Differential Equations and Paths in Lie Groups: The Product Integral.

In what follows, we let G be a Lie group of matrices and L its Lie algebra. We have been careful to consider tangent matrices at the identity I of G. These form the Lie algebra L. What of tangent matrices at any point of G?

Theorem: Let $P{:}\mathbf{R} \to G$. Then $P^{-1}\dfrac{dP}{dt} \in L$. We say that the tangent matrix dP/dt at the point $P(t)$ is transported back to I.

Proof: Let $t_0 \in \mathbf{R}$. Consider the function $Q(t) = P(t_0)^{-1}P(t_0 + t)$. Then $Q(0) = I$, and so $A = \left.\dfrac{dQ}{dt}\right|_{t=0} \in L$. But by calculus, $Q'(t) = P(t_0)^{-1}P'(t_0 + t)$. Thus, $A = Q'(0) = P(t_0)^{-1}P'(t_0) = P(t)^{-1}P'(t)|_{t=t_0}$. Since t_0 is arbitrary, we have the result.

Remark. We write $DP = P^{-1}\dfrac{dP}{dt}$. The same argument shows that $P'P^{-1}$ is also in L. In this case, we say that P' is right transported to I in contrast to the previous left transport. Note that $P'P^{-1} = P(DP)P^{-1}$. Note that that if $P = e^{tA}$, then $P^{-1}P' = P'P^{-1} = A$. So we can think of the exponential curve $P = e^{tA}$ as a curve with constant tangent vector A.

The operator D is similar to a logarithmic derivative. In calculus 1, $f^{-1}f' = f'/f = d(\log f)/dt$. In fact, this is precisely Df when $G = GL_1$. Thus DP generalizes the logarithmic derivative, though it is not in general a derivative of a logarithm.

The following rules are easily proved:

1. $D(cf) = Df$ (where c is constant).
2. $D(g^{-1}) = -g'g^{-1}$ (the negative of the right transport of g).

We are now going to generalize the concept of a one parameter group for some of the classical groups considered before. First, some heuristics.

In a discrete group D, if $g \in D$, the powers g^n of g constitute a subgroup of D, called a cyclic subgroup. For $n \geq 0$, The sequence $h(n) = g^n$ can be defined recursively by:
1. $h(0) = I$ (the identity of the group).
2. $h(n+1) = h(n) \cdot g$ (the recursion).

For a continuous group, G, we take as generator an infinitesimal element $I + Adt$, where A is in the Lie algebra L of G. The powers $h(n)$ will be replaced by the "powers" $H(t)$ and the recursion is replaced by a differential equation as follows:

1. $H(0) = I$ (the initial condition).

$H + dH = H(I + Adt)$, so $H + dH = H + HAdt$, or

2. $dH/dt = HA$ (the differential equation).

We already know that the solution is $H = e^{tA}$. The cyclic group generated by the element g is replaced by the one parameter group generated by the infinitesimal element $A \in L$.

The laws of exponents apply. For example, $g^n g^m = g^{n+m}$, and analogously $e^{tA}e^{sA} = e^{(t+s)A}$. An alternate recursion for powers is $g^{n+1} = g \cdot g^n$. The analogous differential equation is $dH/dt = AH$. In each case, the solution is the same as before.

More generaly, we can define a product $p(n) = \prod_{k=1}^{n} g_n$ of different elements $g_1, g_2, \ldots, g_n$ recursively as follows:

1. $p(0) = I$ (the identity of the group.)
2. $p(n+1) = p(n) \cdot g_{n+1}$ (the recursion.)

For the continuous analog, we replace the sequence g_n by a infinitesimal elements $I + A(t)dt$. where $A(t) \in L$, and $A(t)$ is continuous. We then have the differential equation:

1. $P(0) = I$ (the initial condition.)

$P + dP = P(I + A(t))dt$, so $P + dP = P + PA(t)dt$, or

2. $dP/dt = PA(t)$ (the differential equation.)

(For the analog of the product $g_n g_{n-1} \cdots g_2 g_1$, we would replace (2), by the differential equation $dP/dt = A(t)P$.)

The theory of *product integrals* carries the analogy further. If we break the interval $[0, t]$ into equal pieces using the subdivision

$$0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t, \text{ where } t_{k+1} - t_k = \Delta t = t/n$$

and consider the product $P_n(t) = \prod_{k=1}^{n} (I + A(t_k))\Delta t$, then in analogy with Riemann sums we let $n \to \infty$. It turns out that $P_n(t) \to P(t)$, and it is a solution to the differential equation $dP/dt = PA(t)$. The limit of this product is written $\prod_0^t (1 + A(s)ds)$ or $\prod_0^t e^{A(s)ds}$ obtained by observing that $e^{A(s)ds} = 1 + A(s)ds$ to first order. We do not investigate this further in this course.

Given the differential equation $dP/dt = PA(t)$ with initial condition $P(0) = I$, we expect the solution to be invertible. (It is after all the "product" of infinitesimal invertible elements $I + A(t)dt$.) We now prove this result.

Theorem: Let $P(t)$ be the solution of the differential equation $dP/dt = PA(t)$ with initial condition $P(0) = I$. Then $P(t)$ is invertible.

Proof: Let $Q(t)$ be the solution of $dQ/dt = -A(t)Q$ with initial condition $Q(0) = I$. We claim that $PQ = I$. To see this, compute

$$(PQ)' = PQ' + P'Q = -PAQ + PAQ = 0$$

Thus PQ is constant. Since its value at $t = 0$ is I, we have the result.

(This guess for Q was based on our heuristic approach. The inverse of a product is the product of the inverse in reverse order. The inverse of $I + Adt$ is $I - Adt$ to first order, and the reverse order recursion will have $Q' = -AQ$ rather than $Q' = -QA$. One could also guess how to find the inverse Q, by differentiating $PQ = I$ and seeing what happens.)

Thus, the equation $dP/dt = PA(t)$ can also be rewritten $DP = P^{-1}dP/dt = A(t)$, we see that $P(t)$ is an anti-D-derivative of $A(t)$ with initial value $P(0) = I$.

We are now going to extend the result on exponentials for the various classical groups. We first consider the relationship between the determinant and the trace, in order to generalize the result $\det(e^A) = e^{\text{trace } A}$, given in the notes on exponentials.

We start with this calculus result on taking the derivative of a determinant.

Lemma. Let $A(t)$ be an $n \times n$ differentiable matrix, with columns $a_1(t), \ldots, a_n(t)$. Then

$$\frac{d}{dt} \det A(t) = \sum_{i=1}^{n} \det A_i(t)$$

where $A_i(t)$ is the matrix $A(t)$ with the i-th column replaced by $\dfrac{da_i}{dt}$.

Proof: We have $A = (a_1 \ \cdots \ a_n)$. We know that $\det(A)$ is a multilinear function of $a_1, \ldots, a_n$. We can compute the derivative of A as follows. Let $\Delta A = A(t+h) - A(t)$ with a similar notation for each of the a_i. Then

$$\Delta \det(A) = \det(a_1 + \Delta a_1, \cdots, a_n + \Delta a_n) - \det(a_1, \cdots, a_n)$$

Using multi-linearity, this expression becomes

$$\det(\Delta a_1, a_2, \ldots, a_n) + \det(a_1, \Delta a_2, \ldots, a_n) + \cdots + \det(a_1, a_2, \ldots, \Delta a_n) + R$$

where R consist of similar expressions with two or more Δa_i's as arguments. Now divide by h and let $h \to 0$. We get

$$(\det(A))' = \det(a_1', a_2, \ldots, a_n) + \det(a_1, a_2', \ldots, a_n) + \cdots + \det(a_1, a_2, \ldots, a_n')$$

This is the result. Note: $R/h \to 0$. For example, taking a typical term of R with two Δa_i's as arguments, we have

$$\det(\Delta a_1, \Delta a_2, \ldots, a_n)/h = \det(\Delta a_1/h, \Delta a_2, \ldots, a_n) \to \det(a_1', 0, \ldots, a_n) = 0$$

We can now give an infinitesimal version of the determinant.

Theorem 1. Let $P(t) \in GL_n$ where P is continously differentiable and $P(t_0) = I$. Let $P'(t_0) = A$. Then

$$\frac{d}{dt}\det(P)|_{t=t_0} = \text{trace} P'(t_0). \tag{1}$$

Proof: Let $P = (p_1 \ \dots \ p_n)$ where the p_i are the columns of P. By hypothesis, $p_i(t_0) = e_i$, the standard basis vector in $\mathbf{R}^n$. By the previous result,

$$\frac{d}{dt}\det(P) = \det(p_1', p_2, \dots, p_n) + \dots + \det(p_1, p_2, \dots, p_n').$$

Therfore

$$\begin{aligned} \frac{d}{dt}\det(P)\Big|_{t=t_0} &= \det(p_1'(t_0), e_2, \dots, e_n) + \dots + \det(e_1, e_2, \dots, p_n'(t_0)) \\ &= \det(a_1, e_2, \dots, e_n) + \dots + \det(e_1, e_2, \dots, a_n) \end{aligned}$$

where a_i is the i-th column of $A = P'(t_0)$. But the i-th determinant in this sum can easily be expanded along the i-th row to obtain $a_{11} + \dots + a_{nn} = \text{trace}(A)$. For example,

$$\det(a_1, e_2, \dots, e_n) = \det \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \dots & 1 \end{pmatrix} = a_{11}$$

This proves the result.

With the help of this result, we can prove the following theorem.

Theorem 2. Let $P(t) \in GL_n$ be continously differentiable, with $P(0) = I$. Let $a(t) = \text{trace}(DP) = \text{trace}(P^{-1}P')$. Then

$$\det P(t) = e^{\int_0^t a(s)ds} = e^{\int_0^t \text{trace}[P(s)^{-1}P'(s)]ds} \tag{2}$$

Proof: For any real t_0, the function $P(t_0)^{-1}P(t)$ satisfies the hypothesis of Theorem 1. Taking derivatives at $t = t_0$, we have

$$\frac{d}{dt}\det[P(t_0)^{-1}P(t)]_{t=t_0} = a(t_0)$$

Thus

$$\frac{d}{dt}\det[P(t_0)^{-1}][\det P(t)]_{t=t_0} = a(t_0)$$

$$\det P(t_0)^{-1}\frac{d}{dt}\det P(t)|_{t=t_0} = a(t_0)$$

Since this t_0 is arbitrary, we may replace it by t. This equation is a *scalar equation.* Letting $g(t) = \det P(t)$, this equation becomes

$$g(t)^{-1}g'(t) = a(t)$$

or

$$\frac{d}{dt}\log g = g^{-1}g' = a(t)$$

So $\log g = \int_0^t a(s)ds$. (The constant of integration is 0, since $g(0) = 1$.) Finally we have the required result:

$$g(t) = \det P(t) = e^{\int_0^t a(s)ds} = e^{\int_0^t \mathrm{trace}[P(s)^{-1}P'(s)]ds} = e^{\int_0^t \mathrm{trace}(DP(s))ds}$$

As a simple corollary, we can eliminate the condition $P(0) = I$. If $P(t) \in GL_n$, we take $Q(t) = P(0)^{-1}P(t)$. Note that $DQ = DP$, so we have

$$\det(P(0)^{-1}P(t)) = \det Q(t) = e^{\int_0^t \mathrm{trace}(DP(s))ds}$$

. Therefore, in general, we have

$$\det P(t) = \det P(0)e^{\int_0^t \mathrm{trace}(DP(s))ds}. \tag{3}$$

Curves in SL_n. We have already shown that if $\mathrm{trace}(A) = 0$ then $e^{tA} \in SL_n$ for all t, and conversely. We now go one better and consider arbitrary curves in SL_n.

Theorem 3. If $P(t)$ is a continuously differentiable curve in SL_n, then $\mathrm{trace}(DP) = 0$ for all t. Conversely, if $P(t)$ is a continuously differentiable curve with $P(0) = I$ and $\mathrm{trace}(DP) = 0$ for all t, then in $P(t) \in SL_n$.

Proof: First, assume $P(t) \in SL_n$, so $\det(P(t) = 1$. Using equation (3), this gives $\int_0^t \mathrm{trace}(DP(s))ds = 0$ for all t. Thus, $\mathrm{trace}(DP(t)) = 0$ for all t. The converse is equally clear, since by equation (3), if $\mathrm{trace}DP(t) = 0$ for all t, $\det(P) = $ constant.

This second half of this theorem can be stated in terms of differential equations: If $P(t)$ is an $n \times n$ matrix satisfying the equation $dP/dt = PA(t)$ with $P(0) = I$, and if trace $A(t) = 0$, then $\det\ P(t) = 1$.

As noted above, we have similar results for the equation $dP/dt = AP$. In this case, we use the operator D_r defined by $D_rP = P'(t)P^{-1}$.

Observations on Linear Differential Equations. We consider the system of homogeneous linear differential equations $\frac{dx}{dt} = A(t)x$ with initial condition $x = x_0$. Here x is a column vector. If we solve the matrix differential equation $\frac{dP}{dt} = A(t)P$ with $P(0) = I$, the solution to the vector equation may easily be verified to be $x = P(t)x_0$. For the initial condition is clearly true, and clearly,

$$\frac{dx}{dt} = \frac{d}{dt}(P(t)x_0) = \frac{dP}{dt}x_0 = A(t)Px_0 = A(t)x$$

A way of thinking of the matrix differential equation is that it solves the equation $P'(t) = A(t)P(t)$ for the n linearly independent initial conditions $x_0 = e_i$ for $i = 1$ to n. And any initial condition x_0 is a linear combination of these standard basis, so the corresponding solution is a linear combination of the special solutions. If the equation is $\frac{dx}{dt} = xA(t)$ where x is a row vector, the same argument will use the matrix differential equation $\frac{dP}{dt} = PA(t)$ with initial condition $P(0) = I$. The solution of the vector equation with initial condition $x = x_0$ will then be $x = x_0 P(t)$. Theorem 3 can then be interpreted as follows: If $\text{trace} A(t) = 0$, then for any independent vectors $x_1, \ldots, x_n$, and corresponding solutions $x_i(t)$, the volume $V(x_1, \ldots, x_n)$ is constant.

The Orthogonal Group. We now calculate the Lie Algebra of the Orthogonal group O_n. The result is given in the following theorem. In view of our work with one parameter groups, the result is no surprise.

Theorem 4. The Lie Algebra of O_n consists of all skew symmetric $(n \times n)$ matrices A $(A^t = -A.)$ Let $P(t)$ be a continuously differentiable curve in O_n. Then $DP = P^{-1}P'$ is skew symmetric. Conversely, if DP is skew symmetric and $P(0) = I$, then $P(t) \in O_n$.
Proof. Let $P(t) \in O_n$, with $P(0) = I$. Let $A = P'(0)$. We show that $A + A^t = 0$. To see this, use the definition of O_n, namely $PP^t = I$. Now differentiate to get $PP'^t + P'P^t = 0$. Putting $t = 0$, since $P(0) = P^t(0) = I$, we have $A^t + A = 0$. Conversely, suppose $A(t) + A(t)^t = 0$, and $DP(t) = A(t)$ with $P(0) = I$. Since $DP = P^{-1}P'$, we have $P'(t) = PA(t)$. We compute PP^t by finding its derivative:

$$(PP^t)' = P(P^t)' + P'P^t = P(PA)^t + PAP^t = PA^tP^t + PAP^t = P(A^t + A)P = 0$$

Therefore $PP^t = \text{constant}$, so $PP^t = P(0)P^t(0) = I$. Thus $P^{-1} = P^t$ and the result is proved. (We have used the obvious identity $(P^t)' = (P')^t$.) Note that for fixed skew symmetric matrix A, we get $P' = PA$ with $P(0) = I$ which leads to the solution $P(t) = e^{tA}$ considered in the notes on the exponential function. In particular, any skew symmetric matrix is in the Lie algebra of O_n.