The Matrix Exponential

Adrien Deloro

Şirince ’21 Summer School

These are lecture notes with exercises for a one week course of $6 \times 2 = 12$ hours given in July 2021 at the Nesin Matematik Köyü in Şirince, Turkey. They should be accessible to a 3rd student in mathematics; syllabus and prerequisites are described below. I wish to thank Quentin Dupré for many clever suggestions.

If found with flaws or mistakes, this document compiled on 23rd July 2021 should be returned to adrien.deloro@imj-prg.fr. Please mention the date.

Contents

Introduction

> To solve the linear equation $x''(t) - 2x'(t) + x(t) = 0$, one is told to first solve polynomial $\lambda^2 - 2\lambda + 1 = 0$ (with double root 1) and then write $x(t) = c_1 e^t + x_2 te^t$. *Why?*

A partial answer is of course that:

1. by the Cauchy-Lipschitz (also known as Picard-Lindelöf) theorem, the set of solutions is a vector space of dimension 2; and

2. one can *check* that functions e^t and te^t are solutions, and linearly independent.

While 1. is satisfactory but begs for more linear algebra, 2. remains somehow magical, and simply begs for more linear algebra. We therefore reformulate question 'Why?' as follows.

Question. Is there a way to combine the exponential map with linear algebra, and apply it to linear differential equations?

This class provides a positive answer. In § 1 we recast the tools necessary for § 2, where we define and give the first properties of the matrix exponential. In § 3 we introduce the important *Chevalley decomposition*, which helps compute the matrix exponential; its interest goes beyond that course. Sections 4 and 5 are more advanced: § 4 studies further analytic properties and requires some familiarity with multi-variable calculus; § 5 shows how the matrix exponential allows one to study certain matrix groups. Fortunately both are optional reads. Finally we turn to linear ODE's in § 6.

This class should be accessible to a third year student; it is not completely self-contained, as described below.

Prerequisites

Analysis and topology: We need only basic notions from analysis: distances, continuity, convergence, Cauchy sequences, completeness, series. (Fortunately, one does not need to be skilled in the art of computing series.) Topologically one is expected to know compactness and the 'extreme value theorem' that a real-valued, continuous map on a compact set is bounded and attains its bound.

Functional analysis: The notion of a normed vector space will be re-introduced in § 1.2 but some preliminar familiarity helps: over a real or complex, *finite-dimensional* vector space, all norms are equivalent.

Linear algebra: We must borrow a little more from linear algebra: eigenvalues and eigenvectors, the Cayley-Hamilton theorem, diagonalisation, trigonalisation. Polynomials of matrices play a constant role.

It is good to be trained in thinking 'up to conjugacy'. We put emphasis on abstract properties, not on computing base changes.

Differential geometry (optional): At some point there is a little (optional) differential geometry; in order to read § 4, one must be familiar with the differential (in many variables), diffeomorphisms, and the inverse function theorem.

Quadratic algebra (optional): For § 5, one needs decent training in scalar products and Hermite/complex scalar products; hermitian and unitary matrices, and their diagonalisation properties; or at least, their real analogues.

Differential equations: If you could read the introduction, you know enough.

1 Trying to generalise the complex case

In this section we review the formalisation of the exponential of complex numbers in order to carry it to matrices. In § 1.1 we inspect the classical construction in the complex case. In order to generalise, one needs a vector analogue of the modulus $|\cdot|$, the so-called *norms* (§ 1.2), and then see how norms on matrix spaces interact with matrix multiplication (§ 1.3).

1.1 The complex exponential

As one expects, the field of real numbers is denoted by $\mathbb{R}$. For the set of non-negative real numbers, we use $\mathbb{R}_{\geq 0}$; for positive real numbers, we write $\mathbb{R}_{>0}$. Notation such as $\mathbb{R}^*, \mathbb{R}^+, \mathbb{R}^{++}$ are strongly discouraged; in general, an analyst's notation tend to obscure algebraic structures. The field of complex numbers is denoted by $\mathbb{C}$; there is no ordering on $\mathbb{C}$ compatible with its algebraic structure, so '$\mathbb{C}_{>0}$' is completely meaningless. Most of what we do takes place over $\mathbb{R}$ or $\mathbb{C}$ indifferently; so throughout, $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$. As algebraic structures the fields $\mathbb{R}$ and $\mathbb{C}$ are well-understood, and we add more data, a notion of distance (and hence a topology).

Recall that $\mathbb{R}$ and $\mathbb{C}$ are equipped with

$$|\cdot|: \begin{array}{ccc} \mathbb{C} & \to & \mathbb{R}_{\geq 0} \\ z & \mapsto & |z| \end{array},$$

a map called the *modulus*, or the *absolute value* when the focus is on real numbers. It enjoys familiar properties.

1.1.1. Properties.

- $(\forall z \in \mathbb{C})(|z| = 0 \leftrightarrow z = 0)$.
- $(\forall z_1, z_2 \in \mathbb{C})(|z_1 + z_2| \leq |z_1| + |z_2|)$ *('triangle inequality')*.
- $(\forall z_1, z_2 \in \mathbb{C})(|z_1 \cdot z_2| = |z_1| \cdot |z_2|)$.

In particular, $|\cdot|$ is a distance in the sense of metric topology, thus giving rise to a notion of convergence.

1.1.2. Definition. Let $(z_n) \in \mathbb{C}^{\mathbb{N}}$ be a sequence of complex numbers.

- The sequence *converges to* $\ell \in \mathbb{C}$ if:

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists n_0 \in \mathbb{N})(\forall n \in \mathbb{N})(n \geq n_0 \to |z_n - \ell| < \varepsilon),$$

in which case ℓ is called *the limit* of the sequence; indeed, if it exists, 'the' limit is unique. As usual this is indicated by $z_n \xrightarrow[n\to\infty]{} \ell$ or simply $z_n \longrightarrow \ell$.

- The sequence is a *Cauchy sequence* if:

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists n_0 \in \mathbb{N})(\forall p, q \in \mathbb{N})(p \geq q \geq n_0 \to |z_p - z_q| < \varepsilon).$$

It is easily seen by the triangle inequality that every convergent sequence is a Cauchy sequence. The converse actually holds.

1.1.3. Theorem (admitted from analysis)**.** *The metric space $\mathbb{C}$ is* metrically complete, *viz. every Cauchy sequence is convergent. Since $\mathbb{R}$ is a closed subset of $\mathbb{C}$, the same holds of $\mathbb{R}$.*

We apply this to *series*. A series is just the sequence of partial sums of an 'infinite sum'; due to their historical and technical significance, they are sometimes taught in specific classes by analysts. Suppose we are trying to sum an infinite amount of complex numbers, say the numbers z_n. We then write:

- $\sum z_n$ for the series itself;
- $S_n = \sum_{k=0}^n z_k$ for the partial sum up to n; hence (S_n) is another notation for $\sum z_n$;
- $\sum_{n=0}^\infty z_n$ for the limit of S_n, if it exists.

1.1.4. Definition. Let $\sum z_n$ be a series of complex numbers. Call the series *absolutely convergent* if the real non-negative series $\sum |z_n|$ is convergent.

Notice that since $\sum |z_n|$ is real and non-negative, it is convergent iff bounded above.

1.1.5. Corollary. *Let $\sum z_n$ be a complex series. Suppose it is absolutely convergent. Then it is convergent.*

Proof. Let $S_n = \sum_{k=0}^n z_k$; we must prove that the sequence (S_n) is convergent. The proof uses completeness (theorem 1.1.3): we simply prove that (S_n) is a Cauchy sequence. Let $\varepsilon \in \mathbb{R}_{>0}$; we look for suitable n_0. By assumption, the series $\sum |z_n|$ is convergent. This means that the sequence $T_n = \sum_{k=0}^n |z_k|$ is convergent. We relate (S_n) and (T_n) as follows.

Notice that when $p > q$ are fixed integers, one has:

$$\begin{aligned} |S_p - S_q| &= \left| \sum_{k=0}^p z_k - \sum_{k=0}^q z_k \right| \\ &= \left| \sum_{k=q+1}^p z_k \right| \\ &\leq \sum_{k=q+1}^p |z_k| \\ &= T_p - T_q. \end{aligned}$$

(The same vacuously holds if $p = q$.) Since (T_n) is convergent, it is a Cauchy sequence. It also is non-decreasing (being a series of non-negative real numbers). So there is $n_0 \in \mathbb{N}$ such that for all integers p, q, one has:

$$p \geq q \geq n_0 \to 0 \leq T_p - T_q < \varepsilon.$$

By the above, for $p \geq q \geq n_0$ one has $|S_p - S_q| = T_p - T_q < \varepsilon$, which means that (S_n) itself is a Cauchy sequence. □

This may not look like much but allows us to introduce the complex exponential, a definition we want to generalise to matrices.

1.1.6. Theorem. *For all $z \in \mathbb{C}$, the series $\sum \frac{z^n}{n!}$ converges.*

Proof. It suffices to prove that the series is absolutely convergent, so by multiplicativity of $|\cdot|$ we are left with dealing with $\sum \left|\frac{z^n}{n!}\right| = \sum \frac{|z|^n}{n!}$. So we may work with $|z| \in \mathbb{R}_{\geq 0}$; for simplicity denote it by t. Of course we may assume $t > 0$; it is a fixed, positive, real number. We are studying $\sum \frac{t^n}{n!}$.

Let $n_0 \in \mathbb{N}$ be such that $n_0 \geq 2t$. Then for $n \geq n_0$, one has:

$$\frac{t^{n+1}}{(n+1)!} = \frac{t}{n+1} \cdot \frac{t^n}{n!} \leq \frac{1}{2} \cdot \frac{t^n}{n!} \leq \cdots \leq \left(\frac{1}{2}\right)^{n-n_0} \cdot \frac{t^{n_0}}{n_0!}.$$

This allows us to split the series into two:

$$\sum \frac{t^n}{n!} = \underbrace{\sum_{n<n_0} \frac{t^n}{n!}}_{\Sigma_1} + \underbrace{\sum_{n\geq n_0} \frac{t^n}{n!}}_{\Sigma_2}.$$

The first sum Σ_1 is obtained through finite summation; the series Σ_2 is clearly convergent since:

$$\sum_{n\geq n_0} \frac{t^n}{n!} \leq \sum_{n\geq n_0} \left(\frac{1}{2}\right)^{n-n_0} \cdot \frac{t^{n_0}}{n_0!}.$$

Hence the real series $\sum \frac{t^n}{n!}$ converges, meaning that the complex series $\sum \frac{z^n}{n!}$ is absolutely convergent, hence convergent by Corollary 1.1.5. □

This defines the complex exponential $\exp\colon \mathbb{C} \to \mathbb{C}^\times$. We admit that it is a group homomorphism $(\mathbb{C}, +) \to (\mathbb{C}^\times, \cdot)$ as we do not want to manipulate series. (This is done in exercise 1.4.1.)

1.1.7. Remark. Proper study of the complex exponential belongs to complex analysis, and the theory of holomorphic functions. Indeed exp has fascinating analytic properties, and is arguably deeper a topic than the matrix exponential. However we shall not discuss them. See Rudin's classical *Real and Complex Analysis* for more.

1.1.8. Inspection and discussion. In order to adapt the construction of the complex exponential to the matrix case, we have a couple of tasks:

- $\mathbf{T}_1$. find a suitable analogue of $|\cdot|$ for matrices;
- $\mathbf{T}_2$. obtain convergence criteria for series in matrix spaces, extending Corollary 1.1.5 ;
- $\mathbf{T}_3$. find a substitute for multiplicativity of $|\cdot|$, which may no longer hold with matrices.

Tasks $\mathbf{T}_1$ and $\mathbf{T}_2$ are treated in § 1.2 and task $\mathbf{T}_3$ in § 1.3. Then in § 2 we can start studying the matrix exponential.

1.2 Normed vector spaces

We handle tasks $\mathbf{T}_1$ and $\mathbf{T}_2$ given in discussion § 1.1.8, viz.:

$\mathbf{T}_1$. find a suitable analogue of $|\cdot|$ for matrices;

$\mathbf{T}_2$. obtain convergence criteria for series in matrix spaces, extending Corollary 1.1.5.

Recall that $\mathbb{K}$ stands for $\mathbb{R}$ or $\mathbb{C}$.

1.2.1. Definition. Let V be a $\mathbb{K}$-vector space. A *norm* on V is a map $\|\cdot\|\colon V \to \mathbb{R}_{\geq 0}$ satisfying:

- $(\forall v \in V)(\|v\| = 0 \leftrightarrow v = 0)$;
- $(\forall v_1, v_2 \in V)(\|v_1 + v_2\| \leq \|v_1 + v_2\|)$;
- $(\forall \lambda \in \mathbb{K})(\forall v \in V)(\|\lambda \cdot v\| = |\lambda| \cdot \|v\|)$.

Notice that $|\cdot|$ is actually a norm on $\mathbb{C}$ seen as either a $\mathbb{R}$- or a $\mathbb{C}$-vector space. But there are of course many more examples. We shall focus on finite-dimensional spaces.

1.2.2. Examples. Let $V = \mathbb{K}^n$. Define:

- $\|(x_1, \ldots, x_n)\|_1 = \sum_{k=1}^n |x_k|$;
- $\|(x_1, \ldots, x_n)\|_2 = \sqrt{\sum_{k=1}^n |x_k|^2}$;
- $\|(x_1, \ldots, x_n)\|_\infty = \max_{k=1\ldots n} |x_k|$.

Then each is a norm on V. (For $\|\cdot\|_2$ there is of course something to understand relating to scalar products and the Cauchy-Schwarz inequality; the notion is supposed to be familiar.)

This fulfills task $\mathbf{T}_1$ and we turn to $\mathbf{T}_2$. A norm, being a distance, gives rise to a topology; the notions of a convergent sequence, of a Cauchy sequence (recalled in definition1.1.2) are obtained replacing $|\cdot|$ by $\|\cdot\|$; as usual with metric topologies, a sequence has *at most* one limit. We are trying to prove completeness of matrix spaces when equipped with norms. But we have a situation: there are plenty of norms, so on the face of it there are many distinct topologies coming from norms. Which to choose?

1.2.3. Definition. Let V be a $\mathbb{K}$-vector space. Two norms N, N' are *equivalent* if there is $c \in \mathbb{R}_{>0}$ such that:

$$(\forall v \in V)(N(v) \leq cN'(v) \wedge N'(v) \leq cN(v)).$$

This means that each of the norm bounds the other. One may write the definition in many equivalent ways (with two distinct constants, with $\frac{1}{c}\ldots$); our formulation is symmetric, which is always nice.

1.2.4. Lemma. *Suppose N and N' are equivalent norms on V. Then they define the same topology (and the same notion of boundedness).*

Proof. Let c witness equivalence, viz. $(\forall v \in V)(N(v) \leq cN'(v) \wedge N'(v) \leq cN(v))$. Clearly '$N$-bounded' and '$N'$-bounded' bear the same meaning; we turn to topologies.

Here is a proof if you know what an abstract topology is. We prove that N and N'

give rise to the same open sets. Let $X \subseteq V$ be any subset; it is enough to supppose that X is N-open, and prove that it is N'-open. So let $x \in X$; by assumption there is $\varepsilon \in \mathbb{R}_{>0}$ such that $B_N(x, \varepsilon) \subseteq X$. Then clearly $B_{N'}(x, \frac{1}{c}\varepsilon) \subseteq X$. So X is N'-open. We conclude symmetrically.

In case you don't know what a topology is, we prove that the two norms give rise to the same notion of convergence. So suppose $(v_n) \in V^{\mathbb{N}}$ is a sequence which is convergent with respect to N, viz. there is $\ell \in V$ such that $v_n \longrightarrow \ell\,[N]$ (sometimes this is indicated by $v_n \xrightarrow{N} \ell$). We prove that $v_n \longrightarrow \ell\,[N']$. Let $\varepsilon \in \mathbb{R}_{>0}$. Now $v_n \longrightarrow \ell\,[N]$, so there is n_0 such that:

$$(\forall n \in \mathbb{N}) \left(n \geq n_0 \rightarrow N(v_n - \ell) < \frac{1}{c}\varepsilon \right).$$

Fix such n_0; for any $n \geq n_0$ one then has $N'(v_n - \ell) \leq c \cdot N(v_n - \ell) < \varepsilon$, which proves $v_n \longrightarrow \ell\,[N']$. We conclude symmetrically. □

Hence two equivalent norms give rise to the same topology.

1.2.5. Theorem. *(As always, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.) Let V be a finite-dimensional $\mathbb{K}$-vector space. Then all norms on V are equivalent.*

Proof not covered in class.

Proof. Hopefully this is familiar; the idea is to use as little as possible. Notice that equivalence of norms is a transitive relation. So it suffices to prove that every norm is equivalent to one norm of reference. We shall transfer the problem to $(\mathbb{K}^n, \|\cdot\|_\infty)$, which has good topological properties.

Let $\mathcal{B} = (b_1, \ldots, b_n)$ be a basis of V. By definition, for each $v \in V$ there is a unique tuple $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$ with $v = \sum \lambda_k b_k$. Put $\|v\|_{\mathcal{B}} = \max_{k=1\ldots n} |\lambda_k|$. It is easily seen to be a norm. Now take any other norm N on V; we prove it is equivalent to $\|\cdot\|_{\mathcal{B}}$.

First, in the notation above,

$$N(v) \leq \sum_{k=1}^{n} N(\lambda_k b_k) = \sum_{k=1}^{n} |\lambda_k| N(b_k) \leq n \|v\|_{\mathcal{B}} \cdot \max_{k=1\ldots n} N(b_k).$$

Notice that $c_1 = n \max_{k=1\ldots n} N(b_k)$ is a fixed real number, so we have one desired inequality: $N(v) \leq c_1 \|v\|_{\mathcal{B}}$.

A converse inequality, viz. bounding $\|v\|_{\mathcal{B}}$ by a multiple of $N(v)$, requires some topology. Consider the set:

$$C = \left\{ (\lambda_1, \ldots, \lambda_k) \in \mathbb{K}^n : \max_{k=1\ldots n} |\lambda_k| = 1 \right\},$$

which is the unit sphere in $(\mathbb{K}^n, \|\cdot\|_\infty)$. Thanks to completeness of $\mathbb{K}$ recalled in theorem 1.1.3, one easily gets that C is compact, for instance through the sequential characterisation that every sequence has a converging subsequence. So move to:

$$\begin{array}{rccc} f\colon & C & \rightarrow & \mathbb{R}_{>0} \\ & (\lambda_1, \ldots, \lambda_n) & \mapsto & N(\sum_{k=1}^{n} \lambda_k b_k). \end{array}$$

This map is continuous and C is compact; by the extreme value theorem, f attains a minimum on C. Notice that f does not vanish on C. So there is $\varepsilon \in \mathbb{R}_{>0}$ such that

$(\forall(\lambda_1,\ldots,\lambda_n)\in C)(f(\lambda_1,\ldots,\lambda_n)\geq\varepsilon)$.

We finish the proof. Let $v\in V$. If $v=0$ then $\|v\|_{\mathcal{B}}=0$ and there is nothing to do. Otherwise write $v=\sum_{k=1}^n\lambda_k v_k$ and let:

$$\mu=\max_{k=1\ldots n}|\lambda_k|=\|(\lambda_1,\ldots,\lambda_n)\|_\infty=\|v\|_{\mathcal{B}}>0.$$

Also let $\lambda'_k=\frac{\lambda_k}{\mu}$, so now $\max_{k=1\ldots n}|\lambda'_k|=1$. This means that $(\lambda'_1,\ldots,\lambda'_k)\in C$. Therefore:

$$\begin{aligned}N\left(\frac{1}{\mu}v\right)&=N\left(\sum_{k=1}^n\frac{\lambda_k}{\mu}b_k\right)\\&=N(\sum_{k=1}^n\lambda'_k b_k)\\&=f(\lambda'_1,\ldots,\lambda'_n)\\&\geq\varepsilon,\end{aligned}$$

so that $N(v)\geq\varepsilon\mu=\varepsilon\|v\|_{\mathcal{B}}$ and we are done. □

Consequently, one may simply write $v_n\longrightarrow\ell$ without specifying the norm.

1.2.6. Remarks.

- This phenomenon is typical of finite-dimensional normed spaces. Notice however that it no longer holds in infinite dimension, typically in functional analysis; things are even worse in probability theory. Moreover, it only applies to norms (not to distances).
- In theory, all norms on a finite-dimensional real/complex-vector space are therefore equivalent. But in practice, one norm may be more useful, more adapted to a given problem. In the rest of the class we shall often adopt a norm for a proof, and use another one for another argument; at times we shall even *change norm during the argument*, relying on equivalence to patch paragraphs together. Always pause and think before opting for a norm.

We finally fulfill task $\mathbf{T}_2$, viz. extend Corollary 1.1.5 to matrix spaces.

1.2.7. Definition. Let V be a $\mathbb{K}$-vector space with a norm $\|\cdot\|$. A series $\sum v_n$ is *normally convergent* if the real series $\sum\|v_n\|$ is convergent.

Notice that two equivalent norms give rise to the same notion of normal convergence (not so in general).

1.2.8. Corollary. *Let V be a finite-dimensional $\mathbb{K}$-vector space with a norm $\|\cdot\|$. Then the unit sphere and the closed unit ball $\{v\in V:\|v\|\leq 1\}$ are compact.*

It follows of course that *any* closed ball is compact. (As a side remark, a normed vector space whose closed unit ball is compact is necessarily finite-dimensional, as asserted by a famous theorem by Riesz.)

Proof not covered in class.

Proof. The arguments are in the proof of theorem 1.2.5. Two equivalent norms give rise to the same notion of a compact set, so we may actually choose the norm. Fix a basis $\mathcal{B} = (b_1, \ldots, b_k)$ and decide to work with the norm $\| \cdot \|_{\mathcal{B}}$. We must prove that the unit sphere and the closed unit ball with respect to this norm are compact.

The map:

$$\begin{array}{cccc} \varphi\colon & \mathbb{K}^n & \to & V \\ & (\lambda_1, \ldots, \lambda_n) & \mapsto & \sum_{k=1}^n \lambda_k b_k \end{array}$$

is a vector space isomorphism, and actually more: it is an isometry between $(\mathbb{K}^n, \|\cdot\|_\infty)$ and $(V, \| \cdot \|_{\mathcal{B}})$. In particular it is a homeomorphism; of course the image of the unit sphere is the unit sphere, and likewise for the closed unit ball.

So we have reduced to the case of $(\mathbb{K}^n, \| \cdot \|_\infty)$; there compact is easily equivalent to closed and bounded, and both the unit sphere and the closed unit ball enjoy these properties. □

The desired analogue of Corollary 1.1.5, viz. solution to Task $\mathbf{T}_2$, is as follows.

1.2.9. Corollary. *Let V be a finite-dimensional $\mathbb{K}$-vector space with a norm $\| \cdot \|$. Then $(V, \| \cdot \|)$ is complete, meaning that all Cauchy sequences converge. In particular, if $\sum v_n$ is a normally convergent series, then $\sum v_n$ is convergent.*

Proof. Let (v_n) be a Cauchy sequence (with respect to $\| \cdot \|$); we must prove it converges. Every Cauchy sequence is bounded: return to the definition, with $\varepsilon = 1$. So there is $C > 0$ such that each v_n is in the closed ball with radius C, which is compact. Hence (v_n) has a converging subsequence. It is an exercise to prove that a Cauchy sequence with a converging subsequence is actually convergent.

The implication 'normally convergent ⇒ convergent' follows the proof of Corollary 1.1.5. □

1.2.10. Remark. The proper study of norms in infinite-dimensional vector spaces is carried in functional analysis, a fascinating topic. One usually recommends Rudin's *Functional analysis*; as a student I immensely liked Brezis' book.

1.3 Matrix norms

Matrix spaces, being finite-dimensional, have only one normed topology. But depending on the topic some norms come handier than others. We finally handle the last task on our list from § 1.1.8:

$\mathbf{T}_3$. find a substitute for multiplicativity of $|\cdot|$, which may no longer hold with matrices.

Suppose that a normed $\mathbb{K}$-vector space V bears a bilinear multiplication $\cdot\colon V \times V \to V$; it is then called a $\mathbb{K}$-*algebra*, and this is typically the case with matrix spaces. Then one may ask about the relationship between multiplication and the norm. This gives rise to an important notion.

1.3.1. Definition. Suppose V is a $\mathbb{K}$-algebra. A norm $\|\cdot\|$ is *submultiplicative* if it satisfies:

$$(\forall v_1, v_2 \in V)(\|v_1 \cdot v_2\| \leq \|v_1\| \cdot \|v_2\|).$$

We shall prove that *on matrix algebras there always exist submultiplicative norms.*

1.3.2. Definition. Let $\|\cdot\|$ be a norm on $\mathbb{K}^n$. For $A \in M_n(\mathbb{K})$ let:

$$|\!|\!|A|\!|\!| = \max_{\|X\|=1} \|AX\|,$$

which is well-defined by compactness of the unit sphere in $(\mathbb{K}^n, \|\cdot\|)$ obtained in Corollary 1.2.8. It is easily seen to be a norm, called the *operator norm* associated to $\|\cdot\|$.

1.3.3. Properties. *Let $\|\cdot\|$ be a norm on $\mathbb{K}^n$ and $|\!|\!|\cdot|\!|\!|$ be the associated operator norm. Then:*

(i) $(\forall A \in M_n(\mathbb{K}))(\forall X \in \mathbb{K}^n)(\|AX\| \leq |\!|\!|A|\!|\!| \cdot \|X\|)$;

(ii) $(\forall A, B \in M_n(\mathbb{K}))(|\!|\!|A \cdot B|\!|\!| \leq |\!|\!|A|\!|\!| \cdot |\!|\!|B|\!|\!|)$, *viz.* $|\!|\!|\cdot|\!|\!|$ *is submultiplicative.*

Proof.

(i) If $X = 0$ there is nothing to prove. Otherwise let $Y = \frac{1}{\|X\|}X$, so that $\|Y\| = 1$. Then by definition $\|AY\| \leq |\!|\!|A|\!|\!|$, so multiplying, $\|AX\| \leq |\!|\!|A|\!|\!|\|X\|$.

(ii) Let X have norm 1 and put $X' = BX$. By the above, we find:

$$\|ABX\| = \|AX'\| \leq |\!|\!|A|\!|\!| \cdot \|X'\| \leq |\!|\!|A|\!|\!| \cdot |\!|\!|B|\!|\!| \cdot \|X\| = |\!|\!|A|\!|\!| \cdot |\!|\!|B|\!|\!|.$$

Taking the maximum over the sphere, we get $|\!|\!|AB|\!|\!| \leq |\!|\!|A|\!|\!| \cdot |\!|\!|B|\!|\!|$, as desired. □

1.3.4. Remark. There exist of course matrix norms which are *not* submultiplicative; and they can be useful as well. For instance, $\|A\|_\infty = \max_{i,j=1\ldots n} |a_{i,j}|$ is often very handy; you may check it is not submultiplicative.

1.4 Exercises

1.4.1. Exercise.

1. *Prove* $(\forall a, b \in \mathbb{R}_{\geq 0})(\exp(a+b) = \exp(a) \cdot \exp(b))$. *Deduce the same for complex numbers. Hint: write partial sums up to n; see a square and a triangle as indexing sets.*

2. *Prove* $(\forall a \in \mathbb{R}_{\geq 0})(\lim_{n\to\infty}(1 + \frac{a}{n})^n = \exp(a))$. *Deduce the same for a complex number. Hint: have you tried logarithms? (In the complex case, one needs to know that there is a partial* $\log$ *function defined on a small neighbourhood of* $1 \in \mathbb{C}$*.)*

Solution.

1. Consider partial sums $E_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$.

   On the one hand,

$$E_n(a) \cdot E_n(b) = \left(\sum_{i=0}^n \frac{a^i}{i!}\right) \cdot \left(\sum_{j=0}^n \frac{b^j}{j!}\right) = \sum_{0\leq i,j\leq n} \frac{a^i b^j}{i!j!};$$

on the other hand,

$$E_n(a+b) = \sum_{k=0}^{n} \frac{(a+b)^k}{k!} = \sum_{k=0}^{n}\sum_{i=0}^{k} \frac{1}{k!}\binom{k}{i} a^i b^{k-i}$$
$$= \sum_{\substack{0\leq i,j\leq n:\\ i+j\leq n}} \frac{a^i b^j}{i!j!}.$$

We are summing the same terms, but one index set is bigger than the other. Let us give them names. Let:

- $S_n = \{(i,j) \in \{0,\dots,n\}^2\}$, which is a *square*;
- $T_n = \{(i,j) \in \{0,\dots,n\}^2 : i+j \leq n\}$, which is a triangle.

Hence $E_n(A) \cdot E_n(B) = \sum_{(i,j)\in S_n} C_{i,j}$ while $E_n(A+B) = \sum_{(i,j)\in T_n} C_{i,j}$. Now clearly $S_n \subseteq T_{2n} \subseteq S_{2n}$.

First suppose $a, b \in_b R_{\geq 0}$. Since all terms are non-negative, we get:

$$E_n(a)\cdot E_n(b) \leq E_{2n}(a+b) \leq E_{2n}(a)\cdot E_{2n}(b).$$

Both the leftmost and rightmost members go to $\exp(a)\cdot\exp(b)$, while the middle one goes to $\exp(a+b)$: we are done.

Now to the complex case. By the triangle inequality,

$$|E_n(a)\cdot E_n(b) - E_n(a+b)| = \left| \sum_{(i,j)\in S_n\setminus T_n} \frac{a^i b^j}{i!j!} \right|$$
$$\leq \sum_{(i,j)\in S_n\setminus T_n} \frac{|a|^i |b|^j}{i!j!}$$
$$= E_n(|a|)\cdot E_n(|b|) - E_n(|a|+|b|),$$

which goes to 0 by the first case. Therefore $|E_n(a)\cdot E_n(b) - E_n(a+b)| \longrightarrow 0$; by continuity of field operations, this gives $\exp(a)\cdot\exp(b) = \exp(a+b)$.

2. Suppose $a \in \mathbb{R}_{\geq 0}$. Recalling something about the log function, we have $1 + \frac{a}{n} = \exp(\log(1+\frac{a}{n})$ and therefore:

$$\left(1+\frac{a}{n}\right)^n = \exp\left(n\log\left(1+\frac{a}{n}\right)\right) = \exp\left(n\left(\frac{a}{n} + o\left(\frac{1}{n}\right)\right)\right) = \exp(a + o(1)).$$

By definition, $a + o(1) \xrightarrow[n\to\infty]{} a$. By continuity of exp (something we did not prove but enough analysis!), the sequence goes to $\exp(a)$.

Oddly enough, one can still use the logarithm trick in the complexe case. This is because $1 + \frac{a}{n} \longrightarrow 1$, and therefore remains in a small neighbourhood of 1. In complex analysis one can define a partial log on such a neighbourhood; what matters is to avoid $\mathbb{R}_{\leq 0}$.

1.4.2. Exercise. *Let V be a finite-dimensional $\mathbb{K}$-vector space equipped with a norm. An* accumulation point *of a sequence (v_n) is some $\ell \in V$ such that there is a subsequence converging to ℓ, viz. there exists an increasing map $\varphi\colon \mathbb{N} \to \mathbb{N}$ with $v_{\varphi(n)} \longrightarrow \ell$.*

1. *Prove that a Cauchy sequence with an accumulation point is convergent.*
2. *Prove that a bounded sequence with at most one accumulation point is convergent.*

Solution.

1. Let (v_n) be a Cauchy sequence; suppose $v_{\varphi(n)} \longrightarrow \ell$ for some extraction function $\varphi\colon \mathbb{N} \to \mathbb{N}$. Let $\varepsilon \in \mathbb{R}_{>0}$. Since the sequence is Cauchy, there is $n_1 \in \mathbb{N}$ such that for all integers p, q one has: $p \geq q \geq n_1 \to \|v - p - v_q\| < \frac{\varepsilon}{2}$. Also, there is $n_2 \in \mathbb{N}$ such that for any integer n, one has: $n \geq n_2 \to \|v_n - \ell\| < \frac{\varepsilon}{2}$. Let $n_0 = \max(n_1, n_2)$. Then for any $n \geq n_0$, one has $\varphi(n) \geq n \geq n_0$, so:

$$\|v_n - \ell\| \leq \|v_n - v_{\varphi(n)}\| + \|v_{\varphi(n)} - \ell\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

   proving convergence to ℓ.

2. This is more interesting. Let (v_n) be bounded. Then the sequence remains in some closed ball, which is compact; hence (v_n) has at least one accumulation point, say ℓ. We claim that $v_n \longrightarrow \ell$. Indeed, suppose that it is *not* the case. Then there is $\varepsilon \in \mathbb{R}_{>0}$ such that:

$$(\forall n_0 \in \mathbb{N})(\exists n \in \mathbb{N})(n \geq n_0 \wedge \|v_n - \ell\| \geq \varepsilon).$$

   For each n_0 we let $\varphi(n_0)$ be the least integer $n > \varphi(n_0 - 1)$ with the property. This defines an extraction function φ such that the norm $\|v_{\varphi(n)} - \ell|$ remains $\geq \varepsilon$. But the subsequence $(v_{\varphi(n)})$ remains bounded, so it has an accumulation point μ. But μ is then also an accumulation point of (v_n), hence $\mu = \ell$. Taking limits in $\|v_{\varphi(n)} - \ell\| \geq \varepsilon$ we find $0 = \|\mu - \ell\| \geq \varepsilon$, a contradiction.

Notice that 2. no longer holds if closed balls cease to be compact, viz. in an infinite-dimensional space; as a matter of fact, in functional analysis there are bounded sequences with no accumulation points at all.

1.4.3. Exercise. *Let $\|\cdot\|$ be any matrix norm. Prove that there is $c \in \mathbb{R}_{>0}$ such that for all matrices A, B one has $\|A \cdot B\| \leq c \cdot \|A\| \cdot \|B\|$.*

Solution. Fix one submultiplicative norm N; by equivalence, there is $c \in \mathbb{R}_{>0}$ such that for all matrices, $N(A) \leq c \cdot \|A\|$ and $\|A\| \leq N(A)$. Then for any pair of matrices:

$$\|A \cdot B\| \leq cN(AB) = cN(A)N(B) \leq c^3 \cdot \|A\| \cdot \|B\|,$$

so constant c^3 does the job.

1.4.4. Exercise. *Prove that if $n > 1$ there are no multiplicative norms on $M_n(\mathbb{K})$ (viz. satisfying $N(AB) = N(A)N(B)$ for all matrices).*

Solution. This is because $M_n(\mathbb{K})$ has zero divisors: there exist $A, B \neq 0$ with $AB = 0$, e.g. $A = B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. If a norm were multiplicative it would have $\|A\|^2 = \|A^2\| = \|0\| = 0$, whence $\|A\| = 0$ and $A = 0$: a contradiction.

2 The matrix exponential and first properties

Last reminder: throughout, $\mathbb{K}$ stands for $\mathbb{R}$ or $\mathbb{C}$. We define the matrix exponential as expected in § 2.1; this results in more question. First algebraic properties ($\exp(A)$ is a polynomial in A, invariance under conjugacy, the sum property) are discussed in § 2.2. The derivative of $t \mapsto \exp(tA)$ is computed in § 2.3, but will not play a role before § 6.

Throughout there will be some notational ambiguity: we shall work with sequences or series of matrices. Of course we want to index our sequences by n; of course we also want our matrices to have size $n \times n$. We shall often write 'let $(A_n) \in (M_n(\mathbb{K}))^{\mathbb{N}}$ be a sequence of matrices', and leave it to the reader to decide with n stands for the (fixed) dimension, and which n for the (varying) sequence index. Despite our efforts, this results in no confusion at all.

Certainly the notion of a diagonalisable matrix is familiar. The *spectrum* of a matrix M (viz. the set of its eigenvalues) is denoted by $\mathrm{Sp}(M)$.

2.1 The definition

Having fulfilled tasks $\mathbf{T}_1$–$\mathbf{T}_3$ of 1.1.8 we are now ready to adapt the formalisation of the complex exponential (theorem 1.1.6 of § 1.1) to the matrix case.

2.1.1. Lemma (and definition). *Let $A \in M_n(\mathbb{K})$. Then the series $\sum \frac{A^n}{n!}$ converges to a matrix denoted by* $\exp A$ *and called the* exponential *of A.*

Proof. By § 1.2, more specifically Corollary 1.2.9, it suffices to prove that the series is normally convergent for any norm (since all are equivalent by theorem 1.2.5). We fix a submultiplicative norm; these exist by § 1.3. Then $\left\| \frac{A^n}{n!} \right\| \leq \frac{\|A\|^n}{n!}$. But $\sum \frac{\|A\|^n}{n!}$ is convergent as we saw when proving theorem 1.1.6, so $\sum \frac{A^n}{n!}$ is normally convergent: we are done. □

2.1.2. Questions. The definition raises a number of questions.

Q$_1$. What are the properties of the matrix exponential?

Q$_2$. How to compute $\exp(A)$ in practice?

Q$_3$. What are applications of $\exp(A)$ to differential equations?

2.1.3. Remarks.

- On **Q**$_1$—it is not hard (exercise 2.4.5) to see that the matrix exponential is not injective, unless $n = 1$ and $\mathbb{K} = \mathbb{R}$.

  In particular, there can only be partially defined 'matrix logarithms'. We shall see three natural subsets $\mathcal{S} \subseteq M_n(\mathbb{K})$ such that exp restricts to a bijection $\mathcal{S} \simeq \exp(\mathcal{S})$:

  - $\mathcal{S}$ is the set of nilpotent matrices and $\exp(\mathcal{S})$ the set of 'unipotent' matrices (§ 3.3);
  - $\mathcal{S}$ is some small neighbourhood of 0 and $\exp(\mathcal{S})$ some small neighbourhood of I_n (§ 4.2);
  - $\mathcal{S}$ is the space of hermitian matrices and $\exp(\mathcal{S})$ the set of hermitian definite positive matrices (§ 5.2).

In each case, one could technically define a (partial) logarithm.

- $\mathbf{Q_2}$ is addressed in § 3. So far, only the exponential of a *diagonal* matrix is easy to compute:

$$\exp\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} = \begin{pmatrix}e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n}\end{pmatrix}.$$

- $\mathbf{Q_3}$ is returned to in § 6.

The rest of the section is devoted to attacking questions $\mathbf{Q_1}$ and $\mathbf{Q_2}$; only with more tools shall we later solve some more difficult aspects. Before we start we introduce a general notation.

2.1.4. Notation. Let $E_n(A) = \sum_{k=0}^{n} \frac{A^k}{k!}$ be the partial sum up to n, so that $E_n(A) \longrightarrow \exp(A)$.

2.2 Algebraic properties

Nilpotent matrices; local polynomials

Recall the definition of a nilpotent matrix.

2.2.1. Definition. A matrix $N \in M_n(\mathbb{K})$ is *nilpotent* if there is $k \geq 0$ with $N^k = 0$.

2.2.2. Lemma. *The following are equivalent:*

(i) N *is nilpotent;*

(ii) $N^n = 0$;

(iii) $\operatorname{Sp} N = \{0\}$.

Proof. One must remember that if a matrix T is strictly upper-triangular, viz. has the form:

$$T = \begin{pmatrix}0 & & * \\ & \ddots & \\ & & 0\end{pmatrix},$$

then $T^n = 0$. This is easily seen by computing successive powers of T (and better seen by trials than by induction): in T^k, the diagonal, super-diagonal, … k^{th} super-diagonal, are all 0.

In particular, if $\operatorname{Sp} N = \{0\}$, then conjugating N in triangular form, one get $N^n = 0$, which implies nilpotence. It remains to show that nilpotence implies $\operatorname{Sp} = \{0\}$. Suppose N is nilpotent, and take $k \in \mathbb{N}$ with $N^k = 0$. Let $\lambda \in \operatorname{Sp}(N)$ and $X \in \mathbb{K}^n \setminus \{0\}$ witness it. Then $N^k \cdot X = \lambda^k X = 0$, so $\lambda^k = 0$, which implies $\lambda = 0$. □

2.2.3. Remark. Let N be nilpotent. Then $\exp(N) = \sum_{k=0}^{n-1} \frac{N^k}{k!}$ is a finite sum, hence a polynomial in N.

We can therefore compute the exponential of nilpotent matrices. Surprisingly, although we cannot compute $\exp(A)$ in general, it always is a polynomial in A. We do *not* assume A to be nilpotent.

2.2.4. Lemma ($\exp(A)$ is a polynomial in A)**.** *For every $A \in M_n(\mathbb{K})$, there is a polynomial $P_A(X) \in \mathbb{K}[X]$ such that* $\exp A = P_A(A)$.

Of course $P_A(X)$ depends on A.

Proof. Let $\mathbb{K}[A] = \langle I_n, A, A^2, \dots \rangle$ be the subspace generated by powers of A. (Parenthetically, by the Cayley-Hamilton theorem, one may stop before power $A^n \in \langle I_n, \dots, A^{n-1}\rangle$.) Notice that $\mathbb{K}[A]$ is exactly the set of matrices which are polynomials in A.

Since $\mathbb{K}[A]$ is a linear subspace of finite-dimensional $M_n(\mathbb{K})$, it is a closed subset. Now for all $n \in \mathbb{N}$ one has $E_n(A) \in \mathbb{K}[A]$. Taking the limit and by closedness, we find $\exp(A) \in \mathbb{K}[A]$, as desired. □

Lemma 2.2.4 implies that $A \cdot \exp(A) = \exp(A) \cdot A$; however since the polynomial depends on A, it offers no practical way to compute $\exp(A)$.

Invariance under conjugacy; diagonalisable matrices

The following is the key of many arguments.

2.2.5. Lemma (invariance under conjugacy)**.** *Suppose P is invertible. Then:*

$$\exp(PAP^{-1}) = P \cdot \exp(A) \cdot P^{-1}.$$

Proof. Let n be fixed. Then:

$$E_n(PAP^{-1}) = \sum_{k=0}^{n} \frac{(PAP^{-1})^k}{k!} = \sum_{k=0}^{n} \frac{PA^kP^{-1}}{k!} = P \cdot E_n(A) \cdot P^{-1}.$$

Now let $n \to \infty$. The left-hand goes to $\exp(PAP^{-1})$. By continuity of multiplication, the right-hand goes to $P \cdot \exp(A) \cdot P^{-1}$. □

2.2.6. Remark. We already knew how to compute the exponential of a diagonal matrix (take the exponential of the diagonal); by conjugacy we can now compute $\exp(A)$ for any *diagonalisable* A. Indeed, write $A = PDP^{-1}$ with invertible P and diagonal D; then $\exp(A) = P\exp(D)P^{-1}$.

2.2.7. Example. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$; we compute $\exp(A)$. The matrix has obvious eigenvectors. Let $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, for which $A = P\begin{pmatrix} 1 & \\ & 2 \end{pmatrix}P^{-1}$. Therefore:

$$\begin{aligned}\exp(A) &= \exp\left(P\begin{pmatrix} 1 & \\ & 2 \end{pmatrix}P^{-1}\right) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} e & \\ & e^2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} e & -e \\ 0 & e^2 \end{pmatrix} = \begin{pmatrix} e & e - e^2 \\ 0 & e^2 \end{pmatrix}.\end{aligned}$$

The sum property

Lured by the complex case one could hope that exp realises a group morphism from $(M_n(\mathbb{K}), +)$ to $(\mathrm{GL}_n(\mathbb{K}), \cdot)$. This is however not the case. In general, $\exp(A + B) \neq \exp(A) \cdot \exp(B)$.

2.2.8. Example. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

- One sees: $(\forall n \in \mathbb{N})(n \geq 2 \to A^n = 0)$, so it is trivial to check $\exp(A) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.
- One also sees: $(\forall n \in \mathbb{N})(n \geq 1 \to B^n = B)$. So the series is easily computed, and we find $\exp(B) = \begin{pmatrix} e & 0 \\ 0 & 0 \end{pmatrix}$.
- Turn to $A + B$; see that $(\forall n \in \mathbb{N})(n \geq 1 \to (A + B)^n = A + B)$. Therefore $\exp(A + B) = \begin{pmatrix} e & e \\ 0 & 0 \end{pmatrix}$.
- However, $\exp(A) \cdot \exp(B) = \begin{pmatrix} e & 0 \\ 0 & 0 \end{pmatrix} \neq \exp(A + B)$.

2.2.9. Proposition (the sum property). *If A and B commute (viz. $AB = BA$) then* $\exp(A+B) = \exp(A) \cdot \exp(B) = \exp(B) \cdot \exp(A)$. *In particular,* exp *takes values in* $\mathrm{GL}_n(\mathbb{K})$.

Proof. This is a special case of a general phenomenon which is worth returning to. In the following computations the subscripts (index sets) are more important than what we are actually summing. We fix a submultiplicative norm.

Bear in mind notation $E_n(A) = \sum_{k=0}^n \frac{A^k}{k!}$. So typically,

$$E_n(A) \cdot E_n(B) = \left(\sum_{i=0}^n \frac{A^i}{i!} \right) \cdot \left(\sum_{j=0}^n \frac{B^j}{j!} \right) = \sum_{0 \leq i, j \leq n} \frac{1}{i!j!} A^i B^j.$$

On the other hand A and B commute, so by Newton's expansion one has for any $k \geq 0$:

$$(A + B)^k = \sum_{i=0}^k \binom{k}{i} A^i B^{k-i} = \sum_{\substack{0, \leq i, j \leq k: \\ i+j=k}} \frac{k!}{i!j!} A^i B^j,$$

and therefore:

$$E_n(A + B) = \sum_{k=0}^n \frac{(A + B)^k}{k!} = \sum_{k=0}^n \sum_{\substack{0, \leq i, j \leq k: \\ i+j=k}} \frac{1}{i!j!} A^i B^j = \sum_{\substack{i, j=0 \ldots n: \\ i+j \leq n}} \frac{1}{i!j!} A^i B^j.$$

We are summing the same terms $C_{i,j} = \frac{1}{i!j!} A^i B^j$, though over different index sets. We shall denote these by:

- $S_n = \{(i, j) \in \{0, \ldots, n\}^2\}$, which is a *square*;
- $T_n = \{(i, j) \in \{0, \ldots, n\}^2 : i + j \leq n\}$, which is a triangle.

Hence $E_n(A) \cdot E_n(B) = \sum_{(i,j)\in S_n} C_{i,j}$ while $E_n(A+B) = \sum_{(i,j)\in T_n} C_{i,j}$.

Thus:

$$\begin{aligned}
\|E_n(A) \cdot E_n(B) - E_n(A+B)\| &= \left\| \sum_{(i,j)\in S_n \smallsetminus T_n} C_{i,j} \right\| \\
&\leq \sum_{(i,j)\in S_n \smallsetminus T_n} \|C_{i,j}\| \\
&\leq \sum_{(i,j)\in S_n \smallsetminus T_n} \frac{\|A\|^i \|B\|^j}{i!j!} \\
&= \sum_{(i,j)\in S_n} \frac{\|A\|^i \|B\|^j}{i!j!} - \sum_{(i,j)\in T_n} \frac{\|A\|^i \|B\|^j}{i!j!} \\
&= E_n(\|A\|) \cdot E_n(\|B\|) - E_n(\|A\| + \|B\|) \\
&\longrightarrow \exp(\|A\|) \cdot \exp(\|B\|) - \exp(\|A\| + \|B\|).
\end{aligned}$$

The latter is 0 since $\exp\colon (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$ is a group morphism.

In particular, since A and $-A$ always commute, one finds $\exp(A) \cdot \exp(-A) = \exp(-A) \cdot \exp(A) = \exp(0) = I_n$, so $\exp(A)$ is invertible, with inverse $\exp(-A)$. □

2.3 Analytic properties

2.3.1. Lemma. *Let $A \in M_n(\mathbb{C})$. Then the map*

$$\begin{array}{rccc} f\colon & \mathbb{R} & \to & \mathrm{GL}_n(\mathbb{C}) \\ & t & \mapsto & \exp(tA) \end{array}$$

is differentiable on $\mathbb{R}$, satisfying $f'(t) = A \cdot f(t) = f(t) \cdot A$. (In particular, it is continuous.)

Proof. We hope the reader is familiar with Hardy's notation $o(\cdot)$. Notice that one may use $o(\cdot)$ for vector (or even matrix) functions: a norm $\|\cdot\|$ being fixed, write $X(h) = o(h)$ if $\|X(h)\| = o(h)$ in the real sense. Notice that this does not depend on the norm as soon as all are equivalent.

Fix $t_0 \in \mathbb{R}$; letter h will denote a small real number. Observe that the matrices $t_0 A$ and hA commute, so by the sum property (proposition 2.2.9):

$$f(t_0 + h) = \exp(t_0 A + hA) = \exp(t_0 A) \cdot \exp(hA) = f(t_0) \cdot \exp(hA).$$

Now $\exp(hA) = I_n + hA + R(h)$ where $R(h) = \sum_{n\geq 2} \frac{h^n A^n}{n!}$. In particular, for $|h| \leq 1$ and choosing a submultiplicative norm, one has $\|R(h)\| \leq h^2 \exp(\|A\|)$, so $R(h) \in o(h)$.

Therefore $f(t_0 + h) = f(t_0)(I_n + hA + o(h)) = f(t_0) + hf(t_0)A + o(h)$, which is a Taylor expansion of order 1. This proves that f is differentiable at t_0, with derivative $f(t_0) \cdot A$. □

This formula is generalised in § 4.1, where we treat exp as a multi-variable function.

2.4 Exercises

2.4.1. Exercise. *Let $\|\cdot\|$ be a submultiplicative norm on $M_n(\mathbb{K})$. Prove that $\|\exp(A)\| \leq \exp(\|A\|)$.*

Solution. This is easy. For fixed n one has:

$$\|E_n(A)\| = \left\|\sum_{k=0}^{n} \frac{A^k}{k!}\right\| \leq \sum_{k=0}^{n} \frac{\|A\|^k}{k!} = E_n(\|A\|).$$

By continuity of the norm function, the left-hand goes to $\|\exp(A)\|$; by definition, the right-hand goes to $\exp(\|A\|)$.

2.4.2. Exercise. *Suppose that* $\exp(A) = I_n$*; prove that* $\operatorname{Sp}(A) \subseteq 2i\pi\mathbb{Z}$*. Does the converse hold?*

Solution. If $\exp(A) = I_n$ then trigonalising A, we see that every eigenvalue λ of A must satisfy $e^\lambda = 1$, so $\lambda \in 2i\pi\mathbb{Z}$.

The converse is not true: consider matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, with spectrum $\{0\} \subseteq 2i\pi\mathbb{Z}$. Then $A^2 = 0$ so $\exp(A) = I_2 + A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \neq I_2$.

2.4.3. Exercise. *Prove that N is nilpotent iff for all k* $= 0 \dots n$ *one has* $\operatorname{tr}(N^k) = 0$.

Solution. I would presumably use an ill-named 'Vandermonde' determinant (Vandermonde *never* discussed them).

2.4.4. Exercise. *Prove that* $\left(I + \frac{A}{n}\right)^n \xrightarrow[n\to\infty]{} \exp(A)$*. (You may use the analogue property holding in real numbers.)*

Solution. Fix a submultiplicative norm and compute:

$$\begin{aligned}
\left\|E_n(A) - \left(I + \frac{A}{n}\right)^n\right\| &\leq \sum_{k=0}^{n} \left\|\frac{A^k}{k!} - \binom{n}{k}\frac{A^k}{n^k}\right\| \\
&\leq \sum_{k=0}^{n} \left|\frac{1}{k!} - \binom{n}{k}\frac{1}{n^k}\right| \|A\|^k.
\end{aligned}$$

Let $r_{n,k} = \frac{1}{k!} - \binom{n}{k}\frac{1}{n^k}$ be the real number appearing in the modulus. Notice how:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!},$$

so $r_{n,k} \geq 0$. Hence actually:

$$\begin{aligned}
\left\|E_n(A) - \left(I + \frac{A}{n}\right)^n\right\| &\leq \sum_{k=0}^{n} r_{n,k}\|A\|^k \\
&= E_n(\|A\|) - \left(I + \frac{\|A\|}{n}\right)^n.
\end{aligned}$$

But the latter goes to 0 as $n \to \infty$, since we know $\left(1 + \frac{x}{n}\right)^n \xrightarrow[n\to\infty]{} e^x$ for real x.

One can almost retrieve the sum property from this formula, since for commuting matrices:

$$\left[\left(I + \frac{A}{n}\right)\left(I + \frac{B}{n}\right)\right] = \left(I + \frac{A+B}{n} + \frac{AB}{n^2}\right)^n.$$

However in order to prove convergence of the latter to $\exp(A+B)$ one needs a little more.

2.4.5. Exercise. *Prove that if $n \geq 2$ then $\exp\colon M_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ is* not *injective. (Surjectivity will be dealt with in § 3.3; the answer depends on whether $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.)*

Solution. It is enough to treat the case $n = 2$. Consider the map:

$$\begin{array}{rccc} \varphi\colon & \mathbb{C} & \to & M_2(\mathbb{R}) \\ & a + ib & \mapsto & \begin{pmatrix} a & -b \\ b & a \end{pmatrix}. \end{array}$$

It clearly is injective; it also preserves both the additive and multiplicative structures: it is an embedding of $\mathbb{R}$-algebras. Consequently, for any $a + ib \in \mathbb{C}$ and $n \in \mathbb{N}$:

$$\varphi(E_n(a+ib)) = E_n(\varphi(a+ib)).$$

One needs to know that φ is continuous (which is obvious); therefore taking limits, one obtains $\varphi(\exp(a+ib)) = \exp(\varphi(a+ib))$, or more accurately.

So consider matrix $A = \begin{pmatrix} 0 & -2\pi \\ 2\pi & 0 \end{pmatrix} = \varphi(2i\pi)$. We just saw $\exp(A) = \exp\circ\varphi(2i\pi) = \varphi\circ\exp(2i\pi) = \varphi(1) = I_2$, while $A \neq 0$: injectivity is lost.

2.4.6. Exercise. *Compute* $\exp\begin{pmatrix} 0 & -x \\ x & 0 \end{pmatrix}$ *and* $\exp\begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix}$ *for real x.*

Solution. Let $A_x = \begin{pmatrix} 0 & -x \\ x & 0 \end{pmatrix}$. We can use the above isomorphism $\varphi\colon \mathbb{C} \simeq M_2(\mathbb{R})$ as real algebras to see that $\exp A_x = \varphi(e^{ix}) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$. Now let $B_x = \exp\begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix}$. The complex trick no longer works. But $B_x^2 = x^2 I_2$, then $B_x^3 = B_{x^3}$, and so on. So we can actually compute the series and find:

$$\exp B_x = \begin{pmatrix} 1 + \frac{x^2}{2} + \frac{x^4}{4!} + \dots & x + \frac{x^3}{3!} + \dots \\ x + \frac{x^3}{3!} + \dots & 1 + \frac{x^2}{2} + \frac{x^4}{4!} + \dots \end{pmatrix} = \begin{pmatrix} \cosh x & \sinh x \\ \sinh x & \cosh x \end{pmatrix}.$$

2.4.7. Exercise. *Let* $B = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 3/2 \end{pmatrix}$. *Find* $\exp(B)$.

Solution. The determinant is 1 and trace is 2; clearly B has eigenvalue 1 with multiplicity 2. The eigenspace is:

$$E_1(B) = \ker(B - I) = \ker\begin{pmatrix} -1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix} = \ker\begin{pmatrix} 1 & -1 \end{pmatrix} = \operatorname{Span}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

So let $Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$; it can be checked that $Q^{-1}BQ = \begin{pmatrix} 1 & -1/2 \\ & 1 \end{pmatrix}$. Of course:

$$\begin{aligned} \exp\begin{pmatrix} 1 & -1/2 \\ & 1 \end{pmatrix} &= \exp\left(\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1/2 \\ & 0 \end{pmatrix}\right) = \exp\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \cdot \exp\begin{pmatrix} 0 & -1/2 \\ & 0 \end{pmatrix} \\ &= \begin{pmatrix} e & \\ & e \end{pmatrix} \cdot \begin{pmatrix} 1 & -1/2 \\ & 1 \end{pmatrix} = \begin{pmatrix} e & -1/2e \\ & e \end{pmatrix}, \end{aligned}$$

and therefore:

$$\begin{aligned}\exp(B) &= \exp\left(Q\begin{pmatrix}1 & -1/2\\ & 1\end{pmatrix}Q^{-1}\right) = \begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\cdot\begin{pmatrix}e & -1/2e\\ & e\end{pmatrix}\cdot\begin{pmatrix}0 & 1\\ 1 & -1\end{pmatrix}\\ &= \begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\cdot\begin{pmatrix}-1/2e & 3/2e\\ e & -e\end{pmatrix} = \begin{pmatrix}e/2 & e/2\\ -e/2 & 3/2e\end{pmatrix}.\end{aligned}$$

2.4.8. Exercise. *Prove that* $\det\exp A = \exp\operatorname{tr} A$.

Solution. Trigonalise.

2.4.9. Exercise. *Suppose* $(\forall \in \mathbb{R})(\exp(tA) = \exp(tB))$. *Prove that* $A = B$.

Solution. Differentiate at $t = 0$.

2.4.10. Exercise. *There are direct analytic proofs that* $\exp(tA)' = A \cdot \exp(tA)$ *not using the sum property, proposition 2.2.9 (using normal convergence of function series instead).*

The Cauchy-Lipschitz theorem, also known as the Picard-Lindelöf theorem, implies that an affine differential equation of order 1 with given initial condition has exactly one (global) solution.

Use this to prove the sum property: if $AB = BA$ *then* $\exp(A + B) = \exp(A) \cdot \exp(B)$.

Solution. Let $f(t) = \exp(t(A + B))$ and $g(t) = \exp(tA) \cdot \exp(tB)$. Both are differentiable by usual arguments; let us compute derivatives:

$$f'(t) = (A + B)\exp(t(A + B)) = (A + B)f(t)$$

and, bearing in mind that B and $\exp(tA) \in \mathbb{K}[A]$ commute,

$$\begin{aligned}g'(t) &= A\exp(tA) \cdot \exp(tB) + \exp(tA) \cdot B\exp(tB)\\ &= (A + B) \cdot \exp(tA)\exp(tB)\\ &= (A + B)g(t)\end{aligned}$$

So both f and g are solutions to the same differential equation $M'(t) = (A + B)M(t)$. Now $f(0) = g(0) = I_n$, so the Cauchy-Lipschitz theorem implies $f(t) = g(t)$. In particular $t = 1$ gives the desired formula.

3 The Chevalley '$D + N$' decomposition

We return to the second question from 2.1.2.

Q$_2$. How to compute $\exp(A)$ in practice?

We already know what to do if A is nilpotent, or presented as PDP^{-1} with diagonal D.

The Chevalley decomposition $A = D + N$ is a useful additive decomposition of a matrix into a sum of two commuting parts, one diagonalisable and one nilpotent; it is unique. This is of course used in conjunction with the sum property, so $\exp(A) = \exp(D) \cdot \exp(N)$. But the interest of the Chevalley decomposition goes beyond the matrix exponential.

We give a key tool in § 3.1; the decomposition itself is in § 3.2 and can be used to prove surjectivity *in the complex case* (§ 3.3). Last, § 3.4 deals with practical computation of the decomposition; it is harder and optional to read.

Before reading this section make sure you understand that 'diagonalisable' merely means 'conjugate to a diagonal matrix by an invertible matrix', and that you have some training in eigenspaces and invariant subspaces.

3.1 Coprime kernel lemma

The following is one of the most useful lemmas in linear algebra.

3.1.1. Lemma (coprime kernel lemma). *Let $P, Q \in \mathbb{K}[X]$ be coprime polynomials and $A \in M_n(\mathbb{K})$ be a matrix with $(PQ)(A) = 0$. Then $\mathbb{K}^n = \ker P(A) \oplus \ker Q(A)$. Moreover, the projectors onto* $\ker P(A)$ *parallel to* $\ker Q(A)$, *and vice-versa, are polynomials in A.*

Proof. By Bézout's theorem on coprime polynomials, there exist polynomials $U, V \in \mathbb{K}[X]$ with $PU + QV = 1$. Apply the identity to A, getting $P(A) \cdot U(A) + Q(A) \cdot V(A) = I_n$.

Let $X \in \ker P(A) \cap \ker Q(A)$. Then in particular,

$$\begin{aligned} X &= I_n \cdot X \\ &= (P(A) \cdot U(A) + Q(A) \cdot V(A)) \cdot X \\ &= U(A) \cdot (P(A) \cdot X) + V(A) \cdot (Q(A) \cdot X) \\ &= 0, \end{aligned}$$

proving that the sum is direct.

Now let $X \in \mathbb{K}^n$. Put $X_Q = (P(A)U(A)) \cdot X$ and $X_P = (Q(A)V(A)) \cdot X$; we just noted $X = X_P + X_Q$. Moreover $Q(A) \cdot X_Q = Q(A)P(A) \cdot U(A)X = 0$ so $X_Q \in \ker Q(A)$, and $X_P \in \ker P(A)$ likewise. So $\mathbb{K}^n = \ker P(A) + \ker Q(A) = \ker P(A) \oplus \ker Q(A)$. Finally, X_P is the projection onto $\ker P(A)$ parallel to $Q(A)$; the associated projector is $Q(A)V(A) = (QV)(A)$, a polynomial in A. Likewise on the other space. □

3.1.2. Remark. There is a straightforward generalisation to finitely many polynomials, provided they are *pairwise* coprime.

3.2 The Chevalley decomposition

Science often reflects chauvinisms—the French call 'triangle de Pascal' what the Italians call 'triangolo di Tartaglia'. But here is amusingly different: the French give credit of the Chevalley decomposition to the American mathematician Dunford.

3.2.1. Theorem (Chevalley decomposition). *Let $A \in M_n(\mathbb{K})$ be a matrix. Then there is a unique pair (D, N) of matrices with:*

- *D is diagonalisable over $\overline{\mathbb{K}}$;*
- *N is nilpotent;*
- *D and N commute;*
- *$A = D + N$.*

Moreover, D and N are polynomials in A.

3.2.2. Remarks.

- Do not forget the commutation clause $DN = ND$.

- The matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is diagonalisable over $\mathbb{C}$ but not over $\mathbb{R}$. So when $\mathbb{K} = \mathbb{R}$ the matrix D must be taken diagonalisable *over the algebraic closure* $\overline{\mathbb{K}}$.

  However, D and N are polynomials in A *with coefficients in* $\mathbb{K}$.

- The theorem even holds over any field of characteristic 0 as its proof will show.

- This is an algebraic tool not requiring topology. And if there is some topology (say $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), the Chevalley decomposition need not be continuous: let $A_\varepsilon = \begin{pmatrix} 1-\varepsilon & 1 \\ & 1+\varepsilon \end{pmatrix}$. Then $D_\varepsilon = A_\varepsilon$ and $N_\varepsilon = 0$; however $N_0 = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}$.

3.2.3. Remarks (on the relations with and the advantages over Jordan decompositions)**.** Some more remarks if you know the so-called 'Jordan decomposition' of a matrix. The Chevalley decomposition is *not* the same.

- *A* Jordan decomposition is extrinsic. First, the very definition of what a Jordan normal form is somehow arbitrary. Second, presenting a matrix as $A = PJP^{-1}$ requires finding a change of basis matrix: this is non-canonical.

  The Chevalley decomposition instead is intrinsic as D and N depend only on A.

- Suppose $A = PJP^{-1}$ is *a* Jordan decomposition; write $J = C + M$, with $C = \begin{pmatrix} \lambda_1 I_{d_1} & & \\ & \ddots & \\ & & \lambda_k I_{d_k} \end{pmatrix}$ a block-scalar matrix and M a matrix of Jordan blocks. Then $A = PCP^{-1} + PMP^{-1}$ is *the* Chevalley decomposition.

  Though neither of P, C, M is canonical (for instances the blocks of C can always be permuted), D and N are.

- Do however *not* compute a Jordan decomposition in order to find the Chevalley decomposition. It is a waste of time.

  As a matter of fact, the Chevalley decomposition can be computed effectively and exactly (§ 3.4, while the Jordan decomposition is theoretically impossible by Galois theory.

Proof. It is the kind of argument where existence must be treated first, as we actually prove a little more.

Step 1. Existence as polynomials in A (for $\mathbb{K} = \mathbb{C}$).

Verification. Start with characteristic polynomial of A, which splits into monomials over $\mathbb{C}$, with roots the eigenvalues, say:

$$\chi_A(X) = \prod_{\lambda \in \mathrm{Sp}(A)} (X - \lambda)^{\alpha_\lambda},$$

where α_λ is known as the *algebraic multiplicity* of λ in χ_A.

By the Cayley-Hamilton theorem, $\chi_A(A) = 0$. Notice that the factors $(X - \lambda)^{\alpha_\lambda}$ are mutually copime. This suggest to let $F_\lambda(A) = \ker(A - \lambda I_n)^{\alpha_\lambda}$. Now by the coprime kernel lemma 3.1.1, we find:

$$\mathbb{K}^n = \bigoplus_{\lambda \in \mathrm{Sp}(A)} F_\lambda(A).$$

Moreover, the projectors involved in the decomposition, say π_λ, are polynomials in A. Let:

$$D = \sum_{\mathrm{Sp}\, A} \lambda \pi_\lambda \quad \text{and} \quad N = A - D.$$

Both are polynomials in A; in particular, D and N commute. It remains to prove that D is diagonalisable and N is nilpotent.

On each $F_\lambda(A)$, matrix D acts as scalar λ, while matrix N has spectrum $\{0\}$; this is because A has only eigenvalue λ on $F_\lambda(A)$. Since $\mathbb{K}^n$ is the direct sum of the various $F_\lambda(A)$, we find that D has a global basis of eigenvectors, and N has global spectrum $\{0\}$: so D is diagonalisable and N is nilpotent. ◇

Step 2. Existence as polynomials in A (for $\mathbb{K} = \mathbb{R}$).

Verification. Suppose $A \in M_n(\mathbb{R})$. We want to show that in the notation above, $\sum_{\mathrm{Sp}(A)} \lambda \pi_\lambda$ is a *real* polynomial. Being careful with Bézout identities in the coprime kernel lemma, real eigenvalues do have real projectors $\pi_\lambda \in \mathbb{R}[X]$. But non-real eigenvalues $\lambda \in \mathrm{Sp}(A) \setminus \mathbb{R}$ give rise to non-real projectors π_λ.

So return to step 1, *still diagonalising over* $\mathbb{C}$. Since A is a real matrix, $\chi_A \in \mathbb{R}[X]$. Therefore non-real eigenvalues come in pairs $(\lambda, \overline{\lambda})$, and complex-conjugate eigenvalues have equal multiplicities. We treat one such pair.

Define $F_\lambda(A)$ as a subspace of $\mathbb{C}^n$; our arguments still give $\mathbb{C}^n = \bigoplus_{\mathrm{Sp}(A)} F_\lambda(A)$. Projectors π_λ are polynomials in A with complex coefficients. However, letting $\alpha = \alpha_\lambda = \alpha_{\overline{\lambda}}$, one finds:

$$\begin{aligned} F_{\overline{\lambda}}(A) &= \{X \in \mathbb{C}^n : (A - \overline{\lambda} I_n)^\alpha \cdot X = 0\} \\ &= \{X \in \mathbb{C}^n : \overline{(A - \lambda I_n)^\alpha} \cdot X = 0\} \\ &= \{X \in \mathbb{C}^n : (A - \lambda I_n)^\alpha \cdot \overline{X} = 0\} \\ &= \{\overline{X} : X \in F_\lambda(A)\}. \end{aligned}$$

In particular, $\pi_{\overline{\lambda}} = \overline{\pi_\lambda}$. Therefore $\lambda \pi_\lambda + \overline{\lambda} \pi_{\overline{\lambda}} = 2\,\mathrm{Re}(\lambda \pi_\lambda)$ is a real polynomial. As we said, real eigenvalues create no problems. Therefore $D = \sum_{\mathrm{Sp}(A)} \lambda \pi_\lambda \in \mathbb{R}[X]$, as desired. ◇

Step 3. Uniqueness.

Verification. In either case we work over $\mathbb{C}$. Suppose $A = D' + N'$ is another decomposition. Then D' commutes with D' and N', so D' commutes with A. But D is a polynomial in A, so D' commutes with D as well. And so on: all matrices in the picture commute. Now D and D' are commuting, diagonalisable matrices: they are simultaneously diagonalisable, so $D - D'$ is diagonalisable. Also, N and N' are commuting, nilpotent matrices: so $N - N'$ is nilpotent. Hence $D - D' = N' - N$ is *both* diagonalisable and nilpotent. It can be diagonalised to its spectrum, which is 0:

it can be diagonalised to the 0 matrix, so $D - D' = N' - N = 0$, as desired. ◇

This completes the proof. □

3.2.4. Remarks.

- Thus, if one knows the Chevalley decomposition of a matrix $A = D + N$, computing $\exp(A)$ reduces to computing $\exp(D)$ and $\exp(N)$, and multiplying them.
- It so happens that there is a general algorithmic method to compute Chevalley decompositions (§ 3.4, which is optional and harder).
- One should however not be too enthusiastic: in order to compute $\exp(A)$, one *still needs to diagonalise* D. The latter cannot be done exactly since general, exact determination of $\mathrm{Sp}(D) = \mathrm{Sp}(A)$ amounts to general, exact resolution of polynomials: something proved impossible in Galois theory.

3.3 The *complex* matrix exponential is onto

3.3.1. Theorem. *The map* $\exp\colon M_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ *is onto. Moreover, for* $B \in \mathrm{GL}_n(\mathbb{C})$ *there is* $P_B(X) \in \mathbb{C}[X]$ *with* $\exp(P_B(B)) = B$.

The following definition is important when studying matrix groups.

3.3.2. Definition. A matrix B is *unipotent* if $\mathrm{Sp}(B) = \{1\}$.

(We will tend to avoid letter U for unipotent matrices, in order to prevent confusion with *unitary* matrices studied in § 5.) We also need a classical result from algebra.

3.3.3. Lemma (Lagrange interpolation)**.** *Let* $\mathbb{F}$ *be any field and* $a_0, \ldots, a_k \in \mathbb{F}$ *be* $k+1$ distinct *elements. Let* $b_0, \ldots, b_k \in \mathbb{F}$ *be any elements. Then there is* $P \in \mathbb{F}[X]$ *of degree* $\leq k$ *such that for all* i *one has* $P(a_i) = b_i$.

Proof of Theorem 3.3.1. Let $B \in \mathrm{GL}_n(\mathbb{C})$; we prove that $B \in \exp(\mathbb{C}[B])$.

Step 1. The diagonalisable case.

Verification. Suppose that B is diagonalisable. By conjugacy (lemma 2.2.5), we may suppose that B is diagonal, say

$$B = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

The complex numbers λ_i are not 0 since B is invertible. By surjectivity of the complex exponential $\exp\colon \mathbb{C} \to \mathbb{C}^\times$, there are $\mu_i \in \mathbb{C}$ with $\exp(\mu_i) = \lambda_i$.

The μ_i are not uniquely defined since the complex exponential has a non-trivial kernel. But if we consistently choose them such that $(\forall i, j = 1 \ldots n)(\lambda_i = \lambda_j \to \mu_i = \mu_j)$, we can apply Lagrange interpolation. So there is a polynomial P with $P(\lambda_i) = \mu_i$.

It depends on B. Then clearly:

$$\exp(P(B)) = \exp\begin{pmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_n\end{pmatrix} = \begin{pmatrix}e^{\mu_1} & & \\ & \ddots & \\ & & e^{\mu_n}\end{pmatrix} = B.\diamond$$

Step 2. The unipotent case.

Verification. Suppose that $\mathrm{Sp}(B) = \{1\}$. By conjugacy again, we may suppose that B is upper-triangular with 1's on the diagonal. Consider the following two sets of matrices:

$$\mathcal{N} = \left\{\begin{pmatrix}0 & & * \\ & \ddots & \\ & & 0\end{pmatrix}\right\} \quad\text{and}\quad \mathcal{Y} = \left\{\begin{pmatrix}1 & & * \\ & \ddots & \\ & & 1\end{pmatrix}\right\}.$$

Clearly, $\exp(\mathcal{N}) \subseteq \mathcal{Y}$. We claim that equality holds, and a preimage of $Y \in \mathcal{Y}$ can be taken as a polynomial in Y, which will prove the claim.

Let $\mathrm{lo}(1 + X) = X - \frac{X}{2} + \cdots + (-1)^n \frac{X^{n-1}}{n-1}$, which is the truncation of logarithm to order n. For $N \in \mathcal{N}$ our series are finite sums and $\exp(\mathrm{lo}(I + N)) = I + N$.

In particular, if B is upper-triangular with 1's on the diagonal, apply this to $N = B - I_n$. We get $\exp(\mathrm{lo}(B)) = B$, as desired. $\diamond$

Step 3. The general case.

Verification. Let $B \in \mathrm{GL}_n(\mathbb{R})$ with no assumptions; we want to prove it is in the image of exp. Write the Chevalley decomposition of B, say $B = D + N$. Bear in mind that $D, N \in \mathbb{C}[B]$.

By inversibility, $0 \notin \mathrm{Sp}(B) = \mathrm{Sp}(D)$, so D is invertible and diagonalisable; in particular there is $C \in \mathbb{C}[D]$ with $\exp(C) = D$. Since $D \in \mathbb{C}[B]$, one has $C \in \mathbb{C}[B]$. Also notice that:

$$B = D(I_n + D^{-1}N),$$

and since D and N commute, $D^{-1}N$ is nilpotent. So $I_n + D^{-1}N$ is unipotent. Therefore there is $M \in \mathbb{C}[D^{-1}N]$ with $\exp(M) = I_n + D^{-1}N$. Since $D^{-1}N \in \mathbb{C}[B]$, one has $M \in \mathbb{C}[B]$.

Therefore $C + M \in \mathbb{C}[B]$ so they commute, and by the sum property:

$$\exp(C + M) = \exp(C) \cdot \exp(M) = D \cdot (I_n + D^{-1}N) = D + N = B.\diamond$$

This proves surjectivity *in the complex case.* □

3.4 Effectiveness of the Chevalley decomposition (optional)

Not covered in class.

This subsection is a harder read, but also a more interesting one. One must be comfortable with advanced algebra: ideals in rings, polynomial rings, coprimality.

One should also have seen at least once the Newton method. Contrary to widespread belief, the Newton method is *not* a mere tool in numerical analysis. The p-adic fields in number theory were discovered by Hensel thanks to the Newton method.

Last, we need an easy lemma.

3.4.1. Lemma. *Let $P(X) \in \mathbb{K}[X]$. Then there is $Q(X,Y)$ such that:*

$$P(X+Y) = P(X) + YP'(X) + Y^2 Q(X,Y).$$

Proof. The property is linear in P, so it is enough to treat the case of monomials. Expanding $(X+Y)^k$ following Newton's binomial, the claim is now obvious. □

3.4.2. Theorem. *There is an exact algorithm in at most* $\sim \log_2 n$ *steps giving the Chevalley decomposition of a matrix in* $M_n(\mathbb{K})$.

Proof. The algorithm is contained in the proof.

Let $A \in M_n(\mathbb{K})$ and $\mathbb{A} = \mathbb{K}[A]$ be the subspace of matrices which can be written as polynomials in A. It turns out that $\mathbb{A}$ is not only a vector subspace of $M_n(\mathbb{K})$ but also a subalgebra, meaning $I_n \in \mathbb{A}$ and $\mathbb{A}$ is closed under matrix product.

As always in algebra, $\mathbb{A}^\times$ denotes the group of invertible elements, viz.:

$$\mathbb{A}^\times = \{x \in \mathbb{A} : (\exists y \in \mathbb{A})(xy = yx = I_n)\}.$$

Be very careful that inversion is relative to $\mathbb{A}$ (it so happens, as proved in § 4.3, that here this is more harmless than seems.)

Step 1. There is a polynomial with leading coefficient 1 denoted μ_A such that $\mathbb{A} \simeq \mathbb{K}[X]/(\mu_A)$ as $\mathbb{K}$-algebras.

Verification. Consider the *evaluation map*:

$$\begin{array}{rccc} \mathrm{ev}_A \colon & \mathbb{K}[X] & \to & \mathbb{A} \\ & P & \mapsto & P(A). \end{array}$$

It is a morphism of $\mathbb{K}$-algebras, and is onto by construction (which is the correct explanation for commutativity of $\mathbb{A}$).

Let $I = \ker \mathrm{ev}_A$ be the kernel, which is a ideal of $\mathbb{K}[X]$, so that $\mathbb{A} \simeq \mathbb{K}[X]/I$ as $\mathbb{K}$-algebras. Since the polynomial ring $\mathbb{K}[X]$ is principal, there is a unique polynomial with leading coefficient 1 such that $I = (\mu_A)$. ◇

Reflecting on the proof of the Chevalley decomposition, we see that another polynomial should play a role. Let:

$$P = \prod_{\lambda \in \mathrm{Sp}(A)} (X - \lambda).$$

Step 2. $P'(A) \in \mathbb{A}^\times$ and $P(A)$ is nilpotent.

Verification. Since P has only simple roots, it has no common factor with its derivative P'; in gcd-notation, $P \wedge P' = 1$. Now μ_A and P have the same roots, so again $\mu_A \wedge P' = 1$. By a Bézout relation there are $U, V \in \mathbb{K}[X]$ such that $\mu_A U + P'V = 1$. Applying in A we find $P'(A) \cdot V(A) = 1$ in $\mathbb{A}$, as desired.

Now by the Cayley-Hamilton theorem, $\mu_A | \chi_A | P^n$, so $(P(A))^n = 0$ in $\mathbb{A}$. ◇

We define a sequence of elements of $\mathbb{A}$ as follows:

- $x_0 = A$;
- $x_{n+1} = x_n - \underbrace{P(x_n)[P'(x_n)]^{-1}}_{t_n}$.

This definition requires $P'(x_n)$ to be invertible, which is a consequence of our next step. Keep the definition of t_n in mind.

Step 3. For all integers $n \in \mathbb{N}$:

- $P'(x_n)$ is invertible in $\mathbb{A}$;
- $P(x_n) \in (P(x_0)^{2^n})$. (This is notation for 'the ideal generated by $P(x_0)^{2^n}$'.)

Verification. By induction on n. The case $n = 0$ was dealt with in step 2. By lemma 3.4.1 there is a polynomial $Q(X, Y)$ with $P'(X + Y) = P'(X) + YP''(X) + Q(X, Y)$. In particular since $x_{n+1} = x_n - t_n$ one gets:

$$P'(x_{n+1}) = P'(x_n) - t_nP''(x_n) + t_n^2 Q(x_n, t_n).$$

By induction, $P'(x_n) \in \mathbb{A}^\times$. Now t_n is a multiple of $P(x_n)$, hence a multiple of $P(x_0)$, which is nilpotent: so t_n is nilpotent. Since $\mathbb{A}$ is commutative, $P'(x_n) + t_n$ is invertible. This is the first claim.

Always by lemma 3.4.1 there is a polynomial $R(X, Y)$ with $P(X + Y) = P(X) + YP'(X) + Y^2R(X, Y)$. In particular one has (keeping the definition of t_n in mind):

$$P(x_{n+1}) = P(x_n) - t_nP'(x_n) + t_n^2R(x_n, t_n) = t_n^2R(x_n, t_n).$$

But this is $(t_n^2) \leq (P(x_n))^2 \leq (P(x_0)^{2^n})^2 \leq (P(x_0)^{2^{n+1}})$. This is the second claim. ◇

Step 4. The sequence (x_n) is stationary at some n_0; letting $D = x_{n_0}$ and $N = A - D$ we are done.

Verification. Let n_0 be such that 2^{n_0} is greater than the nilpotence order of $P(x_0)$ in $\mathbb{A}$ (the latter is at most the size of the matrix, so certainly $n_0 \leq n$). Then for all $n \geq n_0$ one has $P(x_n) = 0$, so $t_n = 0$, implying $x_{n+1} = x_n$: the sequence is stationary. By construction, D and N are in $\mathbb{A}$, viz. polynomials in A; they commute. It remains to prove that D is diagonalisable and N is nilpotent.

By construction, $P(x_{n_0}) = 0$, so matrix $D = x_{n_0}$ is killed by a polynomial with simple roots: it is therefore diagonalisable as a simple application of the coprime kernel lemma 3.1.1.

Finally observe:

$$N = A - D = x_0 - x_{n_0} = \underbrace{x_0 - x_1}_{=t_0} + \underbrace{x_1 - x_2}_{=t_1} + \cdots + \underbrace{x_{n_0-1} - x_{n_0}}_{=t_{n_0-1}}.$$

But each t_k is in $(P(x_k))$, hence nilpotent; and therefore so is N. ◇

This completes the (effective) proof. □

3.4.3. Remark. As one sees, the phenomenon takes place in a larger setting than linear algebra. All one needs is:

- a field $\mathbb{K}$ of characteristic 0;
- a proper quotient $\mathbb{A} = \mathbb{K}[X]/I$ by a non-trivial ideal.

The algorithm then produces a decomposition of $x = X \mod I$ as $x = s + n$ where the minimal polynomial of s in any algebraic closure has only simple roots, and n is nilpotent.

As a matter of fact, the characteristic 0 assumption can be weakened into: all irreducible factors of μ (the minimal polynomial of x, viz. 'the' generator of I) have non-zero derivative.

3.5 Exercises

3.5.1. Exercise. *Find where we used $\mathbb{K} = \mathbb{C}$ in the proof of surjectivity (theorem 3.3.1) and spot the failure in the real case. Give a matrix in $\mathrm{GL}_n(\mathbb{R}) \setminus \exp(M_n(\mathbb{R}))$.*

3.5.2. Exercise. *Prove that* exp *is a homeomorphism between the set of nilpotent matrices and the set of unipotent matrices.*

Solution. Continuity is clear. Return to § 3.3; since lo is continuous, if we prove bijectivity we also have continuity of the inverse. But we already proved surjectivity so only injectivity remains.

So let N be a nilpotent matrix with $\exp(N) = I_n$; we must prove $N = 0$; we may suppose that $N \in \mathcal{N}$ is upper-triangular with 0's on the diagonal.

Let $k \in \mathbb{N}$ be minimal with $N^k = 0$ (this exists, by nilpotence). Then:

$$\exp(N) = \sum_{\ell<k} \frac{N^\ell}{\ell!} = I_n + \sum_{0<\ell<k} \frac{N^\ell}{\ell!} = I_n,$$

so $\sum_{0<\ell<k} \frac{N^\ell}{\ell!} = 0$. Now this has the form $N \cdot Q(N) = 0$, where Q is a polynomial with constant term 1. In particular, $Q(N)$ is an upper-triangular matrix with 1's on the diagonal, hence invertible. So $N = 0$, as wanted.

3.5.3. Exercise.

1. *Compute the Chevalley decomposition of* $\exp(A)$ *based on that of* A.
2. *Deduce that* A *is diagonalisable iff* $\exp(A)$ *is.*

Solution.

1. One has $\exp A = \exp(D)+\exp(D)\cdot(\exp(N)-I_n)$. All terms commute since all are polynomials in A. Clearly $\exp(D)$ is diagonalisable, in any eigenbasis for D. Now $\exp(N) - I_n$ is nilpotent. Indeed, one may trigonalise N to see it: up to conjugacy, N is upper-triangular with 0's on the diagonal. So $\exp(N)$ is upper-triangular with 1's on the diagonal, and $\mathrm{Sp}(\exp(N) - I_n) = \{0\}$ again.
2. If A is diagonalisable then so is $\exp(A)$ since it is a polynomial in A. Suppose $\exp(A)$ is diagonalisable; let $A = D + N$ be the Chevalley decomposition of A. Then since $\exp(A)$ has no nilpotent part, $\exp(D) \cdot (\exp(N) - I_n) = 0$; since $\exp(D)$ is invertible, we get $\exp(N) = I_n$, and N is nilpotent. But this implies $N = 0$ as exp induces a bijection between nilpotent and unipotent matrices.

3.5.4. Exercise (image of exp: the real case)**.** *Prove that:*

$$\exp(M_n(\mathbb{R})) = \{A \in \mathrm{GL}_n(\mathbb{R}) : (\exists B \in \mathrm{GL}_n(\mathbb{R}))(A = B^2)\}.$$

Hint: if $A = B^2 \in \mathrm{GL}_n(\mathbb{R})$, *then there is* $Q \in \mathbb{C}[X]$ *with* $\exp(Q(B)) = B$.

Solution. One inclusion is not hard. If $A = \exp(M)$ with $M \in M_n(\mathbb{R})$, then using the sum property one sees that $(\exp(\frac{1}{2}M))^2 = \exp(M) = A$ is a square.

For the converse, let $A \in \mathrm{GL}_n(\mathbb{R})$ be a square, say $A = B^2$. Treat B as a complex matrix: there is $Q(X) \in \mathbb{C}[X]$ with $\exp(Q(B)) = B$. Now B is actually a real matrix, so:

$$B = \overline{B} = \overline{Q(B)} = \overline{Q}(B),$$

and therefore $\exp(\overline{Q}(B)) = B$. Since $Q(B)$ and $\overline{Q}(B)$ are polynomials in B, they commute; we find:

$$\exp(Q(B) + \overline{Q}(B)) = B^2 = A,$$

but $Q(B) + \overline{Q}(B)$ is a real matrix, so $A \in \exp(M_n(\mathbb{R}))$.

3.5.5. Exercise. *Prove that the abstract form of the Chevalley decomposition (remark 3.4.3) fails if there are no assumptions on the characteristic or the factors of* μ.

Hint: over the field $\mathbb{K} = \mathbb{F}_p(T)$, *consider* $\mathbb{A} = \mathbb{K}[X]/(X^p - T)$.

Solution. Using Eisenstein's criterion, see that $\mathbb{A}$ is actually a field: hence has no nilpotent elements. But $x = X \mod I$ has minimal polynomial $X^p - T$, which has a unique root of mulitplicity $p > 1$.

4 Advanced analytic properties

This section is an optional read though it greatly helps clarify matters. It requires knowledge of advanced calculus, that is elementary differential geometry. One must know what the *differential* of a function of a vector variable/several real variables is. We give the definition in § 4.1 but it takes time to understand. Then comes the notion of a C^k-diffeomorphism; using the inverse function theorem, we prove in § 4.2 that the exponential is a local diffeomorphism from a neighbourhood of 0 to one of I_n. This gives another proof of the surjectivity of $\exp\colon M_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ in § 4.3.

4.1 The exponential as a multi-variable map

In order to distinguish points from vectors (technically, points live in the manifold and vectors on the tangent bundle) I usually denote the former by $\underline{a}$ and the latter by $\overline{h}$.

4.1.1. Definition. A map $f\colon \mathbb{R}^n \to \mathbb{R}^k$ is *differentiable* at $\underline{a} \in \mathbb{R}^n$ if there is a linear map $L\colon \mathbb{R}^n \to \mathbb{R}^k$ such that:

$$(\forall \overline{h} \in \mathbb{R}^n)(f(\underline{a} + \overline{h}) = f(\underline{a}) + L(\overline{h}) + o(\overline{h})).$$

The linear map is then unique and called the *differential of* f *at* $\underline{a}$, denoted by $D_{\underline{a}} f$.

With differentiability come of course the notions of C^k and C^∞ maps. With a couple of results from analysis on normally convergent function series, one can easily show that the matrix exponential is C^∞; we do not follow this line.

4.1.2. Proposition. *The matrix exponential is differentiable at* 0, *and* $D_0 \exp$ *is the identity map* $\mathrm{Id}\colon M_n(\mathbb{K}) \to M_n(\mathbb{K})$.

Proof. A consequence of the definition, reproducing the proof of lemma 2.3.1. Indeed let H be a 'small' matrix. Then fixing as always a submultiplicative norm:

$$\begin{aligned}\exp(0+H) &= \exp(H)\\ &= I_n + H + \sum_{n\geq 2} \frac{H^n}{n!}\\ &= \exp(0) + \mathrm{Id}(H) + R(H).\end{aligned}$$

Now $\|R(H)\| \leq \|H\|^2 \exp(\|H\|) = o(H)$, as desired. □

4.1.3. Remark. One can explicitly compute the differential of exp everywhere. For matrices A, B let $\mathrm{ad}_A(B) = AB - BA$. Then exp is differentiable at any A and:

$$D_A \exp = \exp(A) \sum_{n\geq 0} \frac{(-1)^n}{(n+1)!} \mathrm{ad}_A^n .$$

This is either advanced or technical (there exist elementary proofs through differential equations), and would take us straight to classical Lie theory.

4.2 The exponential is a local diffeomorphism

In what follows we tend to be heavy-handed distinguishing *local*, viz. on neighbourhoods, from *global*, viz. on all of the ambient space. Since exp is not globally injective, there is no hope it will be a global diffeomorphism.

4.2.1. Definition. A (local) C^k-diffeomorphism between open subsets $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^n$ is a C^k bijection $f\colon \mathcal{U} \simeq \mathcal{V}$ whose reciprocal mapping is also C^k.

4.2.2. Theorem ('inverse function theorem', admitted from elementary differential geometry). *Suppose* $f\colon \mathbb{R}^n \to \mathbb{R}^n$ *is* C^k *on a neighbourhood of* $\underline{a}$, *and* $D_{\underline{a}} f$ *is invertible as a linear map* $\mathbb{R}^n \to \mathbb{R}^n$. *Then there are neighbourhood* $\mathcal{U}$ *of* $\underline{a}$ *and* $\mathcal{V}$ *of* $f(\underline{a})$ *such that* f *induces a (local)* C^k*-diffeomorphism* $\mathcal{U} \simeq \mathcal{V}$.

4.2.3. Corollary. *The exponential map induces a (local) diffeomorphism from a neighbourhood of* 0 *to one of* I_n. *In particular, there exist* $\varepsilon \in \mathbb{R}_{>0}$ *and a* C^1 *map* $\log_{\mathcal{V}(I_n)}\colon B(I_n, \varepsilon) \to M_n(\mathbb{K})$ *which is a (local) reciprocal mapping of* exp. *In particular,* $\log_{\mathcal{V}(I_n)}$ *is differentiable at* I_n *and* $\log_{\mathcal{V}(I_n)}(I_n + H) = H + o(H)$.

4.2.4. Remark. More explicitly, one can fix a submultiplicative norm $\|\cdot\|$ and see that on $B(0,1)$ the logarithm series converges normally; this gives a series expansion for $\log_{\mathcal{V}(I_n)}$.

One may thus define the matrix logarithm *locally*, with good analytic properties. However it is not defined everywhere; there is no global matrix logarithm (see complex case, as opposed to real case).

4.3 Application to surjectivity

4.3.1. Theorem. *The map* $\exp\colon M_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ *is surjective.*

Proof. Let $A \in M_n(\mathbb{C})$. Let $\mathbb{A} = \mathbb{C}[A]$ be the subspace of polynomials in A; it actually is a subalgebra, viz. contains I_n and is stable under matrix product. Recall that $\mathbb{A}^\times$ denotes the group of invertible elements, viz.:

$$\mathbb{A}^\times = \{x \in \mathbb{A} : (\exists y \in \mathbb{A})(xy = yx = I_n)\};$$

it is indeed a subgroup of $\mathrm{GL}_n(\mathbb{C})$ with respect to multiplication, but it need not be commutative. Bear in mind that $\exp(A) \in \mathbb{A}$ by lemma 2.2.4. We equip $\mathbb{A} \leq M_n(\mathbb{C})$ with its subspace topology.

Step 1. One has $\mathbb{A}^\times = \mathbb{A} \cap \mathrm{GL}_n(\mathbb{C})$, viz. a matrix in $\mathbb{A}$ is invertible *relatively to* $\mathbb{A}$ iff it is invertible.

Verification. One implication is obvious. For the converse, suppose that $M \in \mathbb{A} \cap \mathrm{GL}_n(\mathbb{C})$. Then $0 \notin \mathrm{Sp}(M)$, so the characteristic polynomial has a non-trivial constant term; it has the form $M^n + \cdots + \lambda$ where $\lambda \neq 0$. Now $M^n + \cdots + \lambda = 0$ by the Cayley-Hamilton theorem. Multiplying (on either side) by M^{-1}, we find that M^{-1} is a polynomial in M, hence in $\mathbb{A}$. ◇

Step 2. The restriction $\exp\colon (\mathbb{A}, +) \to (\mathbb{A}^\times, \cdot)$ is a group morphism. Moreover, it is enough to prove that it is onto, viz. $\exp(\mathbb{A}) = \mathbb{A}^\times$.

Verification. Notice that the inclusion $\exp(\mathbb{A}) \subseteq \mathbb{A}^\times$ holds: if $B \in \mathbb{A}$ then B is a polynomial in A. But $\exp(B)$ is a polynomial in B, so again a polynomial in A. Moreover, $\exp(B)$ is invertible. Hence $\exp(\mathbb{A}) \subseteq \mathbb{A} \cap \mathrm{GL}_n(\mathbb{K}) = \mathbb{A}^\times$. So the restriction does take $\mathbb{A}$ to $\mathbb{A}^\times$. Since elements in $\mathbb{A}$ commute pairwise, it is a morphism by the sum property, proposition 2.2.9.

Suppose that the equality $\exp(\mathbb{A}) = \mathbb{A}^\times$ is known and let us prove that exp is onto. Let $B \in \mathrm{GL}_n(\mathbb{C})$. Then applying the result to $\mathbb{B} = \mathbb{C}[B]$, since $B \in \mathbb{B} \cap \mathrm{GL}_n(\mathbb{C}) = \mathbb{B}^\times$ there is $A \in \mathbb{B}$ with $\exp(A) = B$: as desired. ◇

We have reduced the problem to showing that a group homomorphism is surjective: this should be simpler. The proof is a connectedness argument using diffeomorphisms.

Step 3. The space $\mathbb{A}^\times$ is connected.

Verification. To prove connectedness we prove path-connectedness. Let $M \in \mathbb{A}^\times$; we find a continuous path to I_n inside $\mathbb{A}^\times$. Indeed consider $\Gamma(z) = (1 - z)I_n + zM$, a map from $\mathbb{C}$ to $\mathbb{A}$. Also consider $\Delta(z) = \det \Gamma(z)$, which is a polynomial in z. Notice that $\Gamma(0) = I_n$ and $\Gamma(1) = M$ so δ vanishes at neither 0 nor 1. In particular Δ is not the zero polynomial; it has finitely many roots. So there is a continuous path $\zeta\colon [0, 1] \to \mathbb{C}$ taking 0 to 0, 1 to 1, and avoiding the roots of δ. Then $\gamma = \Gamma \circ \zeta\colon [0, 1] \to \mathbb{A}$ does $\gamma(0) = I_n$, $\gamma(1) = M$, and $\det \gamma(t) \neq 0$, so $\gamma(t) \in \mathbb{A} \cap \mathrm{GL}_n(\mathbb{C}) = (\mathbb{A}^\times$: we are done. ◇

Step 4. The subset $\exp(\mathbb{A})$ is open and closed in $\mathbb{A}^\times$.

Verification. To avoid confusion, we shall denote by $\check{\exp}$ the restriction of the exponential map from $\mathbb{A}$ to itself, viz. $\check{\exp}\colon \mathbb{A} \to \mathbb{A}$. Since $\mathbb{A}$ is a subspace of $M_n(\mathbb{C})$, we can treat it as some $\mathbb{R}^k$ (the dimension depends on properties of $\mathbb{A}$) and do calculus.

Since $\exp\colon M_n(\mathbb{R}) \to M_n(\mathbb{R})$ was, $\check{\exp}$ remains C^1 (even, C^∞); moreover the differential at 0 remains $D_0\check{\exp} = \mathrm{Id}\colon \mathbb{A} \to \mathbb{A}$. By the inverse function theorem, there are small neighbourhoods $\mathcal{U}$ of $0 \in \mathbb{A}$ and $\mathcal{V}$ of $I_n \in \mathbb{A}$ such that $\check{\exp}$ induces a (local) homeomorphism $\mathcal{U} \simeq \mathcal{V}$. In particular, $\exp(\mathbb{A}) \supseteq \exp(\mathcal{U}) = \mathcal{V}$ contains a neighbourhood of $I_n \in \mathbb{A}$. So $\exp(\mathbb{A})$ contains a neighbourhood of $I_n \in \mathbb{A}^\times$. We completely forget about $\mathbb{A}$ and focus on $\mathbb{A}^\times$ as a topological group.

A subgroup containing I_n as an interior point must be open, because translations are homeomorphisms. Hence $\exp(\mathbb{A}) \leq \mathbb{A}^\times$ is an open subgroup. Now *every* open subgroup of a topological group is actually closed, because whenever $H \leq G$ is open we have:

$$G = \bigsqcup_{g \in \mathcal{S}} gH$$

for some index set $\mathcal{S} \subseteq G$. Then removing H itself, one has $G \smallsetminus H = \bigsqcup_{g \in \mathcal{S}'} gH$, a union of open subsets of G: so $G \smallsetminus H$ is open in G, meaning that H is closed in G. ◇

By step 4, $\exp(\mathbb{A})$ is both open and closed in $\mathbb{A}^\times$, which is connected by step 3. Therefore $\exp(\mathbb{A}) = \mathbb{A}^\times$ and we are done as seen in step 2. □

4.3.2. Corollary. *For every $B \in \mathrm{GL}_n(\mathbb{C})$ there is $Q_B \in \mathbb{C}[X]$ with $\exp(Q_B(B)) = B$.*

Proof. In step 2 of the proof of theorem 4.3.1 we found a preimage of $B \in \mathrm{GL}_n(\mathbb{C})$ inside $\mathbb{C}[B]$, that is a preimage of the form $Q_B(B)$. □

Analytic properties can be taken much further. The study of (analytic) Lie groups is based on the following result.

4.3.3. Theorem (Cartan, von Neumann). *Let $G \leq \mathrm{GL}_n(\mathbb{K})$ be a closed subgroup. Then its tangent space at I_n is equal to:*

$$\{M \in M_n(\mathbb{K}) : (\forall t \in \mathbb{R})(\exp(tM) \in G)\}.$$

This is out of the scope of this class.

4.4 Exercises

4.4.1. Exercise. *Prove that there exists a neighbourhood of I_n in $\mathrm{GL}_n(\mathbb{K})$ containing no non-trivial subgroup, or equivalently that $\mathrm{GL}_n(\mathbb{K})$ has no arbitrarily small subgroups. (Hint: let $\mathcal{U}$ be a neighbourhood of 0 on which $\exp$ is a C^1-diffeomorphism. Do not consider $\exp(\mathcal{U})$ but $\exp(\frac{1}{2}\mathcal{U})$.)*

Solution. Let $\varepsilon \in \mathbb{R}_{>0}$ and $\mathcal{U} = B(0, \varepsilon)$ be such that $\exp\colon \mathcal{U} \to \exp(\mathcal{U})$ is a (local) C^1-diffeomorphism; in particular $\mathcal{V} = \exp(\mathcal{U})$ is a neighbourhood of I_n. We try to prove that $\mathcal{V}$ has no other subgroup than $\{I_n\}$…and fail.

First attempt: let $H \subseteq \mathcal{V}$ be a subgroup of $\mathrm{GL}_n(\mathbb{K})$; let $h \in H$ and $x \in \mathcal{U}$ with $h = \exp(x)$. One gets $h^n = \exp(nx) \in H = \exp \mathcal{U}$, so one hopes that nx remains in $\mathcal{U}$

while $n \to \infty$ which is a contradiction. However, the argument '$(\exp(nx) \in \exp H) \Rightarrow (nx \in H)$' is faulty, as exp is not globally injective.

Second attempt: take the least n with $nx \notin \mathcal{U}$. Still won't work.

Third and successful attempt: let $\mathcal{W} = \exp\left(\frac{1}{2}(\mathcal{U})\right)$, yet another neighbourhood of I_n. We claim that $\mathcal{W}$ has no other subgroup than $\{I_n\}$. Indeed, let $H \subseteq \mathcal{W}$ be a subgroup of $\mathrm{GL}_n(\mathbb{K})$; let $h \in H \setminus \{I_n\}$. By assumption there is $x \in \frac{1}{2}\mathcal{U}$ with $h = \exp(x)$. Let n be the least integer with $nx \notin \frac{1}{2}\mathcal{U}$; notice however that $(n-1)x \in \frac{1}{2}\mathcal{U}$, and $nx = (n-1)x + x \in \frac{1}{2}\mathcal{U} + \frac{1}{2}\mathcal{U} = \mathcal{U}$. Then $\exp(nx) = h^n \in H \subseteq \exp(\frac{1}{2}\mathcal{U})$ and $nx \in \mathcal{U}$ (hence in a domain of injectivity of exp) force $nx \in \frac{1}{2}\mathcal{U}$, a contradiction.

4.4.2. Exercise.

1. *Let (A_n) be a sequence of matrices converging to A. Prove $\left(I + \frac{A_n}{n}\right)^n \xrightarrow[n\to\infty]{} \exp(A)$. (Hint: have you tried logarithms?)*

2. *Application. Do* not *suppose that A and B commute. Prove:*

$$\left(\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\right)^n \xrightarrow[n\to\infty]{} \exp(A+B).$$

3. *Second application. Prove:*

$$\left(\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\exp\left(\frac{-A}{n}\right)\exp\left(\frac{-B}{n}\right)\right)^{n^2} \xrightarrow[n\to\infty]{} \exp(AB - BA).$$

Solution.

1. Since $A_n \longrightarrow A$, one has $I + \frac{A_n}{n} \longrightarrow I$; so for n large enough, the matrix is in the domain of the logarithm. As the latter function is C^1, there is a Taylor expansion $\log_{\mathcal{V}(I_n)}(I + H) = H + o(H)$. So bearing in mind that (A_n) is bounded,

$$I + \frac{A_n}{n} = \exp\left(\log_{\mathcal{V}(I_n)}\left(I + \frac{A_n}{n}\right)\right) = \exp\left(\frac{A_n}{n}\right) + o\left(\frac{1}{n}\right)$$

and

$$\left(I + \frac{A_n}{n}\right)^n = \exp(A_n + o(1)).$$

By continuity, we find $\exp(A_n + o(1)) \to \exp(A)$, as desired.

2. Let $A_n = n\left(\exp(\frac{A}{n})\exp(\frac{B}{n}) - I\right)$. Notice that $\exp(\frac{A}{n}) = I + \frac{A}{n} + o(\frac{1}{n})$ and likewise for B. Hence:

$$A_n = n\left[\left(I + \frac{A}{n} + o\left(\frac{1}{n}\right)\right)\left(I + \frac{B}{n} + o\left(\frac{1}{n}\right)\right) - I\right] = A + B + o(1).$$

Therefore $A_n \xrightarrow[n\to\infty]{} A + B$, and by the first question we have the result.

3. Similar argument with:

$$\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\exp\left(\frac{-A}{n}\right)\exp\left(\frac{-B}{n}\right) = I + \frac{AB - BA}{n^2} + O\left(\frac{1}{n^3}\right).$$

5 Exponential and hermitian matrices

This section is optional again. We explore further properties of the matrix exponential, in relation with hermitian matrices. The relevant definitions and a conjugacy result are in § 5.1. We prove in § 5.2 that exp induces a homeomorphism between hermitian matrices and hermitian, definite positive matrices. This finds applications to the study of the unitary group in § 5.3.

5.1 Unitary conjugacy and norms

First recall the definition of a hermitian and of a unitary matrix.

5.1.1. Definition. Let $A \in M_n(\mathbb{C})$.

- The *Hermite-conjugate* of A is $A^* = \overline{A}^t = (\overline{a_{j,i}})_{(i,j)}$, which is obtained by complex-conjugating all coefficients, and transposing the matrix (in any order as these two operations commute).

  Observe at once that $(AB)^* = B^* \cdot A^*$.

- A is *Hermite-symmetric*, or *hermitian*, if $A^* = A$.

- A is *unitary* if $AA^* = A^*A = I_n$.

We denote by H_n the real vector space of $n \times n$ hermitian matrices; H_n is *not* a complex subspace of $M_n(\mathbb{R})$. Let HDP_n be the subset of so-called 'hermitian definite positive matrices', viz. hermitian matrices A with $\mathrm{Sp}(A) \subseteq \mathbb{R}_{>0}$. Be careful that it is *not* a linear subspace. Finally, U_n denotes the unitary group; being closed and bounded, it is compact.

5.1.2. Theorem (admitted from linear algebra)**.** *Let A be a hermitian matrix. Then there exist a unitary matrix U such that $U^{-1}AU$ is diagonal. Moreover,* $\mathrm{Sp}(A) \subseteq \mathbb{R}$.

We shall refer to theorem 5.1.2 as 'unitary conjugacy'.

5.1.3. Lemma. *There exist matrix norms which are invariant under conjugacy by unitary matrices, viz. with:*

$$(\forall A \in M_n(\mathbb{K}))(\forall U \in \mathrm{U}_n(\mathbb{K}))(N(UAU^{-1}) = N(A)).$$

Proof. One needs to think. Let $\langle A|B\rangle = \mathrm{tr}(A^*B)$, where tr is the trace. This is a Hermite-linear form, actually a complex scalar product. So $N(A) = \sqrt{\mathrm{tr}(A^*A)}$ is a norm on $M_n(\mathbb{C})$, often called the *Schur norm*. We claim that it is invariant under conjugacy by the unitary group.

Indeed let $U \in \mathrm{U}_n(\mathbb{C})$. By definition, $U^* = U^{-1}$, so:

$$\begin{aligned} N(U^{-1}AU)^2 &= \mathrm{tr}((U^{-1}AU)^*(U^{-1}AU)) \\ &= \mathrm{tr}(U^*A^*U^{-*}U^{-1}AU) \\ &= \mathrm{tr}(U^{-1}A^*AU), \end{aligned}$$

and since tr is conjugacy-invariant it also equals $\mathrm{tr}(A^*A) = N(A)^2$. □

5.1.4. Corollary. U_n *is compact.*

Proof. It is closed as given by continuous equations in the coefficients. For the norm $\sqrt{\operatorname{tr}(A^*A)}$ it is easily seen bounded. □

5.2 A homeomorphism

5.2.1. Theorem. *The restriction* $\exp\colon \mathrm{H}_n \to \mathrm{HDP}_n$ *is a homeomorphism.*

Proof. For the moment only continuity is clear; so we have three things to prove.

Step 1. Injectivity.

Verification. Suppose $A, B \in \mathrm{H}_n$ have $\exp(A) = \exp(B)$; we must prove $A = B$. Notice that $E = \exp(A) = \exp(B)$ commutes with both A and B, but we don't know whether A and B commute.

By unitary conjugacy (theorem 5.1.2), there is a unitary matrix U with $A = UDU^{-1}$ with D a real diagonal matrix, say $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. As we know, $\exp D = \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix}$. Of course $U \exp(D) U^{-1} = \exp(A) = E$.

Recall that the real exponential is injective; so whenever $e^{\lambda_i} = e^{\lambda_j}$, one has $\lambda_i = \lambda_j$. By Lagrange interpolation (lemma 3.3.3), there is $P \in \mathbb{R}[X]$ such that $P(e^{\lambda_i}) = \lambda_i$. Then $P(\exp D) = D$, so $P(E) = P(U\exp(D)U^{-1}) = UDU^{-1} = A$ is a polynomial in E. But $E = \exp(B)$ also is a polynomial in B; hence $A \in \mathbb{C}[B]$, implying that A and B commute. Now $\exp(A - B) = E \cdot E^{-1} = I_n$.

Of course there are non-zero matrices whose exponential is I_n, but here $A - B \in \mathrm{H}_n$. By unitary conjugacy, say $A - B = V \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} V^{-1}$, with real eigenvalues μ_k. Then $\exp(A - B) = V \begin{pmatrix} e^{\mu_1} & & \\ & \ddots & \\ & & e^{\mu_n} \end{pmatrix} V^{-1} = I_n$ implies $\begin{pmatrix} e^{\mu_1} & & \\ & \ddots & \\ & & e^{\mu_n} \end{pmatrix} = I_n$, which can happen only for $\mu_1 = \cdots = \mu_n = 0$. So $A = B$; we proved injectivity. ◇

A more geometric argument, avoiding Lagrange interpolation. It is enough to show that A and B commute. Since they are diagonalisable, it is enough to show that they have the same eigenspaces. Since A and E commute, it is the case that A stabilises (=leaves invariant) the eigenspaces of E, and vice-versa.

For M_1 to stabilise the eigenspaces of M_2 and vice-versa does *not* generally imply that M_1 and M_2 have the same eigenspaces: for instance, $M_1 = I_n$ commutes to every matrix, but has only one eigenspace. The argument is however valid here since A and E have the same multiplicities for their eigenvalues, by injectivity of the real exponential.

So A and E have the same eigenspaces, and B and E have the same eigenspaces.

Therefore A and B have the same eigenspaces: hence they commute. ◇

Step 2. Surjectivity.

Verification. Let $B \in \mathrm{HDP}_n$; we find $A \in \mathrm{H}_n$ with $\exp(A) = B$. This is easy: by unitary diagonalisation, there is a unitary matrix U such that $UBU^{-1} = D$, say $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. By definition, we have $\lambda_i \in \mathbb{R}_{>0}$. So there are $\mu_i \in \mathbb{R}$ with $e^{\mu_i} = \lambda_i$. Then let $A = U^{-1}\begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}U$; one finds:

$$\exp(A) = U^{-1}\exp\begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}U = U^{-1}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}U = U^{-1}DU = B,$$

which proves surjectivity. ◇

Step 3. Continuity of the reciprocal mapping.

Verification. We use sequential continuity. Suppose (B_n) is a sequence in HDP_n converging to $B \in \mathrm{HDP}_n$. Let $A_n \in \mathrm{H}_n$ be such that $\exp(A_n) = B_n$; also let $A \in \mathrm{H}$ with $\exp(A) = B$. We must show $A_n \xrightarrow[n\to\infty]{} A$.

In general, to have *at most* one accumulation point is no sufficient condition for convergence. But it is a basic exercise in topology (exercise 1.4.2) that in finite-dimensional, normed vector spaces, a *bounded* sequence with at most one accumulation point is indeed convergent to that point. So it suffices to prove:

(i) that (A_n) has at most one accumulation point, which is A;

(ii) that (A_n) is bounded.

Item (i) is easy. Indeed, suppose $A_{\varphi(n)} \xrightarrow[n\to\infty]{} C$. Since H_n is a real vector subspace of $M_n(\mathbb{C})$, it is closed; hence $C \in \mathrm{H}_n$. But exp is continuous at C, so:

$$\exp(C) = \lim \exp(A_{\varphi(n)}) = \lim B_{\varphi(n)} = B = \exp A.$$

since $\exp\colon \mathrm{H}_n \to \mathrm{HDP}_n$ is injective on H_n, we get $C = A$. This shows that A is the only *possible* accumulation point for (A_n). Item (ii) (boundedness) will take the rest of the argument. Interestingly, we shall use two distinct norms.

Let $\|\cdot\|$ be a norm invariant under unitary conjugation, viz. for $U \in \mathrm{U}_n$ and $M \in M_n(\mathbb{C})$ one has $\|UMU^{-1}\| = \|M\|$. Since all matrices involved are hermitian, every A_n is unitary-conjugate to a real, diagonal matrix D_n; say $A_n = U_n D_n U_n^{-1}$. By the choice of our norm, $\|A_n\| = \|D_n\|$ and $\|B_n\| = \|\exp(D_n)\|$. It suffices to prove that (D_n) is bounded; but since (B_n) is convergent, $(\exp(D_n))$ is bounded.

Now change norm and consider $\|\cdot\|_\infty$, for which the norm of D_n is the maximum of the absolute values of the diagonal entries. We shall prove that (D_n) is bounded.

Since $(\exp(D_n))$ remains bounded, the sets $\mathrm{Sp}(\exp(D_n))$ remain bounded above in $\mathbb{R}_{>0}$. Since the real exponential is increasing, the latter implies that the sets

$\mathrm{Sp}(D_n)$ remain bounded above in $\mathbb{R}$. This argument does not bound them below, since the limit of the exponential at $-\infty$ is 0.

In order to also bound them below, notice that $B_n^{-1} \longrightarrow B^{-1}$. So our argument also applies to inverses: therefore the sets $\mathrm{Sp}(\exp(D_n)^{-1}) = \mathrm{Sp}(\exp(-D_n))$ remain bounded above in $\mathbb{R}_{>0}$, the sets $\mathrm{Sp}(-D_n)$ remain bounded above in $\mathbb{R}$, and therefore the sets $\mathrm{Sp}(D_n)$ remain bounded *below* in $\mathbb{R}$. So $(\|D_n\|_\infty)$ remains bounded, and so does $\|D_n\| = \|A_n\|$.

We have thus proved that the sequence under study (i) can only have A as an accumulation point and (ii) is bounded. This shows $A_n \xrightarrow[n\to\infty]{} A$, finally proving continuity of the reciprocal mapping. ◇

This completes the proof. □

5.2.2. Corollary. HDP_n *is connected.*

Proof. It is homeomorphic to a real vector space. □

We shall prove connectedness of U_n by introducing more valuable tools.

5.3 The polar map and unitary group

5.3.1. Theorem. U_n *is connected.*

The proof requires two lemmas.

5.3.2. Lemma (square root in HDP_n)**.** *If $H \in \mathrm{HDP}_n$ then there is a unique $K \in \mathrm{HDP}_n$ with $K^2 = H$; we denote it by $\sqrt{H}$. The map $\sqrt{\cdot}\colon \mathrm{HDP}_n \to \mathrm{HDP}_n$ is continuous.*

We give a quick proof building on Theorem 5.2.1; for direct proofs see the exercises.

Proof. The map $\exp\colon \mathrm{H}_n \to \mathrm{HDP}_n$ is a homeomorphism; let $\log_{\mathrm{HDP}}$ be the reciprocal bijection, which is continuous. The idea is to put:

$$\sqrt{H} = \exp(\frac{1}{2}\log_{\mathrm{HDP}}(H)).$$

Let $H \in \mathrm{HDP}_n$. Let $L = \log_{\mathrm{HDP}}(H)$. Then $\frac{1}{2}L \in \mathrm{H}_n$ so $\exp(\frac{1}{2}L) \in \mathrm{HDP}_n$ and by the sum property, $\exp(\frac{1}{2}L)^2 = \exp(L) = H$. Moreover, if $K \in \mathrm{HDP}_n$ satisfies $K^2 = H$ then $2\log_{\mathrm{HDP}}(K) = \log_{\mathrm{HDP}}(H) = L$. This proves not only existence and uniqueness, but also continuity as all functions involved are continuous. □

The second lemma is a key tool when studying real and complex Lie groups. It builds on the simple observation that every complex number $z \neq 0$ can be written uniquely as $z = \rho e^{i\theta}$, with $\rho \in \mathbb{R}_{>0} = \mathrm{HDP}_1$ and $e^{i\theta} \in \mathbb{S}^1 = \mathrm{U}_1$.

5.3.3. Lemma (polar decomposition)**.** *Let $P \in \mathrm{GL}_n(\mathbb{C})$. Then there is a unique pair $(U, H) \in \mathrm{U}_n \times \mathrm{HDP}_n$ such that $P = UH$ (they need not commute). Moreover, multiplication $\cdot\colon \mathrm{U}_n \times \mathrm{HDP}_n \to \mathrm{GL}_n(\mathbb{C})$ is a homeomorphism.*

However, there is no reason for U and H to commute.

Proof. Matrix P^*P is hermitian, definite and positive: Hermite-symmetry is obvious. Now suppose $\lambda \in \mathrm{Sp}(P^*P)$ and let $X \in \mathbb{C}^n \smallsetminus \{0\}$ be an eigenvector. Then $P^*PX = \lambda X$; multiplying on the left by X^* we get:

$$\lambda X^*X = X^*P^*PX = (PX)^*(PX).$$

However $\langle Y|Z\rangle = Y^*Z$ defines a complex scalar product on $\mathbb{C}^n$, so X^*X and $(PX)^*(PX)$ are in $\mathbb{R}_{>0}$ (the latter, since P is invertible). There remains $\lambda > 0$.

Let $H = \sqrt{P^*P} \in \mathrm{HDP}_n$ be its square root; also let $U = P \cdot H^{-1}$. It remains to show that U is unitary, but indeed, using Hermite-symmetry of H:

$$U^*U = H^{-*}P^*PH^{-1} = H^{-1}H^2H^{-1} = I_n,$$

as desired.

It remains to prove uniqueness. Suppose $UH = U_1H_1$ in obvious notation. Then $H_1^*H_1 = H_1^*U_1^*U_1H_1 = H^*U^*UH = H^*H$, so by uniqueness of the square root in HDP_n, we have $H_1 = H$; then $U_1 = U$ follows.

We move to the second claim. Multiplication certainly is continuous. We just proved it is bijective. As a matter of fact, we constructed the reciprocal mapping: $P \mapsto (P \cdot \sqrt{P^*P}^{-1}, \sqrt{P^*P})$. Its continuity comes from that of $\sqrt{\cdot}$. □

Proof of theorem 5.3.1. We even prove that U_n is path-connected. Let $U \in \mathrm{U}_n \subseteq \mathrm{GL}_n(\mathbb{C})$; since the latter space is path-connected, there is $\gamma\colon [0,1] \to \mathrm{GL}_n(\mathbb{C})$ with $\gamma(0) = I_n$ and $\gamma(1) = U$. Now write the polar decomposition $\gamma(t) = \upsilon(t)\eta(t)$ with $\upsilon(t) \in \mathrm{U}_n$ and $\eta(t) \in \mathrm{H}_n(t)$. Then both υ and η are continuous maps. Now $\upsilon\colon [0,1] \to \mathrm{U}_n$ satisfies $\upsilon(0) = I_n$ and $\upsilon(1) = U$: it is a continuous path in U_n, as desired. □

5.4 Exercises

5.4.1. Exercise. *Prove that if $n > 1$ there are no conjugacy-invariant norms (viz. satisfying $N(PAP^{-1}) = N(A)$ for all matrices A and P with P invertible).*

Solution. Matrices $A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $A_\varepsilon \begin{pmatrix} 0 & \varepsilon \\ 0 & 0 \end{pmatrix}$ are conjugate for all $\varepsilon \in \mathbb{R}_{>0}$. But $A_\varepsilon \xrightarrow[\varepsilon \to 0]{} A_0$; by continuity, a conjugacy-invariant norm would have $\|A_1\| = \lim \|A_\varepsilon\| = \|A_0\| = 0$, a contradiction.

5.4.2. Exercise. *Give an algebraic proof of the existence and uniqueness of $\sqrt{\cdot}$, not using that $\exp\colon \mathrm{H}_n \to \mathrm{HDP}_n$ is a homeomorphism. (Use unitary trigonalisation and Lagrange interpolation.)*

Prove that $\sqrt{H} \in \mathbb{C}[H]$ and that uniqueness holds among hermitian matrices with spectrum in $\mathbb{R}_{\geq 0}$.

Solution. Let $H \in \mathrm{HDP}_n$. We first prove existence of $\sqrt{H}$ as a polynomial in H. By unitary trigonalisation, there is a unitary matrix U with $UHU^{-1} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$; by definition, all λ_i are in $\mathbb{R}_{>0}$. So each is a square in $\mathbb{R}_{>0}$, say $\lambda_i = \mu_i^2$. Moreover, using Lagrange interpolation (lemma 3.3.3), there is a polynomial P such that $P(\mu_i) = \lambda_i$. So let $K = P(H) = U^{-1}\begin{pmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_n\end{pmatrix}U$, which is clearly hermitian definite positive, and squares to H. Clearly $K^2 = H$; notice that $K \in \mathbb{C}[H]$.

We turn to uniqueness. Let K' be another candidate; notice that K' and $K'^2 = H$ must commute. Since $K \in \mathbb{C}[H]$ and $H \in \mathbb{C}[K']$, we have $K \in \mathbb{C}[K']$. In particular, when we diagonalise K' we also diagonalise K. Let us do it: we find two diagonal matrices D, D' with spectrum $\subseteq \mathbb{R}_{>0}$ which have the same square. Therefore $D = D'$ and $K = K'$.

5.4.3. Exercise. *Give a topological proof of the continuity of* $\sqrt{\cdot}$*, not using that* $\exp\colon \mathrm{H}_n \to \mathrm{HDP}_n$ *is a homeomorphism. (Follow step 3 of the proof of theorem 5.2.1, viz. use exercise 1.4.2.)*

Solution. First notice that the argument for uniqueness of $\sqrt{\cdot}$ in exercise 5.4.2 proves slightly more: if $K' \in \mathrm{H}_n$ has spectrum in $\mathbb{R}_{\geq 0}$ and satisfies $K'^2 = H$, then $K' = K$. We call this property 'strong uniqueness'.

We use a sequential characterisation. Suppose $H_n \longrightarrow H$ in HDP_n; let $K_n = \sqrt{H_n}$ and $K = \sqrt{H}$; we want to show that $K_n \longrightarrow K$. Using exercise 1.4.2, it is enough to prove that (K_n) is bounded and has at most one accumulation point in $M_n(\mathbb{C})$, namely K.

We prove boundedness. By unitary diagonalisation, say $K_n = U_n D_n U_n^{-1}$ with $U_n \in \mathrm{U}_n$ and D_n is diagonal with spectrum in $\mathbb{R}_{>0}$. Take a norm $\|\cdot\|$ invariant under unitary conjugation, and also norm $\|\cdot\|_\infty$ which is the maximum of moduli of coefficients. Notice that for diagonal D, one has $\|D\|_\infty = \max_{\mathrm{Sp}(D)}|\lambda|$ and therefore $\|D^2\|_\infty = \|D\|_\infty^2$. By equivalence of norms there is $c \in \mathbb{R}_{>0}$ such that for any matrix M one has $\|M\| \leq c\|M\|_\infty$ and $\|M\|_\infty \leq c\|M\|$. Now:

$$\begin{aligned}
\|K_n\| &= \|U_n D_n U_n^{-1}\| \\
&= \|D_n\| \\
&\leq c \cdot \|D_n\|_\infty \\
&= c\sqrt{\|D_n^2\|_\infty} \\
&\leq c\sqrt{c}\sqrt{\|D_n^2\|} \\
&= c\sqrt{c}\sqrt{\|H_n\|},
\end{aligned}$$

which is bounded since (H_n) converges.

We prove uniqueness of the accumulation point. Suppose $K_{\varphi(n)} \longrightarrow L$ for some matrix $L \in M_n(\mathbb{C})$. We must prove $L = K$. By strong uniqueness it suffices to show that $L^2 = H$ and $L \in \mathrm{H}_n$ has spectrum in $\mathbb{R}_{\geq 0}$. The former is clear by continuity of multiplication; we prove the latter. As a limit of hermitian matrices, L itself is hermitian; it remains to show that its spectrum is in $\mathbb{R}_{\geq 0}$. We refrain from using continuity of eigenvalues as a multiset. Instead, return to $K_n = U_n D_n U_n^{-1}$. Since U_n is compact, the sequence (U_n), and even its subsequence $(U_{\varphi(n)})$, has a converging subsequence $U_{\chi(n)} \longrightarrow V \in \mathrm{U}_n$; one still has $K_{\chi(n)} \longrightarrow L$. Then by continuity $D_{\chi(n)} = U_{\chi(n)} K_{\chi(n)} U_{\chi(n)}^{-1} \longrightarrow VLV^{-1}$ is a limit of diagonal matrices with diagonal entries in $\mathbb{R}_{>0}$; therefore it is a diagonal matrix with entries in $\mathbb{R}_{\geq 0}$; this is what we claimed on $\mathrm{Sp}(L)$.

5.4.4. Exercise (a brutal analyst's proof of the continuity of $\sqrt{\cdot}$)**.** *There are many other proofs that* $\sqrt{\cdot}\colon \mathrm{HDP}_n \to \mathrm{HDP}_n$ *is continuous. Here is an elementary one, which however* *requires some computations. Here, $\| \cdot \|_2$ on $\mathbb{C}^n$ denotes the usual norm $\|X\|_2 = \sqrt{X^*X} = \sum_{i=1}^n |x_i|^2$; and $|||\cdot|||_2$ is the associated operator norm.*

1. *Let (H_n) be a sequence of matrices in H_n satisfying:*

$$\max_{\lambda \in \mathrm{Sp}(H_n)} |\lambda| \xrightarrow[n\to\infty]{} 0.$$

Prove that $H_n \longrightarrow 0$.

2. *Let $A \in \mathrm{HDP}_n$ be a hermitian, definite positive matrix and $\lambda \in \mathrm{Sp}(A) \subseteq \mathbb{R}$ be its* least *eigenvalue. Prove that:*

$$\min_{\|X\|_2=1} X^*AX \geq \lambda.$$

3. *Let $A, B \in \mathrm{HDP}_n$ and μ be the* maximal *eigenvalue, in absolute value, of $\sqrt{A} - \sqrt{B}$. Let X be an associated eigenvector with $\|X\|_2 = 1$. Prove that:*

$$\mu X^*(\sqrt{A} + \sqrt{B})X \leq |||A - B|||_2 .$$

Hint: compute $X^*[(\sqrt{A} - \sqrt{B})\sqrt{A} + \sqrt{B}(\sqrt{A} - \sqrt{B})]X$.

4. *Conclude.*

Solution.

1. Since all norms are equivalent and there is a norm invariant under conjugation by U_n, we may suppose that our matrices are diagonal with real spectrum. Then the assumption on the spectrum immediately gives that the sequence goes to 0 with respect to norm $\| \cdot \|_\infty$, hence with respect to any norm.

   (This is completely false if one just assumes diagonalisability of course.)

2. Since A is hermitian there is an orthonormal basis consisting of eigenvectors, say $X_1, \dots, X_n \in \mathbb{C}^n$ with eigenvalues $\lambda_i \in \mathbb{R}_{>0}$. Let X have norm 1. Then $X = \sum_{i=1}^n z_i X_i$ in obvious notation, whence $\sum_{i=1}^n |z_i|^2 = 1$ and using orthonormality:

$$X^*AX = \sum_{i,j=1\dots n} \overline{z_i} X_i^* \lambda_j z_j X_j = \sum_{i=1}^n |z_i|^2 \lambda_i \geq \lambda \sum_{i=1}^n |z_i|^2 = \lambda.$$

3. We have $(\sqrt{A} - \sqrt{B})X = \mu X$, so $X^*(\sqrt{A} - \sqrt{B})^* = \overline{\mu} X^*$. But $\sqrt{A} - \sqrt{B}$ is hermitian and μ is real, so it reduces to $X^*(\sqrt{A} - \sqrt{B}) = \mu X^*$. Therefore:

$$\begin{aligned}\mu X^*(\sqrt{A} + \sqrt{B})X &= X^*\mu\sqrt{A}X + X^*\sqrt{B}\mu X \\ &= X^*[(\sqrt{A} - \sqrt{B})\sqrt{A} + \sqrt{B}(\sqrt{A} - \sqrt{B})]X \\ &= X^*(A - B)X.\end{aligned}$$

   The latter is the complex scalar product $\langle X|(A-B)X\rangle$, which by the Cauchy-Schwarz inequality is at most $\|X\|_2 \cdot \|(A-B)X\|_2 \leq |||A - B|||_2$.

4. Let $A \in \mathrm{HDP}_n$ be fixed and $\varepsilon \in \mathbb{R}_{>0}$ be given. Let $B \in \mathrm{HDP}_n$ be such that $|||A - B|||_2$ is very small. Let λ be the least eigenvalue of A. Let μ be the maximal eigenvalue, in absolute value, of $\sqrt{A} - \sqrt{B}$ and X of norm 1 witness it (these depend on B). Then by question 2 one has $X^* \sqrt{A} X \geq \lambda$ and $X^* \sqrt{B} X \geq 0$, so by question 3:

$$\mu \leq \frac{|||A - B|||_2}{X^*(\sqrt{A} + \sqrt{B})X} \leq \frac{|||A - B|||_2}{\lambda}$$

becomes very small. Recall that μ was the largest eigenvalue (in absolute value) of $\sqrt{A} - \sqrt{B}$. By question 1, if B goes to A, then μ goes to 0 and $\sqrt{B}$ goes to $\sqrt{A}$: this is continuity.

5.4.5. Exercise (a review exercise: the three logarithms).

- *Let* $\log_{\mathcal{Y}}: \mathcal{Y} \to \mathcal{N}$ *be the logarithm from the set* $\mathcal{Y}$ *of unipotent matrices to the set* $\mathcal{N}$ *of nilpotent matrices, as implicitly defined in § 3.3.*

- *Let* $\log_{\mathcal{V}(I_n)}: \mathcal{V}(I_n) \to \mathcal{V}(0)$ *be the logarithm from a small neighbourhood of* $I_n \in \mathrm{GL}_n(\mathbb{C})$ *to a small neighbourhood of* $0 \in M_n(\mathbb{C})$*, as defined in § 4.2.*

- *Let* $\log_{\mathrm{HDP}}: \mathrm{HDP}_n \to \mathrm{H}_n$ *be the logarithm from hermitian definite positive matrices to hermitian matrices, as defined in § 5.2.*

Prove that if a matrix is in the domain of any two of them, then the logarithms compute the same image.

Solution. Let $M \in \mathcal{Y} \cap \mathcal{V}(I_n)$; write $M = I_n + h$ where h is a small matrix (typically $\|h\| < 1$ for some submultiplicative norm). Then by the analytic theory, $\log_{\mathcal{V}(I)}(M) = \sum_{k \geq 1} (-1)^{k+1} \frac{h^k}{k}$. But here $h = M - I_n$ is nilpotent, so the series is a finite sum, and computes exactly $\log_{\mathcal{Y}}(I_n + h) = \log_{\mathcal{Y}}(M)$.

Second, $\mathcal{Y} \cap \mathrm{HDP} = \{I_n\}$. Indeed, if $M \in \mathcal{Y} \cap \mathrm{HDP}$, then by definition of a unipotent matrix, $M = I_n + N$ for some nilpotent matrix, which must be hermitian. In particular N is diagonalisable, and nilpotent, so $N = 0$ and $M = I_n$. Clearly both $\log_{\mathcal{Y}}$ and $\log_{\mathrm{HDP}}$ map I_n to 0.

Last, let $M \in \mathcal{V}(I_n) \cap \mathrm{HDP}$; typically $M = I_n + H$ where H is a small hermitian matrix. Then $\log_{\mathcal{V}(I_n)}(M)$ is given by:

$$\log_{\mathcal{V}(I_n)}(I_n + h) = \sum_{k \geq 1} (-1)^{k+1} \frac{h^k}{k}.$$

Notice that at each finite stage, the partial sum is in H_n, which being a subspace is closed. Therefore $\log_{\mathcal{V}(I_n)}(M) \in \mathrm{H}_n$. Since $\exp: \mathrm{H}_n \to \mathrm{HDP}_n$ is injective, we find $\log_{\mathcal{V}(I_n)}(M) = \log_{\mathrm{HDP}}(M)$.

6 Application to linear systems of differential equations

We finally return to the last qestion of § 2.1.

Q$_3$. What are the applications of $\exp(A)$ to differential equations?

§ 6.1 gives a method to reduce scalar equations of order n to vector equations of order 1. In § 6.2 we see how to use the matrix exponential to solve such equations with constant coefficients. Some further aspects are discussed in §§ 6.3.

6.1 Reducing the order of a linear ODE

Recall a fundamental method. A differential equation is *scalar* if the unknown function $x(t)$ takes values in $\mathbb{K}$; if it takes values in $\mathbb{K}^n$, call it a *vector* differential equation. The *order* of a differential equation is the largest n such that $x^{(n)}$ appears in it. We assume its coefficient to be 1.

6.1.1. Lemma. *Every scalar linear differential equation of order n can be rewritten as a vector linear differential equation of order* 1.

Proof. Never forget this. Consider equation $x^{(n)} = a_0(t)x(t)+\cdots+a_{n-1}(t)x^{(n-1)}(t)$. Then introduce the variable vector:

$$X(t) = \begin{pmatrix} x(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix} \in \mathbb{K}^n.$$

Notice that if $x(t)$ is C^n, then $X(t)$ is C^1; moreover

$$X'(t) = \begin{pmatrix} x'(t) \\ \vdots \\ x^{(n)}(t) \end{pmatrix}$$

now obeys differential equation:

$$X'(t) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ a_0(t) & a_1(t) & \dots & a_{n-2}(t) & a_{n-1}(t) \end{pmatrix} \cdot X(t),$$

viz. $X'(t) = A(t)\cdot X(t)$. □

Be very careful that the matrix exponential works only if coefficients are *constant*.

6.2 Using the matrix exponential

The fundamental result is now easy to state and prove.

6.2.1. Lemma. *If $X(t)\colon \mathbb{R} \to \mathbb{K}^n$ is differentiable and there is $A \in M_n(\mathbb{K})$ with $X'(t) = A\cdot X(t)$, then $X(t) = \exp(tA)\cdot X(0)$.*

Repeated warning: no such thing with a variable matrix $A(t)$. (Return to the formula in remark 4.1.3 to convince yourself that differentiating $\exp(\int A(t)dt)$ is not pleasant.)

Proof. Let $Y(t) = X(t) - \exp(tA)\cdot X(0)$; using the assumptions, it is differentiable with derivative identically 0. So Y is constant and equals $Y(0) = 0$, as claimed. □

With this at hand we can finally explain, and generalise, a basic method.

6.2.2. Example. To solve linear equation $x''(t) - 2x'(t) + x(t) = 0$, one is told to first solve polynomial $\lambda^2 - 2\lambda + 1 = 0$ (with double root 1) and then write $x(t) = c_1e^t + x_2te^t$. We explain why.

Verification. Let $X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix}$. Then $X(t)$ is governed by the differential equation $X'(t) = A \cdot X(t)$ where $A = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$. In particular, vector solutions are of the form $X(t) = \exp(tA) \cdot X(0)$. Let us compute $\exp(tA)$.

Notice that the characteristic polynomial of A is $\lambda^2 - 2\lambda + 1$, with double root 1; the matrix is not diagonalisable, but $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector. Any independent vector will provide a trigonalisation basis; consider $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The relevant change of basis is coded in matrix $P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$; notice how:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}_{P} \cdot \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_{T} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}}_{P^{-1}}.$$

Now $\exp(tT)$ is easily computed since the Chevalley decomposition is obvious here; in particular, for $t \in \mathbb{R}$:

$$\begin{aligned} \exp(tA) &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}_P \cdot \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} (1-t)e^t & te^t \\ -te^t & (1+t)e^t \end{pmatrix}. \end{aligned}$$

By lemma 6.2.1 one has $X(t) = \exp(tA) \cdot X(0)$. So the first coordinate is:

$$x(t) = x(0)(1-t)e^t + x'(0)te^t = x(0)e^t + (x'(0) - x(0))te^t,$$

as predicted by the method. ◇

6.3 A stability lemma

6.3.1. Lemma. *Let* $M \in M_n(\mathbb{C})$*. Then* $\exp(tM) \xrightarrow[t\to+\infty]{} 0$ *iff* $(\forall \lambda \in \operatorname{Sp} M)(\operatorname{Re} \lambda < 0)$.

Proof. There are two implications to prove.

⇒. Suppose $\exp(tM) \longrightarrow 0$ as $t \to +\infty$. Let $\lambda \in \operatorname{Sp}(M)$; let $X \in \mathbb{C}^n \setminus \{0\}$ with $MX = \lambda X$. Fix $t \in \mathbb{R}$ and compute:

$$E_n(tM) \cdot X = \sum_{k=0}^{n} \frac{t^k M^k}{k!} X = \sum_{k=0}^{n} \frac{t^k \lambda^k}{k!} X,$$

so that $E_n(tM) \cdot X \xrightarrow[n\to\infty]{} \exp(\lambda t) \cdot X$. On the other hand, by definition and continuity of $\cdot$, one also has $E_n(tM) \cdot X \xrightarrow[n\to\infty]{} \exp(tM) \cdot X$. Hence:

$$\exp(tM) \cdot X = e^{\lambda t} X,$$

in words: X is an eigenvector for $\exp(tM)$, with eigenvalue $\exp(\lambda t)$.

We now let t go to $+\infty$. By assumption $\exp(tM) \longrightarrow 0$ as $t \to +\infty$. So $e^{\lambda t} X \longrightarrow 0$. But $X \neq 0$, and therefore $e^{\lambda t} \longrightarrow 0$. We know from complex calculus that this happens only when $\operatorname{Re} \lambda < 0$.

$\Leftarrow$. This implication is more subtle. Write $M = D + N$ in Chevalley decomposition. Of course $\operatorname{Sp}(M) = \operatorname{Sp}(D)$. Now $tM = tD + tN$ where the terms commute, so:

$$\exp(tM) = \exp(tD) \cdot \exp(tN).$$

We estimate each. The nilpotent contribution is no surprise:

$$\exp(tN) = \sum_{k=0} n t^k \frac{N^k}{k!} = o(t^{n+1}).$$

In order to estimate the diagonal term we shall work with two different norms in the argument. Let $\lambda \in \operatorname{Sp}(M)$ be the eigenvalue with largest $\operatorname{Re} \lambda$, say $\operatorname{Re} \lambda = -\varepsilon$ with $\varepsilon > 0$ (by assumption). There exists an invertible matrix P such that:

$$D = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1},$$

so that for $t \in \mathbb{R}$:

$$\exp(tD) = P \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} P^{-1}.$$

Now P is fixed; first fixing any submultiplicative norm $\|\cdot\|$ and letting $c_0 = \|P\| \cdot \|P^{-1}\|$ we get:

$$\|\exp(tD)\| \leq c_0 \left\| \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} \right\|.$$

And now we change norm in the middle of a computation: consider $\|A\|_\infty = \max_{i,j=1\ldots n} |a_{i,j}|$. Since all norms on $M_n(\mathbb{C})$ are equivalent, there is a constant c such that $\|A\| \leq c\|A\|_\infty$. Therefore $\|\exp(tD)\| \leq c_0 c \max_{k=1\ldots n} |e^{\lambda_k t}| = O(e^{-\varepsilon t})$.

Altogether, $\exp(tM) = o(t^{n+1} e^{-\varepsilon t}) \longrightarrow 0$ as $t \to +\infty$. □

Analysts may wish to remember that if $\varepsilon \in \mathbb{R}_{>0}$ is such that $(\forall \lambda \in \operatorname{Sp}(A))(\lambda \leq -\varepsilon)$, then there is $c \in \mathbb{R}$ such that $\|\exp(tA)\| \leq c e^{-\varepsilon t}$.

6.4 Exercises

6.4.1. Exercise (radioactive decay)**.** *Suppose three elements A, B, C decay with respect to the transitions:*

$$A \to B \to C,$$

meaning both A and B are radioactive and transform into the next element, while C is stable. Say at time t there is $a(t)$ of element A; define $b(t)$ and $c(t)$ likewise. Element A has a decay rate $\alpha > 0$; define β likewise. (Element C being stable means $\gamma = 0$.)

Decay is rendered by the following equations:

- $a'(t) = -\alpha a(t)$;
- $b'(t) = \alpha a(t) - \beta b(t)$;
- $c'(t) = \beta b(t)$.

Suppose $\alpha > \beta$. Suppose we started with 1 unit of A, viz. $a(0) = 1$ while $b(0) = c(0) = 0$. Find the time when B reaches its maximal quantity, viz. t_0 maximising $b(t_0)$.

Solution. We convert the problem to matrix form. Let $X(t)$ be the vector $\begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix}$. By the equations, one has:

$$X'(t) = \underbrace{\begin{pmatrix} -\alpha & 0 & 0 \\ \alpha & -\beta & 0 \\ 0 & \beta & 0 \end{pmatrix}}_{M} \cdot X(t).$$

We shall find the Chevalley decomposition of M. Since the matrix is lower-triangular, its eigenvalues are $-\alpha, -\beta, 0$ which are distinct by assumption, so M is diagonalisable.

Eigenvectors are readily computed:

$$\begin{aligned}
\ker(M + \alpha I_3) &= \ker \begin{pmatrix} 0 & 0 & 0 \\ \alpha & \alpha - \beta & 0 \\ 0 & \beta & \alpha \end{pmatrix} = \left\langle \begin{pmatrix} \alpha - \beta \\ -\alpha \\ \beta \end{pmatrix} \right\rangle; \\
\ker(M + \beta I_3) &= \ker \begin{pmatrix} \beta - \alpha & 0 & 0 \\ \alpha & 0 & 0 \\ 0 & \beta & \beta \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\rangle; \\
\ker(M - 0 I_3) &= \ker \begin{pmatrix} -\alpha & 0 & 0 \\ \alpha & -\beta & 0 \\ 0 & \beta & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.
\end{aligned}$$

This suggests to form matrix $P = \begin{pmatrix} \alpha - \beta & 0 & 0 \\ -\alpha & 1 & 0 \\ \beta & -1 & 1 \end{pmatrix}$. Then:

$$M = \underbrace{\begin{pmatrix} \alpha - \beta & 0 & 0 \\ -\alpha & 1 & 0 \\ \beta & -1 & 1 \end{pmatrix}}_{P} \cdot \underbrace{\begin{pmatrix} -\alpha & & \\ & -\beta & \\ & & 0 \end{pmatrix}}_{D} \cdot \underbrace{\begin{pmatrix} \frac{1}{\alpha-\beta} & 0 & 0 \\ \frac{\alpha}{\alpha-\beta} & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}}_{P^{-1}}.$$

Therefore for real t:

$$\begin{aligned}\exp(tM) &= \begin{pmatrix} \alpha-\beta & 0 & 0 \\ -\alpha & 1 & 0 \\ \beta & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} e^{-\alpha t} & & \\ & e^{-\beta t} & \\ & & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\alpha-\beta} & 0 & 0 \\ \frac{\alpha}{\alpha-\beta} & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} (\alpha-\beta)e^{-\alpha t} & 0 & 0 \\ -\alpha e^{-\alpha t} & e^{-\beta t} & 0 \\ \beta e^{-\alpha t} & -e^{-\beta t} & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\alpha-\beta} & 0 & 0 \\ \frac{\alpha}{\alpha-\beta} & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{-\alpha t} & 0 & 0 \\ \frac{\alpha}{\alpha-\beta}(e^{-\beta t}-e^{-\alpha t}) & e^{-\beta t} & 0 \\ 1-\frac{\alpha e^{-\beta t}-\beta e^{-\alpha t}}{\alpha-\beta} & 1-e^{-\beta t} & 1 \end{pmatrix}\end{aligned}$$

Moreover, $X(t) = \exp(tM)\cdot X(0) = \exp(tM)\cdot\begin{pmatrix}1\\0\\0\end{pmatrix}$, which is the first column of $\exp(tM)$. We are only interested in $b(t)$, viz.:

$$b(t) = (\exp(tM))_{2,1} = \frac{\alpha}{\alpha-\beta}(e^{-\beta t}-e^{-\alpha t}).$$

Now $b(t)$ is maximal when its derivative vanishes, that is when $\alpha e^{-\alpha t} = \beta e^{-\beta t}$. This happens at:

$$t_0 = \frac{\log\alpha - \log\beta}{\alpha-\beta}.$$

6.4.2. Exercise (one-parameter subgroups)**.** *Let $f\colon\mathbb{R}\to \mathrm{GL}_n(\mathbb{K})$ be a 'one-parameter subgroup', a geometer's terminology for a continuous group homomorphism, viz.:*

$$(\forall s,t\in\mathbb{R})(f(s+t) = f(s)\cdot f(t)).$$

We shall prove that there is $A\in M_n(\mathbb{K})$ such that $(\forall t\in\mathbb{R})(f(t)=\exp(tA))$.

1. *Suppose in addition that f is differentiable. Prove the result.*
2. *We simply assume that f is continuous. Let F be the primitive of f vanishing at 0. Prove that for all $a>0$ sufficiently small, $F(a)\in\mathrm{GL}_n(\mathbb{K})$. (Hint: bound $\|aI_n - F(a)\|$.)*
3. *Conclude that f is differentiable.*

Solution. The assumptions clearly imply $f(0)=f(0)^2$; since $f(0)$ is invertible, one has $f(0)=I_n$.

1. Fix t and differentiate with respect to s, obtaining $f'(s+t) = f'(s)\cdot f(t)$. Now apply to $s=0$, getting $f'(t) = f'(0)\cdot f(t)$; as we know, this solves into $f(t) = \exp(tA)\cdot f(0) = \exp(tA)$.
2. A priori $F\colon\mathbb{R}\to M_n(\mathbb{K})$, and we do not know whether it takes invertible values. However $f(0)=I_n$ and f is continuous. Fix any norm on $M_n(\mathbb{K})$. By continuity

at 0 there is a small open interval, say $U = (0, \delta)$, such that such that for $a \in U$ one has $\|f(a) - I_n\| < 1$. Therefore on U:

$$\begin{aligned}\|F(a) - aI_n\| &= \left\| \int_{t=0}^{a} f(t)dt - aI_n \right\| \\ &\leq \int_{t=0}^{a} \|f(t) - I_n\|\, dt \\ &\leq a.\end{aligned}$$

In particular for sufficiently small $a > 0$ one has $\|F(a) - aI_n\| < 1$. We contend that $F(a)$ is then invertible. Indeed, write $F(a) = aI_n + M$ where $\|M\| < 1$. Then the series

3. Notice that for fixed $a, t \in \mathbb{R}$:

$$F(a+t) - F(t) = \int_{s=0}^{a} f(t+s)ds = \int_{s=0}^{a} f(t) \cdot f(s)ds = f(t) \cdot F(a).$$

By the above, for sufficiently small $a > 0$, one has $f(t) = (F(a+t) - F(t)) \cdot F(a)^{-1}$. But F is C^1, and therefore so is f. We are done.

6.4.3. Exercise. *Let $\mathbb{S}^1$ be the unit circle; write* $R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Let $\varphi : \mathbb{S}^1 \to \mathrm{GL}_n(\mathbb{R})$ be a continuous group homomorphism. Prove that there are $P \in \mathrm{GL}_n(\mathbb{R})$, $d \in \mathbb{N}$, and $k_1, \ldots, k_d \in \mathbb{Z}$ such that for any $t \in \mathbb{R}$:

$$\varphi(e^{it}) = P \begin{pmatrix} R(k_1 t) & & & & & \\ & \ddots & & & & \\ & & R(k_d t) & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} P^{-1}.$$

Solution. Let $\chi(t) = \varphi(e^{it})$, which is a continuous group homomorphism from $\mathbb{R}$ to $\mathrm{GL}_n(\mathbb{R})$, that is, a real 'one-parameter subgroup'. By exercise 6.4.2, there is $A \in M_n(\mathbb{R})$ such that $(\forall t \in \mathbb{R})(\chi(t) = \exp(tA))$.

Notice that:

$$\exp(2\pi A) = \chi(2\pi) = \varphi(e^{2i\pi}) = \varphi(1) = I_n,$$

so $2\pi A$ is diagonalisable (over $\mathbb{C}$), with spectrum $\subseteq 2i\pi\mathbb{Z}$. (Both claims have been proved in exercises; if necessary, prove them.) Therefore A is diagonalisable with spectrum $\subseteq i\mathbb{Z}$. But A is a real matrix, so its non-real eigenvalues come in pairs $\pm ik_j$ with $k_j \in \mathbb{Z}$; the only real eigenvalue is 0.

So far, A is conjugate in $\mathrm{GL}_n(\mathbb{C})$ to the matrix:

$$D = \begin{pmatrix} \begin{bmatrix} ik_1 & \\ & -ik_1 \end{bmatrix} & & & & & \\ & \ddots & & & & \\ & & \begin{bmatrix} ik_d & \\ & -ik_d \end{bmatrix} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

with first blocks $\pm k_j$ and then 0's. But D itself is $\mathrm{GL}_n(\mathbb{C})$-conjugate to:

$$C = \begin{pmatrix} \begin{bmatrix} & -k_1 \\ k_1 & \end{bmatrix} & & & & & \\ & \ddots & & & & \\ & & \begin{bmatrix} & -k_d \\ k_d & \end{bmatrix} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

with first blocks $\begin{pmatrix} & -k_j \\ k_j & \end{pmatrix}$ and then 0's.

Now A and C are real matrices conjugate in $\mathrm{GL}_n(\mathbb{C})$, so it is a standard exercise that there is $P \in \mathrm{GL}_n(\mathbb{R})$ with $A = PCP^{-1}$. In particular, for any real number t one has $\exp(tA) = P\exp(tC)P^{-1}$, and $P \in \mathrm{GL}_n(\mathbb{R})$ does not depend on t. But for any real θ one has:

$$\exp\begin{pmatrix} & -\theta \\ \theta & \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = R(\theta).$$

So finally, computing blockwise,

$$\varphi(e^{it}) = \chi(t) = \exp(tA) = P\exp(tC)P^{-1} = P\begin{pmatrix} R(k_1 t) & & & & \\ & \ddots & & & \\ & & R(k_d t) & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}P^{-1}.$$