SYSTEMS OF DIFFERENTIAL EQUATIONS, EULER'S FORMULA

1. Uniqueness for solutions of differential equations.

We consider the system of differential equations given by

$$\frac{d}{dt}\vec{x} = \vec{v}(\vec{x}) \ , \tag{1}$$

with a given initial condition $\vec{x}(0) = \vec{x}_0$. Here $\vec{x} \in \mathbb{R}^n$ and $\vec{v}$ is a function that maps $\mathbb{R}^n$ into $\mathbb{R}^n$. We shall assume that for any two vectors $\vec{x}_1, \vec{x}_2$ we

$$\|\vec{v}(\vec{x}_1) - \vec{(}\vec{x}_2)\| \le L\|\vec{x}_1 - \vec{x}_2\|$$

where L is some constant, usually called the Lipschitz constant. An example is

$$\vec{(}\vec{x}) = A\vec{x}$$

where A is a constant real $n \times n$ matrix. IWe compute

$$\|A\vec{x}_1 - A\vec{x}_2\|^2 = \|A(\vec{x}_1 - \vec{x}_2)\|^2 = (\vec{x}_1 - \vec{x}_2) \cdot A^T A(\vec{x}_1 - \vec{x}_2) \le \lambda\|(\vec{x}_1 - \vec{x}_2)\|^2$$

where λ is the largest eigenvalue of $A^T A$.

The following is relatively easy to prove.

Theorem 1.1. *The differential equation* (1) *has at most one solution that satisfies the given initial condition.*

Proof. Suppose there are two solutions $\vec{x}_1(t)$ and $\vec{x}_2(t)$ both satisfying $\vec{x}_1(0) = \vec{x}_2(0) = \vec{x}_0$. Integrating we see that both solutions satisfy the equation

$$\vec{x}_i(t) = \vec{x}_0 + \int_0^t \vec{v}(\vec{x}_i(\tau))d\tau \ , i = 1, 2 \ .$$

Hence, noting that the initial condition drops out, we get

$$\|\vec{x}_1(t) - \vec{x}_2(t)\| = \|\int_0^t \vec{v}(\vec{x}_1(\tau))d\tau - \int_0^t \vec{v}(\vec{x}_2(\tau))d\tau\| = \|\int_0^t [\vec{v}(\vec{x}_1(\tau)) - \vec{v}(\vec{x}_2(\tau))]d\tau\|$$

Using the Minkowski inequality which is essentially the triangle inequality we get

$$\|\vec{x}_1(t) - \vec{x}_2(t)\| \le \int_0^t \|\vec{v}(\vec{x}_1(\tau)) - \vec{v}(\vec{x}_2(\tau))\|d\tau$$

and using the Lipschitz condition

$$\|\vec{x}_1(t) - \vec{x}_2(t)\| \le L\int_0^t \|\vec{x}_1(\tau)) - \vec{x}_2(\tau)\|d\tau \ .$$

and this holds for all t as long as the solutions exist. If $t < T$ we have that

$$\|\vec{x}_1(t) - \vec{x}_2(t)\| \le L\int_0^t \|\vec{x}_1(\tau)) - \vec{x}_2(\tau)\|d\tau \le L\int_0^T \|\vec{x}_1(\tau)) - \vec{x}_2(\tau)\|d\tau$$

This inequality implies that for all $t \le T$ that

$$\|\vec{x}_1(t) - \vec{x}_2(t)\| \le LTM(T)$$

where we set $M(T) = \max_{[0,T]} \|\vec{x}_1(t) - \vec{x}_2(t)\|$. Hence we also have that

$$M(T) \le LTM(T)$$

and if we choose T such that $LT < 1$ it follows that $M(T) = 0$. Hence the two solution coincide on the time interval $[0, T]$. Choosing $\vec{x}(T)$ as the new initial condition the solution must coincide on the interval $[T, 2T]$ also and so on. We can argue the same way that for negative times the solutions have to coincide. □

2. Some remarks about the e^{At}

Recall that we defined the exponential of a matrix e^{At} by

$$e^{At} = \sum_{n=0}^{\infty} \frac{A^n t^n}{n!} \ .$$

Here are some facts

Theorem 2.1. *We have*

$$e^{At} e^{As} = e^{A(t+s)}$$

for all $s, t \in \mathbb{R}$.

Proof. Pick any initial condition $\vec{x}_0$. The function

$$\vec{x}(t) = e^{A(t+s)} \vec{x}_0$$

is a solution of the equation $\dot{\vec{x}} = A\vec{x}$. This follows from

$$\frac{d}{dt} e^{A(t+s)} = A e^{A(t+s)} \ .$$

Further the function $\vec{y}(t) = e^{At} e^{As} \vec{x}_0$ is also a solution of the equation $\dot{\vec{x}} = A\vec{x}$. moreover, for $t = 0$ we have that $\vec{x}(0) = e^{As} \vec{x}_0 = \vec{y}(0)$. By uniqueness $\vec{x}((t) = \vec{y}(t)$ and thus

$$e^{At} e^{As} \vec{x}_0 = e^{A(t+s)} \vec{x}_0$$

for all $\vec{x}_0$. Since $\vec{x}_0$ is arbitrary this proves the theorem. □

An interesting consequence of this theorem is that e^{At} is invertible for all t.

$$e^{At} e^{A(-t)} = e^{A(t-t)} = I \ .$$

3. One parameter families of matrices

We say that a family of $n \times n$ matrices $P(t)$ is a one parameter family if

$$P(0) = I$$

and for all $t, s \in \mathbb{R}$,

$$P(t)P(s) = P(t+s) \ .$$

We shall only consider one parameter families that are differentiable.

A particularly useful idea is to consider one parameter families of rotations $R(\phi)$. These are matrices that satisfy $R(\phi)^T R(\phi) = I$. First we compute the derivative

$$\frac{d}{d\phi} R(\phi) = \lim_{\varepsilon \to 0} \frac{R(\phi + \varepsilon) - R(\phi)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{R(\varepsilon) - I}{\varepsilon} R(\phi) = \Omega R(\phi)$$

where we denote

$$\Omega = \lim_{\varepsilon \to 0} \frac{R(\varepsilon) - I}{\varepsilon} = \frac{d}{d\phi} R(0) \ .$$

The matrix Ω is not arbitrary. Indeed, differentiating

$$\frac{d}{d\phi} I R^T(\phi) R(\phi) = \frac{d}{d\phi} I = 0$$

and bu the product rule

$$\frac{d}{d\phi}\Big|_{\phi=0} I R^T(\phi) R(\phi) = \Omega^T + \Omega$$

and we learn that Ω must be a skew symmetric matrix,

$$\Omega^T = -\Omega \ .$$

So far this worked in arbitrary dimensions. We specialize to three dimension and write the general skew symmetric matrix as

$$\Omega = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

note the interesting fact that

$$\Omega \vec{x} = \vec{\omega} \times \vec{x} \ .$$

We also note that $\Omega \vec{\omega} = 0$. Recall that we have the equation

$$R'(\phi) = \Omega R(\phi)$$

and this allows us to compute $R(\phi)$ explicitly. We shall assume that the vector $\vec{\omega}$ is normalized. We have to compute

$$e^{\Omega \phi} = \sum_{n=0}^{\infty} \frac{\Omega^n \phi^n}{n!}$$

Here are some computations:

$$\Omega^2 = \begin{bmatrix} -\omega_2^2 - \omega_3^2 & \omega_1 \omega_2 & \omega_1 \omega_3 \\ \omega_2 \omega_1 & -\omega_3^2 - \omega_1^2 & \omega_2 \omega_3 \\ \omega_3 \omega_1 & \omega_3 \omega_2 & -\omega_1^2 - \omega_3^2 \end{bmatrix}$$

which can be written as

$$\Omega^2 = -I + \vec{\omega} \vec{\omega}^T \ .$$

Here we use that $\vec{\omega}$ is a unit vector. Thus we can start a little table:

$$\Omega \ , \Omega^2 = -I + \vec{\omega} \vec{\omega}^T \ , \Omega^3 = -\Omega \ , \Omega^4 = -\Omega^2 \ldots$$

Thus it makes sense to split

$$e^{\Omega \phi} = \sum_{m=0}^{\infty} \frac{\Omega^{2m} \phi^{2m}}{(2m)!} + \sum_{m=0}^{\infty} \frac{\Omega^{2m+1} \phi^{2m+1}}{(2m+1)!}$$

into even and odd powers. We have that

$$\Omega^{2m+1} = (-1)^m \Omega$$

and hence the second sum reduces to

$$\sum_{m=0}^{\infty} \frac{\Omega^{2m+1}\phi^{2m+1}}{(2m+1)!} = \Omega \sum_{m=0}^{\infty} \frac{(-1)^m \phi^{2m+1}}{(2m+1)!} = \Omega \sin\phi \ .$$

For the even sum have to be careful noting that for $m = 1, 2, \ldots$

$$\Omega^{2m} = (-1)^m (I - \vec{\omega}\vec{\omega}^T) \ .$$

For $m = 0$ we have the identity which we write

$$I = I - \vec{\omega}\vec{\omega}^T + \vec{\omega}\vec{\omega}^T$$

and get that

$$\sum_{m=0}^{\infty} \frac{\Omega^{2m}\phi^{2m}}{(2m)!} = \vec{\omega}\vec{\omega}^T + (I - \vec{\omega}\vec{\omega}^T) \sum_{m=0}^{\infty} \frac{(-1)^m \phi^m}{(2m)!}$$

which equals

$$\vec{\omega}\vec{\omega}^T + (I - \vec{\omega}\vec{\omega}^T)\cos\phi \ .$$

To summarize, we have shown that

$$e^{\Omega\phi} = \cos\phi I + \vec{\omega}\vec{\omega}^T(1 - \cos\phi) + \Omega\sin\phi$$

Let's note a few things: The vector $\vec{\omega}$ is an eigenvector for this matrix with eigenvalue 1. This is the axis of rotation. Take

$$\vec{\omega} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

i.e, the z axis. Then we get the matrix

$$\begin{bmatrix} \cos\phi & 0 & 0 \\ 0 & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sin\phi = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which is precisely a rotation in the positive direction by an angle ϕ. To summarize:

Theorem 3.1. *The rotation about the $\vec{\omega}$ axis by an angle ϕ is given by*

$$R(\phi) = \cos\phi I + (1 - \cos\phi)\vec{\omega}\vec{\omega}^T + \Omega\sin\phi \ ,$$

in particular

$$R(\phi)\vec{x} = \cos\phi\vec{x} + (1 - \cos\phi)(\vec{\omega} \cdot \vec{x})\vec{\omega} + \sin\phi(\vec{\omega} \times \vec{x}) \ .$$

This is Euler's formula. Because

$$\Omega^2 + I = \vec{\omega}\vec{\omega}^T$$

Euler's formula canals be written in the form

$$R(\phi) = \cos\phi I + (1 - \cos\phi)(\Omega^2 + I) + \Omega\sin\phi = I + (1 - \cos\phi)\Omega^2 + \sin\phi\Omega$$

Note that the angle is any value between 0 and 2π. If $\phi < 0$ we may replace ϕ by $-\phi$ which keeps the sign of the cosine function fixed but changes the sign of the sign function. Thus if, additionally we reverse the direction of $\vec{\omega}$ we get back the same rotation. Needless to say that the rotation by an angle $\phi = 0$ or $\phi = 2\pi$ is the identity. Also note that in terms of $R(\phi)$ we have that

$$\frac{1}{2}[R(\phi) + R(\phi)^T] = \cos\phi I + (1 - \cos\phi)\vec{\omega}\vec{\omega}^T$$

and

$$\frac{1}{2}[R(\phi) - R(\phi)^T] = \Omega \sin \phi$$

4. A purely algebraic derivation of Euler's formula

Our previous result concerns solution of the differential equation $R'(\phi) = \Omega R(\phi)$. Suppose now that you are given an arbitrary rotation M. Can we find ϕ and Ω so that

$$M = I + (1 - \cos \phi)\Omega^2 + \sin \phi \Omega \ ?$$

To be more specific we have the following theorem.

Theorem 4.1. *Let M be a 3×3 rotation. Define*

$$\cos \phi = \frac{\mathrm{Tr} M - 1}{2} \ .$$

and

$$\Omega = \frac{1}{2 \sin \phi}[M - M^T]$$

provided that $\phi \neq 0, \pi, 2\pi$. Then

$$M = M = I + (1 - \cos \phi)\Omega^2 + \sin \phi \Omega \ .$$

For $\phi = 0, 2\pi$ we have that $M = I$ and for $\phi = \pi$

$$M = I + 2\Omega^2 \ ,$$

and hence, Euler's formula holds in these cases as well.

Recall that a 3×3 matrix M is a rotation if it satisfies $M^T M = I$ and $\mathrm{det} M = +1$. We would like to show that there exist a unit vector $\vec{\omega}$ and an angle ϕ, $0 \leq \phi \leq 2\pi$ such that

$$M = \cos \phi I + (1 - \cos \phi)\vec{\omega}\vec{\omega}^T + \Omega \sin \phi \ .$$

As usual

$$\Omega = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \ .$$

We first start with a simple Lemma:

Lemma 4.2. *Let M be a rotation in three space, i.e., $M^T M = I$ and $\mathrm{det} M = +1$. Then the matrix M must have the eigenvalue 1. Moreover, the other two eigenvalues must be of the form $e^{\pm i\phi}$ for some $0 \leq \phi \leq 2\pi$.*

Proof. To see this consider

$$\mathrm{det}(M - I) = \mathrm{det} M^T \mathrm{det}(M - I) = \mathrm{det} M^T(M - I)$$
$$= \mathrm{det}(I - M^T) = \mathrm{det}(I - M)^T = \mathrm{det}(I - M) = -\mathrm{det}(M - I) \ .$$

Hence $\mathrm{det}(M - I) = 0$ and 1 is an eigenvalue. If we denote the other two eigenvalues by λ_1 and λ_2 we must have that $\lambda_1 + \lambda_2 + 1 = \mathrm{Tr} M$ and $\lambda_1 \lambda_2 = 1$ (Why?) Hence

$$\lambda_1 + \lambda_2 = \mathrm{Tr} M - 1 \ , \ \lambda_1 \lambda_2 = 1 \ .$$

The best way to solve these equations is to note that $-3 \leq \mathrm{Tr} M \leq 3$ (Why?) Hence we may define

$$\cos \phi = \frac{\mathrm{Tr} M - 1}{2} \ ,$$

and we have to solve the equations $\lambda_1 + \lambda_2 = 2\cos\phi$, $\lambda_1\lambda_2 = 1$. We easily find that $\lambda_1 = e^{i\phi}$ and $\lambda_2 = e^{-i\phi}$. Thus we have the eigenvalues $e^{i\phi}, e^{-i\phi}, 1$. □

Let us assume that $\phi \neq 0, \pi, 2\pi$. These cases we deal with later. Recall that

$$\cos\phi = \frac{\mathrm{Tr}M - 1}{2} \ ,$$

and define

$$\Omega = \frac{1}{2\sin\phi}[M - M^T]$$

Note that this suggests itself from Euler's formula (Why?). We have to check that

$$M = I + (1 - \cos\phi)\Omega^2 + \sin\phi\Omega =: R$$

Cayley's theorem tells us that

$$(M - I)(M - e^{i\phi}I)(M - e^{-i\phi}I) = 0$$

and developing the products yields

$$M^3 - (1 + 2\cos\phi)M^2 + (1 + 2\cos\phi)M - I = 0 \ .$$

Now

$$I + (1 - \cos\phi)\Omega^2 + \sin\phi\Omega = I + \frac{1 - \cos\phi}{4\sin^2\phi}[M - M^T]^2 + \sin\phi\frac{1}{2\sin\phi}[M - M^T]$$

$$= I + \frac{1}{4(1 + \cos\phi)}[M - M^T]^2 + \frac{1}{2}[M - M^T] \ .$$

We further have that

$$[M - M^T]^2 = M^2 + M^{2T} - 2I$$

and by Cayley's theorem

$$M^2 = (1 + 2\cos\phi)M - (1 + 2\cos\phi)I + M^T \ , \ M^{2T} = (1 + 2\cos\phi)M^T - (1 + 2\cos\phi)I + M$$

so that

$$M^2 + M^{2T} - 2I = 2(1 + \cos\phi)[M + M^T] - 4(1 + \cos\phi)I$$

Thus,

$$R = \frac{1}{2}[M + M^T] + \frac{1}{2}[M - M^T] = M \ .$$

The remaining cases are easily dealt with. Assume that $\phi = 0$ or 2π. Then

$$\mathrm{Tr}M = 3 \ .$$

Now the matrix M is of the form

$$[\vec{u}_1, \vec{u}_2, \vec{u}_3]$$

all of them being unit vectors. The trace, therefore is $u_{11} + u_{22} + u_{33} = 3$ since each of these numbers is between -1 and 1 they all must be equals to 1. This means that the rotation matrix must be the identity matrix. The case $\phi = \pi$ implies that -1 must be a two fold eigenvalue. From this we get three facts: $M^2 = I$ and hence $M = M^T$ and $M + I$ has a two dimensional null space. Set

$$P = \frac{M + I}{2}$$

and note that

$$P^2 = P \ , P^T = P$$

Hence P projects the three dimensional space onto a one dimensional space and therefore it must be of the form

$$P = \vec{\omega}\vec{\omega}^T$$

for some unit vector $\vec{\omega}$. Thus,

$$M = -I + 2\vec{\omega}\vec{\omega}^T = I + 2\Omega^2$$

which is what we wanted to show.