Appendix A : Topological groups

(Munkres, Supplementary exercises following $ 22)

As the name suggests, a topological group is a mathematical structure that is both a topological space and a group, with some sort of compatibility between the topological and algebraic structures. Such objects lie at the point where two different areas of pure mathematics meet.

One motivation for looking at such combination structures is pure intellectual curiosity. Scientists are always interested in learning what happens if you combine **A** with **B**. Sometimes the results are not particularly useful (or do not seem so at the time), but very often these combinations can lead to important new insights into our knowledge of the original structures and to powerful new methods for analyzing questions that had previously been relatively difficult to study.

It turns out that topological groups form a rich family of interesting, relatively accessible and fundamentally important topological spaces. One reason for this is that the group structure turns out to impose some severe restrictions on the topology of the underlying space. In the other direction, the structure theory of an important special case of topological groups — namely, the compact connected Lie (pronounced "lee") groups — foreshadowed the classification of finite simple groups that was completed during the second half of the twentieth century (*Note:* A finite group is said to be **SIMPLE** if it is nonabelian and the only normal subgroups are the trivial subgroup and the group itself; the alternating groups on $n \geq 5$ letters are the most basic examples, there is a very large class of such groups that are related to Lie groups, and there is a list of 26 other groups that are called sporadic).

Another reason for the importance of topological groups is that such structures arise in a wide range of mathematical contexts. In particular, many of the most important objects studied in mathematical analysis come from topological groups (usually with some additional structure). Topological groups also arise play crucial roles in many areas of geometry, topology and algebra, and they are quite useful in application of mathematics to the sciences as well, and physics is an especially prominent example.).

Having discussed the role of topological groups, the next step is to describe them formally.

Definition. A *topological group* is a quadruple

$$(G, \mathbf{T}, m, \mathbf{inv})$$

consisting of a nonempty topological space $(G, \mathbf{T})$ and a group $(G, m, \mathbf{inv})$ such that the multiplication map $m : G \times G \to G$ and the inverse map $\mathbf{inv} : G \to G$ are continuous.

EXAMPLES. 1. The real numbers $\mathbf{R}$ form a topological group with respect to the addition operation, and the nonzero real numbers form a topological group with respect to multiplication. Similar statements hold for the complex numbers $\mathbf{C}$. In fact, one can the real and complex numbers with the addition and multiplication operations as examples of a suitably defined *topological field.*

2. The unit circle S^1 is a topological group because it is a subgroup of the multiplicative group of nonzero complex numbers.

3. If $(G, \cdots)$ and $(H, \cdots)$ are topological groups then one can make their product $G \times H$ into a topological space and a group, and it turns out that the product topology and the product group operations define a topological group structure on $G \times H$. If we specialize this to the case

$G = H = S^1$ we obtain the group structure of the 2-*dimensional torus* T^2. By induction the k-fold product of S^1 with itself, which is equivalent to $T^{k-1} \times S^1$ is also a topological group which is called the k-torus and denoted by T^k.

4. Let $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$, and let $\mathbf{GL}(n, \mathbf{F})$ be the group of invertible $n \times n$ matrices over $\mathbf{F}$. Suppose first that $\mathbf{F} = \mathbf{R}$. Then $\mathbf{GL}(n, \mathbf{R})$ is an open subset of the set of all $n \times n$ matrices (with the topology inherited from $\mathbf{R}^{n^2}$, and the coordinates of the multiplication and inverse maps are respectively polynomial and rational functions of the entries of the matrices (multiplication) or matrix (inversion). Therefore the conditions for a topological group are satisfied. Suppose now that $\mathbf{F} = \mathbf{C}$. Then $\mathbf{GL}(n, \mathbf{C})$ consists of all complex matrices whose determinants are nonzero. The natural topology for the space of $n \times n$ matrices over $\mathbf{C}$ is given by identifying the matrices with points of $\mathbf{R}^{2n^2}$, and from this viewpoint the real and imaginary parts of the complex determinant are polynomial functions of the real and imaginary parts of the matrix entries. Therefore we see that $\mathbf{GL}(n, \mathbf{C})$ corresponds to an open subset of $\mathbf{R}^{2n^2}$, and as in the real case it follows that multiplication and inversion are polynomial and rational functions so that $\mathbf{GL}(n, \mathbf{C})$ is also a topological group.

5. The subgroups of *orthogonal* and *unitary* matrices are important compact subgroups of $\mathbf{GL}(n, \mathbf{R})$ and $\mathbf{GL}(n, \mathbf{C})$ respectively. The verification that they are subgroups is carried out in linear algebra courses, Why are they compact? Recall that orthogonal and unitary matrices are characterized by the fact that their rows (equivalently, columns) form an orthonormal set. Since the rows (or columns) are all unit vectors, it follows that the groups O_n and U_n of orthogonal and unitary matrices are subsets of the compact set

$$\prod{}^n S^{\alpha n - 1} \quad \subset \quad \mathbf{R}^{\alpha n^2}$$

where $\alpha = 1$ for $\mathbf{R}$ and $\alpha = 2$ for $\mathbf{C}$. Since the orthonormality conditions reduce to equations involving certain polynomials in the real and imaginary parts of the matrix entries, it follows that the orthogonal and unitary groups are closed subsets of the products of spheres displayed above, and therefore these groups are compact.

6. In Exercises I.1.1 and VI.2.4 we considered a metric $\mathbf{d}_p$ on the integers for each prime p, and we noted that the completion $\widehat{\mathbf{Z}_p}$ was a compact metric space. In fact, one can extend the usual addition and multiplication maps on $\mathbf{Z}$ to continuous maps

$$\widehat{\mu} : \widehat{\mathbf{Z}_p} \times \widehat{\mathbf{Z}_p} \to \widehat{\mathbf{Z}_p}$$

$$\widehat{\alpha} : \widehat{\mathbf{Z}_p} \times \widehat{\mathbf{Z}_p} \to \widehat{\mathbf{Z}_p}$$

that make the space in question into a *compact topological commutative ring with unit.* The existence of these extensions follows directly from the results on extending uniformly continuous functions on metric spaces to the completions of the latter and the uniform continuity of ordinary addition and multiplication on $\mathbf{Z}$ with respect to the metrics $\mathbf{d}_p$ (verify this!). The systems obtained in this manner are known as the **p-adic integers**.

Especially in topology, whenever some type of mathematical structure is defined, one should also define the mappings or morphisms from one such object to another. For topological groups, the notion of *continuous homomorphism* is an obvious choice. Formally, these are continuous functions $\varphi : G \to H$ such that $\varphi(ab) = \varphi(a) \cdot \varphi(b)$ or alternatively satisfy the morphism identity

$$m_H \circ (\varphi \times \varphi) = \varphi \circ m_G \ : G \times G \to H$$

where m_G and m_H are the multiplication maps for G and H respectively.

One particularly important continuous homomorphism is the *exponential map* from the additive topological group of real or complex numbers to the multiplicative topological group of nonzero real or complex numbers. Another example on $\mathbf{GL}(n, \mathbf{F})$ is the map sending an invertible matrix to its transposed inverse. Over the complex numbers one also has the conjugation map on both the additive and nonzero multiplicative groups of complex numbers and the conjugate of the transposed inverse for invertible complex matrices.

All the nonexponential examples in the preceding paragraph are in fact *topological automorphisms* of the groups in question; *i.e.*, they are continuous homomorphisms that have continuous inverses. In fact, for all these examples the map is equal to its own inverse.

Properties of topological groups

We had previously stated that the group structure on a topological group (and especially its continuity!) implies strong restrictions on the topology of the underlying space. We shall discuss the most basic properties here.

HOMOGENEITY OF TOPOLOGICAL GROUPS

A topological space X is said to be *homogeneous* if for each pair of points $u\,v \in X$ there is a homeomorphism $h : X \to X$ such that $h(u) = v$. It is necessary to be somewhat careful when using this term, because the expression "homogeneous space" has a special meaning that is described below. There are many examples of spaces that are homogeneous, and there are many spaces that do not satisfy this condition.

EXAMPLES. 1. Every normed vector space V is homogeneous. Given two vectors $u, v \in V$ the translation map $T(x) = x + (v - u)$ is an isometry and sends u to v.

2. The closed unit interval $[0, 1]$ is not homogeneous; more precisely, it is impossible to construct a homeomorphism taking an end point to a point in the open interval $(0, 1)$. If such a homeomorphism h existed then for all open sets U containing 1 there would be homeomorphisms between $U - \{1\}$ and $h(U) - \{h(1)\}$. On the other hand, sets of the latter type are disconnected if $h(1) \in (0, 1)$ while $U - \{1\}$ is the connected set $[0, 1)$ if $U = [0, 1]$.

3. Since the open interval $(0, 1)$ is homeomorphic to $\mathbf{R}$, it follows from the first example that $(0, 1)$ is homogeneous. In fact, *if U is an open connected subset of $\mathbf{R}^n$, then U is homogeneous.* The proof is relatively elementary, but since the argument is a bit lengthy it will be given separately in Appendix C.

The homogeneity of topological groups is essentially a generalization of the first example given above:

PROPOSITION. *If G is a topological group, let $L_a : G \to G$ be the continuous map $L_a(g) = a \cdot g$ ("left multiplication"), and let $R_a : G \to G$ be the continuous map $R_a(g) = g \cdot a$ ("left multiplication"). Then L_a and R_a are homeomorphisms.*

Sketch of proof. Let $b = a^{-1}$. Then the associativity and continuity of multiplication imply that L_b and R_b are inverses to L_a and R_a respectively.■

COROLLARY. *A topological group is homogeneous.*

Proof. Given distinct points $a\,, b \in G$ let $c = b\,a^{-1}$ and note that $L_c(a) = b$.■

SEPARATION PROPERTIES OF TOPOLOGICAL GROUPS

All of the examples of topological groups that we have explicitly described are Hausdorff. It turns out that weaker separation properties imply the Hausdorff Separation Property, and the Hausdorff Separation Property itself implies even stronger separation properties. We shall begin by proving the first of these statements.

PROPOSITION. *If a topological group is a $\mathbf{T_0}$ space, then it is a Hausdorff (or $\mathbf{T_2}$) space.*

Proof. The proof splits into two parts, one showing the implication $\mathbf{T_0} \implies \mathbf{T_1}$ and the other showing the implication $\mathbf{T_1} \implies \mathbf{T_2}$.

($\mathbf{T_0} \implies \mathbf{T_1}$). It suffices to prove that the one point set $\{1\}$ consisting only of the identity element is closed in G; for every other element g, one can apply the homeomorphism L_g to see that $\{g\} = L_g(\{1\})$ is also closed in G. Therefore we need only show that $G - \{1\}$ is open. If x belongs to the latter, then the $\mathbf{T_0}$ condition implies that there is n open subset W that contains either 1 or x but not both. If $x \in W_x$ let $U_x = W_x$. Suppose now that $1 \in W_x$ but $x \notin W_x$; if we can find a homeomorphism $h : G \to G$ such that h switches 1 and x, then $h(W_x)$ is an open set containing x but not 1, and if we set $U_x = h(W_x)$ this will express $G - \{1\}$ as a union of open sets and thus prove that $G - \{1\}$ is open. The desired homeomorphism is merely the composite $L_x \circ \mathbf{inv}$.

($\mathbf{T_1} \implies \mathbf{T_2}$). It will suffice to show that the diagonal is closed in $G \times G$. Let $D : G \times G \to G$ be the composite $m \circ (\mathbf{id}_G \times \mathbf{inv})$. As noted in Exercise 1 on page 145 of Munkres, this map is continuous. But $\Delta_G = D^{-1}(\{1\})$ and since $\{1\}$ is closed in G by assumption, it follows that Δ_G is closed in $G \times G$ so that G is Hausdorff, which is the same as saying that G is $\mathbf{T_2}$.■

If a topological group is Hausdorff, then one can prove that it satisfies even stronger separation properties. In particular, Exercise 7 on page 146 of Munkres shows that a Hausdorff (equivalently, a $\mathbf{T_0}$) topological group is $\mathbf{T_3}$, and Exercise 10 on pages 213–214 of Munkres shows that such a topological group is also $\mathbf{T_{3\frac{1}{2}}}$.

CLOSED SUBGROUPS OF TOPOLOGICAL GROUPS

In the theory of topological groups it is important to know whether or not a subgroup is closed, and if a subgroup is not closed it is important to have the following additional information about its closure:

PROPOSITION. *If G is topological group and H is a subgroup, then its closure $\overline{H}$ is also a subgroup. Furthermore, if H is normal then so is $\overline{H}$.*

Before proceeding further, here are some examples when $G = \mathbf{R}$ (with addition as the group operation). The rationals are clearly not a closed subgroup (since they are dense in the real line), but the integers are, and one way of seeing this is to note that the complement of integers is equal to the union of the open intervals $(n, n+1)$ where n ranges over all the elements of $\mathbf{Z}$. Over the complex numbers, the additive groups of real an purely imaginary numbers are closed subgroups (why?) and in the groups $\mathbf{GL}(n, \mathbf{F})$ the orthogonal and unitary groups are closed subgroups because the latter are compact.

Proof of proposition. If H is a subgroup then we know that $m(H \times H) = H$ and $\mathbf{inv}(H) = H$. Recall that a function f is continuous if and only if it satisfies the condition

$$f(\overline{A}) \quad \subset \quad \overline{f(A)} \ .$$

If we apply this to multiplication and the inverse map we find that

$$m\left(\overline{H}\times\overline{H}\right) \;=\; m\left(\overline{H\times H}\right) \;\subset\; \overline{m(H\times H)} \;=\; \overline{H}$$

and similarly

$$\mathbf{inv}\left(\overline{H}\right) \;\subset\; \overline{H}$$

so that $\overline{H}$ is a subgroup of G.

Suppose now that H is normal in G; in other words, for all $a\in G$ the inner automorphism $I_a(x)=axa^{-1}$ maps H to itself. This mapping is continuous because it is the composite $L_a\circ R_{a^{-1}}$; by associativity it can also be written as $R_{a^{-1}}\circ L_a$. As in the previous paragraph the continuity of this map and the hypothesis $I_a(H)\subset H$ imply that $I_a\left(\overline{H}\right)\subset\overline{H}$.■

In group theory, if one has a group G and a subgroup H, then it is possible to form the set G/H of cosets, and it H is a normal subgroup then one can construct a quotient group structure on G/H. For topological groups one can carry out the same constructions and topologize G/H using the quotient topology; this quotient is $\mathbf{T_1}$ if and only if H is a closed subgroup. Some additional information on this can be found in Exercises 5 and 6 on page 146 of Munkres (see also Exercise 7(d) on that page). We shall also note that the standard isomorphism theorems known from ordinary group theory have topologized analogs for topological groups.

Example. Although closed subgroups play an extremely important role in the theory of topological groups, the non-closed subgroups are too important to be ignored, so we shall give one more example involving T^2: Given an irrational real number $a>0$, consider the continuous homomorphism

$$\varphi:\mathbf{R}\longrightarrow T^2$$

defined by

$$\varphi(t) \;=\; (\,\exp(2\pi it),\ \exp(2\pi iat)\,)$$

and let H be the image of φ. To see that H is not closed, consider its intersection with $\{1\}\times S^1$, and call this intersection L. By definition it follows that L consists of all elements of the form $(1,\exp(2\pi ian)$ for some $n\in\mathbf{Z}$. If H is a closed subgroup of T^2 then so is L. By the previous description L is countable; we claim it is also infinite; if not, then there would be two integers $p\neq q$ such that $\exp(2\pi iap)=\exp(2\pi iaq)$, and the latter would imply that $a(p-q)$ is an integer, contradicting the irrationality of a. Since T^2 is a compact metric space and L is a closed subset, it follows that L is complete with respect to the subspace metric. Therefore a corollary of Baire's Theorem implies that L has an isolated point. But L is a topological group and therefore every point is isolated by homogeneity. This and compactness force L to be finite, contradicting our previous conclusion that L was infinite. The contradiction arose from the assumption that the subgroup H was closed, and therefore we see that H cannot be a closed subgroup.■

CONNECTED SUBGROUPS OF TOPOLOGICAL GROUPS

It is also important to know whether a subgroup of a topological group is open, but the reasons for this are fundamentally different, and the following result indicates why this is the case:

PROPOSITION. *An open subgroup of a topological group is also closed.*

In contrast, a closed subgroup of a topological group is not necessarily open, and the simplest example is probably the inclusion of $\mathbf{R}\times\{0\}$ in $\mathbf{R}^2$.

Proof. Let H be an open subgroup of G. Then there exist elements $g_\alpha \in G$ such that the cosets Hg_α are pairwise disjoint and their union is G; we may as well suppose that $g_\beta = 1$ if g_β is the unique element such that $H = Hg_\beta$. Since the right translation maps R_{g_α} are homeomorphisms it follows that each coset is open. Therefore

$$H = G - \bigcup_{g_\alpha \neq 1} Hg_\alpha$$

expresses H as the complement of a union of open subsets, and it follows that H is closed.∎

Here are a few consequences of this relatively elementary but far-reaching observation.

PROPOSITION. (*i*) *The smallest subgroup which contains a fixed open subset of a topological group is both open and closed.*

(*ii*) *The connected component of the identity in a topological group is a normal closed subgroup.*

(*iii*) *If a topological group is connected, then it is generated by every open neighborhood of the identity.*

Proof. (*i*) Let W be an open neighborhood of the identity; by taking the union of W with its inverse we may assume that $\mathbf{inv}(W) = W$. Consider the sequence of sets W_n defined inductively so that $W_1 = W$ and $W_{n+1} = W_n \cdot W$, where the raised dot denotes group multiplication. By construction W_n is mapped to itself by the inverse mapping; we claim that it is also open. To see this, note that if U is an open subset of G and A is an arbitrary set then the identity $A \cdot U = \cup_a L_a(U)$ shows that $A \cdot U$ is open. Consider the set $W_\infty = \cup_n W_n$. If H is an arbitrary subgroup of G containing W, then H certainly contains W_∞. We claim that W_∞ itself is a subgroup; at this point we only need to verify that W_∞ is closed under multiplication, but this is an elementary exercise (work out the details!). Thus W_∞ is the smallest subgroup containing W, and it is an open subset. By our previous result it is also a closed subset.

Proof of (*ii*). As usual, write $a \sim b$ if a and b lie in the same connected component of G. Since left multiplication is continuous we know that $a \sim b$ implies $ga \sim gb$ for all $g \in G$. Therefore if $a \sim 1$ and $b \sim 1$ we obtain the relation $ab \sim a$, so that $ab \sim 1$ by transitivity. Therefore the component of the identity is closed under multiplication. Similarly, if $a \sim 1$ then left multiplication by a^{-1} yields $1 \sim a^{-1}$ and hence that the component of the identity is also closed under taking inverses. Thus this component G° is a subgroup, and it is closed because connected components are closed.

Finally to see that G° is normal, let $x \in G$ and let I_x denote conjugation by x. Since I_x is a homeomorphism and takes 1 to itself, it follows that I_x must also take the connected component of 1, which is G°, to itself, and therefore the latter must be a normal subgroup.

Proof of (*iii*). If U is a neighborhood of the identity we have seen that the smallest subgroup H containing U is open and closed. Since $1 \in H$, the set is also nonempty, and therefore by connectedness we must have $H = G$.∎

Note that the additive groups of p-adic integers and the rationals are examples for which the component of the identity is not an open subgroup.

At this point we shall only state and prove one more result which illustrates the strong properties of topological groups.

PROPOSITION. *If G is a connected topological group and D is a discrete normal subgroup of G, then D is central.*

The normality is crucial in this result. The group $\mathbf{GL}(n, \mathbf{C})$ is connected if $n \geq 2$ (see the exercises for hints on proving this) but it contains many discrete non-central subgroups. One simple example is the subgroup with two elements given by the identity and the diagonal matrix with a -1 in the upper left corner and $+1$ in every other diagonal entry.

Proof. Let $J : G \times D \to D$ be the continuous map sending (g, x) to gxg^{-1}; the image lies in D because the latter is a normal subgroup. What is the image of $G \times \{x\}$? We know it is connected and that it contains x. Since $\{x\}$ is the connected component of x in D because the latter is discrete, it follows that J maps $G \times \{x\}$ onto x. The latter means that $gxg^{-1} = x$ for all $g \in G$, which can be rewritten in the form $gx = xg$ for all $g \in G$. Therefore x lies in the center of G. Since x was arbitrary it follows that all of D lies in the center.■

Analysis on topological groups

This subsection discusses how some important concepts from real variables (uniform continuity and a good theory of integration) can be developed for topological groups. A reader who would prefer to continue with the topological and geometrical discussion may skip this section without missing any topological or geometrical material.

As noted in Section I.3 of the notes, topological groups have a uniform structure that allows one to formulate a useful notion of uniform continuity. In principle, the idea is to consider a family of symmetric neighborhoods of the identity (*i.e.*, neighborhoods mapped to themselves under inversion) and to formulate a uniform concept of closeness using these neighborhoods. For example, if we are given two functions f and g from a set X to the topological group G and a symmetric open neighborhood V of the identity, one can say that f and g are uniformly within V of each other if $f(x) \cdot g(x)^{-1} \in V$ for all $x \in X$

In fact, every topological group can be viewed as a uniform space in *two* ways; the left uniformity turns all left multiplications into uniformly continuous maps while the right uniformity turns all right multiplications into uniformly continuous maps. If G is not abelian, then these two structures need not coincide. These uniform structures allow to talk about notions such as completeness, uniform continuity and uniform convergence on topological groups.

INTEGRATION ON LOCALLY COMPACT TOPOLOGICAL GROUPS

In mathematical analysis, the locally compact groups are of particular importance because they admit a natural notion of measure and integral that was introduced by Alfréd Haar in the nineteen thirties. The idea is to assign a "translation invariant volume" to subsets of a locally compact Hausdorff topological group (*e.g.* if one takes a reasonable subset S and considers its left translate $L_g(S)$ for some $g \in G$, then the volumes of the two sets are equal) and subsequently to define an integral for functions on those groups.

If G is a locally compact Hausdorff topological group, one considers the algebra $\mathcal{M}$ of subsets generated by all compact subsets of G and including all countable unions, countable intersections and complements. If g is an element of G and S is a set in $\mathcal{M}$, then the set $L_a(S)$ is also in $\mathcal{M}$. It turns out that there is, up to a positive multiplicative constant, only one left-translation-invariant measure on $\mathcal{M}$ which is finite on all compact sets, and this is the Haar measure on G (*Note:* There is also an essentially unique right-translation-invariant measure on $\mathcal{M}$, but the two measures need not coincide; the difference between these two measures is completely understood). Using the general Lebesgue integration approach, one can then define an integral for all measurable functions $f : G \to \mathbf{R}$ (or $\mathbf{C}$).

The Internet site

`http://www.wikipedia.org/wiki/Haar_measure`

contains further information about integration on locally compact groups and its uses in mathematics (and physics!).

Remarks on Lie groups

We have already mentioned the particular importance of Lie groups in mathematics. It is beyond the scope of these notes to describe the objects fully (this is material for the third course in the geometry/topology sequence), but we have developed enough ideas to discuss a few important points.

We begin with a discussion of the exponential map for $n \times n$ matrices over the real numbers. Our results on norms for finite-dimensional real vector spaces imply that $\mathbf{R}^k$ is complete with respect to every norm. Take $k = n^2$ and view $\mathbf{R}^{n^2}$ as the space of all $n \times n$ matrices over $\mathbf{R}$. The exponential mapping

$$\exp : \mathbf{M}(n;\mathbf{R}) \longrightarrow \mathbf{GL}(n,\mathbf{R})$$

is the map defined by the familiar infinite series

$$\exp(A) = \sum_{n \geq 0} \frac{1}{n!} A^n$$

and this series converges absolutely because

$$\sum_{n \geq 0} \frac{1}{n!} \parallel A^n \parallel \ \leq \ \sum_{n \geq 0} \frac{1}{n!} \parallel A \parallel^n$$

and the right hand side converges to $\exp(\parallel A \parallel)$.

THEOREM. *The exponential map defines a homeomorphism from an open neighborhood of* $0 \in \mathbf{M}(n;\mathbf{R})$ *to an open neighborhood of* $I \in \mathbf{GL}(n,\mathbf{R})$.

Proof. The key to this is the Inverse Function Theorem. By construction we know that

$$\exp(A) \ - \ \exp(0) \ = \ \exp(A) \ - \ I \ = \ A \ + \ A \cdot \theta(A)$$

where $\lim_{\parallel A \parallel \to 0} \theta(A) = 0$. Therefore it follows that $D\exp(0) = I$. If exp is a $\mathbf{C}^1$ mapping then we can apply the Inverse Function Theorem to complete the proof.

One direct way to verify that exp is $\mathbf{C}^1$ is to use power series. If A is an $n \times n$ matrix and B is another matrix of the same size that is close to zero, then we may write

$$\exp(A+B) - \exp(A) = \sum_{n > 0} \frac{1}{n!} \ [(A+B)^n - A^n]$$

where each of the terms in brackets is a sum of degree n monomials in the noncommuting matrices A and B, and each monomial in the summand has at least one factor equal to B. The n^{th} bracketed term can be written as a sum $f_n(A,B) + g_n(A,B)$ where each term in $f_n(A,B)$ has exactly one B

factor and each term in $g_n(A,B)$ has at least two B factors. The first of these may be written in the form

$$f_n(A,B) \quad = \quad \sum_{r=1}^{n} A^r\,B\,A^{n-r-1}$$

and one can check that the series

$$\Phi(A,B) \quad = \quad \sum_{n>1} \frac{1}{n!}\,f_n(A,B)$$

converges absolutely because $\|\ f_n(A,B)\ \| \leq n\,\|A\|^{n-1}\cdot\|B\|$. Furthermore, it follows that Φ is continuous in A (by uniform convergence of the series) and is linear in B. Therefore we will have that $D\exp(A)(B) = \Phi(A,B)$ and exp is $\mathbf{C}^1$ if and only if for $B \neq 0$ we can show that

$$\lim_{B\to 0} \frac{1}{\|B\|} \sum_{n>0} \frac{1}{n!}\,g_n(A,B) \quad = \quad 0$$

An upper estimate for the norm of the right hand side is given by

$$\frac{1}{\|B\|} \sum_{n>0} \frac{1}{n!} \quad \|\ g_n(A,B)\ \|$$

where each term $\|g_n(A,B)\|$ is bounded from above by

$$\sum_{r=2}^{n} \binom{n}{r} \|B\|^r \|A\|^{n-r} \ .$$

The standard considerations then show that the series

$$\Theta(A,B) = \frac{1}{\|B\|} \sum_{n>0} \frac{1}{n!}\,g_n(A,B)$$

converges absolutely for all A and B and that

$$\lim_{B\to 0} \Theta(A,B) = 0$$

if A is held fixed. Thus we have shown directly that the matrix exponential is a $\mathbf{C}^1$ function.∎

This result yields an important structural property.

PROPOSITION. *There is a neighborhood U of the identity in* $\mathbf{GL}(n,\mathbf{R})$ *such that the only subgroup contained in U is the trivial subgroup.*

The conclusion is often abbreviated to say that $\mathbf{GL}(n,\mathbf{R})$ is a *topological group with no small subgroups* (NSS).

Proof. Choose $\delta > 0$ so that exp maps the disk of radius $2\,\delta$ about the origin homeomorphically to an neighborhood W of the identity in $\mathbf{GL}(n,\mathbf{R})$. Suppose that U is the image of the neighborhood of radius δ in $\mathbf{GL}(n,\mathbf{R})$, and let H be a subgroup that lies entirely in U. Suppose further that H is nontrivial, and choose a matrix A so that $\|A\| < \delta$ and $1 \neq \exp(A) \in U$.

Consider the $\mathbf{C}^\infty$ map $\varphi_A : \mathbf{R} \longrightarrow \mathbf{GL}(n, \mathbf{R})$ sending t to $\exp(tA)$. Since $\exp(P+Q) = \exp(P)\exp(Q)$ if P and Q commute, it follows that φ_A is a continuous homomorphism. The sequence $\{ \|nA\| = n\|A\| \}$ is unbounded and therefore there is a first n such that $\|nA\| = n\|A\| > \delta$. Since $\|A\| < \delta$ it follows that $n\|A\| = \|nA\| < 2\,\delta$. But this means that

$$\exp(nA) \quad = \quad \exp(A)^n \quad \in \quad W - U \ ;$$

on the other hand, since $\exp(A)$ was supposed to belong to a subgroup $H \subset U$, we also have that $\exp(A)^n = \exp(nA) \in U$. The contradiction arises from the assumption that there was a nontrivial subgroup H contained in U, and therefore no such subgroup exists.■

COROLLARY. *If H is a subgroup of $\mathbf{GL}(n, \mathbf{R})$, then H has no small subgroups.*■

Results of A. Gleason and (jointly and independently) D. Montgomery and L. Zippin provide *a completely topological characterization of Lie groups as locally compact Hausdorff groups with no small subgroups.* These results are proved in the book Montgomery and Zippin as well as Kaplansky's *Lie Algebras and Locally Compact Groups.*

Note that a locally compact Hausdorff topological group that is locally connected and second countable does not necessarily satisfy the "no small subgroups" condition. A countably infinite product of copies of S^1 is a relatively simple counterexample.

Appendix B : Stereographic projection and inverse geometry

The conformal property of stereographic projections can be established fairly efficiently using the concepts and methods of inverse geometry. This topic is relatively elementary, and it has important connections to complex variables and hyperbolic (= Bolyai-Lobachevsky noneuclidean) geometry.

Definition. Let $r > 0$ be a real number, let $y \in \mathbf{R}^n$, and let $S(r;y)$ be the set of all points $x \in \mathbf{R}^n$ such that $|x - y| = r$. The *inversion map with respect to the sphere* $S(r;y)$ is the map T on $\mathbf{R}^n - \{0\}$ defined by the formula

$$T(x) = y \ + \ \frac{r^2}{|x-y|} \cdot \big(x - y\big) \ .$$

Allternatively, $T(x)$ is defined so that $T(x) - y$ is the unique positive scalar multiple of $x - y$ such that

$$\big|T(x) - y\big) \cdot \big|x - y\big) = r^2 \ .$$

Another way of saying this is that inversion interchanges the exterior points to $S(r;y)$ and the interior points with the center deleted. If $r = 1$ and $y = 0$ then inversion simply takes a nonzero vector x and sends it to the nonzero vector pointing in the same direction with length equal to the reciprocal of $|x|$ (this should explain the term "inversion").

If $n = 2$ then inversion corresponds to the *conjugate* of a complex analytic function. Specifically, if a is the center of the circle and r is the radius, then inversion is given in complex numbers by the formula

$$T(z) \quad = \quad r^2 \cdot \big(\overline{z} - a\big)^{-1} \quad = \quad r^2 \cdot \overline{\big(z - \overline{a}\big)^{-1}}$$

where the last equation holds by the basic properties of complex conjugation. If $r = 1$ and $a = 0$ then inversion is just the conjugate of the analytic map sending z to z^{-1}.

The geometric properties of the analytic inverse map on the complex plane are frequently discussed in complex variables textbooks. In particular, this map has nonzero derivative wherever the function is defined, and accordingly the map is conformal. Furthermore, the map $z \to z^{-1}$ is an involution (its composite with itself is the identity), and it sends circles not containing 0 to circles of the same type. In addition, it sends lines not containing the origin into circles containing the origin and vice versa (*Note:* This means that 0 lies on the circle itself and NOT that 0 is the center of the circle!). Since complex conjugation sends lines and circles to lines and circles, preserves the angles at which curves intersect and also sends 0 to itself, it follows that the inversion map with respect to the unit circle centered at 0 also has all these properties).

It turns out that all inversion maps have similar properties. In particular, they send the every point of sphere $S(r;y)$ to itself and interchange the exterior points of that sphere with all of the interior points except y (where the inversion map is not defined), they preserve the angles at which curves intersect, they are involutions, they send hyperspheres not containing the central point y to circles of the same type, and they send hyperplanes not containing $\{y\}$ into hyperspheres containing $\{y\}$ and vice versa. We shall limit our proofs to the properties that we need to study stereographic projections; the reader is encouraged to work out the proofs of the other assertions.

We begin by recalling(?) some simple observations involving isometries and similarity transformations from $\mathbf{R}^n$ to itself. Proofs or hints for proofs in the cases of isometries can be found in many standard linear algebra texts.

FACT 1. *If $b \in \mathbf{R}^n$ and F is translation by b (formally, $F(x) = x + b$), then F is an isometry from $\mathbf{R}^n$ to itself and F sends the straight line curve from x to y defined by*

$$\alpha(t) = ty + (1-t)x$$

to the straight line curve from $F(x)$ to $F(y)$ defined by

$$\alpha(t) = tF(y) + (1-t)F(x) \ .$$

Definition. If $r > 0$ then a *similarity transformation with ratio of similitude* r on a metric space X is a 1–1 correspondence f from X to itself such that $\mathbf{d}\big(f(x), f(y)\big) = r \cdot \mathbf{d}(x,y)$ for all $x,\ y \in X$.

Note that every isometry (including every identity map) is a similarity transformation with ration of similitude 1 and conversely, the inverse of a similarity transformation with ratio of similitude r is a similarity transformation with ratio of similitude r^{-1}, and the composite of two similarity transformations with ratios of similitude r and s is a similarity transformation with ratio of similitude rs. Of course, if $r > 0$ then the invertible linear transformation rI on $\mathbf{R}^n$ is a similarity transformation with ratio of similitude r.■

FACT 2. *In $\mathbf{R}^n$ every similarity transformation with ratio of similitude r satisfying $F(0) = 0$ has the form $F(x) = rA(x)$ where A is given by an $n \times n$ orthogonal matrix.*■

FACT 3. *In $\mathbf{R}^n$ every similarity transformation F is conformal; specifically, if α and β are differentiable curves in $\mathbf{R}^n$ that are defined on a neighborhood of $0 \in \mathbf{R}$ such that $\alpha(0) = \beta(0)$ such that both $\alpha'(0)$ and $\beta'(0)$ are nonzero, then the angle between $\alpha'(0)$ and $\beta'(0)$ is equal to the angle between $[F \circ \alpha]'(0)$ and $[F \circ \beta]'(0)$.*■

The verification of the third property uses Facts 1 and 2 together with the additional observation that if A is given by an orthogonal transformation then A preserves inner products and hence the cosines of angles between vectors.

The key to relating inversions and stereographic projections is the following result:

PROPOSITION. *Let $e \in \mathbf{R}^n$ be a unit vector, and let T be inversion with respect to the sphere $S(1;0)$. Then T interchanges the hyperplane defined by the equation $\langle x, e \rangle = -1$ with the nonzero points of the sphere $S(\frac{1}{2};\ -\frac{1}{2}e)$.*

Proof. By definition we have

$$T(x) = \frac{1}{\langle x, x \rangle} \cdot x$$

and therefore the proof amounts to finding all x such that

$$\left| T(x) + \tfrac{1}{2}e \right| = \tfrac{1}{2}.$$

This equation is equivalent to

$$\tfrac{1}{4} = \left| T(x) + \tfrac{1}{2}e \right|^2 = \langle T(x) + \tfrac{1}{2}e,\ T(x) + \tfrac{1}{2}e \rangle$$

and the last expression may be rewritten in the form

$$\langle T(x), T(x)\rangle + \langle T(x), e\rangle + \frac{1}{4} =$$

$$\frac{\langle x, x\rangle}{\langle x, x\rangle^2} + \frac{\langle x, e\rangle}{\langle x, x\rangle} + \frac{1}{4}$$

which simplifies to

$$\frac{1}{\langle x, x\rangle} + \frac{\langle x, e\rangle}{\langle x, x\rangle} + \frac{1}{4}\ .$$

Our objective was to determine when this expression is equal to $\frac{1}{4}$, and it follows immediately that the latter is true if and only if $1 + \langle x, e\rangle = 0$; *i.e.*, it holds if and only if $\langle x, e\rangle = -1$.∎

An llustration of the preceding result appears in the files `stereopic1.*` in the course directory.

COROLLARY. *Let e be as above, and let W be the $(n-1)$-dimensional subspace of vectors that are perpendicular to e. Then the stereographic projection map from $S(1;0) - \{e\}$ to W is given by the restriction of the composite*

$$G \circ t \circ H$$

to $S(1;0) - \{e\}$, where H is the similarity transformation $H(u) = \frac{1}{2}(u-e)$ and G is the translation isometry $G(v) = v + e$.

Proof. It will be convenient to talk about the *closed ray* starting at a vector a and passing through a vector b; this is the image of the parametrized curve

$$\gamma(t) = (1-t)a + tb = a + t(b-a)$$

where $t \geq 0$.

First note that the composite $T \circ H$ sends the ray starting at the point e and passing through a point x with $\langle x, e\rangle < 1$ to the ray starting at e and passing through the point $H(x)$, and the latter satisfies $\langle H(x), e\rangle < 0$. This is true for the mapping H by Fact 1 and the equation $H(e) = 0$, and it is true for $T \circ H$ because T is inversion with respect to a sphere centered at 0. Furthermore, by construction H sends the sphere $S(1;0)$ to $S\left(\frac{1}{2};\ -\frac{1}{2}e\right)$, and it also sends the hyperplane P defined by $\langle x, e\rangle = -1$ to itself.

By construction, stereographic construction sends the point $y \in S(1;0) - \{e\}$ to the point $z \in W$ such that $z - e$ is the unique point at which the ray starting at e and passing through y meets the hyperplane P, and inversion sends the point $\eta \in S\left(\frac{1}{2};\ -\frac{1}{2}e\right) - \{0\}$ to the unique point α at which the ray starting at 0 and passing through η meets the hyperplane P.

Combining these, we see that $T \circ H$ maps y to the unique point where the ray passing through 0 and y meets the hyperplane P. This point may be written uniquely in the form $w - e$ where $w \in W$, and in fact we have $w = G \circ T \circ H(\mathrm{y})$. On the other hand, by the preceding paragraph we also know that w is given by the stereographic projection.∎

The conformal property for inversions

The following is an immediate consequence of the Chain Rule and Fact 3 stated above:

PROPOSITION. *Let U be open in $\mathbf{R}^n$, let $x \in U$ and let $f : U \to \mathbf{R}^n$ be a $\mathbf{C}^1$ mapping. Then f is conformal at x if $Df(x)$ is a nonzero scalar multiple of an orthogonal transformation.*

The key observation behind the proposition is that if $\gamma : (-\delta, \delta) \to U$ is a differentiable curve with $\gamma(0) = x$ then

$$[f \circ \gamma]'(0) = Df(x)[\gamma'(0)]$$

by the Chain Rule.∎

We are now ready to prove the result that we wanted to establish.

THEOREM. *Every inversion map is conformal.*

Proof. It is convenient to reduce everything to the case where the sphere is $S(1;0)$. Given an arbitrary sphere $S(r;y)$ there is a similarity transformation sending $S(r;y)$ to $S(1;0)$ that is defined by the formula $F(x) = r^{-1}(x - y)$, and if T' and T are the associated inversions then we have

$$T' = F^{-1} \circ T \circ F \ .$$

By construction $DF(x) = r^{-1}I$ for all x and therefore it follows that $DT' = DT$ (as usual, "D" denotes the derivative of a function). Therefore, by the proposition it will suffice to show that DT is always a scalar multiple of an orthogonal map.

Let $x \neq 0$ and write an arbitrary vector $v \in \mathbf{R}^n$ as a sum $v = cx + u$ where $c \in \mathbf{R}$ and $\langle x, u \rangle = 0$. We have already noted that $T(x) = \rho(x)^{-2} \cdot x$ where $\rho(x) = |x|$, and we wish to use this in order to compute the value of $DT(x)$ at some vector $h \in \mathbf{R}^n$. The appropriate generalization of the Leibniz Rule for products and elementary multivariable calculus show that

$$DT(x)[h] = \rho(x)^{-2}h + \langle (\nabla[\rho(x)]^{-2}, h \rangle \cdot x$$

where

$$\nabla[\rho(x)]^{-2} = -[\rho(x)]^{-4} \cdot \nabla[\rho(x)]^2 = -2[\rho(x)]^{-4}x \ .$$

If $\langle x, h \rangle = 0$ this shows that $DF(x)[h] = [\rho(x)]^{-2}h$, while if $h = x$ it follows that $DF(x)[x] = -[\rho(x)]^{-2}x$. In particular, this shows that there is an orthonormal basis for $\mathbf{R}^n$ consisting of eigenvectors for $DT(x)$ with associated eigenvalues $\pm\rho(x)^{-2}$, and therefore it follows that $DT(x)$ is a positive scalar multiple of an orthogonal map, which in turn implies that T is conformal at x. Since x was arbitrary, this proves the theorem.∎

Appendix C : Homogeneity of open sets in Euclidean spaces

The purpose of this section is to prove a result that was stated in Appendix A:

HOMOGENEITY THEOREM. *If U is an open connected subset of $\mathbf{R}^n$ and a and b are distinct points of U, then there is a homeomorphism $h : U \to U$ such that $h(a) = b$.*

There are two major steps in the proof; the first is to show that connected open sets are locally homogeneous, and the second is to use local homogeneity and connectedness to prove global homogeneity.

A local construction

The following result provides one way to verify local homogeneity:

PROPOSITION. *Let D^n be the solid unit disk in $\mathbf{R}^n$, and let $v \in D^n$ be an interior point with $|v| < 1$. Then there is a homeomorphism $f : D^n \to D^n$ such that f is the identity on S^{n-1} and $f(0) = v$.*

Sketch of Proof. The geometric motivation is simple. Every point on D^n lies on a closed segment joining the origin to a point on S^{n-1}. One maps such a segment linearly to the segment joining v to the same point on S^{n-1}. As is often the case with such geometrical ideas, it takes a fair amount of algebraic manipulation to show that this actually works. An illustration of the basic idea is given in the files `radialproj.*` in the course directory.

The following result will be useful in the course of the proof:

LEMMA. *D^n is homeomorphic to the quotient space of $S^{n-1} \times [0, 1]$ whose equivalence classes are the one point subsets for all points with positive first coordinates together with the set $S^{n-1} \times \{0\}$.*

Proof of lemma. Consider the continuous map $g : S^{n-1} \times [0, 1] \to D^n$ sending (x, t) to tx. The inverse images of points are precisely the classes described above, and since g is a closed and surjective mapping it follows that the quotient space with the given equivalence classes is homeomorphic to the image, which is D^n.∎

Proof of proposition continued. Define $F : S^{n-1} \times [0, 1] \to D^n$ by the formula

$$F(u, t) = tu + (1 - t)v \ .$$

If $t = 0$ then $F(u, t) = F(u, 0) = v$ and therefore the lemma implies that F passes to a continuous mapping $f : D^n \to D^n$. This map corresponds to the geometric idea proposed in the first paragraph of the proof. We need to prove that f is 1–1 and onto.

By construction we have $f(0) = v$; it will be useful to start by proving that no other point maps to v. But suppose that we have $f(tu) = v$ where $t \in (0, 1]$ and $|u| = 1$. By construction this means that

$$v = tu + (1 - t)\, v$$

which is equivalent to the equation $t(u - v) = 0$. Since $|v| < 1 = |u|$ we clearly have $u \neq v$ so this forces the conclusion that $t = 0$, which contradicts our original assumption that $t > 0$.

Suppose now that we have $w \in D^n - \{v\}$. Consider first the case where $w \in S^{n-1}$. Then by construction $w = f(w)$, and there is no other point w^* on S^{n-1} such that $w^* = f(w)$. On the other hand, if $|\xi| < 1$ we claim that $f(\xi) \notin S^{n-1}$ because

$$|f(\xi)| \;=\; \big|\, |\xi|u + (1-|\xi|)v \,\big| \;\leq\; |\xi| + (1-|\xi|)\cdot|v| \;<\; |\xi| + (1-|\xi|) \;=\; 1\ .$$

Therefore we only need to show that *if* $|w| < 1$ *then there is a unique point* ξ *such that* $|\xi| < 1$ *and* $f(\xi) = w$. This in turn reduces to showing that there is a unique point $u \in S^{n-1}$ such that w lies on the open segment joining v and u; the latter is equivalent to showing that *there is a unique real number* t *such that* $t > 1$ *and*

$$\big|v + t(w-v)\big| = 1\ .$$

Rather than prove this by working out an explicit but complicated formula for t in terms of v and w, we shall approach the assertion by analyzing the behavior of the quadratic function

$$f(t) \;=\; \big|v + t(w-v)\big|^2$$

which represents the square of the distance between the point $v + t(w-v)$ and the origin. The coefficient of t^2 for this function is the positive number $|w-v|^2$ and therefore $\lim_{t\to\pm\infty} f(t) = +\infty$. Furthermore, $f(0) = |v|^2 < 1$ and $f(1) = |w|^2 < 1$, and therefore there are exactly two values of t such that $f(t) = 1$, one of which is greater than 1 and one of less is less than 0.

If s is chosen so that $s > 1$ and $f(s) = 1$ in the notation of the preceding paragraphs and we let $u = v + s(v-w)$, then $f(x) = w$ if and only if

$$x \;=\; w \;+\; \left(\frac{1-s}{s}\right) v\ .\blacksquare$$

The local homogeneity of an open subset of $\mathbf{R}^n$ is an direct consequence of the proposition. Note first that the proposition above remains true for every closed disk in $\mathbf{R}^n$ of the form

$$\mathbf{D}(y;r) \;=\; \{\ y \in \mathbf{R}^n \mid |y-a| = r\ \}\ .$$

Local homogeneity

The following local result is an immediate consequence of the preceding proposition:

LOCAL HOMOGENEITY PROPERTY. *If* U *is open in* $\mathbf{R}^n$ *and* $a \in U$, *then there is an open neighborhood* v *of* x *such that* $V \subset U$ *and for all* $x \in V$ *there is a homeomorphism* $h : U \to U$ *sending* a *to* x.$\blacksquare$

Proof. Note first that the proposition above remains valid for every closed disk in $\mathbf{R}^n$ of the form

$$\mathbf{D}(y;r) \;=\; \{\ y \in \mathbf{R}^n \mid |y-a| \leq r\ \}\ .$$

To see this, observe that this disk is the image of D^n under the self-homeomorphism h of $\mathbf{R}^n$ sending x to $rx + a$. Thus if $b \in \mathbf{D}(y;r)$ with $|b-a| < r$ we can construct a homeomorphism sending a to b and fixing the boundary sphere by the formula

$$g(y) = h\left(f\left(h^{-1}(y)\right)\right)$$

where h is the homeomorphism in the proposition sending 0 to $h^{-1}(b)$. Given an open set $U \subset \mathbf{R}^n$ and a point $a \in U$, one can find some $r > 0$ so that $\mathbf{D}(y; r) \subset U$. For each b in the interior of this disk we can construct a homeomorphism of the closed disk sending a to b such that the restriction of the homeomorphism to the boundary is the identity. Given two points $b\lambda$, b' in the interior of this disk one can use such homeomorphisms and their inverses to construct a homeomorphism of the disk that is again fixed on the boundary sphere and sends b to b'. This homeomorphism can be extended to the entire open set U by defining it to be the identity on the complement of the disk (why does this work?). This proves the desired property.■

A global homogeneity theorem

Suppose now that $U \subset \mathbf{R}^n$ is open and connected. The following result allows us to extend the local homogeneity property to a global result:

PROPOSITION. *Let X be a connected topological space, and for each $x \in X$ suppose that there is an open neighborhood V of X such that for each $v \in V$ there is a homeomorphism $h : X \to X$ such that $h(x) = v$. Then X is homogeneous; in other words, for each pair of distinct points $x, y \in X$ there is a homeomorphism $h : X \to X$ such that $h(x) = y$.*

Proof. Consider the binary relation on X given by $a \sim b$ if and only if there is a homeomorphism $h : X \to X$ such that $h(a) = b$. It follows immediately that this is an equivalence relation. Furthermore, if C is an equivalence class of $\sim$ and $x \in C$, then there is an open neighborhood V of x such that $V \subset C$. In particular, it follows that C is open. Since this is true for all equivalence classes it follows that the latter decompose X into pairwise disjoint open subsets. Therefore the union of all equivalence classes except C is also open, and hence C is closed. But X is connected, and therefore we must have $C = X$, so that all points of X are equivalent under $\sim$. By the definition of the latter, it follows that X is homogeneous.■

Appendix D : Normal forms for orthogonal transformations

The Spectral Theorem in linear algebra implies that a normal linear transformation on a complex inner product space (one that commutes with its adjoint) has an orthonormal basis of eigenvectors. In particular, since the adjoint of a unitary transformation is its inverse, the result implies that every unitary transformation has an orthonormal basis of eigenvectors.

It is clear that one cannot have a direct generalization of the preceding result to orthogonal transformations on real inner product spaces. In particular, plane rotations given by matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

do not have such a basis over the reals except in the relatively trivial cases when θ is an integral multiple of π and the matrices reduce to $(-I)^k$ for some integer k. However, if one takes these into account it is possible to prove the following strong result on the existence of a "good" orthonormal basis for a given orthogonal transformation.

NORMAL FORM. *Let V be a finite-dimensional real inner product space, and let $T : V \to V$ be an orthogonal transformation of V. Then there is an orthonormal direct sum decomposition of V into T-invariant subspaces W_i such that the dimension of each W_i is either* 1 *or* 2.

In particular, this result implies that there is an ordered orthonormal basis for V such that the matrix of T with respect to this ordered orthonormal basis is a block sum of 2×2 and 1×1 orthogonal matrices.

It is beyond the scope of these notes to go into detail about the results from a standard linear algebra course that we use in the proof of the result on normal forms. Virtually all of the background information can be found in nearly any linear algebra text that includes a proof of the Spectral Theorem (*e.g.*, the book by Fraleigh and Beauregard cited in the bibliography).

Small dimensional orthogonal transformations

Since orthogonal transformations preserve the lengths of vectors it is clear that a 1-dimensional orthogonal transformation is just multiplication by ± 1. It is also not difficult to describe 2-dimensional orthogonal transformations completely using the fact that their columns must be orthonormal. In particular, it follows that the matrices representing 2-dimensional orthogonal transformations of $\mathbf{R}^2$ with respect to the standard inner product have the form

$$\begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix}$$

for some real number θ. We have already noted that one class of cases corresponds to rotation through θ, and it is an exercise in linear algebra to check that each of the remaining matrices

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

has an orthonormal basis of eigenvectors, with one vector in the basis having eigenvalue -1 and the other having eigenvalue $+1$. The details of verifying this are left to the reader as an exercise (*Hint:* First verify that the characteristic polynomial is $t^2 - 1$).

One can use this to prove a geometrically sharper version of the result on normal form.

STRONG NORMAL FORM. *Let V be a finite-dimensional real inner product space, and let $T : V \to V$ be an orthogonal transformation of V. Then there is an orthonormal direct sum decomposition of V into T-invariant subspaces W_i such that the dimension of each W_i is either 1 or 2 and T operates by a plane rotation on each 2-dimensional summand.*

This is an immediate consequence of the preceding discussion, for if T operates on an invariant 2-dimensional subspace with by a map that is not a rotation, then we can split the subspace into two 1-dimensional eigenspaces.

Complexification

If $V = \mathbf{R}^n$ and T is represented by the orthogonal matrix A, it is clear how one can extend T to a unitary transformation on $\mathbf{C}^n$. We shall need a version of this principle that works for an arbitrary real inner product space V. The idea is simple but a little inelegant; it can be done better if one uses tensor products, but we want to prove the result without introducing them.

One defines the complexification of V formally to be very much as one defines the complex numbers. The underlying set of complex vectors $V_{\mathbf{C}}$ is given by $V \times V$, addition is defined in a coordinatewise fashion, multiplication by a complex scalar $a + bi$ is given by the formula

$$[a + bi](v, w) = (av - bw, bv + aw)$$

and the complexified inner product is defined by

$$\langle\, (v, w), (v', w')\, \rangle_{\text{complex}} =$$

$$(\langle v, v' \rangle + \langle w, w' \rangle) + i \cdot (\langle w, v' \rangle - \langle v, w' \rangle) \ .$$

Both intuitively and formally the pair (v, w) can be viewed as $v + i\,w$ if one identifies a vector v in the original space with $(v, 0)$ in the complexification. Verification that the structure defined above is actually a complex inner product space is essentially an exercise in bookkeeping and will not be carried out here.

If we are given a linear transformation $T : V \to W$ of real inner product spaces, then

$$T \times T : V \times V \to W \times W$$

defines a complex linear transformation $T_{\mathbf{C}}$ on the complexification. Moreover, this construction is compatible with taking adjoints and composites:

$$(T^*)_{\mathbf{C}} = (T_{\mathbf{C}})^*$$

$$(S \circ T)_{\mathbf{C}} = S_{\mathbf{C}} \circ T_{\mathbf{C}}$$

In particular, *if T is orthogonal then $T_{\mathbf{C}}$ is unitary.*

The main ideas behind the proof are

(1) to extract as much information as possible using the Spectral Theorem for the unitary transformation $T_{\mathbf{C}}$,

(2) prove the result by induction on the dimension of V.

In order to do the latter we need the following observation that mirrors one step in the proof of the Spectral Theorem.

PROPOSITION. *Let T be as above, and assume that $W \subset V$ is T-invariant. Then the orthogonal complement $W^\perp$ is also T-invariant.*■

In the Spectral Theorem the induction proof begins by noting that there is an invariant 1-dimensional subspace corresponding to an eigenvector for T. One then obtains a transformation on the orthogonal complement of this subspace and can apply the inductive hypothesis to the associated transformation on this invariant subspace. We would like to do something similar here, and in order to begin the induction we need the following result.

KEY LEMMA. *If V is a finite dimensional real inner product space and $T : V \to V$ is an orthogonal transformation, then there is a subspace $W \subset V$ of dimension 1 or 2 that is T-invariant.*

If we have this subspace, then we can proceed as before, using the induction hypothesis to split $W^\perp$ into an orthogonal direct sum of T-invariant subspaces of dimension 1 or 2, and this will complete the derivation of the normal form.■

Proof of the key lemma. Suppose that λ is an eigenvalue of $T_{\mathbf{C}}$ and x is a nonzero eigenvector; since $T_{\mathbf{C}}$ is unitary we know that $|\lambda| = 1$. Express both λ and x in terms of their real and imaginary components, using the fact that $|\lambda| = 1$:

$$\lambda \quad = \quad \cos\theta + i\,\sin\theta, \qquad x \quad = \quad (v, w) \quad = \quad v + i\,w$$

Then the eigenvalue equation $T_{\mathbf{C}}(x) = \lambda\,x$ may be rewritten in the form

$$T(v) + i\,T(w) \quad = \quad T_{\mathbf{C}}(x) \quad = \quad \lambda\,x \quad = \quad (\cos\theta + i\,\sin\theta)\cdot(v + i\,w) \quad =$$

$$\Big(\cos\theta\,v - \sin\theta\,w\Big) + i\,\Big(\sin\theta\,v + \cos\theta\,w\Big)$$

which yields the following pair of equations:

$$T(v) \quad = \quad \cos\theta\,v - \sin\theta\,w$$

$$T(w) \quad = \quad \sin\theta\,v + \cos\theta\,w$$

It follows that the subspace W spanned by v and w is a T-invariant subspace, and since it is spanned by two vectors its dimension is at most 2. On the other hand, since $x \neq 0$ we also know that at least one of v and w is nonzero and therefore the dimension of W is at least 1.■

References

The following are the texts for the course:

J. R. Munkres. Topology. (Second Edition), *Prentice-Hall, Saddle River NJ*, 2000. ISBN: 0–13–181629–2.

C. H. Edwards, Jr. Advanced Calculus of Several Variables. (Corrected Reprint of 1973 Edition). *Dover, Mineola NY*, 1994. ISBN: 0–496–68336–2.

Further references

Alexandroff, Pavel S.; Hopf, Heinz. Topologie, Band 1. (German. Grundlehren der mathematischen Wissenschaften *etc.* Nr. 45.) *Julius Springer, Berlin*, 1935.

Amadio, Roberto M.; Curien, Pierre-Louis. Domains and Lambda Calculi. (Cambridge Tracts in Theoretical Computer Science, 46.) *Cambridge University Press, Cambridge*, 1998. ISBN: 0-521-62277-8.

Birkhoff, Garrett; MacLane, Saunders. A Survey of Modern Algebra. (Reprint of the Third 1968 Edition.) *Chelsea, New York NY*, 1988. ISBN: 0-023-74310-7.

Birkhoff, Garrett; MacLane, Saunders. A Survey of Modern Algebra. (Fourth Edition.) *Prentice-Hall, Englewood Cliffs NJ*, 1977. ISBN: 0-023-10070-2.

Bourbaki, Nicolas. General Topology, Chapters 1–4. (Translated from the French. Reprint of the 1989 English translation. Elements of Mathematics.) *Springer-Verlag, Berlin-Heidelberg-New York, etc.*, 1998. ISBN: 3-540-64241-2.

Bourbaki, Nicolas. General Topology, Chapters 5–10. (Translated from the French. Reprint of the 1989 English translation. Elements of Mathematics.) *Springer-Verlag, Berlin-Heidelberg-New York, etc.*, 1998. ISBN: 3-540-64563-2 .

Christenson, Charles O.; Voxman, William L. Aspects of Topology. (Pure and applied Mathematics, Vol. 39.) *Marcel Dekker, New York-Basel*, 1977.

Christenson, Charles O.; Voxman, William L. Aspects of Topology. (Second Edition.) *BCS Associates, Moscow ID*, 1998. ISBN: 0-914351-07-9; 0-914351-08-7.

Dieudonné, Jean A. History of algebraic geometry. An outline of the history and development of algebraic geometry. (Translated from French by Judith D. Sally. Wadsworth Mathematics Series.) *Wadsworth International, Belmont, CA*, 1985. ISBN: 0-534-03723-2.

Dugundji, James. Topology. (Reprint of the 1966 Edition. Allyn and Bacon Series in Advanced Mathematics. *Allyn and Bacon, Boston MA-London (U.K.)-Sydney (Austr.)*, 1978. ISBN: 0-205-00271-4.

Fraleigh, John B.; Beauregard, Raymond A. Linear Algebra (Third Edition. Historical notes by Victor J. Katz) *Addison-Wesley, Reading MA etc.*, 1995. ISBN: 0-2-1-52675-1.

Franz, Wolfgang. General Topology. (Translated from the German by Leo F. Boron.) *Frederick Ungar, New York NY*, 1965.

Goffman, Casper; Pedrick, George. First Course in Functional Analysis. (Second Edition.) *Chelsea, New York NY*, 1983. ISBN: 0-828-40319-8.

Hewitt, Edwin; Stromberg, Karl. Real and Abstract Analysis. A Modern Treatment of the Theory of Functions of a Real Variable. (Graduate Texts in Mathematics Vol. 25.) *Springer-Verlag, Berlin-Heidelberg-New York-etc.*, 1997. ISBN: 0-387-90138-8.

Hocking, John G.; Young, Gail S. Topology. (Second edition.) *Dover, New York NY*, 1988. ISBN: 0-486-65676-4.

James, I. M. (*ed.*). History of topology. *North-Holland, Amsterdam (NL)*, 1999. ISBN: 0-444-82375-1.

Jänich, Klaus. Topology. (With a chapter by Theodor Bröcker. Translated from the German by Silvio Levy. Undergraduate Texts in Mathematics.) *Springer-Verlag, Berlin-Heidelberg-New York*, 1984. ISBN: 0-387-90892-7.

Kaplansky, Irving. Set Theory and Metric Spaces. (Second edition.) *Chelsea, New York*, 1977. ISBN 0-8284-0298-1.

Kaplansky, Irving. Lie Algebras and Locally Compact Groups. (Chicago Lectures in Mathematics Series.) *University of Chicago Press, Chicago IL*, 1995. ISBN: 0-226-42453-7.

Kasriel, Robert H. Undergraduate Topology. (Reprint of the 1971 Edition.) *Robert E. Krieger, Huntington NY*, 1977. ISBN: 0-721-65298-0

Kelley, John L. General Topology. (Reprint of the 1955 Edition. Graduate Texts in Mathematics, No. 27.) *Springer-Verlag, Berlin-Heidelberg-New York*, 1975. ISBN: 0-387-90125-6.

Krantz, Steven G. Real Analysis and Foundations. (Studies in Advanced Mathematics.) *CRC Press, Boca Raton FL*, 1991. ISBN: 0-8493-7156-2.

McCarty George S. Topology : An Introduction with Application to Topological Groups. (Second Edition.) *Dover, New York NY*, 1988. ISBN: 0-486-65633-0.

Montgomery, Deane; Zippin, Leo. Topological Transformation Groups. (Reprint of the 1955 Edition.) *Robert E. Krieger, Huntington NY*, 1974. ISBN: 0-88275-169-7

Pontryagin, Lev S. Topological Groups. (Translated from the Second Russian Edition with a preface by Arlen Brown, with additional material translated by P. S. V. Naidu. Classics of Soviet Mathematics) *Taylor and Francis, London (U. K.)*, 1987. ISBN: 2-881-24133-6

Rickart, Charles E. General Theory of Banach Algebras. (Reprinted with Corrections.) *Robert E. Krieger, Huntington NY*, 1974. ISBN: 0-88275-091-7.

Royden, Halsey L. Real Analysis. (Third Edition.) *Macmillan, New York NY*, 1988. ISBN: 0-02-404151-3.

Rudin, Walter. Principles of Mathematical Analysis. (Third Edition. International Series in Pure and Applied Mathematics.) *McGraw-Hill, New York-Auckland-Düsseldorf*, 1976. ISBN: 0-07-054235-X.

Rudin, Walter. Functional Analysis. (Second Edition. International Series in Pure and Applied Mathematics.) *McGraw-Hill, New York NY, etc.*, 1991. ISBN: 0-07-054236-8.

Steen, Lynn Arthur; Seebach, J. Arthur, Jr. Counterexamples in Topology. (Reprint of the Second (1978) Edition.) *Dover, Mineola NY*, 1995. ISBN: 0-486-68735-X.

Taylor,Paul. Practical Foundations of Mathematics. (Cambridge Studies in Advanced Mathematics 59.) *Cambridge University Press, Cambridge*, 1998. ISBN: 0-521-63107-6.

Wolf, Robert S. Proof, Logic, and Conjecture: The Mathematician's Toolbox. *W. H. Freeman, New York NY*, 1998. ISBN: 0-7167-3050-2.

Summary of Files in the Course Directory

The course directory is $\sim$`res/math205A` on the `math.ucr.edu` network.

`braintest.pdf`
`braintest.ps`

This is definitely not a serious piece of course material, but it does illustrate the importance of staying focused on the main points when learning or doing mathematics.

`categories.pdf`
`categories.ps`

A brief survey of category theory, not needed for the course but included as background.

`concat.pdf`
`concat.ps`

A picture illustrating the stringing together, or concatenation, of two curves, where the ending point of the first is the starting point of the second.

`contents.dvi`
`contents.pdf`
`contents.ps`

The title page and table of contents for this document.

`coursehw.dvi`
`coursehw.pdf`
`coursehw.ps`

The electronic files containing the homework assignments for this course; some of these come from the two texts and others are also included. Solutions will be provided in the course directory. Preliminary versions with names of the form `prelimhw.`$*$ will initially be in the course directory.

`coursetext.dvi`
`coursetext.pdf`
`coursetext.ps`

The electronic files for viewing or printing this document. Preliminary versions with names of the form `prelimtext.`$*$ will initially be in the course directory.

`cubicroots.dvi`
`cubicroots.pdf`
`cubicroots.ps`

This note discusses the application of the Contraction Lemma to finding the roots of cubic polynomials with real coefficients.

`foundations1.pdf`
`foundations1.ps`

A brief discussion of basic logic and some very elementary set theory. This is meant as review and/or background.

foundations2.pdf
foundations2.ps

This discusses topics on relations and functions from the perspective of this course to the extent that it differs from the main text; it also includes comments on axioms for the positive integers and discussions of the roles of the Axiom of Choice and the Generalized Continuum Hypothesis in set theory. The latter discussions are not needed for the course but may be helpful for obtaining a broader perspective of the mathematical foundations underlying the course.

general205A.dvi
general205A.pdf
general205A.ps

This is the handout from the first day of class.

math205Afall03.pdf
math205Afall03.ps

The course outline.

nicecurves.dvi
nicecurves.pdf
nicecurves.ps

This an appendix to the course notes, and it proves an assertion from Section III.5 about joining two points in a connected open subset of Euclidean space by a curve that is infinitely differentiable and has nonzero tangent vectors at every point.

ordinals.dvi
ordinals.pdf
ordinals.ps

Background material on the ordinal numbers. The latter may be viewed as equivalence classes of well-ordered sets, and the class of all ordinals determines a natural indexing for the class of all cardinal numbers. This summary, which was written by N. Strickland of the University of Sheffield, is available online from

http://www.shef.ac.uk/ pm1nps/courses/topology/ordinals.∗

and is included in the course directory only for the convenience of students enrolled in this course. The site also includes several other interesting and informative documents related to this course.

polya.pdf
polya.ps

A one page summary of the advice for solving problems in the classic book, *How to Solve It*, by G. Pólya (with an extra piece of advice added at the end).

proper.dvi
proper.pdf
proper.ps

A discussion of the basic facts about proper maps, with emphasis on their properties that are relevant to algebraic geometry. This is supplementary material not covered in the course, and the level of exposition is slightly higher than in the course notes.

`radproj.pdf`
`radproj.ps`

A picture illustrating the radial projection map that is used to prove a result in Appendix C.

`realnumbers.pdf`
`realnumbers.ps`

A summary of the basic properties of the real numbers.

`smirnov.dvi`
`smirnov.pdf`
`smirnov.ps`

A proof of Smirnov's Theorem on constructing global metrics for topological spaces whose topologies locally come from metric spaces. This is supplementary material not covered in the course, and the level of exposition is slightly higher than in the course notes.

`solutions*.dvi`
`solutions*.pdf`
`solutions*.ps`

Solutions to the assigned exercises in the course, where $*$ ranges from 1 to 5. The first file contains solutions through Section II.1, the second contains solutions through the remainder of Unit II, the third contains solutions for Unit III, the fourth contains solutions for Units IV and V, and the fifth contains solutions for Unit VI. No solutions are provided for the exercises listed for the appendices.

`stereopic1.pdf`
`stereopic1.ps`

A picture showing the relationship between stereographic projections and inversions with respect to spheres.

`stereopic2.pdf`
`stereopic2.ps`

More pictures involving stereographic projections.

`swisscheese.pdf`
`swisscheese.ps`

A picture illustrating how a square with two holes can be viewed as a compactification of the Euclidean plane.