**ExactpAdics**: An exact representation
of $p$-adic numbers

Christopher Doris
University of Bristol
christopher.doris@bristol.ac.uk
April 2017

**Abstract**

We describe two new packages `ExactpAdics` and `ExactpAdics2` for
the Magma computer algebra system for working with $p$-adic numbers
exactly, in the sense that numbers are represented lazily to infinite $p$-
adic precision. This has the benefits of increasing user-friendliness and
speeding up some computations, as well as forcibly producing provable
results. The two packages use different methods for lazy evaluation,
which we describe and compare in detail. The intention is that this
article will be of benefit to anyone wanting to implement similar func-
tionality in other languages.

1 Introduction

When dealing with completed fields, such as $\mathbb{R}$ or $\mathbb{Q}_p$, it is generally quite
difficult to represent elements exactly. Instead, the commonest way to rep-
resent elements is by specifying them to some pre-determined precision, and
then performing operations such as arithmetic to this precision also. This is
the foundation of floating point arithmetic. For example, one might repre-
sent the real number $e$ by its approximation $2.718281828$ to a precision of 10
real digits. We say such a representation is inexact because several real num-
bers can have the same representation: $e$, $2.718281828$ and $2.7182818281$ all
have the same representation to 10 digits precision.

Such a representation is also usually zealous meaning that when an op-
eration is performed, such as multiplication, it is immediately computed to
the required precision. For instance, computing $e \times e$ will work to 10 digits
precision and actually compute $2.718281828 \times 2.718281828 = 7.389056096$.
In fact, $e \times e = 7.389056098\ldots$, demonstrating that precision errors can
creep into the results, so that they are in fact less precise than the precision
claims.

An often-suggested alternative to zealous arithmetic is lazy arithmetic,
wherein an operation does not produce an answer per-se, but a “promise
to produce an answer to a desired precision*. That is, calling $e \times e$ would not produce the approximation 2.718281828, but would produce a function which, when called with an integer $k$, returns an approximation to $e \times e$ to $k$ digits precision.

Such a function can be said to be an *exact* representation of a real number, because no two distinct real numbers have the same representation: for a sufficiently large precision $k$, the representing functions will return different approximations.

These comments hold true for $p$-adic numbers too. For instance, an element of $\mathbb{Q}_p$ is generally represented in zealous, inexact arithmetic by its residue class in $\mathbb{Q}_p/p^k\mathbb{Z}_p$ for some *absolute precision* $k$: e.g. $1 + 2^{10}\mathbb{Z}_2$ might represent $1$, $1 + 2^{10}$ or $1 + 5 \times 2^{100}$.

There are numerous implementations of such $p$-adic arithmetic. FLINT [6] provides some low-level arithmetic with elements of $\mathbb{Q}_p$, univariate polynomials over $\mathbb{Q}_p$, and unramified extensions of $\mathbb{Q}_p$. Sage [11] and Magma [1] have more fully-featured implementations, including arbitrary finite extensions of $\mathbb{Q}_p$ and higher-level routines for tasks such as factoring.

Also of note is an implementation in Mathemagix [12] of the so-called *relaxed $p$-adic arithmetic*, which treats elements of $\mathbb{Q}_p$ like an infinite sequence of $p$-adic coefficients, somewhat like $\mathbb{F}_p((t))$, and represents them by a truncated sequence followed by a function to retrieve the next coefficient. This representation is therefore exact, because for different numbers, these streams of digits must eventually diverge. This has specific uses in $p$-adic recursion solving, and in principle is useful in general, but is somewhat more complicated to implement than the lazy arithmetic presented in this article, and as such is less fully featured.

A more in-depth description of different $p$-adic arithmetic systems is given by Caruso [2].

In this article, we present two new implementations of two different lazy, exact $p$-adic arithmetic systems. The implementations are written for the Magma computer algebra system [1] which, as mentioned above, already has a fully-featured implementation of zealous, inexact $p$-adic arithmetic. Our packages, called *ExactpAdics* and *ExactpAdics2*, aim to use the inexact functionality already available as much as possible, in order to provide a more user-friendly wrapper. This allows for rapid addition of new features to the exact arithmetic as soon as they are available inexactly.

To the author’s knowledge, these are the first highly-featured, general-purpose implementations of lazy $p$-adic arithmetic.

This article describes the rationale and the fundamental concepts behind the packages, but does not constitute a user manual. The user-manuals are available online at [https://cjdoris.github.io/ExactpAdics](https://cjdoris.github.io/ExactpAdics) and [https://cjdoris.github.io/ExactpAdics2](https://cjdoris.github.io/ExactpAdics2), and the packages may be downloaded from here also.

At the time of writing, we recommend the typical user to use the *ExactpAdics2*
As an application, these packages have been used to implement the algorithm in [3] to compute the 2-part of the conductor of a hyperelliptic curve of genus 2 defined over a number field. This implementation is available from [https://cjdoris.github.io/Genus2Conductor](https://cjdoris.github.io/Genus2Conductor). It uses such high-level $p$-adic routines as: computing the completion of a number field at a finite place (§9.4); computing the factorization of a univariate polynomial (§9.12) and the fields defined by its factors; and Hensel-lifting roots of a system of multivariate equations (§9.9).

As another application, these packages can optionally be used with the implementation of the algorithms described in [4] for computing the Galois group of a $p$-adic polynomial. This is available from [https://cjdoris.github.io/pAdicGaloisGroup](https://cjdoris.github.io/pAdicGaloisGroup). With either package present the Galois group algorithm becomes provably correct, whereas otherwise with inexact $p$-adics there is no such guarantee. We also find that the algorithms run faster with exact $p$-adics, at least for reasonably high-degree inputs.

**Acknowledgements.** This work was partially supported by a grant from GCHQ.

### 1.1 Terminology

Suppose $K$ is a $p$-adic field (a finite extension of $\mathbb{Q}_p$), with ring of integers $\mathcal{O} = \mathcal{O}_K$ and uniformizing element $\pi = \pi_K$. The $\pi$-adic valuation is denoted $\text{val} = \text{val}_K$ such that $\text{val}(\pi) = 1$.

When we refer to an *inexact* (representation of a) $p$-adic number $x \in K$, we mean a conjugacy class $x + \pi^k\mathcal{O}$. We refer to $k$ as the *absolute precision* of (the representation of) the number.

Equivalently, it may be represented as $\pi^v(y + \pi^r\mathcal{O})$ where $y \in \mathcal{O}$ and $r \geq 0$. We refer to $v$ as the *weak valuation* of $x$; it is a lower bound on the true valuation of $x$. We refer to $r$ as the *relative precision*; it bounds the number of non-zero $\pi$-adic digits of $x$ known. Note that $v + r = k$.

We say that $x$ is *weakly zero* if $y \in \pi^r\mathcal{O}$, that is if the representation is of the form $\pi^{v+r}\mathcal{O}$. Note:

- If $x$ is not weakly zero, then it is not zero.
- If $r = 0$ then $x$ is weakly zero.

We typically enforce the following *normalizing condition*: if $r > 0$ then $y \in \mathcal{O}^\times$. Now note:

- If $x$ is not weakly zero, then its valuation is exactly $v$.
- $x$ is weakly zero if and only if $r = 0$ (and if and only if $k = v$).

Magma’s built-in $p$-adics (FldPad, FldPadExact, etc.) are inexact in this sense, and satisfy the normalizing condition. We note that *prime* $p$-adic
fields — i.e. \( \mathbb{Q}_p \) — as opposed to their elements, can themselves naturally be represented exactly by the prime itself. Extensions of the form \( K(x)/(f(x)) \) are usually represented inexactly via an inexact representation of the polynomial \( f(x) \in K[x] \); however we note that Magma does additionally have a builtin exact representation of extensions, represented by a map \( m : \mathbb{Z} \to K[x] \) such that \( m(k) \) is a defining polynomial to precision \( k \). We refer to this latter representation as *semi-exact*, since the field is represented exactly but its elements are represented inexactly.

The residue class field \( \mathcal{O}/\pi\mathcal{O} \) is denoted \( \mathbb{F} = \mathbb{F}_K \), and \( \bar{x} \in \mathbb{F} \) denotes the residue class of \( x \in \mathcal{O} \).

A polynomial \( f(x) = \sum_{i=0}^{d} f_i x^i \in K[x] \) of degree \( d \) is *Eisenstein* if \( \text{val}(f_0) = 1 \), \( \text{val}(f_i) \geq 1 \) for \( 1 \leq i < d \) and \( \text{val}(f_d) = 0 \). It is irreducible, its roots have valuation \( \frac{1}{d} \), and so it defines a totally ramified extension \( K(x)/(f(x)) \) of degree \( d \) such that \( x + (f(x)) \) is a uniformizer.

A polynomial \( f(x) = \sum_{i=0}^{d} f_i x^i \in \mathcal{O}[x] \) of degree \( d \) is *inertial* if \( \text{val}(f_d) = \text{val}(f_0) = 0 \) and \( \bar{f}(x) = \sum_{i=0}^{d} \bar{f}_i x^i \in \mathbb{F}[x] \) is irreducible over the residue class field \( \mathbb{F} \). It is irreducible, the residue classes of its roots generate an extension of \( \mathbb{F} \) of degree \( d \), and so it defines an unramified extension \( K(x)/(f(x)) \).

### 1.2 Comparison of zealous and lazy arithmetic

#### 1.2.1 Precision

In zealous arithmetic, the user is generally required to choose a precision to work at in advance. Then all computations are performed to that precision, and it may happen that the precision chosen was not sufficient. In this case, the user will probably start the computation over with a higher precision. This process of manually increasing the precision of a computation can be burdensome for the user. In lazy arithmetic, such precision decisions are made automatically as far as possible.

**Example 1.1.** Here is a typical interactive Magma session, using its builtin lazy arithmetic:

```plaintext
> // try to factorize at precision 10
> K := pAdicField(2, 10);
> R<x> := PolynomialRing(K);
> f := my_favourite_polynomial(R);
> Factorization(f);
error: ...
> // try to factorize at precision 20
> K := pAdicField(2, 20);
> R<x> := PolynomialRing(K);
> f := my_favourite_polynomial(R);
> Factorization(f);
error: ...
```
> // try to factorize at precision 40
> K := pAdicField(2, 40);
> R<x> := PolynomialRing(K);
> f := my_favourite_polynomial(R);
> Factorization(f);
> [ <x^10 + ... >, ... ]

Using lazy arithmetic provided by our package, the equivalent session would be the following. Note that there is no explicit mention of precision.

> K := ExactpAdicField(2);
> R<x> := PolynomialRing(K);
> f := my_favourite_polynomial(R);
> Factorization(f);
> [ <x^10 + ... >, ... ]

In lazy arithmetic, each individual computation is performed to approximately the smallest precision it can be, and so precisions are very “local” in the computation. In zealous arithmetic, the precision is generally chosen once at the start of a computation, so each operation is performed to the same precision, and so precisions are more “global”. If there is a single operation requiring a high “global” precision, this increases the precision that all other operations are performed to, which is a performance hit compared to lazy arithmetic.

Example 1.2. An example comes from the conductor algorithm mentioned in the introduction. One portion of this algorithm takes a polynomial \( f(x) \in \mathbb{Q}_2[x] \), computes its factorization, chooses a factor \( g(x) \), computes the extension \( L/\mathbb{Q}_2 \) defined by \( g \), and then finds a root of \( g \) in \( L \). Usually, the precision required for the factorization far exceeds that of the root-finding; however, because the root-finding is over an extension \( L \), if it were to be done at the same high precision as the factorization, its run-time would often dominate.

1.2.2 Correctness and provability

When a \( p \)-adic number \( x \in K \) is represented inexactely as a class \( x + \pi^k \mathcal{O} \), then it can be ambiguous whether it is really representing \( x \) or the class itself. For many operations, the distinction makes no difference; for example since

\[
(x + y) + \pi^k \mathcal{O} = (x + \pi^k \mathcal{O}) + (y + \pi^k \mathcal{O})
\]

then addition works the same in either interpretation. For other operations, Magma can produce potentially misleading answers; for example if \( x \) is represented as \( 0 + \pi^k \mathcal{O} \) then \( \text{Valuation}(x) \) will return \( k \), when in fact all we really know is that \( \text{val}(x) \geq k \).
Definition 1.3. Suppose $F$ is a mathematical function and suppose $\tilde{F}$ is a programmatic function intended to implement $F$, so it takes as inputs representations of the inputs of $F$ and returns as outputs representations of the outputs of $F$. We say that $\tilde{F}$ represents $F$ if for all possible inputs $X$ to $F$ and representations $\tilde{X}$ of $X$, that $\tilde{F}(\tilde{X})$ either does not return successfully or returns a representation of $F(X)$.

Hence if $\tilde{F}$ represents $F$, then its outputs depend only on the inputs being represented, and not on the representation of the inputs themselves. In the case of $p$-adic computation, this means that the output of $\tilde{F}$ should not depend on the precision that its inputs were given to, and therefore is unambiguously a function of the $p$-adic value, and not its representation.

As already indicated, the Valuation intrinsic in Magma does not represent the valuation function. Also equality is not represented, because it actually is equality of the representation: if $x = 1$ and $y = 1 + 2^{10}$ are both represented by $1 + 2^{10}\mathbb{Z}_2$ then $x \text{ eq } y$ will be true. In fact, it is not possible to determine that two $p$-adic numbers are equal when given to any finite precision, and it is only possible to tell that they are unequal if they are given to sufficiently large precision.

As another example, given a polynomial $f(x)$ represented as $\tilde{f}(x) = (1 + \pi^{10}\mathcal{O})x^2 + (0 + \pi^{10}\mathcal{O})$, the Roots intrinsic in Magma will return a double root $\tilde{r} = 0 + \pi^{10}\mathcal{O}$ in $K$. This is correct as a function of the representations themselves, since $\tilde{f}(\tilde{r}) = 0 + \pi^{10}\mathcal{O}$ represents 0, but if $f(x) = x + 2^{11}$ then it is irreducible and therefore has no roots in $K$. Similarly Factorization and GCD do not represent factorization and greatest common divisor of $p$-adic polynomials.

In our packages, if the name of an intrinsic function is the name of a mathematical function, then the intrinsic represents the function. For example, our Valuation intrinsic (see Examples 4.1 and 5.7) will only return the true valuation of the given number; if the input is weakly zero, it may try to increase its precision, and could potentially do this forever (if the input is 0) or raise a precision error, but it is guaranteed that if it returns, its return value is correct.

In some cases, such as Roots (§9.8) and Factorization (§9.12), the correctness of the output is forced by the fact that the outputs are given exactly. That is, if Roots returns a root (exactly), then it by definition comes with a program to compute an approximation to the root to arbitrarily high precision, and therefore assuming the program is correct this is a proof that the root is correct. In the case of Roots, it is Hensel’s lemma which provides this proof.

The intrinsics which do not represent a function, and therefore depend on the representation, are given names which make this clear. The terms Weakly and Definitely are used to denote tests which can give false positives or false negatives; for example IsWeaklyZero is true if its input ap-
pears to be zero up to some precision (but does not guarantee it is zero), and \texttt{IsDefinitelyPrimitive} returns true if its input can be proven to be a primitive element (but if it returns false, this does not imply that its input is not primitive). Similarly the term \texttt{Weak} denotes non-representing functions, so \texttt{WeakValuation} returns the $k$ in $0 + \pi^k \mathcal{O}_K$ and is therefore actually a lower bound on the true valuation; and \texttt{WeakDegree} returns an upper bound on the degree of a polynomial, but which may be incorrect if its top coefficient is actually zero.

1.2.3 Overheads

The main down-sides of lazy arithmetic are the extra time and memory overheads introduced. In lazy arithmetic, $p$-adic values depend on other $p$-adic values, and all these dependencies need to be kept in memory for the duration of a computation. Each time an operation is performed, some dependency tracking and propagation needs to occur, which entails some processing time overhead.

This said, we find that these overheads do not usually dominate the run-time of lazy $p$-adic arithmetic unless one performs a large number of ordinarily very fast operations, such as basic arithmetic. If this is the case, then one can consider implementing the whole sequence of operations as a new atomic $p$-adic operation, which therefore now only contributes a single node to the dependency graph.

1.3 Structure of this article

The first three sections describe the \texttt{ExactpAdics} package.

In §2 we describe the core data types and functionality provided by the package, including a simplified description of the lazy evaluation of $p$-adic numbers.

In §3 we describe the lazy evaluation scheme actually employed by the package, which includes tracking dependencies between different $p$-adic values.

In §4 we describe “precision strategies”, which are a way of programatically avoiding precision errors with minimal input from the user. This is not a core feature, but greatly improves user-friendliness.

Next we describe the \texttt{ExactpAdics2} package and compare.

In §5 we describe the core data types and functionality provided by the package, including a description of the lazy evaluation scheme.

In §6 we compare the merits of the approaches taken by the two packages, including timings on some problems of interest.

The remaining sections describe additional features which either improve user-friendliness or provide more functionality. These features are mainly present in both packages.
In §7 we describe additional structures which are not core functionality, namely multivariate polynomials and tuples.

In §8 we describe our representation of valuations (defined in a generic sense) of $p$-adic objects, and the operations available on them.

Finally in §9 we give an overview of additional features not covered elsewhere. This is largely to demonstrate that these packages are of practical use, since they include features such as root finding, factorization, residue classes and completions. We also provide some implementation notes.

1.4 Pseudocode

As the package is written in Magma, we shall use a simplification of the Magma language to demonstrate concepts. As this article may be useful to implement similar functionality in other languages, we summarise the syntax here.

Every variable has a type. For example a ring of integers has type \texttt{RngInt}, an integer has type \texttt{RngIntElt}, a boolean (true or false) has type \texttt{BoolElt} and an inexact $p$-adic field has type \texttt{FldPad}. New types are defined as

\begin{verbatim}
  type NAME[ELT]: PARENT
\end{verbatim}

where \texttt{NAME} is the name of the type. The part in brackets is optional, but when given the type \texttt{ELT} is also declared, and \texttt{NAME} is actually a structure with elements of type \texttt{ELT}. The part after the colon is optional, but when given the new type is a child of \texttt{PARENT} in the type hierarchy, and in particular the new type inherits the attributes from \texttt{PARENT}.

A type has attributes, which are named pieces of data attached to instances of the type. Attributes may be attached to a type by

\begin{verbatim}
  attributes TYPE: ATTR1, ATTR2, ...
\end{verbatim}

where \texttt{TYPE} is the name of the type, and \texttt{ATTRn} are the names of the attributes.

Instances of the type are created like \texttt{x := New(TYPE)}, and its attributes are accessed like \texttt{x \textasciitilde ATTR}.

There are three types of functions in Magma: \texttt{function}, \texttt{procedure} and \texttt{intrinsic}, all declared in a similar fashion, such as:

\begin{verbatim}
  function example(x, y : z := 0)
    return x + y + z
  end function
\end{verbatim}

The difference between the three is that a function returns a value, and should not have any side-effects, a procedure does not return a value but can have side-effects (in particular an input may be passed by reference like \texttt{\textasciitilde x} and it becomes modifiable), and an intrinsic is a function or procedure
which forms the main user-interface. The $z:=0$ part is an optional parameter named $z$ whose default value is 0. Furthermore, intrinsics may have type declarations on its inputs and outputs, which allows overloading of intrinsics with the same name but different type signatures. For example:

```plaintext
intrinsic ImportantExpression(
    x :: RngIntElt,
    y :: RngIntElt,
    z :: RngIntElt)
-> FldRatElt
return (x^2 + y^2) / z^3
```

is an intrinsic taking three integers and returning a rational.

Note that any pseudocode in this article is illustrative, and does not necessarily match the code in the implementation. The pseudocode is presented as simply as is possible to get the ideas across, whereas the real code will contain more checks and optimizations.

## 2 ExactpAdics: Core structures and elements

### 2.1 Abstract base types

In ExactpAdics we have an abstract base type `StrPadExact` representing any kind of exact $p$-adic structure or set, and such a set has elements of type `PadExactElt`:

```plaintext
type StrPadExact[PadExactElt]
```

We shall later have sub-types representing the field of $p$-adic numbers (§2.2), rings of polynomials over $p$-adic numbers (§2.3), and more (§7).

Such a structure will always have an approximation which is an analogous inexact structure:

```plaintext
attributes StrPadExact: approximation
```

Elements always have a parent structure to which they belong, as well as an approximation and an update function:

```plaintext
attributes PadExactElt: parent, approximation, update
```

The approximation is an element of the approximation of the parent, and so provides a finite-precision approximation to the element. The update function provides the means to update the approximation arbitrarily precisely, and will be described later in this section and in §3. Figure 1 illustrates the relationships between these attributes.

We provide some universal intrinsics to retrieve the parent of an element, its absolute precision and its weak valuation. The latter two are defined to be the absolute precision and weak valuation of the approximation, and therefore can change over time.
2.2 \( p \)-adic fields

An exact \( p \)-adic field is represented by the type \texttt{FldPadExact} (compare the name to the inexact \texttt{FldPad} type in Magma) which derives from \texttt{StrPadExact} (and so inherits its attributes) and has some additional attributes:

\begin{verbatim}
type FldPadExact[FldPadExactElt]: StrPadExact
attributes FldPadExact: xtype, prime, defining_polynomial
\end{verbatim}

The \texttt{xtype} attribute takes one of the special enumerated values:

- \texttt{PRIME}: the field is \( \mathbb{Q}_p \) for some \( p \), and the \texttt{prime} attribute is \( p \).
- \texttt{INERT}: the field is an unramified extension of another exact \( p \)-adic field \( K \), and the \texttt{defining_polynomial} attribute is an inertial polynomial \( f(x) \in K[x] \), defining the extension as \( K(x)/(f(x)) \).
- \texttt{EISEN}: the field is a totally ramified extension of another exact \( p \)-adic field \( K \), and the \texttt{defining_polynomial} is an Eisenstein polynomial \( f(x) \in K[x] \), defining the extension as \( K(x)/(f(x)) \).

The \texttt{approximation} field of an exact \( p \)-adic field is a corresponding inexact field, which in Magma has type \texttt{FldPad}. For the \texttt{PRIME} field \( \mathbb{Q}_p \), this is simply \texttt{pAdicField(p)}.
For extensions (INERT or EISEN) the situation is a little more complicated. In essence, we want the approximation to be the extension defined by an approximation of the defining_polynomial. The problem is that later we may want a more precise approximation, and so we have two choices:

- Replace the approximation field with a more precise approximation whenever it is required. This means that any element of the field may have an approximation lying in an older approximation field, and so will need to be coerced into the latest approximation field at some time.

- Use Magma’s built-in semi-exact representation of $p$-adic extensions (§1.1): ext$K \mid m$ where $m$ is a map taking an integer and returning an approximation of the defining_polynomial to that precision.

We use the second choice because the explicit coercion between different approximation fields in the first choice was found to add a performance hit. It also has the benefit that we can talk of the approximation field, since it does not change in time.

Elements of exact $p$-adic fields are represented by the type FldPadExactElt. The meanings of its attributes are inherited from its parent type PadExactElt but to be explicit:

- parent is the FldPadExact field to which it belongs;
- approximation is an element of the approximation field of its parent, and is therefore a FldPadElt;
- update is its update function, used to update the approximation.

We define coercion so that $K ! <\text{init}, \text{mkupdate}>$ creates an element of $K$ whose initial approximation is $\text{init}$ and whose update function is $\text{mkupdate}(x)$ where $x$ is the element being created, thus allowing the update function to refer to $x$ itself:

```plaintext
intrinsic IsCoercible(K :: FldPadExact, args :: Tup) -> FldPadE_elt
x := New(FldPadE_elt)
x\'parent := K
x\'init := args[1]
x\'update := args[2](x)
return true, x
```

We provide intrinsics to access basic information; intrinsics for inertia degree and ramification degree are defined similarly:

```plaintext
intrinsic IsPrimeField(K :: FldPadExact) -> BoolElt
return K\'xtype eq PRIME
```

11
intrinsic DefiningPolynomial(K :: FldPadExact) -> RngUPolElt_FldPadExact
if IsPrimeField(K) then
  error "not an extension"
else
  return K\text{defining\_polynomial}

intrinsic BaseField(K :: FldPadExact) -> FldPadExact
return BaseRing(DefiningPolynomial(K))

intrinsic Degree(K :: FldPadExact) -> RngIntElt
return Degree(DefiningPolynomial(K))

intrinsic AbsoluteDegree(K :: FldPadExact) -> RngIntElt
if IsPrimeField(K) then
  return 1
else
  return Degree(K) * AbsoluteDegree(BaseField(K))

2.3 Univariate polynomials

A univariate polynomial ring over a \( p \)-adic field is represented by the type RngUPol_FldPadExact (analogous to the inexact type RngUPolElt[FldPad] in Magma) which also derives from StrPadExact:

type RngUPol_FldPadExact[RngUPolElt_FldPadExact]
attributes RngUPol_FldPadExact: base\_ring

Such a ring is defined by its base\_ring, an exact \( p \)-adic field (i.e. of type FldPadExact).

The approximation of such a ring must be the univariate PolynomialRing of the approximation of the base\_ring (i.e. of type RngUPol[FldPad]).

2.4 The update function

The update attribute of a PadExactElt is a means to increase the precision of its approximation to a given absolute precision. Therefore it is natural to define it as a procedure which takes as input an absolute precision \( k \), and whose side-effect is to replace the approximation by one whose precision is at least \( k \). Using this definition will result in a working implementation of exact \( p \)-adics, but as we shall see in \S3 it has some drawbacks and so in reality we use a slightly different definition. For now, however, it suffices to think of the update function in this way.

In the update function, instead of modifying the approximation of an element directly, one should use the following intrinsic which first checks
that the update is consistent with the pre-existing approximation and in reality may perform more checks:

```
intrinsic Update(x :: FldPadExactElt, xx :: FldPadElt)
  assert IsWeaklyEqual(x\'approximation, xx)
  x\'approximation := xx
```

Instead of calling the `update` function directly, we should use the following intrinsic which ensures it is only called when required. In fact, since the `update` function is not a function at all then this intrinsic will actually have a different definition (§3.3), but with the same effect.

```
intrinsic IncreaseAbsolutePrecision(x :: PadExactElt, n)
  if not AbsolutePrecision(x) ge n then
    x\'update(n)
```

In practice, we don’t usually just want to increase the precision of an element, but we want the approximation itself. Hence we also make available an intrinsic `Approximation` to retrieve an approximation to an element to a certain absolute precision. It simply has to increase the absolute precision of the element, then return its approximation, perhaps with its precision decreased to the desired value.

```
intrinsic Approximation(x :: FldPadExactElt, n)
  IncreaseAbsolutePrecision(x, n)
  return ChangeAbsolutePrecision(x\'approximation, n)
```

To increase the precision of an extension field, we just need to increase the precision of its `defining_polynomial` correspondingly. This ensures that the next time the semi-exact `approximation` field retrieves a defining polynomial, it will already be available to the given precision. There is nothing to be done for `PRIME` fields, since the prime is already represented exactly.

```
intrinsic IncreasePrecision(K :: FldPadExact, n)
  if not IsPrimeField(K) then
    IncreaseAbsolutePrecision(K\'defining_polynomial, n)
```

The precision of a polynomial ring is the precision of its base ring, so to increase one we just have to increase the other:

```
intrinsic IncreasePrecision(R :: RngUPol_FldPadExact, n)
  IncreasePrecision(R\'base_ring, n)
```

### 2.5 Examples

#### Example 2.1.

Here is a definition of binary addition on two \(p\)-adic numbers. The initial approximation is simply the sum of the approximations of the inputs. The update function retrieves approximations to the inputs to the
required precision, adds them, and sets this as the new approximation for
the sum.

\begin{verbatim}
intrinsic '+(x :: FldPadExactElt, y :: FldPadExactElt)
 -> FldPadExactElt
init := x\`approximation + y\`approximation
mkupdate := function (z)
  return procedure (n)
    Update(z, Approximation(x, n) + Approximation(y, n))
  return Parent(x) ! <init, mkupdate>
\end{verbatim}

This example is an over-simplification compared to the true implementation
in the following ways:

- The update function is not quite as described. See §3.
- The initial approximation init is computed to the current precision
  of the inputs, which may be overkill if they are both very precise.
  Instead, the implementation adds together approximations to “first
  precision”, i.e.

\begin{verbatim}
init := ChangeAbsolutePrecision(x\`approximation,
  Min(WeakValuation(x)+1, AbsolutePrecision(x)))
 + ChangeAbsolutePrecision(y\`approximation,
  Min(WeakValuation(y)+1, AbsolutePrecision(y)))
\end{verbatim}

As an optimization, most functions will compute the initial approxi-
imation from the inputs to first precision if possible.

- It should be checked that the inputs have the same parent field, or
  can be coerced to a common field.

\begin{example}
Here we give a definition of binary multiplication, which is
very similar to addition. The main change is in computing the precision
required in the approximations.

\begin{verbatim}
intrinsic '*(x :: FldPadExactElt, y :: FldPadExactElt)
 -> FldPadExactElt
init := x\`approximation * y\`approximation
mkupdate := function (z)
  return procedure (n)
    Update(z, Approximation(x, n - WeakValuation(y))
         * Approximation(y, n - WeakValuation(x)))
  return Parent(x) ! <init, mkupdate>
\end{verbatim}
\end{example}
3 ExactpAdics: Dependency tracking

3.1 Motivation

So far we have described a simple scheme for implementing exact $p$-adics, but it has drawbacks.

Example 3.1. Suppose we are given elements $a, b \in \mathbb{Q}_p$, and compute $c = a^3 + a^2b + ab^2 + b^3$, and wish to increase the absolute precision of $c$ to 100.

We therefore require each of the summands $a^3, a^2b, ab^2, b^3$ to absolute precision 100. Now suppose that $\text{val}(a) = 10$ and $\text{val}(b) = 0$, so in fact we require the summands to relative precisions 70, 80, 90, 100 respectively. Hence we require $a$ and $b$ to these same relative precisions.

Therefore, if we increase the precision of each summand in turn to its required value, then we will be updating $a$ first to relative precision 70, then 80, then 90 — i.e. absolute precisions 80, 90 and 100 — which is 3 separate updates. If updating $a$ is an expensive operation, then this could become a performance issue.

Clearly, the right thing to do in this situation is to observe that we only need to update $a$ once to absolute precision 100. With the current description of the update function, this is not possible.

Our solution is to split updates into two steps: the first step identifies which other updates are required to occur first, we call these dependencies; the second step actually performs the update. With this explicit separation, we can find all of the dependencies of a calculation before satisfying any of them, allowing us to remove any redundancy as in the above example.

In the example, $c$ has 4 dependencies, namely the 4 summands. Each of these summands in turn depends on one or both of $a$ and $b$. There is redundancy in these dependencies because $a$ and $b$ each appear three times, and therefore could be merged.

3.2 Getters

We encapsulate these ideas into a new type\footnote{In the package it is actually called ExactpAdics_Gettr}:

\begin{verbatim}
type Getter
attributes Getter: state, get_dependencies, get_value

A Getter represents an exact $p$-adic computation with dependencies. The state attribute is some getter-specific state which is passed by reference (and hence is modifiable) into the other functions.

The get_dependencies attribute is a procedure(-state, -deps) which assigns to deps a list of dependencies. A dependency is a pair $<x,n>$ where
\end{verbatim}
x is some $p$-adic value (i.e. a PadExactElt, such as a $p$-adic number or polynomial) and $n$ is an absolute precision. Such a dependency should be interpreted as the getter saying “I can’t compute my value until these values are to these absolute precisions.” We say a dependency is satisfied if the absolute precision of $x$ is at least $n$.

The $\text{get\_value}$ attribute is a procedure($\text{-state, -value}$) which, assuming that the dependencies previously reported are all satisfied, either assigns something to value or doesn’t. If it does, then this is interpreted as the value of the computation. If it doesn’t, then this is interpreted as the getter having more dependencies, and so $\text{get\_dependencies}$ needs to be called again.

Evaluating a getter means getting the value from the $\text{get\_value}$ procedure. Of course, this requires satisfying the dependencies reported by $\text{get\_dependencies}$ first, and leads to a recursive dependency satisfaction algorithm which we describe shortly.

### 3.3 Update function

With getters defined, we may now define precisely what an update function is: it is a function taking as input an absolute precision $n$ and returning a Getter. This getter, when evaluated, will have the side-effect of increasing the absolute precision of the element to $n$. The value of the getter is ignored.

With this definition, $\text{IncreaseAbsolutePrecision}$ would actually be defined like so:

```plaintext
intrinsic IncreaseAbsolutePrecision(x :: PadExactElt, n)
  if not AbsolutePrecision(x) ge n then
    ignored := Evaluate(x\ update(n))
```

Now to increase the absolute precision of a value, we just need to know how to evaluate a getter.

### 3.4 Evaluating getters

To evaluate a getter requires conceptually three steps: first we retrieve its dependencies from $\text{get\_dependencies}$, then we satisfy those dependencies, then we retrieve the value via $\text{get\_value}$. If $\text{get\_value}$ did not return a value, then we will need to repeat these steps.

Each of the dependencies is a pair $<x, n>$ of a $p$-adic element and an absolute precision. Calling $x\ update(n)$ returns a getter which, on evaluation, increases the absolute precision of $x$ to $n$, which we require. Hence we have reduced the problem of evaluating the original getter to the problem of evaluating these dependent getters. Recursing, we traverse the tree of dependencies all the way to its leaves. Figure 2 illustrates this dependency tree for the motivating example (assuming $a$ and $b$ themselves have no dependencies).
To avoid duplicated work, we want to combine all nodes for the same value together, taking the maximum of their absolute values, resulting in a directed acyclic graph such as in Figure 3.

A sink in this graph is precisely a getter with no dependencies. Therefore it may be evaluated and removed from the graph. Repeating, we will eventually reach the source, which can also be evaluated, and we have succeeded.

In practice, we do not represent these dependencies as a tree at all, but we take advantage of a simple fact: if the value \( x \) was created before value \( y \), then \( x \) cannot possibly depend on \( y \). We therefore keep track of the order of creation of elements by giving them a new attribute

\[ \text{attributes } \text{PadExactElt: id} \]

to which we assign the value of a global counter when the element is created. We now represent the nodes of the graph simply as an associative array, where the node \( <x,n> \) is the value at the index \( x\cdot id \). Adding a new dependency into the tree is a matter of checking if there is already a dependency for this \( x \); if so, then we should combine the old and new absolute precision in the array; if not, then we add the new node into the array. Traversing the tree in dependency order is now a matter of running through the indices of the array in sorted order.
We now present a version of this algorithm. The first procedure `add_dependencies` takes a list of dependency pairs `<x,n>` and recursively adds them and all their own dependencies into the array.

```plaintext
procedure add_dependencies(~array, todo_list)
    while #todo_list gt 0 do
        // pop an item from the todo list
        x, n := Pop(~todo_list)
        // if this is a new dependency, or the target
        // precision is greater than the existing one,
        // and it is not already satisfied, then replace
        // it and compute more dependencies to put in
        // the todo list
        if (x`id not in array or not n le array[x`id][2])
            and not n le AbsolutePrecision(x)
            then
                getter := x`update(n)
                array[x`id] := <x, n, getter>
                getter`get_dependencies(~getter`state, ~deps)
                for dep in deps do
                    Append(~todo_list, dep)

The procedure `satisfy_dependencies` takes a list of dependency pairs and satisfies them all, first by calling `add_dependencies` to make an array of all dependencies, and then by running through the dependencies in order and trying to satisfy them. Here is one possible implementation:

```plaintext
procedure satisfy_dependencies(deps)
    // initially compute all dependencies
    array := AssociativeArray()
    add_dependencies(~array, deps)
    // keep trying until the graph is empty
    while #array gt 0 do
        // traverse the nodes in order
        for i in Sort(Keys(array)) do
            // do the update
            getter := array[i][3]
            getter`get_value(~getter`state, ~value)
            if assigned value then
                // success: remove the entry from the array
                delete array[i]
            else
                // failure: get more dependencies and start over
                getter`get_dependencies(~getter`state, ~deps)
                add_dependencies(~array, deps)
                break
```
The implementation in the `ExactpAdics` package behaves a little differently: if a particular node fails, then instead of immediately jumping back to the bottom of the tree, we continue traversing it to the top, but skipping over nodes which now have unsatisfied dependencies. This is made possible by changing `add_dependencies` to explicitly track the children and parents of each node in the array, and altering `satisfy_dependencies` to:

```plaintext
procedure satisfy_dependencies(deps)
    // initially compute all dependencies
    array := AssociativeArray()
    add_dependencies(~array, deps)
    // keep trying until the graph is empty
    while #array gt 0 do
        // traverse the nodes in order
        for i in Sort(Keys(array)) do
            item := array[i]
            if item has no children then
                // do the update
                getter := item[3]
                getter.get_value(~getter.state, ~value)
                if assigned value then
                    // success: remove the entry from the array
                    for each parent of item do
                        remove item as a child of parent
                    delete array[i]
                else
                    // failure: get more dependencies
                    getter.get_dependencies(~getter.state, ~deps)
                    add_dependencies(~array, deps)
            end if
        end for
    end while
end procedure
```

Observe the main differences are that we now need to check a node has no children before processing it; when a node succeeds, we now need to remove it from the list of children of each of its parents; and when it fails, we no longer break out from looping over nodes.

Which of these two routines is better is arguable. The former routine may suffer from updating an early element of the graph, and then discovering later that the same element needs to be updated again, whereas the latter avoids this problem by performing as many updates on the graph as possible before starting over again. On the other hand, the latter routine may suffer from traversing the whole graph needlessly if an early node failed and is a child of everything else.

In practice, on problems of interest, we found that the latter routine usually performed better. That is, it was sometimes significantly faster and was rarely significantly slower, which is why the latter is used in the current implementation.
We also now define the intrinsic which evaluates a getter:

```plaintext
intrinsic Evaluate(g :: Getter)
loop
  g^\prime get_dependencies(-g^\prime state, -deps)
  satisfy_dependencies(-deps)
  g^\prime get_value(-g^\prime state, -value)
  if assigned value then
    return value
```

### 3.5 Lazy computations

Now that we have a way of representing computations with dependencies, we now define some ways of combining and modifying them to produce more complex computations, with dependency tracking still built-in.

For instance, `Compose` takes as input a getter `g` and a function `f`, and returns a getter `h` such that `Evaluate(h) = f(Evaluate(g))`.

Similarly, `ComposeProcedure` takes as input a getter `g` and a procedure `f`, and returns a getter `h` such that `Evaluate(h)` has the same side-effects as calling `f(Evaluate(g))`.

Similarly, `ComposeGetter` takes as input a getter `g` and a function `f` returning a getter, and returns a getter `h` such that `Evaluate(h) = Evaluate(f(Evaluate(g)))`.

For added convenience, these compose functions can take a sequence of getters instead a single getter. In this case, the arity of the function `f` must equal the length of the sequence. For example, `Evaluate(Compose([g1, g2], f)) = Evaluate(f(Evaluate(g1), Evaluate(g2)))`.

The intrinsic `Flatten` takes as input a sequence of getters, and returns the getter whose value is the sequence of values of the input getters.

There are also intrinsics for defining null getters, which do nothing, and for defining getters directly in terms of the `get_value` and `get_dependencies` functions. It is also possible to define getters directly whose dependencies are themselves getters, instead of `<x, n>` element-precision pairs.

The package also provides “lazy” versions of some intrinsics, which by convention are given the suffix `Lazy`, which returns a getter which when evaluated has the same side-effects and return value as the non-lazy version.

For example the following intrinsic returns a getter `g` such that `Evaluate(g)` has the same side-effects as calling `IncreaseAbsolutePrecision(x, n)` directly:

```plaintext
intrinsic IncreaseAbsolutePrecision_Lazy
  (x :: PadExactElt, n) -> Gettr
  if AbsolutePrecision(x) ge n then
    return NullGetter()
  else
    return x^\prime update(n diff AbsolutePrecision(x))
```
Indeed, one could now define the non-lazy version as

```plaintext
intrinsic IncreaseAbsolutePrecision(x :: PadExactElt, n)
    ignored := Evaluate(IncreaseAbsolutePrecision_Lazy(x,n))
```

Similarly, the intrinsic `Approximation_Lazy(x, n)` returns a getter whose value is an approximation of `x` to absolute precision `n`.

**Example 3.2.** We can now present an implementation of binary addition using these tools. Compare this with the earlier version (Example 2.1), presented in terms of the simplified representation where the update function was simply a procedure; now it is a getter, built using `ComposeProcedure` and `Approximation_Lazy` out of simpler getters.

```plaintext
intrinsic '+' (x :: FldPadExactElt, y :: FldPadExactElt)
    -> FldPadExactElt
    init := x`approximation + y`approximation
    mkupdate := function (z)
        return function (n)
            return ComposeProcedure(
                // lazily computes approximations to x and y to precision n
                [ Approximation_Lazy(x, n), Approximation_Lazy(y, n) ],
                // uses the approximations xx and yy to update
                // the value of z
                procedure (xx, yy)
                    Update(z, xx + yy)
            )
        return Parent(x) ! <init, mkupdate>
```

Most update functions in the package are defined in a similar fashion: firstly they lazily compute approximations to their inputs, and then they use these to update the value. The exceptions to this are mainly when the precision required of the input is not known immediately, and therefore some iteration is required; in such a circumstance, the getter returned by the update function usually needs to be defined directly in terms of its `get_dependencies` and `get_value` procedures.

### 4 ExactpAdics: Precision strategies

#### 4.1 Motivating example

Suppose we want to compute the valuation of a $p$-adic number $x$. If the number is not currently weakly zero, then this is straightforward: the valuation is the weak valuation. Otherwise, we can’t immediately deduce the valuation.
Therefore, we might try increasing the absolute precision of $x$. If the number is now not weakly zero, then we are done. Otherwise, we might increase the absolute precision further, and repeat this process for some time. At some point, if we don’t discover the answer, we might give up.

How do we choose the amount to increase the absolute precision by? How long do we go for before giving up? A simple answer might be to keep doubling the precision forever until we succeed, but this would never terminate if $x = 0$.

On the other hand, the user may want the process to definitely terminate after some amount of effort, and therefore give up after the precision has reached some limit. Or, knowing more about the inputs, it may be more appropriate to increase the precision linearly instead of exponentially, for example. We abstract away such decisions into a precision strategy.

4.2 Definition

A precision strategy is a strictly increasing sequence of non-negative integers. The sequence may be finite or infinite in length.

4.3 Representation

How such a sequence is represented is not too important. In the ExactpAdics package, a precision strategy is represented as one of the following:

- A single non-negative integer $n$, which is a strategy of length 1: $(n)$.
- A list of strategies, which is the concatenation of those strategies.
- A function $m$ which takes an integer and either returns true and a larger value, or returns false. It represents a sequence of integers as follows: let $n_0$ be the previous value in the strategy; for $i = 0, 1, \ldots$, if $m(n_i)$ is false, then terminate the sequence, otherwise it is true and also returns $n_{i+1}$, the next element of the sequence.
- A string, which is interpreted as the global strategy with that name (see below).
- A tuple <"limit",n> which limits the remaining strategy to $n$; that is, it will terminate when the strategy reaches $n$. More precisely, the first time the strategy outputs a number $m \geq n$, it instead outputs $n$ itself and then terminates.
- A tuple <"exp",e> which is equivalent to the function taking $n$ to $\lceil n^e \rceil$. It therefore represents an infinite sequence which grows exponentially.
- A tuple <"random"> which randomises the remaining strategy as follows: if at time $i$ the previous value was $n_i$ and the next value will be
\( n_{i+1} > n_i \), then we replace the next value with a uniform random element in \( \{ n_i + 1, \ldots, n_{i+1} \} \). This aims to dampen any potential issues arising from forcing precisions to come from a small set of values, such as powers of 2.

We provide the user with a global array of named strategies, with procedures to define and retrieve strategies with a particular name. Currently we define three named strategies by default:

- "defaultLimit": \(<"limit", 100>\), not a precision strategy in itself, but can be mixed in to other strategies to limit them.
- "unlimitedDefault": [1, \(<"randomize">, \(<"exp", 2>\)], starts at 1 and keeps doubling forever.
- "default": ["defaultLimit", "unlimitedDefault"], the same as the previous strategy, but with the default limit applied.

### 4.4 Usage and conventions

Any function which makes a non-canonical decision about how to control the precision of its inputs should take one or more precision strategies as parameters to make these decision.

By convention, these parameters all have the word Strategy in their name, to make their purpose clear. Where there are any precision strategies parameters, there will be one with the name precisely Strategy. Its value is used as the default value for the others. Its default value is the string "default", which refers to the global strategy with this name. For example:

```plaintext
intrinsic DoSomething( : Strategy := "default",
    Strategy1 := Strategy,
    Strategy2 := Strategy,
    Strategy2b := Strategy2)
...
```

This means that for the typical user, it suffices to set a global strategy with the name "default" and then forget about precision strategies. If it then turns out that some computations are raising precision errors, the user can consider altering the precision strategy.

Using a precision strategy whose maximum value is 100 is functionally very similar to using inexact \( p \)-adics to precision 100. The difference is that in the inexact case, all computations are done to this precision, whereas in the exact case, 100 is the worst case: if a computation can be done with less precision, then it will be.
4.5 Baseline precision

A function which takes a precision strategy parameter is free to use it in any fashion. It is intended, of course, that it will be interpreted as a sequence of precisions to try computations at, but there are different kinds of precision:

- ** Absolute precision:** the integer \( k \) such that we know the value modulo \( \pi^k \); that is, the approximation is of the form \( x + \pi^k \mathcal{O} \).

- ** Relative precision:** the absolute precision minus the weak valuation; that is, the integer \( r \) such that the approximation is of the form \( \pi^r (x + \pi^r \mathcal{O}) \).

If we interpret the entries of a precision strategy as absolute precisions, then it may be that the \( p \)-adic number actually has a large negative valuation, and so computing it to positive absolute precision is overkill. In a sense, the absolute precision is relative to the valuation 0, and 0 is an arbitrary choice.

If we interpret them as relative precisions, then we lose repeatability because the base-line for the relativeness can move: the weak valuation may increase in time. To demonstrate the issue, suppose we are computing the valuation of \( x = 0 \), initially known to absolute precision 10, and the precision strategy goes up to 100. Interpreting the strategy as relative precisions, we would increase the absolute precision of \( x \) to 110. If we try to compute the valuation again, then we will increase the absolute precision again to 210. There is the potential to keep increasing the absolute precision of \( x \) indefinitely by repeatedly trying to compute its valuation.

We introduce a new kind of precision:

- ** Baseline precision:** the absolute precision minus the “baseline valuation”.

The baseline valuation is any fixed valuation attached to the value. Hence it may depend on the value, but does not change over time. By default, the baseline valuation is set to the weak valuation of the value when it is initially created. If the baseline valuation is set to 0, then we recover the absolute precision.

If we now interpret entries of the precision strategy as baseline precisions, then we avoid the two problems described above.

**Example 4.1.** Hence, a reasonable implementation of a function to compute the valuation is

```
intrisic Valuation(x :: FldPadExctElt : Strategy:="default")
  for n in Strategy do
    IncreaseAbsolutePrecision(x, BaselineValuation(x) + n)
    if not IsWeaklyZero(x) then
      return WeakValuation(x)
  error "precision error"
```


24
5 ExactpAdics2: Core structures and elements

5.1 Overview
Recall that in the ExactpAdics package, our updates are performed in terms of absolute precisions: an update function receives an absolute precision, and updates the approximation accordingly. In order to do this, the update function must compute the absolute precisions required of its dependencies, which are fed into the dependency-tracking framework, and recursively the update functions of the dependencies themselves must compute the absolute precisions required of their dependencies.

In the ExactpAdics2 package, we simplify this procedure by introducing a proxy for absolute precision. This proxy is a single positive integer \( n \) which we refer to as the epoch. At any given time, a \( p \)-adic object has a current epoch meaning that its current approximation is associated to that epoch. The precision of the current approximation must increase with the epoch.

Importantly, by definition the approximation of a \( p \)-adic object at epoch \( n \) depends only on the approximations of its dependencies at epoch \( n \). Hence a \( p \)-adic object is represented by essentially two pieces of information: a list of the other \( p \)-adic objects on which it depends; and an approximation function which takes as input a list of approximations of its dependencies at some epoch \( n \), and returns an approximation which is taken to be the approximation of the object at the same epoch.

Since the approximation function is only given the approximations of its dependencies at a given epoch \( n \), all it must do is return an approximation to the best precision it can given its inputs. It is not aiming to return an approximation to any specific precision.

For example, our representation of \( \mathbb{Q}_p \) has no dependencies, and its representation at epoch \( n \) is the fixed-precision field \( \text{pAdicField}(p, 2^n) \) whose elements are of the form \( \pi^r(y + \pi^r \mathbb{Z}_p) \) for \( r \leq 2^n \). Hence the precision increases exponentially with \( n \), the intention being that one will, in a small number of epochs, be able to increase the precision of a \( p \)-adic object to some desired absolute precision. Since there are only a small number of possible epochs — the user is highly unlikely to go beyond \( n = 20 \) — the dependency-tracking framework should only be invoked relatively infrequently.

As we shall see in §5.6, the dependency-tracking itself is also quite straightforward.

5.2 Abstract base types
As with ExactpAdics, this package uses the abstract types StrPadExact and PadExactElt to represent \( p \)-adic structures (such a fields and rings) and elements respectively. However, these are now also both subtypes of AnyPadExact, which represent any \( p \)-adic object:
type AnyPadExact
attributes AnyPadExact: id, dependencies, approximations, get_approximation

type StrPadExact: AnyPadExact

type PadExactElt: AnyPadExact
attributes PadExactElt: parent

The id attribute, as before, is a unique integer used to identify the object. It is assigned from a global counter, and so each object can only depend on objects with smaller id. This is used to simplify dependency tracking.

The dependencies attribute is a list of other $p$-adic objects (i.e. of type AnyPadExact) on which this one directly depends.

The approximations attribute is a list of approximations of the object. The object at position $n$ in the list is the approximation of the object at epoch $n$. It is analogous to the approximation attribute in the Exact$\mathbb{p}$Adics package, except that we now record all approximations.

The get_approximation attribute is the approximation function, and is analogous to the update function from the Exact$\mathbb{p}$Adics package. It is a function with two inputs: an epoch (a positive integer) and the list of approximations of the dependencies at the given epoch. It must return the approximation of the object at the given epoch.

The parent attribute of an element (a PadExactElt) is the structure (a StrPadExact) containing the element. The approximation of an element at epoch $n$ must be an element of the approximation of the parent at epoch $n$.

Figure 4 illustrates the relationships between these types and their attributes.

5.3 $p$-adic fields

The way in which $p$-adic fields and their elements are built on top of these base types is identical to the Exact$\mathbb{p}$Adics package:

type FldPadExact[FldPadExactElt]: StrPadExact
attributes FldPadExact: xtype, prime, defining_polynomial

where the xtype is either PRIME indicating it is a prime $p$-adic field $\mathbb{Q}_p$, in which case the prime attribute must be set to the prime $p$, or else it is INERT or EISEN indicating an unramified or totally ramified extension, in which case the defining_polynomial attribute must be set to the inert or Eisenstein defining polynomial.

The approximations of a $p$-adic field must be fixed-precision inexact $p$-adic fields, such as pAdicField(2,20) in Magma (representing $\mathbb{Q}_2$, whose elements have relative precision at most 20). However, two such fields in
Magma are considered to be different, even if they only differ in their precision, and yet we will need to coerce approximate $p$-adic numbers between different approximate $p$-adic fields representing the same exact field, which will be manual and slow. Hence we define

attributes $FldPadExact$: infinite_precision_approximation

which is a semi-exact approximation of the field defined via a map, as described in §1.1. The approximations are then fixed-precision versions of this infinite-precision field produced via the $ChangePrecision$ intrinsic in Magma. Since Magma now understands all of these approximations to come from a common underlying field, it performs coercion between them for free.

5.4 Univariate polynomials

The way in which rings of univariate $p$-adic polynomials and their elements are defined is again identical to the ExactpAdics package:

type $RngUPol_FldPadExact[RngUPolElt_FldPadExact]$ attributes $RngUPol_FldPadExact$: base_ring

It is defined by its base_ring, a $p$-adic field (a $FldPadExact$). The approximation of such a ring at epoch $n$ is the univariate polynomial ring over the approximation at epoch $n$ of the base ring.
5.5 Examples

Example 5.1. Here we present an implementation of binary addition of $p$-adic numbers (cf. Example 3.2).

```plaintext
intrinsic '+' (x :: FldPadExactElt, y :: FldPadExactElt)
  -> FldPadExactElt
z := New(FldPadExactElt)
z`parent := x`parent
z`dependencies := [x, y]
z`get_approximation := function (n, xds)
  return xds[1] + xds[2]
return z
```

In fact, most purely arithmetic functions are this simple to implement.

Example 5.2. Here we present an implementation of polynomial resultant. Since the resultant depends on the degree of the polynomials, we need to ensure that the approximations of the inputs have the correct degree using EnsureAllApproximationsAreFullDegree, similar to as in Remark 5.6.

```plaintext
intrinsic Resultant (f :: RngUPolElt_FldPadExact, g :: RngUPolElt_FldPadExact)
  -> RngUPolElt_FldPadExact
EnsureAllApproximationsAreFullDegree(f)
EnsureAllApproximationsAreFullDegree(g)
h := New(RngUPolElt_FldPadExact)
h`parent := f`parent
h`dependencies := [f, g]
h`get_approximation := function (n, xds)
  return Resultant(xds[1], xds[2])
return h
```

5.6 Generating approximations

We now describe how we generate the approximations of a $p$-adic object from its dependencies and get_approximation function.

Suppose we are given a $p$-adic object and an epoch $n$, and we wish to compute the approximation of the object at the given epoch. The intrinsic BringToEpoch does this for us:

```plaintext
intrinsic BringToEpoch(x :: AnyPadExact, n :: RngIntElt)
  if #x`approximations lt n then
    for d in x`dependencies do
      BringToEpoch(d, n)
    xds := [d`approximations[n] : d in x`dependencies]
  ```
xx := x.get_approximation(n, xds)
x.approximations[n] := xx

First it checks if there is already an approximation at this epoch. If not, we run through the dependencies and bring these up to the same epoch recursively. Now we can construct a list of approximations of the dependencies at this epoch, pass this to get_approximation to produce the required approximation, and update the approximations list accordingly.

We also supply the intrinsic EpochApproximation which returns the approximation at a given epoch:

intrinsic EpochApproximation(x :: AnyPadExact, n :: RngIntElt)
        BringToEpoch(x, n)
        return x.approximations[n]

The true implementation of BringToEpoch is slightly more complicated. The following subsections explain how.

5.6.1 Saving the approximation

Instead of simply saving the output of get_approximation as a new approximation directly, we assign it using a generic intrinsic called SetApproximation, which performs some checks. This includes checking that the approximation is of the right type; that, if it is an element, the approximation is an element of the approximation of its parent; and that the approximation is consistent with the current best approximation attached to the object.

Furthermore, our package assumes that if an object has an approximation at epoch n, then it has approximations for all lower epochs. So what do we do if we are jumping from epoch 1 to 10, for example, how do we set the intermediate approximations? For each subtype of AnyPadExact, there must be an intrinsic InterpolateEpochs implemented which takes as input a p-adic object, a range of epochs, and an approximation at the top epoch. It must return a list of approximations for the intermediate epochs.

Example 5.3. The default implementation uses the approximation function to generate the intermediate values, provided we exceed the min_epoch. This is usually sufficient for structures. Note that the dependencies are guaranteed to be at the top epoch already.

intrinsic InterpolateEpochs(x :: AnyPadExact, n1 :: RngIntElt, n2 :: RngIntElt, xx :: FldPadElt)
        if n1 ge x.min_epoch then
            return [x.get_approximation(n, xds)
            where xds := [d.approximations[n] : d in x.dependencies] : n in [n1..n2-1]]
        else
            error "not implemented: InterpolateEpochs with min_epoch>1"
        end
Example 5.4. For \( p \)-adic numbers, we coerce the approximation into the approximations of the parent field at the intermediate epochs.

```plaintext
intrinsic InterpolateEpochs(x :: FldPadExactElt,
    n1 :: RngIntElt, n2 :: RngIntElt, xx :: FldPadElt)
return [x\texttt{parent}\text{'approximations}[n] ! xx : n in [n1..n2-1]]
```

Now if `SetApproximation` is setting some approximation for a high epoch, it will use `InterpolateEpochs` to fill in the gaps.

5.6.2 Minimum epoch

Example 5.5. Suppose we are implementing division of two \( p \)-adic numbers. Here is what looks like a reasonable implementation:

```plaintext
intrinsic '/' (x :: FldPadExactElt, y :: FldPadExactElt)
require IsDefinitelyNonzero(y)
z := New(FldPadExact)
z\texttt{parent} := x\texttt{parent}
z\texttt{dependencies} := [x, y]
z\texttt{get_approximation} := function (n, xds)
    return xds[1] / xds[2]
return z
```

Note, however, that even though we checked that \( y \) is non-zero, we are not guaranteed that all of its approximations are not weakly zero. If some of them are, then the division inside `get_approximation` may raise an error.

To solve this, we have `IsDefinitelyNonzero` return a second value, which is the smallest epoch at which an approximation for \( y \) is not weakly zero. Note that since approximations may not become less precise as epoch increases, this implies that all approximations for \( y \) are not weakly zero above this epoch. We can then set the new attribute

`attributes AnyPadExact: min_epoch`

to this epoch.

The meaning of `min_epoch` is that it is the smallest epoch for which the `get_approximation` function should be called, and hence in our example, the division will only use non weakly zero approximations to \( y \).

To use `min_epoch`, we simply need to insert the following line into `BringToEpoch`

`n := Max(n, x\texttt{min_epoch})`

which ensures that the epoch we are updating to is at least `min_epoch`. 
Remark 5.6. For the particular case of division, and similar functions, we can take a different approach and implement it like so:

```plaintext
intrinsic '/' (x :: FldPadExactElt, y :: FldPadExactElt)
  EnsureAllApproximationsAreNonzero(y)
  z := New(FldPadExactElt)
  z`parent = x`parent
  z`dependencies := [x, y]
  z`get_approximation := function (n, xds)
    return xds[1] / xds[2]
  return z
```

where, as the name suggests, the intrinsic `EnsureAllApproximationsAreNonzero` ensures that all approximations of `y` are not weakly zero. This is achieved by first calling `IsDefinitelyNonzero` to check `y` is nonzero and find an epoch at which its approximation is not weakly zero, and then by using `InterpolateEpochs` and `SetEpochs` to interpolate this approximation down to epoch 1.

5.6.3 Maximum epoch

Analogous to `min_epoch`, there is also

.attributes `AnyPadExact`: `max_epoch`

which is the maximum epoch at which `get_approximation` should be called.

The intention here is that the user can set the maximum epoch on a \( p \)-adic object as a way of limiting the precision to which computations involving that object are performed.

The `BringToEpoch` intrinsic is modified to insert a check that the target epoch is not greater than the `max_epoch`, if it is set. If so, it will raise a precision error:

```plaintext
if assigned x`max_epoch and n > x`max_epoch then
  error "precision error: max_epoch exceeded"
```

We also supply the intrinsic `CanBringToEpoch` which is the same as `BringToEpoch` except that instead of raising a precision error when the `max_epoch` is reached it returns false, and otherwise returns true to signal success.

5.7 Precision strategies

At present, we do not provide functionality analogous to the precision strategies (§4) of the `ExactAdics` package. The only method for controlling precision currently available to the user is the `max_epoch` attribute described in §5.6.3 which will cause an error to be raised if a computation requires too much precision.
Therefore currently, any functions which need to increase the precision of its inputs do so simply by trying each epoch in order. Hence, there are no Strategy parameters in this package, and instead we can think of the sequence $1, 2, \ldots$ as the default strategy where the values are now epochs, not precisions.

**Example 5.7.** Valuation is implemented like this (cf. Example 4.1):

```plaintext
intrinsic Valuation(x :: FldPadExactElt) -> Val_FldPadElt
for n in 1, 2, \ldots do
  BringToEpoch(x, n)
  if not IsWeaklyZero(x) then
    return WeakValuation(x)

◊
```

6 Comparison of ExactpAdics and ExactpAdics2

6.1 Complexity of updates

Compare the procedures `satisfy_dependencies` (§3.4) of ExactpAdics and `BringToEpoch` (§5.6) of ExactpAdics2, which are the underlying means in each package of generating an approximation to a $p$-adic object.

In the latter, we satisfy each dependency recursively immediately. In the former, we perform a backwards pass to gather all dependencies together, followed by a forwards pass to satisfy dependencies.

The rationale for the behaviour of the former was discussed in §3.1, and it comes down to the fact that the same $p$-adic object may appear multiple times in a dependency with different absolute precisions. By performing the backwards pass first, we can merge all such dependencies into one. On the other hand, in `BringToEpoch` all dependencies are being brought to the same epoch, and therefore we can satisfy each dependency immediately without risk of it needing to be brought to a higher epoch later.

Additionally, in ExactpAdics, satisfying a dependency is allowed to fail (i.e. the `get_value` procedure of a Getter is allowed to not return a value), which triggers a new backwards pass to find dependencies of this failed update, and an extra forwards pass will have to occur to satisfy these. Hence there is in principle no bound on the amount of dependency tracking required to update a single element, whereas in ExactpAdics we do a single pass.

This is a necessary feature of the design of ExactpAdics: because the update function must update its target object to a given absolute precision, we must allow it the freedom to take a guess at the precision required of its dependencies, and then try a better guess if it turns out this was too low. This is because there are some operations where it is difficult or impossible to determine the dependency precisions in advance.

On the other hand, in ExactpAdics2, because there is a looser relationship between precisions and epochs, the `get_approximation` function
is not aiming for any specific precision. Instead, it simply needs to produce an approximation to the best precision it can.

6.2 Number of updates

In ExactpAdics2, an “update” can occur at each epoch. Since precisions are exponential in the epoch, then typically there are only a small number of epochs ever considered, rarely going beyond epoch 20. This limits the number of times the dependency tracking framework ever needs to consider a single object, and so the time spent doing dependency tracking is essentially a small constant times the number of variables in a computation.

On the other hand, in ExactpAdics one can in principle increase the precision of an element by 1 many times, and each time the dependency tracking code will be invoked, so there is essentially no bound on the time spent doing this. To mitigate this, one could modify the package so that elements can only increase their precisions by large jumps, such as doubling each time.

6.3 Implementing new functions

To implement a new low-level operation in ExactpAdics, such as addition of two \( p \)-adic numbers, requires implementing a Getter which (a) can compute the precisions to which its dependencies are required; and (b) compute an approximation, given approximations of its dependencies. To do the same in ExactpAdics2 only requires (b), and therefore implementing new functionality in the latter is often much quicker.

Furthermore, actually computing the dependency precisions can be slow:

**Example 6.1.** Let \( h(x) = f(x)g(x) \) be a product of two polynomials. Suppose we want to compute an approximation to \( h \) with the absolute precision of the \( k \)th coefficient \( (h_k) \) at least \( a_k \). Then we need \( f_i \) to absolute precision \( \max_j a_{i+j} - \text{val}(g_j) \) and \( g_j \) to absolute precision \( \max_i a_{i+j} - \text{val}(f_i) \).

Computing these absolute precisions is of the same order of complexity as performing the multiplication itself. On the other hand, the multiplication is implemented in a low level compiled language such as C, whereas our ExactpAdics package is implemented in the high-level interpreted language Magma, and so computing these absolute precisions can be far more expensive.

6.4 Precision optimality

By design, all computations in ExactpAdics are performed to as little precision as is possible to get the answer. In ExactpAdics2, we perform all computations starting from the same initial precision and keep doubling this
precision as necessary. Hence the latter is not optimal in terms of precision used, but is typically within a factor of 2 of optimal.

It is possible for ExactpAdics2 to be worse than this. Suppose $x$ is cheap to compute approximations for, but loses a lot of precision along the way, so if it has an approximation in $\text{pAdicField}(p,2^{\ast}n)$ then its precision is significantly less than $2^n$. Also suppose that $y$ is expensive to compute, and does not lose any precision. Let $z = x + y$. Now because $x$ loses a lot of precision, so does $z$, and therefore computing an approximation to $z$ requires a relatively high epoch. Computing the approximation to $y$ at this epoch is expensive, but also unnecessary because it achieves this required precision at an earlier epoch.

In a sense, the epoch is not really a proxy for the precision of an element, but a proxy for the worst precision of all the dependencies of the element.

For most common applications, the amount of precision lost tends to be bounded and small as epoch increases, and so this effect is minimal.

6.5 Precomputing dependencies

In §9.1 we describe a generic optimization technique which can make updating the approximations of a selected $p$-adic object much quicker. This is done by pre-computing some of its dependency graph so that it can be traversed more efficiently.

This optimization opportunity is only possible in ExactpAdics2. Even if ExactpAdics were redesigned to make the list of dependencies of an object explicit, so that a piece of the dependency graph could be precomputed, we would still need to do a backwards pass to find the minimal precision required of each dependency.

6.6 Precision strategies

In ExactpAdics, whenever a function needs to increase the precision of an object in a non-canonical way, it does so according to a precision strategy (§4), giving fine control over each precision tried.

On the other hand, ExactpAdics2 currently has no such functionality other than setting the max_epoch parameter on an object (§5.7). When a function needs to increase the precision of an object in a non-canonical way, it repeatedly increases the epoch by 1. In practice, precision strategies in ExactpAdics will usually just repeatedly double the precision, which behaviour is almost the same as increasing the epoch by 1 in ExactpAdics2. In principle, the package could have strategies to control which epochs are used, but this is not yet implemented.
| Experiment                      | Time (sec) |
|--------------------------------|------------|
| (1) Builtin                    | 0.949      |
| (2) ExactpAdics                | 174.284 = 6.907 + 167.377 |
| (3) ExactpAdics2               | 7.330 = 0.400 + 6.930 |
| (4) ExactpAdics2 (opt: default)| 7.047 = 0.528 + 6.519 |
| (5) ExactpAdics2 (opt: fast)   | 2.006 = 0.460 + 1.546 |

Table 1: Timings for a highly dependent computation over different implementations, including two optimizations.

6.7 Timings

6.7.1 Dependency tracking

In this section, we describe an experiment designed to stretch the dependency tracking capabilities of our packages. This involves performing a computation which involves thousands of intermediate variables, but the steps themselves are cheap to compute.

In this experiment, we define $x_1 = 1, x_2 = 2 \in \mathbb{Q}_2$ and for $i = 3, \ldots, 10000$ we define $x_i = x_{j_i} + x_{k_i}$ for some randomly chosen $j_i, k_i \in \{1, 2, \ldots, i - 1\}$. Finally, we define $y = \sum_{i=1}^{10000} x_i$. We time how long it takes to compute $y$ to absolute precision $2^n$ for $n = 1, \ldots, 16$, taking the total time — this emulates a typical sequence of increasing the absolute precision of $y$ 16 times. We repeat this 10 times with different random choices and take the mean.

This experiment is repeated using a number of different $p$-adic implementations, with the mean timings given in Table 1. Note that the random choices are made in advance and so are not timed, and we use the same random seed in each experiment, so precisely the same sequence of operations is being compared.

Experiment (1) uses the builtin inexact $p$-adics available in Magma, and so is a reasonable lower bound on what we can expect to achieve. Experiments (2) and (3) use the ExactpAdics and ExactpAdics2 packages, respectively. These timings are broken into two parts, the first part being the time to construct $y$, and the second part being the time to increase its precision to $2, 4, \ldots, 2^{16}$. We can see that the latter package outperforms the former significantly on both counts.

Experiments (4) and (5) are the same as (3), except we use the optimization techniques described in §9.1 to make $y$ directly depend only on $x_1$ and $x_2$. Experiment (4) uses the default version, which gives a small speed-up. Experiment (5) uses the “fast” version, which forgets the intermediate variables and uses the get_approximation functions directly, and achieves a significant speed-up.
6.7.2 Real-world example

We compute the 2-part of the conductor of the hyperelliptic curve

\[ C : y^2 = -2x^6 - 15x^4 - 37x^2 - 30 \]

using our implementation of \[3\] mentioned in §1. This implementation can use either of our packages for its underlying \( p \)-adic computations. The curve has discriminant \( \Delta = -2^{16} \cdot 3 \cdot 5 \) and conductor \( N = 2^{10} \cdot 3 \cdot 5 \).

When using \texttt{ExactpAdics}, this takes 146 seconds, compared to 33 seconds for \texttt{ExactpAdics2}. The time spent in the dependency-tracking portion of code, which includes actually computing approximations, is 124 and 25 seconds respectively, with 22 and 12 seconds respectively left over to other computations.

With \texttt{ExactpAdics}, this 124 seconds spent in dependency tracking is divided equally between generating approximations and tracking dependencies (this includes calling the update function and computing dependencies). About half of the latter is spent computing dependencies, most of the rest being logic comparing absolute precisions.

In fact, of the whole 146 seconds, 39 seconds is spent just constructing our representation of a valuation of a univariate polynomial. Individually this is fast, but we construct 240,000 of them throughout the algorithm. This demonstrates the benefit of using epochs instead of fine absolute precisions in \texttt{ExactpAdics2}.

Note that the number of times the dependency tracking framework is invoked is about 54,000 and 40,000 for the two packages. Given the same algorithms are used in both packages, we expect these numbers to be similar. In this case we do not appear to have the potential issue that the framework is invoked too often. The number of \( p \)-adic objects created is about 11,000 and 7,000 for the two packages.

6.8 Conclusions

Given the above arguments and evidence, we currently recommend the typical user to choose \texttt{ExactpAdics2} over \texttt{ExactpAdics}.

On the other hand, if more of the internal workings were implemented at a lower level than the Magma language and optimized, then it may be that \texttt{ExactpAdics} could be made comparably fast. Indeed, much of the comparative slowness in \texttt{ExactpAdics} comes from the need for a lot of simple arithmetic to compute absolute precisions, which is typically slow in an interpreted language such as Magma.
7 Additional structures

So far we have described the representation of \( p \)-adic numbers and univariate polynomials over \( p \)-adic fields. We now briefly describe two more structures provided by the package.

7.1 Multivariate polynomials

A multivariate polynomial ring over a \( p \)-adic field is represented by the type \texttt{RngMPol\_FldPadExact} (analogous to the inexact type \texttt{RngMPol\_Elt} in Magma) which derives from \texttt{StrPadExact}:

\begin{verbatim}
  type RngMPol\_FldPadExact[RngMPol\_Elt\_FldPadExact]
  attributes RngMPol\_FldPadExact: base\_ring, rank

  Such a ring is defined by its base\_ring, an exact \( p \)-adic field (i.e. of type \texttt{FldPadExact}), and by its rank, the number of indeterminates.

  An approximation of such a ring must be the multivariate \texttt{PolynomialRing} of an approximation of the base\_ring of the same rank.
\end{verbatim}

7.2 Cartesian products

The cartesian product of a number of exact \( p \)-adic structures is itself an exact \( p \)-adic structure, and has the type \texttt{SetCart\_PadExact} analogous to the type \texttt{SetCart} for general cartesian products.

\begin{verbatim}
  type SetCart\_PadExact[Tup\_PadExact]
  attributes SetCart\_PadExact: components

  Such a cartesian product is defined by its components, a list of exact \( p \)-adic structures.

  An approximation of this structure must be the cartesian product (a \texttt{SetCart}) of an approximations of its components.

  Why do we define this specialised form of cartesian products, when a general one exists already? The difference is that a standard tuple of exact \( p \)-adic values treats the component values as completely independent objects, whereas the exact tuple links them together in the sense that they have a single common update/approximation function. Therefore, the exact tuple is an appropriate choice for a collection of \( p \)-adic values which belong to some conceptually higher structure.
\end{verbatim}

Example 7.1. Suppose we wish to implement a Hensel-lifting routine which takes as input a sequence \( F \in K[x_1, \ldots, x_n]^n \) of \( n \) multivariate polynomials of rank \( n \) over some \( p \)-adic field \( K \) and a sequence \( X \in K^n \) of \( n \) elements of \( K \) such that we can apply Hensel’s lemma to deduce there is a root \( Y \in K^n \) of \( F \) close to \( X \), and returns the sequence \( Y \) (as in §9.9).

To update the components of \( Y \) we perform a Hensel-lifting routine which is essentially some \( n \times n \) linear algebra depending on \( F \) and \( X \), the
important point being that this computes all components of \( Y \) to some precision simultaneously; it is not possible to compute one component of \( Y \) to high precision in isolation. Therefore it makes sense to represent \( Y \) as a tuple with a single update function.

By comparison, if we represented \( Y \) as a sequence of independent values, then each component would still need to maintain its own approximation to the whole vector \( Y \) in order to perform Hensel lifting. Worse still, increasing the precision on one component would perform Hensel lifting, but then only update that one component even though the information is available to update all components. Therefore, increasing the precision of all components of \( Y \) would be \( n \) times too slow.

8 Valuations

In our packages, the valuation of a \( p \)-adic element \texttt{PadExactElt} is intended to be the finest measure available of the valuation of the components of the element. Because there are many different types of \( p \)-adic elements (e.g. numbers, polynomials, tuples), there are as many different types of valuations, all needing to be represented somehow. There are some operations common to all valuations, such as addition, so we define a new abstract type to represent all types of valuation:

\[
\texttt{type Val\_PadExactElt}
\]

and we shall later define sub-types corresponding to each \( p \)-adic structure.

Note that the difference of two valuations is also a valuation, corresponding to the division of two \( p \)-adic elements with those valuations. Therefore, all kinds of precisions — absolute, relative and baseline — are also valuations.

In the packages, we use valuations of subtype of \texttt{Val\_PadExactElt} to represent all valuations (including weak valuations) and all precisions. In particular, the input to an update function is a valuation in this form, representing the intended absolute precision.

\[
\texttt{attributes Val\_PadExactElt: value}
\]

The \texttt{value} field of a valuation contains the actual value of the valuation, whose representation is element-dependent.

8.1 Valuations of \( p \)-adic numbers

Ordinarily we think of the valuation of a \( p \)-adic number as an integer, except that:

- The valuation of zero is not an integer, it takes the special value \( \infty \).
• Multiplication of two valuations is not a useful concept: it has no description in terms of the arithmetic of $p$-adic numbers.

• On the other hand addition and infimum do make sense: the sum of two valuations corresponds to the multiplication of $p$-adic numbers, and the minimum of two valuations corresponds to the addition of $p$-adic numbers.

• Multiplication and division of a valuation by an integer or rational number also does make sense, since it corresponds to exponentiation of a $p$-adic number.

• It is useful to be able to talk about the valuation of elements in an extension, and these may be rational numbers.

• Subtraction is useful to define, since a relative or baseline precision is the difference of two valuations. In particular, we also need to include the symbol $-\infty := 0 - \infty$.

• Supremum is also a useful operation: if we increase the absolute precision of a $p$-adic number several times, then its final absolute value is the maximum of the intermediate absolute precisions.

We deduce that the standard ring of integers $(\mathbb{Z}, +, \times)$ is not a useful structure for valuations to reside in; instead, we define the set $\mathbb{Z} := \mathbb{Q} \cup \{\pm\infty\}$, elements of which we represent with the type:

```markdown
**type** Val_FldPadElt: Val_PadExactElt
```

The `value` attribute is either an integer (a `RngIntElt` in Magma), a rational number (a `FldRatElt`) or $\pm\infty$ (a `Infty`).

The following operations are supported:

• Addition: Defined for all pairs of elements of $\mathbb{Z}$, except $\infty + (-\infty)$ is left undefined and will cause an error.

• Subtraction: Defined for all pairs of elements of $\mathbb{Z}$. In particular, $\infty - \infty$ is defined to be 0; this is because the $p$-adic number 0 represented to infinite $p$-adic absolute precision has infinite weak valuation, and so $\infty - \infty$ should be its relative precision, which is 0. While an arbitrary collection of additions and subtractions is not associative by these definitions, in practice if subtraction is only used to compute precisions, then the results will be well-defined.

• Infemum and supremum (which are the operations `meet` and `join` in the Magma language).

• Multiplication and division by rational numbers (which we term *scaling*).
• Equality, inequality, and orderings =, ≠, ≤, <, ≥, >. In particular, \( \mathbb{Z} \) is totally ordered.

• An operation called \( \text{diff} \) which is defined as follows: \( x \text{ diff } y \) is \( x \) if \( x > y \) and otherwise is \( -\infty \). Note that it is the lowest valuation \( z \) such that \( z \Join y = x \Join y \). It has a natural interpretation in our context: if we require an element to have precision \( x \) and its current precision is \( y \) then \( x \text{ diff } y \) is the lowest valuation \( z \) such that increasing the precision to \( z \) suffices. Whilst defining such an operation for single \( p \)-adic numbers may seem like overkill, it turns out to be useful for aggregates.

8.2 Valuations of aggregate structures

All other \( p \)-adic structures in the package are aggregate structures, in the sense that they represent, perhaps recursively, a collection of \( p \)-adic numbers. As defined at the top of the section, a valuation in the package is the finest possible description of the components of a \( p \)-adic element, and therefore we represent valuations of an aggregate as an analogous aggregate of valuations. Specifically:

• Univariate polynomials: A polynomial \( f(x) = \sum_{i=0}^{\infty} f_i x^i \in K[x] \) over a \( p \)-adic field \( K \) may be more simply thought of as the infinite sequence \((f_0, f_1, \ldots)\) of its coefficients, which is zero for all but finitely many places. Correspondingly, its valuation we represent as the infinite sequence \((\text{val}(f_0), \text{val}(f_1), \ldots)\), which is \( \infty \) at all but finitely many places. If we subtract two such valuations pointwise, the result is an infinite sequence which is 0 at all but finitely many places. Most generally then, a valuation of a univariate polynomial is an infinite sequence which takes the same value at all but finitely many places.

In the package, we define the new type \( \text{AssocDflt} \) which represents an associative array with a default value; that is, it has a default value so that if a key is not in the array, then the value of the array at that key is the default. These are useful for representing functions which are constant at all but finitely many places.

Valuations of univariate polynomials are represented by the type:

\[
\text{type } \text{Val_RngUPolElt_FldPad}: \text{Val_PadExactElt}
\]

whose \text{value} is a default associative array \( \text{AssocDflt} \) whose keys are non-negative integers \( i \) and whose values are \( \text{Val_FldPadElt} \)s.

• Multivariate polynomials: A polynomial

\[
f(x_1, \ldots, x_r) = \sum_{e \in \{0,1,\ldots\}^r} f_e x_1^{e_1} \cdots x_r^{e_r} \in K[x_1, \ldots, x_r]
\]
of rank $r$ over $K$ can be thought of as the map $e \mapsto f_e$ taking exponent vectors to the corresponding coefficient. As with univariate polynomials, this map is zero almost everywhere. In analogue with univariate polynomials, we represent the valuation of a multivariate polynomial with the type:

```plaintext
type Val_RngMPolElt_FldPad: Val_PadExactElt
```

whose value is a default associative array $\text{AssocDflt}$ whose keys are exponent vectors $e$ and whose values are the corresponding $\text{Val_FldPadElt}$s.

- Tuples: Valuations of tuples $\text{Tup_PadExactElt}$ of exact $p$-adic elements are represented by the type:

```plaintext
type Val_Tup_PadExactElt: Val_PadExactElt
```

whose value is a corresponding tuple of valuations, representing the valuations of the components of the tuple.

These valuations all support the following operations:

- Addition, subtraction, scaling, infimum ($\text{meet}$), supremum ($\text{join}$), $\text{diff}$: These are all defined point-wise.
- Equality and inequality: two valuations are equal iff they are equal point-wise.
- Ordering: two valuations are ordered if that ordering applies point-wise.

Note that while the set $Z = \mathbb{Q} \cup \{\pm \infty\}$ of valuations for $p$-adic numbers is totally ordered — and this ordering is respected by the ordering, infimum and supremum operations — the valuations for aggregate $p$-adic elements are only partially ordered. For example two tuples in $\mathbb{Q}_p^2$ may have valuations $(1, 2)$ and $(2, 1)$ and so are not ordered relative to each other, or they may have valuations $(1, 2) < (2, 2)$. This partial ordering is respected by infimum and supremum; for example $x \text{ join } y$ is the unique smallest valuation greater than or equal to both $x$ and $y$.

**Example 8.1.** Suppose a univariate polynomial of degree 5 is known to absolute precision $x = (3, 5, 8, 10, 13, 2, \infty, \infty, \ldots)$. The infinite precisions indicate that we know that coefficients 6 upwards are precisely zero. Also suppose we want to increase its absolute precision to at least $y = (10, 10, \ldots)$. Then it suffices to increase it to

```plaintext
y \text{ diff } x = (10 \text{ diff } 3, 10 \text{ diff } 5, 10 \text{ diff } 8, 10 \text{ diff } 10, 10 \text{ diff } 13, 10 \text{ diff } 2, 10 \text{ diff } \infty, \ldots) = (10, 10, 10, -\infty, -\infty, 10, -\infty, \ldots)
```

and so we see it suffices to only increase the precisions of coefficients 0, 1, 2 and 5.  

\[\diamond\]
9 Additional features

We now describe some of the high-level features available in our packages, including notes on how they are implemented. The majority of these features are in both packages, but any pseudo-code in this section will be as in ExactpAdics2.

9.1 Precomputing dependencies

Suppose \( d = (d_1, \ldots, d_k) \) are \( p \)-adic objects, and \( x \) is some complicated expression in \( d \), such as in §6.7.1. Hence \( x \) does not depend directly on \( d \), it depends on some intermediate expressions which recursively ultimately depend on just \( d \). To compute an approximation for \( x \) requires traversing its dependency graph, including all these intermediate expressions, which will be time-consuming. If we do not care about the intermediate expressions, then it could be more efficient to compute approximations to \( x \) directly from \( d \).

In the ExactpAdics2 package, we provide an intrinsic \texttt{WithDependencies} which takes a \( p \)-adic object \( x \) and a list \( d \) of other \( p \)-adic objects and returns a copy of \( x \) whose direct dependencies are precisely \( d \). The basic idea is that we pre-compute the piece of the dependency graph between \( x \) and \( d \), which the \texttt{get_approximation} function can traverse efficiently.

Specifically, starting from \( x \), we recursively traverse its dependencies, gathering them together to form the set of all of its dependencies. Whenever we reach a dependency lying in \( d \), we terminate that branch of the recursion, so that we only find the dependencies between \( x \) and \( d \). Next, we sort these dependencies by \texttt{id} into a list. Since the \texttt{ids} are assigned sequentially, this also sorts according to dependency.

With its default behaviour, \texttt{WithDependencies} also incorporates information about the \texttt{min_epoch} of each dependency into this list: specifically the list is now a list of pairs \((y, m)\) where \( m \) is the maximum \texttt{min_epoch} of \( y \) or anything depending on \( y \). Having precomputed this list, we can define \texttt{get_approximation} to traverse this list in order: given an epoch \( n \), for each \((y, m)\) in the list, we compute an approximation to \( y \) at epoch \( \max(n, m) \) from its dependencies, which will already be at this epoch, and update \( y \) accordingly.

\texttt{WithDependencies} also has a \texttt{Fast} parameter which performs a more aggressive optimization. Note that the default behaviour still explicitly deals with all intermediate dependencies \((y, m)\), and in particular each such \( y \) is updated in the usual manner, which involves a number of consistency checks. The “fast” version ultimately forgets these dependencies entirely and instead just remembers the \texttt{get_approximation} function attached to each one. These are called directly, one by one, with the approximations they return just appended to a temporary list, which is used as input when calling
the next one, and so on. The last item in this list will be the approximation to \(x\) returned by \texttt{get\_approximation}. As a result, there is no caching or consistency checking of each intermediate approximation, which can be a significant speed-up. The final answer, which is used to update \(x\), is still checked in the usual manner so we do not lose any safety.

Note that because the “fast” algorithm does not allow caching of intermediate variables, the \texttt{min\_epoch} of the created object must be the maximum of the \texttt{min\_epochs} of all dependencies. Similarly its \texttt{max\_epoch} must be the minimum of those of its dependencies.

Furthermore, the “fast” algorithm assumes that the approximations of the intermediate variables are all as produced by \texttt{get\_approximation}. Therefore division, which can change the approximations of its dependencies (Remark 5.6), should not be an intermediate expression. For this reason, the \texttt{Fast} parameter is false by default because it is not guaranteed to be safe. We also give division (and other intrinsics with the same issue) a \texttt{Safe} parameter which, when true, does not use this trick and is therefore safe to be an intermediate expression, at the cost of a potentially higher \texttt{min\_epoch}.

Timings demonstrating the benefits of using this optimization technique are given in §6.7.1. As noted in §6.5, this generic optimization is not possible in \texttt{Exact\pAdics}.

9.2 Valuation comparison

A common \(p\)-adic operation is to compare the valuation of a \(p\)-adic number \(x\) with some given valuation \(v\). Consider the following code:

\[
\text{if Valuation}(x) \gt 10 \text{ then} \nonumber \\
\ldots 
\]

The first thing this does is compute the valuation of \(x\) precisely, and then compare the answer with 10. However, this is overkill: since there is no canonical way to increase the precision of \(x\) in order to find its valuation (which may be very high), then \texttt{Valuation} will proceed according to some precision strategy, and therefore could never return an answer, or could raise a precision error.

We provide the following intrinsic:

\[
\text{intrinsic ValuationGe}(x, n) \\
\text{IncreaseAbsolutePrecision}(x, n) \\
\text{return WeakValuation}(x) \ge n 
\]

so that \texttt{ValuationGe}(x, n) is functionally very similar to \texttt{Valuation}(x) \texttt{ ge} n except that now there is a canonical way to increase the precision of \(x\) in order to get the answer, and it is guaranteed to produce a result with as little precision as required.
In reality, the definition of \texttt{ValuationGe} is made a little more complex by checking if the answer is already known without increasing the precision of \( x \).

We similarly provide analogues \texttt{ValuationEq, ValuationNe, ValuationLt, ValuationGe} and \texttt{ValuationGt} for the other comparison operators.

9.3 Residue class fields and higher quotients

Since Magma's inexact \( p \)-adics includes some functionality around residue class fields, we make similar functionality available in our package.

The intrinsic function \texttt{ResidueClassField} takes as input an exact \( p \)-adic field \( K \) (type FldPadExact) and returns its residue class field \( \mathbb{F} \) (type FldFin) and the quotient map \( q : \mathcal{O} \to \mathbb{F} \).

This is implemented by computing the residue class field of the \textit{approximation} field of \( K \) (i.e. \texttt{ResidueClassField(K'approximation)}), which returns \( \mathbb{F} \) and the quotient map \( \tilde{q} : \tilde{\mathcal{O}} \to \mathbb{F} \) where \( \tilde{\mathcal{O}} \) is the integer ring of the \textit{approximation} field. Then \( q \) may be defined in terms of \( \tilde{q} \): given \( x \in K \), increase the absolute precision of \( x \) to at least 1, and then call \( \tilde{q}(\tilde{x}) \).

The quotient map \( \tilde{q} \) also comes with a partial inverse, an embedding \( \tilde{q}^{-1} : \mathbb{F} \hookrightarrow \tilde{\mathcal{O}} \), which we similarly extend to a partial inverse \( q^{-1} : \mathbb{F} \hookrightarrow \mathcal{O} \). In this case, \( q^{-1}(x) \) is always given to absolute precision 1, and cannot have its absolute precision increased; in a sense, it refuses to choose among the many possible pre-images. In order to force such a choice, the intrinsic \texttt{WeakApproximation} is provided, which takes as input an exact \( p \)-adic number, and returns another exact \( p \)-adic number which is equal to the input up to the precision of the input.

In a completely analogous manner, the intrinsic \texttt{Quotient(K, n)} returns the ring \( \mathcal{O}/\pi^n\mathcal{O} \) and the quotient map \( q \), which again has a partial inverse. Hence \texttt{Quotient(K, 1)} and \texttt{ResidueClassField(K)} are equivalent, except that the latter represents the result as a field, and not a more general ring.

9.4 Completions of number fields

Magma's inexact \( p \)-adics includes some functionality around taking completions of number fields at finite primes, so we make similar functionality available in our package.

The procedure \texttt{ExactCompletion} takes as input a number field \( F \) and a finite place \( p \) of \( F \), and returns the completion \( \hat{K} := F_p \) as an exact \( p \)-adic field, and the embedding map \( e : F \hookrightarrow \hat{K} \).

This is implemented around the builtin intrinsic \texttt{Completion} which takes the same inputs, and returns the completion \( \hat{K} \) as a semi-exact \( p \)-adic field, and the embedding \( \hat{e} : F \hookrightarrow \hat{K} \). Then \( K \) is simply an exact \( p \)-adic field whose approximation is \( \hat{K} \), and \( e : F \hookrightarrow K \) returns an element whose update
function uses \( \tilde{e} \) to embed the input element of \( F \) into \( \tilde{K} \) to sufficiently high precision.

### 9.5 Newton polygons

The following definitions and results are standard, if not the notation.

**Definition 9.1.** If \( f(x) = \sum_{i=0}^{d} f_i x^i \in K[x] \) is a polynomial over a \( p \)-adic field \( K \), then its Newton polygon \( \mathcal{N}(f) \) is the lower convex hull in \( \mathbb{Q} \times \mathbb{Q} \) of the points \( (i, \text{val}(f_i)) \). It can also be interpreted as the graph of a function \([0, d] \to \mathbb{Q} \), also denoted by \( \mathcal{N}(f) \). By definition, this function is continuous, convex and piece-wise linear. If \( F \) is a face of the Newton polygon, i.e. a line segment from \((i_0, v_0)\) to \((i_1, v_1)\), then its width is \( w(F) = w = i_1 - i_0 \) and its slope is \( s(F) = \frac{v_1 - v_0}{i_1 - i_0} \). Writing \( s(F) = -\frac{h}{e} \) in lowest terms, then the \textit{ramification degree} of the face is \( e(F) = e \), and the \textit{residual polynomial} is\( r(F)(x) = \sum_{i=0}^{w/e} f_i \pi^{ih} \). Writing \( -s(F) = \frac{h}{e} \) in lowest terms, if \( r \) is such a root, then \( r^n \pi^{-h} \) has valuation 0 and \( r(F)(r^n \pi^{-h}) = 0 \). Furthermore the roots \( r \) of valuation \( h/e \) are in \( e \)-to-1 correspondence with roots (possibly repeated) of \( r(F)(x) \) via \( r \mapsto \frac{r^n \pi^{-h}}{\tilde{e}} \).

Hence the Newton polygon and related quantities provide much information about the roots of a polynomial, and so are invaluable in scenarios such as root-finding or factorization of polynomials.

We provide an intrinsic

```plaintext
intrinsic NewtonPolygon(f :: RngUPolElt_FldPadExact : Support:=<0,Degree(f)>)
  -> NwtnPgon
```

which takes as input a \( p \)-adic polynomial \( f \) and returns its Newton polygon. Since computing this involves computing the valuations of some of its coefficients, which may initially be weakly zero, it takes a \texttt{Strategy} parameter. It also takes a \texttt{Support} parameter which is a pair of integers representing a range, and the returned value will be a sub-polygon of the full Newton polygon supported on at least this range; this can be useful if, for example, the polygon might have a single root at 0, and so it suffices to get the piece of the Newton polygon on \([1, \infty)\).

The Newton polygon is computed as follows. We loop through precisions in the \texttt{Strategy} and for each one, compute a corresponding approximation \( x_f \) of \( f \). We compute the lower weak Newton polygon of \( x_f \), defined to be the lower convex hull of the points \((i, w_i)\) where \( w_i \) is the weak valuation of the \( i \)th coefficient of \( x_f \). We also compute the upper weak Newton polygon of \( x_f \), defined to be the lower convex hull of the points \((i, w_i)\) such that the
Figure 5: Computation of a section of a Newton polygon (heavy line) from lower and upper weak Newton polygons. Circles indicate the weak valuations of weakly zero coefficients, crosses indicate valuations of non weakly zero coefficients. Observe that since each end of the leftmost piece of the Newton polygon is at a vertex of the lower polygon, then these must also be vertices of the Newton polygon; contrast with the rightmost piece, in which the face could extend further to the left.

ith coefficient of $xf$ is not weakly zero, and therefore $w_i = \text{val}(f_i)$. The lower weak Newton polygon lies below the Newton polygon, which in turn lies below the upper weak Newton polygon. Therefore if the weak polygons overlap anywhere, then that overlap is a section of the Newton polygon (see Figure 5). If this section includes all of the Support then we are done, otherwise we move on to the next precision in the strategy.

9.6 Ramification polygons and transition functions

The ramification filtration of $\text{Gal}(L/K)$, the Hasse-Herbrand transition function and the upper-numbering of ramification groups are all standard, and appear for instance in Serre [9, Ch. IV]. The theory extends to non-Galois extensions [7], which we summarise now.

**Definition 9.3.** Given a finite extension $L/K$ of $p$-adic fields, its Galois set $\Gamma(L/K)$ is the set of $K$-embeddings of $L$ into a normal closure — this is a generalization of the Galois group. For $\sigma \in \Gamma$, we define $\text{val}(\sigma) := \min_{x \in \mathcal{O}_L} \text{val}_L(\sigma x - x)$ and $\Gamma_v := \{\sigma \in \Gamma : \text{val}(\sigma) \geq v\}$ for $v \geq 0$. The (lower) ramification breaks of $L/K$ are the $v$ at which the function $v \mapsto |\Gamma_v|$ is discontinuous. We define the transition function

$$\phi_{L/K}(v) = \frac{1}{e(L/K)} \int_0^v |\Gamma_t| \, dt$$

which is continuous, piecewise linear, increasing and hence bijective $[0, \infty) \to$.
[0, \infty), and letting \psi_{L/K} be its inverse, we define \Gamma^u = \Gamma_{\psi(u)}. This defines the upper ramification numbering. We define \(L^u = L_v\) to be the fixed field of \(\Gamma^u = \Gamma_v\) (where \(u = \phi(v)\)).

The following lemma summarizes some key aspects of the theory. In particular, the upper numbering is well-behaved under changing the top field and fixing the base field, much in the way that the lower numbering is well-behaved under changing the base field. It also shows that the Galois correspondence generalizes to the sets \(\Gamma^u\).

**Lemma 9.4 ([7] Prop. 2, Rmk. 3, Prop. 3).**

(a) If \(M/L/K\) then \(\phi_{M/K} = \phi_{L/K} \circ \phi_{M/L}\).

(b) Also \(\Gamma^u_{L/K} = \\{\sigma|_L : \sigma \in \Gamma^u_{M/K}\}\), and in particular \(\Gamma^u_{L/K}\) are restrictions of elements of \(\text{Gal}(L/K)^u\).

(c) \(L : L^u = |\Gamma^u|\) and so in particular \(L^u\) is the subfield of \(L\) fixed by \(\text{Gal}(L/K)^u\).

Computing quantities such as the transition function and upper/lower ramification breaks of an extension \(L/K\) is therefore of use when considering the Galois action of inertia or higher ramification groups. To compute these, we use ramification polygons, detailed decriptions of which appear in [5, §4–5] and [8, §3]. We summarize the key points here.

**Definition 9.5.** Suppose \(U/K\) is unramified, degree \(d\), \(f(x) \in U[x]\) is Eisenstein degree \(e\), defining the totally ramified \(L/U\), with uniformizer \(\pi \in L\) such that \(f(\pi) = 0\). Then the ramification polygon of \(L/K\) is the Newton polygon of the polynomial \(f(x + \pi)\) (which is supported on \([1, e]\)) with an additional horizontal face supported on \([e, ed]\).

**Lemma 9.6.** The lower ramification breaks of \(L/K\) are \(v\) where \(-v\) is a slope of a face of the ramification polygon. The corresponding \(|\Gamma_v|\) is the abscissa of the right hand vertex of the corresponding face. Letting \(v_0 = 0 < \ldots < v_t\) be the lower breaks in sorted order and \(s_i = |\Gamma_{v_i}|\), and letting \(u_0 = 0 < \ldots < u_t\) be the upper breaks (i.e. \(u_i = \phi_{L/K}(v_i)\)) then

\[
\frac{u_{i+1} - u_i}{v_{i+1} - v_i} = \frac{s_i}{e(L/K)}
\]

gives a means to compute any one of these three sequences from the other two.

**Proof.** Since \(\mathcal{O}_L = \mathcal{O}_U[\pi]\), for \(\sigma \in \Gamma(L/U) = \Gamma(L/K)\) we have \(\text{val}(\sigma) = \text{val}(\sigma(\pi) - \pi) > 0\). Now \(\sigma(\pi) - \pi\) are precisely the roots of \(f(x - \pi)\), and so by Lemma 9.2 their valuations correspond to faces of the ramification polygon. Specifically, if \(-v\) is the slope of the face and \(w\) its width, then there are \(w\) elements \(\sigma \in \Gamma(L/U)\) such that \(\text{val}(\sigma) = v\). Accumulating these widths from the left gives the sizes of \(\Gamma_v\), as claimed, for \(v > 0\). The extra
horizontal face by construction has slope 0 and vertex at \( ce = (L : K) = |\Gamma| = |\Gamma_0| \). The formula relating \( v, s, u \) follows from the definition of \( \phi_{L/K} \) as an integral.

Hence the slopes and abscissa of vertices of faces of the Newton polygon correspond to \((v_i, s_i)\) and the vertices of the transition function correspond to \((v_i, u_i)\), and there is a bijective correspondence between these sequences. Therefore we can compute transition functions from Newton polygons and vice versa, provided we represent the transition function by its vertices. We introduce a new type to do so:

```hasseherbtransfunc```

It is easy to evaluate the transition function at a given \( v \) or its inverse at \( u \) by interpolating between the vertices. If we have the transition functions \( \phi_{L/K} \) and \( \phi_{M/L} \), then \( \phi_{M/K} = \phi_{L/K} \circ \phi_{M/L} \) has as its lower breaks the union of: (a) the lower breaks of \( \phi_{M/L} \); and (b) \( \phi_{L/K}^{-1} \) applied to the lower breaks of \( \phi_{L/K} \). The upper breaks are similar, and hence we have the vertices of \( \phi_{M/K} \) and therefore deduce a function to compose transition functions.

Now if we are given such an \( M/L/K \) say, with \( M/L \) and \( L/K \) each defined by an Eisenstein polynomial over an unramified extension, then we can compute the ramification polygons of \( M/L \) and \( L/K \) via the definition above. From this, we can compute the transition functions \( \phi_{M/L} \) and \( \phi_{L/K} \). From these and the composition routine described above, we can compute \( \phi_{M/K} \) and from this compute the ramification polygon of \( M/K \). In this manner, we deduce an intrinsic ```ramificationpolygon``` to compute the ramification polygon of an arbitrary extension of \( p \)-adic fields and ```transitionfunction``` to compute the corresponding transition function.

9.7 Hensel’s lemma for univariate root-finding

Recall Hensel’s classic lemma.

**Lemma 9.7** (Hensel). Suppose \( f(x) \in \mathcal{O}[x] \), \( a \in \mathcal{O} \) such that \( v(f(a)) \geq s > 0 = v(f'(a)) \). Then there exists a unique \( b \in K \) such that \( f(b) = 0 \) and \( v(a-b) \geq s \). More precisely, defining \( a' := a - f(a)/f'(a) \) then \( v(f(a')) \geq 2s \) and \( v(f'(a')) = 0 \), so iterating \( a \mapsto a' \) then \( a \rightarrow b \).

We refer to the iteration process in Hensel’s lemma as “Hensel lifting”. It can be generalized to non-integral inputs:

**Lemma 9.8.** Suppose \( f(x) \in K[x] \), where \( K \) is a \( p \)-adic field, and \( a \in K \) such that among all roots \( b \) of \( f \), \( v(a-b) \) is maximised precisely once. Then iterating \( a \mapsto a - f(a)/f'(a) \) yields \( a \rightarrow b \).
**Proof.** The generalization is actually reducible to the original version.

Consider the polynomial $f(x + a)$. Its roots are $b - a$ where $b$ is a root of $f$, and so its Newton polygon measures the number of times each $v(a - b)$ occurs. Hence the hypothesis is equivalent to saying that the first face of the Newton polygon of $f(x + a)$ has width 1.

Suppose this is true, then in particular the first face has integral slope and so there exist $j, k \in \mathbb{Z}$ so that $g(x) := \pi^j f(\pi^k x + a)$ has integral coefficients, $\text{val}(g_0) > 0$ and $\text{val}(g_1) = 0$. Note that $g_0 = g(0)$ and $g_1 = g'(0)$ so the original version of Hensel’s lemma applies to $g$ and 0. By linearity, Hensel lifting on $g$ is equivalent to Hensel lifting on $f$.

**Remark 9.9.** Krasner’s lemma is a corollary of this form of Hensel’s lemma.

We provide an intrinsic `IsHenselLiftable` which takes as input a polynomial $f(x) \in K[x]$ and an element $a \in K$ and returns true if this generalized version of Hensel’s lemma can be applied to find a root $b$ of $f$ close to $a$. If so, it also returns that root.

The algorithm proceeds by computing $f(x + a)$ to sufficient precision to see if the first face of its Newton polygon has width 1 or not. If so, then the returned root has as its initial approximation the approximation of $a$ truncated to a certain precision determined by Hensel’s lemma, and its update function performs the Hensel lifting iteration above.

```plaintext
intrinsic IsHenselLiftable(
    f :: RngUPolElt_FldPadExact,
    a :: FldPadExactElt
) -> BoolElt, FldPadExactElt

// first determine if Hensel's lemma is applicable
// try successively precise approximations
for n in 1, 2, ... do
    // get an approximation of f and a
    xf := EpochApproximation(f, n)
    xa := EpochApproximation(a, n)
    // approximate f(x+a)
    xf2 := Evaluate(xf, x + xa)
    // this Newton polygon is computed from the *weak*
    // valuations, so is not necessarily correct
    np := NewtonPolygon(xf2)
    face := Faces(np)[1]
    // if the first face has width 1 and the right hand
    // vertex is correct, then there really is a face of
    // width 1
    if Width(face) eq 1
    and not IsWeaklyZero(Coefficient(xf, 1))
```
then break
  // if the face has higher width, and both vertices
  // are correct, then there really is a face of this
  // width
  elif Width(face) ne 1
  and not IsWeaklyZero(Coefficient(xf, 0))
  and not IsWeaklyZero(Coefficient(xf, EndVertices(face)[2][1]))
  then return false
  // else we cannot conclude whether the first face
  // has width 1 or not
  else continue
  // if we get this far, then a is Hensel liftable
  // we omit the implementation of Hensel lifting
  root := ...
  return true, root

9.8 Univariate root finding I

Magma provides an intrinsic \texttt{Roots} to find all of the roots of a univariate polynomial over an inexact \textit{p}-adic field. As discussed in §1.2, perhaps confusingly these are roots “up to precision”, so for example given the polynomial $x^2 + 2^{10}z_2$ over $Q_2$, it will return the root $0 + 2^{10}z_2$ with multiplicity 2. In a sense this is misleading, because it could be that the polynomial is actually $x^2 + 2^{11}$ to absolute precision 10, and this polynomial does not have any roots. Hence, one should not interpret the existence of roots of an inexact polynomial to necessarily be roots of any lift of that polynomial to something more precise.

On the other hand, a \texttt{Roots} intrinsic for exact polynomials should only return genuine roots of the full-precision polynomial. We can use the inexact \texttt{Roots} intrinsic and \texttt{IsHenselLiftable} to achieve the desired result:

\begin{verbatim}
intrinsic Roots(f :: RngUPolElt_FldPadExact) -> []
for n in 1,2,... do
  // get an approximation to f
  xf := EpochApproximation(f, n)
  // compute the roots of f up to precision
  xroots := Roots(xf)
  // check that the roots are all Hensel liftable
  roots := []
  for xroot in xroots do
    // the roots must be distinct, up to precision,
    // to have a chance of succeeding; if not, go
    // to the next precision in the strategy
    if Multiplicity(xroot) ne 1 then
      roots := [...]
    end if
  end for
  return roots
end for
\end{verbatim}
continue n
// see if an approximation to the root is
// Hensel liftable to a genuine root of f
ok, root := IsHenselLiftable(f, xroot)
// if not, then go to the next precision
if not ok then
    continue n
// if we get this far, we have a root
Append(~roots, root)
// if we get this far, we have a full set of roots
return roots
// if we get this far, we have run out of things to try
error "precision error"

Note that this can only succeed if all of the roots over the base field
are simple, because Hensel’s lemma can only detect simple roots. This is
the best possible: if \( f \) has a root \( r \) of multiplicity \( m \), then to any precision
this is indistinguishable from \( f \) having an irreducible factor of degree \( m \),
all of whose roots are very close to \( r \). For example, over \( \mathbb{Q}_2 \), the root 1
to multiplicity \( m \) is indistinguishable to high precision from an irreducible
factor whose roots are \( 1 + 2^{1000} \sqrt[2]{2} \). Hence it is not possible to prove that
a polynomial to any finite precision has repeated roots.

9.9 Hensel’s lemma for multivariate root finding

We are now interested in solving square systems of multivariate polynomi-
als, namely we wish to find the roots of systems of \( n \) polynomials
\( f(x) = (f_1(x), \ldots, f_n(x)) \in K[x]^n \) in \( n \) variables \( x = (x_1, \ldots, x_n) \). A root of such a
system is an element \( r \in K^n \) such that \( f(r) = 0 \).

The following multivariate version of Hensel’s lemma is well-known:

**Lemma 9.10.** Suppose \( f(x) \in \mathcal{O}[x]^n \) is a system of \( n \) polynomials in \( n \) variables, \( a \in \mathcal{O}^n \), \( \text{val}(f(a)) \geq s > 2t = 2 \text{val}(\det J(f)(a)) \) where \( J(f)_{i,j} = \frac{\partial f_i}{\partial x_j} \). Then there is a unique \( b \in K^n \) so that \( f(b) = 0 \) and \( \text{val}(a - b) \geq s - t \).
More precisely, defining \( a' = a - f(a)J(f)(a)^{-1} \), then \( \text{val}(\det J(f)(a')) = t \)
and \( v(f(a')) \geq 2(s - t) \); therefore iterating \( a \mapsto a' \) then \( a \to b \).

We can state a slightly more general version, which says that if we can
apply a linear change to the equations, perhaps over an extension, such that
Hensel’s lemma applies, then Hensel’s lemma also applies to the original
system over the base field:

**Lemma 9.11.** Suppose \( f(x) \in K[x] \) is a system of \( n \) polynomials in \( n \) variables, \( a \in K^n \), \( L/K \) a finite extension, \( M, N \in \text{GL}_n(L) \), \( \tilde{a} := Ma \in \mathcal{O}^n_L \),
\( \tilde{f} := NF(M^{-1}x) \in \mathcal{O}_L[x]^n \), \( \text{val}(\tilde{f}(\tilde{a})) \geq s > 2t = \text{val}(\det J(\tilde{f})(\tilde{a})) \). Then \( a \)
Hensel lifts to a unique root of \( \tilde{f} \) in \( K \).
Proof. Define \( \tilde{a} = Ma, \tilde{a}' = \tilde{a} - \tilde{f}(\tilde{a})J(\tilde{f})(\tilde{a})^{-1} \). We know that iterating \( \tilde{a} \mapsto \tilde{a}' \) then \( \tilde{a} \to b \in L \) a root of \( \tilde{f} \). By linearity we find \( \tilde{a}' = Ma' \) where \( a' = a - f(a)J(f)(a)^{-1} \). We conclude that \( a \to b \) such that \( Mb = \tilde{b} \), and since \( a' \in K \), then \( b \in K \) also.

In the package, we provide an intrinsic \texttt{IsHenselLiftable} which takes as input such a system \( f \) and a near-root \( a \) and returns true if Hensel’s lemma is applicable. If so, it also returns the Hensel-lifted root itself. It also optionally accepts two vectors \( \mu, \nu \in \mathbb{Q}^n \) which define the diagonal matrices \( M \) and \( N \) with diagonal entries \( \pi^\mu \) and \( \pi^\nu \), and uses the more general version of Hensel’s lemma. This allows us to implicitly rescale the equations and variables, so that the inputs need not be integral.

It should be possible to determine whether there exists any such \( \mu \) and \( \nu \) so that Hensel’s lemma is applicable, and therefore recover a completely general and parameterless version of multivariate \texttt{IsHenselLiftable} in analogue with the univariate case. The theory for this has not been completely worked out yet.

Remark 9.12. An algorithm to actually compute the roots or factors of such a square system is work in progress.

### 9.10 Hensel’s lemma for univariate factorization

Suppose \( f(x) \in K[x] \) is a monic univariate polynomial of degree \( n = n_1 + n_2 \). Consider the problem of finding a factorization \( f(x) = g(x)h(x) \) where \( \deg(g) = n_1, \deg(h) = n_2 \) and \( g \) and \( h \) are monic. By treating the \( n \) coefficients of \( 1, x, \ldots, x^{n-1} \) in \( f(x) - g(x)h(x) \) as multivariate polynomials in the \( n_1 \) coefficients of \( g \) and the \( n_2 \) coefficients of \( h \), we have a system of \( n \) multivariate polynomials in \( n \) variables to solve.

We conclude that there is a version of Hensel’s lemma applicable to this situation, provided that a given near-factorization \( f(x) \approx g(x)h(x) \) is sufficiently accurate. How accurate this needs to be is controlled by the determinant of the Jacobian matrix \( J \) in Hensel’s lemma. In this case, the first \( n_1 \) rows of \( J \) correspond to \( \frac{df-gh}{dg} = x'h(x) \), and the next \( n_2 \) rows correspond to \( \frac{df-gh}{dh} = x'g(x) \), with the columns being the coefficients of these polynomials. This is precisely the matrix defining the resultant, and so we conclude that \( \det(J) = \text{Res}(g, h) \).

For example, we get the following version of Hensel’s lemma for factorization, although more general versions analogous to those in previous sections are also possible.

**Lemma 9.13.** Suppose \( f(x), g(x), h(x) \in \mathcal{O}[x] \) are monic of degrees \( n = n_1 + n_2, n_1, n_2 \suchthat \text{val}(f - gh) \geq s > 2t = 2\text{Res}(g, h) \). Then \( g, h \) Hensel-lift uniquely to a factorization of \( f \).
Suppose we are given \( f(x) \) and \( g(x) \) but not \( h(x) \) and want to determine if \( g(x) \) is Hensel liftable to a factor of \( f(x) \). It seems natural to define \( h := f \div g \) and apply Hensel’s lemma to this. The following lemma shows that this is indeed the best choice for \( h \):

**Lemma 9.14.** If \( f(x), g(x) \in K[x] \) have degrees \( n \) and \( n_1 \leq n \) and \( g \) is monic, then among polynomials \( h(x) \in K[x] \) of degree \( n_2 = n - n_1 \), \( \text{val}(f - gh) \) is maximized by \( h = f \div g \).

**Proof.** By definition, \( f - g(f \div g) = f \mod g =: h_0 \). Consider arbitrary \( h = f \div g + d \), then \( f - gh = f - g(f \div g) - gd = h_0 - gd \). Define \( B = \text{val}(h_0) + 1 \) and suppose there exists \( d \) so that \( \text{val}(h_0) - gd \geq B \). In particular \( d \neq 0 \). Fix \( d \) of smallest degree, and let \( m \) be this degree. Then the \((m + n_1)\)th coefficient of \( f - gh \) is \(-d_m\) and so \( \text{val}(d_m) \geq B \). Define \( d' = d - d_m x^m \), then \( \text{val}(h_0) - gd' \geq B \) and \( \deg d' < \deg d \), a contradiction. \( \square \)

The package provides an intrinsic `IsHenselLiftable` which takes as input two polynomials \( f \) and \( g \) and returns true if \( g \) is Hensel-liftable to a factor of \( f \). If so, it also returns the factor itself. In analogue with the multivariate version of `IsHenselLiftable`, this intrinsic takes parameters which implicitly re-scale the polynomials and the variable \( x \) before applying Hensel’s lemma.

### 9.11 Univariate factorization by Newton polygon

An easy application of Hensel’s lemma for univariate factorizations is to factor a polynomial according to its Newton polygon.

Recall that the slopes of faces of the Newton polygon of a polynomial \( f(x) \) correspond to valuations of roots of \( f(x) \), with the width of the face corresponding to the number of roots with this valuation. If two roots of \( f(x) \) come from the same irreducible factor, then they are Galois conjugate and so have the same valuation; we conclude that each face of the Newton polygon corresponds to a factor of \( f \) whose degree is the width of the face.

In fact, we can prove this fact directly using a version of Hensel’s lemma for factoring, seen in the previous section: it is not hard to see that with a suitable choice of rescaling on \( f \) and \( x \) that we may choose \( g \) so that Hensel’s lemma is applicable. Specifically, we rescale so that the selected face of the Newton polygon of \( f \) becomes horizontal and incident with the \( x \)-axis, and take for \( g \) the polynomial formed from the coefficients of \( f \) corresponding to the face.

The package provides a routine `NewtonPolygonFactorization` which takes as input a univariate polynomial \( f \) and returns its factorization according to its Newton polygon. It is implemented essentially by first computing the Newton polygon of \( f \), and then for each face constructing a suitable \( g \) and calling `IsHenselLiftable` to produce a factor.
9.12 Univariate factorization into irreducibles I

We also provide an intrinsic \texttt{Factorization} which returns the full factorization of a polynomial \( f(x) \) into irreducible factors.

It is implemented in a very similar fashion to \texttt{Roots} as described in \S 9.8 it calls Magma’s builtin \texttt{Factorization} routine on an approximation to \( f(x) \), and then checks if each factor returned is Hensel liftable using \texttt{IsHenselLiftable}.

9.13 Univariate root finding and factorization into irreducibles II

Our \texttt{Roots} and \texttt{Factorization} intrinsics actually have a parameter \texttt{Alg} to select between two different algorithms. We have already described \texttt{Alg:="Builtin"} (\S 9.8, \S 9.12) which is a wrapper around the builtin intrinsics for inexact \( p \)-adics.

With the parameter \texttt{Alg:="OM"}, which is now the default, we use our own implementation of an “OM algorithm” for computing “Okutsu invariants” of the input polynomial, which identifies its irreducible factors and some properties of the extensions they define. From these, we can use “single factor lifting” to generate arbitrarily precise approximations to the factors. The algorithm is essentially that described in [10, Ch. VI].

\textit{Remark} 9.15. Although not usually presented as such, “single factor lifting” is nothing but Hensel’s lemma in disguise. Recall in [9.10] that we expressed factoring \( f(x) = g(x)h(x) \) as a multivariate system of equations whose coefficients are the \( n_1 = \deg(g) \) coefficients of \( g \) and the \( n_2 = \deg(h) \) coefficients of \( h \).

We can instead write \( g(x) = x^{n_1} + \sum_{i < n_1} g'_i X_{g,i}(x) \) and \( h(x) = x^{n_2} + \sum_{i < n_2} h'_i X_{h,i}(x) \) where \( X_{g,i}(x) \in K[x] \) are fixed monic polynomials of degree \( i \), and instead consider \( f(x) = g(x)h(x) \) as a system of equations in the variables \( g'_i \) and \( h'_i \). Essentially, we have chosen bases for the vector spaces of monic polynomials of degrees \( n_1 \) and \( n_2 \) different from the usual \( 1, x, x^2, \ldots \).

This is a linear change of variables of the sort considered in Lemma 9.11.

The OM algorithm builds up such a basis for each factor, and the point in the algorithm at which an irreducible factor is identified is precisely the point at which Hensel’s lemma, in terms of this basis, can be invoked.

\textit{Remark} 9.16. The same algorithm is also made available as an intrinsic \texttt{ExactpAdics_Factorization} which can take an inexact \( p \)-adic polynomial. This can be used independently of the package.

References

[1] W. Bosma, J. Cannon, and C. Playoust. The Magma algebra system. I: the user language, 1997. \texttt{http://magma.maths.usyd.edu.au}
[2] X. Caruso. Computations with p-adic numbers. *ArXiv e-prints*, Jan 2017. [https://arxiv.org/abs/1701.06794](https://arxiv.org/abs/1701.06794).

[3] T. Dokchitser and C. Doris. 3-torsion and conductor of genus 2 curves. *ArXiv e-prints*, Jun 2017. [https://arxiv.org/abs/1706.06162](https://arxiv.org/abs/1706.06162).

[4] C. Doris. Computing the Galois group of a polynomial over a p-adic field. In preparation.

[5] C. Greve and S. Pauli. Ramification polygons, splitting fields and Galois groups of Eisenstein polynomials. *Int. J. Number Theory*, 8(6):1401–1424, 2012.

[6] W. Hart, F. Johansson, and S. Pancratz. FLINT: Fast Library for Number Theory. [http://flintlib.org](http://flintlib.org).

[7] C. Helou. Non-Galois ramification theory of local fields. *Algebra Berichte*, 64, 1990.

[8] S. Pauli and B. Sinclair. Enumerating extensions of (π)-adic fields with given invariants. *Int. J. Number Theory*, 13(8), 2017.

[9] J.-P. Serre. *Local Fields*. Springer-Verlag, 1979.

[10] B. Sinclair. *Algorithms for enumerating invariants and extensions of local fields*. PhD thesis, University of North Carolina at Greensboro, 2015.

[11] The Sage Developers. Sagemath, the Sage Mathematics Software System. [http://www.sagemath.org](http://www.sagemath.org).

[12] J. van der Hoeven, G. Lecerf, and B. Mourrain. The Mathemagix computer algebra and analysis system. [http://www.mathemagix.org](http://www.mathemagix.org).