Geometry of Compact Lie groups.

Let G be a compact connected Lie group and $T \subset G$ a maximal torus in G. Recall that this means that T is a connected closed Abelian Lie subgroup of G which is not contained in any other connected Abelian Lie group. A prime example of (T, G) is the subgroup of all diagonal unitary matrices T within the unitary group G.

In lecture on Mar 9, 2006 we sketched the geometric proof of:

Theorem 1 (Maximal torus theorem). *Any $x \in G$ is conjugate to some element of T.*

The proof was based on the existence of a bi-invariant Riemannian metric on G. (If $\mathfrak{g}$ is simple, this metric is defined by a multiple of the Killing form.) We had proved existence of such a metric by averaging in an earlier lecture.

We will write the proof of this theorem here. It is based on the following **basic facts** regarding any bi-invariant metric on G

- BF1 The one-parameter subgroups of G are precisely the geodesics through the origin.
- BF2 The closed subgroups of G are totally geodesic

and the following basic facts valid for any Riemannian geometry (M, d^2s).

- BF3 If A, B are closed compact submanifolds of M then the infimum of the distances $d(a, b), a \in A, b \in B$ is realized by some geodesic. This geodesic is orthogonal to both A and B.
- BF4 if g is an isometry of (M, d^2s) then g maps geodesics to geodesics, and preserves orthogonality between vectors.

Some review of Riemannian geometry.

A Riemannian metric on a manifold M is a smoothly varying family of inner products $\langle,\cdot,\cdot\rangle_m$, $m \in M$ on the tangent spaces T_mM of M. Other notation for a Riemannian metric is d^2s, or the coordinate . coordinate notation is $\Sigma g_{ij}dx^i dx^j$. Then $g_{ij}(x) = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\rangle_x$ where the $\frac{\partial}{\partial x^i}$ are the coordinate vector fields for coordinates x^i on M.

The *length* of a vector $v \in T_mM$ is defined to be $|v|_m = \sqrt{\langle v, v\rangle_m}$. The*length* of a curve $c : [a, b] \to M$ is defined to be $\ell(c) := \int_a^b |\dot{c}(t)|_{c(t)}dt$. This length does not depend on the parameterization of the curve c. The *distance* $d(x, y)$ between two points of M is defined to be the infimum of the lengths of all curves which join the two points:

$$d(x, y) = \inf\{\ell(\gamma) : \gamma \text{ a smooth curve from } x \text{ to } y\}.$$

A curve c joining x to y is called a *minimizing geodesic* between x and y if it realizes this infimum, i.e if $\ell(c) = d(x, y)$. Minimizing geodesics satisfy the geodesic equations, which are 2nd order (nonlinear) ODEs on M usually written:

$$\nabla_{\dot{\gamma}}\dot{\gamma} = 0$$

The object ∇ is the the “Levi-Civita connection” on M, and is a special type of “affine connection” which we now describe.

Connections. On vector bundles. On tangent bundles. The Levi-Civita connection.

The goal of this section is to understand and prove:

THEOREM 1. *The Levi-Civita connection for the bi-invariant metric on G is:* $\nabla_X Y = \frac{1}{2}[X, Y]$ *where X, Y are left-invariant vector fields.*

We begin with vector bundles and connections thereon.

C1. **Vector bundles.** A *smooth vector bundle* E over M is a smoothly varying family of vector spaces E_m to each $m \in M$. $E = \bigcup_{m\in M} E_m$ the total space of the bundle. M is called the base space. Any E_m is called the fiber. If M is connected then $dim(E_m)$ is constant and is called the "rank" of E. We write elements of E as (m, v) with $v \in E_m$. The projection $(m, v) \to m$ is called the vector bundle projection and is a smooth submersion. The tangent bundle of M is one of the main examples of a vector bundle. A *section* of E is a map $s : M \to E$ with $s(m) \in E_m$. In other words, a section is a smooth selection of vector $s(m) \in E_m$. If the domain of the section is an open subset we call s a *local section*. Write $\Gamma(E)$ for the space of all smooth sections. The product of a smooth section s by a smooth function f yields another smooth section fs where $fs(m) = f(m)s(m)$. . So $\Gamma(E)$ is a module over the ring $C^\infty(M)$.

Algebraic aside Vector bundles are to manifolds as projective modules are to rings. The space of sections forms a *projective module* over the ring of smooth functions on the base space. The theory of vector bundles over manifolds can be absracted to projective modules over rings. This point of view is essential to algebraic K-theory and to noncommutative geometry, and is important in algebraic geometry.

C2. **Connections**

DEFINITION 1. *Let E be a vector bundle over M. A connection on E is a map $\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E)$,which is (i) bilinear over $\mathbb{R}$, (ii) linear over $C^\infty(M)$ in the vector slot $X \in \Gamma(TM)$, and (iii) is a derivation in the section slot, $s \in \Gamma(E)$.*

When $(X, s) \in \Gamma(TM) \times \Gamma(E)$ we write $\nabla_X s$ for $\nabla(X, s)$. Let X_1, X_2 be arbitrary vector fields on M, s_1, s_2 sections of E and f a smooth function on M. Then the conditions of the definition are

$$(i)\ \nabla_{X_1+X_2} s = \nabla_{X_1} s + \nabla_{X_2} s, \nabla_X(s_1 + s_2) = \nabla_X s_1 + \nabla_X s_2,$$
$$(ii)\ \nabla_{fX} s = f\nabla_X s$$

and

$$(iii)\ \nabla_X(fs) = f\nabla_X s + (df(X))s.$$

C3. **Vector bundle-valued one-forms.** These axioms imply that if $Y(m) = X(m)$ then $(\nabla_X s)(m) = (\nabla_Y s)(m)$. It follows that if we fix a local section s then $X \mapsto \nabla_X s$ defines a one-form on M with values in E. We denote this E-valued one-form by ∇s. .

C4. **Connections form an affine space.** The difference of two connections is a one-form on M with values in E. The space of all connections on E forms an affine space whose underlying space of translations is the space of one-forms with values in E. In particular we can average two (or more) connections.

C5. **Local frames.** A local frame $e = \{e_a\}_{a=1}^k = \{e_1, \ldots, e_k\}$ for E is a collection of $k = rank(E)$ smoothly varying sections such that for each x in their domain, $e = \{e_a(x)\}_{a=1}^k$ for a basis for E_x. A local frame induces a local trivialization $U \times \mathbb{R}^k \to E$ by sending $(u, (v^1, \ldots, v^k))$ to $\Sigma v^i e_i(u)$. This *local trivialization* is a diffeomorphism onto some neighborhood $\pi^{-1}(U)$. If f is another local frame defined over another open set V, and if $U \cap V \neq \emptyset$ then on $U \cap V$ we have the change of basis matrix $\eta^a_b(u)$ defined by $f_b(u) = \Sigma\eta^a_b(u)e_a(u)$. The matrix-valued function η is called the "clutching function" or 'transition function" and is the key part of the formal definition of a vector bundle. Returning to the fixed frame e, define the "connection one form" associated to the connection ∇ and local frame s by $\nabla e_a = \Sigma\omega^b_a e_b$ for $a = 1, \ldots, k$. The connection one-form uniquely determines the connection in this domain.

C6. **Parallel translation.** Equations of parallel translation. If $c : I \to M$ is a curve in M then a section over c is a smooth assignment $t \mapsto s(t)$ where $s(t) \in E_{c(t)}$. If c is a smooth curve and s is a smooth section over c then its covariant derivative $\nabla_{\dot{c}} s$ is well-defined and is another section over c. In terms of a local frame, e defined in a neighborhood of c we can write $s(t) = \sigma f^a(t)e_a(c(t))$ and then

$$\nabla_{\dot{c}} s = \Sigma(df^a/dt)s_a + \Sigma f^a\omega^b_a(\dot{c})s_b.$$

The equations $\nabla_{\dot{c}} s = 0$ are, relative to any local trivialization, linear 1st order differential equations. If $s(t)$ is a solution to this equation then we say that s is "parallel" along c, or that $s(t)$ is being "parallel transported" along c. If the solution $s(t)$has $v = s(0)$ then we call $s(1)$ the parallel translate of v along c to E_y. Parallel translation defines an invertible linear map $E_x \to E_y$.

C7. **Metric compatibility.** If E is endowed with additional structure, such as a fiber inner product, or complex structure, or volume, then it is natural to insist that the connection respect these structures. For example, in the case of a fiber inner product, we insist that parallel translation be a linear isometry $E_x \to E_y$. If we write $\langle v, w\rangle_x$ for the inner product of vectors $v, w \in E_x$, then if s_1, s_2 are two sections we can form their fiber inner product: $\langle s_1, s_2\rangle$ to obtain a function on M. The condition that the connection respect the fiber inner product is equivalent to the condition that for all such smooth sections we have:

$$d\langle s_1, s_2\rangle = \langle \nabla s_1, s_2\rangle + \langle s_2, \nabla s_2\rangle. \quad (1)$$

We call such connections "metric compatible".

C8. **Covariant constancy and flatness.** A section s is called *covariantly constant* if $\nabla s = 0$. If every point m of M is contained in the domain of a local frame $e = \{e_a\}$ for with the e_a are covariantly constant is called *flat*.

Affine Connections.

C9. **Torsion.** A connection in the case $E = TM$ is called an affine connection on M. The torsion of an affine connection ∇ is the tensor $T = T^\nabla$ defined by $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$. The affine connection is called *torsion-free* if $T_\nabla = 0$.

C10. **Standard flat case.** The standard affine connection for $M = \mathbb{R}^n$ is defined by $\nabla_X Y = DY \cdot X$. In indices: $(\nabla_X Y)^i = \Sigma_j X^j \frac{\partial Y^i}{\partial x^j}$. this standard connection is torsion free, which is a good way to remember the coordinate formula for the Lie bracket: $[X, Y] = DY \cdot X - DX \cdot Y$. With this connection, any constant vector field v is covariantly constant. It follows that parallel translation for this connection is simply the usual translation of vectors by addition, and the picture of parallel translation along a curve c is a family if parallel vectors of equal length, attached along c. More formally, parallel transport from x to y is independent of the curve c joining x to y and is equal the composition $\theta_y^{-1} \circ \theta_x$ where $\theta_x : T_x\mathbb{R}^n \to \mathbb{R}^n$, $\theta_y : T_y\mathbb{R}^n \to \mathbb{R}^n$ are the canonical identifications, valid for any finite-dimensional real vector space V in place of $\mathbb{R}^n$. Taking the constant vector fields v to form a basis for $\mathbb{R}^n$ and refering back to C8 we see that this connection is flat.

C11. **The Levi-Civita connection.** A Riemannian metric is a fiber inner product on TM. Once a Riemannian metric is chosen on M, we can, referring back to C7, talk about affine connections on M which are metric compatible. The metric compatibility condition (**??**) is equivalent to the validity of the equation:

$$Z[\langle X, Y\rangle] = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle \tag{2}$$

for all smooth vector fields X, Y, Z on M. In the left-hand side of this equation, $Z[\langle X, Y\rangle]$ means the derivative of the scalar function $\langle X, Y\rangle$ in the direction of Z.
C11.

THEOREM 2 (Fund thm of Riem Geom). *Given a Riemannian metric on M there is a unique torsion-free metric compatible affine connection on M. It is called the Levi-Civita connnection, and is henceforth denoted ∇.*

C12. **Geodesic equations.**

THEOREM 3. *If c is a minimizing geodesic on M then c is smooth and satisfies the differential equation $\nabla_{\dot{c}}\dot{c} = 0$*

We call any solution to $\nabla_{\dot{c}}\dot{c}$ a *geodesic*. One can prove that all sufficiently short subarcs of a geodesic are minimizing geodesics.

The group case. (Thank you Alan Weinstein for pointing out that the Levi-Civita connection is the right-left averaged connection desribed below.)

On any Lie group there are two canonical flat affine connections, one associated to left translation and denoted ∇^L. the other associated to right translation and denoted ∇^R. We can define ∇^L by insisting that the left-invariant vector fields are covariantly constant:

$$\nabla^L X^L = 0$$

where X^L is any left invariant vector field. Recall that X^L is the left-invariant extension of $X \in T_eG$, so that $X^L(g) = (dL_g)_e X$ and $X^L(e) = X$. And in the matrix case $X^L(g) = gX$.

Parallel translation from h to gh for the connection ∇^L along any curve is left translation by g: $v \mapsto (dL_g)_h v$. The torsion of ∇^L is nonzero and is equalt to $T^L(X^L, Y^L) = [X, Y]^L$.

Similarly, $\nabla^R X^R = 0$, with X^R the right-invariant extension of $X \in T_eG$. In the matrix case $X^R(g) = Xg$. Parallel translation for ∇^R equals right translation. The torsion of ∇^R equals $T^R(X^R, Y^R) = -[X, Y]^R$.

It follows from C4 (see the last sentence there) that we can average the right and left connections to obtain a new affine connection:

$$\nabla^{avg} = \frac{1}{2}\nabla^L + \frac{1}{2}\nabla^R$$

Clearly this averaged connection is torsion free at e, since at e we have $X^L(e) = X^R(e)$.

LEMMA 1.

$$\nabla^{avg}_{X^L} Y^L = \frac{1}{2}[X, Y]^L.$$

Proof of lemma. Matrix case. Since $\nabla^L Y^L = 0$ it suffices to show that $\nabla^R_{X^L} Y^L = [X,Y]^L$. Now in the matrix case we have $X^L(g) = gX = (gXg^{-1})g = (Ad_g X)^R(g)$. It follows from this fact and the fact that for any right invariant vector field Z^R we have $\nabla^R Z^R = 0$ that $\nabla^R X^L = (Ad(dg)X)^R(g))$. Consequently, $\nabla^R_{X^L} Y^L = (Ad(dg(X^L))Y)^R(g)$. Now, again in the matrix case we have $Ad(g)Y = gYg^{-1}$ so that when we differentiate with respect to g we get: $Ad(dg(X^L))Y = (d(Ad(g))(X))(Y) = g(XY - YX)g^{-1}$ from which it follows that $(Ad(dg(X^L))Y)^R(g) = g(XY - YX) = [X,Y]^L(g)$.

Corollary 1. *∇^{avg} is torsion free everywhere.*

Proof. Its torsion on a pair X^L, Y^L of left-invariant vector fields is given by $T^{avg}(X^L, Y^L) = \frac{1}{2}[X,Y]^L - \frac{1}{2}[Y,X]^L - [X,Y]^l = 0$. Now use the fact that the left-invariant vector fields frame G. QED

Restatement of theorem 1: The Levi-Civita connection for a bi-invariant metric on a Lie group G is equal to its averaged connection ∇^{avg} from the lemma.

Proof. We have seen ∇^{avg} is torsion-free. By the fundamental theorem of Riemannian geometry, it remains to prove that ∇^{avg} is metric-compatible. Recall that bi-invariance of the Riemannian metric on G is equivalent to the Ad-invariance of its restriction $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_e$ to the identity e; namely to the equality : $\langle Ad_g X, Ad_g Y \rangle = \langle X, Y \rangle$. valid for all $g \in G$ and for all $X, Y \in \mathfrak{g}$. Differentiate this condition in g, with $\dot{g} = Z$, $g(0) = e$ to obtain the equality $\langle [Z,X], Y \rangle + \langle X, [Z,Y] \rangle = 0$ valid for all $X, Y, Z \in \mathfrak{g}$. By left-invariance of the metric then, $\langle [Z,X]^L, Y^L \rangle + \langle X^L, [Z,Y]^L \rangle = 0$ as a function on G. But by the lemma this last expression is equal to

$$2\{\langle \nabla^{avg}_{Z^L} X^L, Y^L \rangle + \langle X^L, \nabla^{avg}_{Z^L} Y^L \rangle\} = 0.$$

On the other hand, $\langle X^L, Y^L \rangle = \langle X, Y \rangle(e)$ is a constant function on the Lie group so its derivative: $Z^L[\langle X^L, Y^L \rangle] = 0$. It follows that both sides of the metric compatibility equation (**??**) are zero. QED

Proof of BF1: one-parameter subgroups are geodesics.

A one-parameter subgroup $c(t) = exp(tX)$ is the integral curve through e of the left invariant vector fields X^L. Thus $\dot{c}(t) = X^L(c(t))$ and so $\nabla_{\dot{c}}\dot{c} = \nabla_{X^L} X^L$. But $\nabla_{X^L} X^L = \frac{1}{2}[X,X]^L = 0$. QED.

Proof of BF2: Lie subgroups are totally geodesic.

Because multiplication by g is an isometry, the geodesics through an arbitrary g all can be expressed in the form $c(t) = gexp(tX)$. (They can also be written $exp(tX)g$, with a different X.). Let H be a LIe subgroup of G and $g \in G$. Then $\dot{c}(0 \in T_g H$ if and only if $X \in Lie(H)$. In this case $exp(tX) \in H$ and so $c(t) \in H$. QED

Remark. If H is a torus, then $[X,Y] = 0$ for all $X, Y \in Lie(H)$. This, and the formula for the Levi-Civita connection shows that the restriction of the Levi-Civita connection to H is flat for H a torus. And so the torii are geometrically 'flat"− meaning have locally Euclidean Riemannian geometry.

Proof of the Maximal torus theorem.

We consider the action of G on itself by conjugation: $g \cdot h = ghg^{-1}$. (We called this $AD(g)(h)$ earlier on in the class.) This action is by isometries since both right and left translation by g are isometries for a bi-invariant metric. For $x \in G$ we set

$$G \cdot x = \{gxg^{-1} : g \in G\}$$

which is the conjugacy class containing x. Fix a maximal torus T, and let $x \in G$ be arbitrary. The theorem is the assertion that

$$(G \cdot x) \cap T \neq \emptyset. \tag{3}$$

Call an element of T 'singular' if it is of the form $exp(2\pi X)$ where $X \in Lie(T)$ and $\theta(X) \in \mathbb{Z}$ for some root θ. Otherwise call T "generic". (The collection X such that $\theta(X) \in \mathbb{Z}$ for some root is called 'the infinitesimal diagram' by Adams in ch. 5 "The geometry of the Stiefel diagram'). We will prove below that

$$(G \cdot t)^{\perp} = T_t(T) \tag{4}$$

Side Remark.

PROPOSITION 1. *If $t \in T$ then the following are equivalent conditions. (i) t is singular. (ii) t is a singular value of the exponential map. (iii) t lies in more than one maximal torus. (iii) The orbit of t under the action of the Weyl group $W = N(T)/T$ has less that $|W|$ elements in it (iv) the collection $\{t^n : n \in \mathbb{Z}\}$ is NOT dense in T. (v) t^{-1} is singular.*

Nonemptiness (**??**) folllows from (**??**), and BF1-BF3. Indeed,take $A = G \cdot t$, $B = G \cdot x$ in BF3 with t generic. Let $c(s), 0 \leq s \leq 1$ be a minimizing geodesic whose length realizes the distance between A and B. As per BF3, we have $c(0) \in A$, $c(1) \in B$. and $\dot{c}(0)$ is perpindicular to A. We can write $c(0) = g \cdot t$ for some $g \in G$. Since conjugation acts by isometries, the curve $\gamma(s) = g^{-1} \cdot c(s)$ is a geodesic having the same length as c, which starts at t, ends at some point $g^{-1} \cdot c(1) \in B$ and is orthogonal to $G \cdot t$ at t. By **??** we have $\dot{\gamma}(0) \in T_t T$. But γ is a geodesic and by BF2, the torus T is totally geodesic, so that all points of the curve $\gamma(s)$ lie in T. In particular its other endpoint $\gamma(1) \in B \cap T$, establishing the needed nonemptiness (**??**).

QED

It remains to verify (**??**). We have $g \cdot t = gtg^{-1}$. Write $g = exp(tZ)$ and differentiate to get that the infinitesimal generator $Z_G(t)$ of the AD action in the direction Z at $t \in T$ is given by $Z_G(t) = Zt - tZ$ in the matrix case, or $dL_t(Z) - dR_t(Z)$ in the abstract case. The set of all such $Z_G(t)$ forms the tangent space to $G \cdot t$ at t. We are to show then, that the collection of vectors of the form $Z_G(t)$ as Z varies over $\mathfrak{g}$ is the orthogonal complement to T at t. We right translate the situation back to the origin, using that right translation is an isometry. and that $dL_t dR_{t^{-1}} = Ad_t$. We have $dR_{t^{-1}} Z_G(t) = Ad_t - Id$, while $dR_{t^{-1}}(T_t T) = T_e T := \mathfrak{g}_0$ where we use $\mathfrak{g}_0$ to denote the Lie algebra of T¿ Thus, we must show that

$$im(Ad_t - Id)^{\perp} = Lie(T).$$

In order to prove this we use root space decomposition of $\mathfrak{g}$:

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus V_\theta \, ; \, \mathfrak{g}_0 = Lie(T)$$

According to the definition, Ad_t acts on the V_θ by rotation by $\theta(t)$ radians. To say that t is generic is to say that these rotations are NOT the identity for any nonzero root θ. Consequently, the restriction of $Ad_t - Id$ to any nonzero root space V_θ is not zero for t generic. On the other hand, $Ad_t = Id$ on $\mathfrak{g}_0 = LIe(T)$ for any $t \in T$, generic or not. We have proved:

LEMMA 2. *$t \in T$ is generic if and only if $ker(Ad_t - Id) = \mathfrak{g}_0 := Lie(T)$.*

Now, for any map S of an inner product space to itself we have $im(S)^{\perp} = ker(S^*)$ where S^* is the transpose of S. And $Ad_t^* = (Ad_t)^{-1} = Ad_{t^{-1}}$ by the bi-invariance of the metric, so that for $S = Ad_t - Id$ we have $S^* = Ad_{t^{-1}} - Id$. Finally, t is generic if and only if t^{-1} is generic. (See 'side remark, (v) of the proposition there.) It follows that

$$im(Ad_t - Id)^{\perp} = ker(Ad_{t^{-1}} - Id) = \mathfrak{g}_0.$$

QED

Aside: the same argument applies to any $t = exp(2\pi X) \in T$ in order yield that at t we have

$$(G \cdot t)^{\perp} = dR_t(ker(Ad_t - Id)) = dR_t(\mathfrak{g}_0 \oplus \bigoplus_{\theta \in S(t)} V_{\theta}),$$

where $S(t)$ consists of those roots θ such that $\theta(X) \in \mathbb{Z}$ – i.e $S(t)$ labels the roots whose corresponding root affine hyperplanes contain X.

Beginning of Weyl integration formula. For $t = exp(2\pi X)$ and θ a root we have, relative to an oriented orthonormal frame for V_{θ}:

$$(Ad_t - Id)|_{V_{\theta}} = \begin{pmatrix} \cos(2\pi\theta(X)) - 1 & -\sin(2\pi(\theta(X)) \\ \sin(\theta(X)) & (\cos(2\pi\theta(X)) - 1) \end{pmatrix}.$$

.