of X. Choose entourage A with (x,x') not in A, and entourage $B = B^{-1}$ with $BB \subset A$. Then the neighborhoods of x and x' constructed from this B as in theorem 44 do not intersect.]

All we have done so far is to define, motivate, and discuss intuitively the notion of a uniform space. We next wish to introduce the notion of a "structure-preserving mapping" between uniform spaces. Let X and Y be uniform spaces, and let $X \xrightarrow{\varphi} Y$ be a mapping of sets. Let $X \times X \xrightarrow{\tilde{\varphi}} Y \times Y$ be the mapping which sends the point (x,x') of $X \times X$ to the point $(\varphi(x), \varphi(x'))$ of $Y \times Y$. (Note, incidentally, that this takes the diagonal of $X \times X$ to the diagonal of $Y \times Y$.) This $X \xrightarrow{\varphi} Y$ is said to be *uniformly continuous* if, for each entourage A of uniform space Y, $\tilde{\varphi}^{-1}[A]$ is an entourage of uniform space X. (Note how similar this notion is to that of a continuous mapping of topological spaces.)

Example. Let each of X and Y be the uniform space of reals, so $\mathbf{R} \xrightarrow{\varphi} \mathbf{R}$. Then (since the entourages in this case are the supersets of the K_ϵ) $\mathbf{R} \xrightarrow{\varphi} \mathbf{R}$ is uniformly continuous if and only if, for each K_ϵ, $\tilde{\varphi}^{-1}[K_\epsilon]$ is a superset of some $K_{\epsilon'}$. That is (reintroducing the metric), $\mathbf{R} \xrightarrow{\varphi} \mathbf{R}$ is uniformly continuous if and only if, for any positive ϵ, there exists a positive ϵ' such that $|\varphi(r) - \varphi(r')| < \epsilon$ whenever $|r - r'| < \epsilon'$. This will be recognized as the standard definition of uniform continuity for functions of a real variable. (It says that "you tell me how close you want $\varphi(r)$ and $\varphi(r')$ to be, and I will tell you how close r and r' must be to guarantee the desired closeness of $\varphi(r)$ and $\varphi(r')$." Thus, for example, $\mathbf{R} \xrightarrow{\varphi} \mathbf{R}$ with $\varphi(r) = r^2$ (figure 123) is not uniformly continuous (for, given ϵ, no matter how small one chooses ϵ', one can find (very large) r and r' within ϵ' but whose squares are not within ϵ). On the other hand, $\mathbf{R} \xrightarrow{\varphi} \mathbf{R}$ with $\varphi(r) = 1/(1 + r^2)$ is uniformly continuous.

It is immediate from the definition that the composition of two uniformly continuous mappings is uniformly continuous.

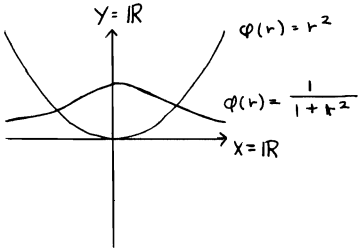

Figure 123

Let the objects be uniform spaces, the morphisms uniformly continuous mappings of uniform spaces, and composition composition of uniformly continuous mappings. We thus obtain a category, called the ***category of uniform spaces.***

We have seen that every uniform space defines a topological space (in fact, a Hausdorff one). It is natural to try to describe this state of affairs by means of a functor from the category of uniform spaces to the category of topological spaces. For X a uniform space, let $\mathbf{F}(X)$ be the corresponding topological space. Next, let $X \xrightarrow{\varphi} Y$ be a uniformly continuous mapping of uniform spaces. Then $X \xrightarrow{\varphi} Y$ is also a continuous mapping of the corresponding topological spaces (since, by theorem 44 and the definition of uniform continuity, inverse images, by φ, of neighborhoods in Y are neighborhoods in X, whence, by theorem 31, φ is continuous). Thus we have a continuous mapping $\mathbf{F}(X) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(Y)$ of topological spaces. That the defining properties of a covariant functor are satisfied is immediate. We shall (because of the "topological viewpoint" toward uniform spaces) regard this as a forgetful functor (you "forget how the neighborhoods are assembled into entourages but not what the neighborhoods are").

Why does one not try to introduce a forgetful functor from the category of metric spaces to the category of uniform spaces? The reason is that there just does not seem to be any well-behaved "category of metric spaces." It is easy to decide what the objects should be for such a category: the problem is to come up with a decent notion of a morphism. If you like, the notion of a uniform space arises from "categorization of that of a metric space." Finally, we remark that many of the constructions available in a metric space are both available and simpler in the underlying uniform space. It often happens in physical applications that one "has a metric space, but the details of the actual numerical values of the metric are essentially gauge." In this situation, the underlying uniform space is almost always what is physically significant.

Exercise 250. Fix $\mathbf{R} \xrightarrow{f} \mathbf{R}$, a mapping of sets. Find a necessary and sufficient condition on f that $d(r,r') = |f(r) - f(r')|$ be a metric on $\mathbf{R}$. Find a necessary and sufficient condition for this metric to yield the uniform space of reals.

Exercise 251. Let X be a uniform space. Prove that each entourage is a superset of an open (in the product topology on $X \times X$) superset of the diagonal.

Exercise 252. Let X be a topological space, and let the entourages be supersets of open (in the product topology on $X \times X$) supersets of the diagonal. Decide whether each of the four defining conditions for a uniform space is satisfied in general by these entourages.

Exercise 253. Find a set X and a uniformity on X which comes from no metric on X.

Exercise 254. Discuss products, sums, isomorphisms, and subobjects in the category of uniform spaces. Prove that the uniform space of the fifth example of this chapter is isomorphic to the subspace $(-\pi/2,\pi/2)$ of the uniform space of reals.

Exercise 255. Let X be a set. Partially order the set of uniformities on X, and introduce coarser and finer uniformities. Does there exist, for any collection of uniformities on X, a finest one coarser than all in this collection?

Exercise 256. Can two distinct uniformities on set X both yield the discrete topology on X? the indiscrete topology?

Exercise 257. Prove that two uniformities on set X that yield the same compact topology on X coincide.

Exercise 258. Let X and Y be uniform spaces, and let $X \xrightarrow{\varphi} Y$ be continuous (on topological spaces). Prove that, if X is compact, φ is uniformly continuous.

39

The Completion of a Uniform Space

Let X be the subspace (0,1) of the real line. One has the intuitive feeling that "within X is somehow recorded the information that X could naturally accept the two points '0' and '1' as 'additional boundary points.'" On what mathematical structure is this intuitive feeling based? It could hardly be just the topological structure of X, for, for example, X is isomorphic, as a topological space, to the real line, while the affixing of two additional points (now at "minus infinity" and "plus infinity") seems somewhat less natural for the real line. Introduce the metric on X given by $d(r,r') = |r - r'|$. One would now like to reformulate the idea "X could naturally accept '0' and '1' as additional boundary points" using this metric, for example, by something like "you can get to the edges of X by traversing only a finite distance." The problem now is that "X does not have actual points at its edges to take the distance from." Fortunately, it is possible to express this idea within X itself (i.e., without already having the boundary points). A Cauchy sequence in metric space X is a sequence, $x_1, x_2, \ldots$, of points of X having the following property: given any positive number ϵ, there is an integer n such that $d(x_{n'}, s_{n''}) < \epsilon$ whenever $n' \geq n$ and $n'' \geq n$. Thus the points of a Cauchy sequence "get closer and closer to each other." In the example above, $x_1 = 1/2$, $x_2 = 3/4$, $x_3 = 7/8, \cdots$ is a Cauchy sequence (which is "trying to converge to the point '1,' except that this is not a point of the metric space X "). Thus Cauchy sequences (note that this is a metric-space concept and not merely a topological one) provide "an internal mechanism for detecting missing points."

Thus we can "detect the absence of natural edge-points" in a metric space and not in a topological space. It is natural to ask how uniform spaces fit into this scheme. We shall see in this chapter that, in the category of uniform spaces, not only "detection," but also "restoration" is possible.

Let X be a uniform space. By a net in X, we mean a directed set Δ together with a mapping from Δ to X. (As in the topological case, we write x_δ for the point of X to which this mapping sends point δ of Δ.) Such a net is called a *Cauchy net* if it has the following property: given any entourage A, there is a δ in Δ such that $(x_{\delta'}, x_{\delta''})$ is in A whenever $\delta' \geq \delta$ and $\delta'' \geq \delta$.

Example. Any Cauchy sequence in a metric space is a Cauchy net (with the directed set Δ the set of positive integers) on the underlying uniform space.

Thus Cauchy nets generalize Cauchy sequences in two directions: one passes from the directed set of positive integers to any directed set, and from metric spaces to uniform spaces.

The following suggests that Cauchy nets appropriately describe the structure in which we are interested.

THEOREM 45. *Let X be a uniform space, and x_δ (δ in Δ) a net in X. Suppose that this net converges, in the underlying topology on X, to point $\underline{x}$ of X. Then this net is a Cauchy net.*

Proof. Let A be any entourage in the uniform space X. Choose entourage B with $B = B^{-1}$ and $BB \subset A$. By theorem 44, the set N of all x with $(\underline{x},x)$ in B is a neighborhood of $\underline{x}$. Since our net converges to $\underline{x}$, there is a δ in Δ with $x_{\delta'}$ in N whenever $\delta' \geq \delta$. That is, $(\underline{x},x_{\delta'})$ is in B whenever $\delta' \geq \delta$. Therefore, for $\delta' \geq \delta$ and $\delta'' \geq \delta$, we have $(\underline{x},x_{\delta'})$ in B and $(\underline{x},x_{\delta''})$ in B, whence (since $B = B^{-1}$ and $BB \subset A$) $(x_{\delta'},x_{\delta''})$ is in A. Thus the net is a Cauchy net. □

It is clear from theorem 45 that a uniform space X normally has many Cauchy nets, for example, the "constant nets," with $x_\delta = \underline{x}$ for all δ, where $\underline{x}$ is some fixed point of X. The question of "missing points" hangs on whether there are Cauchy nets which do not converge to any point of X. A uniform space X is said to be *complete* if every Cauchy net in X converges to some point of X.

Example. The uniform space discussed in the introduction to this chapter is not complete, for we displayed in that introduction a Cauchy net which does not converge. The uniform space of reals is complete. [Sketch of proof: Let x_δ (δ in Δ) be a Cauchy net in the uniform space of reals. Then there exists a δ such that, whenever $\delta' \geq \delta$ and $\delta'' \geq \delta$, $(x_{\delta'},x_{\delta''})$ is in the entourage K_1. Thus all $x_{\delta'}$ with $\delta' \geq \delta$ lie in the closed interval $[x_\delta - 1, x_\delta + 1]$ of the real line. Since this subspace of the real line is compact, the Cauchy net has an accumulation point. The next theorem will show that this net therefore converges.]

A couple of theorems will illustrate that the notion of completeness indeed has the properties one would expect.

THEOREM 46. *Let X be a uniform space, and suppose that the underlying topological space is compact. Then the uniform space X is complete.*

Proof. Let x_δ (δ in Δ) be a Cauchy net in X. Then, since the underlying topology is compact, this net has an accumulation point $\underline{x}$. We shall show that this net in fact converges to $\underline{x}$. Let N be a neighborhood of $\underline{x}$. Choose entourage A such that x is in N whenever $(\underline{x},x)$ is in A, and entourage B with $B = B^{-1}$ and $BB \subset A$. Since our net is Cauchy, there is a δ such that, whenever $\delta' \geq \delta$ and $\delta'' \geq \delta$, $(x_{\delta'},x_{\delta''})$ is in B. Since the net has $\underline{x}$ as an accumulation point, there is a $\underline{\delta} \geq \delta$ with $(\underline{x},x_{\underline{\delta}})$ in B. Now consider any $\delta' \geq \underline{\delta}$ (so, in particular, $\delta' \geq \delta$). Then $(x_{\underline{\delta}},x_{\delta'})$ is in B, while $(\underline{x},x_{\underline{\delta}})$ is also in B.

Hence (since $BB \subset A$), $(\underline{x}, x_{\delta'})$ is in A, that is, $x_{\delta'}$ is in N. We have shown that, for any $\delta' \geq \underline{\delta}$, $x_{\delta'}$ is in N. Hence the net converges to $\underline{x}$. □

Intuitively, since the net is Cauchy, its points "get closer and closer to each other," while, since the net is in a compact space, it "continually comes back and gets close to some $\underline{x}$," with the result that the net must converge to $\underline{x}$.

For the second result illustrating the notion of completeness, we need a definition. Let X be a uniform space, and Y a subset of X. Then $Y \times Y$ is a subset of $X \times X$. Let the entourages for Y be the intersections of the entourages for X with the subset $Y \times Y$ of $X \times X$. It is easily checked that we thus obtain a uniformity on Y. This uniform space Y is called a *uniform subspace* of X.

THEOREM 47. *Let X be a complete uniform space, and let Y be a closed (in the underlying topology on X) subset of X. Then the uniform subspace Y is also complete.*

Proof. Let y_δ (δ in Δ) be a Cauchy net in Y. Then, since Y is a uniform subspace of X, this is also a Cauchy net in X. Since X is complete, it converges, say to the point $\underline{x}$ of X. Since Y is closed, we have by theorem 33 that $\underline{x}$ is actually in Y. Thus, since every Cauchy net in Y converges to a point of Y, Y is complete. □

Again, "since no points are missing from X and since Y, being closed, took every point of X that it possibly could be expected to, there could hardly be points missing from Y." The following statement is perhaps not too misleading: completeness is to uniform spaces as compactness is to topological spaces.

Let the objects be complete uniform spaces, the morphisms uniformly continuous mappings of complete uniform spaces, and composition composition composition of uniformly continuous mappings. We thus obtain the *category of complete uniform spaces.*

We now have the notion of "no points are missing" from a uniform space, namely that of completeness. It is natural to ask whether, given a uniform space which is not complete, there is some unique way to "attach the missing points." Let X be a uniform space. A *completion* of X is a complete uniform space $\overline{X}$, together with a uniformly continuous mapping $X \xrightarrow{\varphi} \overline{X}$, such that the following condition is satisfied: given any complete uniform space Y, together with a uniformly continuous mapping $X \xrightarrow{\alpha} Y$, there is a unique uniformly continuous mapping $\overline{X} \xrightarrow{\varsigma} Y$ such that the diagram of figure 124 commutes. (Note that this is a free construction (via the obvious forgetful functor from the category of complete uniform spaces to the category of uniform spaces). Hence a completion, when one exists, is unique in the appropriate sense.)

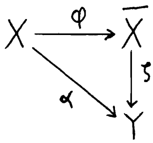

Figure 124

We wish to show that every uniform space possesses a completion. The idea is to let $\bar{X}$ consist of "points of X, together with additional ideal points which serve as limit points for Cauchy nets in X which do not already have limit points in X."

Denote by **C** the collection of all Cauchy nets in uniform space X. For x_δ (δ in Δ) and x_γ (γ in Γ) two such Cauchy nets, write $\{x_\delta\} \approx \{x_\gamma\}$ if the following condition is satisfied: given any entourage A in X, there exists a δ in Δ and a γ in Γ such that $(x_{\delta'},x_{\gamma'})$ is in A whenever $\delta' \geq \delta$ and $\gamma' \geq \gamma$. (Thus "$\approx$" means that "the two Cauchy nets approach each other.") We now claim that this "$\approx$" is in fact an equivalence relation on the set **C** of Cauchy nets. [Proof: (1) Let x_δ (δ in Δ) be a Cauchy net, so, for each entourage A, there exists a δ in Δ such that $(x_{\delta'},x_{\delta''})$ is in A whenever $\delta' \geq \delta$ and $\delta'' \geq \delta$. Hence $\{x_\delta\} \approx \{x_\delta\}$. (2) Let $\{x_\delta\} \approx \{x_\gamma\}$, and let A be any entourage. Then, since A^{-1} is an entourage, there exists a δ and a γ such that $(x_{\delta'},x_{\gamma'})$ is in A^{-1} whenever $\delta' \geq \delta$ and $\gamma' \geq \gamma$. That is, $(x_{\gamma'},x_{\delta'})$ is in A whenever $\delta' \geq \delta$ and $\gamma' \geq \gamma$. Hence $\{x_\gamma\} \approx \{x_\delta\}$. (3) Let $\{x_\delta\} \approx \{x_\gamma\}$ and $\{x_\gamma\} \approx \{x_\kappa\}$. Let A be any entourage. Choose entourage B with $BB \subset A$. By the two assumed instances of "$\approx$" there exists a δ and a $\hat{\gamma}$ such that $(x_{\delta'},x_{\gamma'})$ is in B whenever $\delta' \geq \delta$ and $\gamma' \geq \hat{\gamma}$, and a $\tilde{\gamma}$ and a κ such that $(x_{\gamma'},x_{\kappa'})$ is in B whenever $\gamma' \geq \tilde{\gamma}$ and $\kappa' \geq \kappa$. Fix this δ, this κ, and an element γ of Γ with $\gamma \geq \hat{\gamma}$ and $\gamma \geq \tilde{\gamma}$. Then, for $\delta' \geq \delta$ and $\kappa' \geq \kappa$, we have $(x_{\delta'},x_\gamma)$ in B and $(x_\gamma,x_{\kappa'})$ in B, whence (since $BB \subset A$) $(x_{\delta'},x_{\kappa'})$ is in A. That is, $\{x_\delta\} \approx \{x_\kappa\}$.] (Incidentally, note how, in proofs of this sort, the conditions for a uniform space are just right to give one what one wants. That is how one knows that the definition is "right.")

Denote by $\bar{X}$ the set of equivalence classes of **C** by the equivalence relation above. This $\bar{X}$ is our candidate for the underlying set of the completion of X. Note that an element of $\bar{X}$ is an equivalence class of Cauchy nets in X: we shall call a Cauchy net in a given equivalence class a representative of that element of $\bar{X}$.

The next step is to figure out what we should choose for the mapping φ from X to $\bar{X}$. That is easy: for x a point of X, let $\varphi(x)$ be the element of $\bar{X}$ (i.e., the equivalence class of Cauchy nets) that contains the constant Cauchy net, $x_\delta = x$ for all δ (in some directed set Δ). Note that Cauchy net x_γ (γ in

Γ) is in the equivalence class $\varphi(x)$ if and only if x_γ (γ in Γ) converges to x. [Proof: Cauchy net x_γ (γ in Γ) is equivalent to the constant Cauchy net at x if and only if, for every entourage A, there exists a γ such that $(x,x_{\gamma'})$ is in A whenever $\gamma' \geq \gamma$. But, by theorem 44, this holds if and only if the net x_γ (γ in Γ) converges to x.] Thus the points of $\overline{X}$ are of just two types: those of the form $\varphi(x)$ for some x in X (in which case the point of $\overline{X}$ is just the equivalence class of all Cauchy nets which converge to x), and those not of the form $\varphi(x)$ (in which case the point of $\overline{X}$ is an equivalence class of Cauchy nets none of which converge to any point of X). Furthermore, since the underlying topological space X is Hausdorff, no net therein can converge to more than one point. Hence $\varphi(x) \neq \varphi(x')$ for $x \neq x'$. Thus (as one would certainly have wished) the mapping of sets $X \xrightarrow{\varphi} \overline{X}$ is always one-to-one, and is onto if and only if X is complete.

We so far have the set $\overline{X}$ and the mapping of sets $X \xrightarrow{\varphi} \overline{X}$. The next step is to make the set $\overline{X}$ into a uniform space, that is, to select a suitable collection of entourages. Clearly, we have no choice but to construct the entourages for $\overline{X}$ using those for X. Fix an entourage A for X. For x_δ (δ in Δ) and x_γ (γ in Γ) Cauchy nets in X, we $\{x_\delta\} \approx_A \{x_\gamma\}$ if the following property is satisfied: there exists a δ in Δ and a γ in Γ such that $(x_{\delta'},x_{\gamma'})$ is in A whenever $\delta' \geq \delta$ and $\gamma' \geq \gamma$. Intuitively, $\{x_\delta\} \approx_A \{x_\gamma\}$ means that "the Cauchy nets eventually get and remain within A of each other." Our earlier equivalence relation, for example, is easily formulated in this notation: $\{x_\delta\} \approx \{x_\gamma\}$ means that $\{x_\delta\} \approx_A \{x_\gamma\}$ for every entourage A. (Note also that $\{x_\delta\} \approx_A \{x_\gamma\}$ and $\{x_\gamma\} \approx_B \{x_\kappa\}$ imply $\{x_\delta\} \approx_{AB} \{x_\kappa\}$.) We now wish to "carry entourages over from X to $\overline{X}$." Let A be any entourage for X. We write $\overline{A}$ for the subset of $\overline{X} \times \overline{X}$ consisting of all pairs of elements of $\overline{X}$ such that, for any representatives x_δ (δ in Δ) and x_γ (γ in Γ) of the entries in this pair, we have $\{x_\delta\} \approx_A \{x_\gamma\}$. Now let the entourages for $\overline{X}$ consist of the supersets of the $\overline{A}$.

We claim that this set of entourages for $\overline{X}$ in fact defines a uniformity on the set $\overline{X}$. Proof: (1) Let $\bar{x}$ be in $\overline{X}$, and let x_δ (δ in Δ) and x_γ (γ in Γ) be any representatives of $\bar{x}$. Then, for any entourage A for X, $\{x_\delta\} \approx_A \{x_\gamma\}$ (since these two are in the same equivalence class), whence $(\bar{x},\bar{x})$ is in $\overline{A}$. Hence every entourage for $\overline{X}$ contains the diagonal of $\overline{X} \times \overline{X}$. For $\bar{x}$ and $\bar{x}'$ distinct elements of $\overline{X}$, choose representatives x_δ (δ in Δ) and x_γ (γ in Γ). Then, since these two nets are in different equivalence classes, there is an entourage A for X with $\{x_\delta\} \not\approx_A \{x_\gamma\}$. Hence $(\bar{x},\bar{x}')$ is not in $\overline{A}$. Thus the only elements of $\overline{X} \times \overline{X}$ in every entourage for $\overline{X}$ are those on the diagonal. (2) It is immediate from the definition that, for A any entourage for X, $\overline{A}^{-1} = \overline{A^{-1}}$. Hence, since the entourages for $\overline{X}$ are the supersets of the $\overline{A}$, the inverse of one such entourage is another. (3) Let A be any entourage for X, and choose entourage B with $BB \subset A$. Let $(\bar{x},\bar{x}')$ and $(\bar{x}',\bar{x}'')$ both be in $\overline{B}$, so, for x_δ (δ in Δ), x_γ

(γ in Γ), and x_κ (κ in K) representatives of $\bar{x}$, $\bar{x}'$, and $\bar{x}''$, respectively, we have $\{x_\delta\} \approx_B \{x_\gamma\}$ and $\{x_\gamma\} \approx_B \{x_\kappa\}$. But this implies $\{x_\delta\} \approx_{BB} \{x_\kappa\}$ and hence $\{x_\delta\} \approx_A \{x_\kappa\}$. Thus we have $(\bar{x},\bar{x}'')$ in $\bar{A}$, that is, we have $\bar{B}\bar{B} \subset \bar{A}$. (4) Let A and B be entourages for X. Then, for $(\bar{x},\bar{x}')$ in $\overline{A \cap B}$ and x_δ (δ in Δ) and x_γ (γ in Γ) any representatives of $\bar{x}$ and $\bar{x}'$, respectively, we have $\{x_\delta\} \approx_{A\cap B} \{x_\gamma\}$, whence $\{x_\delta\} \approx_A \{x_\gamma\}$ and $\{x_\delta\} \approx_B \{x_\gamma\}$. Thus $(\bar{x},\bar{x}')$ is in both $\bar{A}$ and $\bar{B}$, and hence in $\bar{A} \cap \bar{B}$. We conclude that $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$, that is, that the intersection of two entourages for $\bar{X}$ is another. It is obvious that any superset of an entourage for $\bar{X}$ is another.

We have so far, starting from the uniform space X, constructed i) a set $\bar{X}$, ii) a mapping of sets, $X \xrightarrow{\varphi} \bar{X}$, and iii) a uniformity on the set $\bar{X}$. The next step is to show that this mapping φ (now of uniform spaces) is uniformly continuous. Let A be any entourage for X, and choose entourage B with $BBB \subset A$. Let (x,x') be a point of B, and let x_δ (δ in Δ) and x_γ (γ in Γ) be representatives of $\varphi(x)$ and $\varphi(x')$, respectively. Then $\{x_\delta\} \approx_B \{x\}$ (since these nets are in the same equivalence class), $\{x\} \approx_B \{x'\}$ (since (x,x') is in B), and $\{x'\} \approx_B \{x_\gamma\}$ (since these nets are in the same equivalence class), whence $\{x_\delta\} \approx_{BBB} \{x_\gamma\}$. But $BBB \subset A$, so $\{x_\delta\} \approx_A \{x_\gamma\}$. That is, $(\varphi(x),\varphi(x'))$ is in $\bar{A}$. We have shown that $\tilde{\varphi}^{-1}[A]$ is a superset of some entourage (namely B) for X, whence the inverse image by $\tilde{\varphi}$ of any entourage for $\bar{X}$ is an entourage for X. Thus $X \xrightarrow{\varphi} \bar{X}$ is uniformly continuous.

Thus, starting from a uniform space X, we have constructed a uniform space $\bar{X}$ and a uniformly continuous mapping $X \xrightarrow{\varphi} \bar{X}$. The next step is to show that the uniform space $\bar{X}$ is in fact complete. Let $\bar{x}_\gamma$ (γ in Γ) be a Cauchy net in $\bar{X}$. We must find a point of $\bar{X}$ to which this net converges. (The idea is to use representatives of the $\bar{x}_\gamma$ to construct a Cauchy net in X which will be a representative of the limit point.) For each γ in Γ, $\bar{x}_\gamma$ is a point of $\bar{X}$: choose a representative, x_{δ_γ} (δ_γ in Δ_γ). (Thus, for each γ, Δ_γ is a directed set and each of these directed sets is associated with a certain Cauchy net in X.) Denote by Ω the set whose elements consist of an element $\underline{\gamma}$ of Γ together with, for each $\gamma \geq \underline{\gamma}$, an element $\underline{\delta}_\gamma$ of that Δ_γ. (That is, an element of Ω is one element each from certain of the Δ_γ, namely all those with $\gamma \geq \underline{\gamma}$.) For $(\underline{\gamma},\underline{\delta}_\gamma)$ and $(\underline{\gamma}',\underline{\delta}_\gamma')$ two elements of Ω, write $(\underline{\gamma},\underline{\delta}_\gamma) \leq (\underline{\gamma}',\underline{\delta}_\gamma')$ if $\underline{\gamma} \leq \underline{\gamma}'$ and, for each $\gamma \geq \underline{\gamma}'$, $\underline{\delta}_\gamma \leq \underline{\delta}_\gamma'$. (Thus "larger" in Ω is "moving farther along Γ before you start to pick elements of the Δ_γ, and farther up each Δ_γ.") Clearly, this set Ω with this ordering is directed (using first the fact that Γ is directed and then the fact that each Δ_γ is directed). We now use this directed set to construct a limit point of our original net in $\underline{X}$. For $\omega = (\underline{\gamma},\underline{\delta}_\gamma)$ an element of this Ω, set $x_\omega = x_{\underline{\delta}_{\underline{\gamma}}}$ (i.e., consider the net associated with $\Delta_{\underline{\gamma}}$, and see where this sends the element $\underline{\delta}_{\underline{\gamma}}$ of $\Delta_{\underline{\gamma}}$). We thus have a mapping from the directed set Ω to X, and hence a net in X. We claim that this net in

X is in fact Cauchy. [Proof: Given entourage A for X, choose B with $BB \subset A$. Choose $\underline{\gamma}$ such that the pair consisting of any two elements of the net $\{\bar{x}_\gamma\}$ with their γ beyond $\underline{\gamma}$ is in $\bar{B}$, and, for each $\gamma \geq \underline{\gamma}$, choose $\underline{\delta}_\gamma$ in Δ_γ such that the pair consisting of any two elements of this net beyond $\underline{\delta}_\gamma$ is in B. Set $\omega = (\underline{\gamma},\underline{\delta}_\gamma)$. Then, for $\omega' \geq \omega$ and $\omega'' \geq \omega$, we have $(x_{\omega'},x_{\omega''})$ in A.] Thus we have constructed a Cauchy net x_ω (ω in Ω) in X. Let $\bar{x}$ be the point of $\bar{X}$ of which this net is a representative. We claim, finally, that our original net, $\bar{x}_\gamma$ (γ in Γ) in $\bar{X}$ converges to this point $\bar{x}$. [Proof: Given entourage A for X, choose B with $BB \subset A$. Then choose γ in Γ with $(\bar{x}_{\gamma'},\bar{x}_{\gamma''})$ in $\bar{B}$ whenever $\gamma' \geq \gamma$ and $\gamma'' \geq \gamma$. Then $(\bar{x},\bar{x}_{\gamma'})$ is in $\bar{A}$ for $\gamma' \geq \gamma$. Thus $\bar{x}_\gamma$ (γ in Γ) converges to $\bar{x}$.]

We are now nearly done. Starting from a uniform space X, we have constructed a complete uniform space $\bar{X}$ and a uniformly continuous mapping $X \xrightarrow{\varphi} \bar{X}$. What we must show, finally, is that this setup satisfies the universal property of the definition of a completion. It is convenient, however, to postpone this final demonstration for a moment to make the following observation.

THEOREM 48. *Let X and Y be uniform spaces, and let $X \xrightarrow{\alpha} Y$ be uniformly continuous. Then, for any Cauchy net x_δ (δ in Δ) in X, the net $\alpha(x_\delta)$ (δ in Δ) in Y is also Cauchy.*

Proof. Let A be any entourage for Y, so $\tilde{\alpha}^{-1}[A]$ is an entourage for X. Then there exists a δ such that $(x_{\delta'},x_{\delta''})$ is in $\tilde{\alpha}^{-1}[A]$ whenever $\delta' \geq \delta$ and $\delta'' \geq \delta$. Hence $(\alpha(x_{\delta'}),\alpha(x_{\delta''}))$ is in A whenever $\delta' \geq \delta$ and $\delta'' \geq \delta$. Thus the net $\alpha(x_\delta)$ (δ in Δ) in Y is Cauchy. □

Now let Y be any complete uniform space, and let $X \xrightarrow{\alpha} Y$ be a uniformly continuous mapping. Let $\bar{x}$ be any point of $\bar{X}$, and let x_δ (δ in Δ) be any representative. Then $\alpha(x_\delta)$ (δ in Δ) is a Cauchy net in Y, whence, since Y is complete, it converges to some element y of Y. Set $\varsigma(\bar{x}) = y$ (noting that, since α is uniformly continuous, this specification of y is independent of representative of x). Thus we have a mapping $\bar{X} \xrightarrow{\varsigma} Y$. This ς is certainly uniformly continuous. (For, for A any entourage for Y, $\alpha^{-1}[A]$ is an entourage for X, whence $\alpha^{-1}[A]$ is an entourage for $\bar{X}$. But $\varsigma^{-1}[A] \subset \overline{\alpha^{-1}[A]}$.) Furthermore, this ς makes the diagram of figure 125 commute, for, for x a point of X, $\varphi(x)$ is the point of X having as representative the constant (at x) net, whence $\varsigma(\varphi(x))$ is just $\alpha(x)$. Finally, this $\bar{X} \xrightarrow{\varsigma} Y$ is the unique uniformly continuous mapping that makes the diagram commute, for any other ς' must agree with ς on elements of $\bar{X}$ of the form $\varphi(x)$ (in order that the diagram commute), while, since the closure of the subset of $\bar{X}$ consisting of elements of this form is all of $\bar{X}$, this fact and continuity of ς determine ς uniquely.

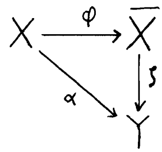

Figure 125

We have now, finally, completed the proof of the following.

THEOREM 49. *Every uniform space possesses a completion.*

One had to choose the completion, $\overline{X}$, to be "large enough to have points for every Cauchy net in X to converge to, but not so large that it contains extra, unrelated points which can destroy the uniqueness of ς in the universal property." The remarkable thing about all this is that the seeds of the original construction and the many verifications were already in the original, very simple, universal definition of a completion as a free object. One did not have to proceed in an aimless way: one knew what one wanted, and needed only a bit of optimism to feel confident that one would get it. Fortunately, the proof of theorem 49 has to be carried out only once: the universal definition of the completion (and the existence of one) is all that one ever will need to know about it.

Finally, we remark that the more familiar notion of the completion of a metric space reduces to that of taking its underlying uniform space and taking the completion of that via theorem 49.

Exercise 259. Show that uniform space X can be regarded as a uniform subspace of its completion $\overline{X}$. Show that the closure of this subset is $\overline{X}$. Is X always an open subset of $\overline{X}$?

Exercise 260. Prove that the completion of the rationals (as a uniform space, with the obvious uniformity) is uniform-isomorphic with the uniform space of reals.

Exercise 261. Let X and Y be uniform spaces. Show that $\overline{X \times Y}$ is naturally isomorphic, as a uniform space, to $\overline{X} \times \overline{Y}$ (where "$\times$" is direct product in the category of uniform spaces).

Exercise 262. Prove that the completion of a connected uniform space is connected.

Exercise 263. Is it true that a uniform space is complete if and only if each of its connected components (as a uniform subspace) is?

Exercise 264. Prove that every discrete uniform space is complete.

Exercise 265. Prove that the image of a complete uniform space, under a uniformly continuous mapping, is complete. (Compare, theorem 38.)

Exercise 266. Let X, $d(\ ,\ )$ be a metric space in which every Cauchy sequence converges. Prove that every Cauchy net in the underlying uniform space converges.

Exercise 267. Do there always exist free objects via the forgetful functor from the category of compact, Hausdorff topological spaces to the category of Hausdorff topological spaces?

Exercise 268. Show that, in the category of complete uniform spaces, the subobjects are closed subsets.

Exercise 269. Is it true that, if a mapping from one uniform space to another takes Cauchy nets to Cauchy nets, then it is uniformly continuous? (Compare, theorem 34.)

Exercise 270. Prove that a compact topological space can have only a finite number of connected components. Can a complete uniform space have more than a finite number of connected components?

Exercise 271. Let X be a set. Discuss the structure of the set of complete uniformities on X in the partially ordered set of all uniformities on X (exercise 255).

Exercise 272. Let X be a topological space, and x_δ (δ in Δ) a net in X. Call this net Cauchy if, for any neighborhood N of the diagonal in $X \times X$, there is a δ with $(x_{\delta'}, x_{\delta''})$ in N whenever $\delta' \geq \delta$ and $\delta'' \geq \delta$. What is wrong with this definition?

Exercise 273. Let A be a uniform subspace of the uniform space of reals, and suppose there exists an isomorphism from uniform space A to uniform space $\mathbf{R}$. Prove that $A = \mathbf{R}$.

Topological Groups

We now begin our study of structures that "mix topology and algebra." The importance of such structures is twofold. First, it is a common situation in physical applications that one winds up with a set on which there is a "notion of closeness" together with "the availability of algebraic combinations of elements of that set." Thus a topological group (and, perhaps even more commonly, a topological vector space) is often just the mathematical structure that is needed to describe what is going on physically. Second, and even more important, a brief look at a few "mixed structures" enables one to define, and obtain the properties of, other such structures when they are needed. A good fraction of the mathematical structures one needs in physical applications must be invented on the spot to suit the problem at hand.

Let G be a group. We write $G \times G \xrightarrow{\pi} G$ for the mapping of sets (where the Cartesian product of sets appears on the left) that sends the element (g,g') of $G \times G$ to the element gg' of G, and $G \xrightarrow{\iota} G$ for the mapping of sets that sends the element g of G to the element g^{-1}. A ***topological group*** consists of three things—i) a set G, ii) a rule which assigns, given any two elements of G, a third, and iii) a collection of subsets of the set G—subject to the following three conditions:

1. The set G, with the product rule ii), is a group.

2. The set G, with the collection iii) of subsets of G (as the open sets), is a Hausdorff topological space.

3. The mappings $G \times G \xrightarrow{\pi} G$ and $G \xrightarrow{\iota} G$ of topological spaces (direct product of topological spaces on the left in the first formula) are continuous.

That is, a topological group is "both a group and a topological space, where these two structures interact by the condition that the group operations be continuous."

Example. Let G be any group, and place on the set G the discrete topology. The first condition is immediate, while the second follows from the fact that any discrete topological space is Hausdorff. For the third condition, note that each mapping $G \times G \xrightarrow{\pi} G$ and $G \xrightarrow{\iota} G$ is from a discrete topological space (since, for the first, the direct product of two discrete spaces is discrete), and hence is necessarily continuous. We thus have a topological group.

Example. Let the set be the set of real numbers, let the product rule be addition of numbers, and let the subsets of R be the open subsets of the real line. The first two conditions are immediate. For the third, note that, for (r,r') any open interval in R, $\pi^{-1}[(r,r')]$ is the subset of $\mathbf{R} \times \mathbf{R}$ consisting of all (a,b) with $a + b$ in (r,r')—certainly an open subset of the topological plane—while $\iota^{-1}[(r,r')] = (-r',-r)$ is also open. Thus π and ι are continuous. We have a topological group, called the *topological group of reals.*

A number of properties of topological groups follow immediately from the definition. The continuity of $G \xrightarrow{\iota} G$ can be stated in terms of neighborhoods as follows: for g an element of G and N a neighborhood of g, the subset of G consisting of all elements g' with g'^{-1} in N is a neighborhood of g^{-1}. Continuity of $G \times G \xrightarrow{\pi} G$ can be restated as follows: given any neighborhood M of gg', there are neighborhoods N of g and N' of g' such that, whenever $\underline{g}$ is in N and $\underline{g}'$ is in N', $\underline{g}\underline{g}'$ is in M. (Intuitively, "if you choose an element close enough to g and another close enough to g', their product can be made as close as you wish to gg'.") Next, fix an element $\underline{g}$ of G, and denote by $\varphi_{\underline{g}}$ the mapping from G to G that sends g to $\underline{g}g$. This mapping is continuous. [Proof: The mapping from G to $G \times G$ which sends g to $(\underline{g},g)$ is continuous by the universal definition of the direct product of topological spaces, while π composed with this mapping is just $\varphi_{\underline{g}}$.] Similarly, $\varphi_{\underline{g}^{-1}}$ is continuous. But $\varphi_{\underline{g}} \circ \varphi_{\underline{g}^{-1}}$ and $\varphi_{\underline{g}^{-1}} \circ \varphi_{\underline{g}}$ are both the identity mapping of G. Thus each $\varphi_{\underline{g}}$ is an isomorphism of topological spaces.

Adjectives for groups and adjectives for topological spaces are applied to topological groups and refer to the appropriate structure. Thus an abelian topological group means that the underlying group is abelian; a compact topological group means that the underlying topological space is compact.

As an example of the interaction between the algebraic and topological structures of a topological group, we prove

THEOREM 50. *Let G be a topological group, and denote by C the connected component of G containing the identity e. Then C is a normal subgroup of G.*

Proof. Since $G \xrightarrow{\iota} G$ is an isomorphism of topological spaces, $\iota[C]$ is connected. But $\iota[C]$ intersects C (namely at e), so $\iota[C]$ is a subset of C. Thus the inverse of any element of C is in C. Let g and g' be elements of C. Then $\varphi_g[C]$ is connected (since φ_g is continuous). But $\varphi_g[C]$ intersects C (namely at g), so $\varphi_g[C] \subset C$. But $\varphi_g[C]$ contains gg', so gg' is in C. Thus C is a subgroup of G. Let g be in C, and let g' be any element of G. Then the mapping from G to G which sends $\underline{g}$ to $g'\underline{g}g'^{-1}$ is continuous, and hence takes C to a connected set. But this connected set intersects C (namely at e), and so is a subset of C. Thus $g'gg'^{-1}$ is in C. That is, C is a normal subgroup. □

Let G and H be topological groups, and let $G \xrightarrow{\varphi} H$ be a mapping of sets.

This $G \xrightarrow{\varphi} H$ is called a ***continuous homomorphism*** if it is both a continuous mapping of topological spaces and a homomorphism of groups. The composition ot two continuous homomorphisms is another (immediately from this fact for continuous mappings and for homomorphisms).

Let the objects be topological groups, the morphisms continuous homomorphisms of topological groups, and composition composition of continuous homomorphisms. We thus obtain a category, the ***category of topological groups.***

What are the forgetful functors from the category of topological groups? There are, of course, two obvious ones: one to the category of groups and one to the category of topological spaces. What we now wish to show is that this last functor can be strengthened somewhat: there is actually a functor from the category of topological groups to the category of uniform spaces. This is exactly what one would have expected. A uniform space is "a topological space in which neighborhoods can be compared with regard to size," while, in a topological group, the action of the group (i.e., the φ_g) yields "topology-preserving motions on the group, which should permit one to bring neighborhoods close to each other for comparison."

Let G be a topological group. For N any neighborhood of the identity e of G, let A_N denote the subset of $G \times G$ consisting of all pairs (g,g') of elements of G with gg'^{-1} in N. Let the entourages be the supersets of the A_N. We claim that these entourages define a uniformity on the set G. [Proof: (1) Each (g,g) is in every A_N, since $gg^{-1} = e$ is in N. For $g \neq g'$, (g,g') is not in A_N for N a neighborhood of e excluding gg'^{-1} (the existence of which is guaranteed by Hausdorffness). (2) For N any neighborhood of e, the subset N' of G consisting of inverses of elements of N is also a neighborhood of e (by continuity of ι). But $A_N^{-1} = A_{N'}$. So inverses of entourages are entourages. (3) Let N be any neighborhood of e, and choose neighborhood M of e such that, whenever g_1 and g_2 are in M, g_1g_2 is in N. (The existence of such an M is guaranteed by continuity of π.) Then, for (g,g') in A_M and (g',g'') in A_M (so gg'^{-1} and $g'g''^{-1}$ are in M), (g,g'') is in A_N (for $gg''^{-1} = (gg'^{-1})(g'g''^{-1})$). That is, $A_MA_M \subset A_N$. (4) That the intersection of two entourages is another is immediate from $A_N \cap A_{N'} = A_{N \cap N'}$. That the superset of any entourage is another is obvious.] Thus, given any topological group G, we obtain a uniformity on the set G, and hence a uniform space, which we denote by $\mathbf{F}(G)$. Note that, from the way we defined the entourages, the underlying topology from this uniformity on set G is just the given (by the original topological group) topology on G.

Next, let $G \xrightarrow{\varphi} H$ be a continuous homomorphism of topological groups. We claim that this mapping φ is also a uniformly continuous mapping of the corresponding uniform spaces. Indeed, let N be a neighborhood of e in H, and A_N the corresponding entourage for H. Then $\tilde{\varphi}^{-1}[A_N]$ consists of all elements

(g,g') of $G \times G$ with $(\varphi(g),\varphi(g'))$ in A_N, that is, all (g,g') with $\varphi(g)[\varphi(g')]^{-1}$ in N, that is, all (g,g') with $\varphi(gg'^{-1})$ in N, that is, all (g,g') with gg'^{-1} in $\varphi^{-1}[N]$. That is, $\tilde{\varphi}^{-1}[A_N] = A_{\varphi^{-1}[N]}$. Uniform continuity is now immediate from the fact that, since $G \xrightarrow{\varphi} H$ is continuous, the inverse image by φ of any neighborhood of e in H is a neighborhood of e in G.

Thus, for G a topological group, let $\mathbf{F}(G)$ be the uniform space obtained above, and, for $G \xrightarrow{\varphi} H$ a continuous homomorphism of topological groups, let $\mathbf{F}(G) \xrightarrow{\mathbf{F}(\varphi)} \mathbf{F}(H)$ be the uniformly continuous mapping obtained above. It is immediate that we thus obtain a covariant functor from the category of topological groups to the category of uniform spaces. We shall regard this functor as forgetful. (You "remember the topology of G, and remember the group structure long enough to turn this topology into a uniformity, but then forget the group structure.")

We may now apply adjectives for uniform spaces to topological groups: they refer to the underlying uniformity. In particular, a *complete topological group* is a topological group whose underlying uniformity is complete. We also have, among others, the *category of complete topological groups.*

One free construction in this business is of particular interest. Let G be a topological group. A *completion* of G consists of a complete topological group $\overline{G}$, together with a continuous homomorphism $G \xrightarrow{\varphi} G$, such that, given any complete topological group H and continuous homomorphism $G \xrightarrow{\alpha} H$, there is a unique continuous homomorphism $\overline{G} \xrightarrow{\varsigma} H$ such that the diagram of figure 126 commutes. We wish to show that every topological group possesses a completion.

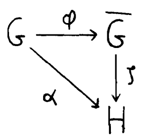

Figure 126

Choose, for the set $\overline{G}$, the completion of G as a uniform space. We next wish to make this set into a group. Let $\overline{g}$ and $\overline{g}'$ be two elements of $\overline{G}$ (so each is an equivalence class of Cauchy nets in G). Let g_δ (δ in Δ) and g_γ (γ in Γ) be representatives. Let Ω be the set $\Delta \times \Gamma$, and, for (δ,γ) and (δ',γ') in this Ω, write $(\delta,\gamma) \leq (\delta',\gamma')$ if $\delta \leq \delta'$ and $\gamma \leq \gamma'$. We thus obtain a directed set. For $\omega = (\delta,\gamma)$ an element of this Ω, let g_ω be the element $g_\delta g_\gamma$ of G. This is a net in G. We claim that it is Cauchy. [Proof: Let M be a neighborhood of e in G, and choose neighborhood N with $NN \subset M$. Then set $\omega =$

(δ,γ), where δ is such that $(g_{\delta'},g_{\delta''})$ is in A_N for $\delta' \geq \delta$ and $\delta'' \geq \delta$, and where γ is such that $(g_{\gamma'},g_{\gamma''})$ is in A_N for $\gamma' \geq \gamma$ and $\gamma'' \geq \gamma$. It follows that, for $\omega' \geq \omega$ and $\omega'' \geq \omega$, $(g_{\omega'},g_{\omega''})$ is in A_M.] We now define $\overline{gg}'$ to be the element of $\overline{G}$ of which this Cauchy net is a representative (noting that the resulting element of $\overline{G}$ is independent of the choices of representatives for $\overline{g}$ and $\overline{g}'$). This set $\overline{G}$, with this product structure, is a group. [Proof: (1) Associativity is immediate from $(g_\delta g_\gamma)g_\kappa = g_\delta(g_\gamma g_\kappa)$. (2) The identity is the constant net at e. (3) The inverse of the element of $\overline{G}$ with representative g_δ (δ in Δ) is the element with representative g_δ^{-1} (δ in Δ).] Thus we have a group $\overline{G}$. For the topology on $\overline{G}$, we choose that which comes from the uniformity on $\overline{G}$ (since, after all, $\overline{G}$ is the completion of uniform space G). Thus this $\overline{G}$ is a complete topological group.

Next, let $G \xrightarrow{\varphi} \overline{G}$ be the mapping which sends g in G to the element of $\overline{G}$ having as representative the constant net at g. This is a continuous homomorphism of topological group. [Proof: It is continuous because $G \xrightarrow{\varphi} \overline{G}$ is just the completion of the uniform space G. It is a homomorphism because, for g and g' in G, $\varphi(g)\varphi(g')$ has as representative the constant net at gg', i.e., is $\varphi(gg')$.]

Finally, we check that this $G \xrightarrow{\varphi} \overline{G}$ satisfies the universal property which defines the completion of a topological group. Let H be any complete topological group, and $G \xrightarrow{\alpha} H$ any continuous homomorphism. Then there is certainly a unique uniformly continuous mapping $\overline{G} \xrightarrow{\varsigma} H$ that makes the diagram commute (since $\overline{G}$ is the completion of G as a uniform space). Hence all that we must check is that this ς is a homomorphism of groups. Let $\overline{g}$ and $\overline{g}'$ be elements of $\overline{G}$, and let g_δ (δ in Δ) and g_γ (γ in Γ) be representatives. Then $\varsigma(\overline{g})$ is the element h of H to which the Cauchy net $\alpha(g_\delta)$ (δ in Δ) in H converges, and, similarly, $\varphi(\overline{g}') = h'$. But $\varphi(\overline{gg}')$ is the element of H to which the Cauchy net $\alpha(g_\delta g_\gamma)$ in H converges, and this element is clearly hh'. Thus $\overline{G} \xrightarrow{\varsigma} H$ is a homomorphism.

Thus every topological group possesses a (necessarily unique) completion.

Exercise 274. Let G be a topological group. Prove that the mappings $G \times G \xrightarrow{\pi} G$ and $G \xrightarrow{\iota} G$ are uniformly continuous.

Exercise 275. Prove that the third condition for a topological group can be replaced by the following: the mapping from $G \times G$ to G that sends (g,g') to gg'^{-1} is continuous.

***Exercise* 276. Show that the completion of the topological group of rationals is the topological group of reals.**

Exercise 277. Prove that both direct products and direct sums exist in the category of topological groups.

Exercise 278. Prove that the completion of an abelian topological group is abelian.

Exercise 279. Let G be a topological group, and let H be a normal subgroup of G that is closed as a subset of G. Show that the quotient group G/H, endowed with the quotient topology, is a topological group. What can one say about this quotient when H is the connected component of G containing the identity?

Exercise 280. Let the group be the additive group C of complex numbers, and let the topology be that given by the metric $d(a, a') = |a - a'|$. Show that this is a topological group.

Exercise 281. Let X be a Hausdorff topological space, and let G be the group of all isomorphisms from X to X, with the compact-open topology. Prove that G is a topological group. When is G complete?

Exercise 282. Find an example of a uniform space that is not the underlying uniform space of any topological group.

Exercise 283. Are there free objects via the forgetful functor from the category of topological groups to the category of topological spaces?

41

Topological Vector Spaces

We now consider our second of two examples of structures that mix topology and algebra. Topological vector spaces are somewhat more useful, a bit more complicated, and considerably richer in structure than topological groups. We regard the previous chapter as just an introduction to the present one.

Let V be a real vector space. We write $V \times V \xrightarrow{\sigma} V$, $\mathbf{R} \times V \xrightarrow{\pi} V$, and $V \xrightarrow{\iota} V$ for the mappings with action $\sigma(v,v') = v + v'$, $\pi(a,v) = av$, and $\iota(v) = -v$, respectively. A (real) ***topological vector space*** consists of four things—i) a set V, ii) a rule that assigns, given two elements of V, a third, iii) a rule that assigns, given a real number and an element of V, an element of V, and iv) a collection of subsets of V—subject to the following three conditions:

1. The set V, with the rules ii) and iii), is a real vector space.

2. The set V, with the collection iv) of subsets, is a Hausdorff topological space.

3. The mappings $V \times V \xrightarrow{\sigma} V$, $\mathbf{R} \times V \xrightarrow{\pi} V$, and $V \xrightarrow{\iota} V$ are all continuous (where $\mathbf{R}$ is assigned the topology of the real line).

Example. Let V be any real vector space (of dimension greater than zero), and assign V the indiscrete topology. The first and third conditions above are satisfied, but the second is not, since V is not Hausdorff. We do not obtain a topological vector space.

Example. Let V be any real vector space (of dimension greater than zero), and assign V the discrete topology. The first and second conditions are immediate, while $V \times V \xrightarrow{\sigma} V$ and $V \xrightarrow{\iota} V$ are continuous (since the left sides are discrete). However, $\mathbf{R} \times V \xrightarrow{\pi} V$ is not continuous. We have $0v = \vec{0}$ for some nonzero v in V (where "$\vec{0}$" denotes that vector). Let M be the neighborhood of $\vec{0}$ in V consisting only of this vector (noting that M is indeed a neighborhood, since V is discrete). Do there exist neighborhoods N of 0 in $\mathbf{R}$, and N' of v in V, such that av' is in M whenever a is in N and v' is in N'? The answer is no, for, since N is a neighborhood of 0 in $\mathbf{R}$, it contains a nonzero number ϵ, while v itself must be in N', and ϵv is not in M (for it is not the zero vector). Thus we do not obtain a topological vector space.

These examples illustrate the point that one must be at least a little subtle in one's choice of topology if one wishes to obtain a topological vector space. Since the topology of the real line enters explicitly into the definition of a topological vector space, the topology on V must be "reminiscent of that of the real line."

Example. Let S be any set, and denote by V the collection of all bounded, real-valued functions on S. Add such functions, and multiply them by real numbers, pointwise. We thus obtain a vector space V. For f and f' two such functions, denote by $d(f,f')$ the upper bound of the function $|f - f'|$ on S. This is a metric on V. We claim that this vector space V, with the topology induced by this metric, is a topological vector space. The first two conditions are immediate. For continuity of σ, let M be the neighborhood of $f + f'$ consisting of functions $\tilde{f}$ with $d(f + f',\tilde{f}) < \epsilon$. Choose N and N' the neighborhoods of f and f', respectively, consisting of $\tilde{f}$ with $d(f,\tilde{f}) < \epsilon/2$ and $d(f',\tilde{f}) < \epsilon/2$, respectively. Then, evidently, the sum of any function in N with any in N' is in M. For continuity of π, let M be the neighborhood of af consisting of functions $\tilde{f}$ with $d(af,\tilde{f}) < \epsilon$. Denote by m an upper bound for f. Let N be the neighborhood of a in $\mathbf{R}$ consisting of $\tilde{a}$ with $|a - \tilde{a}| < \epsilon/2m$, and let N' be the neighborhood of f in V consisting of $\tilde{f}$ with $d(f,\tilde{f}) < \epsilon/(2a + \epsilon/m)$. Then, for $\tilde{a}$ in N and $\tilde{f}$ in N', $\tilde{a}\tilde{f}$ is in M. Continuity of ι is obvious. Thus we obtain a topological vector space.

Example. Let V again be the vector space above, but choose the following topology. A subset of V is open if, for every f in that subset, there is a finite collection, $s_1, \ldots, s_n$, of points of S, and a finite collection, $\epsilon_1, \ldots, \epsilon_n$, of positive numbers, such that every $\tilde{f}$ with $|f(s_1) - \tilde{f}(s_1)| < \epsilon_1, \ldots,$ and $|f(s_n) - \tilde{f}(s_n)| < \epsilon_n$ is also in that subset. Noting that this is a Hausdorff topology on V, the first two conditions for a topological vector space are satisfied. We verify the third condition. Let M be the neighborhood of $f + f'$ defined by $(s_1, \ldots, s_n; \epsilon_1, \ldots, \epsilon_n)$. Choose N and N' the neighborhoods of f and f' defined by $(s_1, \ldots, s_n; \epsilon_1/2, \ldots, \epsilon_n/2)$. Then, for $\tilde{f}$ and $\tilde{f}'$ in N and N', respectively, $\tilde{f} + \tilde{f}'$ is in M. Hence σ is continuous. Let M be the neighborhood of af defined by $(s_1, \ldots, s_n; \epsilon_1, \ldots, \epsilon_n)$. Let ϵ be the smallest of $\epsilon_1, \ldots, \epsilon_n$, and let m be an upper bound for f. Let N be the neighborhood of a in $\mathbf{R}$ consisting of $\tilde{a}$ with $|a - \tilde{a}| < \epsilon/2m$, and let N' be the neighborhood of f in V defined by $(s_1, \ldots, s_n; \epsilon/(2a + \epsilon/m), \ldots, \epsilon/(2a + \epsilon/m))$. Then, for $\tilde{a}$ in N and $\tilde{f}$ in N', $\tilde{a}\tilde{f}$ is in M. Continuity of ι is obvious. Thus we obtain a topological vector space. Note that the topology on V in this example is strictly coarser than that in the previous example (provided S is infinite: the two topologies coincide for finite S).

Example. Let V again be the vector space above, but now choose the following topology. A subset of V is open if, for every f in that subset, there is a positive real-valued function ϵ on S such that every $\tilde{f}$ with $|f(s) - \tilde{f}(s)| < \epsilon(s)$ for all s is also in that subset. (Thus, if we had only admitted constant functions $\epsilon(s)$, this topology would be just that of the earlier example.) Provided only that S has an infinite number of elements, we do not obtain a topological vector space. The problem, as one would expect, is continuity of π. Let f_0 be the function with $f_0(s) = 1$ for every s in S. Then $1\, f_0 = f_0$. Let M be the neighborhood of f_0 defined by some positive function ϵ which assumes values arbitrarily close to zero. (For example, if S is the set of positive integers, we might choose for the function ϵ, $\epsilon(1) = 1$, $\epsilon(2) = 1/2$, $\epsilon(3) = 1/4, \cdots$) Any neighborhood N of 1 in $\mathbf{R}$ contains a number a greater than 1; any neighborhood N' of f_0 in V contains f_0. But af_0 is not in our neighborhood M (for af_0 is the function with constant value a, while the function ϵ assumes arbitrarily small values). For the case of S infinite, this topology on V is strictly finer than that of the third example of this chapter; for the case of S finite, the two topologies coincide.

We conclude from these examples that "the crucial question is normally continuity of $\mathbf{R} \times V \xrightarrow{\pi} V$. Normally, the coarser the topology on V, the more likely one has a topological vector space, provided only that the topology is not so coarse that Hausdorffness is destroyed. Very fine topologies tend not to give a topological vector space."

Let us examine the case of S finite in the examples above in more detail. Let S consist of n elements, $s_1, \ldots, s_n$. Let v_1 be the function (element of V) whose value at s_1 is 1, and zero elsewhere, and similarly for $v_2, v_3, \ldots, v_n$. Then, evidently, every element of V can be written in one and only one way as a linear combination of $v_1, \ldots, v_n$ (the value of f at s_i is the coefficient of v_i in this expansion of f). That is, $v_1, \ldots, v_n$ is a basis for V, so V is n-dimensional. The above topology on V is the following. To obtain a neighborhood of $f = a_1v_1 + \cdots + a_nv_n$, choose a positive number ϵ, and consider all $\tilde{f} = \tilde{a}_1v_1 + \cdots + \tilde{a}_nv_n$ with each $\tilde{a}_i$ within ϵ of a_i. Thus V, as a topological space, is isomorphic with the topological space that is the direct product of the real line with itself n times. We shall call this topology on V, in the finite-dimensional case, the *Euclidean topology*. (Note that the Euclidean topology is independent of the choice of basis.)

It is not just a coincidence that the various topologies on V reduce to the same topology—the Euclidean one—in the finite-dimensional case. In fact, we have

THEOREM 51. *Let V be a finite-dimensional topological vector space. Then the topology on V is the Euclidean topology.*

Proof. Let N be a neighborhood of $\vec{0}$ in V. Fix vector $\underline{v}$ in V. Then, by continuity of $V \times V \xrightarrow{\sigma} V$, the set N' consisting of all elements of V of the form $v + \underline{v}$ with v in N is a neighborhood of $\underline{v}$. By continuity of $\mathbf{R} \times V \xrightarrow{\pi} V$, there is a positive number ϵ such that $a\underline{v}$ is in N' whenever $1 - \epsilon < a < 1 + \epsilon$. Thus $a\underline{v}$ is in N whenever $-\epsilon < a < \epsilon$. Next, note that, by continuity of $\mathbf{R} \times V \xrightarrow{\pi} V$, there is a neighborhood M of $\vec{0}$ in V such that av is in N whenever v is in M and $|a| < 1$. Hence the collection K of all elements of V of the form av with v in M and $|a| < 1$ is a subset of N and a neighborhood of $\vec{0}$

(since K is a superset of neighborhood M). For purposes of this proof, call a subset K of V star-shaped if i) for every v in K and $|a| < 1$, av is in K, and ii) for every vector in V, some nonzero multiple of it is in K. We have just shown above that every neighborhood N of $\vec{0}$ is a superset of some star-shaped neighborhood (namely, the K constructed above). Next, note that, for K star-shaped, $K + K$ (i.e., the set of all $v + v'$ with v and v' in K) is a Euclidean neighborhood of $\vec{0}$. But, by continuity of $V \times V \xrightarrow{\sigma} V$, every neighborhood of $\vec{0}$ in the given topology on V is a superset of some $K + K$ with K star-shaped. Hence every neighborhood of $\vec{0}$ (and hence of any element of V) in the given topology is also a neighborhood in the Euclidean topology: the given topology is at least as coarse as the Euclidean one. We now prove the reverse. Choose basis $v_1, \ldots, v_n$ for V, and denote by S_ϵ the subset of V consisting of $a_1v_1 + \cdots + a_nv_n$ with $(a_1)^2 + \cdots + (a_n)^2 \leq \epsilon$, and by C_ϵ the subset consisting of those with $(a_1)^2 + \cdots + (a_n)^2 = \epsilon$. This C_ϵ is compact in the Euclidean topology, and hence (since the given topology is at least as coarse) is also compact in the given topology. By Hausdorffness of the given topology, there is a neighborhood of $\vec{0}$ in this topology which does not intersect C_ϵ: hence there is a star-shaped one, K. Since this K is star-shaped and does not intersect C_ϵ, K is a subset of S_ϵ. But every neighborhood of $\vec{0}$ in the Euclidean topology contains some S_ϵ—and hence some neighborhood of $\vec{0}$ in the given topology. We conclude that the given topology is at least as fine as the Euclidean one. Thus the topologies coincide. ☐

Thus, for a finite-dimensional vector space, there is no choice of what topology one should choose if one wishes to obtain a topological vector space. Our earlier examples show that this is not true in the infinite-dimensional case.

We are now, finally, in a position to give an intuitive picture of the topology of a topological vector space. Let V be a topological vector space. Theorem 51 shows that every one-dimensional subspace of vector space V (i.e., every "ray") is, as a subspace of topological space V, isomorphic to the real line. The question is "how these different rays are tied to each other topologically." If we "turn a ray through a two-dimensional subspace of V, the topological behavior is just that of the Euclidean case" (since the two-dimensional subspace of V, by theorem 51, has the Euclidean topology). The real question therefore is "how the different rays are tied to each other topologically when they have an infinite number of directions to turn in." Now some freedom becomes available. Suppose we fix an open neighborhood N of $\vec{0}$. This N cuts each ray in a Euclidean-open set. The "set varies as we move the ray about, never becoming zero, since it must be Euclidean-open." When an infinite number of dimensions are available, however, "the open set has a genuine choice as to how small it should become on turning in the various directions." This "choice" represents the possibility for choosing various topologies on V, in the infinite-dimensional case, to obtain a topological vector space.

Next, let V and W be topological vector spaces. A mapping of sets, $V \xrightarrow{\varphi} W$, is called a *continuous linear mapping* if φ is a linear mapping of vector spaces and also a continuous mapping of topological spaces. We see from the examples of this chapter that, at least in the infinite-dimensional case, linearity of φ is not alone sufficient to imply continuity. (In the finite-dimensional case, by contrast, linearity of $V \xrightarrow{\varphi} W$ already implies continuity, for, in this case, V and W must have the Euclidean topology.) Note that the composition of two continuous linear mappings is another.

Let the objects be real topological vector spaces, the morphisms continuous linear mappings of real topological vector spaces, and composition composition of continuous linear mappings. We thus obtain the *category of real topological vector spaces.*

We have immediately two forgetful functors from this category. The first is to the category of abelian topological groups (i.e., we forget how to multiply vectors by numbers, noting that the third condition for a topological vector space already requires that the group operations be continuous). The other forgetful functor is to the category of real vector spaces (i.e., we forget the topology). Other forgetful functors are obtained by composition of forgetful functors with these two. In particular (since a topological vector space defines an abelian topological group, which defines a uniform space), we have a forgetful functor from the category of topological vector spaces to the category of uniform spaces.

A real topological vector space V is said to be *complete* if its underlying uniform space is complete. For V a topological vector space, a *completion* of V is a complete topological vector space $\overline{V}$, together with a continuous linear mapping $V \xrightarrow{\varphi} \overline{V}$, such that the following property is satisfied: given any complete topological vector space W and continuous linear mapping $V \xrightarrow{\alpha} W$, there is a unique continuous linear mapping $\overline{V} \xrightarrow{\varsigma} W$ such that the diagram of figure 127 commutes. We note that this is a free construction (via the forgetful functor from the category of complete topological vector spaces to the category of topological vector spaces), and therefore that the completion is unique if one exists.

Example. The topological vector space of the third example of this chapter is complete. Let f_δ (δ in Δ) be a Cauchy net in V. Then, for each s in S, $f_\delta(s)$ (δ in Δ) is a Cauchy net in $\mathbf{R}$, the uniform space of reals. Since the latter is complete, this net converges to some real number, $f(s)$. It is easily checked that this element f of V is a limit point of our Cauchy net.

Example. The topological vector space of the fourth example of this chapter is also complete, by the argument sketched for the first example.

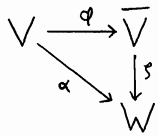

Figure 127

Example. Let V denote the set of all bounded, differentiable functions, $\mathbf{R} \xrightarrow{f} \mathbf{R}$. Add such functions, and multiply them by real numbers, pointwise, so we obtain a vector space. For f and f' two such functions, let $d(f,f')$ be the least upper bound of the function $|f - f'|$ on $\mathbf{R}$. This metric on V induces a topology, so we obtain a topological vector space. This topological vector space is not complete. The sequence of functions $f_1, f_2, \cdots$ illustrated in figure 128 is Cauchy. This Cauchy net has no limit in V (for the limit would "like to be" the function indicated; but this function is not differentiable, and hence does not represent a point of V). It is not difficult to show that the completion of V in this example is isomorphic with the topological vector space of all continuous, bounded functions $\mathbf{R} \xrightarrow{f} \mathbf{R}$.

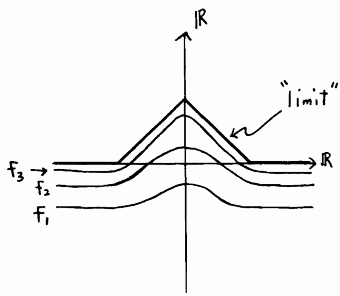

Figure 128

THEOREM 52. *Every topological vector space possesses a completion.*

Sketch of proof. Let V be a topological vector space. Choose for $\overline{V}$ the completion of V as a uniform space, so $\overline{V}$ is a uniform space, and we have $V \xrightarrow{\varphi} \overline{V}$. For $\overline{v}$ and $\overline{v}'$ elements of $\overline{V}$ (with representatives v_δ (δ in Δ) and v_γ (γ in Γ), let $\overline{v} + \overline{v}'$ be the element with representative v_ω (ω in Ω), where Ω is the directed set on $\Delta \times \Gamma$ and where, for $\omega = (\delta,\gamma)$ in Ω, $v_\omega = v_\delta + v_\gamma$. For $\overline{v}$ in $\overline{V}$ (with representative v_δ (δ in Δ)) and a a number, let $a\overline{v}$ be the element of $\overline{V}$ with representative av_δ (δ in Δ). Thus $\overline{V}$ is a complete topological vector space. One checks that it satisfies the universal property. □

A number of facts about the topology of a topological vector space follow immediately from theorem 51 (together with the observation that any vector subspace of a topological vector space, with the induced topology, is also a topological vector space). Thus every topological vector space is connected and simply connected. Only for dimension zero can a topological vector space be compact.

Finally, we remark that all of the discussion above goes through, without change, if "real" is everywhere replaced by "complex," where, for the topology on the set C of complex numbers, we choose that given by the metric $d(a,a') = |a - a'|$.

Exercise 284. A subset A of topological vector space V is said to be convex if, for any v and v' in A, and number a with $0 \leq a \leq 1$, the vector $av + (1 - a)v'$ is in A. Prove that the closure of a convex set is convex.

Exercise 285. A topological vector space V is said to be locally convex if every neighborhood of $\vec{0}$ contains a convex neighborhood. Which of our examples are locally convex?

Exercise 286. Show that the closure of a vector subspace of a topological vector space is again a vector subspace. Why is this neither true in general nor interesting when true for the interior?

Exercise 287. Prove that the homology groups of a topological vector space V re $H_0(V) = Z$, $H_1(V) = 0$, $H_2(V) = 0$, $\cdots$

Exercise 288. Define the direct sum of topological vector spaces, and prove the existence of one. For which topological vector spaces V is V isomorphic to $V + V$?

Exercise 289. Is there a reasonable notion of the tensor product of two topological vector spaces?

Exercise 290. Is it true that every linear mapping from topological vector space V to $\mathbf{R}$ is continuous?

Exercise 291. A subset B of a topological vector space is said to be bounded if, for every neighborhood N of $\vec{0}$, there is a positive number ϵ with $\epsilon B \subset N$. Let V and W be topological vector spaces, and consider Mor(V,W), the set of continuous linear mappings from V to W. Verify that this Mor(V,W) has the structure of a vector space. For B a bounded subset of V, and O an open subset of W, write $K(B,O)$ for the subset of Mor(V,W) consisting of all $V \xrightarrow{\varphi} W$ with $\varphi[B] \subset O$. Let the topology on Mor(V,W) be that generated by these

$K(B, O)$. (This is called the bounded-open topology on Mor(V, W).) Prove that Mor(V, W), with the bounded-open topology, is a topological vector space. Prove that Mor(V, W) is complete if V and W are. Find an example to show that the mapping Mor$(V, W) \times$ Mor$(W, Z) \to$ Mor(V, Z) given by composition is not in general continuous. What is a simple condition that implies continuity?

Categories: Summary

The various categories we have introduced are listed in figure 129. The forgetful functors are indicated by solid arrows, the various free constructions we have discussed by dashed arrows (there are, of course, many more), and the homology functors by the railroad tracks.

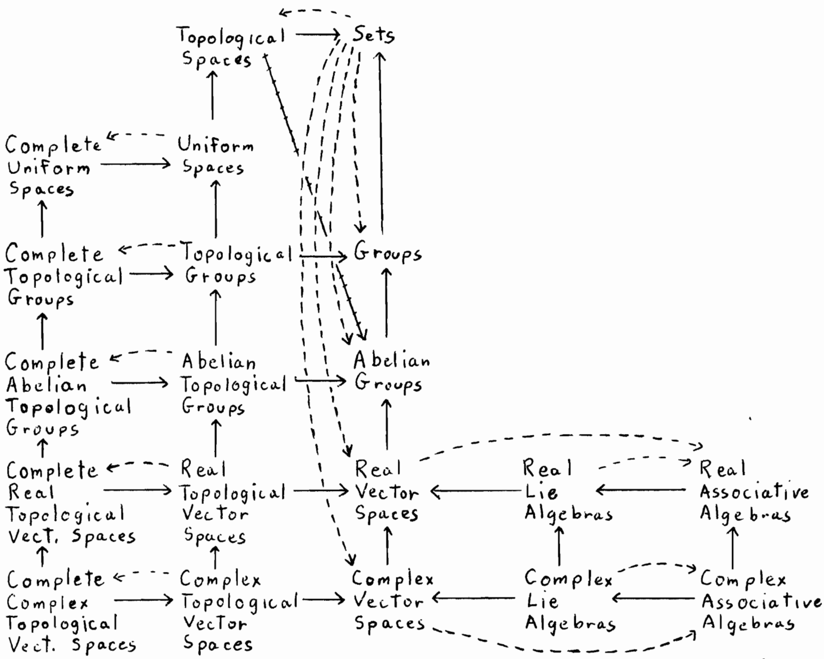

Figure 129

43

Measure Spaces

We now begin our discussion of the theory of integration. We proceed in two steps. In the first, we introduce the notion of measure, which appropriately generalizes that of "area" (e.g., of regions in the plane). In the second, we use this measure to define the notion of an integral.

We first introduce a little notation. Denote by R^* the set consisting of the non-negative real numbers, together with one additional element, which we write ∞. The element ∞ of R^* is called *infinite*, the others *finite*. For a and b elements of R^*, we write $a \leq b$ (or $b \geq a$) if either i) a and b are both finite, and $a \leq b$ as numbers, ii) a is finite and b is infinite, or iii) both are infinite. (Note that this is a partial ordering on R^*: the minimum element is 0, the maximum ∞.) For a and b elements of R^*, we write $a + b$ (called their *sum*) for the element of R^* which is the sum $a + b$ of numbers if both a and b are finite, and which is ∞ if either is infinite. Note that $a + b = b + a$, and $a + (b + c) = (a + b) + c$, but that we do not have a group, because there are no inverses. Finally, for $a_1, a_2, a_3, \cdots$ a sequence of elements of R^*, we write $\sum a_i$ (also called their *sum*) for the element of R^* which is the sum $\sum a_i$ of numbers if each of $a_1, a_2, \cdots$ is finite and this sum converges to some number, and which is ∞ otherwise. Note that $\sum a_i$, so defined, is independent of the order of $a_1, a_2, \cdots$

A *measure space* consists of three things—i) a set X, ii) a collection $\mathbf{M}$ of subsets of X (subsets of X in this collection are called *measurable* sets), and iii) a mapping of sets, $\mathbf{M} \xrightarrow{\mu} R^*$ (for A a measurable set, the element $\mu(A)$ of R^* is called the *measure* of A)—subject to the following four conditions:

1. The empty set $\emptyset$ is measurable, and $\mu(\emptyset) = 0$.
2. For A any measurable set, A^c is also measurable.
3. For $A_1, A_2, \cdots$ measurable sets, $\bigcup A_i$ is also measurable.
4. For $A_1, A_2, \cdots$ measurable and disjoint (i.e., no point of X is in more than one A_i), $\mu(\bigcup A_i) = \sum \mu(A_i)$.

One thinks of the measurable sets as those for which "the notion of area makes sense" and of the measure of such a set as its "area."

A number of facts follow immediately from the definition. The first and second conditions imply that X ($= \emptyset^c$) is measurable. The second and third imply that, for $A_1, A_2, \cdots$ measurable, so is $\bigcap A_i$ ($= (\bigcup A_i^c)^c$). The first and third imply that, for A and B measurable, so is $A \cup B$ (choosing $A_1 = A, A_2 = B, A_3 = \emptyset, A_4 = \emptyset, \cdots$), and similarly for any finite number of measurable sets, and similarly for intersections of a finite number of measurable sets. Hence, for A and B measurable, so is $A - B$ $(= A \cap B^c)$. The first and fourth conditions imply that, for A and B measurable and disjoint, $\mu(A \cup B) = \mu(A) + \mu(B)$ (choosing $A_1, A_2, \cdots$ as above and using $\mu(\emptyset) = 0$), and similarly for any finite number of measurable sets. Finally, for A and B measurable, with A a subset of B, we have $\mu(A) \leq \mu(B)$ (for B is the disjoint union of measurable sets A and $B - A$, whence $\mu(B) = \mu(A) + \mu(B - A)$).

Example. Let S be any set, and let the measurable sets be all subsets of S. Let $\mu(\emptyset) = 0$, and $\mu(A) = \infty$ for A any nonempty subset of S. That the four conditions above are satisfied is immediate, whence we obtain a measure space. (Alternatively, one could set $\mu(A) = 0$ for every subset of S, and obtain a measure space.)

Example. Let S be any set, and let the measurable sets be $\emptyset$ and S. Let $\mu(\emptyset) = 0$, and $\mu(S) = 17$ (or any other element of R^*). We obtain a measure space.

Example. Let the set X consist of the set of positive integers. Choose any sequence, $a_1, a_2, \cdots$ of elements of R^*. Let all subsets of X be measurable. For A a subset of X, set $\mu(A) = \sum_{i \text{ in } A} (a_i)$, where the sum on the right is over the integers i in A. Again, the four conditions for a measure space are immediate.

All of these examples are rather uninteresting (the last perhaps less so). Fortunately, there is a single example which plays a central role in this subject: It both motivates and illustrates the definition, and is by far the most useful for applications. We now give this example.

Let $X = \mathbf{R}$, the set of real numbers. Let A be any subset of $\mathbf{R}$. Consider a countable collection, $I_1 = (a_1, b_1), I_2 = (a_2, b_2), \ldots,$ of open intervals in R (where, for convenience, we allow the empty open interval, $(a,a) = \emptyset$) which covers A (i.e., which is such that every element of A is in at least one of the I_i). For $\mathbf{C}$ such a collection, we write $m(\mathbf{C}) = (b_1 - a_1) + (b_2 - a_2) + \cdots$ Since each term in this sum is a non-negative number, $m(\mathbf{C})$ is an element of R^*. Next, denote by $\bar{\mu}(A)$ the greatest lower bound (in R^*) of elements of R^* of the form $m(\mathbf{C})$, for $\mathbf{C}$ such a countable collection of open intervals which covers A. In this way we associate, with each subset A of $\mathbf{R}$, an element, $\bar{\mu}(A)$, of R^*.

Example. Let A be the open interval $(0,1)$. Let $\mathbf{C}$ be the collection $I_1 = (-3,7)$, $I_2 = (-1,1)$, $I_3 = (0,0)$, $I_4 = (0,0)$, $\cdots$ of open intervals, which covers A. Then $m(\mathbf{C}) = 12$. But we can find such instances of $\mathbf{C}$ with smaller $m(\mathbf{C})$. Let $\mathbf{C}'$ be the collection $I_1 = (0,1)$, $I_2 = (0,0)$, $\cdots$ Then $m(\mathbf{C}') = 1$. Clearly, there is no such $\mathbf{C}$ with $m(\mathbf{C})$ smaller than one. Hence $\bar{\mu}(A) = 1$.

Example. For $A = \mathbf{R}$, $\bar{\mu}(A) = \infty$. Let B consist of the elements $1,2,\cdots$ of R. Then $\bar{\mu}(B) = 0$. Indeed, given any positive number ϵ, let $\mathbf{C}$ be the collection $I_1 = (1-\epsilon, 1+\epsilon)$, $I_2 = (2-\epsilon/2, 2+\epsilon/2)$, $I_3 = (3-\epsilon/4, 3+\epsilon/4)$, $\cdots$ of open intervals, which covers B. Then $m(\mathbf{C}) = 2 + 2(\epsilon/2) + 2(\epsilon/4) + \cdots = 4\epsilon$. Hence $\bar{\mu}(B) \leq 4\epsilon$ for every positive ϵ, whence $\bar{\mu}(B) = 0$.

One might imagine that we could now obtain a measure space by letting the measurable sets all be subsets of $\mathbf{R}$ and letting the measure of such a subset A be $\bar{\mu}(A)$. However, as we shall see shortly, this does not work: one has to proceed in a slightly more subtle way. We first observe that this $\bar{\mu}$ satisfies the following three conditions (where, recall, we have set $X = \mathbf{R}$):

$\bar{1}$. $\bar{\mu}(\emptyset) = 0$.

$\bar{2}$. For A and B subsets of X, with A a subset of B, $\bar{\mu}(A) \leq \bar{\mu}(B)$.

$\bar{3}$. For $A_1, A_2, \cdots$ any subsets of X, $\bar{\mu}(\cup A_i) \leq \sum \bar{\mu}(A_i)$.

That the first condition is satisfied is obvious. For the second, note that, for $I_1, I_2, \cdots$ any collection of open intervals that covers B, this collection also covers A. But $\bar{\mu}(B)$ is the greatest lower bound of $m(\mathbf{C})$ for $\mathbf{C}$ covering B, while $\bar{\mu}(A)$ is the greatest lower bound for covering A. Hence $\bar{\mu}(A) \leq \bar{\mu}(B)$. Finally, we check the third condition. Choose any positive number ϵ. Let $\mathbf{C}^1$ be the collection $I_1^1, I_2^1, I_3^1, \cdots$ of open intervals which covers A_1, with $m(\mathbf{C}^1) \leq \bar{\mu}(A_1) + \epsilon$; let $\mathbf{C}^2$ be the collection $I_1^2, I_2^2, I_3^2, \cdots$ of open intervals which covers A_2, with $m(\mathbf{C}^2) = \bar{\mu}(A_2) + \epsilon/2$; let $\mathbf{C}^3$ be similar, with $m(\mathbf{C}^3) = \bar{\mu}(A_3) + \epsilon/4$, etc. (We can always find such, for, e.g., since $\bar{\mu}(A_1)$ is the greatest lower bound of $m(\mathbf{C})$ with $\mathbf{C}$ covering A_1, there is one such $\mathbf{C}$, which we call $\mathbf{C}^1$, with $m(\mathbf{C}^1)$ only ϵ greater than $\bar{\mu}(A_1)$.) Denote by $\mathbf{C}$ the collection of all these open intervals, I_i^j $(i,j = 1,2,\cdots)$. Then $m(\mathbf{C}) = m(\mathbf{C}^1) + m(\mathbf{C}^2) + m(\mathbf{C}^3) + \cdots \leq \bar{\mu}(A_1) + \epsilon + \bar{\mu}(A_2) + \epsilon/2 + \bar{\mu}(A_3) + \epsilon/4 + \cdots = \sum \bar{\mu}(A_i) + 2\epsilon$. But notice that $\mathbf{C}$ is a collection of open intervals which covers all of $\cup A_i$. Hence $\bar{\mu}(\cup A_i) \leq m(\mathbf{C})$. Thus we have $\bar{\mu}(\cup A_i) \leq \sum \bar{\mu}(A_i) + 2\epsilon$. But this must be true for every positive number ϵ, whence $\bar{\mu}(\cup A_i) \leq \sum \bar{\mu}(A_i)$.

Thus our $\bar{\mu}$ satisfies the three conditions listed above. (Note that these conditions are somewhat different from those for a measure space. In particular, the third condition above does not demand that the A_i be disjoint, but requires only an inequality rather than an equality.) The final step in the construction of a measure space is the use of the result below. Since this result has a number of other applications and since, in particular, we shall have to use it again later, we prove it in a context somewhat broader than needed for this particular example.

THEOREM 53. *Let X be any set, let $\mathbf{P}(X)$ be the collection of all subsets of X, and let $\mathbf{P}(X) \xrightarrow{\bar{\mu}} R^*$ be any mapping satisfying the three conditions just above. Denote by $\mathbf{M}$ the collection of all subsets A of X satisfying the* *following condition:* $\bar{\mu}(E) = \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E)$ *for every subset* E *of* X. *For* A *in* $\mathbf{M}$, *write* $\mu(A) = \bar{\mu}(A)$. *Then* X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ *is a measure space.*

Proof. We verify each of the four conditions for a measure space. (1) For E any subset of X, $\bar{\mu}(\emptyset \cap E) + \bar{\mu}(\emptyset^c \cap E) = \bar{\mu}(\emptyset) + \bar{\mu}(E) = \bar{\mu}(E)$. Thus $\emptyset$ is in $\mathbf{M}$. Also, $\mu(\emptyset) = \bar{\mu}(\emptyset) = 0$. (2) For A in $\mathbf{M}$, we have $\bar{\mu}(E) = \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E)$ for every subset E of X, whence A^c is in $\mathbf{M}$. (3) Suppose first that A and B are in $\mathbf{M}$. Then, for E any subset of X, we have $\bar{\mu}(E) = \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E) = \bar{\mu}(A \cap B \cap E) + \bar{\mu}(A \cap B^c \cap E) + \bar{\mu}(A^c \cap E)$, where we have used A in $\mathbf{M}$ in the first step, and B in $\mathbf{M}$ in the second. But

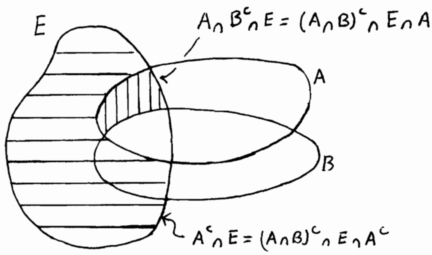

Figure 130

$A \cap B^c \cap E = (A \cap B)^c \cap E \cap A$, and $A^c \cap E = (A \cap B)^c \cap E \cap A^c$ (figure 130). Hence the above becomes $\bar{\mu}(E) = \bar{\mu}(A \cap B \cap E) + \bar{\mu}((A \cap B)^c \cap E \cap A) + \bar{\mu}((A \cap B)^c \cap E \cap A^c) = \bar{\mu}(A \cap B \cap E) + \bar{\mu}(A \cap B)^c \cap E)$, where we used A in $\mathbf{M}$ in the last step. We conclude that $A \cap B$ is also in $\mathbf{M}$, whence $A \cup B$ $(= (A^c \cap B^c)^c)$ is also in $\mathbf{M}$. Thus we have so far verified the third condition for a measure space for a finite number of sets. Next, let A and B be in $\mathbf{M}$, and let these sets be disjoint. Then $\bar{\mu}(E) = \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E)$ for every E, whence, replacing E by $(A \cup B) \cap E$, we obtain $\bar{\mu}((A \cup B) \cap E) = \bar{\mu}(A \cap E) + \bar{\mu}(B \cap E)$. Now let $A_1, A_2, A_3, \cdots$ be elements of $\mathbf{M}$. We might as well take them to be disjoint (for, were they not, A_1, $A_2 - A_1, A_3 - A_2 - A_1, \cdots$ would be disjoint elements of $\mathbf{M}$ with the same union). Then, for E any subset of X, and for any n, we have

$$\begin{aligned}
\bar{\mu}(E) &= \bar{\mu}((\bigcup_1^n A_i) \cap E) + \bar{\mu}((\bigcup_1^n A_i)^c \cap E) \\
&= \sum_1^n \bar{\mu}(A_i \cap E) + \bar{\mu}((\bigcup_1^n A_i)^c \cap E) \\
&\geq \sum_1^n \bar{\mu}(A_i \cap E) + \bar{\mu}((\bigcup_1^\infty A_i)^c \cap E) \\
&\geq \sum_1^\infty \bar{\mu}(A_i \cap E) + \bar{\mu}((\bigcup_1^\infty A_i)^c \cap E)
\end{aligned}$$

$$\geq \bar{\mu}((\bigcup_1^\infty A_i) \cap E) + \bar{\mu}((\bigcup_1^\infty A_i)^c \cap E)$$

$$\geq \bar{\mu}(E)$$

where, in the first step, we have used the fact that the union of any finite number of elements of **M** is in **M**; in the second, we have used the fact that, for A and B disjoint and in **M**, $\bar{\mu}((A \cup B) \cap E) = \bar{\mu}(A \cap E) + \bar{\mu}(B \cap E)$; in the third, we have used property $\bar{2}$ and the fact that $(\bigcup_1^n A_i)^c \supset (\bigcup_1^\infty A_i)^c$; in the fourth, the fact that the previous inequality holds for every n; in the fifth, property $\bar{3}$; and, in the sixth, property $\bar{3}$ again and the fact that $((\bigcup_1^\infty A_i) \cap E) \cup ((\bigcup_1^\infty A_i)^c \cap E) = E$. Since $\bar{\mu}(E)$ appears at each end of this string of inequalities, each "$\geq$" must be an "$=$." Since this must be true—in particular, in the last step—$(\bigcup_1^\infty A_i)$ is in **M**. (4) Let $A_1, A_2, \cdots$ be disjoint elements of **M**. In the string of inequalities above, let $E = (\bigcup_1^\infty A_i)$. In particular, the fourth line must equal $\bar{\mu}(E) = \bar{\mu}((\bigcup_1^\infty A_i))$. Thus we have $\mu(\bigcup A_i) = \sum \mu(A_i)$. □

Applying theorem 53 to our particular example, we obtain a measure space. This measure space is called the *measure space of reals* (or, more commonly, *Lebesque measure* on the reals).

Note how the measurable sets arise in theorem 53. For A and E subsets of X, $A \cap E$ and $A^c \cap E$ are two disjoint sets whose union is E itself. In order that A be measurable, we require that the sum of $\bar{\mu}(A \cap E)$ and $\bar{\mu}(A^c \cap E)$ be $\bar{\mu}(E)$ for every E. In other words, the measurable sets are those which, "when used to divide any other set into two pieces, give a division for which $\bar{\mu}$ is additive" as in figure 131. It is always true, whether or not A is measurable, that $\bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E) \geq \bar{\mu}(E)$ (by condition $\bar{2}$). One thinks of nonmeasurable sets as those "whose points are so densely distributed in X, with so many holes in between, that, whenever you try to cover $A \cap E$ by open intervals, and $A^c \cap E$ by open intervals, the open intervals for the two coverings are forced to overlap—overlap so much that you can actually do better by just covering E itself with open intervals, rather than by taking together these two coverings of $A \cap E$ and $A^c \cap E$."

Example. Let A be the subset of the set **R** of reals consisting of the open interval $(0,1)$. Then A is measurable in the measure space of reals. Indeed, let E be any subset of **R**, and let **C** be a collection, $I_1, I_2, \ldots$, of open intervals that covers E. Fix positive ϵ. Let $I_1 = (a_1, b_1)$, and introduce the three

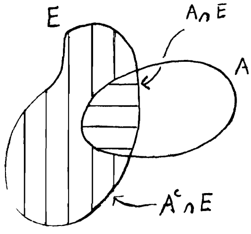

Figure 131

open intervals $\hat{I}_1 = I_1 \cap A$, $\tilde{I}_1 = (a_1,\epsilon)$, and $\tilde{I}_1{}' = (1 - \epsilon, b_1)$ (where, when the first entry in an open interval exceeds the second, it represents the empty interval). Similarly, set $\hat{I}_2 = I_2 \cap A$, $\tilde{I}_2 = (a_2,\epsilon/2)$, $\tilde{I}_2{}' = (1 - \epsilon/2, b_2)$, etc. Let $\hat{\mathbf{C}}$ be the collection $\hat{I}_1, \hat{I}_2, \cdots$ of open intervals, and $\tilde{\mathbf{C}}$ the collection $\tilde{I}_1, \tilde{I}_1{}', \tilde{I}_2, \tilde{I}_2{}', \tilde{I}_3, \cdots$ Then $m(\mathbf{C}) \geq m(\hat{\mathbf{C}}) + m(\tilde{\mathbf{C}}) - 4\epsilon$. But $\hat{\mathbf{C}}$ covers $A \cap E$, whence $\bar{\mu}(A \cap E) \leq m(\hat{\mathbf{C}})$, and $\tilde{\mathbf{C}}$ covers $A^c \cap E$, whence $\bar{\mu}(A^c \cap E) \leq m(\tilde{\mathbf{C}})$. Thus $m(\mathbf{C}) \geq \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E) - 4\epsilon$. Choosing the covering $\mathbf{C}$ of E with $m(\mathbf{C}) \leq \bar{\mu}(E) + \epsilon$, we have $\bar{\mu}(E) \geq \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E) - 5\epsilon$. Since this is true for every ϵ, $\bar{\mu}(E) \geq \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E)$. But $\bar{\mu}(E) \leq \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E)$, whence $\bar{\mu}(E) = \bar{\mu}(A \cap E) + \bar{\mu}(A^c \cap E)$. Of course, the measure of this measurable set A is one.

Example. By the argument above, every open interval is a measurable subset of the measure space of reals. Hence every open subset of the real line (since it is a union of a countable collection of open intervals) is measurable. So every closed subset (as the complement of an open subset) is measurable. In particular, every point of $\mathbf{R}$ is measurable (and has measure zero). Every countable subset of $\mathbf{R}$ (as the countable union of its points) is therefore measurable and has measure zero. Every subset of $\mathbf{R}$ which can be obtained by taking countable unions and intersections of open or closed sets is measurable.

This last example might suggest that perhaps every subset of the measure space of reals is measurable. This is not true, but is "nearly true": it is a rather subtle business even to construct a subset of the measure space of reals which is not measurable.

Example. For a and b numbers in the open interval $(0,1)$, write $a \approx b$ if $a - b$ is rational. This is an equivalence relation. Let A be a subset of $\mathbf{R}$ consisting of exactly one element from each equivalence class. (Thus no two elements of A differ by a rational, while any number in $(0,1)$ differs by a rational from some element of A.) We remark that we are actually using a version of Zorn's lemma (chapter 8) here. We show that the assumption that this A is measurable (with, say, $\mu(A) = \alpha$, an element of R^*) leads to a contradiction.

For each rational number r in (0,1), let A_r consist of all numbers a in (0,1) with $(a + r)$ (mod 1) in A. Then each A_r is measurable, with $\mu(A_r) = \alpha$. Any two A_r are disjoint (since no two elements of A differ by a rational), and the A_r cover (0,1) (since every number in this interval differs by some rational from an element of A). Hence we must have $1 = \mu(0,1) = \mu(\bigcup A_r) = \sum \mu(A_r) = \alpha + \alpha + \alpha + \cdots$ But this is a contradiction (for, if $\alpha = 0$, the sum on the right is zero, while, if $\alpha > 0$, the sum on the right is ∞; in neither case is that sum equal to one).

To summarize, we introduce the notion of a measure space, which represents "sets which have a reasonable notion of area, together with the areas of such sets" (where we regard "reasonable" as the conditions for a measure space). In the case of the reals, one knows what the "area" of an open interval should be. Then, for any subset of the reals, one "covers it with open intervals, takes the sum of their 'areas,' and then keeps trying coverings until this sum takes its smallest value." The result is a certain function $\bar{\mu}$. This $\bar{\mu}$, however, does not satisfy the additivity property necessary for a measure. One therefore takes, for the measurable sets, those which "split any other set into two parts for which $\bar{\mu}$ is additive." These measurable sets give a measure space. For this, the measure space of reals, practically every subset of **R** that one is ever likely to confront is measurable.

Exercise 292. Prove that every isomorphism (of topological spaces) from the real line to itself takes measurable sets to measurable sets. Find one such isomorphism which takes a set of measure zero to one of nonzero measure.

Exercise 293. Let X be a measure space, and let $A_1, A_2, \cdots$ be measurable sets, with $A_1 \subset A_2 \subset A_3 \cdots$ Prove that the limit of $\mu(A_n)$ as n approaches infinity is just $\mu(\bigcup A_i)$.

Exercise 294. Prove that, in the measure space of reals, every subset of a measurable set of measure zero is measurable.

Exercise 295. Consider the following subsets of the measure space of reals: $A_1 = [0,1]$, $A_2 = [0,1/3] \cup [2/3,1]$, $A_3 = [0,1/9] \cup [2/9,3/9] \cup [6/9,7/9] \cup [8/9,1]$, $A_4 = [0,1/27] \cup [2/27,3/27] \cup [5/27,7/27] \cup [8/27,9/27] \cup [18/27,19/27] \cup [20/27,21/27] \cup [24/27,25/27] \cup [26/27,1]$, etc. Let $C = \bigcap A_i$. (This C is called the Cantor set.) Show that C is uncountable, is measurable, and has measure zero. Using these facts and the previous exercise, show that there exist measurable subsets of the measure space of reals that cannot be written as countable unions and intersections of open or closed sets.

Exercise 296. Let X be any measure space. For E any subset of X, let **C** be a collection, $A_1, A_2, \ldots,$ of measurable sets that covers E. Set $m(\mathbf{C}) =$

$\sum \mu(A_i)$, and let $\overline{\mu}(E)$ be the greatest lower bound of these $m(\mathbf{C})$. Prove that this $\overline{\mu}$ satisfies the three conditions $\overline{1}$, $\overline{2}$, and $\overline{3}$. By theorem 53, we obtain a new measure on X. Prove that every subset of X measurable in the original measure space is measurable in this new one. Show that repeating this operation again, starting with the new measure space, yields nothing new.

Exercise 297. Is there some way to make measure spaces into a useful category?

Exercise 298. Does these exist some other way to make a measure space out of the set of reals such that every measurable set in the measure space of reals is also measurable in this new measure and has the same measure?

Exercise 299. Let X be a measure space. For A and B measurable sets, write $A \approx B$ if both $A - B$ and $B - A$ have measure zero. Prove that this is an equivalence relation and that two measurable sets in the same equivalence class have the same measure.

Exercise 300. Let E be a subset of the measure space of reals. Does there exist a measurable superset A of E such that every measurable B with $E \subset B \subset A$ has $\mu(A - B) = 0$?

Constructing Measure Spaces

In this chapter, we describe various techniques that yield measure spaces from measure spaces.

Let X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ be a measure space. Let K be a measurable subset of X. Denote by $\mathbf{N}$ the collection of all subsets of K which, as subsets of X, are measurable (i.e., which are in $\mathbf{M}$). For A in $\mathbf{N}$, set $\nu(A) = \mu(A)$. Then K, $\mathbf{N}$, $\mathbf{N} \xrightarrow{\nu} R^*$ is a measure space. [Proof: The first, third, and fourth conditions for a measure space are immediate from those conditions on measure space X. For the second condition, note that, for A in $\mathbf{N}$, the complement of A in K is $K - A$, which, since K is in $\mathbf{M}$, is itself in $\mathbf{M}$, and hence is in $\mathbf{N}$.] This K, $\mathbf{N}$, $\mathbf{N} \xrightarrow{\nu} R^*$ is called a *measure subspace* of X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$.

Next, let X', $\mathbf{M}'$, $\mathbf{M}' \xrightarrow{\mu'} R^*$ and X'', $\mathbf{M}''$, $\mathbf{M}'' \xrightarrow{\mu''} R^*$ be two measure spaces. We define another. Let $X = X' \cup_d X''$, the disjoint union. Thus a subset of X is a pair, (A', A''), where A' is a subset of X' and A'' is a subset of X''. Denote by $\mathbf{M}$ the collection of all such subsets with A' measurable in X' and A'' measurable in X''. Finally, for the measure of such a subset (A', A'') of X, we take $\mu(A', A'') = \mu'(A') + \mu''(A'')$. We claim that X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ is a measure space. [Proof: (1) $\emptyset = (\emptyset, \emptyset)$ is measurable in X, and $\mu(\emptyset) = \mu'(\emptyset) + \mu''(\emptyset) = 0$. (2) The complement of $A = (A', A'')$ in X is (A'^c, A''^c). (3) The union of $A_1 = (A_1', A_1'')$, $A_2 = (A_2', A_2'')$, $\cdots$ is $(\cup A_1', \cup A_1'')$. (4) Let $A_1 = (A_1', A_1'')$, $A_2 = (A_2', A_2'')$, $\cdots$ be disjoint and measurable in X (so $A_1', A_2', \cdots$ are disjoint and measurable in X', and $A_1'', A_2'', \cdots$ are disjoint and measurable in X''). Then $\sum \mu(A_i) = \sum (\mu'(A_i') + \mu''(A_i'')) = \sum \mu'(A_i') + \sum \mu''(A_i'') = \mu'(\cup A_i') + \mu''(\cup A_i'') = \mu(\cup A_i', \cup A_i'') = \mu(\cup A_i)$.] This measure space is called the ***disjoint union*** of measure spaces X' and X''.

For the third construction, let X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ and X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu'} R^*$ be measure spaces (i.e., they have the same underlying set X and the same measurable sets, but they may assign different measures to these measurable sets). We may write $\mu \leq \mu'$ if $\mu(A) \leq \mu'(A)$ for every measurable set A, noting that this is a partial ordering. Note also that X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu+\mu'} R^*$ is also a measure space, where we set $(\mu + \mu')(A) = \mu(A) + \mu'(A)$ for any measurable set A. This is called the ***sum*** (of these two measures on X). Note that the sum is commutative and associative (i.e., $\mu + \mu' = \mu' + \mu$ and $(\mu + \mu')$

$+ \mu'' = \mu + (\mu' + \mu''))$. Finally, fix any positive number r. Then X, **M**, $\mathbf{M} \xrightarrow{r\mu} R^*$ is a measure space, where, for A measurable, $(a\mu)(A)$ is the element $a\mu(A)$ of R^* if $\mu(A)$ is finite, and is the element ∞ of R^* if $\mu(A)$ is infinite. That the properties one expects to be satisfied are indeed satisfied is immediate: $r(r'\mu) = (rr')\mu$, $r(\mu + \mu') = r\mu + r\mu'$.

The final construction is the most interesting. We first need some notation. Let a and b be elements of R^*. We denote by ab the element of R^* which is i) this product of numbers if both a and b are finite, ii) the element ∞ if one of a or b is infinite, and the other is nonzero (in particular, if both are infinite), and iii) the element 0 if one of a or b is infinite, and the other zero (!). We have, as is easily checked, $ab = ba$, $a(bc) = (ab)c$, and $a(b + c) = ab + ac$.

Now let X', $\mathbf{M}'$, $\mathbf{M}' \xrightarrow{\mu} R^*$ and X'', $\mathbf{M}''$, $\mathbf{M}'' \xrightarrow{\mu''} R^*$ be measure spaces. Let $X = X' \times X''$, Cartesian product, and, for A' in X' and A'' in X'', write $A' \times A''$ for the subset of X consisting of (x',x'') with x' in A' and x'' in A''. Now let E be any subset of X. Let **C** denote a collection of subsets of X of the form $A_1' \times A_1'', A_2' \times A_2'', \ldots$, with $A_1', A_2', \cdots$ measurable in X' and $A_1'', A_2'', \cdots$ measurable in X'', and such that this collection **C** covers E. Set $m(\mathbf{C}) = \mu'(A_1')\mu''(A_1'') + \mu'(A_2')\mu''(A_2'') + \cdots$ (Think of $\mu'(A_1')\mu''(A_1'')$ as the "area of the rectangle $A_1' \times A_1''$ in X," so $m(\mathbf{C})$ is the "sum of the areas of this family of rectangles, which covers E.") Finally, let $\bar{\mu}(E)$ be the greatest lower bound of these $m(\mathbf{C})$. This $\bar{\mu}$ satisfies the three conditions $\bar{1}$, $\bar{2}$, and $\bar{3}$ (by a word-for-word repetition of the argument which follows those conditions). By theorem 53, we obtain a measure space, X, **M**, $\mathbf{M} \xrightarrow{\mu} R^*$, called the *product* of measure spaces X', $\mathbf{M}'$, $\mathbf{M}' \xrightarrow{\mu'} R^*$ and X'', $\mathbf{M}''$, $\mathbf{M}'' \xrightarrow{\mu''} R^*$.

Example. Let each of X' and X'' be the measure space of reals. Then X is the underlying set of the topological plane. Ordinary rectangles in this plane have as their measure their usual area. The measure of a disk of radius r is πr^2, and similarly for other geometrical figures. Thus this product of measure spaces yields the usual notion of "area of figures in the plane."

Exercise 301. Can one take disjoint unions, sums, and products of an infinite collection of measure spaces?

45

Measurable Functions

We now introduce the functions that are "candidates for those that can be integrated over a measure space."

Let X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ be a measure space. Let $X \xrightarrow{f} \mathbf{R}$ be a real-valued function on the set X. We say that this f is *measurable* if inverse images by f of open sets are measurable, that is, if, for each open O in the real line, $f^{-1}[O]$ is measurable in X. A number of facts about measurable functions are immediate. First, note that (since complements of open sets are closed, since inverse images take complements to complements, since complements of measurable sets are measurable), for f measurable, $f^{-1}[C]$ is measurable for each closed C in the real line. In particular (since each point of the real line defines a closed subset), the inverse image, by a measurable function, of any point of the real line is a measurable subset of X. Similarly, the inverse image, by a measurable function, of any subset of the real line which can be written as a countable union or intersection of open or closed sets is measurable in X. The following fact is often useful: $X \xrightarrow{f} \mathbf{R}$ is measurable if and only if, for each real number r, $f^{-1}[(-\infty,r)]$ is measurable. [Proof: The "only if" part is immediate, since $(-\infty,r)$ is open in the real line. For the converse, suppose that each $f^{-1}[(-\infty,r)]$ is measurable. Then each $f^{-1}[[r,\infty)]$ ($= \{f^{-1}[(-\infty,r)]\}^c$) is measurable. Hence each $f^{-1}[(r',\infty)]$ ($= f^{-1}[(r' + 1/2,\infty)] \cup f^{-1}[(r' + 1/4,\infty)] \cup f^{-1}[(r' + 1/8,\infty)] \cup \cdots$) is measurable. Hence each $f^{-1}[(r',r)]$ ($= f^{-1}[(r',\infty)] \cap f^{-1}[(-\infty,r)]$) is measurable. Since every open subset O of the real line can be written as a countable union of such open intervals, each $f^{-1}[O]$ is measurable. Hence f is measurable.]

Example. Let $\mathbf{R} \xrightarrow{f} \mathbf{R}$ be continuous. Then, regarding the $\mathbf{R}$ on the left as the measure space of reals, f is measurable. Indeed, for O open in $\mathbf{R}$, $f^{-1}[O]$ is open in $\mathbf{R}$, and hence measurable, since every open subset of the real line is measurable. The function $\mathbf{R} \xrightarrow{f} \mathbf{R}$ with $f(r)$ one if r is rational, and zero if r is irrational, is measurable, since both the set of rationals, and the set of irrationals, is measurable in the measure space of reals.

Next, let $\mathbf{R} \xrightarrow{F} \mathbf{R}$ be continuous, and let $X \xrightarrow{f} \mathbf{R}$ be measurable. Then $X \xrightarrow{F \circ f} \mathbf{R}$ is measurable, for, for O open in the real line, $F^{-1}[O]$ is open in the real line, whence $f^{-1}[F^{-1}[O]]$ is a measurable subset of X. But $f^{-1}[F^{-1}[O]] = (F \circ f)^{-1}[O]$. In particular, for $X \xrightarrow{f} \mathbf{R}$ measurable, so is $|f|$ (the function on X

whose value at x in X is $|f(x)|$), for $\mathbf{R} \xrightarrow{F} \mathbf{R}$ given by $F(r) = |r|$ is continuous, while $|f| = F \circ f$. More generally, let $\mathbf{R} \times \mathbf{R} \times \cdots \times \mathbf{R} \xrightarrow{F} \mathbf{R}$ be any continuous function of n real variables, and let $f_1, \ldots, f_n$ be measurable functions. Then $F(f_1, \ldots, f_n)$ (the real-valued function on X whose value at x in X is $F(f_1(x), \ldots, f_n(x))$) is measurable. [Proof: For O open in the real line, $F^{-1}[O]$ is open in $\mathbf{R} \times \mathbf{R} \times \cdots \times \mathbf{R}$. This open set can be written as the union of a countable collection of open sets of the form $O_1 \times O_2 \times \cdots \times O_n$, where each of $O_1, \ldots, O_n$ is open in the real line. But the inverse image of this set by $X \xrightarrow{f_1, \ldots, f_n} \mathbf{R} \times \cdots \times \mathbf{R}$ is $f_1^{-1}[O_1] \cap f_2^{-1}[O_2] \cap \cdots \cap f_n^{-1}[O_n]$, and hence is measurable.] Thus, for example, for f and f' measurable, both $f + f'$ (the function whose value at x in X is $f(x) + f'(x)$) and ff' are measurable.

We next consider the behavior of measurable functions under taking limits. Let $f_1, f_2, \cdots$ be measurable functions on measure space X. Let $\lim f_n$ exist, and denote this function by f (that is, let, for each x in X, the sequence $f_1(x), f_2(x), \cdots$ in the real line converge, and let f be the function such that this sequence converges to $f(x)$ for each x in X). We claim that $X \xrightarrow{f} \mathbf{R}$ is therefore measurable. [Proof: Fix number r. For ϵ a positive number, and n a positive integer, set $K(\epsilon,n) = f_n^{-1}[(-\infty, r-\epsilon)] \cup f_{n+1}^{-1}[(-\infty, r-\epsilon)] \cup \cdots$, a measurable subset of X. Set $K(\epsilon) = K(\epsilon,1) \cap K(\epsilon,2) \cap \cdots$, so $K(\epsilon)$ is measurable. Note that $f^{-1}[(-\infty, r-2\epsilon)] \subset K(\epsilon) \subset f^{-1}[(-\infty,r)]$. Hence $f^{-1}[(-\infty,r)]$ $(= K(1) \cup K(1/2) \cup K(1/4) \cup \cdots)$ is measurable. That is, f is measurable.] Thus limits of measurable functions yield measurable functions.

In short, "essentially every finite or countable operation one can think of, applied to measurable functions, yields a function which is again measurable."

A certain, particularly simple class of measurable functions is of special interest. Let X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ be a measure space, and let $X \xrightarrow{f} \mathbf{R}$. This f is said to be a *step function* if i) f is measurable and ii) $f[X]$ is a finite subset of $\mathbf{R}$. Thus, for f a step function, there is a finite number, $r_1, \ldots, r_n$, of numbers, with $A_1 = f^{-1}[r_1], \ldots, A_n = f^{-1}[r_n]$ measurable subsets of X which are disjoint and whose union is X. With the exception of taking limits, most operations, when applied to step functions, yield step functions. For example, if f and f' are step functions, then so are $|f|$, $f + f'$, and ff'. Limits of step functions are, of course, measurable functions, by the remark above.

The interest in step functions lies, in part, in the fact that every measurable function is a limit of step functions. In fact, we have the following slightly stronger result.

THEOREM 54. *Let X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ be a measure space, and let $X \xrightarrow{f} \mathbf{R}$ be a measurable function with $f \geq 0$ (i.e., with $f(x) \geq 0$ for every x in X).*

Then there exists a sequence, $f_1, f_2, \cdots$ *of step functions with* $0 \le f_1 \le f_2 \le \cdots$, *and with* $\lim f_n = f$.

Proof. The proof is by construction. Set

$$f_1(x) = 0 \text{ if } 0 \le f(x) < 1$$
$$f_1(x) = 1 \text{ if } 1 \le f(x)$$

$$f_2(x) = 0 \text{ if } 0 \le f(x) < 1/2$$
$$f_2(x) = 1/2 \text{ if } 1/2 \le f(x) < 1$$
$$f_2(x) = 1 \text{ if } 1 \le f(x) < 3/2$$
$$f_2(x) = 3/2 \text{ if } 3/2 \le f(x) < 2$$
$$f_2(x) = 2 \text{ if } 2 \le f(x)$$

and similarly for f_3 (using a subdivision in intervals of 1/4 from 0 to 3), f_4, etc. It is clear that this is an increasing sequence of step functions, with limit f. □

Now consider a (not necessarily positive) measurable function $X \xrightarrow{f} \mathbf{R}$. Write $f = f_+ - f_-$, where f_+ $(= 1/2(|f| + f))$ and f_- $(= 1/2(|f| - f))$ are non-negative measurable functions. Since each of these is, by the theorem above, a limit of step functions, so is f.

Thus the measurable functions are precisely the functions which can be obtained as limits of step functions.

Exercise 302. Find an example of a function on the measure space of reals that is not measurable.

Exercise 303. Let $f_1, f_2, \cdots$ be measurable functions on measure space X. For each x in X, let the least upper bound of the numbers $f_1(x), f_2(x), \cdots$ exist, and let f be the function such that this bound is $f(x)$ for each x. Prove that f is measurable.

Exercise 304. Is it true that f is measurable if and only if $|f|$ is?

Exercise 305. Let $X \xrightarrow{f} \mathbf{R}$ be a step function, and let $\mathbf{R} \xrightarrow{F} \mathbf{R}$ be any function. Prove that $F \circ f$ is a step function. Does this work in more than one variable?

Exercise 306. Find an example of measurable functions f and f' with $f > f'$ such that there is no step function $\underline{f}$ with $f \ge \underline{f} \ge f'$.

46

Integrals

The "space over which integrals are to be performed" is a measure space. The "things which are to be integrated" are measurable functions. We now define integrals and give a few of their properties.

The idea is to treat the step functions first and then to extend this treatment to all measurable functions. Let X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$ be a measure space, and let $X \xrightarrow{f} \mathbf{R}$ be a step function. Let $a_1, \ldots, a_n$ be the n values that f can assume, and set $A_1 = f^{-1}[a_1], \ldots, A_n = f^{-1}[a_n]$, the (measurable) subsets of X on which f assumes each of its n values. In this case, one has a good idea of what the integral of f "should be," namely $a_1\mu(A_1) + \cdots + a_n\mu(A_n)$ (i.e., the sum of the values of f, each multiplied by the measure of the region on which f assumes that value). There are, however, a couple of problems with this. First, it may happen that, for example, a_1 is nonzero but $\mu(A_1)$ is infinite. Then our "integral" would be infinite. Second, and even more serious, it may happen that, for example, a_1 is positive and a_2 is negative, while both $\mu(A_1)$ and $\mu(A_2)$ are infinite. Then we would want to add "$+\infty$" to "$-\infty$" in the sum above, with, presumably, an indeterminate result. It is convenient to rule out such pathologies at the beginning. Thus a step function f is said to be *integrable* if (with $a_1, \ldots, a_n$ and $A_1, \ldots, A_n$ as above) whenever a_i is nonzero, $\mu(A_i)$ is finite. In this case, the real number $a_1\mu(A_1) + \cdots + a_n\mu(A_n)$ is called the *integral* of f and is written $\int_X f \, d\mu$.

All the properties one would expect to be satisfied by integrals are in fact satisfied. Thus, if f is an integrable step function and a is a number, then af is integrable (since, whenever af is nonzero, so is f) and $\int_X af \, d\mu = a \int_X f \, d\mu$. If f and f' are integrable step functions, then $f + f'$ is an integrable step function (since, whenever $f + f'$ is nonzero, at least one of f or f' is nonzero) and $\int_X (f + f') \, d\mu = \int_X f \, d\mu + \int_X f' \, d\mu$. If f and f' are integrable step functions, with $f \leq f'$, then $\int_X f \, d\mu \leq \int_X f' \, d\mu$.

We thus know how to integrate step functions. Next, one wishes to define the integral of an arbitrary measurable function using, somehow, the fact that this function is necessarily a limit of step functions. There are, however, a number of technical problems which force one, in order to obtain a reasonable definition, to make it with some care. Two examples will illustrate this remark.

Example. Consider, on the measure space of reals, the sequence $f_1, f_2, \cdots$ of functions, with $f_n(x)$ equal to one if $n \leq x \leq n + 1$, and to zero otherwise. These are shown in figure 132. Each is an integrable step function, and, in fact, each has integral one. But lim f_n is the function which is zero everywhere (because the "unit humps move off to infinity"), while this integrable step function has integral zero.

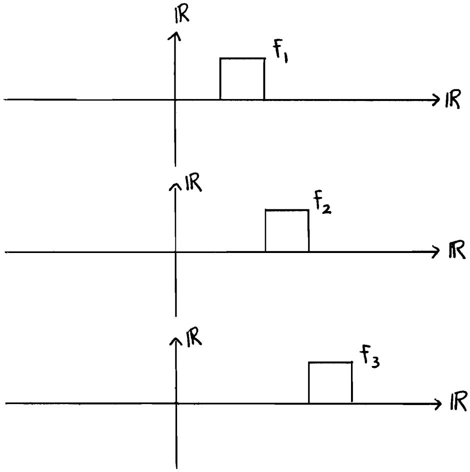

Figure 132

The problem here is that one can have a "hump which contributes to each f_n but which does not contribute to the limit function." In order to rule out this type of behavior, one would like to require that "once such a hump appears in one f_n, it does not move away in later instances of f_n." One might accomplish this, for example, by requiring that $f_1 \leq f_2 \leq \cdots$

Example. Consider, on the measure space of reals, the (measurable) function f with $f(x) = -1/(1 + x_2)$ (figure 133). One would certainly want this function to, ultimately, be integrable (and, in fact, to have integral $-\pi$). On the other hand, no step function f' with $f' \leq f$ is integrable (for, since f takes only negative values, the $a_1, \ldots, a_n$ for f' must all be negative while at least one of $A_1, \ldots, A_n$ must have infinite measure).

It is clear from these examples that what one wants to do is have one's sequence of step functions "increase to f when f is positive and decrease to f when f' is negative."

Let X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} r^*$ be a measure space, and let $X \xrightarrow{f} \mathbf{R}$ be a measurable function. Suppose first that $f \geq 0$. Then this f is said to be *integrable* if every step function f with $f \geq \mathit{f} \geq 0$ is integrable and the set of values of the integrals of such step functions is bounded. The least upper bound of these

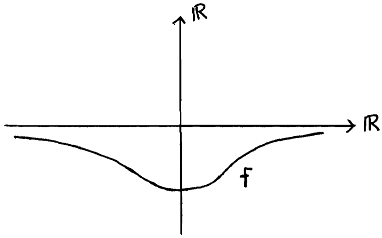

Figure 133

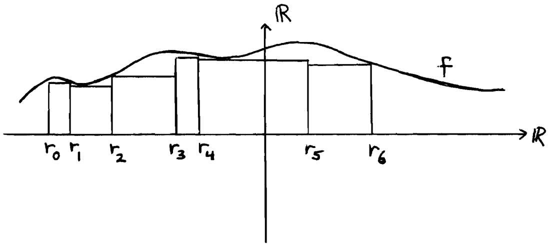

Figure 134

numbers $\int_X \underline{f}\, d\mu$ is called the *integral* of f and is written $\int_X f\, d\mu$. More generally (i.e., not assuming $f \leq 0$), we call $X \xrightarrow{f} \mathbf{R}$ *integrable* if, writing $f = f_+ - f_-$ with f_+ and f_- non-negative, both f_+ and f_- are integrable (in the sense just defined). Of course, the integral of f in this case is just $\int_X f\, d\mu = \int_X f_+\, d\mu - \int_X f_-\, d\mu$. We note that these definitions—of integrable and of the value of the integral—agree when more than one is applicable (e.g., a step function is integrable as a step function if and only if it is integrable as a measurable function, and the two values of the integral then agree).

Let us look at this definition in a little more detail. Let $X \xrightarrow{f} \mathbf{R}$ be measurable, with $f \geq 0$. Then we may "approximate f as closely as we wish by a step function $\underline{f}$ with $f \geq \underline{f} \geq 0$." We think of $\int_X \underline{f}\, d\mu$ as "an approximation to $\int_X f\, d\mu$, an approximation which becomes better and better as the step functions increase to approach f." Let us compare this situation with that of the Riemann integral. In the Riemann case, one must consider only functions on the measure space of reals rather than on an arbitrary measure space. Let $\mathbf{R} \xrightarrow{f} \mathbf{R}$ be measurable, with $f \geq 0$ (figure 134). To obtain the Riemann integral of f, one divides a portion of the real line into intervals, that is, one considers numbers $r_0 < r_1 < \cdots < r_n$. One then approximates f by a function $\underline{f}$, where, for x in the interval $[r_n, r_{n+1}), \underline{f}(x)$ is the greatest lower bound of f on this interval. This, too, is a step function, for it is constant on each of the intervals $(r_0, r_1), \cdots, (r_{n-1}, r_n)$. Thus, in the Riemann case, one approximates by functions which are "constant on intervals." On the other hand, for the integral as defined above, one approximates by step functions which are "constant on arbitrary measurable sets." In short, one broadens one's class of approximating functions and is thus able to integrate more functions. Note also that one has no notion of an "interval" on an arbitrary measure space X (as opposed to on the measure space of reals), and hence has no notion of a "Riemann integral" on an arbitrary measure space.

There is another way to view this difference between the integral as defined above and the Riemann integral. Let $\mathbf{R} \xrightarrow{f} \mathbf{R}$. To obtain the Riemann integral of f, one splits the $\mathbf{R}$ on the left (the measure space of reals) into intervals to obtain an approximating step function for f. In the present case, by contrast, one obtains approximating step functions for f by splitting the $\mathbf{R}$ on the right into intervals (see, e.g., theorem 54). Since the integral of f is to be roughly the "sum of the values of f times the measure of the region on which f has that value," the latter seems more natural.

Example. Let $\mathbf{R} \xrightarrow{f} \mathbf{R}$ be given by $f(x) = 1$ if x is rational and $f(x) = 0$ if x is irrational (figure 135). Then f is integrable (in fact, is a step function) and $\int_X f \, d\mu = 0$. On the other hand, this f is not Riemann integrable.

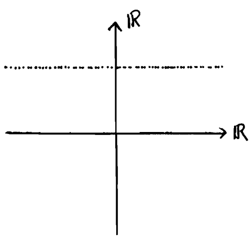

Figure 135

Example. Let $\mathbf{R} \xrightarrow{f} \mathbf{R}$ be given by $f(x) = (x \sin x)/(1 + x^2)$. Then f is measurable but not integrable (for, e.g., f_+ is not). (One sometimes says that such functions are "integrable but not absolutely integrable." This notion, like that of the Riemann integral, refers, e.g., to an ordering of the reals and hence is not applicable for general measure spaces.)

There must be thousands of properties of integrals. One's response to this situation is the usual one: one tries to develop a feeling for what sorts of things are true and what sorts are false, and to become adept at finding proofs and counterexamples when needed. We here illustrate this art with a few examples.

First, note that, if $f \leq f'$ are integrable functions, then $\int_X f\, d\mu \leq \int_X f'\, d\mu$. Hence, in particular, if f is integrable with $f \leq 0$, we have $\int_X f\, d\mu \leq 0$. It is of interest to ask when equality holds. So let f be integrable, with $f \leq 0$. It is clear that, if $f(x) = 0$ except for a subset A of X with $\mu(A) = 0$, then $\int_X f\, d\mu = 0$. We claim that the converse is also true. Indeed, set $A_1 = f^{-1}[(1/2,\infty)]$, $A_2 = f^{-1}[(1/4,\infty)]$, $\cdots$ Then $A_1 \subset A_2 \subset \cdots$ If, for any n, we had $\mu(A_n) \neq 0$, then we would have $\int_X f\, d\mu \neq 0$. Hence each A_n has measure zero, whence $0 = \mu(\cup A_n) = f^{-1}[(0,\infty)]$. That is, f vanishes except on a set of measure zero. It often happens in this subject that some statement about points of X is true except on some measurable set of measure zero. When this is the case, the statement is said to be true *almost everywhere*. Thus we have just shown that a non-negative integrable function has vanishing integral if and only if that function vanishes almost everywhere.

Next, let $X \xrightarrow{f} \mathbf{R}$ be an integrable function, and let A be any measurable subset of X. Then we may regard A as a measure subspace of X and, restricting f to A, obtain a mapping $A \xrightarrow{\tilde{f}} \mathbf{R}$. It is clear that this $\tilde{f}$ is also integrable: we write its integral as $\int_A f\, d\mu$ and call A the *region of integration*. It is immediate from the definition that, if A and B are disjoint measurable sets, then $\int_{A \cup B} f\, d\mu = \int_A f\, d\mu + \int_B f\, d\mu$. This is also true in a certain sense for any countable collection of disjoint measurable sets. To see what the sense is, we first consider an example.

Example. Let $X = \mathbf{R}$, the measure space of reals, and let $X \xrightarrow{f} \mathbf{R}$ have action $f(x) = \sin 2\pi x$. For $n = 1,2, \ldots,$ let $A_n = [n-1,n)$, a measurable set. Then f is integrable on each A_n, and, in fact, $\int_{A_n} f\, d\mu = 0$ for every n. Hence $\sum \int_{A_n} f\, d\mu = 0$. On the other hand, $\int_{\cup A_n} f\, d\mu$ does not even exist, since f is not integrable on $\cup A_n$.

The example suggests what must be done. We claim: if $X \xrightarrow{f} \mathbf{R}$ is integrable, and $A_1,A_2, \cdots$ are measurable disjoint subsets of X, then $\sum \int_{A_n} f\, d\mu = \int_{\cup A_n} f\, d\mu$. [Sketch of proof: It suffices to consider the case $f \geq 0$. Given any step function $\underline{f}$ with $0 \leq \underline{f} \leq f$, this step function, restricted to each A_n, is a step function on that A_n. Hence $\sum \int_{A_n} f\, d\mu \leq \int_{\cup A_n} f\, d\mu$. To prove the reverse inequality, note that, given step functions $\underline{f}_1,\underline{f}_2, \cdots$ on $A_1,A_2, \ldots,$ with $0 \leq \underline{f}_n \leq f$ on A_n, the sum of any finite number of these is a step function $\underline{f}$ on $\cup A_n$ with $0 \leq \underline{f} \leq f$ on $\cup A_n$.]

Consider now a measure space X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$, and a measurable function f with $f \geq 0$. Consider X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\nu} R^*$, where, for A measurable, $\nu(A)$ is

∞ if f is not integrable on A and is $\int_{\substack{A\\ \nu}} f\,d\mu$ if f is integrable on A. The result obtained above shows that X, $\mathbf{M}$, $\mathbf{M} \to R^*$ is a measure space.

We next consider the behavior of the integral under linear combinations of functions. It is immediate that, if f is integrable and a is any number, then af is integrable and $\int_X af\,d\mu = a\int_X f\,d\mu$. We claim, furthermore, that if f and g are integrable, then so is $f + g$, and $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$. [Sketch of proof: It suffices to consider the case $f \geq 0$, $g \geq 0$. Choose positive numbers N and ϵ, and denote by A the subset of X consisting of x with $1/N \leq f(x) \leq N$ and $1/N \leq g(x) \leq N$. Choose step functions $\underline{f}$ and $\underline{g}$ on A with $f \geq \underline{f} \geq (1-\epsilon)f$ and $g \geq \underline{g} \geq (1-\epsilon)g$. Then, for $\underline{h}$ any step function on A with $f + g \geq \underline{h} \geq 0$, we have $\underline{h} \leq f + g \leq \underline{f}/(1-\epsilon) + \underline{g}/(1-\epsilon)$. Hence $\int_A (f+g)\,d\mu \leq \int_A f\,d\mu/(1-\epsilon) + \int_A g\,d\mu/(1-\epsilon)$. Since ϵ is arbitrary, $\int_A (f+g)\,d\mu \leq \int_A f\,d\mu + \int_A g\,d\mu$. Since this is true for every N, $\int_X (f+g)\,d\mu \leq \int_X f\,d\mu + \int_A g\,d\mu$. To obtain the reverse inequality, note that, for $\underline{f}$ and $\underline{g}$ step functions, with $f \geq \underline{f} \geq 0$ and $g \geq \underline{g} \geq 0$, $\underline{f} + \underline{g}$ is a step function, with $f + g \geq \underline{f} + \underline{g} \geq 0$.] We can state these conclusions thus: "take the integral" is a linear mapping from the vector space of integrable functions to the vector space of real numbers.

As a final example of manipulating integrals, we consider limits. Let $f_1, f_2, \cdots$ be integrable functions, and let $f = \lim f_n$ exist. We wish to consider conditions under which $\lim \int_X f_n\,d\mu$ is equal to $\int_X f\,d\mu$. The first example of this chapter shows that at least some additional condition will be necessary. The following example may suggest what this condition should be.

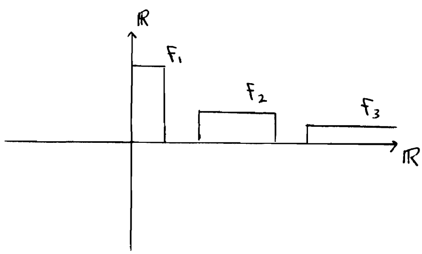

Figure 136

Example. Let $X = \mathbf{R}$. Let $f_1, f_2, \cdots$ be the sequence of functions illustrated in figure 136 (the "hump gets flatter and wider as it moves off to infinity with n, keeping always $\int_X f_n\,d\mu$ equal to one"). The limit of this sequence is $f = 0$, a limit whose integral is zero. Alternatively, one might let $f_1, f_2, \cdots$ be the sequence of functions illustrated in figure 137 (the "hump gets higher and thinner, but approaches, e.g., zero, i.e., it remains within a set of finite measure, keeping $\int_X f_n\,d\mu$ equal to one"). Again, the limit function is

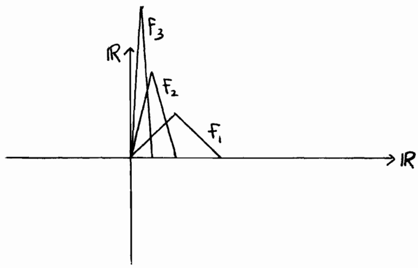

Figure 137

$f = 0$.

One of the more useful results on interchange of limits and integration is this: if $f = \lim f_n$ and if there is an integrable function g with $|f_n| \le g$ for every n, then $\lim \int_X f_n \, d\mu = \int_X f \, d\mu$. [Sketch of proof: For each n, let h_n be the function with $h_n(x) =$ least upper bound of $|f_n(x) - f(x)|, |f_{n+1}(x) - f(x)|, \cdots$ Then, for each n, $h_n \ge 0$, h_n is measurable and, since $h_n \le 2g$, h_n is integrable. We have $\lim h_n = 0$, whence $\lim \int_X h_n \, d\mu = 0$, whence $\lim \int_X |f_n - f| \, d\mu = 0$, whence $\lim |\int_X (f_n - f) \, d\mu| = 0$, whence $\lim \int_X f_n \, d\mu = \int_X f \, d\mu$.]

Exercise 307. Find an example to show that the product of two integrable functions need not be integrable.

Exercise 308. Prove that an integrable function on the measure space of reals is Riemann-integrable if and only if it is continuous almost everywhere.

Exercise 309. Let f be an integrable function on measure space X. Suppose that, for every measurable set A, $\int_A f \, d\mu = 0$. Does it follow that f vanishes almost everywhere?

Exercise 310. Let f be an integrable function on the measure space of reals. For each real number r, let $A_r = (-\infty, r)$, and set $g(r) = \int_{A_r} f \, d\mu$. Prove that the derivative of g exists and is equal to f almost everywhere.

Exercise 311. Let X be a measure space. Denote by V the set of equivalence classes of integrable functions under the equivalence relation "equal almost everywhere." Make this V into a real vector space. Given two elements of V, with representatives f and g, let the distance between them be given by $\int_X |f - g| \, d\mu$. Prove that this is a metric on the set V. Next, show that, with the topology induced by this metric, V becomes a topological vector space.

Prove that the topological vector space V is complete. Finally, check that the closure of the subset of V consisting of the integrable step functions is V itself.

Exercise 312. Prove that f is integrable if and only if $|f|$ is.

Exercise 313. A sequence $f_1, f_2, \cdots$ of measurable functions is said to converge to f in measure if, for any positive ϵ, $\lim \mu(A_n) = 0$, where A_n consists of points x with $|f_n(x) - f(x)| \geq \epsilon$. Give an example to show that $\lim f_n = f$ does not necessarily imply that f_n converges to f in measure. Prove that this implication does hold if $\mu(X)$ is finite.

47

Distributions

There are certain applications for which one wishes to "broaden the class of ordinary functions" to include "generalized functions, which one can think of as ordinary functions for which it is allowed that the function become infinite in an infinitesimally small region," where this type of behavior is restricted only by the requirement that certain integrals converge. These "generalized functions" are called distributions. We shall here give the definition of a distribution, and a few elementary properties. In order to avoid certain irrelevant technical complications, we shall work on a compact subset of the real line. Generalizations to $\mathbf{R} \times \cdots \times \mathbf{R}$, to arbitrary manifolds, and to noncompact subsets are available and involve essentially no new ideas.

Let $\mathbf{R} \xrightarrow{f} \mathbf{R}$ be a mapping of sets. By the *support* of f, we mean the subset $\mathrm{Cl}(f^{-1}[\mathbf{R} - 0])$ of the set of real numbers, that is, the closure of the set of numbers r for which $f(r)$ is nonzero. Note that, for example, the support of $f + g$ is a subset of the union of the support of f and the support of g. Such a function is said to be C^0 if it is continuous. If, in addition, the derivative of f exists (i.e., if the limit, as x approaches zero, of $(f(r + x) - f(r))/x$ exists for every r) and if this derivative, which we denote by f', is also continuous, then f is said to be C^1. If f is C^1 and if the derivative of f' exists and is continuous, then f is said to be C^2, etc. Finally, if all derivatives of f, to all orders, exist and are continuous, then f is said to be C^∞.

Now fix a compact subset K of the real line. We denote by T the collection of all C^∞ functions $\mathbf{R} \to \mathbf{R}$ whose support is a subset of K (that is, T consists of functions that vanish outside of K, and all of whose derivatives exist and are continuous). Clearly, the sum of two such functions is another, and any numerical multiple of such a function is another. Thus T is a real vector space. What we now wish to do is to make this T into a topological vector space. For f and g two functions in T, set

$$d(f,g) = \frac{\max|f - g|}{1 + \max|f - g|} + (1/2)\,\frac{\max|f' - g'|}{1 + \max|f' - g'|} + (1/4)\,\frac{\max|f'' - g''|}{1 + \max|f'' - g''|} + \cdots$$

where "max" refers to the maximal value on K (which always exists, since K is compact). Note also that the sum above necessarily converges, since the first term cannot exceed one, the second 1/2, the third 1/4, etc. Finally, note that this "$d(\ ,\ )$" is a metric on the set T (since each term defines a metric on

T). Thus, since we have a metric on the vector space T, we have a topology on T. (Intuitively, two functions in T are "close" if "their values and all their derivatives are close at every point of K.") It is easily checked that T thus becomes a topological vector space.

A *distribution* is a continuous linear mapping $T \xrightarrow{\varphi} \mathbf{R}$ of topological vector spaces (where, on the right, there appears the obvious topological vector space of reals). That is, a distribution φ assigns, to each C^∞ function f with support in K, a real number $\varphi(f)$, where this assignment is linear (i.e., $\varphi(f + ag) = \varphi(f) + a\varphi(g)$) and continuous (i.e., "functions whose derivatives are all close at every point of K are taken, by φ, to nearby real numbers").

Example. Let $K = [0,1]$. Let $\mathbf{R} \xrightarrow{\alpha} \mathbf{R}$ be any continuous function (just integrable on K would suffice). For f in T, set $\varphi(f) = \int_K \alpha f \, d\mu$, where the measure is that for the measure space of reals. Since the integral is linear, this mapping $T \xrightarrow{\varphi} \mathbf{R}$ is linear. This mapping is also continuous (for, for $\max|f - g| = \epsilon$, $\varphi(f) - \varphi(g) \leq \epsilon \int_K |\alpha| \, d\mu$). Thus we obtain a distribution.

This example is the key to the definition. Ordinary functions yield, by "integrating the product," distributions. It is also easily checked that, if two continuous functions define the same distribution, then they are equal. Thus "the distributions include the ordinary functions as special cases." In order to claim that distributions represent "generalized functions," we must display some distributions which do not result from the construction above.

Example. Let $K = [0,1]$. For f in T, set $\varphi(f) = f(3/4)$ (i.e., $\varphi(f)$ assigns to f the real number that is the value of f at the point 3/4 of [0,1]). This $T \xrightarrow{\varphi} \mathbf{R}$ is clearly a continuous linear mapping. This distribution is called the *Dirac delta distribution* (at 3/4). Let us show that this distribution cannot be expressed as in the previous example. Suppose that there were a continuous function $\mathbf{R} \xrightarrow{\alpha} \mathbf{R}$ such that $f(3/4) = \int_K \alpha f \, d\mu$ for every f in T. We show that this supposition leads to a contradiction. If there were a number r in K different from 3/4 with $\alpha(r)$ nonzero, then, choosing f in T to vanish except near r, and there having the same sign as $\alpha(r)$, we would have $f(3/4) = 0$, and $\int_K \alpha f \, d\mu$ positive, violating $f(3/4) = \int_K \alpha f \, d\mu$. Hence we must have $\alpha(r) = 0$ for every r different from 3/4. Since α is continuous, α must be zero everywhere in K. Hence $\int_K \alpha f \, d\mu$ vanishes for all f, again violating $f(3/4) = \int_K \alpha f \, d\mu$. Hence the Dirac delta distribution does not arise from any "ordinary function."

There is, however, a certain sense in which the Dirac delta distribution arises from "a limit of ordinary functions." Choose a sequence of continuous functions, $\alpha_1, \alpha_2, \cdots$ each "having integral one, and peaked about the value 3/4," such that, as one moves along this sequence of functions, "the peaks become higher and narrower." These are illustrated in figure 138. Then it is

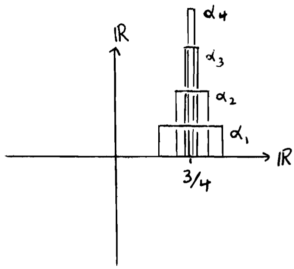

Figure 138

clear intuitively (and also true, provided this intuitive characterization of the α_n is made precise) that $\lim \int_K \alpha_n f \, d\mu = f(3/4)$. Thus one can think of the Dirac delta distribution as "the limit of the sequence of distributions defined by these α_n." Of course, there is no function that is the limit of the sequence $\alpha_1, \alpha_2, \cdots$ of functions.

Example. Let $K = [0,1]$. For f in T, set $\varphi(f) = f'(1/3) + f''(1/2)$. This is a distribution.

Note what a pretty, and simple, idea this is: one "reinterprets" ordinary functions as continuous linear mappings from T to $\mathbf{R}$, an interpretation which leads immediately to "generalized functions."

We now wish to give some examples of properties of distributions. Rule of thumb: first think of things one can do with ordinary functions α, then reexpress that "thing" in terms of the corresponding distribution of the first example in this chapter and thus obtain a "thing" one can do with distributions.

Note that, for α and β continuous functions, and f in T, we have $\int_K \alpha f \, d\mu + \int_K \beta f \, d\mu = \int_K (\alpha + \beta) f \, d\mu$. That is, the linear mapping (from T to $\mathbf{R}$) associated with the sum of functions α and β is the sum of the linear mappings associated with α and β separately. This observation suggests the following definition: for $T \xrightarrow{\varphi} \mathbf{R}$ and $T \xrightarrow{\psi} \mathbf{R}$ two distributions, their *sum*, $\varphi + \psi$, is the distribution given by $(\varphi + \psi)(f) = \varphi(f) + \psi(f)$. Thus addition of distributions generalizes addition of ordinary functions. Similarly, for φ a distribution and a a real number, we denote by $a\,\varphi$ the distribution with action $(a\varphi)(f) = a\varphi(f)$. This generalizes multiplication of functions by numbers. Clearly, under these operations, the set of all distributions (for fixed compact set K) is a vector space.

We can also generalize multiplication of functions by functions. Fix, once and for all, a function h in T. Then note that, for α continuous and f in T, we have $\int_K (h\alpha) f \, d\mu = \int_K \alpha (hf) \, d\mu$. This little formula is the rule for how to relate the distribution defined by α to the distribution defined by $h\alpha$. It suggests how one should define the product of a distribution with h. For $T \xrightarrow{\varphi} \mathbf{R}$ any distribution (and h in T still fixed), let $h\varphi$ denote the distribution with action $(h\varphi)(f) = \varphi(hf)$ (noting that the right side is well defined, for, since h and f are in T, so is hf). This formula indeed gives a continuous linear mapping $T \xrightarrow{h\varphi} \mathbf{R}$, and hence defines a distribution. We call it the *product* of the distribution φ by the function h (in T).

We next define the derivative of a distribution. To see what the definition should be, we again return to the special case of distributions defined by ordinary functions. Suppose now that α is C^1. Then, for f in T, we have $\int_K \alpha' f \, d\mu = -\int_K \alpha f' \, d\mu$ (integrating by parts and using the fact that the support of f is in K to throw away the surface term). Thus: let $T \xrightarrow{\varphi} \mathbf{R}$ be any distribution. We define the *derivative* of this φ, written φ', as the distribution with action $\varphi'(f) = -\varphi(f')$. (The right side makes sense since, for f in T, so is f'.) Note that this derivative is well defined for any distribution—even, for example, one which arises from an ordinary function α which is not differentiable. Does this derivative satisfy the properties one expects of a derivative? It is obvious that, for φ and ψ distributions, and a a number, $(\varphi + a\psi)' = \varphi' + a\psi'$, that is, that the derivative is linear. What about the Leibniz rule? One would like to ask, for example, whether, for two distributions, $(\varphi\psi)' = \varphi\psi' + \psi\varphi'$. Unfortunately, this is not possible, for we have not defined the "product of two distributions" (something which, in fact, can be given no sensible definition). We can, on the other hand, ask the following question: is it true that, for h in T and φ a distribution, $(h\varphi)' = h'\varphi + h\varphi'$? (Note that the right side is a well defined distribution: since h' is in T, $h'\varphi$ is a distribution while $h\varphi'$ is a distribution, and this right side is the sum of distributions.) This is in fact true, as one shows by verifying that the two sides give the same number when applied to any f in T: $(h\varphi)'(f) = -(h\varphi)(f') = \varphi(-hf') = \varphi(h'f - (hf)') = \varphi(h'f) - \varphi(hf)') = (h'\varphi)(f) + \varphi'(hf) = (h'\varphi)(f) + (h\varphi')(f) = (h'\varphi + h\varphi')(f)$.

Example. Let $K = [0,1]$. The derivative of the Dirac delta distribution at 1/3 is the distribution with action $\varphi(f) = -f'(1/3)$.

We next define the support of a distribution. It is convenient to write $T(K)$ instead of T to indicate dependence on the compact set K. Let O be any open subset of K. Then $\mathrm{Cl}(O)$, the closure of O, is a compact subset of K. Just as above, we obtain a topological vector space, $T(\mathrm{Cl}(O))$, of all C^∞ functions with support in $\mathrm{Cl}(O)$. But, since $\mathrm{Cl}(O)$ is a subset of K, every such function also has support in K. Thus every function in $T(\mathrm{Cl}(O))$ is also in $T(K)$, that is, $T(\mathrm{Cl}(O))$ is a subset of $T(K)$ (and, in fact, a closed subspace of the topological vector space $T(K)$). Now suppose we have any distribution, $T(K) \xrightarrow{\varphi} \mathbf{R}$. Then, since $T(\mathrm{Cl}(O))$ is a subset of $T(K)$, we also have a mapping $T(\mathrm{Cl}(O)) \xrightarrow{\varphi} \mathbf{R}$ (in fact, a continuous linear one). We shall say that the distribution φ vanishes on O if this is the zero mapping, that is, if $\varphi(f)$ vanishes for every f in $T(\mathrm{Cl}(O))$. (Of course, this just generalizes the notion of a continuous function α vanishing on O.) The complement of the union of all open subsets O of K on which φ vanishes is called the *support* of φ. (Note that this set, as the complement of an open set, is automatically closed.) It is immediate that this notion of the support of a distribution generalizes that of the support of a function. In fact: let α be a continuous function on K. Then the support of the distribution with action $\int_K \alpha f \, d\mu$ is precisely the support of the function α.

As a final example of structure on the space of distributions, we introduce the notion of the order of a distribution. Fix the compact set K, and write T^{∞} instead of T as above. Next, for n any non-negative integer, write T^n for the vector space of all C^n functions with support in K. We wish to make each of $T^0, T^1, \cdots$ into a topological vector space. We cannot, unfortunately, use the metric of T, for, for example, for f and g in T^2, f''' and g''' may not exist. We therefore choose, for the metric on T^n, that given by the first n terms on the right in the expression near the beginning of this chapter. ("Two C^n functions are close if their values and first n derivatives are close throughout K.) Thus we obtain a sequence, $T^0, T^1, \ldots,$ of topological vector spaces. Next, note that T^{∞} may be regarded as a vector subspace of each of $T^0, T^1, \cdots$ (for a C^{∞} function with support in K is certainly also a C^n function with support in K). Finally, we note that the closure of the subset T^{∞} of T^n is T^n itself (since every C^n function may be approximated, to arbitrary accuracy, by a C^{∞} function).

Now, fix a distribution $T^{\infty} \xrightarrow{\varphi} \mathbf{R}$. Given n, we ask whether it is possible to extend this mapping $T^{\infty} \xrightarrow{\varphi} \mathbf{R}$ to a mapping $T^n \xrightarrow{\tilde{\varphi}} \mathbf{R}$, that is, whether there is a continuous linear mapping $T^n \xrightarrow{\tilde{\varphi}} \mathbf{R}$ such that $\tilde{\varphi}$ agrees with the mapping φ on the subset T^{∞} of T^n. (That is, we ask whether it is possible to extend the action of the distribution, defined originally on C^{∞} functions, to C^n functions.) Such an extension may or may not exist, but if one does exist, it is unique (since the closure of T^{∞} in T^n is all of T^n). Next, we note that, if there is such an extension of $T^{\infty} \xrightarrow{\varphi} \mathbf{R}$ to $T^n \xrightarrow{\tilde{\varphi}} \mathbf{R}$, then there is necessarily an extension to T^{n+1}, T^{n+2}, etc., for, if $T^n \xrightarrow{\tilde{\varphi}} \mathbf{R}$ is an extension to T^n, then, writing $T^{n+1} \xrightarrow{\psi} T^n$ for the natural continuous linear mapping (every C^{n+1} function is also C^n), $T^{n+1} \xrightarrow{\tilde{\varphi}\circ\psi} \mathbf{R}$ is an extension to T^{n+1}. Thus, for every distribution $T^{\infty} \xrightarrow{\varphi} \mathbf{R}$, there is a smallest integer n for which there exists an extension to T^n (or else there may be no extension for any n). We call this n the *order* of the distribution and take the order to be ∞ if no extensions exist. Thus the order of a distribution reflects "how differentiable functions must be in order that the distribution still know how to act on them."

Example. Let $K = [0,1]$. The Dirac delta distribution at some point of K (e.g., the point 1/3) is of order zero, for the mapping $T^{\infty} \xrightarrow{\varphi} \mathbf{R}$ given by $\varphi(f) = f(1/3)$ for f in T^{∞} can be extended to a mapping $T^{0} \xrightarrow{\tilde{\varphi}} \mathbf{R}$ (namely, that given by $\tilde{\varphi}(f) = f(1/3)$ for f in T^{0}). The derivative of the Dirac delta distribution is of order one (for we can extend $\varphi(f) = -f'(1/3)$ for f in T^{1}, but this will not work for an extension to T^{0}, because f' may not exist for f in T^{0}). Any distribution given by $\int_K \alpha f\, d\mu$ with α continuous is of order zero.

The lower the order of a distribution, the "less it requires of the functions on which it is willing to act," that is, the "more nearly like an ordinary function it is." In fact, it is perhaps not too misleading to think of ordinary C^{n} functions as "distributions of order $-n$" (although, mysteriously, the interface between "order 0" and "order -0" does not look right).

Exercise 314. Prove that the order of the sum of two distributions is less than or equal to the maximum of the orders of the summands. Prove that, for φ a distribution and h in T, the order of $h\varphi$ is less than or equal to that of φ. Prove that the order of φ' is one greater than that of φ.

Exercise 315. For which ordinary differential equations is it meaningful to ask for solutions which are distributions?

Exercise 316. Show that the topology on T is independent of the choice of coefficients in the formula for the metric of T.

Exercise 317. Is there a reasonable notion of "$\varphi \leq \psi$" for distributions?

Exercise 318. Are there simple necessary and sufficient conditions on a distribution for it to arise, as in the first example of this chapter, from an ordinary function?

Exercise 319. Let φ be a distribution of order n, and let h be in T^{n}. Define $h\varphi$, and prove that this distribution is of order n.

Exercise 320. Generalize our treatment of distributions to noncompact sets K and to $\mathbf{R} \times \cdots \times \mathbf{R}$.

Exercise 321. Why does one bother to introduce a topology on T at all?

Exercise 322. Define the Fourier transform of a distribution.

Exercise 323. Let $K = [0,1]$, and let φ be a distribution whose support is the single point $1/2$ of K. Does it follow that φ is a linear combination of the Dirac delta distribution at $1/2$ and its derivatives?

Exercise 324. Prove that, for S the support of any continuous function, $\mathrm{Cl}(\mathrm{Int}(S)) = S$. Is this true for the support of any distribution?

48

Hilbert Spaces

A *Hilbert space* consists of two things—i) a complex vector space H, and ii) a rule that assigns, to any two vectors h and h' in H, a complex number (written (h,h') and called the *inner product* of h and h')—subject to the following four conditions:

1. For any vectors h, h', and h'', and any complex number c, we have

$$(h + ch',h'') = (h,h'') + \bar{c}(h',h'')$$

and

$$(h,h' + ch'') = (h,h') + c(h,h'') \quad ,$$

where a bar over a complex number denotes the complex-conjugate of that number.

2. For any vectors h and h', we have

$$\overline{(h,h')} = (h',h) \quad .$$

3. For any nonzero vector h, the real (by the previous condition) number (h,h) is positive.

4. The topological vector space H is complete, as described below.

To complete this definition, we must explain what the fourth condition means. It is convenient, however, to postpone this explanation for a moment in order to make a few observations. For any vector h in H, set $||h|| = [(h,h)]^{1/2}$ (a real number, by condition 3, and, since we now agree to take the positive square root, a non-negative one), and call this number the *norm* of h. This norm satisfies the properties one might expect from its name. For example, for any vectors h and h', we have $|(h,h')| \leq ||h||\ ||h'||$ [proof: for any real number r, we have $0 \leq ||h + rh'||^2 = (h + rh',h + rh') = (h,h) + r(h,h') + r(h',h) + r^2(h',h') \leq ||h||^2 + 2r|(h,h')| + r^2||h'||^2$; thus this polynomial in r can never become negative, a statement equivalent to the desired inequality] and $||h + h'|| \leq ||h|| + ||h'||$ [proof: $||h + h'||^2 = ||h||^2 + (h,h') + (h',h) + ||h'||^2 \leq ||h||^2 + 2|(h,h')| + ||h'||^2 \leq ||h||^2 + 2||h||\ ||h'|| + ||h'||^2 = (||h|| + ||h'||)^2$]. Next, for h and h' any vectors, set $d(h,h') = ||h - h'||$. We claim that this $d(\ ,\ )$ is a metric on the set H. [Proof: (1) By the third condition, $d(h,h')$ is non-negative and vanishes only when $h = h'$. (2) $d(h,h') = (h - h',h - h')^{1/2} = (h' - h,h' - h)^{1/2} = d(h',h)$. (3) $d(h,h') + d(h',h'') = ||h - h'|| + ||h' - h''|| \leq ||(h - h') + (h' - h'')|| = ||h - h''|| = d(h,h'')$.] Thus we now have a metric on the set H, and hence a topology on

H. We now claim, finally, that the complex vector space H, with this topology, is a topological vector space. [Proof: Since, for $d(h,\underline{h}) \leq \epsilon$ and $d(h',\underline{h}') \leq \epsilon$, we have $d(h+h',\underline{h}+\underline{h}') \leq 2\epsilon$, addition is continuous. Since, for $|c-\underline{c}| \leq \epsilon$ and $d(h,\underline{h}) \leq \epsilon$, we have $d(ch,\underline{c}\underline{h}) \leq \epsilon\,||h|| + \epsilon|\underline{c}|$, multiplication of vectors by complex numbers is continuous.] Thus we have a topological vector space H. (Statements within H referring to a topology hereafter refer to this one.)

The fourth condition in our definition requires that the topological vector space H, with the topology just constructed, be complete. One can, in fact, restate this fourth condition (inserting directly all the various definitions—of a topological space, of a topological vector space, of a uniform space, of a complete uniform space—that underlie it) as follows:

4′. Consider any sequence, $h_1,h_2, \ldots$, of vectors in H with the following property: given any positive ϵ, there is a number N such that $||h_n - h_m|| \leq \epsilon$ whenever $n \geq N$ and $m \geq N$. Then there is a vector h in H such that $\lim||h - h_n|| = 0$.

A number of facts follow immediately from the definition. The first condition requires that, for fixed h, the complex number (h,h') is linear in h' and that, for fixed h', the complex number (h,h') is antilinear (i.e., linear, except that one must take complex-conjugates of coefficients in linear combinations) in h. Note also that, for c any complex number and h any vector, $||ch|| = |c|\,||h||$ (for $||ch||^2 = (ch,ch) = c(ch,h) = c\bar{c}(h,h) = |c|^2\,||h||^2$). In particular, the norm of the zero vector is zero (and only the zero vector has this property). It is sometimes convenient to make use of the fact that, if one knows what the norms of the vectors in H are, then one already knows what the inner products are. This claim is immediate from the formula $(h,h') = (1/4)[||h+h'||^2 - ||h-h'||^2 - i||h+ih'||^2 + i||h-ih'||^2]$, which is easily verified by expanding the right side. We next remark that the operation "take the inner product" (regarded as a mapping $H \times H \rightarrow \mathbf{C}$) is continuous. Indeed, for $||h-\underline{h}|| \leq \epsilon$ and $||h'-\underline{h}'|| \leq \epsilon$, we have $|(h,h') - (\underline{h},\underline{h}')| = |(h,h') - (\underline{h},h') + (\underline{h},h') - (\underline{h},\underline{h}')| = |(h-\underline{h},h') - (\underline{h},h'-\underline{h}')| \leq |(h-\underline{h},h')| + |(\underline{h},h'-\underline{h}')| \leq ||h-\underline{h}||\,||h'|| + ||\underline{h}||\,||h'-\underline{h}'|| \leq \epsilon||h'|| + \epsilon||\underline{h}||$. (Intuitively, "if h is close to $\underline{h}$ and h' is close to $\underline{h}'$, then (h,h') is close to $(\underline{h},\underline{h}')$.") Thus, in particular, if $h_1,h_2,\cdots$ is a sequence of vectors that converges (in the topology on H, of course) to vector h, then for any vector h' we have $\lim(h',h_n) = (h',h)$. Note, incidentally, that the statement that $h_1,h_2,\cdots$ converges to h is just the statement that $\lim||h-h_n|| = 0$.

As one might expect (e.g., from the complexity of the definition), examples of Hilbert spaces play an important role.

Example. Fix a measure space X, $\mathbf{M}$, $\mathbf{M} \xrightarrow{\mu} R^*$. Denote by $\mathbf{F}$ the collection of all complex-valued, measurable functions f on the measure space X (i.e., functions $X \xrightarrow{f} \mathbf{C}$ whose real and imaginary parts are both measurable) that are square-integrable (i.e., which are such that the function $\bar{f}f$ is integrable on X). Given two such functions, f and f', in $\mathbf{F}$, write $f \approx f'$ if f and f' are equal almost everywhere. Note that "$\approx$" is an equivalence relation on the set $\mathbf{F}$. Denote by H the set of equivalence classes. Given any complex number c and any elements h and h' of H, with representatives f and f', let $h + ch'$ be the element of H with representative $f + cf'$ [noting that, i) this prescription is independent of choice of representatives, ii) $f + cf'$ is measurable if both f and f' are, and iii) $f + cf'$ is square-integrable if both f and f' are, since $\overline{(f + cf')}(f + cf') \le 4(\bar{f}f + |c|^2\bar{f}'f')$]. With this definition of addition in H and of multiplication of elements of H by complex numbers, we obtain a complex vector space H. We next introduce an inner product. Given elements h and h' of H, with representatives f and f', we set $(h,h') = \int_X \bar{f}f' \, d\mu$, where the integral of a complex-valued function on X is defined as the integral of its real part plus i times the integral of its imaginary part [noting that, i) $\bar{f}f'$ is integrable on X if f and f' are square-integrable, since $|\bar{f}f'| \le \bar{f}f + \bar{f}'f'$, and ii) the value of the integral is independent of the choice of representatives, since functions equal almost everywhere have the same integral]. Thus we have a complex vector space H and, for any two elements h and h' of H, a complex number (h,h').

THEOREM 55. *The H, (,) of the paragraph above is a Hilbert space.*

Proof. (1) The behavior of the inner product under linear combinations is immediate from $\int_X \overline{(f + cf')}f'' \, d\mu = \int_X \bar{f}f'' \, d\mu + \bar{c}\int_X \bar{f}'f'' \, d\mu$ and $\int_X \bar{f}(f' + cf'') \, d\mu = \int_X \bar{f}f' \, d\mu + c\int_X \bar{f}f'' \, d\mu$. (2) The behavior of the inner product under complex conjugation is immediate from $\overline{\int_X \bar{f}f' \, d\mu} = \int_X \bar{f}'f \, d\mu$. (3) Let h be an element of H, with representative f. Then $(h,h) = \int_X \bar{f}f \, d\mu$ is certainly non-negative, since the function $\bar{f}f$ is non-negative. Furthermore, this integral vanishes when and only when $\bar{f}f$ vanishes almost everywhere, that is, only when f vanishes almost everywhere, that is, only when $f \approx 0$, the zero function, that is, only when the equivalence class h contains the zero function, that is, only when $h = 0$. (4) Let $f_1, f_2, \cdots$ be a sequence of measurable, complex-valued, square-integrable functions on X satisfying the following condition: for any positive ϵ, there is a number N such that $\int_X |f_n - f_m|^2 d\mu \le \epsilon$ whenever $n \ge N$ and $m \ge N$. We must find a measurable, complex-valued, square-integrable function f with $\lim \int_X |f - f_n|^2 d\mu = 0$. There is no loss in generality in assuming $\int_X |f_n - f_{n+1}|^2 d\mu \le 1/2^n$, for, were this not satisfied by our original sequence, it is easy to select a subsequence for which this would be satisfied. Fix positive ϵ, and denote by A_ϵ the collection of all points x of X such that, for every number N, $|f_n(x) - f_m(x)| \ge \epsilon$ for some $n \ge N$ and $m \ge N$. Since $\int_X |f_n - f_{n+1}|^2 d\mu \le 1/2^n$, the measure of the set of all x with $|f_n(x) - f_{n+1}(x)| \ge \epsilon$ is less than $1/\epsilon^2 2^n$. Therefore the measure of the set of all x with $|f_n(x) - f_m(x)| \ge \epsilon$ for some $n \ge N$ and $m \ge N$ is less than or equal to

$8/\epsilon^2 2^N$. Thus $\mu(A_\epsilon) \leq 8/\epsilon^2 2^N$ for all N, whence $\mu(A_\epsilon) = 0$. Next, set $A = A_{1/2} \cup A_{1/4} \cup A_{1/8} \cdots$, so $\mu(A) = 0$. But A is precisely the set of points x for which the sequence $f_1(x), f_2(x), \cdots$ of complex numbers is not a Cauchy sequence. That is, the sequence $f_1(x), f_2(x), \cdots$ of complex numbers converges for x not in A. Let $X \xrightarrow{f} \mathbf{C}$ be the function with $f(x)$ this number to which the above sequence converges, if x is not in A, and (say), set $f(x) = 0$ for x in A. This function f is clearly measurable. Next, note that, for $m \geq n$, $\int_X |f_n - f_m|^2 d\mu \leq 2/2^n$, whence, taking the limit $m \to \infty$, $\int_X |f_n - f| d\mu \leq 2/2^n$. Thus f (as the difference of square-integrable functions f_n and $f_n - f$) is square-integrable, and $\lim \int_X |f - f_n|^2 d\mu = 0$. □

As far as I am aware, every Hilbert space which arises in applications is essentially a special case of this example (or, occasionally, a minor modification of the example). In particular, the case when the measure space is the measure space of reals (so the Hilbert space is that of measurable, complex-valued, square-integrable functions on the reals, with two such identified if they agree almost everywhere) is common. Another common special case is the following.

Example. Let X be the set of positive integers. Let every subset of X be measurable, and, for A such a subset, let $\mu(A)$ be the number of points in A (if that number is finite; let $\mu(A) = \infty$ if A contains an infinite number of points). This is a measure space. A complex-valued function $X \xrightarrow{f} \mathbf{C}$ can be represented as a sequence, $(c_1, c_2, \cdots)$, of complex numbers (where c_1 is the value of f at the point "1" of X, etc.). Since every subset of X is measurable, every such function is measurable. The integral of such a function is, of course, the sum of its values. Thus the function represented by $(c_1, c_2, \cdots)$ is square-integrable provided $|c_1|^2 + |c_2|^2 + \cdots$ converges. The Hilbert space H is therefore the set of all sequences, $(c_1, c_2, \cdots)$, of complex numbers for which the sum $|c_1|^2 + |c_2|^2 + \cdots$ converges. (Note that we need not in this example take equivalence classes: since the only subset of X of measure zero is the empty subset, two functions equal almost everywhere are equal everywhere.) For linear combinations of such sequences, one takes the corresponding linear combinations of the corresponding functions: $(c_1, c_2, \cdots) + c(c_1', c_2', \cdots) = (c_1 + cc_1', c_2 + cc_2', \cdots)$. Finally, for the inner product of two such sequences, we have $((c_1, c_2, \cdots), (c_1', c_2', \cdots)) = \overline{c_1} c_1' + \overline{c_2} c_2' + \cdots$ One could, of course, show directly, that is, without recourse to theorem 55, that this is a Hilbert space.

Let H be a Hilbert space. A subset V of H is said to be a ***subspace*** of this Hilbert space if i) V is a subspace of vector space H (i.e., any linear combination of vectors in V is again in V) and ii) V is a closed subset of topological space H.

Example. Let H be the Hilbert space of all square-integrable functions on measure space X, and let A be a measurable subset of X. Denote by V the collection of all elements of H having representatives that vanish on A. Then V is a subspace of Hilbert space H. For H the Hilbert space of all square-integrable functions on the measure space of reals and V the subset consisting of elements of H having representatives which are C^1, V is not a subspace of Hilbert space H (for, although V is a subspace of the vector space, V is not closed).

Next, note that any intersection of subspaces of Hilbert space H is a subspace (since intersections of vector subspaces yield a vector subspace and intersections of closed subsets yield a closed subset). Thus, given any subset K of H, the intersection of all subspaces of H that are supersets of K is a subspace of H; clearly, the smallest one which is a superset of K. This subspace of H is called the subspace *generated* by K. Note also that every subspace of Hilbert space H is itself a Hilbert space, where addition, scalar multiplication, and inner products are those inherited from H. (The first three conditions for a Hilbert space are immediate. The fourth follows from the fact that a closed subset of a complete uniform space is complete.)

Two vectors, h and h', in a Hilbert space are said to be *orthogonal* if $(h,h') = 0$. Thus, for example, the zero vector is orthogonal to every vector and only the zero vector is orthogonal to itself. Now let V be a subspace of Hilbert space H. We denote by $V^\perp$ the set of all vectors h in H that are orthogonal to every vector in V. Perhaps the most important single result about subspaces is the following.

THEOREM 56. *Let V be a subspace of Hilbert space H. Then $V^\perp$ is also a subspace. Furthermore, V and $V^\perp$ are complementary, that is, every vector in H can be written in one and only one way as the sum of one vector from V and one from $V^\perp$.*

Proof. Let h and h' be any vectors in $V^\perp$, and c any complex number. Then, for v in V, $(h + ch',v) = (h,v) + \bar{c}(h',v) = 0$, so $h + ch'$ is in $V^\perp$. Thus $V^\perp$ is a vector subspace of vector space H. Since the inner product is continuous, any limit h of vectors h_λ (λ in Λ) in $V^\perp$ satisfies $(h,v) = 0$ for all v in V, whence this limit is in $V^\perp$. Thus $V^\perp$ is closed, and so $V^\perp$ is a subspace of Hilbert space H. We have only to show that V and $V^\perp$ are complementary. Fix vector h in H, and denote by κ the greatest lower bound of numbers $||h - v||$ for v in V. Choose vectors $v_1, v_2, \cdots$ in V with $\lim ||h - v_n|| = \kappa$ (figure 139). By direct computation, we have $||v_n - v_m||^2 = 2||h - v_n||^2 + 2||h - v_m||^2 - 4||h - (1/2)(v_n + v_m)||^2$. By choosing m and n sufficiently large, we can make each of the first two terms as near $2\kappa^2$ as we wish. Since $(1/2)(v_n + v_m)$ is in V, we have, by definition of κ, that $||h - (1/2)(v_n + v_m)||^2 \geq \kappa^2$. Thus $v_1, v_2, \cdots$ is a Cauchy sequence, whence it must converge to some vector v. This v must be in V (since V is closed), and we have $||h - v||$

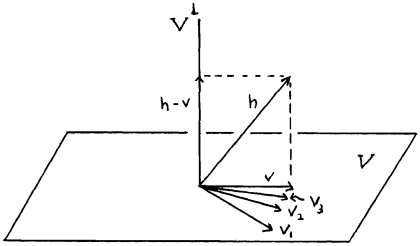

Figure 139

$= \lim||h - v_n|| = \kappa$. We next show that $h - v$ is in $V^\perp$. For any v' in V, set $\alpha = (h - v, v')$. Then, for every number ϵ, we must have $||h - v - \epsilon v'||^2 = \kappa^2 - \epsilon\alpha - \overline{\epsilon\alpha} + |\epsilon|^2||v'||^2 \geq \kappa^2$ by definition of κ. But this inequality can hold for every ϵ only if $\alpha = 0$. Thus every vector h in H can be written as the sum of a vector v (obtained above) in V and a vector, $h - v$, in $V^\perp$. This decomposition is unique, for any vector in both V and $V^\perp$ must be orthogonal to itself, and hence must be the zero vector. □

Example. Let H be the Hilbert space of square-integrable functions on measure space X. Let V be the subspace consisting of elements of H having representatives that vanish on some measurable set A. Then $V^\perp$ is the subspace consisting of elements having representatives that vanish on A^c.

There are many applications of theorem 56, of which we give three as examples. Let V be a subspace of Hilbert space H. Then $V^{\perp\perp} = V$. First note that $V \subset V^{\perp\perp}$ (for every vector in V is certainly orthogonal to everything in $V^\perp$, and hence is in $V^{\perp\perp}$). To obtain the reverse inclusion, let h be a vector in $V^{\perp\perp}$. Then $h = h' + h''$, with h' in V and h'' in $V^\perp$. But $V \subset V^{\perp\perp}$, so we must have h' in $V^{\perp\perp}$. Since $V^\perp$ and $V^{\perp\perp}$ are complementary and since h was already in $V^{\perp\perp}$, we must have $h'' = 0$ and $h' = h$. Hence h is in V. Thus $V^{\perp\perp} \subset V$, whence $V^{\perp\perp} = V$.

Fix a vector $\underline{h}$ in H. Then the mapping from H to $\mathbf{C}$ which sends h to $(\underline{h}, h)$ is certainly continuous and linear. As a second example of theorem 56, we prove the converse: if $H \xrightarrow{\varphi} \mathbf{C}$ is any continuous linear mapping, then there exists a vector $\underline{h}$ with $\varphi(h) = (\underline{h}, h)$. Denote by V the subset of H consisting of all vectors v with $\varphi(v) = 0$. Then V (as the kernel of a linear mapping) is a subspace of vector space H, while V (as the inverse image of a closed set—that consisting only of zero—under a continuous mapping) is closed. Hence V is a subspace of Hilbert space H. Now consider $V^\perp$. If $V^\perp = 0$, then $V = (V^\perp)^\perp = H$, whence φ is the zero mapping, whence $\underline{h} = 0$ will do the job. So suppose that $V^\perp$ is not the zero subspace: choose vector $\tilde{h}$ in V with $\varphi(\tilde{h}) = 1$. Then, for any h' in $V^\perp$, $h' - \varphi(h')\tilde{h}$ is in $V^\perp$. But $\varphi(h' - \varphi(h')\tilde{h}) = \varphi(h') -$

$\varphi(h')\varphi(\tilde{h}) = 0$, so $h' - \varphi(h')\tilde{h}$ is also in V. Hence we must have $h' - \varphi(h')\tilde{h} = 0$, whence h' is a multiple of $\tilde{h}$. In other words, $V^{\perp}$ is one-dimensional. We now claim that $\varphi(h) = (\tilde{h}/||\tilde{h}||,h)$ for every h. Indeed, writing $h = \alpha\tilde{h} + v$, with v in V, we have $\varphi(h) = \varphi(\alpha\tilde{h} + v) = \alpha\varphi(\tilde{h}) + \varphi(v) = \alpha$ and $(\tilde{h}/||\tilde{h}||,h) = (\tilde{h}/||\tilde{h}||,\alpha\tilde{h} + v) = \alpha(\tilde{h}/||\tilde{h}||,\tilde{h}) + (\tilde{h}/||\tilde{h}||,v) = \alpha(\tilde{h}/||\tilde{h}||,\tilde{h}) = \alpha$. This result is called the Riesz representation theorem.

For the third example, we need a definition. Let H be a Hilbert space. A subset of H is called an *orthonormal basis* for H if i) every vector in this subset has unit norm, ii) any two distinct vectors in this subset are orthogonal to each other, and iii) this subset generates the entire Hilbert space H. (In more detail, the last condition means that, given any vector h and any positive number ϵ, there is a vector h' that is a finite linear combination of vectors in our subset and that satisfies $||h - h'|| \leq \epsilon$.)

Example. Consider the Hilbert space of the example following theorem 55. Consider the sequence $h_1 = (1,0,\cdots)$, $h_2 = (0,1,0,\cdots)$, $h_3 = (0,0,1,0,\cdots)$, $\cdots$ in H. This is an orthonormal basis.

Our third example of an application of theorem 56 is this: every Hilbert space possesses an orthonormal basis. To prove this assertion, denote by $\mathbf{K}$ the collection of all subsets K of H such that every vector in K has unit norm and such that any two distinct vectors in K are orthogonal. Order $\mathbf{K}$ by inclusion. Any totally ordered subset of $\mathbf{K}$ has an upper bound (namely, the element of $\mathbf{K}$ obtained by taking the union of all the K in that totally ordered subset). Thus the conditions of Zorn's lemma are satisfied. Let K be a maximal element of $\mathbf{K}$ (so K is an orthonormal subset of H and there is no bigger orthonormal subset). Let V denote the subspace of H generated by this maximal K. Then $V^{\perp}$ must be the zero subspace, for, were there a nonzero vector in $V^{\perp}$, there would be a vector with unit norm and we could include this vector in K, violating maximality. Since $V^{\perp} = 0$, $V = V^{\perp\perp} = H$. Thus K generates H, so K is an orthonormal basis.

This last result already makes a highly nontrivial statement about measure spaces. Let X be a measure space. Then we can find a collection f_λ (λ in Λ) of measurable, square-integrable functions on X such that $\int_X \bar{f}_\lambda f_{\lambda'}\, d\mu$ is one if $\lambda = \lambda'$, and zero otherwise, and such that any measurable, square-integrable function f can be approximated, in the sense of $\int_X |f - f'|^2 d\mu$, to within arbitrary accuracy by some finite linear combination of the f_λ.

Exercise 325. Let H be a Hilbert space, and let $H \xrightarrow{\varsigma} \mathbf{R}$ be given by $\varsigma(h) = ||h||$. Prove that ς is continuous.

Exercise 326. Let K be an orthonormal basis for Hilbert space H. Prove that K is linearly independent. Show that K is a basis for vector space H (in the sense defined in chapter 9) if and only if H is finite-dimensional.

Exercise 327. Let $h_1, h_2, \cdots$ be an orthonormal basis for Hilbert space H. Let h be any vector in H. Prove that there exists a sequence $c_1, c_2, \cdots$ of complex numbers such that the sequence $\underline{h}_1, \underline{h}_2, \cdots$ in H, with $\underline{h}_n = c_1 h_1 + c_2 h_2 + \cdots + c_n h_n$, approaches the vector h. Prove that the choice $c_n = (h_n, h)$ is the unique one for this to be true.

Exercise 328. Find a Hilbert space H and a linear mapping $H \rightarrow \mathbf{C}$ that is not continuous.

Exercise 329. Find a Hilbert space H and a sequence $h_1, h_2, \cdots$ such that $\lim(h_n, h) = 0$ for every h but such that this sequence does not approach the zero vector.

Exercise 330. Let H be the Hilbert space of square-integrable functions on the measure space of reals. Define $H \xrightarrow{\varsigma} \mathbf{C}$ by $\varsigma(f) = \int_{\mathbf{R}} f e^{ikx} d\mu$, where k is any number. This ς is apparently continuous and linear, and apparently not of the form $\varsigma(h) = (\underline{h}, h)$. What is wrong?

Exercise 331. Define a free Hilbert space on a set. (Hint: Let the set be a measure space, with every subset measurable, and the measure of a subset the number of its elements.) Prove that every Hilbert space is free.

Exercise 332. An *isomorphism* of Hilbert spaces, $H \xrightarrow{\varphi} G$, is a linear mapping that is one-to-one and onto, and with $(\varphi(h), \varphi(h')) = (h, h')$. Prove that the free Hilbert spaces on sets K and K' are isomorphic if and only if these sets are isomorphic.

Exercise 333. Let H and G be Hilbert spaces. Consider $H \oplus G$, the direct sum of vector spaces. For (h,g) and (h',g') in $H \oplus G$, set $((h,g),(h',g')) = (h,h') + (g,g')$ (sum of inner products on the right). Prove that $H \oplus G$, (,) is a Hilbert space. (It is called the ***direct sum*** of Hilbert spaces H *and* G.) *Define the tensor product of Hilbert spaces.*

Exercise 334. Prove that every Hilbert space is homogeneous, that is, that, given vectors h and h' with the same norm, there is an isomorphism that sends h to h'.

Exercise 335. Let H be the Hilbert space of square-integrable functions on measure space X. Is it true that every subspace of H arises as in the text, from some measurable subset of X?

Exercise 336. Let H be a complex vector space with an inner product, satisfying the first three conditions in the definition of a Hilbert space. Then H can be regarded as a topological vector space. Show how to extend the inner product to the completion of H (as a topological vector space) and that the result is a Hilbert space.

49

Bounded Operators

One class of things that live in the environment of a Hilbert space—a class we shall be studying for a while—is that of operators. In this chapter we define and discuss the properties of the "most well-behaved" operators, the bounded ones.

Fix a Hilbert space H. A *bounded operator* (on H) is a continuous linear mapping $H \to H$. Denote by **B** the set of all bounded operators on H.

Example. Let X be a measure space, and consider the Hilbert space of measurable, square-integrable, complex-valued functions on X (a Hilbert space we hereafter denote by $L^2(X)$). Choose any bounded, measurable, complex-valued function α on X. With each square-integrable f, associate the function αf. Note that αf is measurable and (since $|\alpha f|^2 \leq M^2 |f|^2$, where M is an upper bound for $|\alpha|$) square-integrable and that, if f and f' are equal almost everywhere, so are αf and $\alpha f'$. Thus this function α defines a mapping, which we denote by A_α, from $L^2(X)$ to $L^2(X)$. Since $\alpha(f + cf') = \alpha f + c\alpha f'$, this mapping is linear. Since $\int_X |f - f'|^2 d\mu \leq \epsilon$ implies $\int_X |\alpha f - \alpha f'|^2 d\mu \leq M^2 \epsilon$ (i.e., since nearby elements of $L^2(X)$ are taken, by A_α, to nearby elements of $L^2(X)$), this mapping A_α is continuous. Thus we have $L^2(X) \xrightarrow{A_\alpha} L^2(X)$, a continuous linear mapping. That is, we have a bounded operator on $L^2(X)$.

It is nowhere near being true that every bounded operator on $L^2(X)$ is one of the A_α obtained above. Nonetheless, these particular bounded operators are simple and (since they refer to bounded, complex, measurable functions on X) rather "explicit." We shall refer to this particular example a number of times in what follows.

Our immediate goal is the following: we wish to find whatever structure we can on this set **B** of bounded operators on Hilbert space H. It turns out that the structure available is remarkably rich: one finds algebraic structure, topological structure, and adjoint structure. Each "type of structure" comes with its own properties, and, in addition, there are properties which relate several structures.

Let A and B be bounded operators on Hilbert space H. Then $A + B$, defined by the action $(A + B)(h) = A(h) + B(h)$, is also a bounded operator. [Proof: This $A + B$ is clearly a linear mapping from H to H. To see that it is continuous, consider $H \to H \times H \to H \times H \to H$, where the first mapping sends h to (h,h), the second (h,h') to $(A(h),B(h'))$, and the third (h,h') to

$h + h'$. Each of these mappings is continuous, and their composition is $A + B$.] Next, for A a bounded operator and c a complex number, let cA be the bounded operator with action $(cA)(h) = c(A(h))$. We can thus add elements of **B** and multiply elements of **B** by complex numbers. Clearly, this set **B** thus acquires the structure of a complex vector space. Next, note that, for A and B bounded operators on H, so is their composition, which we write AB (for the composition of two linear mappings is linear and the composition of two continuous mappings is continuous). It is immediate that this composition is associative (i.e., $A(BC) = (AB)C$) and linear (i.e., $(A + cB)C = AC + cBC$ and $A(B + cC) = AB + cAC$). Thus we have a vector space on which there is given a linear associative product. That is, **B** now has the structure of a (complex) associative algebra. This is the "algebraic structure" on **B**.

Example. Consider the bounded operators A_α on $L^2(X)$. Then, clearly, $A_\alpha + A_\beta = A_{\alpha+\beta}$ and $cA_\alpha = A_{c\alpha}$. In other words, addition of bounded operators, and multiplication of bounded operators by numbers, reflects addition of (bounded, measurable) functions, and multiplication of functions by numbers. Furthermore, $A_\alpha A_\beta = A_{\alpha\beta}$, that is, composition of bounded operators reflects multiplication of functions. (Note, however, that $A_\alpha A_\beta = A_\beta A_\alpha$, while, on the other hand, it is false in general that $AB = BA$ for arbitrary bounded operators.)

Example. Let H be a finite-dimensional Hilbert space, with orthonormal basis $h_1, h_2, \ldots, h_n$. Then any $n \times n$ matrix, with complex entries a_{ij} $(i,j = 1,2, \ldots, n)$ defines a bounded operator A, with the following action. For $h = c_1h_1 + \cdots + c_nh_n$ in H, $A(h) = d_1h_1 + \cdots + d_nh_n$, where $d_i = \sum_{j=1}^{n} a_{ij}c_j$. Then linear combinations of bounded operators correspond to linear combinations of matrices, and composition of operators to multiplication of matrices.

For the next structure on the set **B**, we have to use the fact that our bounded operators are continuous. Let $H \xrightarrow{A} H$ be a bounded operator. For any positive number ϵ, denote by N_ϵ the neighborhood of the zero vector consisting of vectors h with $||h|| \leq \epsilon$. Now, we have $A(0) = 0$ (by linearity), whence, by continuity of A, there is a positive ϵ with $A[N_\epsilon] \subset N_1$. It follows that $A[N_1] \subset N_{1/\epsilon}$ (for, for $||h|| \leq 1$, we have $||\epsilon h|| \leq \epsilon$, whence $||A(\epsilon h)|| \leq 1$, whence $||A(h)|| \leq 1/\epsilon$). The least upper bound of real numbers r with $A[N_1] \subset N_r$ is called the *norm* of A and is written $|A|$. (This least upper bound exists since, by the argument above, $A[N_1] \subset N_r$ for some r.) We can restate this definition in the following form: the value of $|A|$ is the smallest real number r such that $||A(h)|| \leq r||h||$ for every vector h. (Intuitively, $|A|$ is the "maximum expansion factor you can get on any vector by applying A to it.") This definition is the origin of the adjective "bounded" in "bounded operator." Thus, with each element A of **B**, we associate a non-negative real number $|A|$.

Example. Consider the bounded operators A_α on $L^2(X)$. Let M be an upper bound for $|\alpha|$. Then, since $\int_X |\alpha f|^2 d\mu \leq M^2 \int_X |f|^2 d\mu$, we clearly must have $|A_\alpha| \leq M$. One might imagine therefore that the least upper bound of $|\alpha|$ will be the norm of A_α. This is "almost true." Consider, however, a function α which vanishes almost everywhere and which assumes the value 3 on some nonempty set of measure zero. Then the least upper bound of $|\alpha|$ is 3, while $|A_\alpha| = 0$ (since always $\int_X |\alpha f|^2 d\mu = 0$). We therefore proceed as follows. Let α be a measurable, bounded function on X. By the almost everywhere least upper bound of α, we mean the smallest real number M such that $|\alpha| \leq M$ almost everywhere. We show that $|A_\alpha| = M$, where M is the almost everywhere least upper bound of α. Clearly, $|A_\alpha| \leq M$. To establish the reverse inequality, fix any positive ϵ. Let f be the function with $f(x) = 1$ for x with $|\alpha(x)| \geq M - \epsilon$, and zero otherwise. Then, since f takes the value one on a set of nonzero measure, $\int_X |f|^2 d\mu > 0$. Furthermore, since $|\alpha f|^2 \geq (M - \epsilon)^2 |f|^2$, we have $\int_X |\alpha f|^2 d\mu \geq (M - \epsilon)^2 \int_X |f|^2 d\mu$. Thus $|A_\alpha| \geq M - \epsilon$. Since ϵ is arbitrary, $|A_\alpha| \geq M$. We conclude that $|A_\alpha| = M$, that is, that the norm of A_α is precisely the almost everywhere least upper bound of α.

We now wish to obtain a few properties of norms of bounded operators. First, note that, for A and B bounded operators, $|A + B| \leq |A| + |B|$ (for, for h any vector, $||(A + B)(h)|| = ||A(h) + B(h)|| \leq ||A(h)|| + ||B(h)|| \leq |A|\,||h|| + |B|\,||h|| = (|A| + |B|)||h||$). Furthermore, for A any bounded operator and c any number, $|cA| = |c|\,|A|$ (for, for any h, $||(cA)(h)|| = |c|\,||A(h)||$). Finally, for A and B bounded operators, $|AB| \leq |A|\,|B|$ (for, for h any vector, $||AB(h)|| = ||A(B(h))|| \leq |A|\,||B(h)|| \leq |A|\,|B|\,||h||$). Note that the first two of these properties are the same as those for the norms of vectors in our Hilbert space (and the last one would be too, except that we do not know how to take the "product" of two vectors). In particular, these three properties of the norm, together with the previous example, establish the following elementary facts: i) the almost everywhere least upper bound of $\alpha + \beta$ is less than or equal to the almost everywhere least upper bounds of α and β, ii) the almost everywhere least upper bound of $c\alpha$ equals c times the almost everywhere least upper bound of α, iii) the almost everywhere least upper bound of $\alpha\beta$ is less than or equal to the product of the almost everywhere least upper bounds of α and β.

Next, for bounded operators A and B, write $d(A,B) = |A - B|$. We claim that this is a metric on the set **B**. [Proof: (1) If $|A - B| = 0$, $||(A - B)(h)|| \leq |A - B|\,||h|| = 0$ for every h, whence $(A - B)(h) = 0$ for every h, whence $A = B$. (2) $|A - B| = |B - A|$. (3) $|A - B| + |B - C| \leq |(A - B) + (B - C)| = |A - C|$.] Thus we have a metric on the set **B**, so this set has the structure of a topological space. Now this set **B** is both a topological space and a vector space: we claim that it is in fact a topological vector space [the proof is identical to that for Hilbert spaces at the beginning of chapter 48]. We in fact wish to claim even more than this, namely, that the topological vector space **B** is complete. Let $A_1, A_2, \cdots$ be a Cauchy sequence in **B**, so, for every positive ϵ, there is a number N such that $|A_n - A_m| \leq \epsilon$ whenever $n \geq N$ and $m \geq N$. We show that this Cauchy sequence converges to some bounded operator. For each h in H, $||A_n(h) - A_m(h)|| \leq |A_n - A_m| \, ||h||$, whence the sequence $A_1(h), A_2(h), \cdots$ of vectors in H is a Cauchy sequence in H. Since H is complete, this sequence converges to some vector in H, which we write $A(h)$, thus defining a mapping $H \xrightarrow{A} H$. This mapping A is clearly linear. To see that A is continuous, note that $||A(h) - A(h')|| = ||A(h) - A_n(h) + A_n(h) - A_n(h') + A_n(h') - A(h')|| \leq ||A(h) - A_n(h)|| + |A_n| \, ||h - h'|| + ||A_n(h') - A(h')||$. We can make the first and third terms on the right as small as we wish by choosing n sufficiently large, and the second term as small as we wish by choosing $||h - h'||$ sufficiently small (noting that $|A_n|$ is bounded as n varies). Thus this mapping A is in fact a bounded operator. Finally, we must show that A is the limit of $A_1, A_2, \ldots$, that is, that $\lim|A - A_n| = 0$. Given positive ϵ, choose N so $|A_n - A_m| \leq \epsilon$ for $n \geq N$ and $m \geq N$. Then $||(A_n - A)(h)|| = \lim\limits_{m} ||A_n - A_m)(h)|| \leq \lim |A_n - A_m| \, ||h|| \leq \epsilon ||h||$ for $n \geq N$. Thus $|A_n - A| \leq \epsilon$ for $n \geq N$. That is, $\lim |A_n - A| = 0$. Thus the topological vector space **B** is complete. We note in addition that the operation of composition (regarded as a mapping $\mathbf{B} \times \mathbf{B} \to \mathbf{B}$) is continuous, for, for $|A - A'| \leq \epsilon$ and $|B - B'| \leq \epsilon$, $|AB - A'B'| = |AB - A'B + A'B - A'B'| \leq |A - A'| \, |B| + |A'| \, |B - B'| \leq \epsilon|B| + \epsilon|A'|$.

Finally, we remark that $|A|$ also has the following alternative interpretation: $|A|$ is precisely the smallest number r such that $|(A(h),h')| \leq r||h|| \, ||h'||$ for any vectors h and h'. Proof: Let r_0 be this smallest number. Since $|(A(h),h')| \leq ||A(h)|| \, ||h'|| \leq |A| \, ||h|| \, ||h'||$, we have $r_0 \leq |A|$. To obtain the reverse inequality, let ϵ be any positive real. Choose vector h with $||h|| > 0$ and with $||A(h)|| \geq (|A| - \epsilon)||h||$. Then $(|A| - \epsilon)^2||h||^2 \leq (A(h),A(h)) \leq r_0||h|| \, ||A(h)|| \leq r_0|A| \, ||h||^2$, whence $r_0 \geq (|A| - \epsilon)^2/|A|$. Since ϵ is arbitrary, $r_0 \geq |A|$. Hence $r_0 = |A|$.

To summarize, **B** is so far an associative algebra with norm, such that the three algebraic operations (addition, scalar multiplication, and composition) are continuous in the topology of this norm and such that **B**, regarded as a topological vector space, is complete.

We now come to the third, and final, type of structure on **B**. The introduction of this structure begins with the following result.

THEOREM 57. *Let A be a bounded operator on Hilbert space H. Then there is one and only one bounded operator A^* such that $(A(h),h') = (h,A^*(h'))$ for all vectors h and h'.*

Proof. Fix vector h', and, for any vector h, set $\varphi(h) = (h',A(h))$, so $H \xrightarrow{\varphi} \mathbf{C}$. This mapping φ is linear (since $\varphi(h + c\underline{h}) = (h',A(h = c\underline{h})) =$

$(h',A(h) + cA(\underline{h})) = (h',A(h)) + c(h',A(\underline{h})) = \varphi(h) + c\varphi(\underline{h}))$ and continuous (since $|\varphi(h)| = |(h',A(h))| \leq |A|\ ||h'||\ ||h||$). Hence there exists a vector k in H such that $\varphi(h) = (k,h)$ for every h (and this k is unique, for, were there two such, their difference would be orthogonal to every vector). Let A^* be the mapping from H to H with $A^*(h') = k$. Thus we have $(h',A(h)) = (A^*(h'),h)$ for every h and h'. All that remains is to show that A^* is a bounded operator (since this last equation is the complex-conjugate of the equation in the theorem). Since $(A^*(h' + c\underline{h}'),h) = (h' + c\underline{h}',A(h)) = (h',A(h)) + \bar{c}(\underline{h}',A(h)) = (A^*(h'),h) + \bar{c}(A^*(\underline{h}'),h) = (A^*(h') + cA^*(\underline{h}'),h)$ for every h, h', and $\underline{h}'$, A^* is certainly linear. To see that A^* is continuous, first note that $||A^*(h')||^2 = (A^*(h'),A^*(h')) = (AA^*(h'),h') \leq |A|\ ||A^*(h')||\ ||h'||$. Hence (replacing h' by $h' - \underline{h}'$), $||A^*(h') - A^*(\underline{h}')|| \leq |A|\ ||h' - \underline{h}'||$. That is (since A^* takes nearby h' and $\underline{h}'$ to nearby vectors), A^* is continuous. Thus A^* is a bounded operator, completing the proof. □

For A any bounded operator, the operator A^* whose existence is guaranteed by the above theorem is called the *adjoint* of A.

Example. Consider the bounded operators A_α on $L^2(X)$. It is immediate from the fact that $\int_X (\overline{\alpha f}) f'\, d\mu = \int_X f(\overline{\alpha} f')d\mu$ that the adjoint of A_α is $A_\alpha^* = A_{\overline{\alpha}}$. Thus the taking of the adjoint represents, for these examples, the taking of the complex-conjugate of the function α.

What we now wish to do is discover what properties we can of this operation "take the adjoint." There is only one mode of interaction of "adjoint" with itself: for any bounded operator A, $A^{**} = A$. Indeed, applying theorem 57 twice, $(A(h),h') = (h,A^*(h')) = (A^*(h'),h) = (h',A^{**}(h)) = (A^{**}(h),h')$ for any h and h'.

We next consider interaction with the algebraic operations. For A and B bounded operators, and c any complex number, we have $((A + cB)^*(h),h') = (h,(A + cB)(h')) = (h,A(h') + cB(h')) = (h,A(h')) + c(h,B(h')) = (A^*(h),h') + c(B^*(h),h') = ((A^* + \bar{c}B^*)(h),h')$, whence $(A + cB)^* = A^* + \bar{c}B^*$. Thus the operation of taking the adjoint is antilinear under forming linear combinations of bounded operators. We next consider the behavior under composition. We have $((AB)^*(h),h') = (h,(AB)(h')) = (h,A(B(h'))) = (A^*(h),B(h')) = (B^*A^*(h),h')$, whence $(AB)^* = B^*A^*$. That is, the adjoint of the composition is the composition of the adjoints, but in the opposite order. Finally, we consider the relation of adjoints to norms. We claim: for any bounded operator A, $|A| = |A^*|$. But this is immediate from theorem 57 and the fact that $|A|$ is the smallest number r such that $|(Ah,h')| \leq r||h||\ ||h'||$ for every h and h'.

Note that all of these properties of the adjoint could have been guessed from the example above: they reflect the various properties of "take the complex-conjugate function" applied to complex functions.

This completes our discussion of the structure on the set **B** of bounded operators on Hilbert space H. We can summarize this structure by means of the following definition (which we shall, in fact, never use). A *C^*-algebra* consists of three things—i) a complex associative algebra **C**, ii) a rule that assigns, to each element A of **C**, a real number $|A|$, and iii) a rule that assigns, to each element A of **C**, an element A^* of **C**—subject to the following five conditions:

1. For any nonzero A in **C**, $|A|$ is positive.

2. For any A and B in **C**, and any complex number c, $|A + B| \leq |A| + |B|$, $|cA| = |c|\,|A|$, and $|AB| \leq |A|\,|B|$.

3. For any A and B in **C**, and any complex number c, $(A + B)^* = A^* + B^*$, $(cA)^* = \bar{c}A^*$, and $(AB)^* = B^*A^*$.

4. For any A in **C**, $|A^*| = |A|$ and $A^{**} = A$.

5. The topological vector space **C** (i.e., the vector space **C**, with the topology which comes from $|\ |$) is complete.

Thus the set of bounded operators on a Hilbert space has the structure of a C^*-algebra.

There are various special classes of bounded operators which are of particular interest and which one singles out for detailed study. We shall consider three such classes.

A bounded operator A is said to be *Hermitian* if $A = A^*$.

Example. Consider the bounded operators A_α on $L^2(X)$. Then, evidently, A_α is Hermitian if and only if $\alpha = \bar{\alpha}$ almost everywhere.

Since $(A + B)^* = A^* + B^*$, the sum of two Hermitian operators is Hermitian, while, since $(cA)^* = \bar{c}A^*$, cA is Hermitian provided A is and c is real. Of course, the adjoint of a Hermitian operator is Hermitian. Any limit of a sequence of Hermitian operators is Hermitian. Note, however, that it is false in general that, for A and B Hermitian, AB is Hermitian. [The obvious "proof" will not work, for $(AB)^* = B^*A^* = BA$, and not AB.] Note, however, that, for A and B Hermitian, $AB + BA$ and $i(AB - BA)$ are both Hermitian. The following remark makes stronger the analogy between Hermitian operators and real numbers: every bounded operator can be written (in fact, uniquely) as the sum of a Hermitian operator and i times another Hermitian operator. The proof consists of the observation that $A = (1/2)(A + A^*) + (i/2)(-iA + iA^*)$ and that each bounded operator in parentheses on the right is Hermitian.

A bounded operator P is said to be a *projection operator* if P is Hermitian and if, in addition, P satisfies $P = PP$. (Purists like to write these two conditions as one: $P = P^*P$.)

Example. Consider the bounded operators A_α on $L^2(X)$. Then, evidently, A_α is a projection operator if and only if α is real almost everywhere and $\alpha = \alpha^2$ almost everywhere. That is, A_α is a projection operator if and only if the function α takes one of the two values zero or one almost everywhere.

There is a sense in which one understands completely the structure of projection operators. Let H be a Hilbert space, and V any subspace of H. Then any vector h in H can be written uniquely in the form $h = v + \underline{h}$, with v in V and $\underline{h}$ in $V^\perp$. Write P_V for the mapping from H to H that sends h in H to $P_V(h) = v$. We now have

THEOREM 58. *Let H be a Hilbert space. Then every mapping $H \to H$, with V a subspace of H, is a projection operator. Furthermore, every projection operator P is a P_V for some subspace V.*

Proof. It is obvious that, for V a subspace of H, P_V is linear. Furthermore, P_V is continuous, for, for any $h = v + \underline{h}$ in H, with v in V and $\underline{h}$ in $V^\perp$, $||h||^2 = ||v + \underline{h}||^2 = ||v||^2 + (v,\underline{h}) + (\underline{h},v) + ||\underline{h}||^2 = ||v||^2 + ||\underline{h}||^2 \geq ||v||^2 = ||P_V(h)||^2$. Thus P_V is a bounded operator. To see that P_V is Hermitian, consider $h = v + \underline{h}$ and $h' = v' + \underline{h}'$, with v and v' in V and $\underline{h}$ and $\underline{h}'$ in $V^\perp$. Then $(P_V(h),h') = (v,v' + \underline{h}') = (v,v') + (v,\underline{h}') = (v,v')$, while $(h,P_V(h'))$ $= (v + \underline{h},v') = (v,v')$. Thus P_V is Hermitian. Finally, note that, for $h = v + \underline{h}$, with v in V and $\underline{h}$ in $V^\perp$, $P_V(h) = v$, whence $P_VP_V(h) = v = P_V(h)$. Thus $P_VP_V = P_V$. Hence P_V is indeed a projection operator. To prove the converse, let P be any projection operator. Let V be the subspace of H consisting of all vectors of the form $P(h)$ with h in H. Then, for $\underline{h}$ in $V^\perp$, we have $(\underline{h},P(h)) = 0$ for every h, whence $(P(\underline{h}),h) = 0$ for every h, whence $P(\underline{h}) = 0$. Furthermore, for v in V (say, $v = P(h)$ for some h), $P(v) = P(P(h)) = P(h) = v$. Now let h be any vector, and write $h = v + \underline{h}$ with v in V and $\underline{h}$ in $V^\perp$. Then $P(h) = P(v + \underline{h}) = P(v) + P(\underline{h}) = v$. That is, $P = P_V$. □

That is, projection operators are a fancy way of talking about subspaces.

We now obtain a few properties of projection operators. First, note that, if P is a projection operator, then so is $I - P$, where I is the identity bounded operator on H. Indeed, $I - P$ is certainly Hermitian, while $(I - P)(I - P) = I - 2P + PP = I - 2P + P = I - P$. In fact, $I - P_V = P_{V^\perp}$. Furthermore, $P_VP_W = P_WP_V = P_V$ when and only when $V \subset W$. Thus the ordering of subspaces by inclusion is reflected in the corresponding projection operators. Next, note that $P_VP_W = P_WP_V = 0$ when and only when $V \subset W^\perp$ (or, what is the same thing, $W \subset V^\perp$). In general, the sum of two projection operators is not another (for, although this sum is Hermitian, we will not have $(P + P')(P + P') = (P + P')$). However, this sum is a projection operator if $PP' = P'P = 0$. Numerical multiples of projection operators are not in general projection operators, while limits are.

It is particularly easy to find the norm of a projection operator P. First, note that necessarily $|P| \leq 1$. If P is the zero operator (that's a projection operator), then $|P| = 0$. Otherwise, there is a nonzero vector h with $P(h) = h$ ($h = P$(any vector), if nonzero, will do), whence $||P(h)|| = ||h||$. Thus, for a nonzero projection operator P, we must have $|P| = 1$.

A bounded operator U is said to be ***unitary*** if $U^*U = UU^* = I$.

Example. Consider the bounded operators A_α on $L^2(X)$. Then, evidently, A_α is unitary if and only if $|\alpha|^2 = 1$ almost everywhere.

Thus the unitary operators are those whose adjoints are their inverses. To see what this definition means, consider any two vectors, h and h'. Then $(h,h') = (I(h),h') = (U^*U(h),h') = (U(h),U(h'))$. Thus the unitary operators are the "inner product–preserving linear mappings from H to H." In other words, the unitary operators are precisely the isomorphisms from the Hilbert space H to itself.

The sum of two unitary operators is not in general unitary. For U unitary, cU is unitary provided $|c| = 1$. The composition of two unitary operators is, however, unitary, for, for U and U' unitary, $(UU')(UU')^* = UU'U'^*U^* = UIU^* = UU^* = I$, and, similarly, $(UU')^*(UU') = I$. The adjoint of a unitary operator is unitary. For any unitary operator U, we have $|U| = 1$, for, for any h, $||U(h)||^2 = (U(h),U(h)) = (U^*U(h),h) = (h,h) = ||h||^2$.

Exercise 337. Prove that, if P is a projection operator, and U a unitary operator, then UPU^* is a projection operator. Show that the only operator which is both a projection and unitary is the identity operator.

Exercise 338. Find a unitary operator U with $(U(h),h) = 0$ for all h.

Exercise 339. Let X be a measure space, and consider a one-to-one, onto mapping $X \xrightarrow{\kappa} X$ with $\kappa[A]$ measurable if and only if A is, and with $\mu(\kappa[A]) = \mu(A)$. Consider the mapping from $L^2(X)$ to $L^2(X)$ that sends f to $f \circ \kappa$. Show that this is a unitary operator.

Exercise 340. When is the composition of projection operators a projection?

Exercise 341. When does $|A + B| = |A| + |B|$? When does $|AB| = |A|\,|B|$?

Exercise 342. Can one find an inner product, (A,B), on the set $\mathbf{B}$ such that $|A| = (A,A)^{1/2}$? Does the method in chapter 48 work?

The Spectrum of a Bounded Operator

An important tool for unraveling the internal structure of a bounded operator is the study of what is called its spectrum. In this chapter, we motivate, define, and discuss the spectrum of a bounded operator. It turns out that the spectrum tells us more about certain types of operators than others, with the Hermitian operators perhaps the most revealing in this respect.

Let H be a Hilbert space, and A a bounded operator on H. A nonzero vector k in H for which $A(k) = \kappa k$, where κ is a complex number, is called an *eigenvector* of A. The complex number κ is called the corresponding *eigenvalue*. Thus the eigenvectors of A are vectors that "remain invariant, except for scaling" under the action of A.

Example. Consider the bounded operators A_α on $L^2(X)$. Let κ be any complex number. Let the function α on X be such that α takes the value κ

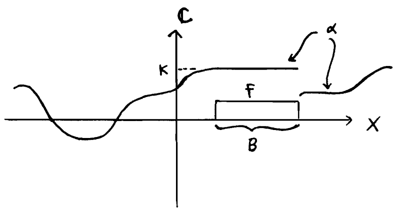

Figure 140

on some measurable set B with $\mu(B)$ neither zero nor infinite (figure 140). Then, we claim, κ is an eigenvalue of A_α. Indeed, let f be the function on X with $f(x) = 1$ if x is in B, and $f(x) = 0$ otherwise. Then f is measurable (since B is), is square-integrable (since $\mu(B) \neq \infty$), and does not define the zero element of $L^2(X)$ (since $\mu(B) \neq 0$). Furthermore, $\alpha f = \kappa f$. Thus the element of $L^2(X)$ defined by this f is an eigenvector of A, with eigenvalue κ.

It is clear from this example that the eigenvectors and eigenvalues of a bounded operator tell one something about the operator. The eigenvalues of a bounded operator, in particular, satisfy the various conditions one might expect. A few such conditions follow.

Let A be a bounded operator, k an eigenvector of A, and κ the corresponding eigenvalue, so $A(k) = \kappa k$. Taking the norm of each side, $||A(k)|| = |\kappa|\ ||k||$. But $||A(k)|| \leq |A|\ ||k||$. Hence $|\kappa| \leq |A|$. The absolute value of each eigenvalue of A is less than or equal to the norm of A. Now suppose, in addition, that A is Hermitian. Then, taking the inner product of each side of $A(k) = \kappa k$ with k, we obtain $(k,A(k)) = (k,\kappa k) = \kappa(k,k)$. But, since A is Hermitian, $(k,A(k))$ is real (for $(k,A(k)) = (A(k),k) = \overline{(k,A(k))}$), while (k,k) is real and nonzero. Thus κ must be real. Each eigenvalue of a Hermitian operator is real. Next, suppose that A is in fact a projection operator. Then, applying A to $A(k) = \kappa k$, we obtain $A(A(k)) = A(\kappa k)$. But $A(A(k)) = A(k) = \kappa k$ (since, for the first step, A is a projection operator), while $A(\kappa k) = \kappa A(k) = \kappa^2 k$. Thus $\kappa k = \kappa^2 k$, whence, since $k \neq 0$, $\kappa = \kappa^2$. The only possible eigenvalues of a projection operator are zero and one. Now suppose, finally, that A is a unitary operator. Taking the inner product of $A(k) = \kappa k$ with itself, $(A(k),A(k)) = (\kappa k,\kappa k)$. But $(A(k),A(k)) = (A^*A(k),k) = (I(k),k) = (k,k)$, while $(\kappa k,\kappa k) = \kappa(\kappa k,k) = \bar{\kappa}\kappa(k,k)$. Hence, since (k,k) is nonzero, $\bar{\kappa}\kappa = 1$. Each eigenvalue of a unitary operator has absolute value one.

In the light of these observations, one might imagine using the set of all eigenvalues of bounded operator A as a means of studying the structure of A. The following example illustrates the problem with such a program.

Example. Consider the bounded operators A_α on $L^2(X)$. Let, for example, X be the measure space of reals, and let α be the function with action $\alpha(x) = 1/(1 + x^2)$. We claim that the (in fact, Hermitian) operator A_α has no eigenvectors whatever. Suppose there were one, k, with eigenvalue κ, and let f be a representative of k. Then, since $A_\alpha(k) = \kappa k$, we must have $\alpha f = \kappa f$ almost everywhere. But $\alpha(x)f(x) = \kappa f(x)$ can hold at a point x only if either $\alpha(x) = \kappa$ or $f(x) = 0$. But, for fixed κ, $\alpha(x) = \kappa$ can hold for at most two of x (e.g., from the graph of figure 141). Therefore we must have $f(x) = 0$

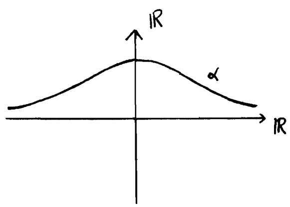

Figure 141

almost everywhere. But this means that $k = 0$, and so k cannot be an eigenvector. The following assertion should, from this and the preceding example, be clear: the complex number κ is an eigenvalue of A_α if and only if the function α assumes the value κ on a measurable set of measure neither zero nor infinite.

Thus the problem is that a bounded operator need not have "enough" eigenvectors or eigenvalues. What we want is some less strict definition, along the same general lines, which will open things up a bit—which will give us more information about the structure of a bounded operator than its eigenvalues do. We now introduce the appropriate notion.

It is convenient to have available the following definition. A bounded operator A is said to be *invertible* if there exists a bounded operator (written A^{-1}) such that $AA^{-1} = A^{-1}A = I$. The (easily checked to be unique) bounded operator A^{-1} is called the *inverse* of A. Thus, for example, $(A^{-1})^{-1} = A$, $(A^*)^{-1} = (A^{-1})^*$, and $(AB)^{-1} = B^{-1}A^{-1}$.

Now fix a bounded operator A on Hilbert space H. A complex number κ is said to be a *regular value* of A if the bounded operator $(A - \kappa I)$ is invertible. The set of complex numbers κ that are not regular (i.e., the set of κ for which $(A - \kappa I)$ is not invertible) is called the *spectrum* of A and is written $\sum(A)$. Thus, for A a bounded operator, $\sum(A)$ is some subset of the set $\mathbf{C}$ of complex numbers.

Example. Let A be a bounded operator on Hilbert space H, and let κ be an eigenvalue of A, with eigenvector k. Then, we claim, κ is in the spectrum of A. Indeed, $(A - \kappa I)$ could hardly be invertible, for we have $(A - \kappa I)(k) = 0$, whence, if $(A - \kappa I)^{-1}$ existed, we would have $(A - \kappa I)^{-1}(A - \kappa I)(k) = (A - \kappa I)^{-1}(0)$, and therefore $k = 0$. This would contradict k's being an eigenvector.

Thus $\sum(A)$ includes, in particular, all of the eigenvalues of A. To decide whether the spectrum of A can include anything else, we return to our earlier example.

Example. Consider the bounded operator A_α on $L^2(\mathbf{R})$ of the previous example, where α is the function with action $\alpha(x) = 1/(1 + x^2)$. Let κ be any real number with $\kappa > 1$. Then κ is a regular value of A_α and, in fact, $(A_\alpha - \kappa I)^{-1} = A_\beta$, where β is the function $\beta = (\alpha - \kappa)^{-1}$ (noting that β is bounded). Similarly, any real $\kappa < 0$, and any complex κ, with nonzero imaginary part, is a regular value. Now consider a real κ in the closed interval $[0,1]$, for example, $\kappa = 1/2$. We claim that κ is in the spectrum of A_α. Choose a sequence $h_1, h_2, \cdots$ of vectors, represented by functions as in figure 142, with $||h_n|| = 1$ for each n. Then, setting $\underline{h}_n = (A_\alpha - \kappa I)(h_n)$, we have, evidently, $\lim||\underline{h}_n|| = 0$. Now suppose that $(A - \kappa I)$ were invertible. Then we would have $\lim||(A_\alpha - \kappa I)^{-1}(\underline{h}_n)|| = \lim||h_n|| = 1$. But $||(A_\alpha - \kappa I)^{-1}(h)|| \le |(A_\alpha - \kappa I)^{-1}|\ ||h||$ for all h, which contradicts $\lim||\underline{h}_n|| = 0$ and $\lim||(A - \kappa I)^{-1}(\underline{h}_n)|| = 1$. This contradiction shows that $(A - \kappa I)$ is not invertible, and hence that this κ is in $\sum(A)$. Thus the spectrum of this A consists of all real κ in the closed interval $[0,1]$.

By simply noticing what properties were actually used in this example, one concludes: complex number κ is in the spectrum of bounded operator A_α on $L^2(X)$ if and only if the following property is satisfied: for every positive ϵ,

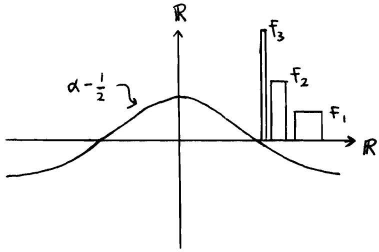

Figure 142

there is a measurable subset B of X, with $\mu(B)$ neither 0 nor ∞, and with $|\alpha(x) - \kappa| \leq \epsilon$ whenever x is in B.

These examples show that the spectrum of a bounded operator A is in general larger than just the set of eigenvalues and is larger in an "appropriate way." One's picture is that the bounded operator A "reacts allergically to elements κ of its spectrum," not necessarily in the most violent way—that $(A - \kappa I)$ annihilates a nonzero vector k—but perhaps in a more gentle way—that $(A - \kappa I)$ refuses to be invertible.

Our program for the remainder of this chapter is to say what we can about the spectrum of a bounded operator, first for the general case and then for the particular types of operators we have introduced.

The basic fact, which gets one started in the discussion of the spectrum in the general case, is the following.

THEOREM 59. *Let A be a bounded operator on Hilbert space H. Then, for any complex number κ in the spectrum of A, we have $|\kappa| \leq |A|$.*

Proof. Let $|\kappa| > |A|$; we show that $(A - \kappa I)$ is invertible. Consider the sequence of bounded operators $B_1 = -\kappa^{-1}(I), B_2 = -\kappa^{-1}(I + A/\kappa), B_3 = -\kappa^{-1}(I + A/\kappa + AA/\kappa^2), \cdots$ Then $B_{n+1} - B_n = -\kappa^{-1}(A^n/\kappa^n)$, where A^n means $AA \cdots A$ (n times). Hence $|B_{n+1} - B_n| \leq |\kappa^{-1}|a^n$, where we have set $a = |A|/|\kappa| < 1$. Therefore, for $m \geq n$, we have $|B_m - B_n| \leq |\kappa^{-1}|(a^n + a^{n+1} + \cdots) = |\kappa^{-1}|a^n/(1 - a)$. We conclude that $B_1, B_2, \cdots$ is a Cauchy sequence in the set **B** of bounded operators. Hence it converges to some bounded operator B. The proof is completed by showing that this B is the inverse of $(A - \kappa I)$. First, note that, by direct computation, $B_n(A - \kappa I) = I - A^n/\kappa^n$. Hence $|B_n(A - \kappa I) - I| = |A^n/\kappa^n| \leq a^n$. That is, $\lim(B_n(A - \kappa I)) = I$. Since composition of operators is continuous, $(\lim B_n)(A - \kappa I) = I$, that is, $B(A - \kappa I) = I$. Similarly, $(A - \kappa I)B = I$. Thus B is indeed the inverse of $(A - \kappa I)$. □

Note that the proof consists of noticing, with a little care, that $(A - \kappa I)^{-1} = -\kappa^{-1}(I - A/\kappa) = \kappa^{-1}(I + A/\kappa + A^2/\kappa^2 + \cdots)$.

Geometrically, theorem 59 says that $\sum(A)$ lies entirely within the circle in the complex plane with radius $|A|$, as illustrated in figure 143.

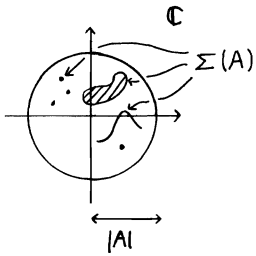

Figure 143

We can actually use theorem 59 again to obtain another little fact about the spectrum. Let A be a bounded operator on Hilbert space H, and let κ be a regular value of A. Let κ' be a complex number with $|\kappa' - \kappa| < 1/|(A - \kappa I)^{-1}|$. We claim that κ' is then also a regular value of A. Consider the bounded operator $I - (A - \kappa I)^{-1}(A - \kappa' I)$. For its norm, we have $|I - (A - \kappa I)^{-1}(A - \kappa' I)| = |(A - \kappa I)^{-1}((A - \kappa I) - (A - \kappa' I))| = |(A - \kappa I)^{-1}((\kappa' - \kappa)I)| \leq |\kappa' - \kappa|\,|(A - \kappa I)^{-1}| < 1$. By theorem 59, "1" is a regular value of this operator, that is, $(I - (A - \kappa I)^{-1}(A - \kappa' I)) - I = -(A - \kappa I)^{-1}(A - \kappa' I)$ is invertible. Call its inverse B, so $-B(A - \kappa I)^{-1}(A - \kappa' I) = I$. That is, $-B(A - \kappa I)^{-1}$ is the inverse of $(A - \kappa' I)$. That is, $(A - \kappa' I)$ is invertible, so κ' is a regular value of A. We conclude in particular from the above remark that every complex number sufficiently close to a regular value of A is itself a regular value of A. That is, the set of regular values of A (regarded as a subset of **C**) is open. The spectrum of A, as the complement of this set, is therefore closed. Thus, for any bounded operator A, $\sum(A)$ is compact, for it is a closed subset of the compact set consisting of all κ with $|\kappa| \leq |A|$.

The above is about all one can say easily about the spectrum of a general bounded operator. (Note that there are some gaping holes. For example, we do not now know (although it is true) that every bounded operator even has a number in its spectrum.) We now turn to the study of the spectrum of the various special classes of operators: Hermitian, projection, and unitary. Our expectations will be confirmed: in each case, the spectrum satisfies the conditions we were able to show earlier for the eigenvalues. It is convenient first to dispense with the two easier cases, those of projection and unitary operators.

Let P be a projection operator on Hilbert space H. Then, we claim, $\sum(P)$ can contain no more than the two numbers zero and one. Indeed, let κ

be a complex number neither zero nor one. Then $(P - \kappa I)(P/\kappa(1 - \kappa) - I/\kappa) = P^2/\kappa(1 - \kappa) - P/(1 - \kappa) - P/\kappa + I = I$, where we used $P^2 = P$ in the last step. Similarly, $(P/\kappa(1 - \kappa) - P/\kappa)(P - \kappa I) = I$. Thus $(P/\kappa(1 - \kappa) - P/\kappa)$ is the inverse of $(P - \kappa I)$. We conclude that this κ must be a regular value of P. Note furthermore that, for P neither the zero nor the identity operator, $\sum(P)$ actually contains both zero and one, for these are both eigenvalues.

Next, let U be a unitary operator on Hilbert space H. We claim that every κ in $\sum(U)$ satisfies $|\kappa| = 1$. First, note that, since $|U| = 1$, every such κ satisfies $|\kappa| \leq 1$. Next, note that U itself is invertible and, in fact, that $U^{-1} = U^*$. Consider the bounded operator $I - U^{-1}(U - \kappa I)$, where κ is any number with $|\kappa| < 1$. For its norm, $|I - U^{-1}(U - \kappa I)| = |U^{-1}(U - (U - \kappa I))| = |U^{-1}\kappa| = |U^*\kappa| = |U^*|\,|\kappa| = |\kappa| < 1$. Hence, by theorem 59, "1" is a regular value of this operator, that is, $(I - U^{-1}(U - \kappa I)) - I = -U^{-1}(U - \kappa I)$ is invertible. Call its inverse B, so $-BU^{-1}(U - \kappa I) = -U^{-1}(U - \kappa I)B = I$. Thus $-BU^{-1}$ is the inverse of $(U - \kappa I)$. Since $(U - \kappa I)$ is invertible, κ is a regular value of U. Since every κ with $|\kappa| < 1$ is a regular value of U and since every κ in the spectrum of U satisfies $|\kappa| \leq 1$, every κ in $\sum(U)$ must satisfy $|\kappa| = 1$.

Finally, we come to the (slightly more difficult) Hermitian case. Fix a Hermitian operator A on Hilbert space H, and fix a complex number $\kappa = a + ib$, with a and b real and b nonzero. We shall have to use the result of the following little calculation several times. Let h be any vector in H. Then $||(A - \kappa I)(h)||^2 = ((A - aI - ibI)(h),(A - aI - ibI)(h)) = ((A - aI)(h),(A - aI)(h)) + ((A - aI)(h),-ibh) + (-ibh,(A - aI)(h)) + (-ibh,-ibh) = ||(A - aI)(h)||^2 - ib((A - aI)(h),h) + ib(h,(A - aI)(h)) + b^2||h||^2 = ||(A - aI)(h)||^2 + b^2||h||^2$, where, in the last step, we used the fact that $(A - aI)$ is Hermitian. Comparing the first and last expressions, we have the desired result:

$$||(A - \kappa I)(h)||^2 \geq b^2||h||^2 \ .$$

Now consider the bounded operator $(A - \kappa I)$. We first note that it is one-to-one, for, if $(A - \kappa i)(h) = 0$, then, by the inequality above, $||h||^2 = 0$, whence $h = 0$. Denote by V the subset $(A - \kappa I)[H]$ of H, that is, the subset consisting of all vectors of the form $(A - \kappa I)(h)$ with h in H. Since $(A - \kappa I)$ is linear, this V is certainly a vector subspace of vector space H. We next show that V is closed. Let $\underline{h}_1 = (A - \kappa I)(h_1), \underline{h}_2 = (A - \kappa I)(h_2), \cdots$ be a sequence of vectors in V, with limit $\underline{h} = \lim \underline{h}_n$. We must show that this $\underline{h}$ is also in V. Since $\underline{h}_1, \underline{h}_2, \cdots$ converges, it is a Cauchy sequence. But, by our inequality above, $b^2||h_n - h_m||^2 \leq ||(A - \kappa I)(h_n - h_m)||^2 = ||\underline{h}_n - \underline{h}_m||^2$, so $h_1, h_2, \cdots$ is also a Cauchy sequence. Hence $h_1, h_2, \cdots$ converges to some vector h. Since $(A - \kappa I)$ is continuous, $(A - \kappa I)(h) = \lim(A - \kappa I)(h_n) = \lim \underline{h}_n = \underline{h}$. Thus $\underline{h}$ is indeed in V. Thus V is closed. We conclude that V, as a vector subspace of vector space H which is also closed, is a subspace of Hilbert space H. The next step is to show that $V = H$. Let h be any vector in $V^{\perp}$, so $((A - \kappa I)(h'),h) = 0$ for every h'. Then $0 = (h',(A - \kappa I)^*(h)) = (h',(A - \bar{\kappa} I)(h))$ for every h'. In particular, choosing $h' = (A - \bar{\kappa} I)(h)$, we obtain $(A - \bar{\kappa} I)(h) = 0$. But, since the Hermitian operator A can have no eigenvectors with nonreal eigenvalue $\bar{\kappa}$, we must have $h = 0$. We have shown that $V^{\perp}$ contains only the zero vector, whence $V = V^{\perp\perp} = H$.

The statement that $V = H$ is precisely the statement that $(A - \kappa I)$ is onto. Since $(A - \kappa I)$ is one-to-one and onto, there exists a (clearly linear) mapping $H \overset{B}{\rightarrow} H$ with $B(A - \kappa I) = (A - \kappa I)B = I$. We next want to show that B is a bounded operator, that is, that this B is continuous. Let h be any vector in H, and set $h = (A - \kappa I)(h')$. Then $||B(h)||^2 = ||B(A - \kappa I)(h')||^2 = ||h'||^2 \leq b^{-2}||(A - \kappa I)(h')||^2 = b^{-2}||h||^2$, where, in the second to last step, we once again used our inequality. We conclude that B is continuous (and, in fact, that $|B| \leq b^{-1}$).

The previous two paragraphs are the proof that, if κ has nonzero imaginary part, then $(A - \kappa I)$ is invertible. That is, every such κ is a regular value of A. We conclude: the spectrum of a Hermitian operator A consists only of real numbers.

One can actually say a bit more about the spectrum in the Hermitian case: for A a Hermitian operator on Hilbert space H, either $|A|$ or $-|A|$ (or possibly both) is in $\sum(A)$. Fix any positive number ϵ, and let h be a vector with $||h|| = 1$ and with $||A(h)||^2 \geq |A|^2 - \epsilon^2$. Let g be any vector with $||g|| = 1$ and $(h,g) = 0$. Then, for any complex number λ, we have $||A(h + \lambda g)||^2 = ||A(h)||^2 + \lambda(A(h),A(g) + \bar{\lambda}(A(g),A(h)) + |\lambda|^2||A(g)||^2 \geq |A|^2 - \epsilon^2 + \lambda(A(h),A(g)) + \bar{\lambda}(A(g),A(h))$. But, by definition of the norm of A, $||A(h + \lambda g)||^2 \leq |A|^2||h + \lambda g||^2 = |A|^2(1 + |\lambda|^2)$ where, in the second step, we used $||h|| = ||g|| = 1$ and $(h,g) = 0$. Comparing these two inequalities, we have $-\epsilon^2 + \bar{\lambda}(A(g),A(h)) + \lambda(A(h),A(g)) - |A|^2|\lambda|^2 \leq 0$. Since this quadratic expression in λ can assume only nonpositive values, we must have $|(A(g),A(h))| \leq \epsilon|A|$, or, since A is Hermitian, $|(g,A^2(h))| \leq \epsilon|A|$, or, since $(g,h) = 0$, $(g,(A^2 - |A|^2 I)(h)) \leq \epsilon|A|$. To summarize, we have shown so far that, for any g with $||g|| = 1$ and $(h,g) = 0$, $|(g,(A^2 - |A|^2 I)(h))| \leq \epsilon|A|$. Next, note that $|(h,(A^2 - |A|^2 I)(h))| = |(h,A^2(h)) - |A|^2||h||^2| = |(A(h),A(h)) - |A|^2||h||^2| \leq \epsilon^2$. Thus, for any vector h' with $||h'|| = 1$, we have $|(h',(A^2 - |A|^2 I)(h))| \leq \epsilon^2 + |A|\epsilon$ (since this is true for h' either parallel or orthogonal to h). In particular, choosing h' a multiple of $(A^2 - |A|^2 I)(h)$, we conclude that $||(A^2 - |A|^2 I)(h)|| \leq \epsilon^2 + |A|\epsilon$.

To summarize, we have shown in the paragraph above that, for any positive ϵ, there is a vector h with $||h|| = 1$ and with $||(A - |A|^2 I)(h)|| \leq \epsilon$. Intuitively, this means that "h is within ϵ of being an eigenvector of A^2 with eigenvalue $|A|^2$." This intuitive picture suggests the next step: we want to show that $|A|^2$ is in the spectrum of A^2. Choose a sequence, $h_1,h_2, \ldots,$ of vectors with $||h_n||^2 = 1$ and with $\lim||(A^2 - |A|^2 I)(h_n)|| = 0$. If $(A^2 - |A|^2 I)$

were invertible, we would have, setting $\underline{h}_n = (A^2 - |A|^2I)(H_n)$, $\lim||\underline{h}_n|| = 0$, while $\lim||(A^2 - |A|^2I)^{-1}(\underline{h}_n)|| = \lim||h_n|| = 1$, contradicting $||(A^2 - |A|^2I)(h)|| \le |(A^2 - |A|^2I)^{-1}|\ ||h||$ for all h. Thus $(A^2 - |A|^2I)$ cannot be invertible. We conclude that the number $|A|^2$ is in the spectrum of the bounded operator A^2.

Now suppose that both $|A|$ and $-|A|$ were regular values of A, so both $(A - |A|I)$ and $(A + |A|I)$ are invertible. Then their product, $(A - |A|I)(A + |A|I) = A^2 - |A|^2I$, would have to be invertible. But we have just shown that $A^2 - |A|^2I$ is not invertible. Hence at least one of $|A|$ or $-|A|$ must be in $\sum(A)$.

Note, as a consequence of the result above, the spectrum of a Hermitian operator is not empty. Note also that, for a Hermitian operator, the spectrum actually determines the norm: $|A|$ is the smallest number such that $\sum(A) \subset [-|A|,|A|]$.

We summarize these conclusions about the spectrum in the Hermitian case.

THEOREM 60. *Let A be a Hermitian operator on Hilbert space H. Then $\sum(A)$ is a closed subset of the closed interval $[-|A|,|A|]$, including at least one endpoint.*

Exercise 343. Let C be any compact subset of the complex plane **C**. Prove that there exists a Hilbert space H, and bounded operator A on H, such that $\sum(A) = C$.

Exercise 344. Find an example of a Hilbert space H, and bounded operator A on H, such that the set of eigenvalues of A is not closed.

Exercise 345. Is it true that a bounded operator with a real spectrum is Hermitian? with a spectrum including only zero and one is a projection? with every element of its spectrum having absolute value one is unitary?

Exercise 346. Consider the Hilbert space $L^2(X)$ given in chapter 48. Let A be the operator with action $A(c_1,c_2,\cdots) = (0,c_1,c_2,\cdots)$. Find the norm of A, its eigenvalues and eigenvectors, and its spectrum.

Exercise 347. Say what one can, given the spectra of A and B, about the spectra of AB, $A + B$, and cA.

Exercise 348. Prove that $\overline{\sum(A)} = \sum(A^*)$.

Exercise 349. Prove that two eigenvectors of a Hermitian operator, with different eigenvalues, are orthogonal.

Exercise 350. Does there exist an example of a bounded operator A such that $|\kappa| < |A|$ for every κ in $\sum(A)$?

Exercise 351. The argument earlier in this chapter shows that, if A is Hermitian and $||A(h)|| \leq b||h||$ for all h, with $b < 0$, then A is invertible. Find an example to show that the requirement that A be Hermitian is actually necessary.

Exercise 352. Relate the spectrum of an invertible Hermitian operator A to that of A^{-1}.

Exercise 353. Let A be a bounded operator on Hilbert space H. Prove that, if $A^* = -A$, then the spectrum of A includes only purely imaginary elements; if $A^2 = I$, then the spectrum of A contains no more than $+1$ and -1.

Exercise 354. Let A be a Hermitian operator. State and prove: $U = I + iA + (iA)^2/2! + (iA)^3/3! + \cdots$ exists and is unitary.

51

The Spectral Theorem: Finite-dimensional Case

Our ultimate goal is to state and prove a certain theorem, called the spectral theorem, which provides "a complete and unique description of the internal structure" of any Hermitian operator on any Hilbert space. A special case occurs when the Hilbert space is finite-dimensional: in this particular case the spectral theorem is easy to prove, and its statement in fact reduces to a well-known fact about matrices. Nonetheless, the finite-dimensional case is of interest, for it serves as both a guide to the proof and a source of motivation for the (far more difficult) general case. In this chapter, we shall discuss the spectral theorem on a finite-dimensional Hilbert space.

The finite-dimensional spectral theorem is the following:

THEOREM 61. *Let A be a Hermitian operator on a finite-dimensional Hilbert space H. Then there exists a finite collection, $\kappa_1 < \kappa_2 < \cdots < \kappa_n$, of real numbers, together with a collection, $P_1, \ldots, P_n$, of nonzero projection operators on H, satisfying the following three conditions: i) $I = P_1 + \cdots + P_n$, ii) $P_iP_j = P_jP_i = 0$ for $i \neq j$, and iii) $A = \kappa_1 P_1 + \cdots + \kappa_n P_n$. Furthermore, this $\kappa_1, \ldots, \kappa_n$, $P_1, \ldots, P_n$ is unique.*

Proof. Denote by S the (topological) subspace of topological space H consisting of all vectors h with $||h|| = 1$. Since H is finite-dimensional, S is compact. Consider $S \xrightarrow{\varphi} \mathbf{R}$ with action $\varphi(h) = (h,A(h))$, for h is S (noting that, since A is Hermitian, $\varphi(h)$ is necessarily real). Since A is continuous and since the inner product is continuous, φ is continuous. Therefore φ achieves its minimum, κ_1, that is, there is a vector k with $||k|| = 1$ and $(k,A(k)) = \kappa_1$, such that $(h,A(h)) \geq \kappa_1$ for every h with $||h|| = 1$. We claim that k is therefore an eigenvector of A. Indeed, letting g be any vector with $(k,g) = 0$, and λ any complex number, $(k + \lambda g)/(1 + \lambda\bar{\lambda}(g,g))^{1/2}$ is a vector with unit norm, whence $(1 + \lambda\bar{\lambda}(g,g))^{-1}(k + \lambda g, A(k + \lambda g)) \geq \kappa_1$. Expanding, we have $\lambda(g,A(k)) + \overline{\lambda(g,A(k))} \geq -\lambda\bar{\lambda}(g,A(h)) + \kappa_1\lambda\bar{\lambda}(g,g)$, an inequality which can hold for all λ only if $(g,A(k)) = 0$. Thus every vector orthogonal to k is also orthogonal to $A(k)$, whence $A(k)$ must be a multiple of k, whence k must be an eigenvector of A. Since $(k,A(k)) = \kappa_1$, the corresponding eigenvalue is κ_1. Let V_1 be the subspace of H consisting of all eigenvectors with eigenvalue κ_1, and set $P_1 = P_{V_1}$, a nonzero projection operator.

Now consider the Hilbert space $V_1^{\perp}$. For h in $V_1^{\perp}$ (so $(v,h) = 0$ for all v in V_1), $A(h)$ is also in $V_1^{\perp}$ (for, for any v in V_1, $(v,A(h)) = (A(v),h) = (\kappa_1 v,h) = \overline{\kappa_1}(v,h) = 0$). Thus A can be regarded as a Hermitian operator on the Hilbert space $V_1^{\perp}$. Repeating the argument above, we obtain a real number $\kappa_2 > \kappa_1$, a subspace V_2 of $V_1^{\perp}$ consisting of eigenvectors of A with eigenvalue κ_2, and a projection operator $P_2 = P_{V_2}$. Repeating the argument again, beginning with Hilbert space $(V_1 + V_2)^{\perp}$, we obtain $\kappa_3 > \kappa_2$ and $P_3 = P_{V_3}$, etc. Since H is finite-dimensional, this process terminates (when $(V_1 + \cdots + V_n)^{\perp} = 0$) after a finite number of steps, resulting in $\kappa_1 < \kappa_2 < \cdots < \kappa_n$ and $P_1, \ldots, P_n$.

Since $V_i \subset V_j^{\perp}$ for $i \neq j$, $P_iP_j = P_jP_i = 0$ for $i \neq j$. Since $V_1, \ldots, V_n$ together generate the entire Hilbert space H, $P_1 + \cdots + P_n = I$. For any h in H, $P_i(h)$ is in V_i, whence $A(P_i(h)) = \kappa_i P_i(h)$. Therefore, for any h in H, $A(h) = AI(h) = A(P_1 + \cdots + P_n)(h) = AP_1(h) + \cdots + AP_n(h) = \kappa_1 P_1(h) + \cdots + \kappa_n P_n(h) = (\kappa_1 P_1 + \cdots + \kappa_n P_n)(h)$. That is, $A = \kappa_1 P_1 + \cdots + \kappa_n P_n$.

To prove uniqueness, let $\kappa_1, \ldots, \kappa_n, P_1, \ldots, P_n$ be as in the theorem. Let k be any eigenvector of A, with eigenvalue κ. Then, applying P_i to $A(k) = \kappa k$, we have $\kappa P_i(k) = P_i A(k) = P_i(\kappa_1 P_1 + \cdots + \kappa_n P_n)(k) = \kappa_i P_i(k)$. We cannot have $P_i(k) = 0$ for all i, since $I = P_1 + \cdots + P_n$, so κ must be one of the κ_i. Then, for $j \neq i$, $P_j(k) = 0$, whence $P_i(k) = (I - P_1 - \cdots - P_{i-1} - P_{i+1} - \cdots - P_n)(k) = I(k) = k$. We conclude: for $\kappa_1, \ldots, \kappa_n, P_1, \ldots, P_n$ as in the theorem, the $\kappa_1, \ldots, \kappa_n$ are precisely the eigenvalues of A, and each P_i is precisely the projection operator onto the subspace consisting of eigenvectors with eigenvalue κ_i. It is immediate from this characterization that $\kappa_1, \ldots, \kappa_n, P_1, \ldots, P_n$ is unique. $\square$

Theorem 61 will be recognized as a fancy way of saying the familiar fact that every Hermitian matrix can be diagonalized, with the eigenvalues "along the diagonal." In fact, the language of theorem 61 is normally the most convenient expression of this idea. For example, $h = I(h) = (P_1 + \cdots + P_n)(h) = P_1(h) + \cdots + P_n(h)$ states that every vector can be written as a sum of eigenvectors of A. Since $P_iP_j = 0$ for $i \neq j$, eigenvectors with distinct eigenvalues are orthogonal, etc.

Thus every Hermitian operator on a finite-dimensional Hilbert space can be decomposed in terms of simple things—projection operators—by $A = \kappa_1 P_1 + \cdots + \kappa_n P_n$. Furthermore, given $P_1, \ldots, P_n$ satisfying conditions i) and ii) of theorem 61, and given real numbers $\kappa_1, \ldots, \kappa_n$, $A = \kappa_1 P_1 + \cdots + \kappa_n P_n$ is, clearly, Hermitian. In this sense, then, one "understands completely the structure of a Hermitian operator on a finite-dimensional Hilbert space."

The spectral theorem is a generalization of theorem 61 to the infinite-dimensional case. There are essentially three things involved in the decomposition $A = \kappa_1 P_1 + \cdots + \kappa_n P_n$: the κ_i, the P_i, and the sum. It is of interest to ask how each of these three things will be generalized in the final spectral theorem.

The set of κ_i—the set of eigenvalues of A in the finite-dimensional case—is generalized to the spectrum, $\sum(A)$, of A. From our earlier remarks concerning the spectrum, this seems reasonable. Note also that such a generalization is consistent with the finite-dimensional case. In fact, we claim: for A a Hermitian operator on finite-dimensional H, $\sum(A)$ consists precisely of the n numbers $\kappa_1, \ldots, \kappa_n$. Proof: For κ different from all κ_i, the bounded operator $P_1/(\kappa - \kappa_1) + \cdots + P_n/(\kappa - \kappa_n)$ is the inverse of $(A - \kappa I)$. (Note, incidentally, that we therefore have, in the finite-dimensional case, that $|A|$ is the maximum of $|\kappa_1|, |\kappa_2|, \ldots, |\kappa_n|$.)

The fact that a sum appears in the decomposition of Hermitian A in the finite-dimensional case is a reflection of the fact that $\sum(A)$ is finite in this case. In the infinite-dimensional case, on the other hand, $\sum(A)$ may be infinite and could even be, for example, an entire closed interval in the reals. Thus we must discover some way to "sum over a continuum of values." But we already have such a "way": the notion of an integral. One might imagine therefore that Hermitian A will be represented, in the infinite-dimensional case, by a suitable integral rather than a sum.

The projection operators—the P_i—of the finite-dimensional decomposition will essentially remain as projection operators in the infinite-dimensional case. But one thing, at least, must change: the relation of the P_i to A. In the finite-dimensional case, P_i is the projection operator onto the subspace consisting of eigenvectors with eigenvalue κ_i. This could hardly be the situation in the infinite-dimensional case, however, for we need not, in this case, have any eigenvectors at all. Let us first replace the P_i by another set of projection operators. For each real number κ, let $P_\kappa = P_1 + \cdots + P_m$, where the sum extends over all P_i with $\kappa_i \leq \kappa$. Thus, for $\kappa < \kappa_1$, $P_\kappa = 0$; for κ in $[\kappa_1, \kappa_2)$, $P_\kappa = P_1$; for $\kappa \geq \kappa_n$, $P_\kappa = I$. Each P_κ is a projection operator, an operator which "projects onto a larger and larger subspace as κ increases, where the subspace jumps a dimension or two as κ passes an eigenvalue of A." One might expect, in the infinite-dimensional case, for the P_κ to again "project onto larger and larger subspaces as κ increases, although the size of the subspace may increase more continuously with κ as κ increases through the spectrum of A."

However, these remarks do not as yet suggest how one is to recover the projection operators from A without reference to eigenvectors. To see how this is to be done, first note that, in the finite-dimensional case, $A^2 = (\kappa_1 P_1 + \cdots + \kappa_n P_n)(\kappa_1 P_1 + \cdots + \kappa_n P_n) = \sum_{i,j} \kappa_i \kappa_j P_i P_j =$

$(\kappa_1)^2 P_1 + \cdots + (\kappa_n)^2 P_n$. More generally, for $p = a_0 + a_1 x + \cdots + a_s x^2$ a polynomial in variable x, let $p(A) = a_0 + a_1 A + \cdots + a_s A^2$, a bounded operator. Then, evidently, we have $p(A) = p(\kappa_1)P_1 + \cdots + p(\kappa_n)P_n$. In other words, to evaluate a polynomial of A, one just applies that polynomial to each κ_i, multiplies the resulting number by the corresponding P_i, and sums. Now suppose that it were somehow possible to apply more exotic functions—not just polynomials—to A. One might still expect to have, in the finite-dimensional case, $f(A) = f(\kappa_1)P_1 + \cdots + f(\kappa_n)P_n$. What one would like to do is find a suitable f so that, when applied to A, it produces precisely the P_κ above. What function should we choose? For real κ, let θ_κ be the function with $\theta_\kappa(x)$ one for $x \leq \kappa$, and $\theta_\kappa(x)$ zero otherwise. Then, from the remarks above, one would expect that $\theta_\kappa(A)$ will be precisely P_κ.

Thus, if one can find a reasonable meaning for "$f(A)$," where f begins as a real function of a real variable—and if one can do this for sufficiently exotic instances of f—then one might have a hope of obtaining the projection operators appropriate for a decomposition of a Hermitian operator in the infinite-dimensional case. It turns out to be convenient to proceed in two steps: first for continuous f and then for others. We now begin this program.

Continuous Functions of a Hermitian Operator

In this chapter, we show that, given a continuous real-valued function of one real variable, a sensible meaning can be attached to the sentence "Apply that function to a Hermitian operator A."

Fix, once and for all, a Hilbert space H and a Hermitian operator A on H. Denote by **C** the collection of all continuous functions $\sum(A) \rightarrow \mathbf{R}$, where the topological space $\sum(A)$ on the left is that subspace of the real line. Our first goal is to decide what is the structure of this set **C**. Note that, for f and f' such continuous functions, and r any real number, $f + f'$, ff', and rf are all continuous functions on $\sum(A)$. Thus **C** has the structure of an associative algebra. Next, for f and f' in **C**, write $d(f,f') = \max_{x \text{ in } \Sigma(A)} |f(x) - f'(x)|$. This is clearly a metric on the set **C**, whence **C** has the structure of a topological space. Note that the three algebraic operations in **C** are all continuous: addition of functions, multiplication of functions, and multiplication of functions by real numbers. In particular, the vector space **C**, with the topology above, is a topological vector space.

Now let f be an element of **C**. What we are trying to do is define a certain bounded operator, associated with this f, which we may write $f(A)$. In other words, we want to associate, with each element f of **C**, an element, $f(A)$, of the set **B** of all bounded operators on the Hilbert space H. In other words, we want to find a certain mapping $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$. If we regard **C** and **B** as just sets, then there will certainly be many mappings $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$ of sets. We will have no particular criterion to decide which mapping is the "right one," and hence will have no "natural" $f(A)$ associated with f. The idea, then, is to impose on this mapping $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$ certain "niceness conditions," with the hope that these conditions will make the mapping unique. What conditions should we impose? Note that we can add continuous functions (i.e., elements of **C**) and that we can add bounded operators (i.e., elements of **B**). It would certainly be convenient if addition were preserved by ψ, that is, if $(f + f')(A) = f(A) + f'(A)$ for all f and f' in **C**. Similarly, we can multiply continuous functions by numbers and multiply bounded operators by numbers; we can multiply continuous functions by such functions and multiply (compose) bounded operators, and one might like to ask that these operations be preserved by $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$. We can express the same idea in terms of the structure of **C** and **B**. Each is an associative algebra, and we might require that $\mathbf{C} \xrightarrow{\psi}$

$\mathbf{B}$ be a homomorphism of associative algebras. Similarly, one might like to require that, if f and f' are "nearby" continuous functions, then $f(A)$ and $f'(A)$ be "nearby" bounded operators. That is, one might like to require that $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$ be a continuous mapping of topological spaces.

One might imagine that some such set of conditions would give us a unique mapping $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$, that is, would define a unique bounded operator $f(A)$ for each f in $\mathbf{C}$. That is, one suspects that some result along the following lines will be true.

THEOREM 62. *Let A be a Hermitian operator on Hilbert space H, and denote by $\mathbf{C}$ the set of all continuous, real-valued functions on $\sum(A)$, so $\mathbf{C}$ is a topological space and an associative algebra. Then there is one and only one mapping $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$ satisfying the following three conditions: i) ψ is a homomorphism of associative algebras, ii) ψ is a continuous mapping of topological spaces, and iii) for f the element of $\mathbf{C}$ with action $f(x) = a_0 + a_1 x$ (a_0 and a_1 real numbers), $\psi(f)$ is the bounded operator $a_0 I + a_1 A$.*

$f(A) := \psi(f)$

In fact, theorem 62 is true: we now proceed with a rather leisurely proof.

By condition iii), for $f(x) = a_0 + a_1 x$, $f(A) = a_0 I + a_1 A$. In other words, for linear polynomials f, $f(A)$ is the obvious bounded operator. Let us next consider a quadratic polynomial, that is, f in $\mathbf{C}$ with action $f(x) = a_0 + a_1 x + a_2 x^2$. Then, letting g have action $g(x) = x$, we have $f = a_0 + a_1 g + a_2 g^2$. Since, by condition *i*, ψ is to be a homomorphism, we must have $f(A) = a_0 I + a_1 g(A) + a_2 g(A)^2$. But $g(A) = A$, so $f(A) = a_0 I + a_1 A + a_2 A^2$. Similarly for polynomials of higher order. We conclude: for f in $\mathbf{C}$ with action $f(x) = a_0 + a_1 x + \cdots + a_n x^n$, with $a_0, \ldots, a_n$ real numbers, we must, by conditions i) and iii), have $f(A) = a_0 I + a_1 A + \cdots + a_n A^n$.

This observation does not, however, complete the proof, for there will, in general, be functions f in $\mathbf{C}$ which cannot be expressed as polynomials. It is also clear that conditions i) and iii) will not tell us what $f(A)$ should be for such functions, for such an f cannot be written as a linear combination of products of polynomials of the form $a_0 + a_1 x$. In short, we have essentially exhausted what we can deduce from conditions i) and iii): to decide what $f(A)$ is to be for a nonpolynomial f, we must use condition ii). The idea is to "approximate f by polynomials, which yield bounded operators as above, which will serve as approximations to $f(A)$."

We first make the following observation. Let f in $\mathbf{C}$ be a polynomial: $f(x) = a_0 + a_1 x + \cdots + a_n x^n$. Then $\sum(f(A)) = f(\sum(A))$, that is, the spectrum of the bounded operator $f(A) = a_0 I + \cdots + a_n A^n$ is precisely the set of real numbers of the form $f(\kappa)$ with κ in $\sum(A)$. Indeed, let κ be in $\sum(A)$. Then, since $x = \kappa$ is a root of the polynomial $f(x) - f(\kappa)$, we have $f(x) - f(\kappa) =$

$(x - \kappa)p(x)$ for some polynomial p. Hence $f(A) - f(\kappa)I = (A - \kappa I)p(A)$. But now $f(A) - f(\kappa)I$ could hardly be invertible, for had it an inverse, B, $Bp(A)$ would, by the equation above, be an inverse of $(A - \kappa I)$, violating the assumption that κ is in the spectrum of A. Thus $f(\kappa)$ is in the spectrum of the bounded operator $f(A)$. We have shown that $\sum (f(A)) \supset f(\sum (A))$. To obtain the reverse inclusion, let λ be in the spectrum of $f(A)$. We have (by the fundamental theorem of algebra) $f(x) - \lambda = r(x - \kappa_1) \cdots (x - \kappa_n)$, where r is a real number and $\kappa_1, \ldots, \kappa_n$ are the n complex numbers with $f(\kappa_i) = \lambda$. Therefore $f(A) - \lambda I = r(A - \kappa_1 i) \cdots (A - \kappa_n I)$. It cannot be true that each of $(A - \kappa_1 I) \cdots, (A - \kappa_n I)$ is invertible, for that would imply that $f(A) - \lambda I$ is invertible, contradicting the assumption that λ is in $\sum (f(A))$. Hence at least one of $\kappa_1, \ldots, \kappa_n$ must be in the spectrum of A. Since $f(\kappa_i) = \lambda$ for each i, this implies that λ is in $f(\sum (A))$. Thus $\sum (f(A)) \subset f(\sum (A))$. We conclude that $\sum (f(A)) = f(\sum (A))$.

Thus the operation "apply a polynomial to A" has the expected action on the spectrum of A. But from this we can deduce the action on the norm. Indeed, for A any Hermitian operator, $|A|$ is the maximum of $|\kappa|$, for κ in $\sum (A)$. Hence $|f(A)|$ is the maximum value of $|\lambda|$ for λ in $\sum (f(A))$, which (since $\sum (f(A)) = f(\sum (A))$) is the maximum value of $|\lambda|$ for λ in $f(\sum (A))$, that is, the maximum value of $|f(\kappa)|$ for κ in $\sum (A)$. That is (using the metric $d(\ ,\ )$ on **C**), $|f(A)|$ is just $d(f,0)$. Replacing f by $f - f'$, we conclude: for f and f' polynomials in **C**, $|f(A) - f'(A)| = d(f,f')$. In other words, "nearby polynomials" (in terms of the metric of **C**) yield "nearby bounded operators" (in terms of the norm on **B**).

Now consider a function f in **C**, not necessarily a polynomial. Suppose we can find a sequence, $f_1,f_2, \ldots,$ of polynomials in **C** with $\lim f_n = f$ in **C**, that is, with $\lim d(f,f_n) = 0$. Then, in particular, $f_1,f_2, \cdots$ is a Cauchy sequence in **C**. Now, for each of the polynomials $f_1,f_2, \ldots,$ we have a corresponding bounded operator, $f_1(A),f_2(A), \cdots$ Since $f_1,f_2, \cdots$ is a Cauchy in **C** and since $|f_n(A) - f_m(A)| = d(f_n,f_m)$, the sequence $f_1(A),f_2(A), \cdots$ of bounded operators is Cauchy in **B**. Hence this sequence converges to some bounded operator, which we may write $f(A)$. This $f(A)$ is clearly independent of the choice of Cauchy sequence (for, if $f_1',f_2', \cdots$ converges to f, we must have, for every positive ϵ, $d(f_n,f_n') \leq \epsilon$ for all sufficiently large n, whence $|f_n(A) - f_n'(A)| \leq \epsilon$ for all sufficiently large n, whence $f_1(A), \cdots$ must converge to the same bounded operator as $f_1'(A), \cdots$).

That every continuous function f on $\sum (A)$ can be approximated as above by polynomials is a consequence of the following: Let C be any compact subset of the real line, f any continuous function on C, and ϵ any positive number. Then there exists a polynomial f' such that $|f(x) - f'(x)| \leq \epsilon$ for all x in C. This is the Weierstrass approximation theorem (whose proof we omit, since the theorem is not unreasonable intuitively, since the proof is technical and not very difficult, and since almost every textbook in mathematics gives a proof).

We now have our mapping $\mathbf{C} \xrightarrow{\psi} \mathbf{B}$—by construction, the unique one which is even a candidate for satisfying the three conditions of theorem 62 What remains is to verify that these three conditions are in fact satisfied by this ψ. For condition i), consider f and f' in $\mathbf{C}$, and let $f = \lim f_n$ and $f' = \lim f_n'$, with $f_1, f_2, \cdots$ and $f_1', f_2', \cdots$ polynomials. Then $\lim f_n f_n' = ff'$, whence $\lim f_n(A) f_n'(A) = f(A)f'(A)$, whence $(ff')(A) = f(A)f'(A)$. Similarly, $(f+f')(A) = f(A) + f'(A)$, and $(rf)(A) = rf(A)$. Thus ψ is a homomorphism of associative algebras. For condition ii), let f and f' be in $\mathbf{C}$, and let $f = \lim f_n$ and $f' = \lim f_n'$, with $f_1, f_2, \cdots$ and $f_1', f_2', \cdots$ polynomials. Then $\lim d(f_n, f_n') = d(f, f')$, whence $\lim |f_n(A) - f_n'(A)| = |f(A) - f'(A)|$, whence $|f(A) - f'(A)| = d(f, f')$. Thus ψ is distance-preserving and hence continuous. Condition iii) is obvious.

This completes the proof of theorem 62.

Note that, for any f in $\mathbf{C}$, $f(A)$ is necessarily Hermitian (since this is certainly true for f a polynomial and since any limit of Hermitian operators is Hermitian). One can also define $f(A)$ for f a continuous, complex-valued function: set $f = f_r + if_i$, with f_r and f_i real (and necessarily continuous), and then set $f(A) = f_r(A) + if_i(A)$. Thus, for example, for A a Hermitian operator, we know what $\sin A$, 2^A, and e^{iA} mean. If the spectrum of A has only non-negative values, we know what $\sqrt{A}$ means.

This taking of continuous functions of a Hermitian operator is often of practical interest in itself, that is, without reference to applications to the spectral theorem.

Exercise 355. Consider the bounded operators A_α on $L^2(X)$. State and prove: $f(A_\alpha) = A_{f\circ\alpha}$.

Exercise 356. Let A and A' be Hermitian operators on Hilbert space H, with $AA' = A'A$. Prove that, for any continuous functions f and f', $f(A)f'(A') = f'(A')f(A)$.

Exercise 357. Consider the function f with action $f(x) = x^{-1}$. Show that, for A a Hermitian operator, with zero not in the spectrum of A, $f(A)$ exists and is A^{-1}.

Exercise 358. Let A be a Hermitian operator with positive spectrum. Prove that $A^m A^n = A^{m+n}$ and $(A^m)^n = A^{mn}$ for (not necessarily integral) real m and n.

Exercise 359. Let A be a Hermitian operator. Let f be a continuous function on the spectrum of A, with $f(x)$ always either zero or one for x in $\sum(A)$.

Prove that $f(A)$ is a projection operator.

Exercise 360. Let A be a Hermitian operator. Let f be a continuous, complex-valued function on $\sum(A)$, with $|f(x)| = 1$ for x in $\sum(A)$. Prove that $f(A)$ is unitary. (In particular, e^{iA} is unitary.)

Exercise 361. Why cannot one easily introduce continuous functions of arbitrary bounded operators?

Exercise 362. Prove that, if f is a positive continuous function and A is Hermitian, then $f(A)$ is invertible.

Exercise 363. Prove that any Hermitian operator whose spectrum does not include zero can be written as the difference of two Hermitian operators with positive spectrum.

Exercise 364. Prove that, if f and g are continuous real functions, then $f(g(A)) = (f \circ g)(A)$.

53

Other Functions of a Hermitian Operator

The purpose of this chapter is to show that one can give reasonable meaning to the sentence "Apply the function f to Hermitian operator A" for certain real, not necessarily continuous, functions f on $\sum(A)$. It is convenient to precede this generalization of theorem 62 by the introduction of certain additional structure on the set of Hermitian operators on a Hilbert space.

Fix a Hilbert space H, and denote by $\mathbf{H}$ the collection of all Hermitian operators on H. Then, of course, this $\mathbf{H}$ has the structure of a real topological vector space (where the topology on $\mathbf{H}$ is that which comes from the norm). In fact, there is still more structure on this $\mathbf{H}$, structure we now introduce. For A and B Hermitian operators on H, write $A \leq B$ if $(h,A(h)) \leq (h,B(h))$ for every vector h in H (noting that $(h,A(h))$ and $(h,B(h))$ are real, since A and B are Hermitian). We claim that this "$\leq$" is a partial ordering on the set $\mathbf{H}$. It is obvious that $A \leq A$ for every A in $\mathbf{H}$, and that $A \leq B$ and $B \leq C$ imply $A \leq C$. Thus we have only to show that $A \leq B$ and $B \leq A$ (i.e., that $(h,A(h)) = (h,B(h))$ for all h) implies $A = B$. Since $(h,A(h)) = (h,B(h))$ for all h, we have $(h + h',A(h + h')) = (h + h',B(h + h'))$ for all h and h', that is, $(h,A(h)) + (h',A(h)) + (h,A(h')) + (h',A(h')) = (h,B(h)) + (h',B(h)) + (h,B(h')) + (h',B(h'))$, whence $(h',A(h)) + (h,A(h')) = (h',B(h)) + (h,B(h'))$ for all h and h'. Comparing this last equation with the result of replacing h' by ih' therein, we have $(h',A(h)) = (h',B(h))$, or, what is the same thing, $(h',(A - B)(h)) = 0$ for all h and h'. Since h' is arbitrary, $(A - B)(h) = 0$ for all h. That is, $A = B$. We conclude that "$\leq$" is indeed a partial ordering.

Thus the set $\mathbf{H}$ of Hermitian operators on Hilbert space H has the structure of a partially ordered set. We next wish to see how this partial ordering interacts with the other structure on $\mathbf{H}$.

Example. Consider the bounded operators A_α on $L^2(X)$. For α and β real (bounded, measurable) functions on X, A_α and A_β are Hermitian. We have $A_\alpha \leq A_\beta$ if and only if $\alpha \leq \beta$ almost everywhere.

The interaction with the algebraic operations is quite simple. If A, B, and C are Hermitian operators, with $A \leq B$, then $A + C \leq B + C$. [Proof: $(h,A(h)) + (h,C(h)) \leq (h,B(h)) + (h,C(h))$ for all h.] Furthermore, for $A \leq B$ and r a real number, $rA \leq rB$ if r is positive, and $rA \geq rB$ if r is negative. [Proof: $(h,(rA)(h)) = r(h,A(h))$ and $(h,(rB)(h)) = r(h,B(h))$.] Note also that, for A Hermitian, so is A^2, and $A^2 \geq 0$ (for, for any h, $(h,A^2(h)) = (A(h),A(h)) \geq 0$).

The interaction of the partial ordering with the spectrum is more subtle. We claim: for any A in **H**, $A \geq 0$ if and only if $\sum(A)$ contains only non-negative numbers. First, note that, if $\sum(A)$ contains only non-negative numbers, then $B = \sqrt{A}$ exists, with $B^2 = A$. Hence $A \geq 0$. To prove the converse, let $A \geq 0$, and let b be any positive real. We have only to show that $A + bI$ is invertible (for this will imply that $-b$ is a regular value of A and hence that $\sum(A)$ contains only non-negative numbers). First, note that, for any h, $||(A + bI)(h)||^2 = (A(h),A(h)) + 2b(h,A(h)) + b^2(h,h) \geq b^2||h||^2$, where we used $A \geq 0$ in the last step. The rest of the proof is identical to the argument regarding the spectrum of a Hermitian operator at the end of chapter 50. The mapping $H \xrightarrow{(A+bI)} H$ is one-to-one (since, by the above inequality, $(A + bI)$ cannot annihilate a nonzero vector). Set $V = (A + bI)[H]$, a vector subspace of vector space H. To prove that V is closed, let $\underline{h}_1 = (A + bI)(h_1)$, $\underline{h}_2 = (A + bI)(h_2)$, $\cdots$ be a sequence in V converging to $\underline{h}$. Then $\underline{h}_1,\underline{h}_2,\cdots$ is a Cauchy sequence in H, whence, by the inequality above, $h_1,h_2,\cdots$ is a Cauchy sequence, whence it converges to some vector h. Therefore $\underline{h} = (A + bI)(h)$ is in V. That is, V is closed. For h in $V^{\perp}$, $(h,(A + bI)(h')) = 0$ for every h', whence $((A + bI)(h),h') = 0$ for every h', whence $(A + bI)(h) = 0$, whence (since $(A + bI)$ is one-to-one) $h = 0$. Thus $V = H$, so $(A + bI)$ is both one-to-one and onto. Therefore there exists a linear mapping from H to H which is the inverse of $(A + bI)$, a mapping which, again by the inequality above, is continuous. Thus $(A + bI)$ is invertible. This proves the assertion at the beginning of this paragraph. Note that this relationship between the partial ordering and the spectrum permits us to define "$\geq$" in terms of spectra: $A \geq B$ if and only if $\sum(A - B)$ contains only non-negative numbers.

Since "$\geq$" has now been tied to the behavior of spectra, we can relate this partial ordering to norms. For example: if $A \geq B \geq 0$, then $|A| \geq |B|$. [Proof: Let $\kappa > |A|$, so, for some positive ϵ, $(\kappa - \epsilon) \geq |A|$. Then $(\kappa - \epsilon)I \geq A$, whence $(\kappa - \epsilon)I \geq B$. Since $-B + \kappa I - \epsilon I \geq 0$, the spectrum of $-B + \kappa I$ includes only positive numbers, whence $-B + \kappa I$ is invertible. Since every $\kappa > |A|$ is a regular value of B, $|A| \geq |B|$.]

There is one further respect in which the partial ordering on **H** is analogous to the partial ordering on the reals. Every nonincreasing sequence of real numbers, bounded below, converges to some real number. Similarly,

THEOREM 63. *Let $A_1,A_2,\cdots$ and B be in* **H**, *with $A_1 \geq A_2 \geq \cdots \geq B$. Then there exists one and only one Hermitian operator A such that* lim $A_n(h) = A(h)$ *for every h in H.*

Proof. Let $m \geq n$, so $A_n - A_m \geq 0$. Then there exists a Hermitian operator C with $C^2 = A_n - A_m$. For any vector h, we have $||(A_n - A_m)(h)||^2 = ||C^2(h)||^2 = ||C(C(h))||^2 \leq |C|^2||C(h)||^2 = |C|^2(C(h),C(h)) =$

$|C|^2(h,C^2(h)) = |C|^2(h,(A_n - A_m)(h))$. But $|C|^2 = |C^2| \leq |A_1 - B|$, where, in the first step, we have used $\sum (C^2) = (\sum (C))^2$ and, in the second, $C^2 \leq A_1 - B$. We conclude: $||(A_n - A_m)(h)||^2 \leq |A_1 - B|(h,(A_n - A_m)(h))$. Now, for fixed h, $(h,A_1(h)),(h,A_2(h)), \cdots$ is a nonincreasing sequence of real numbers, bounded below (by $(h,B(h))$), whence it is a Cauchy sequence. By the inequality just obtained, $A_1(h),A_2(h), \cdots$ is a Cauchy sequence in H. Hence it converges to some vector, which we write $A(h)$, thus defining a mapping $H \xrightarrow{A} H$. This mapping A is clearly unique. What remains, therefore, is to show that A is a Hermitian operator. This A is certainly linear, for, for h and h' in H and c a complex number, $A(h + ch') = \lim A_n(h + ch') = \lim(A_n(h) + cA_n(h')) = \lim A_n(h) + c \lim A_n(h') = A(h) + cA(h')$, where we have used continuity of addition and scalar multiplication in H. To see that A is continuous, set $n = 1$, and take the limit of infinite m in $||(A_n - A_m)(h)||^2 \leq |A_1 - B|(h,(A_n - A_m)(h))$ to obtain $||(A_1 - A)(h)||^2 \leq |A_1 - B|(h,(A_1 - A)(h)) \leq |A_1 - B|(h,(A_1 - B)(h)) \leq |A_1 - B|^2||h||^2$. Hence, for every h, $||A(h)|| \leq (|A_1| + |A_1 - B|)||h||$. Finally, to see that A is Hermitian, note that, for any vectors h and h', $(h,A(h')) = (h, \lim A_n(h')) = \lim(h,A_n(h')) = \lim(A_n(h),h') = (\lim A_n(h),h') = (A(h),h')$, where we have used continuity of the inner product. ☐

One might well wonder why the statement of theorem 63 ends with "$\cdots$ such that $\lim A_n(h) = A(h)$ for every h in H" rather than simply "$\cdots$ such that $\lim A_n = A$." The reason is that the theorem would be made false by such an alteration.

Example. Let H be the Hilbert space of all sequences, $(c_1,c_2, \cdots)$, of complex numbers with $|c_1|^2 + |c_2|^2 + \cdots$ finite. For $n = 1,2, \ldots,$ let A_n be the bounded operator on H with action $A_n(c_1,c_2, \cdots) = (0, \ldots ,0,c_n,c_{n+1}, \cdots)$. Then each A_n is Hermitian (in fact, a projection operator), with $A_n \geq 0$. Furthermore, $A_1 \geq A_2 \geq \cdots$ Clearly, for every vector h, $\lim A_n(h) = 0$, so $A = 0$ is the bounded operator whose existence is guaranteed by theorem 63. Does $\lim A_n = 0$, that is, does $\lim |A_n| = 0$? The answer is no, for, since each A_n is a nonzero projection operator, $|A_n| = 1$ for every n.

As this example shows, theorem 63 states that "A is the limit of $A_1,A_2, \cdots$" in a sense somewhat different from that which comes from the (norm) topology on **H**. This alternative notion of "limit of a sequence of Hermitian operators" will play an important role in what follows.

These preliminaries out of the way, we now return to the question at hand: the application of (not necessarily continuous) functions to a Hermitian operator. The idea is to "approximate such a function by a nonincreasing sequence of continuous functions, and then use theorems 62 and 63." It is convenient to first settle on the class of functions with which we shall be concerned.

Fix a Hermitian operator A on Hilbert space H, and consider $\sum(A)$. A real-valued function f on $\sum(A)$ is said to be ***upper semicontinuous*** if there exists a sequence, $f_1, f_2, \ldots$, of continuous, real-valued functions on $\sum(A)$, with $f_1 \geq f_2 \geq \cdots$, and with $\lim f_n(x) = f(x)$ for every x in $\sum(A)$.

Example. Fix a real number κ, and let θ_κ be the function with $\theta_\kappa(x)$ one if $x \leq \kappa$, and zero otherwise (figure 144). Then θ_κ is upper semicontinuous. On the other hand, $-\theta_\kappa$ is not upper semicontinuous.

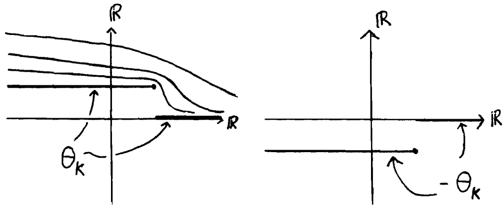

Figure 144

We note that every continuous function f is upper semicontinuous (choosing $f_1 = f_2 = \cdots = f$) and that the sum of upper semicontinuous functions f and f' is upper semicontinuous (for, if $f_1, f_2, \cdots$ and $f'_1, f'_2, \cdots$ are the corresponding sequences of continuous functions, $(f_1 + f'_1) \geq (f_2 + f'_2) \geq \cdots$, and $\lim(f_n + f'_n)(x) = (f + f')(x)$ for each x in $\sum(A)$). Furthermore, if f and f' are non-negative upper semicontinuous functions, then so is ff' (for, if $f_1, f_2, \cdots$ and $f'_1, f'_2, \cdots$ are the corresponding sequences of continuous functions, then $f_1 f'_1 \geq f_2 f'_2 \geq \cdots$, and $\lim(f_n f'_n)(x) = (ff')(x)$ for each x in $\sum(A)$). The only problem with the set of upper semicontinuous functions is that it does not have a very rich algebraic structure, for example, it does not form a vector space (as the example above shows). We recover a more satisfactory structure (and thus make theorems easier to state) by brute force.

Denote by $\tilde{\mathbf{C}}$ the collection of all real-valued functions f on $\sum(A)$ that can be written in the form $f = u - u'$, with each of u and u' upper semicontinuous and non-negative. This $\tilde{\mathbf{C}}$ is the set of functions with which we shall be concerned: we wish to decide what structure it has. First, note that every continuous function on $\sum(A)$ is in $\tilde{\mathbf{C}}$, so we have $\mathbf{C} \subset \tilde{\mathbf{C}}$. Next, note that the sum of two functions in $\tilde{\mathbf{C}}$ is in $\tilde{\mathbf{C}}$, for, if $f = u - u'$ and $g = v - v'$, then $f + g = (u + v) - (u' + v')$. Furthermore, for $f = u - u'$ in $\tilde{\mathbf{C}}$, and r a real number, rf is in $\tilde{\mathbf{C}}$ [proof: write $rf = (ru) - (ru')$ if r is positive and $rf = (-ru') - (-ru)$ if r is negative]. In short, $\tilde{\mathbf{C}}$ has the structure of a real vector space. We next note that the product of two functions, $f = u - u'$ and $g = v - v'$, in $\tilde{\mathbf{C}}$ is in $\tilde{\mathbf{C}}$ (for $fg = (uv + u'v') - (uv' + u'v)$). Thus this $\tilde{\mathbf{C}}$ has the structure of an associative algebra. Note, incidentally, that **C** is a subalgebra of $\tilde{\mathbf{C}}$. We next wish to define a topology on this associative algebra $\tilde{\mathbf{C}}$. First, note that, for u a non-negative, upper semicontinuous function on $\sum(A)$, u is bounded (for, for $f_1, f_2, \cdots$ the corresponding sequence of continuous functions, $f_1 \geq u \geq 0$, while f_1, as a continuous function on a compact set, is bounded). For f and g in $\tilde{\mathbf{C}}$, set $d(f,g) =$ least upper bound of $|f(x) - g(x)|$, for x in $\sum(A)$ (noting that, by the previous remark, this least upper bound exists). This is a metric on $\tilde{\mathbf{C}}$, and hence defines a topology on $\tilde{\mathbf{C}}$. Note that this topology is "the same" as that which we introduced on the continuous functions preceding theorem 62. More precisely, **C** is a subspace of topological space $\tilde{\mathbf{C}}$. To summarize: $\tilde{\mathbf{C}}$ is an associative algebra with a topology.

The purpose of this chapter is to establish the following generalization of theorem 62:

THEOREM 64. *Let A be a Hermitian operator on Hilbert space H. Then there exists one and only one mapping $\tilde{\mathbf{C}} \xrightarrow{\tilde{\psi}} \mathbf{B}$ satisfying the following four conditions: i) $\tilde{\psi}$ is a homomorphism of associative algebras, ii) $\tilde{\psi}$ is a continuous mapping of topological spaces, iii) for f the element of $\tilde{\mathbf{C}}$ with action $f(x) = a_0 + a_1 x$ (a_0 and a_1 real numbers), $\tilde{\psi}(f)$ is the bounded operator $a_0 I + a_1 A$, and iv) for $f_1 \geq f_2 \geq \cdots \geq f \geq 0$ in $\tilde{\mathbf{C}}$, with each f_n continuous and with* $\lim f_n(x) = f(x)$ *for every x in $\sum(A)$, we have* $\lim \tilde{\psi}(f_n)(h) = \tilde{\psi}(f)(h)$ *for every h in H.*

It is convenient, as was the case with theorem 62, to mix the motivation with the proof.

Let us first agree to use the more suggestive notation $f(A)$ rather than $\tilde{\psi}(f)$. Now, the first three conditions of the theorem are identical with the three conditions of theorem 62, and, furthermore, $\mathbf{C} \subset \tilde{\mathbf{C}}$. Hence we must have $\tilde{\psi} = \psi$ on **C** (i.e., whenever both mappings are defined). In other words, we "already know" what $f(A)$ means for f continuous. The thrust of the theorem, therefore, is that condition iv) permits one to "extend the definition of $f(A)$ (uniquely)" to functions f which need not be continuous, namely, to functions in $\tilde{\mathbf{C}}$. Suppose first that f and f' are continuous functions, with $f \geq f'$. Then $f - f' \geq 0$, whence there exists a Hermitian operator B with $(f - f')(A) = B^2$, whence, for every vector h, $(h,(f - f')(A)(h)) = (h, B^2(h)) = (B(h), B(h)) \geq 0$. That is, $f(A) \geq f'(A)$. In other words, "larger continuous functions, when applied to A, give larger Hermitian operators." Now let f be upper semicontinuous (we might as well assume $f \geq 0$, since this could always be accomplished by adding a constant). Let $f_1, f_2, \cdots$ be a sequence of continuous functions, with $f_1 \geq f_2 \geq \cdots$ and with $\lim f_n(x) = f(x)$ for every x in $\sum(A)$. Then, by the remark above, we have $f_1(A) \geq f_2(A) \geq \cdots \geq 0$. We now use theorem 63: there is a unique Hermitian operator C such that lim

$f_n(A)(h) = C(h)$ for every h in H. This C is our candidate for $f(A)$. That is, we "know how to apply $f_1, f_2, \cdots$ to A (since these functions are continuous), but we do not yet know how to apply f (not necessarily being continuous) to A." We use theorem 63 to get a candidate.

The next step is to show that this "candidate" depends only on f itself and not on the choice of the sequence $f_1, f_2, \cdots$ to this end, let $f'_1 \geq f'_2 \geq \cdots$ be another sequence of continuous functions, with $\lim f'_n(x) = f(x)$ for every x in $\sum(A)$, and let C' be the corresponding Hermitian operator. We must show $C = C'$. Fix the integer n, and fix positive ϵ, and consider the function $f_n + \epsilon$. For each integer m, let K_m consist of all points x of $\sum(A)$ for which $f'_m(x) \geq f_n(x) + \epsilon$. Then each K_m (as a closed subset of compact set $\sum(A)$) is compact. Since $f'_1 \geq f'_2 \geq \cdots$, we have $K_1 \supset K_2 \supset \cdots$ Furthermore, since $\lim f'_m(x) \leq f(x) < f_n(x) + \epsilon$, we have $\bigcap K_m = \emptyset$. We claim that these properties imply that some K_m is empty. [Proof: If not, choose x_1 in K_1, x_2 in K_2, etc. Then $x_1, x_2, \ldots$, as a sequence in a compact set, has an accumulation point x. By construction, this x is in every K_m, whence x is in $\bigcap K_m$, violating $\bigcap K_m = \emptyset$.] Thus, for some m, $f'_m \leq f_n + \epsilon$, whence $f'_m(A) \leq f_n(A) + \epsilon I$. Therefore $C' \leq f_n(A) + \epsilon I$, whence, since ϵ is arbitrary, $C' \leq f_n(A)$, whence, since n is arbitrary, $C' \leq C$. Similarly, $C \leq C'$, so $C = C'$.

What we have shown so far is that, if there is to be any $\tilde{\psi}$ satisfying the conditions of theorem 64, it must agree with ψ on the continuous functions and it must have the action described above for non-negative, upper semicontinuous functions. Next, let f be any function in $\tilde{\mathbf{C}}$, and write $f = u - u'$, with u and u' non-negative and upper semicontinuous, Then, since $\tilde{\psi}$ is to be a homomorphism, we must choose, for our candidate for $f(A)$, $f(A) = u(A) - u'(A)$ (noting that $u(A)$ and $u'(A)$ were defined above). We have next to show that this candidate is independent of the choice of u and u', that is, that, if $f = v - v'$ with each of v and v' non-negative and upper semicontinuous, then $u(A) - u'(A) = v(A) - v'(A)$. Noting that $u - u' = v - v'$ implies $u + v' = u' + v$, and hence $(u + v')(A) = (u' + v)(A)$, it suffices to prove that $(u + v')(A) = u(A) + v'(A)$. But this is immediate, for, for $f_1 \geq f_2 \geq \cdots$ and $f'_1 \geq f'_2 \geq \cdots$ continuous, with $\lim f_n(x) = u(x)$ and $\lim f'_n(x) = v'(x)$ for each x in $\sum(A)$, we have $(u + v')(A)(h) = \lim(f_n + f'_n)(A)(h) = \lim(f_n(A)(h) + f'_n(A)(h)) = \lim f_n(A)(h) + \lim f'_n(A)(h) = u(A)(h) + v(A)(h)$. Thus $f(A) = u(A) - u'(A)$ is independent of the decomposition, $f = u - u'$, of f.

To summarize: we have defined a mapping $\tilde{\mathbf{C}} \xrightarrow{\tilde{\psi}} \mathbf{B}$ and have shown that this is the only mapping which is even a candidate for satisfying the four conditions of theorem 64. What remains, therefore, is to show that this $\tilde{\psi}$ does in fact satisfy conditions i)–iv). For condition i), first, note that, for u and v non-negative and upper semicontinuous, $(u + v)(A) = u(A) + v(A)$ (as shown in the previous paragraph). It is obvious, furthermore, that, for u non-negative and upper semicontinuous, and r a real number, $(ru)(A) = ru(A)$. Thus $\tilde{\psi}$ is indeed a linear mapping of real vector spaces. To show that $\tilde{\psi}$ is a homomorphism, we must show that, for u and v non-negative and upper semicontinuous, $(uv)(A) = u(A)v(A)$. But this is immediate, for, for $f_1 \geq f_2 \geq \cdots$ and $f'_1 \geq f'_2 \geq \cdots$ the corresponding sequences of continuous functions, $(uv)(A)(h) = \lim(f_n f'_n)(A)(h) = \lim f_n(A)f'_n(A)(h) = \lim f_n(A)(\lim f'_n(a)(h)) = \lim f_n(A)(v(A)(h)) = u(A)v(A)(h)$. Hence $\tilde{\psi}$ is a homomorphism of real associative algebras. For condition ii), we must show that, for f upper semicontinuous and non-negative, with $|f| \leq \epsilon$, $|f(A)| \leq \epsilon$. Let δ be any positive number, and choose continuous function g on $\sum(A)$ with $g \geq f$ and $|g| \leq \epsilon$. Then $|g(A)| \leq \epsilon + \delta$. But $g \geq f \geq 0$, whence $g(A) \geq f(A) \geq 0$, so $|f(A)| \leq |g(A)| \leq \epsilon + \delta$. Since δ is arbitrary, $|f(A)| \leq \epsilon$. Thus $\tilde{\psi}$ is continuous. Conditions iii) and iv) are obvious.

This completes the proof of theorem 64. Note, incidentally, that each bounded operator $f(A)$ is Hermitian (although we could not replace "**B**" by "**H**" in theorem 64, for **H** is not an associative algebra). Furthermore, in the proof of continuity, we actually proved somewhat more: for f and f' in $\tilde{\mathbf{C}}$, with $d(f,f') \leq \epsilon$, $|f(A) - f'(A)| \leq \epsilon$.

A few remarks about theorems 62 and 64 may better explain what is going on. First, theorem 62 is rather useful for a variety of applications: one occasionally wishes to evaluate a continuous function of a Hermitian operator. Theorem 64, on the other hand, is somewhat less useful, for it is more rare that one wants a discontinuous function of a Hermitian operator. Furthermore, even when theorem 64 is used, it is normally only the first three properties which are of interest. What, then, is the role of the fourth condition? Consider the statement which is theorem 64 with condition iv) omitted.

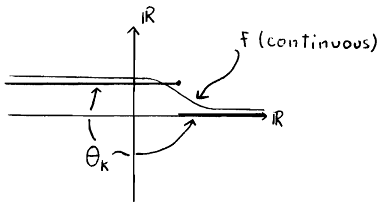

Figure 145

There certainly exists a mapping $\tilde{\mathbf{C}} \xrightarrow{\tilde{\psi}} \mathbf{B}$ satisfying the remaining three conditions (namely, the mapping whose existence is guaranteed by theorem 64 itself). Is this $\tilde{\psi}$ unique? In fact, the answer is no, a point we may best illustrate by the following observation. Consider the function θ_κ in $\tilde{\mathbf{C}}$ (figure 145). The best chance of having a unique bounded operator $\theta_\kappa(A)$ would be if we could suitably approximate θ_κ by a continuous function. What notion of "suitably approximate" must we use? We are forced, since our statement refers only to the topology of $\tilde{\mathbf{C}}$, to use that topology. We thus ask: is it true that, for every positive ϵ, there is a continuous f with $|\theta_\kappa - f| \leq \epsilon$? It is clear that there is none (e.g., for $\epsilon = 1/3$). What happens in theorem 64, then, is that one introduces a weaker notion of "nearby functions" (pointwise limits of nonincreasing functions) and a weaker notion of "nearby operators" (vector-wise limits on H), and thereby becomes able to sensibly speak of $f(A)$ for a wider class of functions f. This extension—from continuous functions to a wider class—nonetheless continues to satisfy the conditions (homomorphism, continuity) that were satisfied in the continuous case.

Exercise 365. Prove that, for f in $\tilde{\mathbf{C}}$, $\sum (f(A)) = f(\sum (A))$.

Exercise 366. Consider the following topology on $\mathbf{B}$. To define a neighborhood of A in $\mathbf{B}$, fix n vectors, $h_1, \ldots, h_n$, and positive ϵ. Consider all B in $\mathbf{B}$ with $\|(A - B)(h_1)\| \leq \epsilon, \ldots, \|(A - B)(h_n)\| \leq \epsilon$. Prove that this defines a topology on $\mathbf{B}$ and that theorem 63 refers to convergence in this topology.

Exercise 367. Fix Hermitian operator A. Consider the mapping from $\mathbf{C}$ to the set of all compact subsets of the real line which associates, with f in $\mathbf{C}$, $\sum (f(A))$. Find a suitable topology on this set of compact subsets for which this mapping is continuous.

Exercise 368. Is there some natural "maximal" set of functions which could replace $\tilde{\mathbf{C}}$ for which theorem 64 would still work?

Exercise 369. State and prove: $\tilde{\psi}$ is order-preserving. Could such a condition satisfactorily replace condition iv) in theorem 64?

Exercise 370. On the set $\mathbf{R}$ of reals, let the open sets be $\emptyset$, $\mathbf{R}$ itself, and sets of the form $(-\infty, a)$. Prove that this is a topology on the set $\mathbf{R}$. Prove that a function $\mathbf{R} \to \mathbf{R}$ is upper semicontinuous if and only if it is continuous, where the $\mathbf{R}$ on the left is the real line, and that on the right the topological space above. Define lower semicontinuous, and prove that a function, if both upper semicontinuous and lower semicontinuous, is continuous.

Exercise 371. Find an example of a measurable function not in $\tilde{\mathbf{C}}$.

Exercise 372. Prove that every function in $\tilde{\mathbf{C}}$ that is a limit (in the topology of $\tilde{\mathbf{C}}$) of a net of continuous functions is continuous.

Exercise 373. Define the absolute value of Hermitian operator A, and compare its norm with that of A.

The Spectral Theorem

We establish, in this chapter, an infinite-dimensional generalization of theorem 61, a result for which the previous two chapters are the groundwork. As it turns out, one requires only a single fact from these two chapters, namely, theorem 64.

Fix a Hilbert space H, and a Hermitian operator A on H. For each real number κ, let θ_κ be the real-valued function with action $\theta_\kappa(x)$ one if $x \leq \kappa$, and zero otherwise. Then θ_κ, regarded as a function on $\sum(A)$, is in $\tilde{\mathbf{C}}$ (and, in fact, is itself upper semicontinuous). Thus, by theorem 64, there is a Hermitian operator $P_\kappa = \theta_\kappa(A)$. For each real κ, we obtain a bounded operator P_κ: this family of bounded operators is called the *spectral family* of A.

Example. Consider the bounded operators A_α on $L^2(X)$. Fix a real-valued, measurable, founded function α on X, so A_α is a Hermitian operator on $L^2(X)$. Fix a real number κ. Then $\theta_\kappa(A_\alpha) = A_{\theta_\kappa \circ \alpha}$. Now, $\theta_\kappa \circ \alpha$ is the

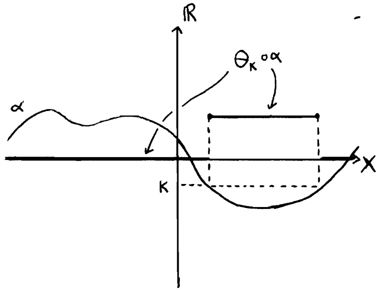

Figure 146

function on X whose value is one for x with $\alpha(x) \leq \kappa$, and zero otherwise (figure 146). Thus the action of $P_\kappa = \theta_\kappa(A)$ is the following: for f a representative of an element of $L^2(X)$, P_κ sends this f to $(\theta_\kappa \circ \alpha)f$, that is, sends this f to the function whose value is $f(x)$ for x with $\alpha(x) \leq \kappa$, and zero otherwise (figure 147). This is the structure of the spectral family of A_α.

There are three important properties of the spectral family of Hermitian operator A. First, note that, for any real κ, the function θ_κ satisfies $\theta_\kappa\theta_\kappa = \theta_\kappa$ (for this function takes only the values zero and one). Since "apply functions to A" is a homomorphism of associative algebras, we have $\theta_\kappa(A)\theta_\kappa(A) =$

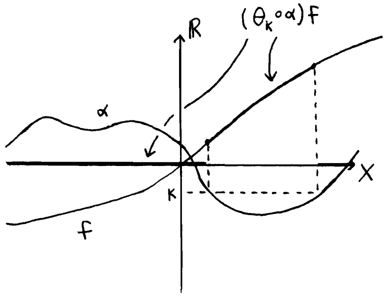

Figure 147

$\theta_\kappa(A)$, that is, $P_\kappa P_\kappa = P_\kappa$. That is, each P_κ is a projection operator. Since, furthermore, for $\kappa < \kappa'$, we have $\theta_\kappa \theta_{\kappa'} = \theta_\kappa$ for functions, it follows that $P_\kappa P_{\kappa'} = P_{\kappa'} P_\kappa = P_\kappa$ for the spectral family. Finally, for $\kappa < \sum(A)$ (i.e., for κ less than each number in the spectrum of A), the function θ_κ vanishes on $\sum(A)$, whence $\theta_\kappa(A) = 0$, that is, $P_\kappa = 0$. Similarly, for $\kappa \geq \sum(A)$, $P_\kappa = I$. These algebraic properties can be expressed more geometrically. Since each P_κ is a projection operator, each is the projection onto some subspace, V_κ, of the Hilbert space H. For $\kappa < \kappa'$, we have $P_\kappa P_{\kappa'} = P_{\kappa'} P_\kappa = P_\kappa$: we claim that this means that $V_\kappa \subset V_{\kappa'}$. Indeed, let v be a vector in V_κ, so $P_\kappa(v) = v$. Then $P_{\kappa'}(v) = P_{\kappa'}(P_\kappa(v)) = P_{\kappa'} P_\kappa(v) = P_\kappa(v) = v$, so v is also in $V_{\kappa'}$. Finally, for $k < \sum(A)$, $P_\kappa = 0$, so V_κ is the zero subspace of H, while, for $\kappa \geq \sum(A)$, V_κ is the entire Hilbert space H.

Consider now the behavior of V_κ, the subspace onto which P_κ projects, as κ increases. For $\kappa < \sum(A)$, $V_\kappa = 0$. As κ begins to increase through the spectrum of A, V_κ becomes larger than the zero subspace. As κ increases, the size of V_κ increases. Finally, when κ has finally passed through the spectrum, that is, when $\kappa \geq \sum(A)$, V_κ has increased to become the entire Hilbert space. "The size of V_κ normally increases continuously as κ increases." Suppose that this were not the case, that is, let k be a nonzero vector and κ_0 a real number with $P_{\kappa_0}(k) = k$, and with $P_\kappa(k) = 0$ for $\kappa < \kappa_0$. Then, as we shall show shortly, k is an eigenvector of A with eigenvalue κ_0. Thus "sudden increases in the size of the subspace V_κ with κ indicate eigenvectors for that point of the spectrum, while gradual increases allow more and more vectors to be included in V_κ as κ increases without these vectors having to be eigenvectors." In this way "eigenvector-like structure" of A is described without a commitment to A's actually having eigenvectors.

It should be noted that all these (general) remarks about the spectral family are precisely what one knows (explicitly) to be true in the example above as well as in the finite-dimensional case (chapter 51).

We now come, finally, to the spectral theorem for Hermitian operators.

THEOREM 65. *Let A be a Hermitian operator on Hilbert space H. Then, for every positive number* ϵ *there exists a number* δ *with the following property: for any real numbers* $\kappa_1 < \kappa_2 < \cdots < \kappa_{n+1}$ *with* $\kappa_1 < \sum(A) < \kappa_{n+1}$ *and* $|\kappa_{i+1} - \kappa_i| \leq \delta$ *for* $i = 1, 2, \ldots, n$, *we have* $|A - \sum_1^n \kappa_i(P_{\kappa_{i+1}} - P_{\kappa_i})| \leq \epsilon$.

Proof. Let f be the function on $\sum(A)$ with action $f(x) = x$, so $f(A) = A$. By continuity in theorem 64, there exists, given ϵ, a δ such that, for any function g in $\tilde{\mathbf{C}}$ with $|f - g| \leq \delta$, $|f(A) - g(A)| \leq \epsilon$. (In fact, by the remark following theorem 64, $\delta = \epsilon$ would do.) Fix this δ, and, for $\kappa_1, \ldots, \kappa_{n+1}$ as in the theorem, let $g = \sum_1^n \kappa_i(\theta_{\kappa_{i+1}} - \theta_{\kappa_i})$ (figure 148). Then $|f - g| \leq \delta$, whence $|f(A) - g(A)| \leq \epsilon$. But $f(A) = A$, and $g(A) = \sum \kappa_i(\theta_{\kappa_{i+1}}(A) - \theta_{\kappa_i}(A)) = \sum \kappa_i(P_{\kappa_{i+1}} - P_{\kappa_i})$. □

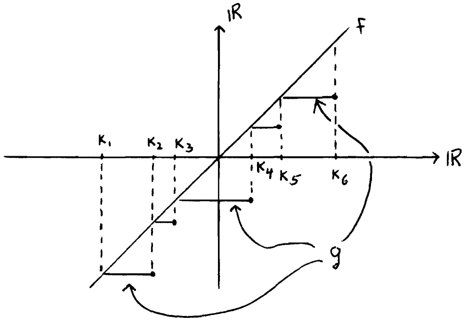

Figure 148

The thrust of the theorem is that "A can be approximated, as closely as one wishes, by an appropriate linear combination of the operators in the spectral family of A." Since one "understands projection operators completely," one, in some sense, "understands completely the structure of Hermitian operators." Note that this theorem, together with the properties of the spectral family, represents an infinite-dimensional generalization of theorem 61.

The second sentence in the statement of theorem 65 is often reexpressed as follows: "Then $A = \int_{\sum(A)} \kappa \, dP_\kappa$." (The notation will be recognized by those familiar with the Riemann-Stieltjes integral.) One's immediate reaction to the symbol "$\int$" is "What is the measure space in terms of which all this would be most neatly formulated?" The underlying set for this measure space would be $\sum(A)$, the spectrum of A. The problem, however, is that the "measure of a set" must be, not an element of R^*, but rather a certain projection operator. Then, it would be claimed, "the integral of the real-valued function with action $f(\kappa) = \kappa$, over this 'measure space,' is precisely the original Hermitian operator A." In fact, such a program could be carried out. One has to redo the subject of measure spaces and integrals to allow "the measure to be valued in more general spaces, for example, the space of projection operators on a Hilbert space." Such an ambitious program is, perhaps, not worthwhile if its only purpose is to simplify the statement of a single theorem.

Let f be any continuous function on $\sum(A)$. A check of the proof of theorem 65 shows that the theorem remains valid if the last equation is replaced by $|f(A) - \sum f(\kappa_i)(P_{\kappa_{i+1}} - P_{\kappa_i})| \leq \epsilon$. In integral form, $f(A) = \int_{\sum(A)} f(\kappa) dP_\kappa$. Thus, taking continuous functions of A has the expected effect on the spectral decomposition: one just applies the function to κ.

Let us now use theorem 65 to prove a statement we claimed earlier without proof: if k is a nonzero vector and κ_0 a real number with $P_{\kappa_0}(k) = k$, and $P_\kappa(k) = 0$ for $\kappa < \kappa_0$, then k is an eigenvector of A with eigenvalue κ_0. Fix positive ϵ, and let $\kappa_1, \ldots, \kappa_{n+1}$ be as in theorem 65. Then, from the conclusion of that theorem, $||A(k) - \sum \kappa_i(P_{\kappa_{i+1}}(k) - P_{\kappa_i}(k))|| \leq \epsilon||k||$. But, by assumption, $\sum \kappa_i(P_{\kappa_{i+1}}(k) - P_{\kappa_i}(k)) = \kappa_i k$, where i is the integer with κ_0 in $(\kappa_i, \kappa_{i+1}]$. Thus $||A(k) - \kappa_i k|| \leq \epsilon||k||$. We conclude: for every positive ϵ there is a δ such that, whenever $|\kappa - \kappa_0| \leq \delta$, $||A(k) - \kappa k|| \leq \epsilon||k||$. Clearly, this is impossible unless $A(k)$ happens to be the vector $\kappa_0 k$.

The spectral theorem represents a rather brute-force way of dealing with Hermitian operators. If one wants to say something about a Hermitian operator, one first approximates it by an appropriate linear combination of the bounded operators in its spectral family, then tries to make one's statement about these approximations, and finally, using the fact that one can "make the approximation as good as one wants," tries to extend the validity of one's statement to the original Hermitian operator itself. For example, almost every statement we have made about Hermitian operators can now be proven, more or less directly, from the spectral theorem ("directly" in the sense that one does not need clever ideas; usually, but not always, "easily"). We give one further example of this viewpoint.

In chapter 23, we proved Schur's theorem in the finite-dimensional case and remarked that there are infinite-dimensional generalizations. We now establish one such. We claim: let B_λ (λ in Λ) be a collection of bounded operators on Hilbert space H, such that the only subspaces V of H with $B_\lambda(v)$ in V for every v in V and λ in Λ are $V = 0$ and $V = H$. Then a Hermitian operator A with $AB_\lambda = B_\lambda A$ for every λ must be a multiple of the identity operator. Proof: Since $AB_\lambda = B_\lambda A$ for each λ, this is true if A is replaced by any polynomial in A, hence if A is replaced by any continuous function of A, hence if A is replaced by any function (in $\tilde{\mathbf{C}}$) of A. Hence $P_\kappa B_\lambda = B_\lambda P_\kappa$ for each κ. For v in the subspace V_κ on which P_κ projects, $P_\kappa(B_\lambda(v)) = B_\lambda(P_\kappa(v)) = B_\lambda(v)$ for each λ. Hence V_κ is an invariant subspace of the B_λ, whence each V_κ is either zero or H. We conclude: there is a number κ_0 such that $P_\kappa = I$ for $\kappa \geq \kappa_0$ and $P_\kappa = 0$ for $\kappa < \kappa_0$. Hence $\sum \kappa_i(P_{\kappa_{i+1}} - P_{\kappa_i}) = \kappa_i I$, where i is the integer with κ_0 in $(\kappa_i, \kappa_{i+1}]$. By theorem 65, for every positive ϵ, there is a δ such that, whenever $|\kappa - \kappa_0| \leq \delta$, $|A - \kappa I| \leq \epsilon$. This is possible, clearly, only if $A = \kappa_0 I$.

Exercise 374. Let A be a Hermitian operator on Hilbert space H. A point κ of $\sum(A)$ is said to be in the discrete spectrum of A if some neighborhood of κ (in the real line) contains no other points of $\sum(A)$. Prove that every point of the discrete spectrum is an eigenvalue. Is every eigenvalue a point of the discrete spectrum?

Exercise 375. Prove, from the spectral theorem, that the spectrum of A is a closed subset of $[-|A|,|A|]$, including at least one endpoint.

Exercise 376. Find an infinite-dimensional version of the finite-dimensional fact that two commuting Hermitian matrices can be simultaneously diagonalized.

Exercise 377. Prove theorem 62 directly from the spectral theorem.

Exercise 378. Prove theorem 61 directly from the spectral theorem.

Exercise 379. Prove, from the spectral theorem, that, for continuous f and Hermitian A, $\sum(f(A)) = f(\sum(A))$.

Exercise 380. Let there be a point of $\sum(A)$ in $(\kappa, \kappa']$. Prove that $P_\kappa \neq P_{\kappa'}$.

Exercise 381. Find a spectral decomposition for a bounded operator that satisfies $AA^* = A^*A$. (Hint: Real and imaginary parts, together with exercise 376.)

Exercise 382. Prove that two Hermitian operators with the same spectral family are equal.

Exercise 383. Let A and B be Hermitian operators. Find a necessary and sufficient condition, in terms of the spectral families, that there exists a continuous f with $B = f(A)$.

55

Operators (Not Necessarily Bounded)

The spectral theorem for Hermitian operators is a pretty and subtle—and perhaps even rather surprising—result. The sad truth, however, is that this theorem is not very useful for physical applications. An example from quantum mechanics will illustrate this point.

Consider a (spin-zero) one-dimensional particle (e.g., in some potential). The set of "classical configurations" of the particle is represented by the points of the real line **R**. The set of "quantum states" of the particle, therefore, would be represented by $L^2(\mathbf{R})$. Consider an element of $L^2(\mathbf{R})$, represented by a complex-valued, measurable, square-integrable function ψ on the measure space of reals. This ψ is called the wave function of the particle (the thing which describes the quantum state of the particle). One often hears the assertion that "observables on the quantum system are to be represented by Hermitian operators on its Hilbert space of quantum states." Thus the configuration observable should have the following action: the state represented by wave function ψ is to be sent to the state represented by the wave function with action $x\psi(x)$. Unfortunately, there is a problem: it is not true in general that, for ψ complex-valued, measurable, and square-integrable, so is $x\psi$. (Although $x\psi$ is certainly complex-valued and measurable, it is not in general square-integrable, e.g., for ψ with action $\psi(x) = (1 + |x|)^{-1}$.) Thus these instructions do not even define a mapping from the Hilbert space $L^2(\mathbf{R})$ to itself, much less a Hermitian operator. Similarly, for the momentum observable, one wishes the following action: ψ is to be sent to the function with action $i(d/dx)\psi(x)$ (where we have suppressed Planck's constant). But this is not well defined (as a mapping from $L^2(\mathbf{R})$ to $L^2(\mathbf{R})$) either, for complex-valued, measurable, square-integrable ψ need not be differentiable (and, in fact, need not even be continuous). Thus, whatever the configuration observable and momentum observable are to be, they do not seem to be bounded operators on $L^2(X)$.

This may seem like a rather technical issue. Why does one allow all those exotic wave functions (i.e., those not differentiable, those which do not die off quickly enough at infinity) to be in the Hilbert space? Why not try to make a Hilbert space out of, say, the C^∞ complex-valued functions of compact support? The set of such functions certainly forms a complex vector space. Furthermore, for two such functions, f and f', ff' is certainly integrable, so we could define the inner product by $(f,f') = \int_{\mathbf{R}} \bar{f}f'\, d\mu$. This inner product has the appropriate linearity and positivity properties for a Hilbert space. Unfortunately, we do not obtain in this way a Hilbert space, for the resulting topological vector space is not complete (a Cauchy sequence of C^∞ functions of compact support need not converge to a C^∞ function of compact support). The completeness property of Hilbert spaces has, of course, played a central role in our treatment of these spaces. One would not like to give up all the technology that has been developed.

The remarks above suggest that the interesting things, at least for applications to quantum mechanics, will not be Hermitian (or even bounded) operators, but will rather be "things which are somewhat like Hermitian operators, except that their action is not defined on the entire Hilbert space." These physical ideas suggest that one introduce and study the following definition.

Let H be a Hilbert space. An *operator* on H consists of a vector subspace D_A (called the *domain* of A) of vector space H, together with a linear mapping $D_A \xrightarrow{A} H$ of complex vector spaces.

Example. Let A be a bounded operator on Hilbert space H. Then, setting $D_A = H$, this A is also an operator on our Hilbert space.

Example. Consider the Hilbert space $L^2(\mathbf{R})$. Let D_A consist of all elements of $L^2(\mathbf{R})$ having, as a representative, a C^∞ function of compact support. Then (since linear combinations of such functions are again such), D_A is a vector subspace of vector space $L^2(\mathbf{R})$. For h an element of D_A with representative f (C^∞, compact support), let $A(h)$ be the element of H with representative xf (noting that xf is necessarily complex-valued, measurable, and square-integrable). Thus we obtain an operator on H.

Example. Consider the Hilbert space $L^2(\mathbf{R})$, and let D_A consist of all elements of $L^2(\mathbf{R})$ having, as a representative, a C^∞ function of compact support. For h an element of D, with representative f (C^∞, compact support), let $A(h)$ be the element of H with representative $i(d/dx)f(x)$ (noting that this function is necessarily complex-valued, measurable (since it is continuous), and square-integrable). Thus we obtain an operator on H.

Two operators, A, D_A and B, D_B, on H are said to be *equal* if $D_A = D_B$ and $A = B$ on this common domain. Note that any modification of the domain (even a rather minor one, e.g., throwing away a few vectors) gives a different operator. It is convenient, in fact, to have available the following definition: operator A, D_A is said to be an *extension* of operator B, D_B if $D_A \supset D_B$, and $A = B$ whenever both are defined (i.e., on D_B). Thus, for example, two operators are equal if and only if each is an extension of the other.

It should be emphasized that two things have happened in the passage from bounded operators to operators. First, the requirement $D_A = H$ (for the bounded case) has been dropped. Second, we no longer require that $D_A \rightarrow H$ be continuous (where D_A is to be given the topology as a subspace of topological space H). It is of interest to ask why these two properties were discarded simultaneously and not one at a time (leading, e.g., to a definition of a "quasi-bounded operator"). It turns out that operators with $D_A = H$, but which are not continuous, seldom arise in applications, and, even when they do arise, one gains little from the fact that $D_A = H$. Little is lost by regarding such operators as just plain operators. The other possibility is $D_A \neq H$, but $D_A \xrightarrow{A} H$ continuous. This possibility warrants further consideration.

A subset of a topological space is said to be *dense* if its closure is the entire topological space (so, e.g., the set of rationals is dense in the real line). In particular, the domain D_A of an operator is dense in H if and only if, for every h in H and every positive ϵ, there exists a vector h' in D_A with $||h - h'|| \leq \epsilon$. It turns out that essentially every operator of interest has dense domain (and, in fact, we would have included this condition in the definition of an operator, except that it would make certain assertions awkward).

Example. The domain D_A of the last two examples is dense in $L^2(\mathbf{R})$. We must show that, given any complex-valued, measurable, square-integrable function f and any positive ϵ, there is a complex-valued, C^∞, function f' of compact support with $\int_{\mathbf{R}} |f - f'|^2 d\mu \leq \epsilon$. Sketch of proof: There is a step function, f_1, with $\int_{\mathbf{R}} |f - f_1|^2 d\mu \leq \epsilon$. Given any measurable set K, there is a measurable set K', the union of a finite collection of open intervals, with $\mu(K - K') \leq \epsilon$ and $\mu(K' - K) \leq \epsilon$. Hence there is a step function f_2, with each $f_2^1[r]$, for $r \neq 0$, at most the union of a finite collection of open intervals, and with $\int_{\mathbf{R}} |f_1 - f_2|^2 d\mu \leq \epsilon$ (figure 149). Now choose C^∞ function f', with compact support, such that $\int_{\mathbf{R}} |f_2 - f'|^2 d\mu \leq \epsilon$.

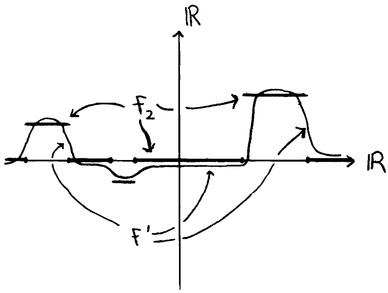

Figure 149

The fact that an operator has dense domain is often used in arguments, for example, via the following: if D_A is dense in H and if h is a vector in H with $(h,h') = 0$ for every h', in D_A, then $h = 0$. Proof: Fix vector $\underline{h}$ in H. Given any positive ϵ, choose h' in D_A with $||\underline{h} - h'|| \leq \epsilon$. Then $|(h,\underline{h})| = |(h,\underline{h}) - (h,h')| = |(h,\underline{h} - h')| \leq ||h||\,||\underline{h} - h'|| \leq \epsilon||h||$. Since this is true for every ϵ, $(h,\underline{h}) = 0$. Since this is true for every $\underline{h}$ in H, $h = 0$.

The "other possibility" referred to above can now be dealt with as follows. We claim: a continuous operator A with dense domain D_A has one and only one extension to a bounded operator. Proof: Since A is continuous, there is some number r such that $||A(h')|| \leq r||h'||$ for every h' in D_A. Let h be any vector in H, and let $h_1,h_2, \cdots$ be any sequence of vectors in D_A, converging to h (the existence of such a sequence is guaranteed by the fact that D_A is dense). Then $h_1,h_2, \cdots$ is a Cauchy sequence in H, whence, by $||A(h')|| \leq r||h'||$, $A(h_1)$, $A(h_2)$, $\cdots$ is a Cauchy sequence in H. Denote by $\tilde{A}(h)$ the vector to which this sequence converges (noting that this $\tilde{A}(h)$ is independent of the choice of $h_1,h_2, \cdots$), thus defining a mapping $H \xrightarrow{\tilde{A}} H$. This mapping is clearly linear, and hence is an operator which is an extension of A, D_A. Finally, $||A(h')|| \leq r||h'||$ for every h' in D_A implies $||\tilde{A}(h')|| \leq r||h'||$ for every h' in H, so $\tilde{A}$ is a bounded operator. Uniqueness is clear from the construction. Thus there is no point in considering continuous operators with dense domain: one might just as well consider their unique bounded extensions.

Since it is apparently operators (not necessarily bounded) that play an important role in applications, the natural next question is, What, of all the structure that was available on the set of bounded operators, is still available on the set of operators? The answer is, Very little indeed. First, note that, for A, D_A an operator and c a number, $D_B = D_A$ and action $B(h) = cA(h)$ for h in D_B defines an operator, which we may regard as cA. We next consider addition of operators A, D_A and B, D_B. We presumably wish to have action $(A + B)(h) = A(h) + B(h)$, so $A + B$ will only "know how to act" on h if both A and B do. Thus we must set $D_{A+B} = D_A \cap D_B$ (noting that this intersection of vector subspaces is a vector subspace) and define the action of $(A + B)$ on this domain as above. Note that the domain "continually gets smaller as one adds operators" (and, indeed, we could even have $D_A \cap D_B = \emptyset$ although both D_A and D_B are dense). In particular, the set of operators does not have the structure of a vector space (since, e.g., $A + (-A)$ has domain D_A rather than H). A similar problem with domains occurs for composition. For A, D_A and B, D_B operators, AB has action $(AB)(h) = A(B(h))$, but its domain consists of all h in D_B for which $B(h)$ is in D_A (indeed, a vector subspace). Finally, unbounded operators do not have norms or adjoints (at least, for the latter, not via theorem 57), for these notions make use of continuity. In particular, we have no obvious topology on the set of operators on Hilbert space H.

In short, we seem to be forced, by physical considerations, to consider a set (that of all operators) on which there is little useful structure and in terms of which we cannot even begin the program which, for bounded operators, led to the spectral theorem. One is apparently faced with the situation that what

one wants physically and what one can do mathematically do not meet on common ground.

Exercise 384. Let A, D_A, B, D_B, and C, D_C be operators. Does $(A+B)+C = A+(B+C)$? $A(BC)=(AB)C$? $A(B+C)=AB+AC$?

Exercise 385. Prove that every operator has an extension to an operator defined everywhere on H and that this extension is unique only if the original operator was defined everywhere.

Exercise 386. Is "is an extension of" a partial ordering on the set of operators on a Hilbert space?

Exercise 387. Find an example of an operator A, D_A, which is not continuous, but with $D_A = H$.

Exercise 388. Find explicitly the sum and composition of the two operators given in the second and third examples of this chapter.

Exercise 389. Consider the Hilbert space H of sequences $(c_1, c_2, \cdots)$ of complex numbers. Let D_A consist of sequences that are zero after some entry. For $(c_1, c_2, \ldots, c_n, 0, 0, \cdots)$ in D_A, let $A(c_1, \ldots, c_n, 0, \cdots) = (c_1, 2c_2, 3c_3, \ldots, nc_n, 0, \cdots)$. Prove that this is an operator. Is it densely defined? Suppose, instead, we had defined the action of A by $A(c_1, \ldots, c_n, 0, \cdots) = (c_1, \ldots, c_n, 1, 0, 0, \cdots)$. Is this an operator?

Exercise 390. Find an example of two operators, each densely defined, whose sum is not. Prove that, nonetheless, this sum has an extension that is an extension of each of the summands.

56

Self-Adjoint Operators

It turns out that there is a way to get around the difficulty described in the previous chapter: there exists a satisfactory mathematical treatment for a large class of operators, a class which includes many of physical interest. The key observation that makes such a treatment possible is the following: what one is actually interested in is not arbitrary (not necessarily bounded) operators, but rather certain ones which satisfy some condition analogous to Hermiticity in the bounded case. The idea is to choose very carefully this "condition analogous to Hermiticity."

Let H be a Hilbert space, and let D_A, A be an operator on H, with D_A dense in H. Denote by D_{A^*} the collection of all vectors h in H for which there exists a h' in H such that $(h,A(\underline{h})) = (h',\underline{h})$ for every $\underline{h}$ in D_A (i.e., for every $\underline{h}$ for which this formula makes sense). In other words: D_{A^*} consists of all vectors h in H for which the linear mapping $D_A \rightarrow \mathbf{C}$ which sends $\underline{h}$ in D_A to the complex number $(h,A(\underline{h}))$ is continuous (for continuity of this mapping is completely equivalent to the existence of a vector h' in H such that $(h,A(\underline{h})) = (h',\underline{h})$). Note next that this D_{A^*} is necessarily a vector subspace of H, for, for h and g in D_{A^*} (say, with $(h,A(\underline{h})) = (h',\underline{h})$ and $(g,A(\underline{h})) = (g',\underline{h})$ for every $\underline{h}$ in D_A), we have $(h + cg,A(\underline{h})) = (h' + cg',\underline{h})$ for every $\underline{h}$ in D_A, whence $h + cg$ is also in D_{A^*}.

We next show that this vector subspace D_{A^*} of H is the domain of a certain operator. First, note that, for h in D_{A^*}, so there exists an h' in H with $(h,A(\underline{h})) = (h',\underline{h})$ for every $\underline{h}$ in D_A, this h' is unique (for, were there two, their difference would be orthogonal to every $\underline{h}$ in D_A, whence, since D_A is dense, their difference would be zero). Let $D_{A^*} \xrightarrow{A^*} H$ be the mapping with the following action: for h in D_{A^*} (so there exists a unique h' in H with $(h,A(\underline{h})) = (h',\underline{h})$ for every $\underline{h}$ in D_A), set $A^*(h) = h'$. It is immediate that this mapping A^* is linear.

Thus, starting from an operator D_A, A on H, with D_A dense in H, we obtain an operator D_{A^*}, A^*. This D_{A^*}, A^* is called the *adjoint* of D_A, A. Thus: for h in D_{A^*} and $\underline{h}$ in D_A, we have $(h,A(\underline{h})) = (A^*(h),\underline{h})$. Note that this definition agrees with that of the adjoint for the case of a bounded operator.

Comparison of the discussion above with the proof of theorem 57 shows that all we have done is to repeat, for the unbounded case, the construction of the adjoint in the bounded case, taking appropriate care to get the domains right. In fact, we may summarize the above in a form analogous to that of theorem 57.

THEOREM 66. *Let D_A, A be an operator on Hilbert space H, with D_A dense in H. Then there exists one and only one operator D_{A^*}, A^* on H satisfying the following two conditions: i) for every h in D_{A^*} and $\underline{h}$ in D_A, $(h,A(\underline{h})) = (A^*(h),\underline{h})$, and ii) D_{A^*}, A^* is an extension of every operator satisfying condition i).*

One of the key facts about adjoints is this: if A, D_A is an extension of D_B, B (each with dense domain), then D_{B^*}, B^* is an extension of D_{A^*}, A^*. [Proof: Let h be in D_{A^*}, so, for some h', $(h,A(\underline{h})) = (h',\underline{h})$ for every $\underline{h}$ in D_A. Then, since A is an extension of B, we certainly have $(h,A(\underline{h})) = (h',\underline{h})$ for every $\underline{h}$ in D_B, and therefore $(h,B(\underline{h})) = (h',\underline{h})$ for every $\underline{h}$ in D_B. Thus h is also in D_{B^*} and, furthermore, $B^*(h) = A^*(h)$.] That is, as the domain of an operator gets larger, the domain of its adjoint gets smaller.

Let D_A, A be an operator, with dense domain, on Hilbert space H. This operator is said to be *self-adjoint* if it is equal to its adjoint, that is, if $D_{A^*} = D_A$, and $A = A^*$ on this common domain. Thus, for example, every Hermitian operator is self-adjoint (but, as we shall see shortly, the converse is false).

The concept of a self-adjoint operator is a central one in this subject: it represents a common ground between the mathematics and the physics of Hilbert spaces. In order to justify this assertion, we must deal with two issues. On the one hand, we must make a case that the standard operators that arise in physical applications are (or at least can be made to be) self-adjoint. On the other hand, we must show that mathematical tools are available for dealing with self-adjoint operators. Neither is obvious. One must in each case pay attention to the apparently irrelevant technical details (i.e., the question of domains) of the definition of self-adjoint. We now discuss briefly each of these two issues.

Example. Consider the second example of chapter 55 (the operator "multiplication by x on C^∞ functions of compact support"). First, note that, for f and $\underline{f}$ any two C^∞ functions of compact support, we have $\int_{\mathrm{R}} \bar{f}(x\underline{f})\,d\mu = \int_{\mathrm{R}} (\overline{xf})\underline{f}\,d\mu$. That is, for h in D_A, we have $(h,A(\underline{h})) = (A(h),\underline{h})$ for every $\underline{h}$ in D_A. We conclude that this h is therefore necessarily in D_{A^*} and, furthermore, $A^*(h) = A(h)$. In other words, we have shown that the adjoint of D_A, A is an extension of $D_A A$. Is this operator, or is it not, self-adjoint? To answer this question, we must decide whether there are any vectors in D_{A^*} which are not in D_A. In other words, we must decide whether h and h' in H and $(h,A(\underline{h})) = (h',\underline{h})$ for all $\underline{h}$ in D_A implies that h itself is in D_A. In still other words, we must decide whether f and f' measurable and square-integrable, and $\int_{\mathbf{R}} \bar{f}(x\underline{f})\,d\mu = \int_{\mathbf{R}} \overline{f'}\,\underline{f}\,d\mu$ for all $C^\infty \underline{f}$ of compact support, implies that f itself is C^∞ with compact support. Let, for example, f be the (not C^∞) function illustrated in figure 150, and let $f' = xf$. Then, clearly, both f and f' are measurable and square-integrable. Furthermore, we have $\int_{\mathbf{R}} \bar{f}(x\underline{f})\,d\mu = \int_{\mathbf{R}} \overline{f'}\,\underline{f}\,d\mu$ for every C^∞ function $\underline{f}$ of compact support. That is, we have h and h' in H, with h not in D_A but with $(h,A(\underline{h})) = (h',\underline{h})$ for every $\underline{h}$ in D_A. That is, we have a vector h in D_{A^*} with h not in D_A. That is, the adjoint of D_A, A is not only an extension of D_A, A, but has a strictly larger domain: $D_{A^*} \supset D_A$, and not equality. That is, this operator D_A, A is not self-adjoint.

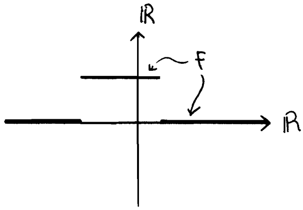

Figure 150

Our "physically interesting" operator D_A, A turns out not to be self-adjoint, for the domain of its adjoint is larger than D_A. What has gone wrong? Recall that increasing the size of the domain of an operator decreases the size of the domain of its adjoint. Thus, if we are ever going to obtain a self-adjoint operator out of all this, the best chance would be to increase the size of D_A, thereby decreasing the size of D_{A^*}, hoping to "get the two into line." It would be instructive to carry out this "enlargement of the size of the domain of A" in small steps, for example, by first dropping the C^∞ condition and then compactness of the support. We shall, however, jump directly to the "right answer."

Example. Let H be the Hilbert space $L^2(\mathbf{R})$. Let D_B consist of all vectors in H having, as representative, a function f with xf square-integrable. We note that D_B is a vector subspace of H (for, for xf and xf' square-integrable, so is $x(f+f') = xf + xf'$). Since D_B includes, in particular, the C^∞ functions of compact support, this D_B is, by chapter 52, dense in H. For h an element of D_B, with representative f, let $B(h)$ be the vector in H with representative xf (noting that, by definition of the domain D_B, this xf is indeed in H).

Clearly, $D_B \xrightarrow{B} H$ is linear. Thus we have an operator (called the *configuration operator* on $L^2(\mathbf{R})$). Is this operator self-adjoint? First, note that, for f and $\underline{f}$ measurable and square-integrable, with both xf and $x\underline{f}$ square-integrable, we have $\int_{\mathbf{R}} \bar{f}(x\underline{f})\,d\mu = \int_{\mathbf{R}} (\overline{xf})\underline{f}\,d\mu$. Thus, for both h and $\underline{h}$ in D_B, we have $(h, B(\underline{h})) = (B(h), \underline{h})$. That is, every vector h in D_B is also in D_{B^*}, and, for such an h, $B^*(h) = B(h)$. The adjoint of D_B, B is an extension of D_B, B. Are they equal? Let f and f' be measurable and square-integrable, and let $\int_{\mathbf{R}} \bar{f}(x\underline{f})\,d\mu = \int_{\mathbf{R}} \overline{f'}\underline{f}\,d\mu$ for every $\underline{f}$ representing an element of D_B. (That is, let f represent an element of D_{B^*}.) Then, clearly, we must have $f' = xf$ almost everywhere. Hence, since f' is square-integrable, xf must be square-integrable. That is, this f must represent an element of D_B. We have shown, in other words, that every element of D_{B^*} is also in D_B. Since the adjoint of D_B, B is an extension of this operator, we conclude that D_B, B is identical to its adjoint. The configuration operator is self-adjoint.

Note what has happened here. The configuration operator B of the example above is an extension of the operator A of the earlier example. By taking such an extension, we obtain a self-adjoint operator. What domain did the trick? One had to choose "the largest reasonable domain on which multiplication by x could be expected to act," that is, the most "natural-looking domain for 'multiplication by x.' " It is easy to understand why such should be the proper choice. Note that the only self-adjoint extension of the configuration operator is this operator itself, and the only self-adjoint operator of which the configuration operator is an extension is again the configuration operator. [Proof: The larger the domain of an operator, the smaller the domain of its adjoint.] Self-adjointness "knows how big the domain must be," and this proper domain is normally the "most natural one."

The pattern of this example is typical. We give a second example.

Example. Consider the third example of chapter 55 (the operator "i times the derivative on C^∞ functions of compact support"). First, note that, for f and $\underline{f}$ any two C^∞ functions of compact support, we have $\int_{\mathbf{R}} \bar{f}(i\,d\underline{f}/dx)\,d\mu = \int_{\mathbf{R}} (\overline{i\,df/dx})\underline{f}\,d\mu$ (integrating by parts and noting that, by compact supports, the surface terms vanish). Thus the adjoint of D_A, A is again an extension of D_A, A. To test for self-adjointness, let f be the function with action $f(x) = x$ for x in $[0,1)$, $f(x) = 2 - x$ for x in $[1,2]$, and $f(x) = 0$ otherwise. Let f' be the function with $f'(x) = i$ for x in $(0,1)$, $f'(x) = -i$ for x in $(1,2)$, and $f'(x) = 0$ otherwise. (That is, both f and f' are measurable and square-integrable, though neither is C^∞—or even differentiable. The function f' is "doing its best to be i times the derivative of f.") We now claim: for any $C^\infty \underline{f}$ of compact support, $\int_{\mathbf{R}} \bar{f}(i\,d\underline{f}/dx)\,d\mu = \int f'\underline{f}\,d\mu$ (a claim which is easily verified by splitting each integral into four, over $(-\infty,0)$, $(0,1)$, $(1,2)$, and $(2,\infty)$, respectively). Thus this f is a representative of an element h of D_{A^*}, although h is not in D_A (since f is certainly not C^{∞}). We conclude that this operator is not self-adjoint.

As usual, we want to correct our operator by enlargement of its domain. It is convenient to first introduce a few facts about functions. A complex-valued function f on R is said to be absolutely continuous if it is an indefinite integral, that is, if there is a measurable function g, integrable on closed intervals, such that f has the following action: $f(x) = c + \int_0^x g\, d\mu$ for $x \geq 0$, and $f(x) = c - \int_{-x}^0 g\, d\mu$ for $x < 0$, for some constant c. Note that this g, when one exists, is unique almost everywhere: we write $g = Df$. Thus, for example, every absolutely continuous function is continuous; for f C^1, Df is the derivative of f. Note, however, that the function f of the example above is also absolutely continuous and, in fact, iDf is the function f' of that example.

Example. Let H be the Hilbert space $L^2(\mathbf{R})$. Denote by D_B the collection of all elements of H having an absolutely continuous representative f with Df square-integrable. Then (since, for f and f' absolutely continuous, so is $f + cf'$, and $D(f + cf') = Df + cDf'$) D_B is a vector subspace of H. For h in D_B (with absolutely continuous representative f), let $B(h)$ be the element of H with representative iDf (noting that, by definition of D_B, iDf is square-integrable). Thus $D_B \xrightarrow{B} H$ is linear. This operator D_B, B is called the *momentum operator* on $L^2(\mathbf{R})$. We claim that this momentum operator is self-adjoint. Suppose first that f and $\underline{f}$ are representatives of elements of D_B. Then $\int_{\mathbf{R}} \bar{f}(iD\underline{f})\, d\mu - \int_{\mathbf{R}} (\overline{iDf})\underline{f}\, d\mu = i \int_{\mathbf{R}} D(\bar{f}\underline{f})\, d\mu$. For any positive r, we have $\int_0^r D(\bar{f}\underline{f})\, d\mu = \bar{f}(r)\underline{f}(r) - \bar{f}(0)\underline{f}(0)$. Since the limit, as r approaches infinity, of the left side exists, $\bar{f}(r)\underline{f}(r)$ must approach a constant, which, by square-integrability of f and $\underline{f}$, must be zero. Thus $\int_{\mathbf{R}} D(\bar{f}\underline{f})\, d\mu = \lim \int_{-r}^r D(\bar{f}\underline{f})\, d\mu = 0$. We conclude that every element h of D_B is also in D_{B^*}, and then $B^*(h) = B(h)$. That is, the adjoint of D_B, B is an extension of D_B, B. To prove that D_B, B is self-adjoint, we must show that every vector in D_{B^*} is also in D_B. Thus let f and f' be measurable and square-integrable, with $\int_{\mathbf{R}} \bar{f}(iD\underline{f})\, d\mu = \int_{\mathbf{R}} \overline{f'}\underline{f}\, d\mu$ for every $\underline{f}$ a representative of D_B. We must show that f is absolutely continuous, with $f' = iDf$. Given positive r and ϵ, let $\underline{f}$ be the C^{∞} function illustrated in figure 151. Then, since $\int_{\mathbf{R}} \bar{f}(iD\underline{f})\, d\mu = \int_{\mathbf{R}} \overline{f'}\underline{f}\, d\mu$ for every ϵ, we have, taking the limit as ϵ approaches zero, $i(\overline{f(r) - f(0)}) = \int_0^r \overline{f'}\, d\mu$. Since this is true for every r, f is indeed absolutely continuous, with $f' = iDf$. The momentum operator is self-adjoint.

These examples at least make the point that one can attack directly the problem of constructing a self-adjoint operator, given the physical idea of what the operator ought, roughly, to be. We emphasize, however, that this issue must apparently be faced on a case-by-case basis: one must find a suitable domain and actually check self-adjointness. For example, the domains above will certainly not do for an operator involving second derivatives. Note

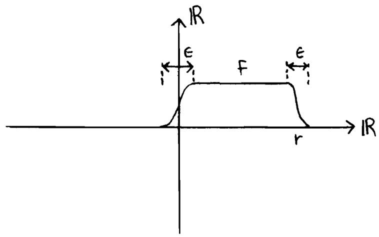

Figure 151

also that the sum of two self-adjoint operators is not in general self-adjoint—for example, the sum of the configuration and momentum operators above is not self-adjoint. It is mysterious why this sort of activity should have any relevance to physical problems.

This completes the first half of our program: the demonstration that the self-adjoint operators include many of physical interest. For the second half, we must argue that self-adjoint operators can be dealt with, in some reasonable way, mathematically. The basic "niceness feature" of self-adjoint operators—the feature which gets one started—is the following:

THEOREM 67. *Let D_A, A be a self-adjoint operator on Hilbert space H. Then the linear mapping $D_A \xrightarrow{A+iI} H$ is one-to-one and onto.*

Proof. For any h in D_A, we have $\|(A + iI)(h)\|^2 = (A(h),A(h)) + (A(h),ih) + (ih,A(h)) + (ih,ih) = (A(h),A(h)) + i(A(h),h) - i(h,A(h)) + (h,h) = (A(h),A(h)) + (h,h) \geq \|h\|^2$. We shall several times use this inequality.

By the inequality above, for any h in D_A with $(A + iI)(h) = 0$, we have $h = 0$. Hence the mapping $D_A \xrightarrow{A+iI} H$ is one-to-one. Denote by V the vector subspace $(A + iI)[D_A]$ of H. We shall show that this V is closed. Let $h_1,h_2, \cdots$ be a sequence of vectors in V (so there are vectors $\underline{h}_1,\underline{h}_2, \cdots$ in D_A with $h_1 = (A + iI)(\underline{h}_1)$, $h_2 = (A + iI)(\underline{h}_2)$, etc.) converging to some vector h. We must show that h is also in V. Note first that $h_1,h_2, \cdots$ is a Cauchy sequence. But, by our inequality, $\|h_n - h_m\|^2 = \|(A + iI)(\underline{h}_n - \underline{h}_m)\|^2 \geq \|\underline{h}_n - \underline{h}_m\|^2$, and so $\underline{h}_1,\underline{h}_2, \cdots$ is also a Cauchy sequence. Hence this sequence converges to some vector $\underline{h}$. We have, for any h' in D_A, $(\underline{h},A(h')) = \lim(\underline{h}_n,A(h')) = \lim(A(\underline{h}_n),h') = \lim(h_n - i\underline{h}_n,h') = (h - i\underline{h},h')$, where we have used self-adjointness in the second step and $A(\underline{h}_n) = h_n - i\underline{h}_n$ in the third. But this equation is precisely the statement that $\underline{h}$ is in D_{A^*}, with $A^*(\underline{h}) = h - i\underline{h}$. Since D_A, A is self-adjoint, $\underline{h}$ is also in D_A, with $A(\underline{h}) = h - i\underline{h}$. That is, $(A + iI)(\underline{h}) = h$. We have just shown that h is also in V. Thus V is indeed closed. We next wish to show that $V = H$ (i.e., that $(A + iI)$ is onto). It suffices, since V is closed, to prove that the only vector h in $V^{\perp}$ is $h = 0$. For h in $V^{\perp}$, we have, for every $\underline{h}$ in D_A, $(h,(A + iI)(\underline{h})) = 0$. That is, we have $(h,A(\underline{h})) = (ih,\underline{h})$ for every $\underline{h}$ in D_A. But this is precisely the statement that h is in D_{A^*}, with $A^*(h) = ih$. By self-adjointness, h is in D_A, with $(A - iI)(h) = 0$. Therefore, $0 = (h,(A - iI)(h)) = (h,A(h)) - i(h,h)$. Since the first term on the right is real, and the second imaginary, $(h,h) = 0$, that is, $h = 0$. Thus $V = H$, completing the proof. ☐

The proof will be recognized as a more careful version—keeping track of domains—of the earlier proof that the spectrum of a Hermitian operator is real. Note the essential use made of self-adjointness.

Example. Consider the configuration operator on $L^2(\mathbf{R})$. Theorem 67, in this case, reduces to the following statement. Given any measurable, square-integrable function f, there is one and only one measurable, square-integrable function g, with xg square-integrable and with $(x + i)g = f$. It is easy to convince oneself directly that this statement is true.

Example. Consider the momentum operator on $L^2(\mathbf{R})$. Theorem 67, in this case, reduces to the following statement. Given any measurable, square-integrable function f, there is one and only one absolutely continuous function g, with Dg square-integrable and with $(iDg + ig) = f$ almost everywhere. This is a—by no means obvious—assertion about existence and uniqueness of solutions of a certain differential equation.

Theorem 67 states that, for D_A, A self-adjoint, D_A is a "vector-space copy" of H, with A the isomorphism of vector spaces. Self-adjointness is a rather strong and "tight" condition. It is immediate from theorem 67, incidentally, that no alteration of the domain of a self-adjoint operator can preserve self-adjointness (for such an alteration could hardly preserve the one-to-one, onto character of $(A + iI)$). Of course, theorem 67 is equally true for $(A - iI)$.

Now let D_A, A be a self-adjoint operator on Hilbert space H. Set $U = (A - iI)(A + iI)^{-1}$. Then, since $(A + iI)^{-1}$ is a one-to-one, onto mapping from H to D_A while $(A - iI)$ is a one-to-one, onto mapping from D_A to H, U is a (clearly linear) one-to-one, onto mapping from H to H. In short, U is an operator with domain H. This observation suggests that we try to show that U is in fact a bounded operator. First, note that, for $\underline{h}$ in D_A, $||(A + iI)(\underline{h})||^2 = ||A(\underline{h})||^2 + ||\underline{h}||^2 = ||(A - iI)(\underline{h})||^2$. Now let h be any vector in H, and set $h = (A + iI)(\underline{h})$ with $\underline{h}$ in D_A (always possible, by theorem 67). Then $U(h) = (A - iI)(A + iI)^{-1}(h) = (A - iI)(A + iI)^{-1}(A + iI)(\underline{h}) = (A - iI)(\underline{h})$. Hence $||U(h)||^2 = ||(A - iI)(\underline{h})||^2 = ||(A + iI)(\underline{h})||^2 = ||h||^2$. Thus U is in fact a bounded operator: its norm is $|U| = 1$. Furthermore, this U is actually a unitary operator. Replacing h in $||U(h)||^2 = ||h||^2$ successively by $h + ih'$ and $h - ih'$, we obtain $(U(h),U(h')) = (h,h')$ for all h and h'. That is, $(U^*U(h),h')$

$= (h,h')$ for all h and h'. Since h' is arbitrary, $U^*U = I$. Since U is one-to-one and onto, this also implies $UU^* = I$.

Thus every self-adjoint operator defines, by the construction above, a unitary operator. Let us look a bit at the mechanism of this construction.

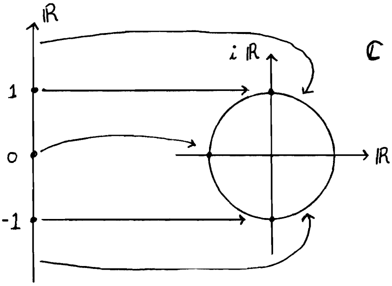

Figure 152

Consider the mapping $\mathbf{R} \xrightarrow{\alpha} \mathbf{C}$ with action $\alpha(r) = (r - i)(r + i)^{-1}$ (figure 152). Then, for each real r, $|\alpha(r)| = 1$, so α maps the set of reals to the unit circle in the complex plane. The real line is "wrapped once around the unit circle in the complex plane," with $+1$ sent to $-i$, -1 sent to $+i$, etc. "If $+\infty$ and $-\infty$ were real numbers, these would both be sent to one. Now, the construction above amounts essentially to setting "$U = \alpha(A)$." (This intuitive picture, however, ignores some important technical details. For example, had we set $U = (A + iI)^{-1}(A - iI)$ rather than $U = (A - iI)(A + iI)^{-1}$, then U would be a one-to-one, onto mapping from D_A to D_A rather than a one-to-one, onto mapping from H to H.) One thinks of a self-adjoint operator A as "essentially a Hermitian operator, except that, for certain vectors h (namely, those not in D_A), we have $A(h) = \infty h$." Behavior of this sort is unacceptable for a bounded operator, which is why the domain of A ends up not being the entire Hilbert space. Fortunately, the function α "knows exactly what to do with ∞, namely, send it to the finite number one." Thus "U knows how to act even on those vectors on which A itself refuses to act." In this way U becomes a bounded operator. One also sees in this picture why U ends up as a unitary operator: the function α sends the real line to the unit circle in the complex plane.

What fact is analogous to the fact that $\alpha(r) = 1$ for no r? It is this: the number one is not an eigenvalue of U. To prove this last assertion, suppose there were a nonzero vector h with $U(h) = h$. By theorem 67, there is a $\underline{h}$ in D_A with $(A + iI)(\underline{h}) = h$. Then $U(h) = (A - iI)(A + iI)^{-1}(h) = (A - iI)(A + iI)^{-1}(A + iI)(\underline{h}) = (A - iI)(\underline{h})$. Thus $U(h) = h$ becomes $(A - iI)(\underline{h}) = (A + iI)(\underline{h})$, which implies immediately that $\underline{h} = 0$ and hence that $0 = (A + iI)(\underline{h}) = h$. Thus this h cannot be an eigenvector.

To summarize, we have shown so far that every self-adjoint operator leads, via $U = (A - iI)(A + iI)^{-1}$, to a unitary operator not having one as an eigenvalue. There is, in fact, a converse to this statement. In order to guess what it should be, one solves $c = (r - i)/(r + i)$ for r in terms of c, obtaining $r = i(1 + c)/(1 - c)$. This suggests

THEOREM 68. *Let U be a unitary operator on Hilbert space H, and suppose that one is not an eigenvalue of U. Set $D_A = (I - U)[H]$. Then $A = i(I + U)(I - U)^{-1}$ exists as a mapping from D_A to H, and D_A, A is a self-adjoint operator.*

Proof. First, note that D_A, as the image of a vector space under a linear mapping, is a vector subspace of H. Next, note that, since one is not an eigenvalue of U, $H \xrightarrow{I-U} D_A$ is one-to-one and onto, whence $D_A \xrightarrow{(I-U)^{-1}} H$ exists. Hence we indeed have $D_A \xrightarrow{A} H$, a well-defined linear mapping. That is, A, D_A is certainly an operator.

To prove that D_A is dense in H, it suffices to show that the only vector h orthogonal to every vector in D_A is the zero vector. Let h be such a vector, so $(h,(I - U)(h')) = 0$ for every h' in H. Then $((I - U^*)(h),h') = 0$ for every h', whence $(I - U^*)(h) = 0$. That is, $U^*(h) = h$, or, applying U to this equation, $h = U(h)$. Since one is not an eigenvalue of U, it follows that $h = 0$. Thus D_A is dense in H. We next show that the adjoint of D_A, A is an extension of D_A, A. We must show that, for h and h' in D_A, $(h,A(h')) = (A(h),h')$. Set $h = (I - U)(g)$ and $h' = (I - U)(g')$, so $A(h) = i(I + U)(g)$ and $A(h') = i(I + U)(g')$. Then

$$
\begin{aligned}
(h,A(h')) &= ((I - U)(g),i(I + U)(g')) \\
&= ((I + U^*)(I - U)(g),ig') \\
&= ((I + U^{-1})(I - U)(g),ig') \\
&= ((U + I)U^{-1}(I - U)(g),ig') \\
&= ((I + U)(U^{-1} - I)(g),ig') \\
&= ((I + U)(U^* - I)(g),ig') \\
&= ((U^* - I)(I + U)(g),ig') \\
&= ((I + U)(g),i(U - I)(g') \\
&= (i(I + U)(g),(I - U)(g')) \\
&= (A(h),h') .
\end{aligned}
$$

The proof is completed by showing that every vector in D_{A*} is also in D_A.

Thus let h and h' be vectors, with $(h, A(\underline{h})) = (h', \underline{h})$ for every $\underline{h}$ in D_A. We must show that h itself is in D_A. Set $\underline{h} = (I - U)(g)$, with g in H, so $A(\underline{h}) = i(I + U)(g)$. Then we have $(h, i(I + U)(g)) = (h', (I - U)(g))$ for every g in H, that is, $(-i(I + U^*)(h), g) = ((I - U^*)(h'), g)$ for every g in H. Since g is arbitrary, $-i(I + U^*)(h) = (I - U^*)(h')$, or, applying U, $i(I + U)(h) = (I - U)(h')$. But this equation can be rewritten $h = (I - U)(-ih'/2 + h/2)$, which shows that h is indeed in D_A. □

This straightforward, computational proof involves no new ideas.

Thus, not only does every self-adjoint operator define a unitary operator for which one is not an eigenvalue, but every such unitary operator defines a self-adjoint operator. Furthermore, these two constructions invert each other, that is, if we begin with self-adjoint A, D_A and then construct the corresponding unitary U and then the corresponding self-adjoint operator, the result is just the original self-adjoint operator. [Proof: Let D_A, A be self-adjoint, and set $U = (A - iI)(A + iI)^{-1}$. Then $I - U = I - (A - iI)(A + iI)^{-1} = ((A + iI) - (A - iI))(A + iI)^{-1} = 2i(A + iI)^{-1}$. By theorem 67, the domain of the self-adjoint operator $\tilde{A}$ constructed from U (via theorem 68) is precisely D_A. Since, furthermore, $I + U = I + (A - iI)(A + iI)^{-1} = ((A + iI) + (A - iI))(A + iI)^{-1} = 2A(A + iI)^{-1}$, we have $\tilde{A} = i(I + U)(I - U)^{-1} = 2iA(A + iI)^{-1}(1/2i)(A + iI) = A$.] Thus we have obtained a one-to-one correspondence between the set of all self-adjoint operators on Hilbert space H and the set of all unitary operators on H not having one as an eigenvalue. But unitary operators are bounded: we know how to deal with them. The idea, then, is to describe various properties of our self-adjoint operators in terms of the corresponding properties of the unitary operator. Some examples of this approach follow.

How can one tell, from U, whether A is actually Hermitian (i.e., bounded)? We claim that A is Hermitian if and only if one is a regular value of U. First, note that, if one is a regular value of U, so $(I - U)^{-1}$ is bounded, then $A = i(I + U)(I - U)^{-1}$ is certainly bounded. If, conversely, A is bounded, then, since $(I - U)^{-1} = (1/2i)(A + iI)$, one must be a regular value of U. Several other facts follow from the (easily checked) formula $(A - \kappa I) = (\kappa + i)[U - (\kappa - i)/(\kappa + i)I](I - U)^{-1}$, applied to vectors in D_A. Suppose that h in D_A is an eigenvector of A with eigenvalue κ, that is, that $A(h) = \kappa h$. Then, from the formula above, $[U - (\kappa - i)/(\kappa + i)](I - U)^{-1}(h) = 0$. Now, $(I - U)^{-1}$ certainly cannot annihilate h (for this is the inverse of some mapping), so $[U - (\kappa - i)/(\kappa + i)I]$ must annihilate the nonzero vector $(I - U)^{-1}(h)$. In other words, $(\kappa - i)/(\kappa + i)$ must be an eigenvalue of U. Conversely, if h' is an eigenvalue of U with eigenvalue $(\kappa - i)/(\kappa + i)$, then $(I - U)(h')$ is in D_A and is an eigenvector of A with eigenvalue κ. We have shown: κ is an eigenvalue of A if and only if $(\kappa - i)/(\kappa + i)$ is an eigenvalue of U. The function α takes eigenvalues to eigenvalues.

A similar remark applies to the spectra. A number κ will be said to be in the *spectrum* of self-adjoint operator D_A, A if there exists no bounded operator B with $B(A - \kappa I)$ the identity (on D_A) and $(A - \kappa I)B$ the identity (on H). Then (again by the formula above), κ is in the spectrum of A if and only if $(\kappa - i)/(\kappa + i)$ is in the spectrum of U.

We have in some sense now completed the second half of our program. We have made available to ourselves mathematical tools which should allow us to do reasonable things with, and make interesting statements about, self-adjoint operators. We claim that the self-adjoint operators are "mathematically tame," at least far more so than general (not necessarily bounded) operators.

It is natural to ask how "mathematically tame" the self-adjoint operators really are. There is, for example, a spectral theorem for Hermitian operators; is there such a theorem for self-adjoint operators? In fact, the answer is yes. As a final example of the present technique, we discuss how this theorem goes. One first defines continuous functions, and then "other functions," of a unitary operator. This is done along the same general lines as for Hermitian operators, except that one constantly uses $U^* = U^{-1}$ (rather than $A^* = A$, in the Hermitian case). Next, for each point ω on the unit circle in the complex plane, one considers the function θ_ω shown in figure 153. Set $\tilde{P}_\omega = \theta_\omega(U)$.

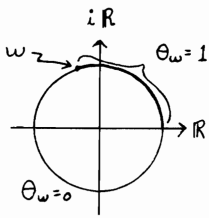

Figure 153

This family of projection operators is analogous to the spectral family for a Hermitian operator. One next shows $U = \int_{\text{circle}} \omega \, d\tilde{P}_\omega$, that is, one obtains a spectral theorem for unitary operators. Set $P_\kappa = \tilde{P}_{(\kappa-i)/(\kappa+i)}$ for each real κ. Then, using the relationship between self-adjoint A and unitary U, one shows that A can be expressed as $A = \int_{\mathbf{R}} \kappa \, dP_\kappa$. Essentially the only change in the spectral theorem in the passage from the Hermitian to the self-adjoint case is that the range of integration (the spectrum of A) is compact in the former and not in the latter.

It would perhaps be of interest to understand the following issues better: i) why does this notion of self-adjoint-ness play such a central role?[1] and ii)

1. Operators correspond to observables in quantum mechanics, and the spectrum of an operator corresponds to, in some sense, "the values that can be obtained on measuring

why does one describe the structure of a (say, physically interesting) self-adjoint operator in terms of a unitary operator which, apparently, has no physical significance?

Exercise 391. Consider the sum of the configuration and momentum operators on $L^2(\mathbf{R})$. Show that the adjoint of this sum is an extension of the sum but that the sum itself is not self-adjoint. Find a self-adjoint extension.

Exercise 392. Find explicitly the adjoints of the operators in the first and third examples of this chapter.

Exercise 393. An operator D_A, A is said to be closed if it satisfies the following condition: whenever $h_1, h_2, \cdots$ in D_A converges to h, and $A(h_1), A(h_2), \cdots$ converges to h', h is itself in D_A and $A(h) = h'$. Prove that every self-adjoint operator is closed.

Exercise 394. Consider the Hilbert space $L^2(\mathbf{R} \times \mathbf{R} \times \mathbf{R})$ (products of measure spaces). The momentum operator in this Hilbert space should, roughly speaking, be "i times the gradient." Is there some suitable generalization of absolutely continuous to this case? Does there exist, on this Hilbert space, a self-adjoint momentum operator?

Exercise 395. Can two self-adjoint operators have only the zero vector in the intersection of their domains?

Exercise 396. Does the sum of two self-adjoint operators always have a self-adjoint extension? Is the adjoint of such a sum always an extension of the sum?

Exercise 397. Find explicitly the action of the unitary operators associated with the configuration and momentum operators.

Exercise 398. Let H be the Hilbert space $L^2[0,1]$ (where $[0,1]$ is that subspace of the measure space of reals). Let D_A consist of all elements of H having representatives f with i) f absolutely continuous, ii) Df square-integrable, and iii) $f(0) = f(1) = 0$. Let the action of A be $A(f) = iDf$. Prove that D_{A^*}, A^* is an extension of D_A, A and that D_{A^*} consists precisely of elements of H having representatives satisfying i) and ii) above. (Thus D_A, A is not self-adjoint.) Prove that the adjoint of D_{A^*}, A^* is D_A, A (so D_{A^*}, A^* is not self-adjoint either). Is it true that there is no self-adjoint extension of D_A, A? What does this mean physically?

---

the observable." Thus the physical role of Hermiticity in the bounded case is to give a real spectrum. But why, physically, is there all this concern about domains?

Exercise 399. Is it true in general that the adjoint of the adjoint of an operator with dense domain is that operator?

Exercise 400. Consider the Hilbert space $L^2(\mathbf{R})$. The Hamiltonian on this Hilbert space is to be "$-d^2/dx^2 + V$," where V (the potential) is a real function on $\mathbf{R}$. Under what conditions on V can one construct a self-adjoint operator along these lines?

Exercise 401. Let U be a unitary operator, and let ω be a complex number with $|\omega| = 1$ and with ω a regular value of U. Then one is a regular value of the unitary operator $\omega^{-1}U$. Use the spectral theorem for the Hermitian operator associated with $\omega^{-1}U$ to obtain directly a spectral theorem for U.

Exercise 402. Let D_A, A be self-adjoint, with real κ not in its spectrum. Then $(\kappa - i)/(\kappa + i)$ is a regular value of the corresponding U. Use exercise 401 to obtain directly a spectral theorem for D_A, A.

Exercise 403. What are the spectra of the configuration and momentum operators?

Exercise 404. Why does one not consider "not necessarily bounded unitary or projection operators"?

Index of Defined Terms