LINEAR SYSTEMS OF DIFFERENTIAL EQUATIONS DONE RIGHT

KYLE ORMSBY

1. Introduction to the problem

Consider a continous function $F : \mathbb{R}^n \to \mathbb{R}^n$. We will think of F as a *vector field*, and can think of $F(x)$ as a velocity vector positioned at $x \in \mathbb{R}^n$. Our goal is to find a path $\gamma : \mathbb{R} \to \mathbb{R}^n$ which starts at a point $p \in \mathbb{R}^n$ (so $\gamma(0) = p$, our *initial condition*) and, for each time $t \in \mathbb{R}$, satisfies the *differential equation*

$$\gamma'(t) = F(\gamma(t)). \tag{1}$$

If $\gamma = (x_1, x_2, \ldots, x_n)$, then we may rewrite (1) as the system of scalar equations

$$\begin{aligned} x_1' &= F_1(x_1, \ldots, x_n) \\ x_2' &= F_2(x_1, \ldots, x_n) \\ &\vdots \\ x_n' &= F_n(x_1, \ldots, x_n). \end{aligned}$$

We may think of this system as the time-derivatives of n dependent variables simultaneously satisfying n position-dependent equations. A solution γ is then a *flow line*, the path of a particle pushed around by the vector field F so that its velocity matches F at any given point.

Example 1. As an example, consider the vector field $F : \mathbb{R}^2 \to \mathbb{R}^2$ taking (x, y) to $F(x, y) = (x, -y)$. If our initial position is (p_1, p_2), then we are seeking $\gamma : \mathbb{R} \to \mathbb{R}^2$ taking t to $\gamma(t) = (x(t), y(t))$ satisfying $x(0) = p_1$, $y(0) = p_2$, $x'(t) = x(t)$, and $y'(t) = -y(t)$.

Figure 1 depicts the vector field F by drawing scaled down versions of the vectors $F(x, y)$ at a sampling of points in the square $[0, 2] \times [0, 2]$. The curve drawn within the figure is a solution satisfying the initial condition $\gamma(0) = (-1, 1)$.

This particular system of differential equations may be solved one variable at a time since x' only depends on x and y' only on y. Indeed, separation of variables quickly gives the solution $x = p_1 e^t$, $y = p_2 e^{-t}$. We may then note that $xy = p_1 p_2$ is a constant, so the curve traced by γ is a hyperbola.

The above example fits into a class of examples of the form

$$\begin{aligned} x' &= a_{11} x + a_{12} y + b_1 \\ y' &= a_{21} x + a_{22} y + b_1 \end{aligned}$$

in which the a_{ij} and b_k are all scalar constants. This may then be rephrased in matrix form as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

This admits the further generalization to n variables

$$x' = Ax + b$$

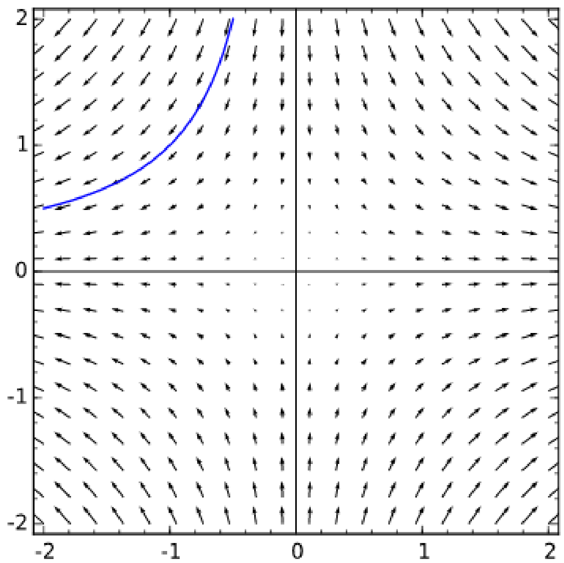

FIGURE 1. The vector field $F(x, y) = (x, -y)$ and a flow line with initial position $(-1, 1)$.

where

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

and x' is the column vector consisting of the time derivatives of $x_1, \ldots, x_n$.

This is a very special sort of differential equation in which the vector field F is an affine transformation. If we specialize yet further and take $b = 0$ (so that F is linear), then we arrive at our current object of study.

Definition 2. A *(homogeneous) linear system of differential equations* is one of the form

$$x' = Ax$$

where $x = (x_1, \ldots, x_n)$ and A is an $n \times n$ matrix.

Remark 3. When $b \neq 0$, the literature calls the system $x' = Ax + b$ an *inhomogeneous* linear system of differential equations. From our point of view, it would be nicer to call this an *affine* system of differential equations, but we shall quail in the face of centuries worth of tradition. In these notes (but perhaps not elsewhere!) we shall assume that all linear systems are homogeneous unless explicitly noted otherwise.

Example 4. Here is one of those patently ridiculous "practical" examples that should nonetheless illustrate the types of problems which fit into the framework of linear systems of differential equations:

A picklemaker[1] has two brine tanks, one containing 100 gallons of brine, the other containing 200 gallons of brine. A system of pipes and pumps connects the tanks so that the following processes happen:

(a) fresh water enters the first tank at a rate of 20 gallons per minute,
(b) solution moves from the first tank to the second at a rate of 30 gallons per minute,

[1] one who makes pickles

(c) solution moves (via a separate pipe) from the second tank to the first at a rate of 10 gallons per minute, and
(d) solution is dumped from the second tank at a rate of 20 gallons per minute.

Note that total solution volume is conserved in this system.

If the first tank contains x pounds of salt and the second tank contains y pounds of salt, then the development of the system may be described by the following system of equations:

$$x' = -30\frac{x}{100} + 10\frac{y}{200}$$
$$y' = 30\frac{x}{100} - 30\frac{y}{200}.$$

Simplifying fractions and translating into matrices, we have

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -3/10 & 1/20 \\ 3/10 & -3/20 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

The picklemaker resets this system at midnight on January 1, 2016 so that the first tank contains 10 pounds of salt and the second contains 5 pounds of salt. If he lets the system run for a year, how much salt will be in the tanks?

2. The operator norm and matrix exponentiation

In order to help the picklemaker, we are going to develop a theory of matrix exponentiation. In order to ensure that these methods are legitimate (the picklemaker appreciates our help but demands authenticity), we will take a detour through the world of operator norms.

Recall that the (complex) exponential function takes any (complex) number t to

$$e^t = \sum_{k=0}^{\infty} \frac{1}{k!} t^k.$$

If $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, define $L^0 = \mathrm{id}$, the identity linear transformation, and inductively define

$$L^k = L \circ L^{k-1}$$

for $k > 0$. Thus $L^1 = L$, $L^2 = L \circ L$, $L^3 = L \circ (L \circ L)$, *etc.*

Definition 5. The *exponential* of a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$ is

$$e^L := \sum_{k=0}^{\infty} \frac{1}{k!} L^k.$$

Some comments on interpretation are necessary, some trivial, some deep. First, note that each factor $1/k!$ is a scalar, and if M is a linear transformation and λ is a scalar, then λM takes v to $\lambda M(v)$; this gives meaning to the terms $\frac{1}{k!}L^k$. Also recall that we are perfectly comfortable adding together finitely many linear transformations: $L + M : v \mapsto L(v) + M(v)$. Thus for any integer N, the sum

$$\sum_{k=0}^{N} \frac{1}{k!} L^k$$

makes sense as a linear transformation. What, though, is meant when we let $N \to \infty$?

In order to make sense of such a limit, we will need a notion of distance between linear transformations. This notion is provided by the operator norm. In order to motivate it, recall the following lemma which we proved earlier in the term.

Lemma 6. *For any linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$, there exists a constant $c \in \mathbb{R}$ such that for all $v \in \mathbb{R}^n$, $|L(v)| \leq c|v|$.*

By taking the smallest possible such c, we measure the maximal factor by which L dilates vectors; this is the operator norm, defined formally below.

Definition 7. The *operator norm* of a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ is

$$||L|| := \inf\{c \geq 0 \mid |L(v)| \leq c|v| \text{ for all } v \in \mathbb{R}^n\}.$$

It follows immediately that for any $v \in \mathbb{R}^n$,

$$|L(v)| \leq ||L|| \cdot |v|$$

In your homework, you will prove that $||L||$ may be computed in the following ways as well.

Proposition 8. *For any linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$,*

$$\begin{aligned} ||L|| &= \sup\{|L(v)| \mid |v| \leq 1\} \\ &= \sup\{|L(v)| \mid |v| = 1\} \\ &= \sup\{|L(v)|/|v| \mid v \in \mathbb{R}^n \setminus \{0\}\}. \end{aligned}$$

The operator norm is, in fact, a *norm*, meaning that it satisfies the following properties as well.

Proposition 9. *For any linear transformations $L, M : \mathbb{R}^n \to \mathbb{R}^m$ and any scalar $\lambda \in \mathbb{R}$,*

(i) $||L|| \geq 0$ *with equality if and only if* $L = 0$,
(ii) $||\lambda L|| = |\lambda| \cdot ||L||$,
(iii) $||L + M|| \leq ||L|| + ||M||$.

Proof. The first property is obvious, and you will prove the second one in your homework. For part (iii), observe that for any $v \in \mathbb{R}^n$,

$$\begin{aligned} |(L + M)(v)| &= |L(v) + M(v)| \\ &\leq |L(v)| + |M(v)| \\ &\leq ||L|| \cdot |v| + ||M|| \cdot |v| \\ &= (||L|| + ||M||)|v|. \end{aligned}$$

Since $||L+M||$ is the greatest lower bound on all $c \geq 0$ such that $|(L+M)(v)| \leq c|v|$ and $||L||+||M||$ is such a c, we see that we must have

$$||L + M|| \leq ||L|| + ||M||.$$

□

Whenever we have a norm, we get a notion of distance. (This is exactly how we developed the notion of distance in $\mathbb{R}^n$.) Indeed, with the following definition and the properties listed below in Proposition 11, we have all the necessary notions needed to talk intelligently about convergence of the series defining e^L.

Definition 10. The *operator distance* between two linear transformations $L, M : \mathbb{R}^n \to \mathbb{R}^m$ is

$$d(L, M) := ||L - M||.$$

Proposition 11. *For any linear transformations $L, M, N : \mathbb{R}^n \to \mathbb{R}^m$ we have*

(i) $d(L, M) \geq 0$ *with equality if and only if* $L = M$,
(ii) $d(L, M) = d(M, L)$, *and*
(iii) $d(L, N) \leq d(L, M) + d(M, N)$.

Proof. These properties follow immediately from their analogues in Proposition 9. □

Additionally, the operator norm has an extra feature that will prove especially useful in our analysis of the exponential on linear operators; namely, the operator norm is *submultiplicative*.

Theorem 12. *For any linear transformations $L : \mathbb{R}^n \to \mathbb{R}^m$ and $M : \mathbb{R}^k \to \mathbb{R}^n$,*

$$||L \circ M|| \leq ||L|| \cdot ||M||.$$

Thinking in terms of dilation factors, this theorem should seem fairly obvious. The proof is a one-liner using Proposition 8.

Proof. We have

$$||L \circ M|| = \sup_{|v| \leq 1} |L(M(v))| \leq \sup_{|v| \leq 1} ||L|| \cdot |M(v)| = ||L|| \sup_{|v| \leq 1} |M(v)| = ||L|| \cdot ||M||.$$

Here the first inequality follows from our old observation that $|L(w)| \leq ||L|| \cdot |w|$ with $w = M(v)$. The other manipulations are standard. □

We now come upon a surprising property of convergence with respect to the operator distance. We will only sketch the proof since a full treatment depends on Cauchy completeness of the operator metric space, and we can only do so much in the short time we've been given.

Proposition 13. *Let $(L_k : \mathbb{R}^n \to \mathbb{R}^n)_{k \geq 0}$ be a sequence of linear transformations. If $\sum_{k \geq 0} ||L_k||$ converges (as a sequence of real numbers), then $\sum_{k \geq 0} L_k$ converges with respect to the operator distance.*

Sketch of proof. We would like to show that there is a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$ such that for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\left|\left| \sum_{k=0}^{n} L_k - L \right|\right| < \varepsilon$$

whenever $n > N$. Lacking a candidate for L, we will instead show that

$$\left|\left| \sum_{k=0}^{n'} L_k - \sum_{k=0}^{n} L_k \right|\right| < \varepsilon$$

whenever n, n' are sufficiently large. In other words, the sequence of partial sums gets arbitrarily close to itself for large indices. This is in fact an equivalent condition by Cauchy completeness of the operator metric space, but we will not prove it.

Take $\varepsilon > 0$. By hypothesis, $\sum_{k \geq 0} ||L_k||$ converges, so by Cauchy completeness of the real numbers, we may choose $N \in \mathbb{N}$ such that when $n' \geq n > N$,

$$\varepsilon > \sum_{k=0}^{n'} ||L_k|| - \sum_{k=0}^{n} ||L_k|| = \sum_{k=n+1}^{n'} ||L_k||.$$

For the same n, n',

$$\left|\left| \sum_{k=0}^{n'} L_k - \sum_{k=0}^{n} L_k \right|\right| = \left|\left| \sum_{k=n+1}^{n'} L_k \right|\right| \leq \sum_{k=n+1}^{n'} ||L_k|| < \varepsilon,$$

so $\sum L_k$ converges with respect to the operator distance. □

We are now well-poised to prove convergence of the operator exponential.

Theorem 14. *For any linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$,*

$$e^L = \sum_{k=0}^{\infty} \frac{1}{k!} L^k$$

converges to a linear transformation.

Proof. By Proposition 13, it suffices to prove that $\sum_{k \geq 0} \left|\left| \frac{1}{k!} L^k \right|\right|$ converges. By submultiplicativity (Theorem 12) and Proposition 9(ii),

$$\left|\left| \frac{1}{k!} L^k \right|\right| \leq \frac{1}{k!} ||L||^k.$$

But

$$\sum_{k=0}^{\infty} \frac{1}{k!} ||L||^k = e^{||L||}$$

converges so the comparison test for series implies that $\sum_{k \geq 0} \left|\left| \frac{1}{k!} L^k \right|\right|$ converges and we are done. □

Implicit in our sketch proof of Proposition 13 is the fact that linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ form a complete metric space under the distance function induced by the operator norm. In other words, Cauchy sequences (relative to the operator norm) of linear transformations converge to linear transformations. Given Theorem 14, this means that e^L is a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$ for any linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$. As such, e^L has a matrix, which we provide notation for below.

Definition 15. If $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with matrix A, then we let e^A denote the matrix associated with e^L and call it the *exponential matrix* of A.

Of course, e^A has a series description as

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

where A^k denotes iterated matrix multiplication. But we must be careful in how we interpret this series — as it stands, it only makes sense when we translate back to linear transformations and consider convergence relative to the operator norm. It is natural to ask for more.

Question 16. Suppose (L_k) is a sequence of linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ with associated sequence of $m \times n$ matrices (A_k). Further suppose that $(L_k) \to L$ as $k \to \infty$ relative to the operator norm. Let a_k^{ij} denote the ij-entry of A_k, and let a^{ij} denote the ij-entry of the matrix A associated with L. Under these circumstances, is

$$\lim_{k \to \infty} a_k^{ij} = a^{ij}?$$

Answer 17. Yes. It turns out that all norms on finite dimensional vector spaces induce the same topology, and thus have the same convergent sequences. We will return to these ideas at the end of these notes, time permitting.

We now compute a couple of easy examples of exponential matrices.

Example 18. Suppose $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ is a 2×2 real matrix. Then $A^k = \begin{pmatrix} a^k & 0 \\ 0 & b^k \end{pmatrix}$ (check this) and

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = \sum_{k=0}^{\infty} \begin{pmatrix} a^k/k! & 0 \\ 0 & b^k/k! \end{pmatrix} = \begin{pmatrix} e^a & 0 \\ 0 & e^b \end{pmatrix}.$$

Similarly, if A is an $n \times n$ diagonal matrix with diagonal entries $a_1, \dots, a_n$, then e^A is a diagonal matrix with diagonal entries $e^{a_1}, \dots, e^{a_n}$.

Example 19. An $n \times n$ matrix A is called *nilpotent* if there is a positive integer ℓ such that $A^\ell = 0$. (Here 0 means the $n \times n$ matrix with all 0 entries.) If A is nilpotent, then the infinite sum giving e^A has only finitely many nonzero terms and may be computed explicitly by just adding up those terms, *i.e.*,

$$e^A = \sum_{k=0}^{\ell-1} \frac{1}{k!} A^k.$$

For instance, if

$$A = \begin{pmatrix} 0 & 3 & 4 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix},$$

then

$$A^2 = \begin{pmatrix} 0 & 0 & 18 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus

$$e^A = I + A + \frac{1}{2}A^2 = \begin{pmatrix} 1 & 3 & 13 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Solving linear systems of differential equations with exponential matrices

3.1. **The main theorem.** Let $M_{n\times n}(\mathbb{R})$ denote the set of $n \times n$ real matrices. For a fixed $A \in M_{n\times n}(\mathbb{R})$, we may define a matrix-valued function

$$\begin{aligned} \gamma_A : \mathbb{R} &\longrightarrow M_{n\times n}(\mathbb{R}) \\ t &\longmapsto e^{At} \end{aligned}$$

where At is the scalar multiple of A by t. Thus

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} (At)^k = \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k.$$

(Here we are using the fact that scalar multiplication commutes with matrix multiplication.) Identifying $M_{n\times n}(\mathbb{R})$ with $\mathbb{R}^{n^2}$ (why can we do this?), we may view γ_A as a path. As such, it makes sense to differentiate:

$$\begin{aligned}\gamma_A'(t) &= \frac{d}{dt}e^{At}\\ &= \frac{d}{dt}\sum_{k=0}^{\infty}\frac{1}{k!}A^k t^k\\ &= \sum_{k=0}^{\infty}\frac{1}{k!}A^k\frac{d}{dt}(t^k)\\ &= \sum_{k=0}^{\infty}\frac{k}{k!}A^k t^{k-1}\\ &= \sum_{k=1}^{\infty}\frac{1}{(k-1)!}A^k t^{k-1}\\ &= A\sum_{k=1}^{\infty}\frac{1}{(k-1)!}A^{k-1}t^{k-1}\\ &= Ae^{At}.\end{aligned}$$

Several steps here require justification. First and foremost, we have not justified the third equality, which claims that differentiation commutes with infinite series. This is a deeper point that we will return to later. The final step also deserves attention. Observe that k ranges from 1 to ∞, but all of the appearances of k in the sum are in fact $(k-1)$'s. This recovers the exponential function.

Given this computation, we may prove the following theorem.

Theorem 20. *Let $A \in M_{n\times n}(\mathbb{R})$ and let $p_0 \in \mathbb{R}^n$. Then the function*

$$\begin{aligned}x : \mathbb{R} &\longrightarrow \mathbb{R}^n\\ t &\longmapsto e^{At}p_0\end{aligned}$$

is a solution to the differential equation $x' = Ax$ with initial condition $x(0) = p_0$.

Proof. First note that by Example 18, $e^0 = I$ where 0 is the $n \times n$ matrix of all 0's. Now observe that $x(0) = e^{A\cdot 0}p_0 = Ip_0 = p_0$, so x satisfies the initial condition. Finally, we may compute

$$x' = \frac{d}{dt}(e^{At}p_0) = Ae^{At}p_0 = Ax,$$

as desired. □

Example 21. Consider the matrix

$$A = \begin{pmatrix}0 & 3 & 4\\ 0 & 0 & 6\\ 0 & 0 & 0\end{pmatrix}$$

of Example 19. It is easy to compute

$$e^{At} = I + At + \frac{1}{2}A^2t^2 = \begin{pmatrix}1 & 3t & 4t+9t^2\\ 0 & 1 & 6t\\ 0 & 0 & 1\end{pmatrix}.$$

As such, we can use Theorem 20 to solve the differential equation

$$\begin{aligned}x' &= 3y + 4z\\ y' &= 6z\\ z' &= 0.\end{aligned}$$

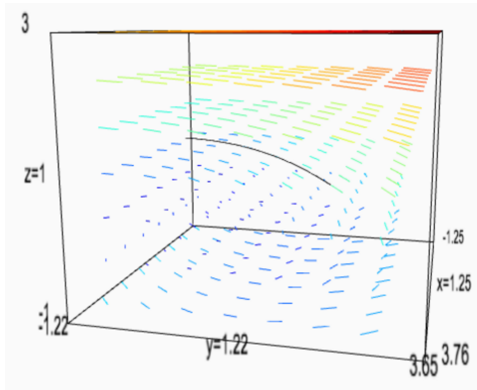

FIGURE 2. A vector field and flow line given by exponentiation of a nilpotent matrix

In particular, if our initial condition is $\gamma(0) = (1, 1, 1)$, then

$$\gamma(t) = e^{At} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (1 + 7t + 9t^2, 1 + 6t, 1)$$

is a flow line for this system. Figure 2 sketches the vector field $A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and the flow line γ.

3.2. **Properties of the matrix exponential.** We will use the following theorem to compute more complicated exponential matrices (and thus solve more complicated linear systems of differential equations).

Theorem 22. *If $A, B \in M_{n\times n}(\mathbb{R})$ and $AB = BA$, then*

$$e^{A+B} = e^A e^B.$$

Proof. The essential observation is that when A and B commute, the binomial theorem applies so that

$$(A + B)^k = \sum_{\ell=0}^{k} \binom{k}{\ell} A^\ell B^{\ell-k} = \sum_{\ell=0}^{k} \frac{k!}{\ell!(k-\ell)!} A^\ell B^{\ell-k}$$

Hence

$$\begin{aligned} e^{A+B} &= \sum_{k=0}^{\infty} \frac{1}{k!}(A + B)^k \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{\ell=0}^{k} \frac{k!}{\ell!(k-\ell)!} A^\ell B^{\ell-k} \right) \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{1}{\ell!} A^\ell \cdot \frac{1}{(k-\ell)!} B^{\ell-k} \\ &= \sum_{k=0}^{\infty} \sum_{r+s=k} \frac{1}{r!} A^r \cdot \frac{1}{s!} B^s. \end{aligned}$$

Observe that in this final double sum, the pairs (r, s) range through all of $\mathbb{N} \times \mathbb{N}$ exactly once. Thus (leaving some algebraic verifications to the reader) we may write

$$e^{A+B} = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{1}{r!} A^r \frac{1}{s!} B^s = \left(\sum_{r=0}^{\infty} \frac{1}{r!} A^r \right) \left(\sum_{s=0}^{\infty} \frac{1}{s!} B^s \right) = e^A e^B,$$

as desired. □

Remark 23. The above theorem is generally false if A and B do not commute!

Corollary 24. *For all $A \in M_{n\times n}(\mathbb{R})$, e^A is invertible with*

$$e^A e^{-A} = I.$$

Proof. We have $A(-A) = (-A)A$, so Theorem 22 implies that

$$I = e^0 = e^{A-A} = e^A e^{-A}.$$

□

Example 25. Let's use this theorem to compute the exponential of $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$. First note that $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$ and

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & ab \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

Thus

$$e^{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}} = e^{\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}} e^{\begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}}.$$

Since $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ is diagonal, we have

$$e^{\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}} = \begin{pmatrix} e^a & 0 \\ 0 & e^a \end{pmatrix}.$$

Since $\begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}^2 = 0$, we have

$$e^{\begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}} = I + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}.$$

Thus we may conclude that

$$e^{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}} = \begin{pmatrix} e^a & 0 \\ 0 & e^a \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^a & e^a b \\ 0 & e^a \end{pmatrix}.$$

More generally,

$$e^{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} t} = \begin{pmatrix} e^{at} & e^{at} bt \\ 0 & e^{at} \end{pmatrix}.$$

Thus the differential equation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

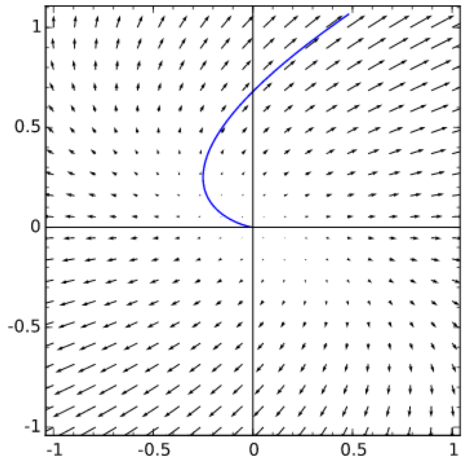

FIGURE 3. Vector field and flow line when $a = b = 1$, $p_1 = -1/4$, and $p_2 = 1/4$.

with initial condition (p_1, p_2) at time 0 has solution

$$e^{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} t} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} e^{at} & e^{at}bt \\ 0 & e^{at} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} p_1 e^{at} + p_2 e^{at} bt \\ p_2 e^{at} \end{pmatrix}.$$

In Figure 3, we set $a = b = 1$. The plot depicts the vector field $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ along with a flow line passing through $(-1/4, 1/4)$ at $t = 0$. Explicitly, this solution is given by $t \mapsto (-e^t/4 + e^t t/4, e^t/4)$.

3.3. **Periodic matrices.**

Example 26. We will now consider a system of differential equations we have seen before:

$$x' = -y$$
$$y' = x.$$

We have already seen that the solutions to this system are circles. Does the matrix exponential recover these solutions?

We may re-express the above system as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Observe that

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since the fourth power of this matrix is the identity matrix, the pattern now repeats, *i.e.*, if $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, then $A^{4k+j} = A^j$ where $k \in \mathbb{N}$ and j is 0, 1, 2, or 3. Thus we may compute the exponential

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} t^{2k} & -\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} t^{2k+1} \\ \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} t^{2k+1} & \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} t^{2k} \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

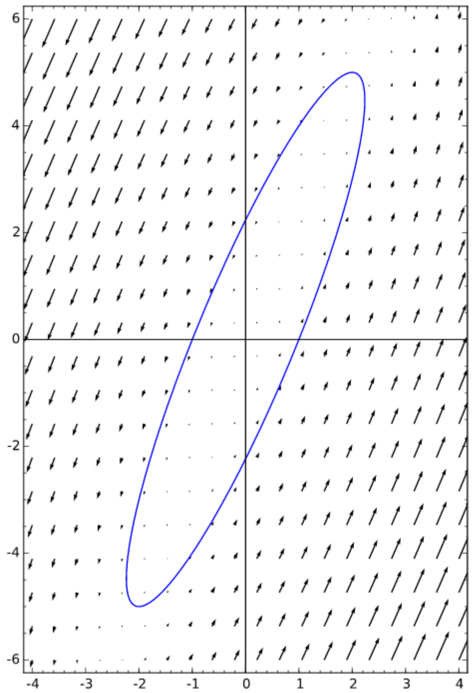

FIGURE 4. An ellipse as a solution to a homogeneous linear system of differential equations?

(The sums here look complicated, but they are just recording the various ± 1's, 0's, factorials, and powers of t which appear in the summation.)

If at time 0 we are at the point $(r, 0)$, then Theorem 20 tells us that our system has solution

$$t \longmapsto e^{At}\begin{pmatrix} r \\ 0 \end{pmatrix} = \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix}.$$

This precisely describes a circle of radius r centered at the origin, which matches our old solution of this problem.

Question 27. Do you think there are homogeneous systems of linear equations whose solutions are ellipses? If so, how might you form them? (If this seems mysterious right now, come back to this question after you have learned what a change of basis is in Linear Algebra.)

Let $M = \begin{pmatrix} 1/2 & 0 \\ 1 & 1/2 \end{pmatrix}$ which has inverse $M^{-1} = \begin{pmatrix} 2 & 0 \\ -4 & 2 \end{pmatrix}$ (check this!) and consider the system $\begin{pmatrix} x' \\ y' \end{pmatrix} = MAM^{-1}\begin{pmatrix} x \\ y \end{pmatrix}$. Figure 4 plots the vector field $MAM^{-1}\begin{pmatrix} x \\ y \end{pmatrix}$ and a solution to this system passing through $(1, 0)$ at $t = 0$ given by $t \mapsto (\cos(t) + 2\sin(t), 5\sin(t))$. This should seem quite tantalizing.

3.4. **A linear algebraic approach to the general** 2×2 **case.** Let A be an $n \times n$ matrix and suppose $A = QBQ^{-1}$ for Q, B $n \times n$ matrices with Q invertible. Then

$$A^2 = (QBQ^{-1})(QBQ^{-1}) = QB(QQ^{-1})BQ^{-1} = QB^2Q^{-1}$$

since $QQ^{-1} = I$. Similarly, for any $k \geq 0$,

$$A^k = QB^kQ^{-1}.$$

Thus

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k = \sum_{k=0}^{\infty} \frac{1}{k!} Q B^k Q^{-1} t^k = Q \left(\sum_{k=0}^{\infty} \frac{1}{k!} B^k t^k \right) Q^{-1} = Q e^{Bt} Q^{-1}.$$

As such, if we can find B and Q such that e^{Bt} is computable and $A = QBQ^{-1}$, then we will be able to compute e^{At}. In the 2×2 case, the following theorem permits us to do just this.

Theorem 28. *Let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *be a* 2×2 *matrix. Then* $A = QBQ^{-1}$ *where* B *is of the form*

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \textit{or} \quad \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

Here the (real or) complex numbers λ_i *are the roots of the* characteristic polynomial *of* A*:*

$$x^2 - (a+d)x + (ad - bc).$$

Proof. This is the 2×2 case of a powerful theorem from linear algebra about so-called *Jordan canonical form* of matrices. We will not comment on the proof here. □

Remark 29. As stated, the theorem does not tell us anything about how to find Q. In the first case, where $B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ this is relatively simple. In this case, it is possible to find a basis (of $\mathbb{R}^2$) of *eigenvectors* of B. An eigenvector of B is a vector v such that $Bv = \lambda v$ for some scalar λ; the scalar λ in such an equation is called an *eigenvalue*. The eigenvalues of $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ are precisely λ_1 and λ_2 (check this!). If v_1 and v_2 are linearly independent vectors such that $Bv_1 = \lambda_1 v_1$ and $Bv_2 = \lambda_2 v_2$, then we may take Q to be the matrix with columns v_1 and v_2.

In the second case, in which $B = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$, the matrix Q is constructed in a different fashion. The above method does not work because B has only one eigenvalue, λ_1, and is not already diagonal, and thus cannot have a basis of eigenvectors. While it is not particularly complicated, we will not study this method in these notes.

Exercise 1. Use these methods to compute the exponential matrix of $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. You will find $\lambda_1 = i$, $\lambda_2 = -i$, so that $e^{Bt} = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$. After finding an invertible matrix Q such that $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = QBQ^{-1}$ and computing $e^{At} = Qe^{Bt}Q^{-1}$ you will be confronted with a matrix of the form

$$\begin{pmatrix} \frac{1}{2}(e^{it} + e^{-it}) & \frac{i}{2}(e^{it} - e^{-it}) \\ \frac{i}{2}(e^{-it} + e^{-it}) & \frac{1}{2}(e^{it} + e^{-it}) \end{pmatrix}.$$

Recalling the identity

$$e^{it} = \cos(t) + i \sin(t)$$

you should see how this matrix simplifies to previous computation of e^{At}.

Example 30. Consider the system of differential equations

$$x' = 4x + 2y$$
$$y' = 3x - y$$

which is induced by the matrix $A = \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix}$. This matrix has characteristic polynomial

$$x^2 - 3x - 10 = (\lambda + 2)(\lambda - 5)$$

with roots $\lambda_1 = -2$, $\lambda_2 = 5$. Thus

$$B = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix} \quad \text{and} \quad e^{Bt} = \begin{pmatrix} e^{-2t} & 0 \\ 0 & e^{5t} \end{pmatrix}.$$

We can compute Q and Q^{-1} to be

$$Q = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix} \quad \text{and} \quad Q^{-1} = \begin{pmatrix} 1/7 & -2/7 \\ 3/7 & 1/7 \end{pmatrix}$$

so

$$e^{At} = Qe^{Bt}Q^{-1} = \frac{1}{7}\begin{pmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{pmatrix}.$$

The solution to the system of differential equations with initial conditions $x(0) = 7$ and $y(0) = 0$ is thus

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{1}{7}\begin{pmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{pmatrix}\begin{pmatrix} 7 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{-2t} + 6e^{5t} \\ -3e^{-2t} + 3e^{5t} \end{pmatrix}.$$

It would be nice to additionally understand the curve described by this solution. To derive this, first note that the system given by B with initial condition $(p_1, p_2$ has solution

$$(x(t), y(t)) = (p_1 e^{-2t}, p_2 e^{5t}).$$

In this case,

$$x^5 y^2 = p_1^5 p_2^2$$

which gives us a handle on the curve described.

The next thing to observe is that $\gamma(t)$ is a solution to the system given by B if and only if $Q\gamma(t)$ is a solution for the system given by A. (Why is this the case?) Hence the solution curves for A differ from the ones for B via the linear change of coordinates Q. (Now would be an opportune moment to return to Question 27 and to consider what curves are described when a linear change of coordinates is applied to a circle.)

4. Theoretical odds and ends

Given the matrix exponential's facility in solving differential equations, we can now see why it was important to prove its convergence. Recall, though, that we only proved convergence with respect to the operator norm? Is this the same thing as convergence with respect to a more traditional norm, like the one induced by considering $n \times n$ matrices as points in $\mathbb{R}^{n^2}$?

We will sketch the proof of a theorem which, like Voltaire's Pangloss, assures us that we live in the best of all possible worlds. Loosely speaking, it tells us that all topologies on finite dimensional real vector spaces which are induced by norms are equivalent. Thus convergence with respect to one norm is equivalent to convergence with respect to all other norms. In order to put meat on this statement, let us recall the definition of a norm.

Definition 31. A *norm* on $\mathbb{R}^n$ is a function $|| \; || : \mathbb{R}^n \to \mathbb{R}$ satisfying the following properties:

(1) for all $x \in \mathbb{R}^n$, $||x|| \geq 0$ with equality if and only if $x = 0$,
(2) for all $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, $||\lambda x|| = |\lambda| \cdot ||x||$, and
(3) for all $x, y \in \mathbb{R}^n$, $||x + y|| \leq ||x|| + ||y||$.

The Euclidean norm on $\mathbb{R}^n$ is, of course, a norm. *A priori*, the operator norm is defined on the set of linear transformations $\mathbb{R}^n \to \mathbb{R}^n$. Using our dictionary between linear transformations and matrices, though, we see that each linear map $\mathbb{R}^n \to \mathbb{R}^n$ corresponds to a unique point in $\mathbb{R}^{n^2}$. Thus it makes sense to think of the operator norm as a norm on $\mathbb{R}^{n^2}$.

Whenever we have a norm, it induces a distance function $d_{|| \, ||}(x, y) = ||x - y||$, and we can define open balls and a topology with respect to this distance function. We can define limits of sequences with respect to this norm via the usual ε-N formalism.

Definition 32. Two norms $|| \, ||$ and $|| \, ||'$ are *equivalent* if there are positive constants c and C such that

$$c||x||' \leq ||x|| \leq C||x||'.$$

Remark 33. The apparent asymmetry in this definition is an illusion. Given the chain of inequalities $c||x||' \leq ||x|| \leq C||x||'$, we get

$$\frac{1}{C}||x|| \leq ||x||' \leq \frac{1}{c}||x||$$

for free. As such, equivalence of norms is in fact an equivalence relation on the set of norms (check this!).

The significance of this definition is made apparent by the following theorem.

Theorem 34. *If two norms $|| \, ||$ and $|| \, ||'$ on $\mathbb{R}^n$ are equivalent, then*

(1) a sequence a_k is convergent with respect to $|| \, ||$ if and only if it is convergent with respect to $|| \, ||'$, and
(2) a subset of $\mathbb{R}^n$ is open with respect to $|| \, ||$ if and only if it is open with respect to $|| \, ||'$.

The proof follows directly from the definitions (check this!) and we will not present it here.
We can now properly appreciate our Panglossian theorem:

Theorem 35. *For $n \geq 0$, any two norms on $\mathbb{R}^n$ are equivalent.*

Corollary 36. *The matrix exponential converges with respect to all norms on $\mathbb{R}^{n^2}$ (including the usual Euclidean one).*

Since it will be essential in our proof of the theorem, let us recall the Bolzano-Weierstrass theorem (which you learned in some for in Math 112) before proceeding.

Theorem 37 (Bolzano-Weierstrass)**.** *Every bounded sequence in $\mathbb{R}$ has a convergent subsequence. (Here we mean convergence with respect to the standard absolute value on $\mathbb{R}$.)*

For $x \in \mathbb{R}^n$, we define $||x||_1 = |x_1| + \cdots + |x_n|$; it is called the ℓ^1 norm or the taxi-cab norm (think of the distance your taxi must take to travel between points in Manhattan). Check that $|| \, ||_1$ is in fact a norm. We will need the following enhanced version of the Bolzano-Weierstrass theorem.

Theorem 38 (Bolzano-Weierstrass on $\mathbb{R}^n$ with respect to the $|| \, ||_1$ norm)**.** *In $\mathbb{R}^n$, any $|| \, ||_1$-bounded sequence has a $|| \, ||_1$-convergent subsequence.*

Proof. Let $\{x_k\}$ be a $|| \, ||_1$-bounded sequence in $\mathbb{R}^n$. Let x_k^j be the j-th coordinate of x_k, $1 \leq j \leq n$. Note that $|x_k^1| \leq ||x_k||_1$, so $\{x_k^1\}$ is a bounded sequence in $\mathbb{R}$ and thus has a convergent subsequence $\{x_{\ell_k}\}$ by Theorem 37. Now consider the sequence $\{x_{\ell_k}^2\}$. We again have that $|x_{\ell_k}^2| \leq ||x_{\ell_k}||_1$ is bounded, so there is a convergent subsequence $x_{h_{\ell_k}}$. If $n = 2$, we are done with convergent subsequence $(x_{h_{\ell_k}}^1, x_{h_{\ell_k}}^2)$ (check this!). Otherwise, keep repeating the process until it has been done to all of the coordinates. ☐

Remark 39. It is a consequence of Theorem 35 that the Bolzano-Weierstrass theorem on $\mathbb{R}^n$ holds with respect to any norm.

We are now prepared to prove our main theorem.

Proof. Since equivalence of norms is transitive, it suffices to show that an arbitrary norm $|| \ ||$ is equivalent to $|| \ ||_1$.

We first check that there is a constant c such that $c||x|| \leq ||x||_\infty$ for all $x \in \mathbb{R}^n$. Observe that $x = \sum x_i e_i$, so the triangle inequality implies that $||x|| \leq \sum |x_i| \cdot ||e_i||$. If $m = \max\{||e_i||\}$, then we further have $||x|| \leq \sum |x_i| \cdot m = m||x||_1$. Hence we may take $c = 1/m$ and derive that $c||x|| \leq ||x||_1$.

It remains to show that there is a positive constant C such that $||x||_1 \leq C||x||$ for all $x \in \mathbb{R}^n$. Suppose for contradiction that no such C exists. In that case, for each $k \in \mathbb{N}$ we can find $x_k \in \mathbb{R}^n$ such that

$$||x_k||_1 > k||x_k||.$$

Let $y_k = x_k/||x_k||_1$. Then $||y_k||_1 = 1$ so the sequence $\{y_k\}$ is bounded in the $|| \ ||_1$ norm and thus has a subsequence y_{ℓ_k} converging in $|| \ ||_1$ to an element y_∞. Since $||y_{\ell_k}||_1 = 1$, we also have $||y_\infty||_1 = 1$ (check this!) and, in particular, $y_\infty \neq 0$. Note, though, that $||y_{\ell_k}|| < 1/\ell_k$, so $||y_\infty|| = 0$, whence $y_\infty = 0$, a contradiction. Thus there must be some $C > 0$ such that $||x||_1 \leq C||x||$ for all x, completing our proof. □