p-Adic Stability In Linear Algebra

Xavier Caruso
Université Rennes 1
xavier.caruso@normalesup.org

David Roe
University of British Columbia
roed.math@gmail.com

Tristan Vaccon
Université Rennes 1
tristan.vaccon@univ-rennes1.fr

ABSTRACT

Using the differential precision methods developed previously by the same authors, we study the p-adic stability of standard operations on matrices and vector spaces. We demonstrate that lattice-based methods surpass naive methods in many applications, such as matrix multiplication and sums and intersections of subspaces. We also analyze determinants, characteristic polynomials and LU factorization using these differential methods. We supplement our observations with numerical experiments.

Categories and Subject Descriptors

I.1.2 [**Computing Methodologies**]: Symbolic and Algebraic Manipulation – *Algebraic Algorithms*

General Terms

Algorithms, Theory

Keywords

p-adic precision; linear algebra; ultrametric analysis

1. INTRODUCTION

For about twenty years, the use of p-adic methods in symbolic computation has been gaining popularity. Such methods were used to compute composed products of polynomials [2], to produce hyperelliptic curves of genus 2 with complex multiplication [4], to compute isogenies between elliptic curves [8] and to count points on varieties using p-adic cohomology theories (*cf.* [5, 7] and many followers). However, a general framework allowing a precise study of p-adic precision — the main issue encountered when computing with p-adic numbers — was designed only recently in [3].

In [3], we advocate the use of lattices to track the precision of vectors and points on p-adic manifolds. The main result of *loc. cit.*, recalled in Proposition 2.1, allows for the propagation of precision using differentials. While representing precision by lattices carries a high cost in space and time requirements, the reduced precision loss can sometimes overwhelm these costs within a larger algorithm. In this paper, we apply the ideas of [3] to certain linear algebraic tasks.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
ISSAC'15, July 6–9, 2015, Bath, United Kingdom.
Copyright © 2015 ACM 978-1-4503-3435-8/15/07 ...$15.00
DOI: http://dx.doi.org/10.1145/2755996.2756655.

Main Results. We give a number of contexts where lattice-based precision methods outperform the standard coordinate-wise methods. The two most striking examples are an analysis of matrix multiplication and of sums and intersections of subspaces. In Proposition 3.1, we give a formula for the lattice-precision of the product of two matrices. In Figure 2, we describe the precision loss in multiplying n matrices, which appears linear in n when using standard methods and logarithmic in n when using lattice methods. We also give an example in Figure 4 where lattice methods actually yield an *increase* in precision as the computation proceeds.

Organization of the paper. Section 2.1 recalls the theory of precision developed in [3] and defines a notion of diffuse precision for comparing to coordinate-wise methods. In particular, we recall Proposition 2.1, which allows the computation of precision using differentials. Proposition 2.5 in Section 2.2 is a technical result that will be used to determine the applicability of Proposition 2.1.

In Section 3.1, we analyze matrix multiplication and report on experiments that demonstrate the utility of lattice-based precision tracking. Section 3.2 finds the precision of the determinant of a matrix, and Section 3.3 applies the resulting formula to characteristic polynomials. We define the precision polygon of a matrix, which gives a lower bound on the precision of the characteristic polynomial. This polygon lies above the Hodge polygon of the matrix; we give statistics on the difference and the amount of diffuse precision. Finally, we apply Proposition 2.1 to LU factorization and describe experiments with lattice-based precision analysis. Section 4.1 reviews the geometry of Grassmannians, which we apply in Section 4.2 to differentiating the direct image, inverse image, sum and intersection maps between Grassmannians. We then report in Section 4.3 on tracking the precision of subspace arithmetic in practice. In the appendix, we give a proof of Proposition 2.4.

The code used to make experiments presented in this paper is available at `https://github.com/CETHop/padicprec`.

Notation. Throughout the paper, K will refer to a complete discrete valuation field. Usual examples are finite extensions of $\mathbb{Q}_p$ and Laurent series fields over a field. We denote by $\mathrm{val} : K \twoheadrightarrow \mathbb{Z} \cup \{+\infty\}$ the valuation on K, by $\mathcal{O}_K$ the ring of integers and by $\pi \in K$ an element of valuation 1. We let $\|\cdot\|$ be the norm associated to val.

2. THE THEORY OF P-ADIC PRECISION

The aim of this section is to briefly summarize the content of [3] and fill in certain details.

2.1 Lattices as precision data

In [3], we suggest the use of lattices to represent the precision of elements in K-vector spaces. We shall contrast this approach with the *coordinate-wise method* used in Sage, where the precision of an element is specified by giving the precision of each coordinate separately and is updated after each basic operation.

Consider a finite dimensional[1] normed vector space E defined over K. We use the notation $\|\cdot\|_E$ for the norm on E and $B_E^-(r)$ (resp. $B_E(r)$) for the open (resp. closed) ball of radius r centered at the origin. A *lattice* $L \subset E$ is a sub-$\mathcal{O}_K$-module which generates E over K. Since we are working in a ultrametric world, the balls $B_E(r)$ and $B_E^-(r)$ are examples of lattices. Actually, lattices should be thought of as special neighborhoods of 0 and therefore are good candidates to model precision data. Moreover, as revealed in [3], they behave quite well under (strictly) differentiable maps:

Proposition 2.1. *Let E and F be two finite dimensional normed vector spaces over K and $f : U \to F$ be a function defined on an open subset U of E. We assume that f is differentiable at some point $v_0 \in U$ and that the differential $f'(v_0)$ is surjective. Then, for all $\rho \in (0,1]$, there exists a positive real number δ such that, for all $r \in (0,\delta)$, any lattice H such that $B_E^-(\rho r) \subset H \subset B_E(r)$ satisfies:*

$$f(v_0 + H) = f(v_0) + f'(v_0)(H). \tag{1}$$

This proposition enables the *lattice method* of tracking precision, where the precision of the input is specified as a lattice H and precision is tracked via differentials of the steps within a given algorithm. The equality sign in Eq. (1) shows that this method yields the optimum possible precision. We refer to [3, §4.1] for a more complete exposition.

In [3], we also explained that if f is locally analytic, then the constant δ appearing in Proposition 2.1 can be expressed in terms of the growing function $\Lambda(f)$ defined by

$$\Lambda(f)(v) = \log\left(\sup_{h \in B_E^-(e^v)} \|f(h)\|\right)$$

with the convention that $\Lambda(f)(v) = +\infty$ if f does not converge on $B_E^-(e^v)$. We refer to [3, Proposition 3.12] for the precise statement. We state here the case of integral polynomial functions. A function $f : E \to F$ is said to be *integral polynomial* if it is given by multivariate polynomial functions with coefficients in $\mathcal{O}_K$ in any (equivalently all) system of coordinates associated to a $\mathcal{O}_K$-basis of $B_E(1)$.

Proposition 2.2. *We keep the notation of Proposition 2.1 and assume in addition that f is integral polynomial. Let C be a positive real number such that $B_F(1) \subset f'(v_0)(B_E(C))$. Then Proposition 2.1 holds with $\delta = C \cdot \rho^{-1}$.*

In [3, Appendix A], the theory is extended to manifolds over K, where the precision datum at some point x is a lattice in the tangent space at x. Propositions 2.1 and 2.2 have analogues obtained by working in charts. We use this extension to compute with vector spaces in §4.

[1]The framework of [3] is actually those of Banach spaces. However, we will not need infinite dimensional spaces here.

We will use the following definition in contrasting lattice and coordinate-wise methods. Suppose E is equipped with a basis $(e_1, \dots, e_n)$ and write $\pi_i : E \to Ke_i$ for the projections.

Definition 2.3. Let $H \subset E$ be a lattice. The number of *diffused digits of precision* of H is the length of H_0/H where $H_0 = \pi_1(H) \oplus \cdots \oplus \pi_n(H)$.

If H represents the actual precision of some object, then H_0 is the smallest diagonal lattice containing H. Since coordinate-wise methods cannot yield a precision better than H_0, k provides a lower bound on the number of p-adic digits gained by lattice methods over standard methods.

2.2 A bound on a growing function

In the next sections, we will compute the derivative of several standard operations and sometimes give a simple expression in term of the input and the output. In other words, the function f modeling such an operation satisfies a differential equation of the form $f' = g \circ (f, \mathrm{id})$ where g is a given — and hopefully rather simple — function. The aim of this subsection is to study this differential equation and to derive from it certain bounds on the growing function $\Lambda(f)$. *We will assume that K has characteristic* 0.

Let E, F and G be finite-dimensional normed vector spaces with $U \subseteq E$ and $V \subseteq F$ and $W \subset G$ open subsets. Generalizing the setting above, we consider the differential equation:

$$f' = g \circ (f, h). \tag{2}$$

Here $g : V \times W \to \mathrm{Hom}(E, F)$ and $h : U \to W$ are known locally analytic functions and $f : U \to V$ is the unknown locally analytic function. In what follows, we always assume that V and W contain the origin, $f(0) = 0$, $h(0) = 0$ and $g(0) \neq 0$. These assumptions are harmless for two reasons: first, we can always shift f and h (and g accordingly) so that they both vanish at 0, and second, in order to apply Proposition 2.2 the derivative $f'(0)$ needs to be surjective and therefore *a fortiori* nonzero.

We assume that we are given in addition two nondecreasing convex functions Λ_g and Λ_h such that $\Lambda(g) \leq \Lambda_g$ and $\Lambda(h) \leq \Lambda_h$. We suppose further that there exists ν such that Λ_g is constant on the interval $(-\infty, \nu]$[2]. We introduce the functions τ_ν and Λ_f defined by:

$$\tau_\nu(x) = \begin{cases} x & \text{if } x \leq \nu, \\ +\infty & \text{otherwise;} \end{cases}$$

$$\text{and } \Lambda_f(x) = \tau_\nu \circ (\mathrm{id} + \Lambda_g \circ \Lambda_h)(x + \alpha),$$

where α is a real number satisfying $\|n!\| \geq e^{-\alpha n}$ for all n. If p is the characteristic of the residue field, a suitable value for α is $\alpha = -\frac{p}{p-1} \cdot \log \|p\|$ if $p > 0$ and $\alpha = 0$ if $p = 0$. The next Proposition is proved in Appendix A.

Proposition 2.4. *We have $\Lambda(f) \leq \Lambda_f$.*

Figure 1 illustrates Proposition 2.4. The blue plain line represents the graph of the function Λ_f. A quick computation shows that, on a neighborhood of $-\infty$, this function is given by $\Lambda_f(x) = x + \alpha + \mu$ where μ is the value that Λ_g takes on the interval $(-\infty, \nu]$. Proposition 2.4 says that the graph of $\Lambda(f)$ lies below the plain blue line. Moreover, we remark that the Taylor expansion of $f(z)$ starts

[2]We note that this assumption is fulfilled if we take $\Lambda_g = \Lambda(g)$ because we have assumed that $g(0)$ does not vanish.

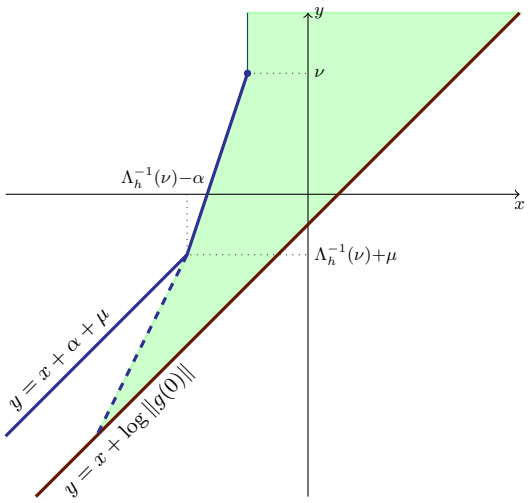

Figure 1: Admissible region for the graph of $\Lambda(f)$

with the term $g(0)z$. Hence, on a neighborhood of $-\infty$, we have $\Lambda(f)(x) = x + \log\|g(0)\|$. Using convexity, we get: $\Lambda(f)(x) \geq x + \log\|g(0)\|$ for all $x \in \mathbb{R}$. In other words, the graph of $\Lambda(f)$ lies above the brown line. Furthermore, we know that the slopes of $\Lambda(f)$ are all integral because f is locally analytic. Hence, $\Lambda(f)$ cannot lie above the dashed blue line defined as the line of slope 2 passing through the first break point of the blue plain line, which has coordinates $(y_0 - \alpha - \mu, y_0)$ with $y_0 = \min(\Lambda_h^{-1}(\nu) + \mu, \nu)$. As a conclusion, we have proved that the graph of $\Lambda(f)$ must coincide with the brown line until it meets the dashed blue line and then has to stay in the green area.

As a consequence, we derive the following proposition, which can be combined with Proposition 3.12 of [3]. Following [3], if φ is a convex function and $v \in \mathbb{R}$, we define

$$\varphi_{\geq v} : x \mapsto \inf_{y \geq 0} \big(\varphi(x+y) - vy\big).$$

It is the highest convex function with $\varphi_{\geq v} \leq \varphi$ and $\varphi'_{\geq v} \geq v$.

Proposition 2.5. *Keeping the above notation, we have:*

$$\Lambda(f)_{\geq 2}(x) \leq 2(x + \alpha + \mu) - \min(\Lambda_h^{-1}(\nu) + \mu,\ \nu)$$

for all $x \leq \min(\Lambda_h^{-1}(\nu) - \alpha,\ \nu - \mu - \alpha)$.

Proof. The inequality follows from the fact that $y = 2(x+\alpha+\mu) - y_0$ is the equation of the dashed blue line. □

Remark 2.6. In certain situations, it may happen that the function f is solution of a simpler differential equation of the form $f' = g \circ f$. If this holds, Proposition 2.5 gives the bound $\Lambda(f)_{\geq 2}(x) \leq 2(x + \alpha + \mu) - \nu$ for $x \leq \nu - \mu - \alpha$.

Beyond this particular case, we recommend choosing the function h by endowing F with the second norm $\|x\|'_F = e^{\mu} \cdot \|x\|_F$ $(x \in F)$ and taking $h : (F, \|\cdot\|_F) \to (F, \|\cdot\|'_F)$ to be the identity on the underlying vector spaces. The function $\Lambda(h) : \mathbb{R} \to \mathbb{R}$ then maps x to $x + \mu$ and we can choose $\Lambda_h = \Lambda(h)$.

3. MATRICES

Let $M_{m,n}(K)$ denote the space of $m \times n$ matrices over K. We will repeatedly use the Smith decomposition for $M \in M_{m,n}(K)$, which is $M = U_M \cdot \Delta_M \cdot V_M$ with U_M and V_M unimodular and Δ_M diagonal. Write $\sigma_i(M)$ for the valuation of the (i,i)-th entry of Δ_M, and by convention set $\sigma_i(M) = +\infty$ if $i > \min(m,n)$. Order the $\sigma_i(M)$ so that $\sigma_i(M) \leq \sigma_{i+1}(M)$.

3.1 Multiplication

To begin with, we want to study the behavior of the precision when performing a matrix multiplication. Let r, s and t be three positive integers and assume that we want to multiply a matrix $A \in M_{r,s}(K)$ by a matrix $B \in M_{s,t}(K)$. This operation is of course modeled by the integral polynomial function:

$$\begin{array}{rccc} \mathcal{P}_{r,s,t} : & M_{r,s}(K) \times M_{s,t}(K) & \to & M_{r,t}(K) \\ & (A, B) & \mapsto & AB. \end{array}$$

According to Proposition 2.1, the behavior of the precision when computing AB is governed by $\mathcal{P}'_{r,s,t}(A,B)$, the linear mapping that takes a pair (dA, dB) to $A \cdot dB + dA \cdot B$.

To fix ideas, let us assume from now that the entries of A and B all lie in $\mathcal{O}_K$ and are known at the same precision $O(\pi^N)$. In order to apply Propositions 2.1 and 2.2, we then need to compute the image of the standard lattice $\mathcal{L}_0 = M_{r,s}(\mathcal{O}_K) \times M_{s,t}(\mathcal{O}_K)$ under $\mathcal{P}'_{r,s,t}(A,B)$. It is of course contained in $M_{r,t}(\mathcal{O}_K)$; this reflects the obvious fact that each entry of the product AB is also known with precision at least $O(\pi^N)$. Nonetheless, it may happen that the above inclusion is strict, meaning that we are *gaining* precision in those cases.

Set $a_i = \sigma_i(A)$ and $b_i = \sigma_i(B)$, and define $M_{r,t}((a_i),(b_j))$ as the sublattice of $M_{r,t}(\mathcal{O}_K)$ consisting of matrices $M = (M_{i,j})$ such that $\mathrm{val}(M_{i,j}) \geq \min(a_i, b_j)$ for all (i,j).

Proposition 3.1. *With the above notation, we have*

$$\mathcal{P}'_{r,s,t}(A,B)(\mathcal{L}_0) = U_A \cdot M_{r,t}((a_i),(b_j)) \cdot V_B$$

$$\textit{and} \quad \mathrm{length}\Big(\tfrac{M_{r,t}(\mathcal{O}_K)}{\mathcal{P}'_{r,s,t}(A,B)(\mathcal{L}_0)}\Big) = \sum_{i,j} \min(a_i, b_j)$$

Proof. We write $A \cdot dB + dA \cdot B = U_A \cdot M \cdot V_B$ with

$$M = \Delta_A \cdot V_A \cdot dA \cdot V_B^{-1} + U_A^{-1} \cdot dB \cdot U_B \cdot \Delta_B.$$

When dA varies in $M_{a,b}(\mathcal{O}_K)$ so does $V_A \cdot dA \cdot V_B^{-1}$ and therefore the first summand in M varies in the subspace of $M_{r,t}(\mathcal{O}_K)$ consisting of matrices whose entries on the i-th row have valuation at least a_i. Arguing similarly for the second summand, we deduce the first statement of the Proposition. The second statement is now clear. □

From the perspective of precision, the second statement of Proposition 3.1 means that the computation of AB gains $\sum_{i,j} \min(a_i, b_j)$ significant digits in absolute precision[3] as soon as $N > \min(a_r, b_t)$ (*cf.* Proposition 2.2). However, many of these digits are diffused in the sense of Definition 2.3. To change bases in order to make this increased precision visible with coordinates, write $AB = U_A \cdot P \cdot V_B$ with $P = \Delta_A \cdot V_A \cdot U_B \cdot \Delta_B$. Tracking precision in the usual way, the (i,j)-th entry of P is known at precision $O(\pi^{N+\min(a_i,b_j)})$. Multiplication by U_A and V_B then diffuses the precision across the entries of AB.

[3]We note that, on the other side, the valuation of the entries of AB may increase, meaning that we are also losing some significant digits if we are reasoning in relative precision.

d	n	Average loss of precision	
		Coord-wise method	Lattice method
2	10	2.8	2.4
2	100	16.7	5.0
2	1000	157.8	7.9
3	10	2.2	1.9
3	100	12.8	4.0
3	1000	122.5	7.0

Results for a sample of 1000 random inputs in $M_{d,d}(\mathbb{Z}_2)^n$

Figure 2: Average loss of precision in Algorithm 1

We now consider the impact of tracking this diffuse precision. Although the benefit is not substantial for a single product of matrices, it accumulates as we multiply a large number of matrices. We illustrate this phenomenon with the following simple example.

Algorithm 1: `example_product`
Input: a list $(M_1, \dots, M_n)$ of square matrices of size d.

1. Set P to the identity matrix of size d
2. **for** $j = 1, \dots, n$ **do compute** $P = PM_i$
3. **return** the top left entry of P

Figure 2 compares the number of significant digits in *relative* precision we are losing on the output of Algorithm 1 when we are using, on the one hand, a standard coordinate-wise track of precision and, on the other hand, a lattice-based method to handle precision. We observe that, in the first case, the number of lost digits seems to grow linearly with respect to the number of multiplications we are performing (that is n) whereas, in the second case, the growth seems to be only logarithmic. It would be nice to have a precise formulation (and proof) of this heuristic.

Note that multiplication of many random matrices plays a central role in the theory of random walks on homogeneous spaces [1]. Better stability in computing such products helps estimate Lyapunov exponents in that context.

3.2 Determinant

The computation of the differential of $\det : M_{n,n}(K) \to K$ is classical: at a point M it is the linear map

$$\det{}'(M) : dM \mapsto \mathrm{Tr}(\mathrm{Com}(M) \cdot dM),$$

where $\mathrm{Com}(M)$ stands for the comatrix of M, which is $\det(M)M^{-1}$ when M is invertible. If $\mathrm{rank}(M) < n-1$, then $\det{}'(M)$ is not surjective and we cannot apply Proposition 2.1. Therefore, we suppose that $\mathrm{rank}(M) \geq n-1$ for the rest of this section.

As with matrix multiplication, we first determine the image of the standard lattice $\mathcal{L}_0 = M_{n,n}(\mathcal{O}_K)$ under $\det{}'(M)$.

Proposition 3.2. *Setting $v = \sigma_1(M) + \cdots + \sigma_{n-1}(M)$, we have $\det{}'(M)(\mathcal{L}_0) = \pi^v \mathcal{O}_K$.*

Proof. From the description of $\det{}'(M)$, we see that it is enough to prove that the smallest valuation of an entry of $\mathrm{Com}(M)$ is v or, equivalently, that the ideal of $\mathcal{O}_K$ generated by all minors of M of size $(n-1)$ is $\pi^v \mathcal{O}_K$. But this ideal remains unchanged when we multiply M on the left or on the right by a unimodular matrix. Thus we may assume that $M = \Delta_M$, and the result becomes clear. □

In terms of precision, Proposition 3.2 implies that if M is given at flat precision $O(\pi^N)$ with $N > v$, then $\det(M)$ is known at precision $O(\pi^{N+v})$. Thus we are gaining v digits in absolute precision or, equivalently, losing $\sigma_n(M)$ digits of relative precision. Furthermore, one may compute $\det(M)$ with this optimal precision by finding an approximate Smith decomposition with Δ_M known at precision $O(\pi^N)$ and multiplying its diagonal entries.

3.3 Characteristic polynomials

Write $\mathrm{char} : M_{n,n}(K) \to K[X]$ for the characteristic polynomial, and $K_{<n}[X] \subset K[X]$ for the subspace consisting of polynomials of degree less than n. Then the differential of char at a point M is

$$\mathrm{char}'(M) : dM \mapsto \mathrm{Tr}(\mathrm{Com}(X - M) \cdot dM).$$

The image is the K-span of the entries of $\mathrm{Com}(X-M)$, which is clearly contained within $K_{<n}[X]$. In fact, the image will equal $K_{<n}[X]$ as long as M does not have two Jordan blocks with the same generalized eigenvalue. For now on, we make this assumption.

Recall that the *Newton polygon* $\mathrm{NP}(f)$ of a polynomial $f(X) = \sum_i a_i X^i$ is the lower convex hull of the points $(i, \mathrm{val}(a_i))$ and the *Newton polygon* $\mathrm{NP}(M)$ of a matrix M is $\mathrm{NP}(\mathrm{char}(M))$. The *Hodge polygon* $\mathrm{HP}(M)$ of M is the lower convex hull of the points $(i, \sum_{j=1}^{n-i} \sigma_j(M))$. For any matrix M, the polygon $\mathrm{NP}(M)$ lies above $\mathrm{HP}(M)$ [6, Thm. 4.3.11].

Such polygons arise naturally in tracking the precision of polynomials [3, §4.2]. Any such polygon P yields a lattice $\mathcal{L}_P$ in $K_{<n}[X]$ consisting of polynomials whose Newton polygons lie above P. This lattice is generated by monomials $a_i X^i$, where $\mathrm{val}(a_i)$ is the ceiling of the height of P at i. These polygons are used in a coordinate-wise precision tracking for polynomial arithmetic. We now introduce another polygon, bounded between $\mathrm{NP}(M)$ and $\mathrm{HP}(M)$, that will provide an estimate on the precision of $\mathrm{char}(M)$.

Definition 3.3. The *precision polygon* $\mathrm{PP}(M)$ of M is the lower convex hull of the Newton polygon of the entries of $\mathrm{Com}(X-M)$.

It is clear from the definition $\mathcal{L}_{\mathrm{PP}(M)} \subset \mathrm{char}'(M)(\mathcal{L}_0)$ where $\mathcal{L}_0$ is the standard lattice $M_{n,n}(\mathcal{O}_K)$. More precisely, $\mathrm{PP}(M)$ is the smallest polygon P for which the inclusion $\mathcal{L}_P \subset \mathrm{char}'(M)(\mathcal{L}_0)$ holds. By Proposition 2.1, the precision polygon determines the minimal precision losses possible when encoding polynomial precision using polygons.

It turns out that the precision polygon is related to the Hodge and Newton polygons. If a polygon P has vertices (x_i, y_i), we let $T_n(P)$ be the translated polygon with vertices $(x_i - n, y_i)$.

Proposition 3.4. *The precision polygon $\mathrm{PP}(M)$ lies between $T_1(\mathrm{HP}(M))$ and $\mathrm{NP}(M)$.*
Moreover, $\mathrm{PP}(M)$ and $T_1(\mathrm{HP}(M))$ meet at 0 and $n-1$.

Proof. The coefficients of $\mathrm{char}(M)$ can be expressed as traces of exterior powers: the coefficient of X^i is $\mathrm{Tr}(\Lambda^i(M))$, which is $\mathrm{Tr}(\Lambda^i(U_M)\Lambda^i(\Delta_M)\Lambda^i(V_M))$. Computing $\Lambda^i(\Delta_M)$, we get the first statement of the Proposition. For $i = 1$, we further find that $\mathrm{PP}(M)$ vanishes at the abscissa $n-1$. By definition so does $T_1(\mathrm{HP}(M))$. The fact that $\mathrm{PP}(M)$ and $T_1(\mathrm{HP}(M))$ meet at abscissa 0 follows from Proposition 3.2.

It remains to prove the comparison with the Newton polygon. Set $f = \mathrm{char}(M)$, set $m_{i,j}$ as the (i,j)-th entry of

M, $f_{i,j}$ as the (i,j)-th entry of $\mathrm{Com}(X-M)$ and $\mu_{i,j} = \mathrm{val}(m_{i,j})$. We write $f[k]$ for valuation of the coefficient of X^k in f, and set $f[-1] = +\infty$. The equation $(X-M)\cdot \mathrm{Com}(X-M) = f\cdot I$ yields, for all i and k,

$$\begin{aligned} f[k] &\geq \inf(f_{i,i}[k-1], \mu_{i,0} + f_{0,i}[k], \dots, \mu_{i,n} + f_{n,i}[k]) \\ &\geq \inf(f_{i,i}[k-1], f_{j,i}[k]), \end{aligned}$$

with the infimum over j. Taking lower convex hulls and noting that $\mathrm{PP}(M)$ is nonincreasing, which follows from the comparison with the Hodge polygon, we get the result. □

Remark 3.5. Experiments actually support the following stronger result: $\mathrm{PP}(M)$ is bounded above by $T_1(\mathrm{NP}(M))$.

For many matrices, $\mathrm{PP}(M) = T_1(\mathrm{HP}(M))$. For random matrices over $\mathbb{Z}_2$, the 2-adic precision polygon is equal to the Hodge polygon in 99.5% of cases in dimension 4, down to 99.1% in dimension 8. Over $\mathbb{Z}_3$, the fraction rises to 99.98%, with no clear dependence on dimension. Empirically, $\mathrm{PP}(M)$ seems most likely to differ from $T_1(\mathrm{HP}(M))$ at 1, corresponding to the precision of the linear term of the characteristic polynomial.

Of course, the precision lattice $\mathcal{E} = \mathrm{char}'(M)(\mathcal{L}_0)$ may contain diffuse precision which is not encapsulated in $\mathrm{PP}(M)$. Diffuse precision arises in 11% of cases in dimension 3, up to 15% of cases in dimension 8. This percentage increases as $\mathrm{val}(\det(M))$ increases, reaching 34% in dimension 9 for matrices constrained to have determinant with 2-adic valuation 12. Moreover, one can construct examples with arbitrarily large amounts of diffuse precision. Suppose the $\sigma_i(M)$ are large. Proposition 3.4 implies that $\mathcal{E}$ is contained within $\mathcal{O}_{K,<n}[X]$ with index at least $\sum_{i=1}^{n-1}\sigma_i(M)$. The precision lattice of $1+M$ is obtained from $\mathcal{E}$ via the transformation $X \mapsto 1+X$, but $\mathrm{PP}(1+M)$ is now flat with height 0.

For randomly chosen matrices, approximating $\mathcal{E}$ using the Hodge polygon loses only a small amount of precision. However, if the $\sigma_i(M)$'s are large or if M is a translate of such a matrix, using lattice precision can be very useful.

3.4 LU factorization

In this section, we denote by $\|\cdot\|$ the subordinate matrix norm on $M_n(K)$ and, given a positive real number C, we let $B(C)$ be the closed ball in $M_n(K)$ centered at the origin of radius C. We consider the following subsets of $M_n(K)$:

- O_n is the open subset of matrices whose principal minors do not vanish (we recall that the latest condition implies the existence and the uniqueness of a LU factorization);
- U_n is the sub-vector space of upper-triangular matrices;
- L_n^0 (resp. L_n^u) is the sub-affine space of nilpotent (resp. unipotent) lower-triangular matrices.

Calculus and precision. We choose to normalize the LU factorization by requiring that L is unipotent and denote by $D : O_n \to L_n^u \times U_n$ the function modeling this decomposition. The computation of the differential of D has already been done in [3, Appendix B]. For $M \in \mathcal{O}_n$ with $D(M) = (L,U)$, the linear mapping $D'(M)$ is given by:

$$dM \mapsto (L\cdot \mathrm{low}(dX), \mathrm{up}(dX)\cdot U) \text{ with } dX = L^{-1}\cdot dM\cdot U^{-1}$$

where low (resp. up) denotes the canonical projection of $M_n(K)$ onto L_n^0 (resp. U_n). It is easily checked that $D'(M)$ is bijective with inverse given by $(A,B)\mapsto AU+LB$.

We now want to apply Proposition 2.5 in order to derive a concrete result on precision. *We then assume that K has characteristic 0.* We pick $M_0 \in O_n$ and write $D(M_0) = (L_0, U_0)$. We consider the translated function f taking M to $D(M_0+M) - D(M_0)$. We then have $f(0) = 0$ and $f'(M) = D'(M_0+M)$. Using the explicit description of the inverse of $D'(M_0+M)$, we find $B(1) \subset f'(0)\cdot B(C)$ for $C = \max(\|U_0\|, \|L_0\|)$. Moreover, f satisfies the differential equation $f' = g\circ f$ where g is defined by:

$$\begin{aligned} g(A,B)(X) = \big((L_0+A)\cdot \mathrm{low}(Y), \mathrm{up}(Y)\cdot(U_0+B)\big) \\ \text{with } Y = (L_0+A)^{-1}\cdot X\cdot (U_0+B)^{-1}. \end{aligned} \tag{3}$$

Let $\kappa(S) = \|S\|\cdot\|S^{-1}\|$ denote the condition number of a matrix S. Remarking that $\|S+T\| = \|S\|$ if $\|T\| < \|S\|$ and $\|(S+T)^{-1}\| = \|S^{-1}\|$ if $\|T\| < \|S^{-1}\|^{-1}$, we deduce from (3) that $\|g(A,B)\| \leq \max\left(\kappa(L_0)\|U_0^{-1}\|, \kappa(U_0)\|L_0^{-1}\|\right)$ provided that $\|A\| < \|L_0^{-1}\|^{-1}$ and $\|B\| < \|U_0^{-1}\|^{-1}$. Combining this with Proposition 2.5 and [3, Proposition. 3.12], we finally find that Eq. (1) holds as soon as

$$\frac{\rho}{r} > \|p\|^{-\frac{2p}{p-1}}\cdot \max(\|L_0\|,\|U_0\|)\cdot \max(\|L_0^{-1}\|,\|U_0^{-1}\|)\cdot \max\left(\kappa(L_0)\|U_0^{-1}\|, \kappa(U_0)\|L_0^{-1}\|\right)^2.$$

Numerical experiments. Let $B_n = (E_{i,j})_{1\leq i,j\leq n}$ be the canonical basis of $M_n(K)$. It can be naturally seen as a basis of $L_n^0 \times U_n$ as well. For a given $M \in O_n$, let us abuse notation and write $D'(M)$ for the matrix of this linear mapping in the above basis. We remark that $D'(M)$ is lower-triangular in this basis. Projecting $D'(M)$ onto each coordinate, we find the best coordinate-wise loss of precision we can hope for the computation of D is given by $\sum_u (\max_v (\mathrm{val}(D'(M)_{u,v})))$. This number should be compared to $\mathrm{val}(\det(D'(M)))$, which is precision lost in the lattice method. The number of diffused digits of precision of $D'(M)(M_{n,n}(\mathcal{O}_K))$ is then the difference between these two numbers. Figure 3 summarizes the mean and standard deviation of those losses for a sample of 2000 random matrices in $M_{d,d}(\mathbb{Z}_2)$.

matrix size	Loss of precision in LU decomposition			
	coord-wise method		lattice method	
	mean	deviation	mean	deviation
2	3.0	5	1.5	1.4
3	9.4	11	2.3	2.3
4	20	20	3.8	3.1

Results for a sample of 2000 instances

Figure 3: Loss of precision for LU factorization

4. VECTOR SPACES

Vector spaces are generally represented as subspaces of K^n for some n and hence naturally appear as points on Grassmannians. Therefore, one can use the framework of [3, Appendix A] to study p-adic precision in this context.

4.1 Geometry of Grassmannians

Given E, a finite dimensional vector space over K, and d, an integer in the range $[0, \dim E]$, we write $\mathrm{Grass}(E,d)$ for the Grassmannian of d-dimensional subspaces of E. It is well-known that $\mathrm{Grass}(E,d)$ has the natural structure of a K-manifold. The aim of this subsection is to recall standard facts about its geometry. In what follows, we set $n = \dim E$ and equip E with a distinguished basis $(e_1,\dots,e_n)$.

Description and tangent space. Let V denote a fixed subspace of E of dimension d. The Grassmannian $\mathrm{Grass}(E,d)$ can be viewed as the quotient of the set of linear embeddings $f : V \hookrightarrow E$ modulo the action (by precomposition) of $\mathrm{GL}(V)$: the mapping f represents its image $f(V)$. It follows from this description that the tangent space of $\mathrm{Grass}(E,d)$ is canonically isomorphic to $\mathrm{Hom}(V,E)/\mathrm{End}(V)$, which is $\mathrm{Hom}(V,E/V)$.

Charts. Let V and V^{c} be two complementary subspaces of E (*i.e.* $V \oplus V^{\mathrm{c}} = E$). We assume that V has dimension d and denote by π the projection $E \to V$ corresponding to the above decomposition. We introduce the set $\mathcal{U}_{V,V^{\mathrm{c}}}$ of all embeddings $f : V \hookrightarrow E$ such that $\pi \circ f = \mathrm{id}_V$. Clearly, it is an affine space over $\mathrm{Hom}(V,V^{\mathrm{c}})$. Furthermore, we can embed it into $\mathrm{Grass}(E,d)$ by taking f as above to its image. This way, $\mathcal{U}_{V,V^{\mathrm{c}}}$ appears as an open subset of $\mathrm{Grass}(E,d)$ consisting exactly of those subspaces W such that $W \cap V^{\mathrm{c}} = 0$. As a consequence, the tangent space at each such W becomes isomorphic to $\mathrm{Hom}(V,V^{\mathrm{c}})$. The identification $\mathrm{Hom}(V,V^{\mathrm{c}}) \to \mathrm{Hom}(W,E/W)$ is given by $du \mapsto (du \circ f^{-1}) \bmod W$ where $f : V \xrightarrow{\sim} W$ is the linear mapping defining W.

When the pair (V,V^{c}) varies, the open subsets $\mathcal{U}_{V,V^{\mathrm{c}}}$ cover the whole Grassmannian and define an atlas. When implementing vector spaces on a computer, we usually restrict ourselves to the subatlas consisting of all charts of the form $(V_I, V_{I^{\mathrm{c}}})$ where I runs over the family of subsets of $\{1,\dots,n\}$ of cardinality d and V_I is the subspace spanned by the e_i's with $i \in I$. A subspace $W \in E$ then belongs to at least one $\mathcal{U}_{V_I,V_{I^{\mathrm{c}}}}$ and, given a family of generators of W, we can determine such an I together with the corresponding embedding $f : V_I \hookrightarrow E$ by row reducing the matrix of generators of W.

A variant. Alternatively, one can describe $\mathrm{Grass}(E,d)$ as the set of linear surjective morphisms $f : E \to E/V$ modulo the action (by postcomposition) of $\mathrm{GL}(E/V)$. This identification presents the tangent space at a given point V as the quotient $\mathrm{Hom}(E,E/V)/\mathrm{End}(E/V) \simeq \mathrm{Hom}(V,E/V)$. Given a decomposition $E = V \oplus V^{\mathrm{c}}$ as above, we let $\mathcal{U}^{\star}_{V,V^{\mathrm{c}}}$ denote the set of surjective linear maps $f : E \to V^{\mathrm{c}}$ whose restriction to V^{c} is the identity. It is an affine space over $\mathrm{Hom}(V,V^{\mathrm{c}})$ which can be identified with an open subset of $\mathrm{Grass}(E,d)$ *via* the map $f \mapsto \ker f$.

It is easily seen that $\mathcal{U}_{V,V^{\mathrm{c}}}$ and $\mathcal{U}^{\star}_{V,V^{\mathrm{c}}}$ define the same open subset in $\mathrm{Grass}(E,d)$. Indeed, given $f \in \mathcal{U}_{V,V^{\mathrm{c}}}$, one can write $f = \mathrm{id}_V + h$ with $h \in \mathrm{Hom}(V,V^{\mathrm{c}})$ and define the morphism $g = \mathrm{id}_E - h \circ \pi \in \mathcal{U}^{\star}_{V,V^{\mathrm{c}}}$. The association $f \mapsto g$ then defines a bijection $\mathcal{U}_{V,V^{\mathrm{c}}} \to \mathcal{U}^{\star}_{V,V^{\mathrm{c}}}$ which commutes with the embeddings into the Grassmannian.

Duality. If E is a finite dimensional vector space over K, we use the notation $E^{\star}$ for its dual (*i.e.* $E^{\star} = \mathrm{Hom}(E,K)$). If we are also given a subspace $V \subset E$, we denote by $V^{\perp}$ the subspace of $E^{\star}$ consisting of linear maps that vanish on V. We recall that the dual of $V^{\perp}$ (resp. $E^{\star}/V^{\perp}$) is canonically isomorphic to E/V (resp. V). For all d, the association $V \mapsto V^{\perp}$ defines a continuous morphism $\psi_E : \mathrm{Grass}(E,d) \to \mathrm{Grass}(E^{\star}, n-d)$. The action of ψ_E on tangent spaces is easily described. Indeed, the differential of ψ_E at V is nothing but the canonical identification between $\mathrm{Hom}(V,E/V)$ and $\mathrm{Hom}(V^{\perp}, E^{\star}/V^{\perp})$ induced by transposition. Furthermore, we observe that ψ_E respects the charts we have defined above, in the sense that it maps bijectively $\mathcal{U}_{V,V^{\mathrm{c}}}$ to $\mathcal{U}^{\star}_{V^{\perp},(V^{\mathrm{c}})^{\perp}} \simeq \mathcal{U}_{V^{\perp},(V^{\mathrm{c}})^{\perp}}$.

4.2 Differential computations

In this subsection, we compute the differential of various operations on vector spaces. For brevity, we skip the estimation of the corresponding growing functions (but this can be done using Proposition 2.5 as before if $\mathrm{char}(K) = 0$).

Direct images. Let E and F be two finite dimensional K-vector spaces of dimension n and m, respectively. Let d be an integer in $[0,n]$. We are interested in the direct image function DI defined on $\mathcal{M} = \mathrm{Hom}(E,F) \times \mathrm{Grass}(E,d)$ that takes the pair (f,V) to $f(V)$. Since the dimension of $f(V)$ may vary, the map DI does not take its values in a well-defined Grassmannian. We therefore stratify $\mathcal{M}$ as follows: for each integer $r \in [0,d]$, let $\mathcal{M}_r \subset \mathcal{M}$ be the subset of pairs (f,V) for which $f(V)$ has dimension r. The $\mathcal{M}_r$'s are locally closed in $\mathcal{M}$ and are therefore submanifolds. Moreover, DI induces differentiable functions $\mathrm{DI}_r : \mathcal{M}_r \to \mathrm{Grass}(F,r)$.

We would like to differentiate DI_r around some point $(f,V) \in \mathcal{M}_r$. To do so, we use the first description of the Grassmannians we gave above: we see points in $\mathrm{Grass}(E,d)$ (resp. $\mathrm{Grass}(F,d)$) as embeddings $V \hookrightarrow E$ (resp. $W \hookrightarrow F$) modulo the action of $\mathrm{GL}(V)$ (resp. $\mathrm{GL}(W)$). The point $V \in \mathrm{Grass}(E,d)$ is then represented by the canonical inclusion $v : V \to E$ whereas a representative w of W satisfies $w \circ \varphi = f \circ v$ where $\varphi : V \to W$ is the linear mapping induced by f. The previous relation still holds if (f,v) is replaced by a pair $(f',v') \in \mathcal{M}_r$ sufficiently close to (f,v). Differentiating it and passing to the quotient we find, first, that the tangent space of $\mathcal{M}_r$ at (f,v) consists of pairs $(df,dv) \in \mathrm{Hom}(E,F) \times \mathrm{Hom}(V,E/V)$ such that

$$d\tilde{w} = \big((df \circ v + f \circ dv) \bmod W\big) \in \mathrm{Hom}(V,F/W)$$

factors through φ (*i.e.* vanishes on $\ker \varphi = V \cap \ker f$) and, second, that the differential of DI_r at (f,V) is the linear mapping sending (df,dv) as above to the unique element $dw \in \mathrm{Hom}(W,F/W)$ such that $dw \circ \varphi = d\tilde{w}$.

Inverse images. We now consider the inverse image mapping II sending a pair $(f,W) \in \mathcal{W} = \mathrm{Hom}(E,F) \times \mathrm{Grass}(F,d)$ to $f^{-1}(W)$. As before, this map does not take values in a single Grassmannian, so we need to stratify $\mathcal{W}$ in order to get differentiable functions. For each integer $s \in [0,n]$, we introduce the submanifold $\mathcal{W}_s$ of $\mathcal{W}$ consisting of those pairs (f,W) such that $\dim f^{-1}(W) = s$. For all s, II induces a continuous function $\mathrm{II}_s : \mathcal{W}_s \to \mathrm{Grass}(E,s)$. Pick $(f,W) \in \mathcal{W}_s$. Set $V = f^{-1}(W)$ and denote by $w : F \to F/W$ the canonical projection. Similarly to what we have done for direct images, one can prove that the tangent space of $\mathcal{W}_s$ at some point $(f,W) \in \mathcal{W}_s$ is the subspace of $\mathrm{Hom}(E,F) \times \mathrm{Hom}(W,F/W)$ consisting of pairs (df,dw) such that $d\tilde{v} = (w \circ df + dw \circ f)_{|W}$ factors through the linear mapping $\varphi : E/V \to F/W$ induced by f. Furthermore II_s is differentiable at (f,W) and its differential is the linear mapping that takes (df,dw) as above to the unique element $dv \in \mathrm{Hom}(V,E/V)$ satisfying $\varphi \circ dv = d\tilde{v}$.

Direct images and inverse images are related by duality as follows: if $f : E \to F$ is any linear map and W is a subspace of F, then $f^{\star}(W^{\perp}) = f^{-1}(W)^{\perp}$. We can thus deduce the differentials of DI_s from those of II_s and *vice versa*.

Sums and intersections. Let d_1 and d_2 be two nonnegative integers. We consider the function Σ defined on the manifold $\mathcal{C} = \mathrm{Grass}(E,d_1) \times \mathrm{Grass}(E,d_2)$ by $\Sigma(V_1,V_2) = V_1 + V_2$. As before, in order to study Σ, we stratify $\mathcal{C}$ according to the dimension of the sum: for each integer $d \in [0, d_1 +$

$d_2]$, we define $\mathcal{C}_d$ as the submanifold of $\mathcal{C}$ consisting of those pairs (V_1, V_2) such that $\dim(V_1 + V_2) = d$. We get a well-defined mapping $\mathcal{C}_d \to \mathrm{Grass}(E, d)$ whose differential can be computed as before. The tangent space of $\mathcal{C}_d$ at a given point (V_1, V_2) consists of pairs $(dv_1, dv_2) \in \mathrm{Hom}(V_1, E/V_1) \times \mathrm{Hom}(V_2, E/V_2)$ such that $dv_1 \equiv dv_2 \pmod{V_1 + V_2}$ on the intersection $V_1 \cap V_2$, and the differential of Σ at (V_1, V_2) maps (dv_1, dv_2) to $dv \in \mathrm{Hom}(V, E/V)$ (with $V = V_1 + V_2$) defined by $dv(v_1 + v_2) = dv_1(v_1) + dv_2(v_2)$ $(v_1 \in V_1,\ v_2 \in V_2)$.

Using duality, we derive a similar result for the mapping $(V_1, V_2) \mapsto V_1 \cap V_2$ (left to the reader).

4.3 Implementation and experiments

Standard representation of vector spaces. One commonly represents subspaces of K^n using the charts $\mathcal{U}_{V_I, V_{I^c}}$ (where I is a subset of $\{1, \ldots, n\}$) introduced above. More concretely, a subspace $V \subset K^n$ is represented as the span of the rows of a matrix G_V having the following extra property: there exists some $I \subset \{1, \ldots, n\}$ such that the submatrix of G_V obtained by keeping only columns with index in I is the identity matrix. We recall that such a representation always exists and, when the set of indices I is fixed, at most one G_V satisfies the above condition. Given a family of generators of V, one can compute G_V and I as above by performing standard row reduction. Choosing the first non-vanishing pivot at every stage provides a canonical choice for I, but in the context of inexact base fields, choosing the pivot with the maximal norm yields a more stable algorithm.

The dual representation. Of course, one may alternatively use the charts $\mathcal{U}^\star_{V_I, V_{I^c}}$. Concretely, this means that we represent V as the left kernel of a matrix H_V having the following extra property: there exists some $I \subset \{1, \ldots, n\}$ such that the submatrix of H_V obtained by *deleting* rows with index in I is the identity matrix. As above, we can then compute I and H_V by performing column reduction.

Note that switching representations is cheap and stable. If $I = \{1, \ldots, d\}$ with $d = \dim V$ and I_d is the identity matrix of size d, the matrix G_V has the form $(I_d \quad G'_V)$. One can represent V with the same I and the matrix $H_V = \begin{pmatrix} -G'_V \\ I_{n-d} \end{pmatrix}$. A similar formula exists for general I.

Operations on vector spaces. The first representation we gave is well suited for the computation of direct images and sums. For instance, to compute $f(V)$ we apply f to each row of G_V, obtaining a family of generators of $f(V)$, and then row reduce. Dually, the second representation works well for computing inverse images, including kernels, and intersections. Since translating between the two dual representations is straightforward, we get algorithms for solving both problems using either representation.

Some experiments. Let us consider the example computation given in the following algorithm.

Algorithm 2: `example_vector_space`
Input: two integers n and N

1. Set $L_0 = \langle (1 + O(2^N), O(2^N), O(2^N)) \rangle \subset \mathbb{Q}_2^3$
2. **for** $i = 0, \ldots, n-1$
3. pick randomly $\alpha, \beta, \gamma, \delta \in M_{3,3}(\mathbb{Z}_2)$ with high precision
4. **compute** $L_{i+1} = \big(\alpha(L_i) + \beta(L_i)\big) \cap \big(\gamma(L_i) + \delta(L_i)\big)$
5. **return** L_n

---

n	Average loss of precision		
	Coord-wise method	Lattice method	
		Projected	Diffused
10	7.3	2.7	-2.4×2
20	14.8	5.5	-4.7×2
50	38.6	13.1	-12.0×2
100	78.1	26.5	-23.5×2

Results for a sample of 1000 executions (with $N \gg n$)

Figure 4: Average loss of precision in Algorithm 2

The expression *high precision* on line 3 means that the precision on α, β, γ and δ is set in such a way that it does not affect the resulting precision on L_{i+1}. Figure 4 shows the losses of precision when executing Algorithm 2 with various inputs n (the input N is always chosen sufficiently large so that it does not affect the behavior of precision). The *Coord-wise* column corresponds to the standard way of tracking precision. On the other hand, in the two last columns, the precision is tracked using lattices. The *Diffused* column gives the amount of diffused precision, factored to be comparible to the Coord-wise column. The fact that only negative numbers appear in this column means that we are actually always gaining precision with this model! Finally, the *Projected* column gives the precision loss after projecting the lattice precision onto coordinates.

APPENDIX

A. PROOF OF PROPOSITION 2.4

We prove Proposition 2.4 in the slightly more general context of K-Banach spaces.

A.1 Composite of locally analytic functions

Let U, V and W be three open subsets in K-Banach spaces E, F and G, respectively. We assume that $0 \in U$, $0 \in V$. Let $f : U \to V$ and $g : V \to W$ be two locally analytic functions around 0 with $f(0) = 0$. The composition $h = g \circ f$ is then locally analytic around 0 as well. Let $f = \sum_{n \geq 0} f_n$ $g = \sum_{n \geq 0} g_n$ and $h = \sum_{n \geq 0} h_n$ be the analytic expansions of f, g and h. Here f_n, g_n and h_n are the restrictions to the diagonal of some symmetric n-linear forms F_n, G_n and H_n, respectively. The aim of this subsection is to prove the following intermediate result.

Proposition A.1. *With the above notation, we have*

$$\|h_r\| \leq \sup_{m, (n_i)} \|g_m\| \cdot \|f_{n_1}\| \cdots \|f_{n_m}\|$$

for all nonnegative integers r, where the supremum is taken over all pairs $(m, (n_i))$ where m is a nonnegative integer and $(n_i)_{1 \leq i \leq m}$ is a sequence of length m of nonnegative integers such that $n_1 + \ldots + n_m = r$.

A computation gives the following expansion for $g \circ f$:

$$\sum \binom{m}{k_1 \ \cdots \ k_\ell} G_m(f_{n_1}, \ldots, f_{n_1}, \ldots, f_{n_\ell}, \ldots, f_{n_\ell}) \quad (4)$$

where $\binom{m}{k_1 \ \cdots \ k_\ell}$ denotes the multinomial coefficient and the sum runs over:
(a) all finite sequences (k_i) of positive integers whose length (resp. sum) is denoted by ℓ (resp. m), and

(b) all finite sequences (n_i) of positive integers of length ℓ. Moreover, in the argument of G_m, the variable f_{n_i} is repeated k_i times.

The degree of $G_m(f_{n_1},\dots,f_{n_1},\dots,f_{n_\ell},\dots,f_{n_\ell})$ is $r = k_1n_1+\dots+k_\ell n_\ell$ and then contributes to h_r. As a consequence h_r is equal to (4) where the sum is restricted to sequences (k_i), (n_i) such that $k_1n_1+\dots+k_\ell n_\ell = r$. Proposition A.1 now follows from the next lemma.

Lemma A.2. *Let E be a K-vector space. Let $\varphi : E^m \to K$ be a symmetric m-linear form and $\psi : E \to K$ defined by $\psi(x) = \varphi(x,x,\dots,x)$. Given positive integers $k_1,\dots,k_\ell$ whose sum is m and $x_1,\dots,x_\ell \in E$, we have*

$$\left\| \binom{m}{k_1\ k_2\ \cdots\ k_\ell} \cdot \varphi(x_1,\dots,x_1,\dots,x_\ell,\dots,x_\ell) \right\| \le \|\psi\| \cdot \|x_1\|^{k_1} \cdots \|x_\ell\|^{k_\ell}$$

where, in the LHS, the variable x_i is repeated k_i times.

Proof. It is enough to prove that

$$\left\| \binom{m}{k_1\ k_2\ \cdots\ k_\ell} \cdot \varphi(x_1,\dots,x_1,\dots,x_\ell,\dots,x_\ell) \right\| \le \|\psi\|$$

provided that all the x_i's have norm at most 1. We proceed by induction on ℓ. The $\ell = 1$ case follows directly from the definition of $\|\psi\|$. We now pick $(\ell+1)$ integers $k_1,\dots,k_{\ell+1}$ whose sum equals m, together with $(\ell+1)$ elements $x_1,\dots,x_{\ell+1}$ lying in the unit ball of E. We also consider a new variable λ varying in $\mathcal{O}_K$. We set $x'_i = x_i$, $k'_i = k_i$ when $i < \ell$ and $x'_\ell = x_\ell + \lambda x_{\ell+1}$ and $k'_\ell = k_\ell + k_{\ell+1}$. By the induction hypothesis, we know that the inequality

$$\left\| \binom{m}{k'_1\ \cdots\ k'_\ell} \cdot \varphi(x'_1,\dots,x'_1,\dots,x'_\ell,\dots,x'_\ell) \right\| \le \|\psi\|$$

holds for all $\lambda \in K$. Furthermore, the LHS of the inequality is a polynomial $P(\lambda)$ of degree k'_ℓ whose coefficient in λ^j is

$$\binom{m}{k'_1\ \cdots\ k'_\ell} \cdot \binom{k'_\ell}{j} \cdot \varphi(\underline{x}_j) = \binom{m}{k_1\ \cdots\ k_{\ell-1}\ j} \cdot \varphi(\underline{x}_j)$$

with $\underline{x}_j = (x_1,\dots,x_1,\dots,x_{\ell+1},\dots,x_{\ell+1})$ where x_i is repeated k_i times if $i < \ell$ and x_ℓ (resp. $x_{\ell+1}$) is repeated j times (resp. $k'_\ell - j$ times). Since $\|P(\lambda)\| \le \|\psi\|$ for all λ in the unit ball, the norm of all its coefficients must also be at most $\|\psi\|$. From the coefficient of λ^{k_ℓ}, the result follows. □

A.2 Bounding a growing function

We return to the setting of Proposition 2.4. Let $f = \sum_{n\ge0} f_n$, $g = \sum_{n\ge0} g_n$ and $h = \sum_{n\ge0} h_n$ be the analytic expansions of f, g and h. Here f_n, g_n and h_n are the restrictions to the diagonal of some symmetric n-linear forms F_n, G_n and H_n, respectively. We recall that $\Lambda(f)$ is the Legendre transform of the Newton polygon $\mathrm{NP}(f)$ defined in Section 3 [3, Proposition 3.9], and that α is a real number such that $\|n!\| \ge e^{-\alpha n}$ for all positive integers n.

Lemma A.3. *We keep the above notation. If (a,b) satisfies $b \ge a + \Lambda(g)\big(\max(b,\Lambda(h)(a))\big)$ then $b \ge \Lambda(f)(a-\alpha)$.*

Proof. We have $f' = \sum_{n\ge0} f'_n$ where

$$f'_n : U \to \mathcal{L}(E,F), \quad x \mapsto \big(h \mapsto n\cdot F_n(h,x,x,\dots,x)\big).$$

Taking $h = x$, we find $\|f'_n\| \ge \|nf_n\| = \|n\| \cdot \|f_n\|$. Combining this with Proposition A.1, we get

$$\|(r+1)f_{r+1}\| \le \sup_{m,(n_i)} \|g_m\| \cdot \prod_{i=1}^{m} \max(\|f_{n_i}\|, \|h_{n_i}\|)$$

for all nonnegative integers r, where the supremum runs over all pairs $(m,(n_i))$ where m is a nonnegative integer and $(n_i)_{1\le i\le m}$ is a sequence of length m of nonnegative integers such that $n_1+\dots+n_m = r$. We set $u_r = \|r!f_r\|$. Multiplying the above inequality by $\|r!\|$, we obtain:

$$u_{r+1} \le \sup_{m,(n_i)} \|g_m\| \cdot \prod_{i=1}^{m} \max(u_{n_i}, \|n_i!h_{n_i}\|) \qquad (5)$$

since the multinomial coefficient $\binom{r}{n_1\ \cdots\ n_m}$ is an integer and hence has norm at most 1. We now pick two real numbers a and b satisfying the hypothesis of the Lemma. Set $d = \Lambda(h)(a)$. Going back to the definitions of $\Lambda(h)$ and Legendre transform, we get $\|h_n\| \le e^{-an+d}$ for all n. Similarly, from our hypothesis on (a,b), we find $\|g_m\| \le e^{-\max(b,d)\cdot m+b-a}$ for all m. We are now ready to prove $u_r \le e^{-ar+b}$ by induction on r. When $r = 0$, it is obvious because u_0 vanishes. Otherwise, the induction follows from

$$\begin{aligned} u_{r+1} &\le \sup_{m,(n_i)} e^{-\max(b,d)\cdot m+b-a+\sum_{i=1}^m(-an_i+\max(b,d))} \\ &= e^{b-a-ar} = e^{-a(r+1)+b}. \end{aligned}$$

From the definition of u_r, we obtain $\|f_r\| \le u_r \cdot \|r!\|^{-1} \le e^{-(a-\alpha)r+b}$. Thus $b \ge \Lambda(f)(a-\alpha)$. □

We can now conclude the proof of Proposition 2.4 as follows. Given $a \in \mathbb{R}$ and $b = a + \Lambda_g \circ \Lambda_h(a)$, we have to prove that $\Lambda(f)(a-\alpha) \le b$ provided that $b \le \nu$. Thanks to Proposition 2.4, it is enough to check that such pairs (a,b) satisfy the hypothesis of Lemma A.3. Clearly: $b \ge a+\Lambda(g)\circ\Lambda(h)(a)$ since $\Lambda_g \ge \Lambda(g)$, $\Lambda_h \ge \Lambda(h)$ and Λ_g is nondecreasing. Furthermore, from $b \le \nu$, we get $\Lambda_g(b) = \min_{x\in\mathbb{R}} \Lambda_g(x) \le \Lambda_g \circ \Lambda_h(a)$, from which we derive $a + \Lambda_g(b) \le a + \Lambda_g \circ \Lambda_h(a) = b$.

References

[1] Yves Benoist and Jean-Francois Quint, *Introduction to random walks on homogeneous spaces*, Japanese Journal of Mathematics **7** (2012), no. 2, 135–166.

[2] Alin Bostan, Laureano González-Vega, Hervé Perdry, and Éric Schost, *From Newton sums to coefficients: complexity issues in characteristic p*, MEGA'05, 2005.

[3] Xavier Caruso, David Roe, and Tristan Vaccon, *Tracking p-adic precision*, LMS Journal of Computation and Mathematics **17** (2014), no. A, 274–294.

[4] Pierrick Gaudry, Thomas Houtmann, Annegret Weng, Christophe Ritzenthaler, and David Kohel, *The 2-adic CM method for genus 2 curves with application to cryptography*, Asiacrypt 2006, 2006, pp. 114–129.

[5] Kiran S. Kedlaya, *Counting points on hyperelliptic curves using monsky–washnitzer cohomology*, J. Ramanujan Math. Soc. **16** (2001), 323–338.

[6] ———, *p-adic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 125, Cambridge UP, Cambridge, UK, 2010.

[7] Alan Lauder, *Deformation theory and the computation of zeta functions*, Proc. London Math. Soc. **88** (2004), no. 3, 565–602.

[8] Reynald Lercier and Thomas Sirvent, *On Elkies subgroups of ℓ-torsion points in elliptic curves defined over a finite field*, J. Théorie des Nombres des Bordeaux **20** (2008), 783–797.