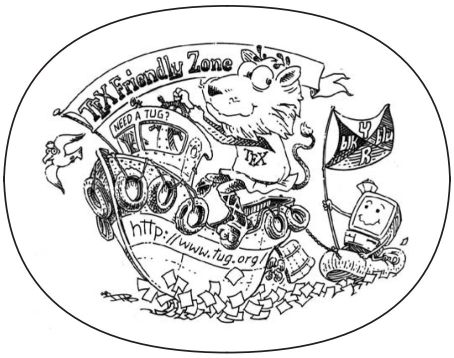
A CLEAN TITLE

LUCAS KERBS



A Fun Subtitle

February 2022 – LucasThesis v1

Ohana means family.
Family means nobody gets left behind, or forgotten.

— Lilo & Stitch

Dedicated to the loving memory of Rudolf Miede.

1939 – 2005

ABSTRACT

Short summary of the contents in English. . . a great guide by Kent Beck how to write good abstracts can be found here:

https://plg.uwaterloo.ca/~migod/research/beckOOPSLA.html

We have seen that computer programming is an art,
because it applies accumulated knowledge to the world,
because it requires skill and ingenuity, and especially
because it produces objects of beauty.

— Donald E. Knuth [3]

ACKNOWLEDGMENTS

Put your acknowledgments here.

Many thanks to everybody who already sent me a postcard!

Regarding the typography and other help, many thanks go to Marco Kuhlmann, Philipp Lehman, Lothar Schlesier, Jim Young, Lorenzo Pantieri and Enrico Gregorio[1], Jörg Sommer, Joachim Köstler, Daniel Gottschlag, Denis Aydin, Paride Legovini, Steffen Prochnow, Nicolas Repp, Hinrich Harms, Roland Winkler, Jörg Weber, Henri Menke, Claus Lahiri, Clemens Niederberger, Stefano Bragaglia, Jörn Hees, Scott Lowe, Dave Howcroft, José M. Alcaide, David Carlisle, Ulrike Fischer, Hugues de Lassus, Csaba Hajdu, Dave Howcroft, and the whole LaTeX-community for support, ideas and some great software.

Regarding LyX: The LyX port was intially done by *Nicholas Mariette* in March 2009 and continued by *Ivo Pletikosić* in 2011. Thank you very much for your work and for the contributions to the original style.

1 Members of GuIT (Gruppo Italiano Utilizzatori di TeX e LaTeX)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

OBJECTS AND THE MAPS BETWEEN THEM

1 A FIRST ATTEMPT

1.1 FUNCTIONAL CALCULUS

Functional Calculus refers to the process of extending the domain of a function on $\mathbb{R}$ to include matrices (or in some cases operators). The most basic formulation uses the fact that the space $n \times n$ matrices forms a ring and so there is a natural way to evaluate polynomials $f \in \mathbb{C}[x]$. If we require that $A \in M_n(\mathbb{C})$ is self-adjoint—and hence diagonalizable as $A = U\Lambda U^*$—then it is a standard result that:

$$\begin{aligned} f(A) &= a_n A^n + \cdots + a_1 A + a_0 \\ &= a_n \left(U\Lambda U^*\right)^n + \cdots + a_1 U\Lambda U^* + a_0 \\ &= a_n U\Lambda^n U^* + \cdots + a_1 U\Lambda U^* + a_0 \\ &= U\left(a_n\Lambda^n + \cdots + a_1\Lambda + a_0\right) U^* \\ &= U\left(f(\Lambda)\right) U^* \end{aligned}$$

Further, since Λ is diagonal and f is a polynomial,

$$f\left(\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}\right) = \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix}$$

Therefore, given a self-adjoint matrix A and a polynomial $f \in \mathbb{C}[x]$

$$f(A) = Uf(\Lambda)U^* = U \operatorname{diag}\{f(\lambda_1), \ldots, f(\lambda_n)\}\, U^*$$

Since self-adjoint matrices play such a vital role in free analysis, we will let $\mathbb{H}_n \subset M_n(\mathbb{C})$ denote the set of $n \times n$-matrices over $\mathbb{C}$. With the polynomial case in mind, we can extend a function $g : [a, b] \to \mathbb{C}$ to a function on self adjoint (normal?) matrices with their spectrum in $[a, b]$. Let A be such a matrix (diagonalized by the unitary matrix U), and define

$$g(A) = U \begin{bmatrix} g(\lambda_1) & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} U^*$$

Thus, for each $n \in \mathbb{N}$, g induces a function on the self-adjoint $n \times n$ matrices with spectrum in $[a, b]$. The natural ordering explain why natural? on self-adjoint matrices is called the **Loewner Order**:

Definition i.1 (Loewner Ordering)**.** *For like size self-adjoint matrices, we say that $A \preceq B$ if $B - A$ is positive semidefinite and $A \prec B$ is $B - A$ is positive definite.*

With this ordering in place, we can extend many of the familiar function theoretic properties (monotonicity, convexity) to these matrix-values functions. In fact, these properties are defined identically to their classical counterpart: We say that a function is *matrix-monotone* if $A \preceq B$ implies that $f(A) \preceq f(B)$ and *matrix-convex* (or *nc-convex*) if

$$f\left(\frac{X+Y}{2}\right) \preceq \frac{f(X)+f(Y)}{2}$$

for every pair of like-size matrice for which f is defined. These condition are rather restrictive (since the must hold for matrices of *all* sizes) so many functions which are convex/monotone (in the traditional sense) fail to be matrix-convex/monotone. For example, $f(x) = x^4$ fails to be nc-convex. Below is an example from FCAC (Helton). What is the best way to refer to this?

Indeed, if

$$X = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

Then

$$\frac{X^4+Y^4}{2} - \left(\frac{1}{2}X + \frac{1}{2}Y\right)^4 = \begin{bmatrix} 164 & 120 \\ 120 & 84 \end{bmatrix}$$

Which is not positive definite! Thus x^4 fails to be convex on even 2×2 matrices.

nc-positive if positive for all matrices

Further, a number of the standard constructions lift identically in this functional calculus.

Definition i.2 (Directional Derivative)**.** *The derivative of f in the direction H is*

$$Df(X)[H] = \lim_{t\to 0} \frac{f(X+tH)-f(X)}{t}$$

where H and X are like-size self-adjoint matices.

Often, the best way to compute these directional derivatives is via an equivalent formulation:

$$Df(X)[H] = \left.\frac{df(X+tH)}{dt}\right|_{t=0}$$

This version allows us to more easily define higher order derivatives

$$D^{(k)}f(X)[H] = \left.\frac{d^{(k)}f(X+tH)}{d^{(k)}t}\right|_{t=0}$$

Despite nc-convexity being so restrictive, Lemma 12 in [2] shows that the standard characterization of convexity via the second derivative: a function f is convex if and only if $D^2 f(X)[H]$ is nc-positive. Unlike the classical case, however, the only convex polynomials are of degree 2[1].

1 See [2] for details.

1.2 EXTENDING MULTI-VARIABLE FUNCTIONS

We can extend this same functional calculus to functions of several variables, although the details are a bit more subtle. We could simply "plug in" at tuple of matrices to a standard multivariable polynomial ring over $\mathbb{R}$ or $\mathbb{C}$, but this ignores the noncommutativity of $M_n(\mathbb{C})$. In light of this, let $x = (x_1, \ldots x_g)$ be a g-tuples of noncommuting formal variables. The formal variables $x_1, \ldots, x_n$ are *free* in the sense that there are no nontrivial relations between them.[2] A **word** in x is a product of these variables (e. g.$x_1x_3x_1x_4^2$ or $x_1^2x_5^3$). An **nc-polynomial** in x is a formal finite linear combination of words in x with coefficients in your favorite field. We use $\mathbb{R}\langle x\rangle$ and $\mathbb{C}\langle x\rangle$ to denote the set of nc-polynomials in x over $\mathbb{R}$ or $\mathbb{C}$ respectively.

With $\mathbb{C}\langle x\rangle$ constructed, we can define the functional calculus. Given a word $w(x) = x_{i_1}^{p_1} \cdots x_{i_d}^{p_d}$ and a g-tuple of self-adjoint matrices, X, we can evaluate w on X via $w(X) = X_{i_1}^{p_1} \cdots X_{i_d}^{p_d}$. Since our nc-polynomials are linear combinations of these words, we can extend this evaluation to evaluation of entire polynomials. Algebraically, we have a natural evaluation map: Given some $f \in \mathbb{C}\langle x\rangle$ and $X = (X_1, \ldots, X_n)$ a g-tuple of self-adjoint matrices, define

$$\begin{aligned} \varepsilon_f : \mathbb{H}_n{}^g &\longrightarrow M_n(\mathbb{C}) \\ X &\longmapsto f(X). \end{aligned}$$

In the context of these multivariate functions, our definition of the Directional Derivative still makes sense (although our direction H now becomes a tuple of directions). We also inherit (from multi-variable calculus) a notion of the **gradient** of a function—but this will require a bit more work.

Maybe an example of an evaluation?

1.3 THE NATURAL INVOLUTION ON NC-POLYNOMIALS

Given our ring of nc polynomials, we may define an involution * which we may view as an extension of the conjugate transpose. Let * reverse the order of words (i. e.$(x_1x_3x_2^2)^* = x_2^2x_3x_1$) and extend linearly to all of $\mathbb{R}\langle x\rangle$. We consider the formal variables $x_1, \ldots, x_n$ *symmetric* in the sense that $x_i^* = x_i$. We say that a polynomial $p \in \mathbb{R}\langle x\rangle$ is symmetric if $p^* = p$. For example, if

$$p(x) = 5x_1^2x_3x_2 + x_3x_2x_3 \qquad q(x) = 3x_2x_1x_2 + x_3^2 - x_1,$$

then a cursory inspection tells that q is symmetric while p is not.

Notice that the majority of the previous two sections breaks down if we try to extend functions to non-self-adjoint matrices. While the

2 This becomes important in the eventual functional calculus—matrices *do* have nontrivial relations. See section [ALGEBRAIC CONSTRUCTION] for the details.

idea of "plugging in" a tuple of matrices so some element of $\mathbb{R}\langle x\rangle$ via the same functional calculus described above, but we can actually *add* structure to $\mathbb{R}\langle x\rangle$ and made evaluation more natural.

Let $x = (x_1, \ldots, x_g)$ be formal variables and let $x^* = (x_1^*, \ldots, x_g^*)$ denote their formal adjoints. Once again, we let the ring $\mathbb{R}\langle x, x^*\rangle$ be the finite formal sums of words in $x_1, x_1^*, \ldots, x_g, x_g^*$ with coefficients in $\mathbb{R}$. Endow $\mathbb{R}\langle x, x^*\rangle$ with an involution $*$ which sends $x_i \mapsto x_i^*$ and $x_i^* \mapsto x_i$ and reverses the order of words extended linearly. Notice that this involution behaves identically to the adjoint with respect to products and sums of matrices. This new ring inherits a natural functional calculus just like that in section 1.2 except it can accept *any* matrix as an input instead of simply self-adjoint matrices.

As discussed in baby linear algebra, we can view the self-adjoint matrices as somehow analogous to the to real numbers while non-self-adjoint mattrices are analogous to complex numbers with nonzero imaginary part how?. In light of this, if we view $\mathbb{R}\langle x\rangle$ as a the nc analogue of polynomials with real variables while $\mathbb{R}\langle x, x^*\rangle$ is the nc analogue of polynomials in complex variables! Techically the analogue is *normal* matrices but honestly I don't wannt think about that right now.

Short section on the involution. Could also put this in the algebra errata

2 A SECOND ATTEMPT

2.1 MATRIX UNIVERSES

Beyond the functional calculus, it becomes useful to construct general functions on spaces of matrices—to do so, we must make this idea of "spaces of matrices" concrete. The largest such space is the so-called **Matrix Universe**—consisting of g-tuples of matrices of all sizes:

$$\mathcal{M}^g = \bigcup_{n=1}^{\infty} (M_n(\mathbb{C}))^g$$

By convention, when we consider some $X = (X_1, \ldots, X_g) \in \mathcal{M}^g$, we require that the X_i are all the same size. Since $\mathcal{M}^g$ is such a large set, we often want to deal with subsets that still carry some of the implicit structure of $\mathcal{M}^g$.

Definition i.3 (Free Set)**.** *We say $D \subset \mathcal{M}^g$ is a **free set** if it is closed with respect to direct sums and unitary conjugation. That it*

1. *$X, Y \in D$ means $X \oplus Y \in D$.*
2. *For X, U like-size matrices with U unitary and $X \in D$, then $UXU^* = (UX_1U^*, \ldots, UX_gU^*) \in D$.*

For the remainder of this text, D will denote some free set. Using the terminology of [4], let $D_n = D \cap M_n(\mathbb{C})^g$ be the level-wise slice of all $n \times n$ matrices in D. We say that D is **open**[1] (resp. **connected**, **simply connected**) if each D_n is open (resp. connected, simply connected). Finally, we say that D is **differentiable** if each D_n is an open C^1 manifold where the complex tangent space of every $X \in D_n$ is all of $M_n(\mathbb{C})^g$.

In the context of sections 1.1 1.2 how to ref multiple sections, the domains in the functional calculus were $\mathbb{H}^g = \bigcup_{n=1}^{\infty} \mathbb{H}_n{}^g$. $\mathbb{H}^g$ is a differentiable a free set with two connected components.

On $\mathcal{M}^g$, we define a product that resembles the inner product on $\mathbb{C}^n$. Given $A, B \in \mathcal{M}^g$ which are g-tuples of $n \times n$ matrices:

$$\cdot : \mathcal{M}^g \times \mathcal{M}^g \longrightarrow M_n(\mathbb{C})$$
$$\cdot(A, B) = A \cdot B \longmapsto \sum_{i=1}^{g} A_i B_i$$

1 The topology of $\mathcal{M}^g$ is still in flux and there is not a canonical topology. See section NUMBER for the details

James uses this product, but like what in the world is going on with it???

$tr(A \otimes Id)$ But its more complicated than that bc A is a "row vector" of sorts

2.2 THE TOPOLOGY OF MATRIX UNIVERSES

How much detail here? Just what we are doing or the basics of the other topologies? This could also talk about the nc varieties/singular sets

2.3 TRACIAL FUNCTIONS AND UNIQUENESS OF THE GRADIENT

Now that we have $\mathcal{M}^d$, we can work with general functions on our matrix universe. As a whole, free analysis is concerned with so-called *free functions*, which respect the direct sums and unitary conjugation. Do they need to be graded?

Definition i.4 (Free Function)**.** *A function $f : D \to \mathcal{M}^{something}$ is called **free** if*

1. $f(X \oplus Y) = f(X) \oplus f(Y)$
2. $f(UXU^*) = f(U)f(X)f(U^*)$ *where X and U are like-size and U is unitary.*

The two other classes of functions we are concerned with are those that act like the trace and the determinant:

Definition i.5 (Determinantal Free Function)**.** *A function $f : D \to \mathbb{C}$ is a **determinantal free function** if*

1. $f(X \oplus Y) = f(X)f(Y)$
2. $f(UXU^*) = f(X)$ *where X and U are like-size and U is unitary.*

Definition i.6 (Tracial Free Function)**.** *A function $f : D \to \mathbb{C}$ is a **tracial free function** if*

1. $f(X \oplus Y) = f(X) + f(Y)$
2. $f(UXU^*) = f(X)$ *where X and U are like-size and U is unitary.*

Given a free function of any type, we can define the directional derivative (Definition i.2) identically. It is worth noting that, while they share the moniker of *free*, determinantal and tracial functions are *not* free functions. It is only these tracial functions which inherit the gradient mentioned above. Similarly to traditional multivariable calculus we define the gradient via its relationship to the directional derivative:

Definition i.7 (Free Gradient)**.** *Given a tracial free function f, the free gradient, ∇f, is the unique free function satisfying*

$$\operatorname{tr}(H \cdot \nabla f(X)) = \operatorname{tr} Df(X)[H]$$

It is not-at-all obvious that such a ∇f should be unique—after all any linear combination of commutator is has trace zero. should I explain this? In the case that f is a single-variable function we can replace ∇f with the traditional derivative, f', as seen in [5, Thm 3.3].

Theorem i.8. *Let $f : (a, b) \to \mathbb{R}$ be a C^1 function. Then*

$$\operatorname{tr} Df(X)[H] = \operatorname{tr}\left(Hf'(X)\right)$$

The proof in [5] simply asserts the uniqueness of a function $g(X)$ and then shows that $g(x) = f'(x)$ for $x \in (a, b)$. Instead, we can construct such a g and recover the theorem along the way:

Proof. We start with a construction from Bhatia's Matrix Analysis: Let $f \in C^1(I)$ and define $f^{[1]}$ on $I \times I$ by

$$f^{[1]}(\lambda, \mu) = \begin{cases} \frac{f(\lambda)-f(\mu)}{\lambda-\mu} & \lambda \neq \mu \\ f'(\lambda) & \lambda = \mu. \end{cases}$$

We call $f^{[1]}(\lambda, \mu)$ the *first divided difference* of f at (λ, μ). If Λ is a diagonal matrix with entries $\{\lambda_i\}$, We may extend f to accept Λ by defining the (i, j)-entry of $f^{[1]}(\Lambda)$ to be $f^{[1]}(\lambda_i, \lambda_j)$. If A is a self adjoint matrix with $A = U\Lambda U^*$, then we define $f^{[1]}(A) = Uf^{[1]}(\Lambda)U^*$. Now we borrow a theorem from Bhatia [1]:

Theorem i.9 (Bhatia V.3.3)**.** *Theorem numbering? Let $f \in C^1(I)$ and let A be a self adjoint matrix with all eigenvalues in I. Then*

$$Df(A)[H] = f^{[1]}(A) \circ H,$$

where $\circ$ denotes the Schur-product[2] in a basis where A is diagonal.

That is, if $A = U\Lambda U^*$, then

$$Df(A)[H] = U\left(f^{[1]}(\Lambda) \circ (U^*HU)\right)U^*.$$

To prove our claim, we need to take the trace of both sides. Since trace is invariant under a change of basis, it is clear that

$$\operatorname{tr} Df(A)[H] = \operatorname{tr}\left(f^{[1]}(\Lambda) \circ (U^*HU)\right).$$

If $U = u_{ij}$, $U^* = \overline{u}_{ij}$ and $H = h_{ij}$, then the (i, j)-entry of U^*HU is

$$(U^*HU)_{ij} = \overline{u}_{ik}h_{k\ell}u_{\ell j}$$

2 Entrywise

Where we sum over the duplicate indices k and ℓ. While the structure of $f^{[1]}(\Lambda)$ is a bit unruly, our diagonal entries are $f'(\lambda)$. This means that when we take the trace of the Schur product, we have

$$\sum_k \sum_\ell \sum_i f'(\lambda_i)\overline{u}_{ik} h_{k\ell} u_{\ell i}$$

Now consider the matrix product $U \operatorname{diag}\{f'(\lambda_1), \dots, f'(\lambda_n)\} U^* H$. Since one of our terms is diagonal, the trace of this multiplication is simple:

$$\operatorname{tr} U \operatorname{diag}\{f'(\lambda_1), \dots, f'(\lambda_n)\} U^* H = \sum_k \sum_\ell \sum_i u_{ik} f'(\lambda_k)\overline{u}_{k\ell} h_{\ell i}$$

Since $u_{ik}, \overline{u}_{k\ell}, h_{\ell i} \in \mathbb{C}$ they commute. We can then relabel our indices $i \mapsto \ell$ $\ell \mapsto k$ $k \mapsto i$ to get

$$\operatorname{tr} U \operatorname{diag}\{f'(\lambda_1), \dots, f'(\lambda_n)\} U^* H = \sum_k \sum_\ell \sum_i f'(\lambda_i)\overline{u}_{ik} h_{k\ell} u_{\ell i},$$

So, for every direction H, we have that $\operatorname{tr}(U \operatorname{diag}\{f'(\lambda_1), \dots, f'(\lambda_n)\} U^* H) = \operatorname{tr}\left(f^{[1]}(\Lambda) \circ (U^* H U)\right)$. overfull hbox :eyeroll: By picking the "correct" H[3], we conclude that our unique quantity $g(X)$ is $U \operatorname{diag}\{f'(\lambda_1), \dots, f'(\lambda_n)\} U^*$. But, recall that $X = U\Lambda U$ so, in the functional calculus, $g(X) = f'(X)$. This recovers theorem 3.3 of [5] as we have constructed a g such that

$$\operatorname{tr} Df(X)[H] = \operatorname{tr} H g(X)$$

■

Connect this to the gradient and this section is done

3 See example EXAMPLE NUMBER for details

Part II

THE SHOWCASE

You can put some informational part preamble text here. Illo principalmente su nos. Non message *occidental* anglo-romanic da. Debitas effortio simplificate sia se, auxiliar summarios da que, se avantiate publicationes via. Pan in terra summarios, capital interlingua se que. Al via multo esser specimen, campo responder que da. Le usate medical addresses pro, europa origine sanctificate nos se.

Part III

APPENDIX

[1] Rajendra Bhatia. *Matrix Analysis*. Vol. 169. Graduate Texts in Mathematics. New York, NY: Springer New York, 1997. ISBN: 978-1-4612-6857-4 978-1-4612-0653-8. DOI: 10.1007/978-1-4612-0653-8.

[2] J. William Helton, Igor Klep, and Scott McCullough. "Free Convex Algebraic Geometry." In: (Apr. 2013). DOI: 10.48550/arXiv.1304.4272.

[3] Donald E. Knuth. "Computer Programming as an Art." In: *Communications of the ACM* 17.12 (1974), pp. 667–673.

[4] J. E. Pascoe. "Free Noncommutative Principal Divisors and Commutativity of the Tracial Fundamental Group." In: (Oct. 2020).

[5] J. E. Pascoe. "Trace Minmax Functions and the Radical Laguerre-P\'olya Class." In: (Aug. 2020). DOI: 10.48550/arXiv.2008.05469.