CHALMERS

Dynamic construction of aglebraic closure and a coinductive proof of Hensel's lemma

Master of Science Thesis in the Master Degree Programme, Computer Science: Algorithms, Languages and Logic

BASSEL MANNAA

CHALMERS UNIVERSITY OF TECHNOLOGY
Department of Computer Science and Engineering
Division of Computer Science
Göteborg, Sweden, May 2010

Dynamic construction of algebraic closure and a Coinductive proof of Hensel's lemma

Bassel Mannaa

Department of Computer science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2010

Dynamic construction of algebraic closure and a Coinductive proof of Hensel's lemma

Bassel Mannaa

Examiner: Prof. Thierry Coquand

Chalmers University of Technology
University of Gothenburg
Department of Computer Science and Engineering
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

ACKNOWLEDGMENTS

I'd like to thank Thierry Coquand for all the guidance and support he provided throughout the work on this thesis. Also, for lots of inspiring and interesting discussions.
I'd also like to thank Anders Mörtberg for the many useful comments and suggestions and for the nice Haskell library.

ABSTRACT

In this thesis we present a dynamic construction of the algebraic closure of a zero characteristic field implemented in the functional programming language Haskell based on Duval's dynamic evaluation method. We also present a complete formalization of the ring of formal power series. Based on that we present a coinductive proof Hensel's lemma. As an application we present an implementation of Newton algorithm for factorization of polynomials with power series coefficients.

CONTENTS

CHAPTER I
INTRODUCTION

> *Whenever you can settle a question by explicit construction, be not satisfied with purely existential arguments.*
>
> Herman Weyl[9]

Although it has become more of a side issue rather than the ferocious battle it once was, the debate over the foundations of mathematics that started in the 1920's seems to be unsettled today as it was then. Considering the enormous body of mathematics developed classically, it is not difficult to see why, however strong the case for constructivism is, mathematicians are reluctant to fully embrace it. On the other hand, Computer science was born outside the paradise [1], for the debate seems to be already settled a priori by the correspondence between the notion of *explicit construction* and computable functions. As Dirk Van Dalen put it "Although mathematics and computer science were confronted with similar problems, mathematics could afford to ignore the issue; classical mathematics could and can very well survive without racking its brain about effectiveness"[16].

1.1 Motivation

Computer algebra systems have been in use of at least 3 decades. However, many of these systems suffer from serious lacuna, that is, the algorithms they employ are not correct as they stand, they usually need a separate -paper and pencil- proof of correctness.

Constructive interpretation to logical propositions, for example BHK interpretation of intuitionistic logic[2], asserts the computational nature of proofs. Modern

[1] The reference here is to Hilbert's famous statement "We will not be driven out of the paradise Cantor has created for us".

[2] Curry-Howard isomorphism is another example, although in its original form it only applied

type theories such as Per Martin-Löf type theory, Coquand's calculus of construction provide a formalization of such interpretation. With the availability of proof assistants (i.e. Agda, Coq) based on these formal systems, the elements for a formal type theoretic computer algebra system are in place.

Although important on its own right, a type theoretic formalization of algebra has more to offer than just providing provably correct algorithms for a computer algebra system. Deeper results can be obtained from such attempt, for example not all algebraic notions are constructively sensible, providing such classification is undoubtedly very important. This can be seen as analogous to Bishop's important work in analysis [4].

The work presented in hereby can be regarded as a part of this attempt of type theoretic formalization of mathematics. The original goal of the thesis was to provide a constructive formalization of Newton theorem. Unfortunately, this goal was not attained. However, as will be outlined next, some parts of this construction have been completed.

1.2 Outline of the thesis

The thesis aimed at providing a formalization of a proof of Newton theorem[3] based on a proof by S.S.Abhyankar [1]. The two major elements of the formalization are as follows:

1. Formal proof of Hensel's lemma

2. Formal construction of algebraic closure of a 0 characteristic field.

Each element is further divided into implementation/formalization phases. Where the implementation is conducted in the functional programming language Haskell. A

to intuitionistic propositional logic; For this logic both interpretations are quite similar except that Curry-Howard isomorphism explicitly defines the class of functions as those of the simply typed lambda calculus.

[3] The theorem is widely known as Newton-Puiseux algorithm. However, the one version under consideration is substantially different in that it does not use Newton polygon.

Haskell implementation of Hensel's lemma, the algebraic closure of a zero characteristic field and Newton algorithm is now complete. In addition a full formal constructive proof of Hensel's lemma is presented. This leaves this thesis short of its original goal by the formalization of the algebraic closure of a 0 characteristic field and Newton theorem.

Remark 1.2.1. *The Haskell code of this thesis is integrated in the constructive-algebra library* `http://hackage.haskell.org/package/constructive-algebra`[4].

1.3 Outline of this document

In chapter II we explain the construction of the algebraic closure as implemented in this thesis. Chapter III contain with a brief introduction to coinduction and the coninductive proof principle. In Chapter IV a formalization of the formal power series ring is provided. In chapter V we give a coinductive proof of Hensel's lemma. In chapter VI Newton theorem is explained briefly followed by some examples from the Haskell implementation of the algorithm. The last chapter contains the concluding remarks and directions of future work.

[4] As of May 2010, only a fraction of the library has been moved to this location.

Part I

Algebraic Closure

CHAPTER II

Dynamic construction of algebraic closure

To the absolute number multiplied by four times the [coefficient of the] square, add the square of the [coefficient of the] middle term; the square root of the same, less the [coefficient of the] middle term, being divided by twice the [coefficient of the] square is the value.

Brahmasphuta-siddhanta[5]

Finding the roots of a polynomial is one of the oldest mathematical problems. Methods to solve certain forms of quadratic equations were developed in many ancient civilizations. The Persian mathematician al-Khwarizmi was probably the first one to provide a full solution to the general quadratic equation. The search for general solution for higher degree equation occupied mathematician since then with formulas for cubic and quartic equations developed in the 16^{th} century[1]. In the 19^{th} century Abel–Ruffini theorem showed that there is no formula for general quintic equation[2].

2.1 Kronecker model

We know by the *fundamental theory of algebra* that the roots of a polynomial $p(X) \in \mathbb{C}[X]$ exist in $\mathbb{C}$ [3]. However, this fact is useless since it doesn't -and sure can't- provide a way to compute these roots. Leopold Kronecker escape of this *Cul-de-sac* was to change the meaning of "solving" a polynomial from the intuitive one of

[1] A solution then meant computation of the roots involving arithmetic operations and radicals on the coefficients.

[2] More modern proofs of the same fact are stated in terms of Galois groups, while the Galois group of a polynomial of degrees less than 5 is solvable, this is not always the case for polynomials of higher degrees.

[3] We assume the reader is familiar with basic algebra. For a reference see Serge Lang, *Algebra*,Springer 2002 and Nathan Jacobson, *Basic Algebra I*, Dover 2009.

computing the roots of the polynomial to the less restrictive meaning of computing *with* the roots of the polynomial[13]. Kronecker idea was simply as follows, given an irreducible polynomial $p(X) \in k[X]$ we can construct the field $k[X]/(p(X))$. If we let $\pi : k[X] \rightarrow k[X]/p(X)$ be the canonical homomorphism then $p(\pi(X)) = \pi(p(X)) = 0$ and hence $\alpha = \pi(X)$ is a root of p. Hence, doing all computation in $k[X]$ modulo $p(X)$ is essentially the same as doing them in $k(\alpha)$ which is the finite extension of k by the root α. Having this in mind it is possible to construct a splitting field $K \supseteq k$ of a polynomial $f(X) \in k[X]$ inductively provided that we have a way to factorize polynomials over the base field as well as the extensions we obtain at each step.

Theorem 2.1.1. *Given a polynomial $f(X) \in k[X]$, then there exist a field $K \supseteq k$ in which $f(X)$ splits (factorize into linear factors)*

Proof. Let $deg(f) = n > 0$ and g_0 be an irreducible factor of f. If g_0 is linear, $g_0 = (X - a)$ then the root a is in k and $f(X) = (X - a)f_1(X)$ with $f_1(X) \in k[X]$. We let $k_1 = k$ and $\alpha_1 = a$. Otherwise, we build a field $k_1 = k[X]/(g_0)$. If we let $\alpha_1 = \pi_{01}(X)$ where $\pi_{01} : k[X] \rightarrow k[X]/(g_0)$ is the canonical projection then we know that $f(X) = (X - \alpha_1)f_1(X)$ for some $f_1(X) \in k_1[X]$. Inductively applying the same argument to f_1 we obtain a tower of fields $k = k_0 \subseteq k_1 \subseteq ... \subseteq k_{n-1} = K$ and K is a splitting field of f, i.e. the minimal extension of k containing the roots of f.

More over it can be shown that splitting fields for a given polynomial are isomorphic[4], hence if K is some splitting field of $p \in \mathbb{Q}[X]$ then $\mathbb{C}$ contains an isomorphic copy of K.

In 1930 Van Der Waerden showed that there is no *general* method for factorization of polynomials in $k[X]$ for any explicitly given field k[18]. Later on Fröhlich and Shepherdson strengthened this result by giving a particular field k_0 for which irreducibility of elements in $k_0[X]$ is undecidable, and hence no factorization algorithm exist [10].

[4] This result can be shown classically, constructively however, it is not always possible to construct the isomorphism[14] p.152.

Although Van Der Waerden result points to a limitation of Kronecker's model. The situation is not so bleak since general factorization algorithms exist for polynomials over finite fields and any finite extension of $\mathbb{Q}$, which are indeed the interesting fields for almost all practical purposes. The main disadvantage of Kronecker's model lies in the high computational cost of factorization algorithms. Duval's model on which we based our approach to constructing the algebraic closure overcomes this problem by avoiding factorization completely.

2.2 Dynamic evaluation

Duval's method[8] stems from the simple observation that for any two polynomials $p(X)$ and $q(X)$ in $\mathbb{Q}[X]$ such that $p(X)$ is square free, $q(X)$ is zero modulo $gcd(p, q)$ and is invertible modulo $p/gcd(p, q)$[12]. So instead of constructing the splitting field of p explicitly, we perform computations in $\mathbb{Q}\langle\alpha\rangle = \mathbb{Q}[X]/(p(X))$ as long as we can *as if* it is the splitting field of p. In other words, α represents all roots of p until we are forced to decide which factor of p it belongs to. In turn this decision is done in a lazy manner; for example if we need to answer a question in the form "$q(\alpha) = 0$?" we split $\mathbb{Q}\langle\alpha\rangle$ into $\mathbb{Q}\langle\alpha_1\rangle = \mathbb{Q}[X]/(p_1(X))$ and $\mathbb{Q}\langle\alpha_2\rangle = \mathbb{Q}[X]/(p_2(X))$ where $p_1 = gcd(p, q)$ and $p_2 = p/gcd(p, q)$ (p_1 and p_2 are consequently coprime but not necessarily irreducible). Then the answer to the question is "*yes*" in the first branch and "*no*" in the second.

The method just illustrated is called dynamic evaluation, in the rest of this section we explain it in more details and present the Haskell implementation[5].

2.2.1 Preliminaries

Theorem 2.2.1. *If $p(X), q(X) \in k[X]$ are two coprime polynomials and $\pi : k[X] \to k[X]/(p(X))$ is the canonical homomorphism, then $\pi(q)$ is unit.*

Proof. since p and q are coprime then we can find r and s such that $r\ p + s\ q = 1$.

[5] The first implementation of this approach was on the D5 system which is written in R-Lisp and has been used with Reduce.

Then $\pi(r\ p + s\ q) = \pi(s)\ \pi(q) = 1$, hence $\pi(q)$ is a unit.

Corollary 2.2.2. *If $p(X) \in k[X]$ is irreducible then $k[X]/(p(X))$ is a field.*

Remark 2.2.3. *When we say that $q(X) \in k[X]$ is a unit (or zero) in $k[X]/(p(X))$ for some p we mean that the canonical projection of q in $k[X]/(p(X))$ is a unit (or zero) respectively.*

Definition 2.2.4. *A polynomial $f \in k[X]$ is said to be square free if and only if $g^2 \nmid f$ for all non-constant $g \in k[X]$.*

Hence, If f is square free and $f = \prod_i p_i^{e_i}$ is its factorization into irreducible factors in $k[X]$, then each e_i is either 0 or 1.

Theorem 2.2.5 (The Chinese remainder theorem). *Let $p \in k[X]$ be a polynomial such that $p = p_1\ p_2$ and $gcd(p_1, p_2) = 1$, then $k[X]/(p) \cong k[X]/(p_1) \oplus k[X]/(p_2)$*

Proof. omitted.

Corollary 2.2.6. *If $p \in k[X]$ be a polynomial with factorization $p = \prod_{i=1}^{n} p_i^{e_i}$. Let $R = k[X]/(p)$ and $R_i = k[X]/(p_i^{ei})$. Then $R \cong \bigoplus R_i$. Moreover, if p is square free then each R_i is a field.*

Computing in the splitting field of a square free polynomial $p \in k[X]$ dynamically amounts to computing in $k[X]/(p)$ instead of computing in the fields $k[X]/(p_i)$ for each irreducible factor p_i of p. However, since not all computations are decidable in $k[X]/(p)$ (for example, a question in the form "$q(X) = 0$?") the internal representation of the root X evolves with such computations. This evolution is mainly a minimal branching (split) of the ring $k[X]/(p)$ such that the question is decidable in each branch. In terms of the Chinese remainder theorem this branching is a decomposition of the ring into the smallest -with regard to the number of terms- possible direct sum of rings in which the question is decidable. Of course we may have more than one polynomial to begin with, i.e. a ring $((k[X_1]/(p_1))[X_2]/(p_2))...[X_n]/(p_n)$ such that each $p_i(X_1, .., X_i)$ is square free in $(k[X_1, .., X_{i-1}]/(p_{i-1}))[X_i]$. For brevity we call this ring $k\langle p_1, ..., p_n\rangle$.

Implementation note II.1 (State Monad)**:** The evolution of the internal representation of the roots is managed through a state monad in the form

```
newtype ST s a = ST { runState :: s -> [(s,a)] }


instance Monad (ST s) where
 (ST p) >>= k = ST (\s0 -> let as = p s0 in
                             concatMap (\(st,vl) -> runState (k vl) st) as)
 return a = ST (\s -> [(s,a)])
```

A state is just a list of multivariate polynomials over some field k. With the following conditions for each element p_i in the list

- It must have at least one term $a\ X_i^{e_i}$ with both $a \neq 0$ and $e_i > 0$
- The coefficients of $X_j^{e_j}$ for all $j > i$ are zero.
- It must be square free as a univariate polynomial in $k\langle p_0, ..., p_{i-1}\rangle[X]$

```
type R k = MPoly k Len
type S k = [R k]
```

[6] Intuitively, in each tuple `(s,a)`, the state `s` contains a refined version of the original state and `a` is the result of the computation which is not decidable in the original state but decidable in `s`. Consequently all the computations are done monadically (in parallel) with respect to some state.

2.2.2 The square free condition

The square free condition guarantees that for any q, if $p \nmid q$ is square free and both have degree greater than 0, then $gcd(p, q)$ and $p/gcd(p, q)$ are coprime.

[6] The code presented here uses a Haskell library initiated by Anders Mörtberg, in fact the work of this thesis is currently part of this library. Most of the basic algebraic structures (such as `MPoly`) are due to Anders.

Theorem 2.2.7. *For two non constant polynomials $q(X)$ and $p(X)$ in $k[X]$, such that p is square free, then*

1. *If $p \mid q$ then q is a zero in $k[X]/(p(X))$*
2. *If $gcd(p, q) = 1$ then q is a unit in $k[X]/(p(X))$*
3. *If $p \nmid q$ and $deg(gcd(p, q)) > 0$, let $g = gcd(p, q)$ and $h = p/gcd(p, q)$ then q is zero in $k[X]/(g(X))$ and a unit in $k[X]/(h(X))$.*

Proof. (1) Let $\pi : k[X] \to k[X]/(p(X))$ be the cannonical projection. Since $p \mid q$ then $q = r\ p$, hence $\pi(q) = \pi(r\ p) = \pi(r)\pi(p) = 0$
(2) Theorem 2.2.1
(3) Since $p \nmid q$, then $deg(h) > 0$. Because p is square free we have $p = g\ h = \prod_i p_i$ where each p_i is irreducible and $p_k \neq p_j$ for $k \neq j$ and hence g and h are coprime and we can find t and u such that $t\ g + u\ h = 1$. From this we know that g is a unit in $k[X]/(h(X))$ with $\pi(t)$ as the where $\pi : k[X] \to k[X]/(h(X))$ is the cannonical porjection. We also have $r\ p + s\ q = g$ for some polynoimals r and s. Hence the projection of q is a unit with the inverse $\pi(q)^{-1} = \pi(s)\ \pi(g)^{-1} = \pi(s)\ \pi(t)$. The fact that q is zero in $k[X]/(g(X))$ follows from the proof of (1).

Corollary 2.2.8. *With the same notation of theorem 2.2.7*

1. *If $deg(g) = 0$ then q is invertible in $k[X]/(p)$*
2. *If $deg(g) > 0$ then q is invertible in $k[X]/(p_1)$ and is zero in $k[X]/(g)$*

Theorem 2.2.9. *For a 0 characteristic field k and $f(X) \in k[X]$ a non-constant polynomial. f is square free if and only if $gcd(f, f') = 1$, where f' is the derivative of f with respect to X.*

Proof. ($\Rightarrow$) Let $f = \prod_{i=1}^{n} f_i$ be the factorization of f with each $f_i(X)$ an irreducible polynomial in $k[X]$. We do induction on n. If $n = 1$ then f is irreducible. Because f is non constant and $char(k) = 0$ we have $f'(X) \neq 0$, then $gcd(f, f')$ is either 1 or f,

but since $deg(f') < deg(f)$ then $gcd(f, f') = 1$.
Assuming the the theorem for some n, Let $h = \prod_{i=1}^{n} f_i$ and $f = f_{n+1}\ h$ with $f_i(X) \in k[X]$ irreducible and $f_i \neq f_j$ for $i \neq j$. Then
$f' = f'_{n+1}\ h + f_{n+1}\ h'$
Since $gcd(f_{n+1}, f'_{n+1}) = 1$ and by induction hypothesis we have $gcd(h, h') = 1$ it follows that $gcd(f, f') = 1$.
($\Leftarrow$) Let $g(X) \in k[X]$ be an irreducible factor of f with multiplicity r. Then
$f' = r\ g^{r-1}\ g'\ h + g^r\ h'$
if $r > 1$ then g^{r-1} divides both f and f' and $gcd(f, f') \neq 1$.

Definition 2.2.10. *Let $f \in k[X]$ and $f = \prod_{i=1}^{n} f_i^{e_i}$ be its factorization in $k[X]$. By the square free associate of f we mean $\prod_{i=1}^{n} f_i$, i.e. the product of the factors of f each taken with multiplicity 1. By a distinct power factorization of f we mean a list where the i^{th} element is the product of factors of f of multiplicity i in $k[X]$.*

Proposition 2.2.11. *Let $f \in k[X]$, $p = gcd(f, f')$, $q = f/p$ then q is the square free associate of f.*

Proof. [7] putting f in its distinct power factorization form we have for some g such that $gcd(f, g) = 1$

$$\begin{array}{lclllll}
f & = & & f_1 & f_2^2 & ... & f_i^i & ... \\
f' & = & \mathrm{g} & & f_2 & ... & f_i^{i-1} & ... \\
p & = & & & f_2 & ... & f_i^{i-1} & ... \\
q & = & & f_1 & f_2 & ... & f_i & ...
\end{array}$$

Implementation note II.2 (Square free associate)**:** In the implementation we restrict ourselves to char 0 fields. We then compute the square free associate of f as follows.

```
sqfr :: (Field k, Eq k) => UPoly (R k) x -> ST (S k) (UPoly (R k) x)
sqfr f = do let f’ = deriv f
```

[7] from Teo Mora, [13].

```
            (_,_,_,d,_) <- iGCD f f'
            return d
```

Where `iGCD` is gcd function such that `iGCD (p,q) = (r,s,g,t,u)` such that `r p + s q = g` and `t g = p, u g = q`.
We also compute the distinct power factorization of f.

```
 sqfrDec :: (Field k) => UPoly (R k) x -> ST (S k) [UPoly (R k) x]
```

2.2.3 The Splitting Field

We have given an example of the kind of problems undecidable in $k\langle p\rangle$ for some $p(X) \in k[X]$ that are decidable in the splitting field of p. Computing in $k\langle p\rangle$ *as if* it is a field dynamically can be seen as a *deferred* extension of $k\langle p\rangle$[8] with the axiom schema of fields:
$\forall x.\ x = 0 \vee \exists y.\ x\ y = 1$.
The extension is delayed in the sense that we do not enforce the axiom until we have to, and even in this case we only enforce an instance of the axiom (an instantiation of x to an element of the ring). If we cannot enforce the axiom in the current ring (i.e. cannot decide on whether the element is invertible or zero) we split/branch into two subrings in which the axiom instance for this element is enforced (0 in one and invertible in the other).

Implementation note II.3 (The type of algebraic closure)**:** We use the type

```
 type R k = MPoly k Len
```

as the type of algebraic closure of a the field `k`. It is essentially `k` extended with countably infinite set of roots `{`x_0`,  `x_1`,....}`. An extension with a finite set `{`x_0`,  `x_1`,...,`x_n`}` can be viewed as representing a subfield of the algebraic closure. However, this distinction is not present in the code.

[8] Note that the ring $k\langle p\rangle$ is not necessarily entire (integral domain).

Implementation note II.4 (The inverse function)**:** The witness that the type `R k` is a field is the `inverse` function, which given an element in `R k` answers with either `()` in which case the element is zero or returns the inverse of the element.

```
 inverse :: (Field k, Eq k) => R k -> ST (S k) (Either () (R k))
```

Since computations are done with respect to the state (monadically), at each point of the computation we can use `R k` as a genuine field for which the field axiom $\forall x.\ x = 0 \vee \exists y.\ x\ y = 1$ is enforced by the `inverse` function. However, in reality computation is performed in all branches. It is worthwhile to compare this approach with the approach in [17].

The `inverse` function have 3 cases to consider:

- The element is zero; for example $z^2 - 1$ is zero in $\mathbb{Q}\langle x^2-2, y+3, z+1\rangle$
- The element is unit; for example $y - 1$ is invertible in $\mathbb{Q}\langle x^2-2, y+3, z+1\rangle$
- The element can be either a unit or zero, in which case branching occur; for example it is undecidable whether $y-1$ is zero in $\mathbb{Q}\langle x^2-2, y^2-1, z+1\rangle$, But $y-1$ is zero in $\mathbb{Q}\langle x^2-2, y-1, z+1\rangle$ and a unit in $\mathbb{Q}\langle x^2-2, y+1, z+1\rangle$.

```
(r,s,g,t,u)    <- iGCD p q
(g,g_deg)      <- iDeg g
(_,p_deg)      <- iDeg p
case g_deg > 0 of
 True -> case g_deg == p_deg of
          False -> do .......
                      putD  s1 s2 (Left ()) (Right qinv)
           True -> return $ Left ()
 False -> do ....
             return $ Right $ qinv
```

2.2.4 The algebraic closure

In the previous section we showed that dynamic evaluation corresponds to delayed extension of some ring with the axiom schema of fields. Now we look at the axiom schema for the algebraic closure of a field:

$\forall p(X) \in k[X].\ \exists \alpha \in k.\ \ p(\alpha) = 0$

Now given the ring $k\langle p_1, ..., p_n\rangle$ and given a polynomial $q \in k\langle p_1, ..., p_n\rangle[X]$. We can enforce the axiom above by building the ring $k\langle p_1, ..., p_n, q(\alpha)\rangle$ and take α as a root of q where α is a fresh variable (formal root). This might look peculiar at first glance, but indeed α is a root of q by isomorphism of splitting fields, or rather α is a presentation of all the roots of q.[9]. In fact we compute $\hat{q}$ the square free associate of q and perform the previous step on $\hat{q}$ for the reasons we stated earlier. Again this can be viewed as a delayed enforcing of the axiom schema of algebraically closed field. We enforce an instance of the schema only when we want to compute the root of some polynomial.

Implementation note II.5: The witness that the type `R k` is algebraically closed is the function `root`

```
root :: (Num k, Field k, Eq k, Show k) => UPoly (R k) x -> ST (S k) (R k)
root p = do ....
            --To a multivariate polynomial in the right indeterminates
            let p1 = poly $ mpoly (m+1) p
            --square free associate
            q <- sqfr p1
            let s = mpoly (m+1) q
            --the root x_m+1
            let r  = toMPoly [(one, replicate (fromInteger m) 0 ++ [1])]
```

[9] In [13], p.47-52. Teo Mora discusses this rather baffling tautological definition, namely "the roots of f are the roots of f". For example we accept that $\sqrt{2}$ as a root of $X^2 - 2$, but $\sqrt{2}$ can only be defined as "a root of $X^2 - 2$".

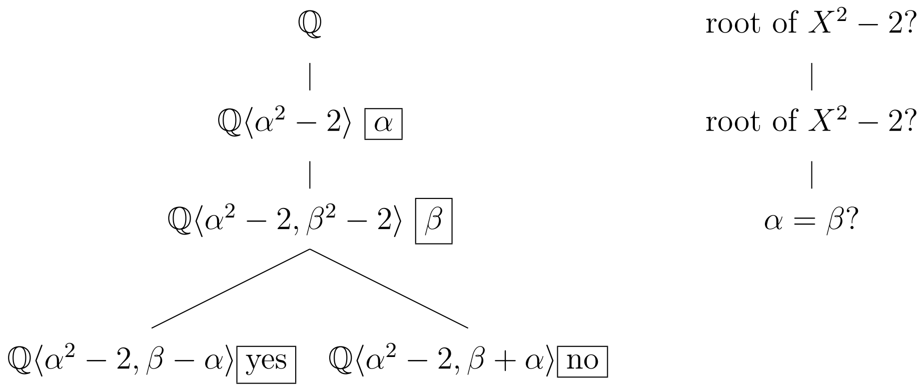
```
            appendToState s
            return $ r
```

Having enforced both axiom schemata in this delayed manner we have a pair (structure, computation) that serves as an algebraic closure. In fact the future plan is to prove that this is a genuine algebraic closure.[10]

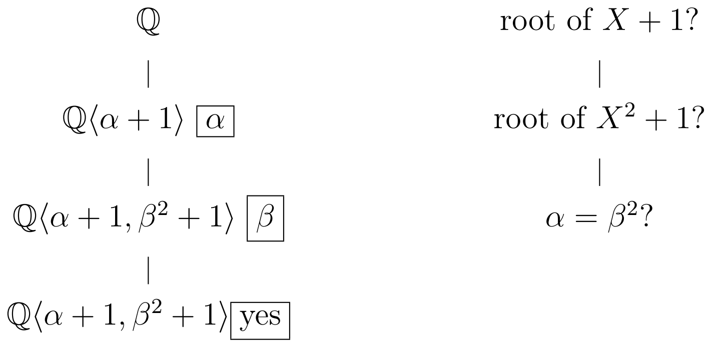
2.2.5 An example

Here we give a simple example of dynamic evaluation. In which we start with the field $\mathbb{Q}$, and see how our approach allows us to treat it as if it was algebraically closed, which we call $\overline{\mathbb{Q}}$. As shown in the figure below, at first $\overline{\mathbb{Q}}$ is internally identical to $\mathbb{Q}$. If we start by asking for the root of $X^2 - 2$ we get α as an answer and at this point the internal representation of $\overline{\mathbb{Q}}$ changes from $\mathbb{Q}$ to $\mathbb{Q}\langle\alpha^2 - 2\rangle$. Then if we repeat the question asking for the root of $X^2 - 2$ again we get a fresh formal root β as an answer and again the internal representation of $\overline{\mathbb{Q}}$ changes to $\mathbb{Q}\langle\alpha^2 - 2, \beta^2 - 2\rangle$. Now since the two polynomials in the two questions are the same it is natural to ask whether $\alpha = \beta$ or we can put the same question in another way by asking for the inverse of $\alpha - \beta$ since according to the field axiom and element is either a unit or zero; in the first case $\alpha = \beta$ and in the second $\alpha \neq \beta$. Since $X^2 - 2$ has two distinct roots the question is undecidable in the current representation of $\overline{\mathbb{Q}}$ and a branching must occur in this case and $\overline{\mathbb{Q}}$ decompose into $\mathbb{Q}\langle\alpha^2 - 2, \beta - \alpha\rangle$ and $\mathbb{Q}\langle\alpha^2 - 2, \beta + \alpha\rangle$. In the first component $\mathbb{Q}\langle\alpha^2 - 2, \beta - \alpha\rangle$ the answer is "yes" and for the second $\mathbb{Q}\langle\alpha^2 - 2, \beta + \alpha\rangle$ the answer is "no". In fact the answer in the second component is actually $\alpha/4$ which is the inverse of $\alpha - \beta$ testifying to the fact that $\alpha - \beta \neq 0$.

[10] We note that the model we present here corresponds to the generic model building approach discussed in [7]. However, it is not clear how we can make use of the consistency proof presented to prove our model correct. It seems to me that the main obstacle is the lack of typing in that proof, for example it is not clear what is the type of x for the predicate $Z(x)$ in [7]p.9 because the type of x evolves with the computation.


$\mathbb{Q}$
$\mathbb{Q}\langle \alpha^2 - 2\rangle$ $\boxed{\alpha}$
$\mathbb{Q}\langle \alpha^2 - 2, \beta^2 - 2\rangle$ $\boxed{\beta}$
$\mathbb{Q}\langle \alpha^2 - 2, \beta - \alpha\rangle\boxed{\text{yes}}$
$\mathbb{Q}\langle \alpha^2 - 2, \beta + \alpha\rangle\boxed{\text{no}}$
root of $X^2 - 2$?
root of $X^2 - 2$?
$\alpha = \beta$?

The two answers correspond to $\alpha = \beta = \pm\sqrt{2}$ and $\alpha = \pm\sqrt{2} \neq \beta = \mp\sqrt{2}$. Compare this to the tree below.


$\mathbb{Q}$
$\mathbb{Q}\langle \alpha + 1\rangle$ $\boxed{\alpha}$
$\mathbb{Q}\langle \alpha + 1, \beta^2 + 1\rangle$ $\boxed{\beta}$
$\mathbb{Q}\langle \alpha + 1, \beta^2 + 1\rangle\boxed{\text{yes}}$
root of $X + 1$?
root of $X^2 + 1$?
$\alpha = \beta^2$?

Implementation note II.6: For our Haskell program the questions as in the example above are answered by applying the inverse function. The previous example translates to

```
ex p = do alpha <- root p
          beta  <- root p
          inverse $ alpha - beta
```

if we let $p = x^2 - 2$ and call `ex` with an argument p

```
 runState (ex p1) []
```

Where the empty list means the current field is the base field with no extensions (i.e. $\mathbb{Q}\langle\rangle = \mathbb{Q}$), we get the following result

```
[([-1/2x^2+1,-y+x],Left ()),([-1/2x^2+1,-1/2y-1/2x],Right 1/4x)]
```

Where `Left ()` means the element is not a unit (i.e. is 0) in the extension of $\mathbb{Q}$ with `[-1/2x^2+1,-y+x]` and `Right 1/4 x` means that the element is a unit and $x/4$ is its inverse in the extension of $\mathbb{Q}$ with `[-1/2x^2+1,-1/2y-1/2x]`.

Part II

Hensel's lemma

CHAPTER III

CoInduction

To prove Hensel's lemma as we will do in the next chapters, we need a way to prove properties of infinite mathematical objects called *formal power series.* Intuitively formal power series can be viewed as polynomials with infinite length.

3.1 Datatypes and Codatatypes

Mathematically formal power series can be defined as follows

Definition 3.1.1 (Formal power series)**.** *Let G be a monoid of functions from the singleton set $\{X\}$ to $\mathbb{N}$. The formal power series over a ring R denoted $R[[X]]$ is the set of functions from G to R . If we denote by X^n the element of G that has a value n at X, then we can write an element in $R[[X]]$ as $\sum_{i=0}^{\infty} a_i\ X^i$ to denote the element that have a value a_i at X^i.*

One representation of formal power series is in the form of streams. For example the type of streams of elements of type `a` is be defined

```
Stream a = Cons a (Stream a)
```

Note that this type is non-well-founded. Such types are called Codatatype as opposed to the well-founded Datatypes[1]. The axiomatization of non-well-founded sets was given by Aczel[2].

The usual proof by induction does not apply to codatatypes. Instead the so-called coinductive proof principle based on the notion of bisimilarity (which we introduce in the next section) is used.

Remark 3.1.2. *Reasoning about non-well-founded sets is problematic for constructive mathematics. In fact their very existence is challenged by constructivists[2]. So*

[1] The prefix “co” originates from the category theoretic view of these types as final objects in the category of coalgebra of some functor.

[2] The following note is attributed to Kronecker “The general concept of an infinite series itself. for example a power series, is in my judgement permissible only with the reservation that in each

we cannot for example define them in Per Martin-Löf type theory. However, some extension allow for this, see for example [6].

3.2 The coinductive proof principle

The coinductive proof principle is usually laid out in category theoretic terms as follows[3].

Definition 3.2.1. *For a T-coalgebra $c : U \to T(U)$ a bisimulation on U is a relation R on U for which there exist a T-coalgebra structure $\gamma : R \to T(R)$ such that the two project functions $\pi_1 : R \to U$ and $pi_2 : R \to U$ are T-coalgebras homomorphism.[11]*

$$\begin{tikzcd}
U \arrow[d, "c"'] & R \arrow[l, "\pi_2"'] \arrow[r, "\pi_1"] \arrow[d, "\gamma"] & U \arrow[d, "c"] \\
T(U) & T(R) \arrow[l, "T(\pi_1)"] \arrow[r, "T(\pi_2)"'] & T(U)
\end{tikzcd}$$

Figure 3.1: commutes iff R is bisimulation

Theorem 3.2.2 (Coinductive proof principle)**.** *For a final T-coalgebra $c : Z \to T(Z)$, for all $z, z' \in Z$ if $R(z, z')$ for some bisimulation R then $z = z'$.[11]*

It is easy to see that $A^{\mathbb{N}}$ of streams of elements from a set A is a final colagebra of the functor $T(X) = A \times X$ with a colgebra structure given by $\langle head, tail \rangle : A^{\mathbb{N}} \to A \times A^{\mathbb{N}}$ where *head* and *tail* are the familiar functions on streams. A bisimulation on streams can then be defined as follows.

Definition 3.2.3. *A relation $R : A^{\mathbb{N}} \times A^{\mathbb{N}}$ is a bisimulation if*

$$\forall \alpha, \beta \in A^{\mathbb{N}}.\ \ R(\alpha, \beta) \to head(\alpha) = head(\beta) \wedge R(tail(\alpha), tail(\beta)).$$

particular case the arithmetical rule by which the terms are given satisfies conditions which make it possible to deal with the series as though it were finite, and thus make it unnecessary, strictly speaking, to go beyond the notion of a finite series."[15] p.71.

[3] If the reader is not familiar with category theory s/he can skip the definition 3.2.1 and theorem 3.2.2 and start from the definition of bisimulation for streams.

This is the principle we use to prove equalities of formal power series. Namely to show that two formal power series are equal we show that they are related by some bisimulation that we construct explicitly. In this case we say that the two series (streams) are bisimilar.

CHAPTER IV
The Ring of formal power series

A ring R can be represented as a record consisting of:

- An underlying set of elements $|R|$
- An equivalence relation $\overset{R}{\approx} \subseteq |R| \times |R|$
- Two binary operations $\overset{R}{+}, \overset{R}{*} \in |R| \times |R| \to |R|$, denoted addition and multiplication respectively.
- Two designated elements $0_R, 1_R \in |R|$ which are the additive and multiplicative identities respectively
- A unary operation $\overset{R}{-} \in |R| \to |R|$ which given an element in $|R|$ returns the additive inverse of that element

Along with a set of proofs of basic ring axioms (i.e. associativity of the binary operations, neutrality of identities with regard to their respective binary operations...etc).

In this chapter we give a construction of the ring of formal power series $R[[X]]$ over a given commutative ring R with decidable equality by using streams as the underlying datatype, i.e. the underlying elements set $R[[X]]$ is the type (Stream $|R|$), the set of all streams with elements in $|R|$. We also prove the correctness of our construction by coinduction.

Notation 4.0.4. *We adorn the different operation (addition, multiplication,...etc) with the ring name (i.e. R, $R[[X]]$). We do the same for some ring elements, namely the identities to avoid confusion.* hd , tl *,*:*, and* map *are the usual head, tail, cons, and map functions (resp) on streams, we leave them undefined here. We also use the ring (i.e. $R[[X]]$) to denote both the ring itself and the underlying set of elements (i.e. $R[[X]]$).*

4.1 Equality $\overset{R[[X]]}{\approx}$

Definition 4.1.1. *Two formal power series $\alpha, \beta \in R[[X]]$ are equal if and only if there exist a bisimulation $\mathcal{R}$ such that $\alpha \mathrel{\mathcal{R}} \beta$* [1]

4.2 Addition $\overset{R[[X]]}{+}$

Definition 4.2.1 (Addition)**.** *The addition operator $\overset{R[[X]]}{+} \in R[[X]] \times R[[X]] \to R[[X]]$ is defined as follows*

$$\mathrm{hd}\ (\alpha \overset{R[[X]]}{+} \beta) = \mathrm{hd}\ \alpha \overset{R}{+} \mathrm{hd}\ \beta$$
$$\mathrm{tl}\ (\alpha \overset{R[[X]]}{+} \beta) = \mathrm{tl}\ \alpha \overset{R[[X]]}{+} \mathrm{tl}\ \beta$$

Definition 4.2.2. *The additive identity $0_{R[[X]]}$ is defined coinductively as $0_{R[[X]]} = 0_R : 0_{R[[X]]}$*

Definition 4.2.3. *The additive inverse function $\overset{R[[X]]}{-} \in R[[X]] \to R[[X]]$ is defined as*

$$\mathrm{hd}\ (\overset{R[[X]]}{-} \beta) =\overset{R}{-} \mathrm{hd}\ \beta$$
$$\mathrm{tl}\ (\overset{R[[X]]}{-} \beta) = \overset{R[[X]]}{-} \mathrm{tl}\ \beta$$

Now we prove some properties related to the operator $(\overset{R[[X]]}{+})$

Proposition 4.2.4 ($\overset{R[[X]]}{+}$ is associative)**.**

$$\forall \alpha, \beta, \gamma \in R[[X]].\ (\alpha \overset{R[[X]]}{+} \beta) \overset{R[[X]]}{+} \gamma \overset{R[[X]]}{\approx} \alpha \overset{R[[X]]}{+} (\beta \overset{R[[X]]}{+} \gamma)$$

Proof. Let $\mathcal{R}_{+assoc} = \{\langle(\alpha \overset{R[[X]]}{+} \beta) \overset{R[[X]]}{+} \gamma\ ,\ \alpha \overset{R[[X]]}{+} (\beta \overset{R[[X]]}{+} \gamma)\rangle \mid \alpha, \beta, \gamma \in R[[X]]\}$, if $\delta\ \mathcal{R}_{+assoc}\ \epsilon$, then $\delta = (\alpha \overset{R[[X]]}{+} \beta) \overset{R[[X]]}{+} \gamma$ and $\epsilon = \alpha \overset{R[[X]]}{+} (\beta \overset{R[[X]]}{+} \gamma)$ for some

[1] Of course, equality is not decidable in the ring $R[[X]]$.

$\alpha, \beta, \gamma \in R[[X]]$. Then

1	$\text{hd}\ \delta = (\text{hd}\ \alpha \overset{R}{+} \text{hd}\ \beta) \overset{R}{+} \text{hd}\ \gamma$	Def of $\overset{R[[X]]}{+}$
2	$\text{hd}\ \delta = \text{hd}\ \alpha \overset{R}{+} (\text{hd}\ \beta \overset{R}{+} \text{hd}\ \gamma)$	Associativity of $\overset{R}{+}$
3	$\text{hd}\ \delta = \text{hd}\ (\alpha \overset{R[[X]]}{+} (\beta \overset{R[[X]]}{+} \gamma)) = \text{hd}\ \epsilon$	2,Def of $\overset{R[[X]]}{+}$
4	$\text{tl}\ \delta = (\text{tl}\ \alpha \overset{R[[X]]}{+} \text{tl}\ \beta) \overset{R[[X]]}{+} \text{tl}\ \gamma$	Def of $\overset{R[[X]]}{+}$
5	$\text{tl}\ \epsilon = \text{tl}\ \alpha \overset{R[[X]]}{+} (\text{tl}\ \beta \overset{R[[X]]}{+} \text{tl}\ \gamma)$	Def of $\overset{R[[X]]}{+}$
6	$\text{tl}\ \delta\ \mathcal{R}_{+assoc}\ \text{tl}\ \epsilon$	4,5, Def of $\mathcal{R}_{+assoc}$

Having proved that $\mathcal{R}_{+assoc}$ is a bisimulation, by *cpp* the associativity result follows directly.

Proposition 4.2.5 ($\overset{R[[X]]}{+}$ is commutative)**.**

$$\forall \alpha, \beta \in R[[X]].\ (\alpha \overset{R[[X]]}{+} \beta) \overset{R[[X]]}{\approx} (\beta \overset{R[[X]]}{+} \alpha)$$

Proof. Let $\mathcal{R}_{+comm} = \{\langle (\alpha \overset{R[[X]]}{+} \beta)\ ,\ (\beta \overset{R[[X]]}{+} \alpha) \rangle \mid \alpha, \beta \in R[[X]]\}$. If $\delta\ \mathcal{R}_{+comm}\ \epsilon$ then $\delta = \alpha \overset{R[[X]]}{+} \beta$ and $\epsilon = \beta \overset{R[[X]]}{+} \alpha$ for some $\alpha, \beta \in R[[X]]$. Then

1	$\text{hd}\ \delta = \text{hd}\ \alpha \overset{R}{+} \text{hd}\ \beta$	Def of $\overset{R[[X]]}{+}$
2	$\text{hd}\ \delta = \text{hd}\ \beta \overset{R}{+} \text{hd}\ \alpha$	1, commutativity of $\overset{R}{+}$
3	$\text{hd}\ \delta = \text{hd}\ (\beta \overset{R[[X]]}{+} \alpha) = \text{hd}\ \epsilon$	2,by Def $\overset{R[[X]]}{+}$
4	$\text{tl}\ \delta = \text{tl}\ \alpha \overset{R[[X]]}{+} \text{tl}\ \beta$	by Def $\overset{R[[X]]}{+}$
5	$\text{tl}\ \epsilon = \text{tl}\ \beta \overset{R[[X]]}{+} \text{tl}\ \alpha$	Def $\overset{R[[X]]}{+}$
6	$(\text{tl}\ \delta)\ \mathcal{R}_{+assoc}\ (\text{tl}\ \epsilon)$	8 ,9 ,Def $\mathcal{R}_{+comm}$

Proposition 4.2.6 ($0_{R[[X]]}$ is the additive identity)**.**

$$\forall \alpha \in R[[X]].\ \alpha \overset{R[[X]]}{+} 0_{R[[X]]} \overset{R[[X]]}{\approx} \alpha$$

Proof. Let $\mathcal{R}_{+id} = \{\langle \alpha \overset{R[[X]]}{+} 0_{R[[X]]}\ ,\ \alpha \rangle \mid \alpha \in R[[X]]\}$.
If $\delta\ \mathcal{R}_{+id}\ \epsilon$ then $\delta = \epsilon \overset{R[[X]]}{+} 0_{R[[X]]}$, we have

1	$\text{hd}\ \delta = \text{hd}\ \epsilon \overset{R}{+} \text{hd}\ 0_{R[[X]]}$	Def of $\overset{R[[X]]}{+}$
2	$\text{hd}\ \delta = \text{hd}\ \epsilon \overset{R}{+} 0_R = \text{hd}\ \epsilon$	1, Def of $0_{R[[X]]}$
3	$\text{tl}\ \delta = \text{tl}\ \epsilon \overset{R[[X]]}{+} \text{tl}\ 0_{R[[X]]} = \text{tl}\ \epsilon \overset{R[[X]]}{+} 0_{R[[X]]}$	Def $\overset{R[[X]]}{+}$ and $0_{R[[X]]}$
4	$(\text{tl}\ \delta)\ \mathcal{R}_{+id}\ (\text{tl}\ \epsilon)$	3, Def of $\mathcal{R}_{+id}$

Proving that $0_{R[[X]]}$ is also a left identity is similar.

Proposition 4.2.7 ($\overset{R[[X]]}{-}$ is the additive inverse)**.**

$$\forall \alpha \in R[[X]].\ \alpha \overset{R[[X]]}{-} \alpha \overset{R[[X]]}{\approx} 0_{R[[X]]}$$

Proof. (Omitted). The proof is trivial and follows similar approach to the ones above.

Proposition 4.2.8 ($\overset{R[[X]]}{+}$ preserves equality)**.**

$$\forall \alpha, \beta, \gamma, \delta \in R[[X]].\ (\alpha \overset{R[[X]]}{\approx} \beta) \wedge (\gamma \overset{R[[X]]}{\approx} \delta) \to \alpha \overset{R[[X]]}{+} \gamma \overset{R[[X]]}{\approx} \beta \overset{R[[X]]}{+} \delta$$

Proof. The proof is trivial and is omitted.

4.3 Multiplication $\overset{R[[X]]}{*}$

Definition 4.3.1. *The multiplication operator* $\overset{R[[X]]}{*} \in R[[X]] \times R[[X]] \to R[[X]]$ *is defined*

$$\begin{aligned}
&\mathrm{hd}\ (\alpha \overset{R[[X]]}{*} \beta) = \mathrm{hd}\ \alpha \overset{R}{*} \mathrm{hd}\ \beta \\
&\mathrm{tl}\ (\alpha \overset{R[[X]]}{*} \beta) = ((\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \mathrm{tl}\ \beta) \overset{R[[X]]}{+} (\mathrm{tl}\ \alpha \overset{R[[X]]}{*} \beta)
\end{aligned}$$

Proposition 4.3.2 ($\overset{R[[X]]}{*}$ preserves equality)**.**

$$\forall \alpha, \beta, \gamma, \delta \in R[[X]].\ (\alpha \overset{R[[X]]}{\approx} \beta) \wedge (\gamma \overset{R[[X]]}{\approx} \delta) \to \alpha \overset{R[[X]]}{*} \gamma \overset{R[[X]]}{\approx} \beta \overset{R[[X]]}{*} \delta$$

We define $\mathcal{R}_{*pres\approx}$ *inductively as follows*

1. *if* $\alpha \overset{R[[X]]}{\approx} \beta$ *then* $\langle \alpha\ ,\ \beta \rangle \in \mathcal{R}_{*pres\approx}$

2. *if* $\alpha \overset{R[[X]]}{\approx} \beta \wedge \gamma \overset{R[[X]]}{\approx} \delta$
   then $\langle (\alpha \overset{R[[X]]}{*} \gamma)\ ,\ (\beta \overset{R[[X]]}{*} \delta) \rangle \in \mathcal{R}_{*pres\approx}$

3. *if* $\langle \alpha, \beta \rangle \in \mathcal{R}_{*pres\approx} \wedge \langle \gamma, \delta \rangle \in \mathcal{R}_{*pres\approx}$
   then $\langle \alpha \overset{R[[X]]}{+} \gamma\ ,\ \beta \overset{R[[X]]}{+} \delta \rangle \in \mathcal{R}_{*pres\approx}$

Proof. We proceed by doing a proof by induction on constructor clauses of $\mathcal{R}_{*pres\approx}$, the proof being trivial for the 1^{st}, we prove for the 2^{nd} and 3^{rd}.

proof for 2^{nd} clause

If $\zeta \; \mathcal{R}_{*pres\approx} \; \eta$ then $\zeta = \alpha \overset{R[[X]]}{*} \gamma$ and $\eta = \beta \overset{R[[X]]}{*} \delta$ where $\alpha \overset{R[[X]]}{\approx} \beta$ and $\gamma \overset{R[[X]]}{\approx} \delta$, then.

1	$\text{hd}\,\alpha = \text{hd}\,\beta \wedge \text{hd}\,\gamma = \text{hd}\,\delta$	Def of $\overset{R[[X]]}{\approx}$
2	$\text{tl}\,\alpha \overset{R[[X]]}{\approx} \text{tl}\,\beta \wedge \text{tl}\,\gamma \overset{R[[X]]}{\approx} \text{tl}\,\delta$	Def of $\overset{R[[X]]}{\approx}$
3	$\text{hd}\,\zeta = \text{hd}\,\alpha \overset{R}{*} \text{hd}\,\gamma = \text{hd}\,\beta \overset{R}{*} \text{hd}\,\delta = \text{hd}\,\eta$	1,$\overset{R}{*}$ pres equality
4	$\text{tl}\,\zeta = ((\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\,\gamma) \overset{R[[X]]}{+} (\text{tl}\,\alpha \overset{R[[X]]}{*} \gamma)$	Def of $\overset{R[[X]]}{*}$
5	$\text{tl}\,\eta = ((\text{hd}\,\beta : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\,\delta) \overset{R[[X]]}{+} (\text{tl}\,\beta \overset{R[[X]]}{*} \delta)$	Def of $\overset{R[[X]]}{*}$
6	$(\text{tl}\,\alpha \overset{R[[X]]}{*} \gamma) \; \mathcal{R}_{*pres\approx} \; (\text{tl}\,\beta \overset{R[[X]]}{*} \delta)$	2, $\mathcal{R}_{*pres\approx}{}^{2^{nd}}$
7	$(\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{\approx} (\text{hd}\,\beta : 0_{R[[X]]})$	from 1 trivially
8	$((\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\,\gamma) \; \mathcal{R}_{*pres\approx} \; ((\text{hd}\,\beta : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\,\delta)$	7,2, IH $\mathcal{R}_{+pres\approx}{}^{2^{nd}}$
9	$(\text{tl}\,\zeta) \; \mathcal{R}_{+pres\approx} \; (\text{tl}\,\eta)$	8,6, $\mathcal{R}_{*pres\approx}{}^{3^{rd}}$

<u>proof for 3^{rd} clause</u>

If $\zeta \; \mathcal{R}_{*pres\approx} \; \eta$ then $\zeta = \alpha \overset{R[[X]]}{+} \gamma$ and $\eta = \beta \overset{R[[X]]}{+} \delta$ where $\alpha \; \mathcal{R}_{*pres\approx} \; \beta$ and $\gamma \; \mathcal{R}_{*pres\approx} \; \delta$, then.

1	$\text{hd}\,\alpha = \text{hd}\,\beta \wedge \text{hd}\,\gamma = \text{hd}\,\delta$	by IH
2	$(\text{tl}\,\alpha) \; \mathcal{R}_{*pres\approx} \; (\text{tl}\,\beta) \; \wedge \; (\text{tl}\,\gamma) \; \mathcal{R}_{*pres\approx} \; (\text{tl}\,\delta)$	by IH
3	$\text{hd}\,\zeta = \text{hd}\,\alpha \overset{R}{+} \text{hd}\,\gamma \overset{R}{\approx} \text{hd}\,\beta \overset{R}{+} \text{hd}\,\delta = \text{hd}\,\eta$	1,Def of $\overset{R[[X]]}{+}$, $\overset{R}{+}$ pres equality
4	$\text{tl}\,\zeta = \text{tl}\,\alpha \overset{R[[X]]}{+} \text{tl}\,\gamma \wedge \text{tl}\,\eta = \text{tl}\,\beta \overset{R[[X]]}{+} \text{tl}\,\delta$	Def of $\overset{R[[X]]}{+}$
5	$\text{tl}\,\zeta \; \mathcal{R}_{*pres\approx} \; \text{tl}\,\eta$	4, 2, $\mathcal{R}_{*pres\approx}{}^{3^{rd}}$

Proposition 4.3.3 ($\overset{R[[X]]}{*}$ is distributive over $\overset{R[[X]]}{+}$).

$$\forall \alpha, \beta, \gamma \in R[[X]].\ \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{+} \gamma) \overset{R[[X]]}{\approx} (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{+} (\alpha \overset{R[[X]]}{*} \gamma)$$

Proof. We define $\mathcal{R}_{*dist+}$ inductively as follows

1. if $\alpha \overset{R[[X]]}{\approx} \beta$ then $\langle \alpha\ ,\ \beta \rangle \in \mathcal{R}_{*dist+}$
2. $\forall \alpha, \beta, \gamma \in R[[X]].\ \langle \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{+} \gamma)\ ,\ (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{+} (\alpha \overset{R[[X]]}{*} \gamma) \rangle \in \mathcal{R}_{*dist+}$
3. $\forall \alpha, \beta, \gamma, \delta \in R[[X]]$. if $\langle \alpha, \beta \rangle \in \mathcal{R}_{*dist+} \wedge \langle \gamma, \delta \rangle \in \mathcal{R}_{*dist+}$
   then $\langle \alpha \overset{R[[X]]}{+} \gamma\ ,\ \beta \overset{R[[X]]}{+} \delta \rangle \in \mathcal{R}_{*dist+}$

We now prove that $\mathcal{R}_{*dist+}$ is a bisimulation by induction, the proof for the 1^{st} clause is trivial (omitted). The proof for the 3^rd clause is similar to the one of proposition 4.3.2 and is omitted here.

proof for the 2^{nd} clause

Let $\delta\ \mathcal{R}_{*dist+}\ \epsilon$ where $\delta = \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{+} \gamma)$ and $\epsilon = (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{+} (\alpha \overset{R[[X]]}{*} \gamma)$ for some $\alpha, \beta, \gamma \in R[[X]]$, then

1	$\text{hd } \delta = \text{hd } \alpha \overset{R}{*} (\text{hd } \beta \overset{R}{+} \text{hd } \gamma)$	Def $\overset{R[[X]]}{+}$, $\overset{R[[X]]}{*}$
2	$\text{hd } \delta = (\text{hd } \alpha \overset{R}{*} \text{hd } \beta) \overset{R}{+} (\text{hd } \alpha \overset{R}{*} \text{hd } \gamma)$	1, $\overset{R}{*}$ dist over $\overset{R}{+}$
3	$\text{hd } \delta \overset{R}{\approx} \text{hd } (\alpha \overset{R[[X]]}{*} \beta) \overset{R}{+} \text{hd } (\alpha \overset{R[[X]]}{*} \gamma)$	2, Def of $\overset{R[[X]]}{*}$
4	$\text{hd } \delta \overset{R}{\approx} \text{hd } ((\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{+} (\alpha \overset{R[[X]]}{*} \gamma)) = \text{hd } \epsilon$	3, Def of $\overset{R[[X]]}{+}$

1	$\mathrm{tl}\ \delta = (\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} (\mathrm{tl}\ \beta \overset{R[[X]]}{+} \mathrm{tl}\ \gamma) \overset{R[[X]]}{+} \mathrm{tl}\ \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{+} \gamma)$	Def of $\overset{R[[X]]}{+}, \overset{R[[X]]}{*}$
2	$\mathrm{tl}\ \epsilon = \ (\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \mathrm{tl}\ \beta \overset{R[[X]]}{+} \mathrm{tl}\ \alpha \overset{R[[X]]}{*} \beta$ $\overset{R[[X]]}{+} (\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \mathrm{tl}\ \gamma \overset{R[[X]]}{+} \mathrm{tl}\ \alpha \overset{R[[X]]}{*} \gamma$	Def of $\overset{R[[X]]}{+}, \overset{R[[X]]}{*}$
3	$\mathrm{tl}\ \epsilon \overset{R[[X]]}{\approx} \ (\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \mathrm{tl}\ \beta \overset{R[[X]]}{+} (\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \mathrm{tl}\ \gamma$ $\overset{R[[X]]}{+} \mathrm{tl}\ \alpha \overset{R[[X]]}{*} \beta \overset{R[[X]]}{+} \mathrm{tl}\ \alpha \overset{R[[X]]}{*} \gamma$	11, $\overset{R[[X]]}{+}$ commutative
4	$\big((\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \mathrm{tl}\ \beta \overset{R[[X]]}{+} (\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \mathrm{tl}\ \gamma\big)$ $\mathcal{R}_{*dist+}$ $\big((\mathrm{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} (\mathrm{tl}\ \beta \overset{R[[X]]}{+} \mathrm{tl}\ \gamma)\big)$	$\mathcal{R}_{*dist+}{}^{1^{st}}$
5	$\big(\mathrm{tl}\ \alpha \overset{R[[X]]}{*} \beta \overset{R[[X]]}{+} \mathrm{tl}\ \alpha \overset{R[[X]]}{*} \gamma\big)\ \mathcal{R}_{*dist+}\ \big(\mathrm{tl}\ \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{+} \gamma)\big)$	$\mathcal{R}_{*dist+}{}^{1^{st}}$
6	$\mathrm{tl}\ \delta\ \mathcal{R}_{*dist+}\ \mathrm{tl}\ \epsilon$	4, 5, 3, 1, $\mathcal{R}_{*dist+}{}^{3^{rd}}$

We note that we used the fact that $\overset{R[[X]]}{*}$ preserves equality and $\overset{R[[X]]}{+}$ preserves equality implicitly many times in the above derivation. Also in the rewriting cases of tl ϵ and tl δ we implicitly used the property $x \overset{R[[X]]}{\approx} y \wedge y\ \mathcal{R}_{*dist+}\ z \to x\ \mathcal{R}_{*dist+}\ z$ for which the proof is also trivial. The left distributivity is similar.

Proposition 4.3.4 ($\overset{R[[X]]}{*}$ is associative)**.**

$$\forall \alpha, \beta, \gamma \in R[[X]].\ (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{*} \gamma \overset{R[[X]]}{\approx} \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{*} \gamma)$$

Proof. We define our relation $\mathcal{R}_{*assoc}$ inductively as follows

1. $\forall \alpha, \beta \in R[[X]].\ \langle \alpha\ ,\ \beta \rangle \in \mathcal{R}_{*assoc}$ if $\alpha \overset{R[[X]]}{\approx} \beta$

2. $\forall \alpha, \beta, \gamma \in R[[X]].\ \langle (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{*} \gamma\ ,\ \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{*} \gamma) \rangle \in \mathcal{R}_{*assoc}$

3. $\forall \alpha, \beta, \gamma, \delta \in R[[X]]$. if $\langle \alpha, \beta \rangle \in \mathcal{R}_{*assoc} \wedge \langle \gamma, \delta \rangle \in \mathcal{R}_{*assoc}$
   then $\langle \alpha \overset{R[[X]]}{+} \gamma \ , \ \beta \overset{R[[X]]}{+} \delta \rangle \in \mathcal{R}_{*assoc}$

Now we prove $\mathcal{R}_{*assoc}$ to be a bisimulation. The proof for elements generated by the 1^{st} and 3^{rd} clauses straight forward, we proceed with a proof for elements generated by the 2^{nd} clause.

Proof for 2^{nd} clause:

Let $\delta \ \mathcal{R}_{*assoc} \ \epsilon$, then $\delta = (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{*} \gamma$ and $\epsilon = \alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{*} \gamma)$ for some $\alpha, \beta, \gamma \in R[[X]]$

1	$\text{hd } \delta = (\text{hd } \alpha \overset{R}{*} \text{hd } \beta) \overset{R}{*} \text{hd } \gamma = \text{hd } \alpha \overset{R}{*} (\text{hd } \beta \overset{R}{*} \text{hd } \gamma)$	Def of $\overset{R[[X]]}{*}$, $\overset{R}{*}$ assoc
2	$\text{hd } \delta = \text{hd } (\alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{*} \gamma)) = \text{hd } \epsilon$	1, Def $\overset{R[[X]]}{*}$
3	$\text{tl } \delta = \ (\text{hd } (\alpha \overset{R[[X]]}{*} \beta) : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \gamma$ $\overset{R[[X]]}{+} (\text{tl } (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{*} \gamma)$	Def of $\overset{R[[X]]}{*}$
4	$\text{tl } \delta = \ (\text{hd } (\alpha \overset{R[[X]]}{*} \beta) : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \gamma$ $\overset{R[[X]]}{+} (\text{tl } (\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{*} \gamma)$	Def of $\overset{R[[X]]}{*}$
5	$\text{hd } (\alpha \overset{R[[X]]}{*} \beta) : 0_{R[[X]]} \overset{R[[X]]}{\approx} (\text{hd } \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} (\text{hd } \beta : 0_{R[[X]]})$	trivial
6	$\text{tl } \delta \overset{R[[X]]}{\approx} \ ((\text{hd } \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} (\text{hd } \beta : 0_{R[[X]]})) \overset{R[[X]]}{*} \text{tl } \gamma$ $\overset{R[[X]]}{+} ((\text{hd } \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \beta) \overset{R[[X]]}{*} \gamma$ $\overset{R[[X]]}{+} (\text{tl } \alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{*} \gamma$	4,5, $\overset{R[[X]]}{*}$ *dist* $\overset{R[[X]]}{+}$

7	$\text{tl}\,\epsilon \overset{R[[X]]}{\approx} (\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} ((\text{hd}\,\beta : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\,\gamma)$ $\overset{R[[X]]}{+} (\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} (\text{tl}\,\beta \overset{R[[X]]}{*} \gamma)$ $\overset{R[[X]]}{+} \text{tl}\,\alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{*} \gamma)$	similar to 6
8	$\big((\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} (\text{hd}\,\beta : 0_{R[[X]]})\big) \overset{R[[X]]}{*} \text{tl}\,\gamma$ $\mathcal{R}_{*assoc}$ $(\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \big((\text{hd}\,\beta : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\,\gamma\big)$	by $\mathcal{R}_{*assoc}{}^{2^{nd}}$
9	$\big(((\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\,\beta) \overset{R[[X]]}{*} \gamma\big)$ $\mathcal{R}_{*assoc}$ $\big((\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} (\text{tl}\,\beta \overset{R[[X]]}{*} \gamma)\big)$	by $\mathcal{R}_{*assoc}{}^{2^{nd}}$
10	$\big((\text{tl}\,\alpha \overset{R[[X]]}{*} \beta) \overset{R[[X]]}{*} \gamma\big)\ \mathcal{R}_{*assoc}\ \big(\text{tl}\,\alpha \overset{R[[X]]}{*} (\beta \overset{R[[X]]}{*} \gamma)\big)$	by $\mathcal{R}_{*assoc}{}^{2^{nd}}$
11	$\text{tl}\,\delta\ \mathcal{R}_{*assoc}\ \text{tl}\,\epsilon$	8,9,10, by $\mathcal{R}_{*assoc}{}^{3^{rd}}$

Proposition 4.3.5 (multiplicative identity)**.** $1_{R[[X]]} = 1_R : 1_{R[[X]]}$ *is the multiplicative identity.*

Proof. Proof is trivial, omitted.

Proposition 4.3.6. $0_{R[[X]]}$ *is annihilator of multiplication*

Proof. Proof is trivial, omitted.

Lemma 4.3.7. *Let* $\mathcal{L}$ *be defined as follows*

1. $\forall \alpha, \beta \in R[[X]]$
   $\langle\big((\text{hd}\,\alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \beta \overset{R[[X]]}{+} \text{tl}\,\alpha \overset{R[[X]]}{*} (0_R : \beta)\big), \alpha \overset{R[[X]]}{*} \beta\rangle \in \mathcal{L}$

2. $\forall \alpha, \beta, \gamma, \delta \in R[[X]].$ *if* $\langle \alpha, \beta \rangle \in \mathcal{L} \wedge \langle \gamma, \delta \rangle \in \mathcal{L}$
then $\langle \alpha \overset{R[[X]]}{+} \gamma \, , \, \beta \overset{R[[X]]}{+} \delta \rangle \in \mathcal{L}$

Then $\mathcal{L}$ *is a bisimulation*

Proof. The second clause being straight forward we prove only for the 1^{st} clause. Let $\delta \, \mathcal{L} \, \epsilon$ such that $\delta = (\text{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \beta \overset{R[[X]]}{+} \text{tl}\ \alpha \overset{R[[X]]}{*} (0_R : \beta)$ and $\epsilon = \alpha \overset{R[[X]]}{*} \beta$, then

1	$\text{hd}\ \delta \overset{R[[X]]}{\approx} \text{hd}\ \alpha \overset{R}{*} \text{hd}\ \beta = \text{hd}\ \epsilon$	Def of $\overset{R[[X]]}{+*}$, 0_R kills $\overset{R}{*}$
2	$\text{tl}\ \delta \overset{R[[X]]}{\approx} (\text{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\ \beta$ $\overset{R[[X]]}{+} (\text{hd tl}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \beta$ $\overset{R[[X]]}{+} \text{tl tl}\ \alpha \overset{R[[X]]}{*} (0_R : \beta)$	Def of $\overset{R[[X]]}{*}$, $0_{R[[X]]}$ kills $\overset{R[[X]]}{*}$
3	$(\text{tl}\ \delta) \, \mathcal{L} \, \big((\text{hd}\ \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl}\ \beta \overset{R[[X]]}{+} (\text{tl}\ \alpha \overset{R[[X]]}{*} \beta)\big)$	2, by $\mathcal{L}^{1^{st}}$, $\mathcal{L}^{2^{nd}}$
4	$(\text{tl}\ \delta) \, \mathcal{L} \, (\text{tl}\ \epsilon)$	3, Def $\overset{R[[X]]}{*}$

Proposition 4.3.8 ($\overset{R[[X]]}{*}$ is commutative)**.** *If* R *is commutative then*

$$\forall \alpha, \beta \in R[[X]]. \ \alpha \overset{R[[X]]}{*} \beta \overset{R[[X]]}{\approx} \beta \overset{R[[X]]}{*} \alpha$$

Proof. We define $\mathcal{R}_{*comm}$ inductively as follows

1. $\forall \alpha, \beta \in R[[X]].$ if $\alpha \overset{R[[X]]}{\approx} \beta$ then $\langle \alpha \, , \, \beta \rangle \in \mathcal{R}_{*comm}$

2. $\forall \alpha, \beta \in R[[X]]. \ \langle \alpha \overset{R[[X]]}{*} \beta \, , \, \beta \overset{R[[X]]}{*} \alpha \rangle \in \mathcal{R}_{*comm}$

3. $\forall \alpha, \beta, \gamma, \delta \in R[[X]].$ if $(\alpha \, \mathcal{R}_{*comm} \, \beta) \wedge (\gamma \, \mathcal{R}_{*comm} \, \delta)$ then $\langle \alpha \overset{R[[X]]}{+} \gamma \, , \, \beta \overset{R[[X]]}{+} \delta \rangle \in \mathcal{R}_{*comm}$

We proceed by induction as usual, The proofs for the 1^{st} and 3^{rd} clauses are straight forward and are omitted here.

Proof for the 2^{nd} clause

Let $\delta \mathcal{R}_{*assoc} \epsilon$ such that $\delta = \alpha \overset{R[[X]]}{*} \beta$ and $\epsilon = \beta \overset{R[[X]]}{*} \alpha$

1	$\text{hd } \delta = (\text{hd } \alpha \overset{R}{*} \text{hd } \beta)$	Def of $\overset{R[[X]]}{*}$
2	$\text{hd } \delta = \text{hd } \beta \overset{R}{*} \text{hd } \alpha = \text{hd } \epsilon$	1,$\overset{R}{*}$ commutative
3	$\text{tl } \delta = (\text{hd } \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \beta \overset{R[[X]]}{+} \text{tl } \alpha \overset{R[[X]]}{*} \beta$	Def of $\overset{R[[X]]}{*}$
4	$\text{tl } \delta \overset{R[[X]]}{\approx} (\text{hd } \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \beta$ $\overset{R[[X]]}{+} \text{tl } \alpha \overset{R[[X]]}{*} (\text{hd } \beta : 0_{R[[X]]} \overset{R[[X]]}{+} (0_R : \text{tl } \beta))$	3, $\overset{R[[X]]}{+}$ pres $\overset{R[[X]]}{\approx}$
5	$\text{tl } \delta \overset{R[[X]]}{\approx} (\text{hd } \alpha : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \beta$ $\overset{R[[X]]}{+} \text{tl } \alpha \overset{R[[X]]}{*} (\text{hd } \beta : 0_{R[[X]]})$ $\overset{R[[X]]}{+} \text{tl } \alpha \overset{R[[X]]}{*} (0_R : \text{tl } \beta))$	4, $\overset{R[[X]]}{*}$ dist $\overset{R[[X]]}{+}$
6	$\text{tl } \delta \overset{R[[X]]}{\approx} \text{tl } \alpha \overset{R[[X]]}{*} (\text{hd } \beta : 0_{R[[X]]}) \overset{R[[X]]}{+} \alpha \overset{R[[X]]}{*} \text{tl } \beta$	$\overset{R[[X]]}{+}$ *comm*, $\mathcal{L}$ bisimulation
7	$(\text{tl } \alpha \overset{R[[X]]}{*} (\text{hd } \beta : 0_{R[[X]]})) \mathcal{R}_{*comm} ((\text{hd } \beta : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \alpha)$	IH $\mathcal{R}_{*comm}{}^{2^{nd}}$
8	$(\alpha \overset{R[[X]]}{*} \text{tl } \beta) \mathcal{R}_{*comm} (\alpha \overset{R[[X]]}{*} \text{tl } \beta)$	IH $\mathcal{R}_{*comm}{}^{2^{nd}}$
9	$\text{tl } \epsilon = (\text{hd } \beta : 0_{R[[X]]}) \overset{R[[X]]}{*} \text{tl } \alpha \overset{R[[X]]}{+} \text{tl } \beta \overset{R[[X]]}{*} \alpha$	Def of $\overset{R[[X]]}{*}$
10	$(\text{tl } \delta) \mathcal{R}_{*comm} (\text{tl } \epsilon)$	6,7,8,9, IH $\mathcal{R}_{*comm}{}^{3^{rd}}$

Implementation note IV.1 (Monads and streams)**:** The type of streams in haskell is

```
 data Stream a = Cons a (Stream a)
```

In the first implementation the type of power series over a ring a looked like this

```
newtype CommutativeRing a => PSeries a x = PS (Stream a) deriving (Ord)
```

However, this type failed to work properly with the type `type R k = MPoly k Len`, i.e. the type dynamic algebraic closure of a field `k`. To see this, imagine that we want to map a monadic operation on the stream type. The map function looks as follows

```
mapMS :: (Monad m) => (a -> m b) -> Stream a -> m (Stream b)
mapMS f g = do y  <- f (head g)
               ys <- map f (tail g)
               return $ Cons (y, ys)
```

Since the bind operator `>>=` of the state monad (see note II.1) forces the inspection of each object in the list, this `mapMS` function would not terminate if called on a monadic operation `f` that branches. This forced us to use a different datatype for streams

```
 newtype (Monad m)=>  Stream m a = Cons (a ,m (Stream m a))
 newtype (CommutativeRing a, Monad m) => PSeries m a x = PS (Stream m a)
```

With this type we can define the monadic `map` as follows

```
mapMS :: (Monad m) => (a -> m b) -> Stream m a -> m (Stream m b)
mapMS f g = do y <- f (head g)
               ys <- tail g
               return $ Cons (y, mapMS f ys)
```

In which the computation on the tail of the stream is delayed.

The two implementations are now part of the library.

CHAPTER V

Hensel's lemma

The statement of Hensel's lemma is as follows

Theorem 5.0.1 (Hensel's lemma)**.** *Let k be a field and let*

$$F(X,Y) = Y^n + a_1(X)Y^{n-1} + ... + a_n(X) \quad \in k[[X]][Y]$$

be a monic polynomial of degree $n > 0$ in Y with coefficients $a_i(X) \in k[[X]]$. Assume that

$$F(0,Y) = \bar{G}(Y)\bar{H}(Y)$$

and

$$\bar{G}(Y) = Y^r + \bar{b}_1 Y^{r-1} + ... + \bar{b}_r \quad \in k[Y]$$
$$\bar{H}(Y) = Y^s + \bar{c}_1 Y^{s-1} + ... + \bar{c}_s \quad \in k[Y]$$

are monic polynomials of degrees $r, s > 0$ in Y with coefficients in k. Such that $\bar{G}(Y)$ and $\bar{H}(Y)$ are coprime.
Then there exist unique monic polynomials

$$G(X,Y) = Y^r + b_1(X)Y^{r-1} + ... + b_r(X) \quad \in k[[X]][Y]$$
$$H(X,Y) = Y^s + c_1(X)Y^{s-1} + ... + c_r(X) \quad \in k[[X]][Y]$$

of degrees r and s in Y with coefficients in $k[[X]]$, such that $G(0,Y) = \bar{G}(Y)$ and $H(0,Y) = \bar{H}(Y)$ and $F(X,Y) = G(X,Y)H(X,Y)$.

Proof. . 5.3

In the following sections we give a formal proof of the lemma. In the formalization we follow a proof presented by S.S.Abhyankar [1].

Remark 5.0.2. *For the purpose of this document we assume a complete formalization of the ring of polynomials over commutative rings. Also, a forall quantification is assumed for free variables.*

Notation 5.0.3. *To distinguish between list and stream functions that usually have the same name, we adorn the stream functions with* $\widetilde{\ }$ *. For example,* $\mathrm{tl}\ , \mathrm{hd}\ , \mathrm{map}$ *are the familiar tail,head, and map operations on lists while their stream counterparts are* $\widetilde{\mathrm{tl}}\ , \widetilde{\mathrm{hd}}\ , \widetilde{\mathrm{map}}\ .$

5.1 The homomorphism $k[[X]][Y] \to k[Y][[X]]$

Definition 5.1.1. *The mapping* $\phi : k[[X]][Y] \to k[Y][[X]]$ *is defined as follows*

$$\begin{aligned}\widetilde{\mathrm{hd}}\ \phi(as) &= \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ as \\ \widetilde{\mathrm{tl}}\ \phi(as) &= \phi(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ as)\end{aligned}$$

We prove some properties for the $(\mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ )$ and $(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ )$

Proposition 5.1.2. $(\mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ ) : k[[X]][Y] \to k[Y]$ *is* $\overset{k[[X]][Y]}{+}$ *homomorphism*

$$\mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ (as \overset{k[[X]][Y]}{+} bs) \overset{k[Y]}{\approx} \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ as \overset{k[Y]}{+} \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ bs$$

Proof. By induction on $\overset{k[[X]][Y]}{+}$

<u>Base case:</u> $[a] \overset{k[[X]][Y]}{+} [b] = [a \overset{K[[X]]}{+} b]$

$$\begin{array}{r|lr}
1 & \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ ([a] \overset{k[[X]][Y]}{+} [b]) = \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ [a \overset{K[[X]]}{+} b] & \text{Def } \overset{k[[X]][Y]}{+} \\
2 & \quad = [\widetilde{\mathrm{hd}}\ (a \overset{K[[X]]}{+} b)] = [\widetilde{\mathrm{hd}}\ a \overset{K}{+} \widetilde{\mathrm{hd}}\ b] & \text{Def } \overset{K[[X]]}{+}, \mathrm{map} \\
3 & \quad = [\widetilde{\mathrm{hd}}\ a] \overset{k[Y]}{+} [\widetilde{\mathrm{hd}}\ b] = \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ [a] \overset{k[Y]}{+} \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ [b] & \text{Def } \overset{k[Y]}{+}, \mathrm{map}
\end{array}$$

Base case: $a : as \overset{k[[X]][Y]}{+} [b] = (a \overset{K[[X]]}{+} b) : as$

$$\begin{array}{r|ll}
4 & \text{map } \widetilde{\text{hd}} \ (a : as \overset{k[[X]][Y]}{+} [b]) = \text{map } \widetilde{\text{hd}} \ ((a \overset{K[[X]]}{+} b) : as) & \text{Def } \overset{k[[X]][Y]}{+} \\
5 & \quad = \widetilde{\text{hd}} \, (a \overset{K[[X]]}{+} b) : (\text{map } \widetilde{\text{hd}} \ as) = (\widetilde{\text{hd}} \, a \overset{K}{+} \widetilde{\text{hd}} \, b) : (\text{map } \widetilde{\text{hd}} \ as) & \text{Def } \overset{K[[X]]}{+}, \text{map} \\
6 & \quad = \widetilde{\text{hd}} \, a : (\text{map } \widetilde{\text{hd}} \ as) \overset{k[Y]}{+} [\widetilde{\text{hd}} \, b] & \text{Def } \overset{k[Y]}{+} \\
7 & \quad = \text{map } \widetilde{\text{hd}} \ (a : as) \overset{k[Y]}{+} \text{map } \widetilde{\text{hd}} \ [b] & \text{Def map}
\end{array}$$

Base case: $[a] \overset{k[[X]][Y]}{+} b : bs = (a \overset{K[[X]]}{+} b) : bs$

Similar to the previous case

Step: $(a : as) \overset{k[[X]][Y]}{+} (b : bs) = (a \overset{K[[X]]}{+} b) : (as \overset{k[[X]][Y]}{+} bs)$

$$\begin{array}{r|ll}
8 & \text{map } \widetilde{\text{hd}} \ (a : as \overset{k[[X]][Y]}{+} b : bs) & \\
9 & \quad = \text{map } \widetilde{\text{hd}} \ ((a \overset{K[[X]]}{+} b) : (as \overset{k[[X]][Y]}{+} bs)) & \text{Def } \overset{k[[X]][Y]}{+} \\
10 & \quad = \widetilde{\text{hd}} \, (a \overset{K[[X]]}{+} b) : (\text{map } \widetilde{\text{hd}} \ (as \overset{k[[X]][Y]}{+} bs)) & \text{Def map} \\
11 & \quad = (\widetilde{\text{hd}} \, a \overset{K}{+} \widetilde{\text{hd}} \, b) : (\text{map } \widetilde{\text{hd}} \ (as \overset{k[[X]][Y]}{+} bs)) & \text{Def } \overset{K[[X]]}{+} \\
12 & \quad = (\widetilde{\text{hd}} \, a \overset{K}{+} \widetilde{\text{hd}} \, b) : (\text{map } \widetilde{\text{hd}} \ as \overset{k[Y]}{+} \text{map } \widetilde{\text{hd}} \ bs) & \text{Induction Hyp (IH)} \\
13 & \quad = (\widetilde{\text{hd}} \, a : \text{map } \widetilde{\text{hd}} \ as) \overset{k[Y]}{+} (\widetilde{\text{hd}} \, b : \text{map } \widetilde{\text{hd}} \ bs) & \text{Def } \overset{k[Y]}{+} \\
14 & \quad = \text{map } \widetilde{\text{hd}} \ (a : as) \overset{k[Y]}{+} \text{map } \widetilde{\text{hd}} \ (b : bs) & \text{Def map}
\end{array}$$

Proposition 5.1.3. $(\text{map } \widetilde{\text{tl}} \ ) : k[[X]][Y] \to k[[X]][Y]$ *is* $\overset{k[[X]][Y]}{+}$ *homomorphism*

$$\text{map } \widetilde{\text{tl}} \ (as \overset{k[[X]][Y]}{+} bs) \overset{k[[X]][Y]}{\approx} \text{map } \widetilde{\text{tl}} \ as \overset{k[[X]][Y]}{+} \text{map } \widetilde{\text{tl}} \ bs$$

Proof. By induction on $\overset{k[[X]][Y]}{+}$

<u>Base case:</u> $[a] \overset{k[[X]][Y]}{+} [b] = [a \overset{K[[X]]}{+} b]$

$$\begin{array}{r|ll}
1 & \text{map } \widetilde{\text{tl}} \ ([a] \overset{k[[X]][Y]}{+} [b]) = \text{map } \widetilde{\text{tl}} \ [a \overset{K[[X]]}{+} b] & \text{Def } \overset{k[[X]][Y]}{+} \\
2 & \quad = [\widetilde{\text{tl}}\,(a \overset{K[[X]]}{+} b)] = [(\widetilde{\text{tl}}\,a \overset{K[[X]]}{+} \widetilde{\text{tl}}\,b)] & \text{Def } \overset{K[[X]]}{+}, \text{map} \\
3 & \quad = [\widetilde{\text{tl}}\,a] \overset{k[[X]][Y]}{+} [\widetilde{\text{tl}}\,b] = \text{map } \widetilde{\text{tl}} \ [a] \overset{k[[X]][Y]}{+} \text{map } \widetilde{\text{tl}} \ [b] & \text{Def } \overset{k[[X]][Y]}{+}, \text{map}
\end{array}$$

<u>Base case:</u> $a : as \overset{k[[X]][Y]}{+} [b] = (a \overset{K[[X]]}{+} b) : as$

$$\begin{array}{r|ll}
4 & \text{map } \widetilde{\text{tl}} \ (a : as \overset{k[[X]][Y]}{+} [b]) = \text{map } \widetilde{\text{tl}} \ ((a \overset{K[[X]]}{+} b) : as) & \text{Def } \overset{k[[X]][Y]}{+} \\
5 & \quad = \widetilde{\text{tl}}\,(a \overset{K[[X]]}{+} b) : (\text{map } \widetilde{\text{tl}} \ as) = (\widetilde{\text{tl}}\,a \overset{K[[X]]}{+} \widetilde{\text{tl}}\,b) : (\text{map } \widetilde{\text{tl}} \ as) & \text{Def } \overset{K[[X]]}{+}, \text{map} \\
6 & \quad = (\widetilde{\text{tl}}\,a : \text{map } \widetilde{\text{tl}} \ as) \overset{k[[X]][Y]}{+} [\widetilde{\text{tl}}\,b] & \text{Def } \overset{k[[X]][Y]}{+} \\
7 & \quad = \text{map } \widetilde{\text{tl}} \ (a : as) \overset{k[[X]][Y]}{+} \text{map } \widetilde{\text{tl}} \ [b] & \text{Def map}
\end{array}$$

<u>Base case:</u> $[a] \overset{k[[X]][Y]}{+} b : bs = (a \overset{K[[X]]}{+} b) : bs$

Similar to the previous case

Step: $(a : as) \overset{k[[X]][Y]}{+} (b : bs) = (a \overset{K[[X]]}{+} b) : (as \overset{k[[X]][Y]}{+} bs)$

8 $\text{map } \widetilde{\text{tl}} \ (a : as \overset{k[[X]][Y]}{+} b : bs)$

9 $= \text{map } \widetilde{\text{tl}} \ ((a \overset{K[[X]]}{+} b) : (as \overset{k[[X]][Y]}{+} bs))$ Def $\overset{k[[X]][Y]}{+}$

10 $= \widetilde{\text{tl}} \, (a \overset{K[[X]]}{+} b) : (\text{map } \widetilde{\text{tl}} \ (as \overset{k[[X]][Y]}{+} bs))$ Def map

11 $= (\widetilde{\text{tl}} \, a \overset{K[[X]]}{+} \widetilde{\text{tl}} \, b) : (\text{map } \widetilde{\text{tl}} \ (as \overset{k[[X]][Y]}{+} bs))$ Def $\overset{K[[X]]}{+}$

12 $= (\widetilde{\text{tl}} \, a \overset{K[[X]]}{+} \widetilde{\text{tl}} \, b) : (\text{map } \widetilde{\text{tl}} \ as \overset{k[[X]][Y]}{+} \text{map } \widetilde{\text{tl}} \ bs)$ (IH)

13 $= (\widetilde{\text{tl}} \, a : \widetilde{\text{tl}} \, as) \overset{k[[X]][Y]}{+} (\widetilde{\text{tl}} \, b : \text{map } \widetilde{\text{tl}} \ bs)$ Def $\overset{k[[X]][Y]}{+}$

14 $= \text{map } \widetilde{\text{tl}} \ (a : as) \overset{k[[X]][Y]}{+} \text{map } \widetilde{\text{tl}} \ (b : bs)$ Def map

Proposition 5.1.4. (map $\widetilde{\text{hd}}$) $: k[[X]][Y] \to k[Y]$ *is* $\overset{k[[X]][Y]}{*}$ *homomorphism*

$$\text{map } \widetilde{\text{hd}} \ (as \overset{k[[X]][Y]}{*} bs) \overset{k[Y]}{\approx} \text{map } \widetilde{\text{hd}} \ as \overset{k[Y]}{*} \text{map } \widetilde{\text{hd}} \ bs$$

Proof. By induction on $\overset{k[[X]][Y]}{*}$

Base case: $[a] \overset{k[[X]][Y]}{*} [b] = [a \overset{K[[X]]}{*} b]$

1 $\text{map } \widetilde{\text{hd}} \ ([a] \overset{k[[X]][Y]}{*} [b]) = \text{map } \widetilde{\text{hd}} \ [a \overset{K[[X]]}{*} b]$ Def $\overset{k[[X]][Y]}{*}$

2 $= [\widetilde{\text{hd}} \, (a \overset{K[[X]]}{*} b)] = [(\widetilde{\text{hd}} \, a \overset{K}{*} \widetilde{\text{hd}} \, b)]$ Def $\overset{K[[X]]}{*}$, map

3 $= [\widetilde{\text{hd}} \, a] \overset{k[Y]}{*} [\widetilde{\text{hd}} \, b] = \text{map } \widetilde{\text{hd}} \ [a] \overset{k[Y]}{*} \text{map } \widetilde{\text{hd}} \ [b]$ Def $\overset{k[Y]}{*}$, map

Step : $a : as \overset{k[[X]][Y]}{*} [b] = (a \overset{K[[X]]}{*} b) : (as \overset{k[[X]][Y]}{*} [b])$

$$\begin{array}{r|ll}
4 & \text{map } \widetilde{\text{hd}} \ (a : as \overset{k[[X]][Y]}{*} [b]) & \\
5 & \quad = \text{map } \widetilde{\text{hd}} \ ((a \overset{K[[X]]}{*} b) : (as \overset{k[[X]][Y]}{*} [b])) & \text{Def } \overset{k[[X]][Y]}{*} \\
6 & \quad = \widetilde{\text{hd}} \ (a \overset{K[[X]]}{*} b) : \text{map } \widetilde{\text{hd}} \ (as \overset{k[[X]][Y]}{*} [b]) & \text{Def map} \\
7 & \quad = (\widetilde{\text{hd}} \, a \overset{K}{*} \widetilde{\text{hd}} \, b) : \text{map } \widetilde{\text{hd}} \ (as \overset{k[[X]][Y]}{*} [b]) & \text{Def } \overset{K[[X]]}{*} \\
8 & \quad = (\widetilde{\text{hd}} \, a \overset{K}{*} \widetilde{\text{hd}} \, b) : (\text{map } \widetilde{\text{hd}} \ as) \overset{k[Y]}{*} [\widetilde{\text{hd}} \, b]) & \text{(IH), Def map} \\
9 & \quad = (\widetilde{\text{hd}} \, a : \text{map } \widetilde{\text{hd}} \ as) \overset{k[Y]}{*} [\widetilde{\text{hd}} \, b] & \text{Def } \overset{k[Y]}{*} \\
10 & \quad = \text{map } \widetilde{\text{hd}} \ (a : as) \overset{k[Y]}{*} \text{map } \widetilde{\text{hd}} \ [b] & \text{Def map}
\end{array}$$

Step : $[a] \overset{k[[X]][Y]}{*} b : bs = (a \overset{K[[X]]}{*} b) : bs$

Similar to the previous case

Step:

$$\begin{aligned}
(a : as) \overset{k[[X]][Y]}{*} (b : bs) = & [a \overset{K[[X]]}{*} b] \\
& \overset{k[[X]][Y]}{+} (0_{K[[X]]} : [a] \overset{k[[X]][Y]}{*} bs) \\
& \overset{k[[X]][Y]}{+} (0_{K[[X]]} : as \overset{k[[X]][Y]}{*} [b]) \\
& \overset{k[[X]][Y]}{+} (0_{K[[X]]} : 0_{K[[X]]} : as \overset{k[[X]][Y]}{*} bs)
\end{aligned}$$

11	$\text{map } \widetilde{\text{hd}} \ (a : as \overset{k[[X]][Y]}{*} b : bs)$	
12	$= \text{map } \widetilde{\text{hd}} \ [a \overset{K[[X]]}{*} b]$ $\overset{k[Y]}{+} \text{map } \widetilde{\text{hd}} \ (0_{K[[X]]} : [a] \overset{k[[X]][Y]}{*} bs)$ $\overset{k[Y]}{+} \text{map } \widetilde{\text{hd}} \ (0_{K[[X]]} : as \overset{k[[X]][Y]}{*} [b])$ $\overset{k[Y]}{+} \text{map } \widetilde{\text{hd}} \ (0_{K[[X]]} : 0_{K[[X]]} : as \overset{k[[X]][Y]}{*} bs)$	map $\widetilde{\text{hd}}$ is $\overset{k[[X]][Y]}{+}$ hom
13	$= [\widetilde{\text{hd}}\, a \overset{K}{*} \widetilde{\text{hd}}\, b]$ $\overset{k[Y]}{+} 0_k : \text{map } \widetilde{\text{hd}} \ ([a] \overset{k[[X]][Y]}{*} bs)$ $\overset{k[Y]}{+} 0_k : \text{map } \widetilde{\text{hd}} \ (as \overset{k[[X]][Y]}{*} [b])$ $\overset{k[Y]}{+} 0_k : 0_k : \text{map } \widetilde{\text{hd}} \ (as \overset{k[[X]][Y]}{*} bs)$	Def map , $\overset{K[[X]]}{*}$
14	$= [\widetilde{\text{hd}}\, a \overset{K}{*} \widetilde{\text{hd}}\, b]$ $\overset{k[Y]}{+} 0_k : ([\widetilde{\text{hd}}\, a] \overset{k[Y]}{*} \text{map } \widetilde{\text{hd}} \ bs)$ $\overset{k[Y]}{+} 0_k : (\text{map } \widetilde{\text{hd}} \ as \overset{k[Y]}{*} [\widetilde{\text{hd}}\, b])$ $\overset{k[Y]}{+} 0_k : 0_k : (\text{map } \widetilde{\text{hd}} \ as \overset{k[Y]}{*} \text{map } \widetilde{\text{hd}} \ bs)$	(IH)
15	$= (\widetilde{\text{hd}}\, a : \text{map } \widetilde{\text{hd}} \ as) \overset{k[Y]}{*} (\widetilde{\text{hd}}\, b : \text{map } \widetilde{\text{hd}} \ bs)$	Def $\overset{k[Y]}{*}$
16	$= \text{map } \widetilde{\text{hd}} \ (a : as) \overset{k[Y]}{*} \text{map } \widetilde{\text{hd}} \ (b : bs)$	Def map

Proposition 5.1.5 (a property of map $\widetilde{\text{tl}}$).

$$\text{map } \widetilde{\text{tl}} \ (as \overset{k[[X]][Y]}{*} bs) = \text{map } (\lambda z \to \text{hd } z : 0_{K[[X]]}) \ as \overset{k[[X]][Y]}{*} \text{map } \widetilde{\text{tl}} \ bs \overset{k[[X]][Y]}{+} \text{map } \widetilde{\text{tl}} \ as \overset{k[[X]][Y]}{*} bs$$

Proof. the proof is quite complicated, we postpone it for the appendix.

Proposition 5.1.6. $\phi(0_{k[[X]][Y]}) \overset{k[Y][[X]]}{\approx} 0_{k[Y][[X]]}$

Proof. let $\mathcal{R}_0 = \{\langle \phi(0_{k[[X]][Y]}), 0_{k[Y][[X]]} \rangle\}$

1	$\widetilde{\text{hd}}\ \phi(0_{k[[X]][Y]}) = \widetilde{\text{hd}}\ \phi([0_{K[[X]]}]) = [\widetilde{\text{hd}}\ 0_{K[[X]]}]$	Def of $\phi, 0_{k[[X]][Y]}$
2	$= [0_k] = \widetilde{\text{hd}}\ 0_{k[Y][[X]]}$	1, Def $0_k, 0_{k[Y][[X]]}$
3	$\widetilde{\text{tl}}\ \phi(0_{k[[X]][Y]}) = \phi([\widetilde{\text{tl}}\ 0_{K[[X]]}]) = \phi([0_{K[[X]]}]) = \phi(0_{k[[X]][Y]})$	Def of $\phi, 0_{k[[X]][Y]}$
4	$\widetilde{\text{tl}}\ 0_{k[Y][[X]]} = 0_{k[Y][[X]]}$	Def of $0_{k[Y][[X]]}$
5	$\widetilde{\text{tl}}\ \phi(0_{k[[X]][Y]})\ \mathcal{R}_0\ \widetilde{\text{tl}}\ 0_{k[Y][[X]]}$	3,4

Proposition 5.1.7. $\phi(1_{k[[X]][Y]}) \overset{k[Y][[X]]}{\approx} 1_{k[Y][[X]]}$

Proof. similar to proof of proposition 5.1.6 (omitted).

Proposition 5.1.8 (ϕ additive homomorphism)**.**

$$\phi(f_1 \overset{k[[X]][Y]}{+} f_2) \overset{k[Y][[X]]}{\approx} \phi(f_1) \overset{k[Y][[X]]}{+} \phi(f_2)$$

Proof. Let $\mathcal{R}_{+hom}$ be defined as follows

$$\mathcal{R}_{+hom} = \{\langle \phi(f_1 \overset{k[[X]][Y]}{+} f_2)\ ,\ \phi(f_1) \overset{k[Y][[X]]}{+} \phi(f_2) \rangle \mid f_1, f_2 \in k[[X]][Y]\}$$

We prove that $\mathcal{R}_{+hom}$ is a bisimulation.

1	$\widetilde{\mathrm{hd}}\,\phi(f_1 \overset{k[[X]][Y]}{+} f_2) = \mathrm{map}\ \widetilde{\mathrm{hd}}\ (f_1 \overset{k[[X]][Y]}{+} f_2)$	Def ϕ
2	$= (\mathrm{map}\ \widetilde{\mathrm{hd}}\ f_1) \overset{k[Y]}{+} (\mathrm{map}\ \widetilde{\mathrm{hd}}\ f_2)$	5.1.2
3	$= \widetilde{\mathrm{hd}}\,\phi(f_1) \overset{k[Y]}{+} \widetilde{\mathrm{hd}}\,\phi(f_2)$	Def ϕ
4	$= \widetilde{\mathrm{hd}}\,(\phi(f_1) \overset{k[Y][[X]]}{+} \phi(f_2))$	Def $\overset{k[Y][[X]]}{+}$
5	$\widetilde{\mathrm{tl}}\,\phi(as \overset{k[[X]][Y]}{+} bs)$	
6	$= \phi(\mathrm{map}\ \widetilde{\mathrm{tl}}\ (as \overset{k[[X]][Y]}{+} bs))$	Def ϕ
7	$= \phi((\mathrm{map}\ \widetilde{\mathrm{tl}}\ as) \overset{k[[X]][Y]}{+} (\mathrm{map}\ \widetilde{\mathrm{tl}}\ bs))$	5.1.3
8	$\widetilde{\mathrm{tl}}\,(\phi(as) \overset{k[Y][[X]]}{+} \phi(bs))$	
9	$= \widetilde{\mathrm{tl}}\,\phi(as) \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,\phi(bs)$	Def $\overset{k[[X]][Y]}{+}$
10	$= \phi(\mathrm{map}\ \widetilde{\mathrm{tl}}\ as) \overset{k[Y][[X]]}{+} \phi(\mathrm{map}\ \widetilde{\mathrm{tl}}\ bs)$	Def ϕ
11	$\widetilde{\mathrm{tl}}\,\phi(as \overset{k[[X]][Y]}{+} bs)\ \mathcal{R}_{+hom}\ \widetilde{\mathrm{tl}}\,(\phi(as) \overset{k[Y][[X]]}{+} \phi(bs))$	10,7

Lemma 5.1.9.

$$\phi(\mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ f_1) \overset{k[Y][[X]]}{\approx} (\widetilde{\mathrm{hd}}\,\phi(f_1) : 0_{k[Y][[X]]})$$

Proof. Trivial. (omitted)

Proposition 5.1.10 (ϕ multiplicative homomorphism)**.** *We want to prove*

$$\phi(f_1 \overset{k[[X]][Y]}{*} f_2) \overset{k[Y][[X]]}{\approx} \phi(f_1) \overset{k[Y][[X]]}{*} \phi(f_2)$$

Proof. Let $\mathcal{R}_{*hom}$ inductively defined as follows

1. $\forall f_1, f_2 \in k[[X]][Y].\ \langle \phi(f_1 \overset{k[[X]][Y]}{*} f_2)\ ,\ \phi(f_1) \overset{k[Y][[X]]}{*} \phi(f_2)\rangle \in\ \mathcal{R}_{*hom}$

2. $\forall f_1, f_2, f_3, f_4 \in k[Y][[X]]$. if $f_1\ \mathcal{R}_{*hom}\ f_2$ and $f_3\ \mathcal{R}_{*hom}\ f_4$ then $(f_1 \overset{k[Y][[X]]}{+} f_3)\ \mathcal{R}_{*hom}\ (f_2 \overset{k[Y][[X]]}{+} f_4)$

3. $\forall f_1, f_2 \in k[Y][[X]]$. if $f_1 \overset{k[Y][[X]]}{\approx} f_2$ then $f_1\ \mathcal{R}_{*hom}\ f_2$

4. $\forall f_1, f_2, f_3 \in k[Y][[X]]$. if $f_1\ \mathcal{R}_{*hom}\ f_2$ and $f_2 \overset{k[Y][[X]]}{\approx} f_3$ then $f_1\ \mathcal{R}_{*hom}\ f_3$

We prove that $\mathcal{R}_{*hom}$ is a bisimulation by induction on the constructor clauses of $\mathcal{R}_{*hom}$. We do the proof for the first clause only, the proof for the other clauses is straightforward. First we show that the heads are equal

1	$\widetilde{\text{hd}}\,\phi(f_1 \overset{k[[X]][Y]}{*} f_2) = \text{map}\ \ \widetilde{\text{hd}}\ \ (f_1 \overset{k[[X]][Y]}{*} f_2)$	Def of ϕ
2	$= (\text{map}\ \ \widetilde{\text{hd}}\ \ f_1) \overset{k[Y]}{*} (\text{map}\ \ \widetilde{\text{hd}}\ \ f_2)$	5.1.4
3	$= \widetilde{\text{hd}}\,\phi(f_1) \overset{k[Y]}{*} \widetilde{\text{hd}}\,\phi(f_2)$	Def of ϕ
4	$= \widetilde{\text{hd}}\,(\phi(f_1) \overset{k[Y][[X]]}{*} \phi(f_2))$	Def of $\overset{k[Y][[X]]}{*}$

Now we show that the tails are in $\mathcal{R}_{*hom}$

5	$\widetilde{\text{tl}}\,\phi(f_1 \overset{k[[X]][Y]}{*} f_2) = \phi(\text{map}\ \ \widetilde{\text{tl}}\ \ (f_1 \overset{k[[X]][Y]}{*} f_2))$	Def of ϕ
6	$= \phi\big(\text{map}\ (\lambda z \to \widetilde{\text{hd}}\ z : 0_{K[[X]]})\ f_1 \overset{k[[X]][Y]}{*} \text{map}\ \ \widetilde{\text{tl}}\ \ f_2$ $\overset{k[[X]][Y]}{+} \text{map}\ \ \widetilde{\text{tl}}\ \ f_1 \overset{k[[X]][Y]}{*} f_2\big)$	5.1.5
7	$= \phi\big(\text{map}\ (\lambda z \to \widetilde{\text{hd}}\ z : 0_{K[[X]]})\ f_1 \overset{k[[X]][Y]}{*} \text{map}\ \ \widetilde{\text{tl}}\ \ f_2\big)$ $\overset{k[Y][[X]]}{+} \phi\big(\text{map}\ \ \widetilde{\text{tl}}\ \ f_1 \overset{k[[X]][Y]}{*} f_2\big)$	5.1.8

8	$\widetilde{\mathrm{tl}}\,(\phi(f_1) \overset{k[Y][[X]]}{*} \phi(f_2))$	
9	$= (\widetilde{\mathrm{hd}}\,\phi(f_1) : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,\phi(f_2) \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,\phi(f_1) \overset{k[Y][[X]]}{*} \phi(f_2)$	Def of $\overset{k[Y][[X]]}{*}$
10	$= (\widetilde{\mathrm{hd}}\,\phi(f_1) : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \phi(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_2)$ $\overset{k[Y][[X]]}{+} \phi(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_1) \overset{k[Y][[X]]}{*} \phi(f_2)$	Def of ϕ
11	$\phi\big(\mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ f_1 \overset{k[[X]][Y]}{*} \mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_2\big)$ $\mathcal{R}_{*hom}$ $\phi\big(\mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ f_1\big) \overset{k[Y][[X]]}{*} \phi\big(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_2\big)$	Def $\mathcal{R}_{*hom}{}^{1st}$
12	$\phi\big(\mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ f_1\big) \overset{k[Y][[X]]}{*} \phi\big(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_2\big)$ $\overset{k[Y][[X]]}{\approx}$ $(\widetilde{\mathrm{hd}}\,\phi(f_1) : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \phi\big(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_2\big)$	5.1.9, $\overset{k[Y][[X]]}{*}$ *pres* $\overset{k[Y][[X]]}{\approx}$
13	$\phi\big(\mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ f_1 \overset{k[[X]][Y]}{*} \mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_2\big)$ $\mathcal{R}_{*hom}$ $(\widetilde{\mathrm{hd}}\,\phi(f_1) : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \phi\big(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_2\big)$	11, 12, IH $\mathcal{R}_{*hom}{}^{4th}$
14	$\phi\big(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_1 \overset{k[[X]][Y]}{*} f_2\big)$ $\mathcal{R}_{*hom}$ $\phi(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f_1) \overset{k[Y][[X]]}{*} \phi(f_2)$	IH $\mathcal{R}_{*hom}{}^{1st}$
15	$\widetilde{\mathrm{tl}}\,\phi(f_1 \overset{k[[X]][Y]}{*} f_2)\ \mathcal{R}_{*hom}\ \mathrm{tl}\,'(\phi(f_1) \overset{k[Y][[X]]}{*} \phi(f_2))$	7,10,13,14, Def of $\mathcal{R}_{*hom}{}^{2nd}$

Definition 5.1.11. *Equality on the type $K[[X]][Y]$ is defined inductively as follows*

1. *if $\alpha \overset{K[[X]]}{\approx} \beta$ then $[\alpha] \overset{k[[X]][Y]}{\approx} [\beta]$*
2. *if $(\alpha \overset{K[[X]]}{\approx} \beta) \wedge (A \overset{k[[X]][Y]}{\approx} B)$ then $\alpha : A \overset{k[[X]][Y]}{\approx} \beta : B$*

Lemma 5.1.12.

$$\forall \alpha, \beta \in K[[X]].\ \ \phi([\alpha]) \overset{k[Y][[X]]}{\approx} \phi([\beta]) \rightarrow \alpha \overset{K[[X]]}{\approx} \beta$$

Proof. let $\mathcal{R}_{\mathit{INJ0}}$ be defined as follows

- If $\alpha, \beta \in K[[X]]$ such that $\phi([\alpha]) \overset{k[Y][[X]]}{\approx} \phi([\beta])$ then $\alpha \; \mathcal{R}_{\mathit{INJ0}} \; \beta$

We claim that $\mathcal{R}_{\mathit{INJ0}}$ is a bisimulation. Following is the proof

$$\begin{array}{r|l|l}
1 & \phi([\alpha]) \overset{k[Y][[X]]}{\approx} \phi([\beta]) & \\ \hline
2 & \quad \widetilde{\text{hd}}\; \phi([\alpha]) \overset{k[Y]}{\approx} \widetilde{\text{hd}}\; \phi([\beta]) & \text{from 1} \\
3 & \quad [\widetilde{\text{hd}}\; \alpha] \overset{k[Y]}{\approx} [\widetilde{\text{hd}}\; \beta] & \text{Def of } \phi \\
4 & \quad \widetilde{\text{hd}}\; \alpha \overset{K}{\approx} \widetilde{\text{hd}}\; \beta & \text{Def of } \overset{k[Y]}{\approx} \\
5 & \quad \tilde{\text{tl}}\; \phi([\alpha]) \overset{k[Y][[X]]}{\approx} \tilde{\text{tl}}\; \phi([\beta]) & \text{from 1} \\
6 & \quad \phi([\tilde{\text{tl}}\; \alpha]) \overset{k[Y][[X]]}{\approx} \phi([\tilde{\text{tl}}\; \beta]) & \text{Def of } \phi \\
7 & \quad \tilde{\text{tl}}\; \alpha \; \mathcal{R}_{\mathit{INJ0}} \; \tilde{\text{tl}}\; \beta & \text{Def of } \mathcal{R}_{\mathit{INJ0}}
\end{array}$$

Lemma 5.1.13.

$$\forall A, B \in K[[X]][Y].\ \ \phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \phi(\beta : B) \rightarrow \alpha \overset{K[[X]]}{\approx} \beta$$

Proof. let $\mathcal{R}_{\mathit{INJ1}}$ be defined as follows

- If $\exists A, B \in K[[X]][Y]$ such that $\phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \phi(\beta : B)$ then $\alpha \; \mathcal{R}_{\mathit{INJ1}} \; \beta$

we prove that $\mathcal{R}_{\mathit{INJ1}}$ is a bisimulation.

1	$\phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \phi(\beta : B)$	
2	$\widetilde{\text{hd}}\,\phi(\alpha : A) \overset{k[Y]}{\approx} \widetilde{\text{hd}}\,\phi(\beta : B)$	from 1
3	$(\widetilde{\text{hd}}\,\alpha : \text{map}\ \ \widetilde{\text{hd}}\ \ A) \overset{k[Y]}{\approx} (\widetilde{\text{hd}}\,\beta : \text{map}\ \ \widetilde{\text{hd}}\ \ B)$	Def of ϕ
4	$\widetilde{\text{hd}}\,\alpha \overset{K}{\approx} \widetilde{\text{hd}}\,\beta$	Def of $\overset{k[Y]}{\approx}$
5	$\widetilde{\text{tl}}\,\phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \widetilde{\text{tl}}\,\phi(\beta : B)$	from 1
6	$\phi(\widetilde{\text{tl}}\,\alpha : \text{map}\ \ \widetilde{\text{tl}}\ \ A) \overset{k[Y][[X]]}{\approx} \phi(\widetilde{\text{tl}}\,\beta : \text{map}\ \ \widetilde{\text{tl}}\ \ B)$	Def of ϕ
7	$\widetilde{\text{tl}}\,\alpha\ \mathcal{R}_{\mathit{INJ1}}\ \widetilde{\text{tl}}\,\beta$	Def of $\mathcal{R}_{\mathit{INJ1}}$

Lemma 5.1.14.

$$\forall A, B \in K[[X]][Y],\ \forall \alpha, \beta \in K[[X]].\ \ \alpha \overset{K[[X]]}{\approx} \beta\ \wedge\ \phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \phi(\beta : B) \rightarrow \phi(A) \overset{K[[X]]}{\approx} \phi(B)$$

Proof. let $\mathcal{R}_{\mathit{INJ2}}$ be defined as follows

- If $\exists \alpha, \beta \in K[[X]]$ such that $\alpha \overset{K[[X]]}{\approx} \beta \quad \wedge \quad \phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \phi(\beta : B)$ then $\phi(A)\ \mathcal{R}_{\mathit{INJ2}}\ \phi(B)$

We prove that $\mathcal{R}_{\mathcal{INJ}2}$ is a bisimulation

1	$\phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \phi(\beta : B)$	
2	$\alpha \overset{K[[X]]}{\approx} \beta$	
3	$\widetilde{\mathrm{hd}}\,\phi(\alpha : A) \overset{k[Y]}{\approx} \widetilde{\mathrm{hd}}\,\phi(\beta : B)$	from 1
4	$(\widetilde{\mathrm{hd}}\,\alpha : \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ A) \overset{k[Y]}{\approx} (\widetilde{\mathrm{hd}}\,\beta : \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ B)$	Def of ϕ
5	$\mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ A \overset{k[Y]}{\approx} \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ B$	2,4, Def $\overset{k[Y]}{\approx}$
6	$\widetilde{\mathrm{hd}}\,\phi(A) \overset{k[Y]}{\approx} \widetilde{\mathrm{hd}}\,\phi(B)$	2,4, Def ϕ
7	$\widetilde{\mathrm{tl}}\,\phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \widetilde{\mathrm{tl}}\,\phi(\beta : B)$	from 1
8	$\phi(\widetilde{\mathrm{tl}}\,\alpha : \mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ A) \overset{k[Y][[X]]}{\approx} \phi(\widetilde{\mathrm{tl}}\,\beta : \mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ B)$	Def of ϕ
9	$\widetilde{\mathrm{tl}}\,\alpha \overset{K[[X]]}{\approx} \widetilde{\mathrm{tl}}\,\beta$	from 2
10	$\phi(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ A)\ \mathcal{R}_{\mathcal{INJ}2}\ \phi(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ B)$	8,9, Def of $\mathcal{R}_{\mathcal{INJ}2}$
11	$\widetilde{\mathrm{tl}}\,\phi(A)\ \mathcal{R}_{\mathcal{INJ}2}\ \widetilde{\mathrm{tl}}\,\phi(B)$	10, Def of ϕ

Proposition 5.1.15 (proof ϕ is injective)**.**

$$\forall A, B \in k[[X]][Y].\ \ \phi(A) \overset{k[Y][[X]]}{\approx} \phi(B) \to A \overset{k[[X]][Y]}{\approx} B$$

Proof. We do the proof by induction on A and B [1]

[1] induction is done on non-empty list, the base case is singleton list.

<u>base case:</u> $C = [\alpha]$, $D = [\beta]$

$$\begin{array}{r|ll}
1 & \underline{\phi([\alpha]) \overset{k[Y][[X]]}{\approx} \phi([\beta])} & \\
2 & \quad \vert \ \alpha \overset{K[[X]]}{\approx} \beta & \text{by 5.1.12} \\
3 & \quad \vert \ [\alpha] \overset{k[[X]][Y]}{\approx} [\beta] & \text{2, Def of } \overset{k[[X]][Y]}{\approx}
\end{array}$$

<u>base case:</u> $C = [\alpha]$, $D = \beta : B$

$$\begin{array}{r|ll}
1 & \underline{\phi([\alpha]) \overset{k[Y][[X]]}{\approx} \phi(\beta : B)} & \\
2 & \quad \vert \ [\text{hd } \alpha] \overset{k[[X]][Y]}{\approx} (\text{hd } \beta : \text{map } (\lambda x \to \text{hd } x)B) & \text{Def of } \alpha \\
3 & \quad \vert \ \bot & \text{2, Def of } \overset{k[[X]][Y]}{\approx} \\
4 & \neg(\phi([\alpha]) \overset{k[Y][[X]]}{\approx} \phi(\beta : B)) &
\end{array}$$

<u>step:</u> $C = \alpha : A$, $D = \beta : B$

$$\begin{array}{r|ll}
1 & \underline{\phi(\alpha : A) \overset{k[Y][[X]]}{\approx} \phi(\beta : B)} & \\
2 & \quad \vert \ \alpha \overset{K[[X]]}{\approx} \beta & \text{from 1, by 5.1.13} \\
3 & \quad \vert \ \phi(A) \overset{k[Y][[X]]}{\approx} \phi(B) & \text{from 1, by 5.1.14} \\
4 & \quad \vert \ A \overset{k[[X]][Y]}{\approx} B & \text{from 3, by IH} \\
5 & \quad \vert \ (\alpha : A) \overset{k[[X]][Y]}{\approx} (\beta : B) & \text{2,4, Def of } \overset{k[[X]][Y]}{\approx}
\end{array}$$

5.2 The lemma

Now that we have the injective homomorphism ϕ we can re-write the Y polynomial $F(X,Y)$ as a formal power series in X with coefficients in $k[Y]$. We let $F = \phi(F(X,Y)) = F_0(Y) + F_1(Y)\ X + ... + F_q\ X^q + ... \in k[Y][[X]]$. Now given two coprime monic polynomials $\bar{G}(Y)$ and $\bar{H}(Y)$ of degrees r and s respectively, such that $F_0(Y) = \bar{G}(Y) \overset{k[Y]}{*} \bar{H}(Y)$; we want to find two power series $G = \sum G_i X^i$ and $H = \sum H_i X^i$. such that $G_0 = \bar{G}(Y)$ and $H_0 = \bar{H}(Y)$. with deg $G_i < r$ for all $i > 0$ and deg $H_i < s$ for all $j > 0$. such that $F = G \overset{k[Y][[X]]}{*} H$. We first define the following system of equations, assuming we have G_p and H_p such that $G_0 \overset{k[Y]}{*} H_p \overset{k[Y]}{+} H_0 \overset{k[Y]}{*} G_p = 1_{k[Y]}$.

$$\begin{aligned}
&U = \widetilde{\mathrm{hd}}\ (\widetilde{\mathrm{tl}}\ F) : (\widetilde{\mathrm{tl}}\ \widetilde{\mathrm{tl}}\ F \overset{k[Y][[X]]}{-} (\widetilde{\mathrm{tl}}\ G \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\ H))\\
&H^* = (H_p : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} U\\
&G^* = (G_p : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} U \qquad\qquad (*)\\
&H = H_0 : \left(\widetilde{\mathrm{map}}\ (\lambda x \to x \bmod\ H_0)\ H^*\right)\\
&E = \widetilde{\mathrm{map}}\ (\lambda x \to x \text{ quot }\ H_0)\ H^*\\
&G = G_0 : \left(G^* \overset{k[Y][[X]]}{+} E \overset{k[Y][[X]]}{*} (G_0 : 0_{k[Y][[X]]})\right)
\end{aligned}$$

Where quot and mod are functions returning the quotient and remainder of Y polynomials division respectively. Now we start by proving that $G \overset{k[Y][[X]]}{*} H \overset{k[Y][[X]]}{\approx} F$ as usual by construction of a bisimulation.

Notation 5.2.1. *For some $C \in k[Y]$, let $\widetilde{C} \in k[Y][[X]]$ denote $(C : 0_{k[Y][[X]]})$. for example: $\widetilde{G_0} = G_0 : 0_{k[Y][[X]]}$, $\widetilde{H_p} = H_p : 0_{k[Y][[X]]}$, $\widetilde{H_0} = H_0 : 0_{k[Y][[X]]}$, and $\widetilde{G_p} = G_p : 0_{k[Y][[X]]}$.*

Lemma 5.2.2. *With the notation above*

$$\widetilde{G_0} \overset{k[Y][[X]]}{*} \widetilde{H_p} \overset{k[Y][[X]]}{+} \widetilde{H_0} \overset{k[Y][[X]]}{*} \widetilde{G_p} \overset{k[Y][[X]]}{\approx} 1_{k[Y][[X]]}$$

Proof. Trivial, (omitted).

Lemma 5.2.3. *For all* $C \in k[Y]$ *and* $\beta \in k[Y][[X]]$

$$\beta \overset{k[Y][[X]]}{\approx} \widetilde{\text{map}}\,(\lambda x \to x \text{ quot } C)\,\beta \overset{k[Y][[X]]}{*} (C : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{+} \widetilde{\text{map}}\,(\lambda x \to x \bmod C)\,\beta$$

Proof. Trivial, (omitted).

Proposition 5.2.4. *For the system of equations* (*)

$$G \overset{k[Y][[X]]}{*} H \overset{k[Y][[X]]}{\approx} F$$

Proof. Let $\mathcal{R}_{\mathit{Lemm}}$ be inductively as follows

1. $\forall \alpha, \beta \in k[Y][[X]]$. if $\alpha \overset{k[Y][[X]]}{\approx} \beta$ then $\alpha \; \mathcal{R}_{\mathit{Lemm}} \; \beta$

2. $\forall \alpha, \beta, \gamma \in k[Y][[X]]$. if $\alpha \overset{k[Y][[X]]}{\approx} \beta$ then $(\gamma \overset{k[Y][[X]]}{+} \alpha) \; \mathcal{R}_{\mathit{Lemm}} \; (\gamma \overset{k[Y][[X]]}{+} \beta)$

3. $\forall \alpha, \beta, \in k[Y][[X]].A \in k[Y]$. if $\alpha \; \mathcal{R}_{\mathit{Lemm}} \; \beta$ then $(A : \gamma) \; \mathcal{R}_{\mathit{Lemm}} \; (A : \alpha)$

4. $\forall \alpha, \beta, \gamma \in k[Y][[X]]$. if $\alpha \; \mathcal{R}_{\mathit{Lemm}} \; \beta$ and $\beta \; \mathcal{R}_{\mathit{Lemm}} \; \gamma$ then $\alpha \; \mathcal{R}_{\mathit{Lemm}} \; \gamma$

It is straight forward to check that $\mathcal{R}_{\mathit{Lemm}}$ is a bisimulation. Now we prove that $(G \overset{k[Y][[X]]}{*} H) \; \mathcal{R}_{\mathit{Lemm}} \; F$.

proof $\widetilde{\mathrm{tl}}\,\widetilde{\mathrm{tl}}\,(G \overset{k[Y][[X]]}{*} H) \; \mathcal{R}_{\mathit{lemm}} \; \widetilde{\mathrm{tl}}\,\widetilde{\mathrm{tl}}\,F$

1

$$\widetilde{\mathrm{tl}}\,\widetilde{\mathrm{tl}}\,(G \overset{k[Y][[X]]}{*} H) = \widetilde{\mathrm{tl}}\,(\widetilde{\mathrm{hd}\,G} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,H \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,G \overset{k[Y][[X]]}{*} H)$$

Def of $\overset{k[Y][[X]]}{*}$

2

$$\begin{aligned} &= \widetilde{\mathrm{hd}\,G} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,\widetilde{\mathrm{tl}}\,H \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{\mathrm{hd}\,H} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,\widetilde{\mathrm{tl}}\,G \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,G \end{aligned}$$

Def of $\overset{k[Y][[X]]}{+, \; *}$

3

$$\begin{aligned} &= \widetilde{G_0} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{map}}\,(\lambda x \to x \bmod H_0)\,\widetilde{\mathrm{tl}}\,H^* \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{H_0} \overset{k[Y][[X]]}{*} \big(\widetilde{\mathrm{tl}}\,G^* \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,E \overset{k[Y][[X]]}{*} \widetilde{G_0}\big) \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,G \end{aligned}$$

Def of G, H

4

$$\begin{aligned} &\overset{k[[X]][Y]}{\approx} \widetilde{G_0} \overset{k[Y][[X]]}{*} \big(\widetilde{\mathrm{map}}\,(\lambda x \to x \bmod H_0)\,\widetilde{\mathrm{tl}}\,H^* \overset{k[Y][[X]]}{+} \widetilde{H_0} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,E\big) \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{H_0} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,G^* \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,G \end{aligned}$$

4.3.3

5

$$\begin{aligned} &= \widetilde{G_0} \overset{k[Y][[X]]}{*} \big(\widetilde{\mathrm{map}}\,(\lambda x \to x \bmod H_0)\,\widetilde{\mathrm{tl}}\,H^* \overset{k[Y][[X]]}{+} \\ &\qquad\qquad \widetilde{H_0} \overset{k[Y][[X]]}{*} (\lambda x \to x \;\mathrm{quot}\; H_0)\,\widetilde{\mathrm{tl}}\,H^*\big) \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{H_0} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,G^* \\ &\quad \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,G \end{aligned}$$

Def of E

6	$\mathcal{R}_{\mathit{Lemm}} \quad \widetilde{G_0} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, H^* \overset{k[Y][[X]]}{+} \widetilde{H_0} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, G^* \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\, H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, G$	(5.2.3), $\mathcal{R}_{\mathit{Lemm}}$ 2nd
	$= \widetilde{G_0} \overset{k[Y][[X]]}{*} \widetilde{G_p} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, U$	
7	$\overset{k[Y][[X]]}{+} \widetilde{H_0} \overset{k[Y][[X]]}{*} \widetilde{H_p} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, U$	Def of G^*, H^*
	$\overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\, H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, G$	
8	$\mathcal{R}_{\mathit{Lemm}}\, \widetilde{\mathrm{tl}}\, U \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\, H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, G$	(5.2.2), $\mathcal{R}_{\mathit{Lemm}}$ 2nd
9	$= \widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, F$	Def of U
10	$\widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, (G \overset{k[Y][[X]]}{*} H)\ \mathcal{R}_{\mathit{Lemm}}\ \widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, F$	By 1-9, $\mathcal{R}_{\mathit{Lemm}}$ 4th

proof $\widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,(G \overset{k[Y][[X]]}{*} H) = \widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,F$

1	$\widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,(G \overset{k[Y][[X]]}{*} H) = \widetilde{\mathrm{hd}}\left(\widetilde{G_0} \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\,H \overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\,G \overset{k[Y][[X]]}{*} H\right)$	Def of $\overset{k[Y][[X]]}{*}$
2	$= G_0 \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,H \overset{k[Y]}{+} \widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,G \overset{k[Y]}{*} H_0$	Def of $\overset{k[Y][[X]]}{*}$
3	$= G_0 \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,H^* \bmod H_0$ $\overset{k[Y]}{+} (\widetilde{\mathrm{hd}}\,G^* \overset{k[Y]}{+} \widetilde{\mathrm{hd}}\,E \overset{k[Y]}{*} G_0) \overset{k[Y]}{*} H_0$	Def of H, G
4	$= G_0 \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,H^* \bmod H_0$ $\overset{k[Y]}{+} (G_p \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,U + G_0 \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,H^* \operatorname{quot} H_0) \overset{k[Y]}{*} H_0$	Def E
5	$= G_0 \overset{k[Y]}{*} \left((\widetilde{\mathrm{hd}}\,H^* \operatorname{quot} H_0) \overset{k[Y]}{*} H_0 \overset{k[Y]}{+} \widetilde{\mathrm{hd}}\,H^* \bmod H_0\right)$ $\overset{k[Y]}{+} G_p \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,U \overset{k[Y]}{*} H_0$	rewriting
6	$= G_0 \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,H^* \overset{k[Y]}{+} G_p \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,U \overset{k[Y]}{*} H_0$	prop of polynomial div
7	$= G_0 \overset{k[Y]}{*} H_p \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,U \overset{k[Y]}{+} G_p \overset{k[Y]}{*} \widetilde{\mathrm{hd}}\,U \overset{k[Y]}{*} H_0$	Def of H^*
8	$= \widetilde{\mathrm{hd}}\,U$	$G_0\,H_p + H_0\,G_p = 1$
9	$= \widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,F$	Def of U

Since $\widetilde{\mathrm{tl}}\,\widetilde{\mathrm{tl}}\,(G \overset{k[Y][[X]]}{*} H)\ \mathcal{R}_{\mathit{lemm}}\ \widetilde{\mathrm{tl}}\,\widetilde{\mathrm{tl}}\,F$ and $\widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,(G \overset{k[Y][[X]]}{*} H) = \widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}\,F$, then by $\mathcal{R}_{\mathit{lemm}}$ 3^{rd} we get that $\widetilde{\mathrm{tl}}\,(G \overset{k[Y][[X]]}{*} H)\ \mathcal{R}_{\mathit{lemm}}\ \widetilde{\mathrm{tl}}\,F$. By the assumption of the lemma we have $\widetilde{\mathrm{hd}}\,(G \overset{k[Y][[X]]}{*} H) = G_0 \overset{k[Y]}{*} H_0 = \overset{k[Y]}{\approx} \widetilde{\mathrm{hd}}\,F$. Thus again by inductive step $\mathcal{R}_{\mathit{lemm}}$ 3^{rd} we get that $G \overset{k[Y][[X]]}{*} H\ \mathcal{R}_{\mathit{lemma}}\ F$. Since $\mathcal{R}_{\mathit{lemma}}$ is a bisimulation we have $G \overset{k[Y][[X]]}{*} H \overset{k[Y][[X]]}{\approx} F$.

Implementation note V.1: For the same reasons stated in note IV.1. The implementation of Hensel's lemma comes in two flavors; a monadic one that uses the type of streams

```
 newtype (Monad m)=>  Stream m a = Cons (a ,m (Stream m a))
```

and a purely functional one that uses the usual type Haskell type of streams

```
 data Stream a = Cons a (Stream a)
```

5.3 The inverse morphism $k[Y][[X]] \to k[[X]][Y]$

We just proved that $F(X,Y)$ seen as a power series $(\phi(F(X,Y)))$ is a product of two power series $G, H \in k[Y][[X]]$. However, we want to show that there exist two polynomials $G_z, H_z \in k[[X]][Y]$ such that there product is equal to $F(X,Y)$. Since ϕ is an embedding, then we know that $k[Y][[X]]$ contains an isomorphic copy of the ring $k[[X]][Y]$. Then what we need to do amounts to finding the inverse image of G and H in $k[[X]][Y]$. In what follows we construct a function $\psi : k[Y][[X]] \times \mathbb{N} \to k[[X]][Y]$ and establish that ψ is really the inverse morphism of ϕ if applied to elements in the image of ϕ.

Definition 5.3.1. *The function $\psi : k[Y][[X]] \times \mathbb{N} \to k[[X]][Y]$ is defined as follows*

$$\psi(\alpha, 0) = [\widetilde{\mathrm{map}}\, \mathrm{h\bar{d}}\ \alpha]$$

$$\psi(\alpha, n) = \widetilde{\mathrm{map}}\, \mathrm{h\bar{d}}\ \alpha : \psi(\widetilde{\mathrm{map}}\, \mathrm{t\bar{l}}\ \alpha, n-1)$$

Where $\mathrm{h\bar{d}}\ : k[Y] \to k$

$$\mathrm{h\bar{d}}\ \ [\,] = 0_k$$

$$\mathrm{h\bar{d}}\ \ a : as = a$$

and $\bar{\text{tl}}\ : k[Y] \to k[Y]$

$$\begin{aligned} &\bar{\text{tl}}\ [\,] = [\,] \\ &\bar{\text{tl}}\ \ a : as = as \end{aligned}$$

Lemma 5.3.2.

$$\forall \alpha \in k[Y][[X]].\ \ \widetilde{\text{tl}}\ \text{map}\ \bar{\text{tl}}\ \alpha = \text{map}\ \bar{\text{tl}}\ \widetilde{\text{tl}}\ \alpha$$

Proof. trivial (omitted)

Lemma 5.3.3.

$$\forall \alpha \in k[Y][[X]]\ \forall n \in \mathbb{N}.\ \ \text{map}\ \ \widetilde{\text{tl}}\ \ \psi(\alpha, n) = \psi(\widetilde{\text{tl}}\ \alpha, n)$$

Proof. . By induction on n.

base case: $n = 0$.

1	$\text{map}\ \ \widetilde{\text{tl}}\ \ \psi(\alpha, 0) = \text{map}\ \ \widetilde{\text{tl}}\ \ [\widetilde{\text{map}}\ \bar{\text{hd}}\ \alpha]$	Def of ψ
2	$\text{map}\ \ \widetilde{\text{tl}}\ \ \psi(\alpha, 0) = [\widetilde{\text{map}}\ \bar{\text{hd}}\ (\widetilde{\text{tl}}\ \alpha)] = \psi(\widetilde{\text{tl}}\ \alpha, 0)$	1, Def of map

step: Assuming the lemma for n

1	$\text{map}\ \ \widetilde{\text{tl}}\ \ \psi(\alpha, n+1) = \text{map}\ \ \widetilde{\text{tl}}\ \ (\widetilde{\text{map}}\ \bar{\text{hd}}\ \alpha : \psi(\text{map}\ \bar{\text{tl}}\ \alpha, n))$	Def of ψ
2	$= \widetilde{\text{map}}\ \bar{\text{hd}}\ (\widetilde{\text{tl}}\ \alpha) : (\text{map}\ \ \widetilde{\text{tl}}\ \ \psi(\text{map}\ \bar{\text{tl}}\ \alpha, n))$	1, Trivial
3	$= \widetilde{\text{map}}\ \bar{\text{hd}}\ (\widetilde{\text{tl}}\ \alpha) : \psi(\widetilde{\text{tl}}\ (\text{map}\ \bar{\text{tl}}\ \alpha), n)$	2, by IH
4	$= \widetilde{\text{map}}\ \bar{\text{hd}}\ (\widetilde{\text{tl}}\ \alpha) : \psi(\text{map}\ \bar{\text{tl}}\ (\widetilde{\text{tl}}\ \alpha), n)$	3, by 5.3.2
5	$= \psi(\widetilde{\text{tl}}\ \alpha, n+1)$	4, Def of ψ

Definition 5.3.4. *We define equality on the type $k[Y]$ by extending the intensional equality on lists as follows*

$$[\,] \overset{k[Y]}{\approx} [\,]$$

$$[\,] \overset{k[Y]}{\approx} 0_{k[Y]}$$

$$0_{k[Y]} \overset{k[Y]}{\approx} [\,]$$

$$\forall a, b \in k\ \forall as, bs \in k[Y].\ a = b \wedge as \overset{k[Y]}{\approx} bs \rightarrow (a : as) \overset{k[Y]}{\approx} (b : bs)$$

where $0_{k[Y]} = [0_k]$

Definition 5.3.5. *The notion of degree on $k[Y]$ can be defined as a function $k[Y] \rightarrow \mathbb{N}$*

$$deg_{k[Y]}[\,] = 0$$

$$deg_{k[Y]}(x : xs) = \text{if}\ \ xs \overset{k[Y]}{\approx} 0_{k[Y]}\ \text{then}\ \ 0\ \text{else}\ \ 1 + deg_{k[Y]}(xs)$$

Lemma 5.3.6.

$$\forall \alpha \in k[Y][[X]].\ \widetilde{\mathrm{hd}}\ \alpha \overset{k[Y]}{\approx} \bar{\mathrm{hd}}\ (\widetilde{\mathrm{hd}}\ \alpha) : \widetilde{\mathrm{hd}}\ (\widetilde{\mathrm{map}}\ \bar{\mathrm{tl}}\ \alpha)$$

Proof. proof is by simple structural induction on $\widetilde{\mathrm{hd}}\ \alpha$

<u>base case:</u> $\widetilde{\mathrm{hd}}\ \alpha = [\,]$

1	$\bar{\mathrm{hd}}\ [\,] : \widetilde{\mathrm{hd}}\ (\widetilde{\mathrm{map}}\ \bar{\mathrm{tl}}\ \alpha) = 0_k : \widetilde{\mathrm{hd}}\ (\bar{\mathrm{tl}}\ [\,] : \widetilde{\mathrm{map}}\ \bar{\mathrm{tl}}\ (\tilde{\mathrm{tl}}\ \alpha))$	Def of $\bar{\mathrm{hd}}$,$\widetilde{\mathrm{map}}$
2	$= 0_k : \bar{\mathrm{tl}}\ [\,] = 0_k = [\,]$	Def of $\bar{\mathrm{tl}}$, equality on $k[Y]$

<u>setp:</u> $\widetilde{\mathrm{hd}}\ \alpha = a : as$

1	$\bar{\mathrm{hd}}\ (a : as) : \widetilde{\mathrm{hd}}\ (\widetilde{\mathrm{map}}\ \bar{\mathrm{tl}}\ \alpha) = a : \widetilde{\mathrm{hd}}\ (\bar{\mathrm{tl}}\ (a : as) : \widetilde{\mathrm{map}}\ \bar{\mathrm{tl}}\ (\tilde{\mathrm{tl}}\ \alpha))$	Def of $\bar{\mathrm{hd}}$,$\widetilde{\mathrm{map}}$
2	$= a : as$	Def of $\bar{\mathrm{tl}}$

Lemma 5.3.7. *For all $\alpha \in k[Y][[X]]$ if $deg_{k[Y]}(\widetilde{\mathrm{hd}}\ \alpha) \leq n$ then*

$$\widetilde{\mathrm{hd}}\ \phi(\psi(\alpha, n)) = \widetilde{\mathrm{hd}}\ \alpha$$

Proof. By induction on n

<u>base case:</u> $n = 0$.

$$\begin{array}{r|ll}
1 & \widetilde{\mathrm{hd}}\ \phi(\psi(\alpha, 0)) = \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ (\psi(\alpha, 0)) & \text{Def of } \phi \\
2 & = \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ [\widetilde{\mathrm{map}}\ \mathrm{h\bar{d}}\ \alpha] = [\widetilde{\mathrm{hd}}\ (\widetilde{\mathrm{map}}\ \mathrm{h\bar{d}}\ \alpha)] & \text{1, Def of } \psi, \mathrm{map} \\
3 & = [\mathrm{h\bar{d}}\ (\widetilde{\mathrm{hd}}\ \alpha)] & \text{trivial}
\end{array}$$

Since $deg_{k[Y]}(\widetilde{\mathrm{hd}}\ \alpha) \leq n$ by definition 5.3.5 we consider two cases

- $\widetilde{\mathrm{hd}}\ \alpha = [\,]$ then from step 3 above we get

  $$\widetilde{\mathrm{hd}}\ \phi(\psi(\alpha, 0)) = [\mathrm{h\bar{d}}\ [\,]] == [0_k] = [\,]$$

  by definition of $\mathrm{h\bar{d}}$, $\overset{k[Y]}{\approx}$

- $\widetilde{\mathrm{hd}}\ \alpha = a : as$ and $as \overset{k[Y]}{\approx} 0_{k[Y]}$ then from setp 3 we get

  $$\widetilde{\mathrm{hd}}\ \phi(\psi(\alpha, 0)) = [\mathrm{h\bar{d}}\ (a : as)] = [a] = a : as$$

  by def of $\mathrm{h\bar{d}}$ and $\overset{k[Y]}{\approx}$

setp: Assuming the statement is true for some n

1	$\widetilde{\mathrm{hd}}\,\phi(\psi(\alpha, n+1)) = \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ (\psi(\alpha, n+1))$	Def of ϕ
2	$= \widetilde{\mathrm{hd}}\ (\widetilde{\mathrm{map}}\,\mathrm{h\bar{d}}\ \alpha) : \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ (\psi(\widetilde{\mathrm{map}}\,\mathrm{t\bar{l}}\ \alpha, n))$	Def of ψ, $\widetilde{\mathrm{map}}$
3	$= \mathrm{h\bar{d}}\ (\widetilde{\mathrm{hd}}\ \alpha) : \mathrm{map}\ \ \widetilde{\mathrm{hd}}\ \ (\psi(\widetilde{\mathrm{map}}\,\mathrm{t\bar{l}}\ \alpha, n))$	trivial
4	$= \mathrm{h\bar{d}}\ (\widetilde{\mathrm{hd}}\ \alpha) : \widetilde{\mathrm{hd}}\ (\widetilde{\mathrm{map}}\,\mathrm{t\bar{l}}\ \alpha)$	by IH
5	$= \widetilde{\mathrm{hd}}\ \alpha$	by 5.3.6

Proposition 5.3.8. *For $\alpha \in k[Y][[X]]$, If $\exists n \in \mathbb{N}.\ \forall j \in \mathbb{N}.\ \ deg_{k[Y]}(\widetilde{\mathrm{hd}}\ {\widetilde{\mathrm{tl}}}^{\,j} \alpha) \leq n$ then $\phi(\psi(\alpha, n)) \overset{k[Y][[X]]}{\approx} \alpha$*

Proof. we define $\mathcal{S}$ as

- $\exists n \in \mathbb{N}.\ \forall j \in \mathbb{N}.\ \ deg_{k[Y]}(\widetilde{\mathrm{hd}}\ {\widetilde{\mathrm{tl}}}^{\,j} \alpha) \leq n$
  then $\phi(\psi(\alpha, n))\ \mathcal{S}\ \alpha$

The heads are equal by 5.3.7, now we show assuming that $\phi(\psi(\alpha, n))\ \mathcal{S}\ \alpha$ we have $\phi(\psi(\widetilde{\mathrm{tl}}\ \alpha, n))\ \mathcal{S}\ \widetilde{\mathrm{tl}}\ \alpha$. for $n = 0$ the proof is trivial, we consider when $n > 0$

1	$\widetilde{\mathrm{tl}}\,\phi(\psi(\alpha, n)) = \phi(\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ \psi(\alpha, n))$	Def of ϕ
2	$= \phi(\widetilde{\mathrm{tl}}\ (\widetilde{\mathrm{map}}\,\mathrm{h\bar{d}}\ \alpha) : \mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ \psi(\widetilde{\mathrm{map}}\,\mathrm{t\bar{l}}\ \alpha, n-1))$	Def of ψ
3	$= \phi(\widetilde{\mathrm{map}}\,\mathrm{h\bar{d}}\ (\widetilde{\mathrm{tl}}\ \alpha) : \mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ \psi(\widetilde{\mathrm{map}}\,\mathrm{t\bar{l}}\ \alpha, n-1))$	trivial
4	$= \phi(\widetilde{\mathrm{map}}\,\mathrm{h\bar{d}}\ (\widetilde{\mathrm{tl}}\ \alpha) : \psi(\widetilde{\mathrm{tl}}\ (\widetilde{\mathrm{map}}\,\mathrm{t\bar{l}}\ \alpha), n-1))$	5.3.3
5	$= \phi(\widetilde{\mathrm{map}}\,\mathrm{h\bar{d}}\ (\widetilde{\mathrm{tl}}\ \alpha) : \psi(\widetilde{\mathrm{map}}\,\mathrm{t\bar{l}}\ (\widetilde{\mathrm{tl}}\ \alpha), n-1))$	5.3.2
6	$= \phi(\psi(\widetilde{\mathrm{tl}}\ \alpha), n)$	Def of ψ
7	$\mathcal{S}\ \ (\widetilde{\mathrm{tl}}\ \alpha)$	Def of $\mathcal{S}$

In the last step implicitly uses the fact that if $\forall j \in \mathbb{N}.\ deg_{k[Y]}(\widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}^{\,j}\alpha) \leq n$ the $\forall j \in \mathbb{N}.\ deg_{k[Y]}(\widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}^{\,j}(\widetilde{\mathrm{tl}}\,\alpha)) \leq n$

Now if we can prove that the property $\forall j \in \mathbb{N}.\ deg_{k[Y]}(\widetilde{\mathrm{hd}}\,\widetilde{\mathrm{tl}}^{\,j}(\widetilde{\mathrm{tl}}\,\alpha)) \leq n$ is true for H and G we obtained in (*), then we can use ψ to get their pre-image in $k[[X]][Y]$.

Lemma 5.3.9.

$$\forall n \in \mathbb{N}.\forall \alpha \in k[Y][[X]].\forall A \in k[Y].$$
$$\deg_{k[Y]}(A) > 0 \to deg_{k[Y]}(\widetilde{\mathrm{hd}}\ \ \widetilde{\mathrm{tl}}^{\,n}\ \widetilde{\mathrm{map}}\,(\lambda z \to z \bmod A)\ \alpha) < deg_{k[Y]}(A)$$

Proof. By induction on n. Let $deg_{k[Y]} = m > 0$

base case: $n = 0$, by property of polynomial division

$$deg_{k[Y]}(\widetilde{\mathrm{hd}}\ \ \widetilde{\mathrm{map}}\,(\lambda z \to z \bmod A)\ \alpha) = deg_{k[Y]}((\widetilde{\mathrm{hd}}\ \alpha) \bmod\ A) < m$$

step: assuming the statement is true for n.

$$\begin{aligned}
& deg_{k[Y]}(\widetilde{\mathrm{hd}}\ \ \widetilde{\mathrm{tl}}^{\,n+1}\ \widetilde{\mathrm{map}}\,(\lambda z \to z \bmod A)\ \alpha) \\
&= deg_{k[Y]}(\widetilde{\mathrm{hd}}\ \ \widetilde{\mathrm{tl}}^{\,n}\ \widetilde{\mathrm{map}}\,(\lambda z \to z \bmod A)\ \widetilde{\mathrm{tl}}\ \alpha) && \text{(by def of map)} \\
&< m && \text{(by IH)}
\end{aligned}$$

We cannot define the notion of degree for $k[[X]][Y]$ the way we did for $k[Y]$ since we need to assume that equality on $k[[X]]$ is decidable. Which is not. We use the following definition to

Definition 5.3.10. *For some $L \in k[[X]][Y]$ We say that* $\mathrm{Dg}\ (L) \leq n$ *for some* $n \in \mathbb{N}$ *if and only if* $\forall n < i < \mathrm{len}\ (L).\ L(i) \overset{k[[X]][Y]}{\approx} 0_{k[[X]][Y]}$. *Where* len *is the length function on lists and $L(i)$ is the element of index i in L.*

Lemma 5.3.11. *For $f \in k[[X]][Y]$,* $\mathrm{Dg}\ (\mathrm{map}\ \ \widetilde{\mathrm{tl}}\ \ f) \leq \mathrm{Dg}\ (f)$

Proof. The proof is straight forward by noting that len $(f) =$ len (map $\widetilde{\text{tl}}$ f) and (map $\widetilde{\text{tl}}$ $f)(i) = \widetilde{\text{tl}}\, f(i)$. We omit it here.

Corollary 5.3.12. *If $f \in k[[X]][Y]$ is monic of degree n,* Dg ($\widetilde{\text{map}}$ tl f) $< n$

Lemma 5.3.13. *For $f \in k[[X]][Y]$, if* Dg $(f) \leq n$ *then*

$$\forall j \in \mathbb{N}.\ deg_{k[Y]}(\widetilde{\text{hd}}\,\widetilde{\text{tl}}^{\,j}\phi(f)) \leq n$$

Proof. The proof is straight forward by noting that len $(f) =$ len (map $\widetilde{\text{tl}}$ f) and (map $\widetilde{\text{tl}}$ $f)(i) = \widetilde{\text{tl}}\, f(i)$. We omit it here.

Proposition 5.3.14. *For the equation system* (*)

$$U \overset{k[Y][[X]]}{\approx} (G_0 : 0_{k[[X]][Y]}) \overset{k[[X]][Y]}{*} \widetilde{\text{tl}}\, H \overset{k[[X]][Y]}{+} (H_0 : 0_{k[[X]][Y]}) \overset{k[[X]][Y]}{*} \widetilde{\text{tl}}\, G$$

Proof. let $\mathcal{T}$ be defined

1. For all $\alpha, \beta \in k[Y][[X]]$ If $\alpha \overset{k[Y][[X]]}{\approx} \beta$ then $\alpha\ \mathcal{T}\ \beta$
2. For all $\alpha, \beta, \gamma \in k[Y][[X]]$ If $\alpha\ \mathcal{T}\ \beta$ then $(\alpha \overset{k[Y][[X]]}{-} \gamma)\ \mathcal{T}\ (\beta \overset{k[Y][[X]]}{-} \gamma)$
3. For all $\alpha, \beta, \gamma \in k[Y][[X]]$ If $\alpha\ \mathcal{T}\ \beta$ and $\beta\ \mathcal{T}\ \gamma$ then $\alpha\ \mathcal{T}\ \gamma$

Clearly $\mathcal{T}$ is a bisimulation. We proof that
$U\ \mathcal{T}\ (G_0 : 0_{k[[X]][Y]}) \overset{k[[X]][Y]}{*} \widetilde{\text{tl}}\, H \overset{k[[X]][Y]}{+} (H_0 : 0_{k[[X]][Y]}) \overset{k[[X]][Y]}{*} \widetilde{\text{tl}}\, G$

1	$\widetilde{\text{hd}}\, U = \widetilde{\text{hd}}\, \widetilde{\text{tl}}\, F$	Def of U
2	$= \widetilde{\text{hd}}\, ((G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\text{tl}}\, H \overset{k[Y][[X]]}{+} \widetilde{\text{tl}}\, G \overset{k[Y][[X]]}{*} H)$	Def of U
3	$= \widetilde{\text{hd}}\, ((G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\text{tl}}\, H) \overset{k[Y]}{+} \widetilde{\text{hd}}\, \widetilde{\text{tl}}\, G \overset{k[Y]}{*} H_0$	Def of $\widetilde{\text{hd}}$, H
4	$= \widetilde{\text{hd}}\, ((G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\text{tl}}\, H) \overset{k[Y]}{+} \widetilde{\text{hd}}\, \widetilde{\text{tl}}\, G \overset{k[Y]}{*} \widetilde{\text{hd}}\, (H_0 : 0_{k[Y][[X]]})$	rewriting
5	$= \widetilde{\text{hd}}\, ((G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\text{tl}}\, H \overset{k[Y][[X]]}{+} (H_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\text{tl}}\, G)$	Def of $\widetilde{\text{hd}}$

1	$\widetilde{\mathrm{tl}}\, U = \widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, F \overset{k[Y][[X]]}{-} \widetilde{\mathrm{tl}}\, G \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, H$	Def of U
2	$\mathcal{T} \left(\widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, (G \overset{k[Y][[X]]}{*} H) \overset{k[Y][[X]]}{-} \widetilde{\mathrm{tl}}\, G \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, H\right)$	5.2.4, IH $\mathcal{T}$
	$\overset{k[Y][[X]]}{\approx} (G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, H$	
3	$\overset{k[Y][[X]]}{+} (H_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, G$	Def of $\overset{k[Y][[X]]}{*}$, 4.3.8
	$\overset{k[Y][[X]]}{+} \widetilde{\mathrm{tl}}\, H \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, G \overset{k[Y][[X]]}{-} \widetilde{\mathrm{tl}}\, G \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, H$	
4	$= (G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, H \overset{k[Y][[X]]}{+} (H_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, \widetilde{\mathrm{tl}}\, G$	rewriting
5	$= \widetilde{\mathrm{tl}} \left((G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, H) \overset{k[Y][[X]]}{+} (H_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} G\right)$	Def of $\overset{k[Y][[X]]}{*+}$
6	$\widetilde{\mathrm{tl}}\, U\ \mathcal{T}\ \widetilde{\mathrm{tl}} \left((G_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} \widetilde{\mathrm{tl}}\, H) \overset{k[Y][[X]]}{+} (H_0 : 0_{k[Y][[X]]}) \overset{k[Y][[X]]}{*} G\right)$	1-5, Def of $\mathcal{T}$

Corollary 5.3.15. *rewriting the previous result we get*

$$\forall q.\, U(q) \overset{k[Y]}{\approx} H_0 \overset{k[Y]}{*} G(q+1) + G_0 \overset{k[Y]}{*} H(q+1)$$

Lemma 5.3.16. *For the equation system* (*)

$T_1(q):\ \ \forall q.\ deg_{k[Y]}(U(q)) < deg_{k[Y]}(F(0))$

$T_2(q):\ \ \forall q.\ deg_{k[Y]}(G(q+1)) < deg_{k[Y]}(G_0)\quad T_3(q):\ \ \forall q.\ deg_{k[Y]}(H(q+1)) < deg_{k[Y]}(H_0)$

Proof. The proof is by mutual induction on q

By the premise of the lemma we know that $deg_{k[Y]}(G_0) = r$ and $deg_{k[Y]}(H_0) = s$ where $r + s = n = deg_{k[Y]}(F(X,Y))$

Base case: $q = 0$

Since $F(X,Y)$ is monic of degree n by 5.3.12 and 5.3.13 we know that $deg_{k[Y]}(F(1)) < n$. By definition of U, $deg_{k[Y]}(U(0)) = deg_{k[Y]}(F(1))$. Hence By definition of U,

$deg_{k[Y]}(U(0)) < n$. By 5.3.15 and properties of polynomials $deg_{k[Y]}(G(1)) = deg_{k[Y]}(U(0) \overset{k[Y]}{-} G_0 \overset{k[Y]}{*} H(1)) - deg_{k[Y]}(H_0)$.
By 5.3.9 and definition of H we get that $deg_{k[Y]}(H(1)) < s$, hence $deg_{k[Y]}(G_0 \overset{k[Y]}{*} H(1)) < r + s = n$. Since $deg_{k[Y]}(U(0)) < n$ we get that $deg_{k[Y]}(U(0) \overset{k[Y]}{-} G_0 \overset{k[Y]}{*} H(1)) < n$ again by properties of polynomials. Thus $deg_{k[Y]}(G(1)) < r = deg_{k[Y]}(G_0)$

Inductive step: Assuming $T_1(n)$, $T_2(n)$ and $T_3(n)$for some n
By similar argument $deg_{k[Y]}(G(n+1)) < r = deg_{k[Y]}(G_0)$ hence $T_2(n+1)$ and by 5.3.9 and definition of H we get $deg_{k[Y]}(H(n+1)) < s = deg_{k[Y]}(H_0)$ hence $T_3(n+1)$. Then $T_1(n+1)$ follows by 5.3.15 and simple properties of polynomials.

Now we combine the previous results to give a proof of the Hensel's lemma as stated in theorem 5.0.1.

proof of theorem 5.0.1. By 5.3.8 and 5.3.16 we get that $\phi(\psi(G, r)) = G$ and $\phi(\psi(H, s)) = H$. Since ϕ is a homomorphism and by proposition 5.2.4 we get that

$$\phi(\psi(G, r) \overset{k[[X]][Y]}{*} \psi(H, s)) \overset{k[Y][[X]]}{\approx} G \overset{k[Y][[X]]}{*} H \overset{k[[X]][Y]}{\approx} \phi(F(X, Y))$$

By injectivity of ϕ (proposition 5.1.15) we get that

$$\psi(G, r) \overset{k[[X]][Y]}{*} \psi(H, s) \overset{k[[X]][Y]}{\approx} F(X, Y)$$

Completing the proof Hensel's lemma. □

Part III

Finale

CHAPTER VI
Newton theorem

In this chapter we present Newton theorem as an application to the work developed in parts I and II.

6.1 Newton theorem

The statement of Newton theorem is as follows

Theorem 6.1.1 (Newton theorem). *For an algebraically closed field k of zero characteristic, given a monic polynomial of degree n*

$$F(X,Y) = Y^n + a_1(X)\,Y^{n-1} + ... + a_n(X) \in k[[x]][Y]$$

there exist a positive integer m such that

$$F(X,Y) = \prod_{i=1}^{n} \left(Y - \eta_i(X^{1/m})\right) \quad , \eta_i(X^{1/m}) \in k[[X^{1/m}]]$$

The standard method for obtaining this linear factorization is the famous Newton-Puiseux algorithm. The approach we follow is a new approach presented by Abhyankar [1]. The main idea of the proof is reducing the polynomial to one on which Hensel's lemma can be applied. Then by application of Hensel's lemma factors of smaller degrees are obtained. Repeating this process at most n times we obtain the linear factors. In its general form as presented in [1] the proof is not fully constructive. It relies on the ability to test if a polynomial in $k[[X]][Y]$ is the zero polynomial which is obviously undecidable. To remedy we consider only polynomials that satisfy some condition.

Theorem 6.1.2 (Weak Newton theorem). *For an algebraically closed field k of zero* *characteristic, Let*

$$F(X,Y) = Y^n + a_1(X)\; Y^{n-1} + ... + a_n(X) \in k[[X]][Y]$$

be a a monic polynomial of degree $n > 1$ *such that* $gcd(F, F_Y) = 1$ *in* $k((X))[Y]$, *where* F_Y *is the* Y *derivative of* F*. Then there exist a positive integer* m *such that*

$$F(X,Y) = \prod_{i=1}^{n} \big(Y - \eta_i(X^{1/m})\big) \quad , \eta_i(X^{1/m}) \in k[[X^{1/m}]]$$

Proof. The strategy is to find two coprime factors of $F(0,Y)$ then use Hensel's lemma to lift those to factors of $F(X,Y)$ in $k[[X]][Y]$. Then we inductively repeat the process for the obtained factors. We need to ensure that these factors exist, i.e. $F(0,Y) \neq (Y+a)^n$ for all $a \in k$ and $n > 1$. Now if $F(0,Y) = (Y+a)^n$ then either $F(0,Y) = Y^n$ when $a = 0$ or , since k is of characteristic 0, $F(0,Y) = Y^n + naY^{n-1} + ... + a^n$ $(na \neq 0)$. To exclude the latter we kill the $(n-1)^{th}$ coefficient in $F(X,Y)$, this amounts to adding $a_1(X)/n$ to the roots of $F(X,Y)$. Doing this we get

$$F_1(X,Y) = F(X, Y - a_1(X)/n) = Y^n + \tilde{a_1}(X)Y^{n-1} + ... + \tilde{a_n}(X)$$

such that $\tilde{a_1}(X) = 0$. Clearly, $F(0, Y - a_1(X)/n) \neq (Y+a)^n$ for $a \neq 0$. To ensure that $F(0, Y - a_1(X)/n) \neq Y^n$, we multiply the roots of $F(X, Y - a_1(X)/n)$ such that one of $\tilde{a_i}(X) \neq 0$ at $X = 0$. To multiply by X^{-d} we construct

$$F_2(X,Y) = X^{-dn}F_1(X, X^dY) = Y^n + \tilde{a_1}(X)X^{-d}Y^{n-1} + ... + \tilde{a_i}(X)X^{-di}Y^{n-i} + ... + \tilde{a_n}(X)X^{-dn}$$

We need to adjust the value d such that $\tilde{a_i}(0) \neq 0$ for some i and for all j $\tilde{a_j}(X) \in k[[X]]$. Thus we take $d = \min_{i=1}^{n} \big(\text{ord } \tilde{a_i}(X)/i\big)$. Since d might be a rational number, say $d = d_1/d_2$ yet another change of the polynomial is due. We change the indeterminate in the coefficients to $\hat{X}$ where $\hat{X}^{d_2} = X$ to get

$$F_3(\hat{X}, Y) = \hat{X}^{-d_1 n} F_1(\hat{X}^{d_2}, \hat{X}^{d_1} Y) = Y^n + \hat{a_1}(\hat{X})Y^{n-1} + ... + \hat{a_i}(\hat{X})Y^{n-i} + ... + \hat{a_n}(\hat{X})$$

The condition that $gcd(F, F_Y) = 1$ guarantees that $d \neq \infty$, i.e. that for some i we have $\tilde{a}_i(X) \neq 0$. Since otherwise, $F(X, Y - a_1(X)/n) = Y^n$ $(n > 1)$. Hence $F(X,Y) = (Y + a_1(X)/n)^n$ which is impossible for the square free $F(X,Y)$. This means that for all a we have $F_3(0,Y) \neq (Y + a)^n$ for $n > 1$ and we can find two coprime factors of $F_3(0,Y)$. We apply Hensel's lemma to lift those factors to factors of $F_3(X,Y)$ and then repeat the process for the obtained factors until linear factors are obtained. The process is guaranteed to end in at most $n-1$ steps. The linear factors obtained would be in the form $Y - b(X_1)$ where $b(X_1) \in k[[X_1]]$ and $X_1^m = X$ for some m. Note that since $gcd(F, F_Y) = 1$ then for any $G \mid F$ we have $gcd(G, G_Y) = 1$. Thus the same argument holds as we proceed inductively.

Remark 6.1.3. *We note that the proof of theorem 6.1.2 is not fully constructive. Markov's principle was used to justify that if $F_1(X,Y) \neq Y^n$ then $a_i(X) \neq 0$ for some i. However, this use of the principal is not essential to the proof.*

Implementation note VI.1 (Newton theorem)**:** The processes on the root (addition, multiplication..etc) lend themselves to another state monad. For example addition of d to the roots of a polynomial $F(X,Y)$ can be done by changing the polynomial indeterminate from Y to Z where $Z = Y + d$ to get $F(X,Z)$. Then we can keep track of the such change be adding $Z = Y + d$ to the state. With the current base library however, this was not possible since the indeterminates are just phantom types. So a polynomial $F_1 = Z$ and a polynomial $F_2 = Y$ are identical; a list `[0,1]`. As a solution each recursive call reverts the changes to the roots it introduced. However, not all changes can be reverted. In particular the change of the coefficients indeterminates from X to some X' where $X'^d = X$ cannot be reflected in the power series directly. This is because the power of the indeterminate must be a natural number, in fact it is the position of the coefficient in the stream that determines the power of the indeterminate. As a work around the program returns with the list of linear factors a natural number d mainly saying that the roots are power series in $X^{1/d}$. Hence the function signature

```
newton1 :: (Field k, Eq k) => F (ST (S k)) (R k) x y
                         -> ST (S k) (Integer, [F (ST (S k)) (R k) x y])
```

Where `F (ST (S k)) (R k) x y` is a type synonym for polynomials in `y` with power series coefficients in `x` over `R k`. `R k` is the type of algebraic closure for a field `k` and `ST (S k)` is the state monad where the state `(S k)` is a list of polynomials denoting the algebraic numbers obtained. To recall these definitions in more details see implementation notes II.1 and II.3.

After altering the roots we use the algebraic closure function `root` to obtain a root p of $F(0, Y)$. Then divide by p iteratively until we get q such that $F(0, Y) = p^n q$ and $p \nmid q$.

```
twoCopDec :: (Field k, Eq k) => UPoly (R k) y
                         -> ST (S k) (UPoly (R k) y, UPoly (R k) y)
```

Then by applying Hensel's lemma we get two factors of $F(X, Y)$ on which we recursively call the function `newton1`. As a result we get two tuples `(d1, factors1)` and `(d2, factors2)`. Since it might be the case that `d1` and `d2` are not equal, i.e. in `factor1` the coefficients are power series in X^{1/d_1} while in `factor2` they are in X^{1/d_2} ($d_1 \neq d_2$). We have to process these factors such that the coefficients in both are power series in $X^{1/\mathrm{lcm}\,(d_1,d_2)}$.

6.2 Examples

Here we show two examples of Newton theorem at work in which we find the branches of two simple plane curves.

Implementation note VI.2 (UI)**:** We also coded a small parser[1] and pretty printer for Newton algorithm. The output is printed as a pair `State` and `Result`. The state specifies the the algebraic closure (see II) while the result is the linear factors of the polynomial.

Example 6.2.1 (The node)**:** Figure 6.1 shows the node plane curve $Y^2 + X^2(X+1) = 0$.

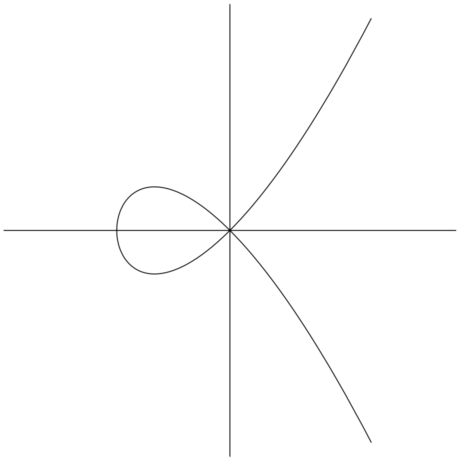


Figure 6.1: $Y^2 + X^2(X+1) = 0$

Applying the algorithm to this curve we get

```
> Y^2+X^3+X^2
State: a^2+1=0
Result: (Y-aX-1/2aX^2+1/8aX^3-1/16aX^4+ ...) (Y+aX+1/2aX^2-1/8aX^3+1/16aX^4+...)
```

Example 6.2.2 (The folium of Descartes)**:** The folium of Descartes plane curve is defined by the equation $Y^3 + X^3 - 3XY = 0$. The curve is shown in figure 6.2.

[1] The parser was built using BNF converter `www.digitalgrammars.com/bnfc` (visited May 2010).

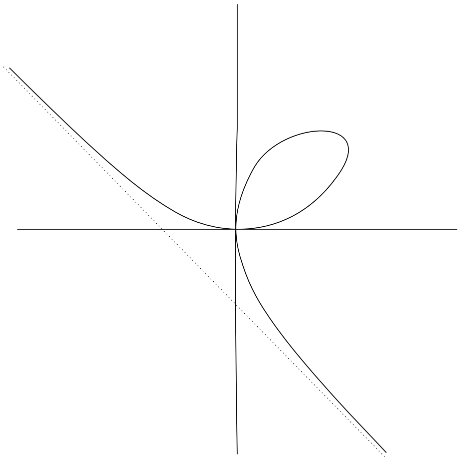


Figure 6.2: $Y^3 + X^3 - 3XY = 0$

Applying the algorithm to this polynomial we get

```
> Y^3+X^3-3XY

State: a=0,b^2-3=0

Result: (Y-1/3X^2+ ...) (Y-bX^1/2+1/6X^2+ ...) (Y+bX^1/2+1/6X^2+ ...)


State: a^2-3=0,b^2-3/4=0

Result: (Y-aX^1/2+1/6X^2+ ...) (Y+(-b+1/2a)X^1/2+(-1/6ab-1/12)X^2+ ...)

                                       (Y+(b+1/2a)X^1/2+(1/6ab-1/12)X^2+ ...)
```

CHAPTER VII
Conclusion

7.1 Discussion

Here we give some final remarks about the project and highlight some issues that deserves further investigation from our point of view.

7.1.1 Regularity property for Hensel's lemma

As the reader might have noticed reading part II, coinductive proofs even for simple propositions tend to be very long. They sometimes also become quite complicated; see for example the proofs in sections 5.1 and 5.3. It is worthwhile to see if starting from a polynomial $F(X, Y) \in k[X][Y]$ the coefficients of the lifted factors generated by one or more repeated application of Hensel's lemma admit some regularity property. In other words if these coefficients are recognizable[1] power series. In this case it would be possible to do without coinduction since these series could be finitely represented.

7.1.2 Complexity

No complexity analysis of the constructions presented have been done. The experimental observation of Newton algorithm indicate a high complexity. The program also tend to occupy a big stack space, most probably exponential in the degree of the input polynomial. The stack space problem can hopefully be remedies by some optimization; mainly by re-implementing the functions tail recursively.

7.1.3 Haskell

Haskell is a good choice for mathematical programming. In its purely functional style, there is a conspicuous correspondence between the resulting algorithms and

[1] By recognizable we mean recognizable by a *finite* k-nondeterministic automata.

the mathematical (constructive) proof. This correspondence becomes less clear when monadic computation is involved.

The lack of dependent types in Haskell causes some limitations. For example it is not obvious how the characteristic of a field can be encoded in the field definition. Another example is the problem of variable change of a polynomial discussed in implementation note VI.1 which can be easily solved if dependent types were available.

7.2 Future work

In the near future the plan is to optimize the code and analyze the complexity of the algorithms involved. One possible extension is to implement Hensel's lemma over more general/abstract structure than polynomials with power series coefficients, see [3] .

The next step would be constructing a logical model for the algebraic closure and a formal proof of Newton theorem based on the proof of theorem 6.1.2. Then it would be possible to implement the whole program in a proof assistant such as Agda or Coq.

BIBLIOGRAPHY

[1] S S Abhyankar. Algebraic geometry for scientists and engineers. American Mathematical Society, 1990.

[2] P Aczel. Non-well-founded sets. number 14. In *Notes. Stanford University*, 1988.

[3] Maria Emilia Alonso, Henri Lombardi, and Hervé Perdry. Elementary constructive theory of henselian local rings. *Math. Log. Q.*, 54(3):253–271, 2008.

[4] Errett Bishop. Foundations of constructive analysis. McGraw-Hill, 1967.

[5] Henry Thomas Colebrooke. *Algebra with Arithmetic of Brahmagupta and Bhaskara*. London, 1817.

[6] Thierry Coquand. Infinite objects in type theory. In *In Types for Proofs and Programs, International Workshop TYPES'93, volume 806 of LNCS*, pages 62–78, 1994.

[7] Thierry Coquand. A completeness proof for geometrical logic. In Westerstahl Hajek, Valdes-Villuaneva, editor, *Logic, Methodology and Philosophy of Sciences. Preceedings of the Twelfth International Congress*, pages 79–90, 2005.

[8] D.Duval. *Diverses questions relatives au calcul formel avec des nombres algébriques.* PhD thesis, Institut Fourier, Grenoble, 1987.

[9] P. Duren et al. A century of mathematics in america: Part ii. American Mathematical Society, 1989.

[10] A. Fröhlich and J. C. Shepherdson. On the factorisation of polynomials in a finite number of steps. *Mathematische Zeitschrift*, 62:331–334.

[11] Bart Jacobs and Jan Rutten. A tutorial on (co)algebras and (co)induction. *EATCS Bulletin*, 62:62–222, 1997.

[12] D.Duval J.Della Dora, C.Dicrescenzo. About a new method for computing in algebraic number fields. In *EUROCAL '85*, pages 289–290. 1985.

[13] Teo Mora. *Solving Polynomial Equation Systems I.* Cambridge University Press, 2002.

[14] Wim Ruitenburg Ray Mines, Fred Richman. *A Course in Constructive Algebra.* Universitext. Springer, 1987.

[15] David E. Rowe and John McCleary. *The History of Modern Mathematics: Ideas and their reception.* Academic Pr, 1990.

[16] Werner DePauli-Schimanovich; Eckehart Köhler; F. Stadler. The foundational debate: Complexity and constructivity in mathematics and physics. Vienna Circle Institute Yearbook. Springer, 1995.

[17] Allan K. Steel. Computing with algebraically closed fields. *J. Symb. Comput.*, 45(3):342–372, 2010.

[18] Van Der Waerden. Eine bemerkung über die unzerlegbarkeit von polynomen. *Math. Annalen*, 102:738–739, 1930.

APPENDIX
Proof of proposition 5.1.5

Proof. By induction on $\overset{k[[X]][Y]}{*}$

<u>Base case:</u> $[a] \overset{k[[X]][Y]}{*} [b] = [a \overset{K[[X]]}{*} b]$

1	$\mathrm{map}\ \widetilde{\mathrm{tl}}\ ([a] \overset{k[[X]][Y]}{*} [b]) = \mathrm{map}\ \widetilde{\mathrm{tl}}\ [a \overset{K[[X]]}{*} b]$	Def of $\overset{k[[X]][Y]}{*}$
2	$= [\widetilde{\mathrm{tl}}\,(a \overset{K[[X]]}{*} b)] = [(\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]} \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b) \overset{K[[X]]}{+} \widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b]$	Def of $\overset{K[[X]]}{*}$, map
3	$= [\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]} \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b] \overset{k[[X]][Y]}{+} [\widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b]$	Def of $\overset{k[[X]][Y]}{+}$
4	$= [\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]}] \overset{k[[X]][Y]}{*} [\widetilde{\mathrm{tl}}\, b] \overset{k[[X]][Y]}{+} [\widetilde{\mathrm{tl}}\, a] \overset{k[[X]][Y]}{*} [b]$	Def of $\overset{k[[X]][Y]}{*}$
5	$= \mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})[a] \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ [b]$ $\overset{k[[X]][Y]}{+} \mathrm{map}\ \widetilde{\mathrm{tl}}\ [a] \overset{k[[X]][Y]}{*} [b]$	Def of map

<u>Step :</u> $a : as \overset{k[[X]][Y]}{*} [b] = (a \overset{K[[X]]}{*} b) : (as \overset{k[[X]][Y]}{*} [b])$

6	$\mathrm{map}\ \widetilde{\mathrm{tl}}\ (a : as \overset{k[[X]][Y]}{*} [b])$	
7	$= \mathrm{map}\ \widetilde{\mathrm{tl}}\ ((a \overset{K[[X]]}{*} b) : (as \overset{k[[X]][Y]}{*} [b]))$	Def of $\overset{k[[X]][Y]}{*}$
8	$= (\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]} \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b \overset{K[[X]]}{+} \widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b)$ $:\ \mathrm{map}\ \widetilde{\mathrm{tl}}\ (as \overset{k[[X]][Y]}{*} [b])$	Def of map , $\widetilde{\mathrm{tl}}$
9	$= [\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]} \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b] \overset{k[[X]][Y]}{+} [\widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b]$ $\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \mathrm{map}\ \widetilde{\mathrm{tl}}\ (as \overset{k[[X]][Y]}{*} [b])$	Def of $\overset{k[[X]][Y]}{+}$
10	$= [\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]} \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b] \overset{k[[X]][Y]}{+} [\widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b]$ $\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ [b]$ $\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as \overset{k[[X]][Y]}{*} [b]$	(IH), Def of $\overset{k[[X]][Y]}{+}$
11	$= (\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]} \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b) : \mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} [\widetilde{\mathrm{tl}}\, b]$ $\overset{k[[X]][Y]}{+} (\widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b) : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as \overset{k[[X]][Y]}{*} [b]$	Def of $\overset{k[[X]][Y]}{+}$
12	$= ((\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]}) : \mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as) \overset{k[[X]][Y]}{*} [\widetilde{\mathrm{tl}}\, b]$ $\overset{k[[X]][Y]}{+} (\widetilde{\mathrm{tl}}\, a : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as) \overset{k[[X]][Y]}{*} [b]$	Def of $\overset{k[[X]][Y]}{*}$
13	$= \mathrm{map}\ (\lambda z \to \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ (a : as) \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ [b]$ $\overset{k[[X]][Y]}{+} \mathrm{map}\ \widetilde{\mathrm{tl}}\ (a : as) \overset{k[[X]][Y]}{*} [b]$	Def of map

Step : $[a] \overset{k[[X]][Y]}{*} b : bs = (a \overset{K[[X]]}{*} b) : bs$
Similar to the previous case
Step:

$$\begin{aligned}(a : as) \overset{k[[X]][Y]}{*} (b : bs) =&[a \overset{K[[X]]}{*} b]\\ &\overset{k[[X]][Y]}{+} (0_{K[[X]]} : [a] \overset{k[[X]][Y]}{*} bs)\\ &\overset{k[[X]][Y]}{+} (0_{K[[X]]} : as \overset{k[[X]][Y]}{*} [b])\\ &\overset{k[[X]][Y]}{+} (0_{K[[X]]} : 0_{K[[X]]} : as \overset{k[[X]][Y]}{*} bs)\end{aligned}$$

14 $$\operatorname{map}\ \widetilde{\text{tl}}\ (a : as \overset{k[[X]][Y]}{*} b : bs)$$

15 $$\begin{aligned}&= \operatorname{map}\ \widetilde{\text{tl}}\ [a \overset{K[[X]]}{*} b]\\ &\overset{k[[X]][Y]}{+} \operatorname{map}\ \widetilde{\text{tl}}\ (0_{K[[X]]} : [a] \overset{k[[X]][Y]}{*} bs)\\ &\overset{k[[X]][Y]}{+} \operatorname{map}\ \widetilde{\text{tl}}\ (0_{K[[X]]} : as \overset{k[[X]][Y]}{*} [b])\\ &\overset{k[[X]][Y]}{+} \operatorname{map}\ \widetilde{\text{tl}}\ (0_{K[[X]]} : 0_{K[[X]]} : as \overset{k[[X]][Y]}{*} bs)\end{aligned}$$ 5.1.3

16 $$\begin{aligned}&= [(\widetilde{\text{hd}}\, a : 0_{K[[X]]}) \overset{K[[X]]}{*} \widetilde{\text{tl}}\, b] \overset{k[Y][[X]]}{+} [\widetilde{\text{tl}}\, a \overset{K[[X]]}{*} b]\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \operatorname{map}\ \widetilde{\text{tl}}\ ([a] \overset{k[[X]][Y]}{*} bs)\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \operatorname{map}\ \widetilde{\text{tl}}\ (as \overset{k[[X]][Y]}{*} [b])\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : 0_{K[[X]]} : \operatorname{map}\ \widetilde{\text{tl}}\ (as \overset{k[[X]][Y]}{*} bs)\end{aligned}$$ Def of $\overset{K[[X]]}{+} *$, map ,

17 $$\begin{aligned}&= [(\widetilde{\text{hd}}\, a : 0_{K[[X]]}) \overset{K[[X]]}{*} \widetilde{\text{tl}}\, b] \overset{k[Y][[X]]}{+} [\widetilde{\text{tl}}\, a \overset{K[[X]]}{*} b]\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : [a : 0_{K[[X]]}] \overset{k[[X]][Y]}{*} \operatorname{map}\ \widetilde{\text{tl}}\ bs\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : [\widetilde{\text{tl}}\, a] \overset{k[[X]][Y]}{*} bs\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \operatorname{map}\ (\lambda z \rightarrow \widetilde{\text{hd}}\, z : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} [\widetilde{\text{tl}}\, b]\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \operatorname{map}\ \widetilde{\text{tl}}\ as \overset{k[[X]][Y]}{*} [b])\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : 0_{K[[X]]} : \operatorname{map}\ (\lambda z \rightarrow \widetilde{\text{hd}}\, z : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} \operatorname{map}\ \widetilde{\text{tl}}\ bs\\ &\overset{k[[X]][Y]}{+} 0_{K[[X]]} : 0_{K[[X]]} : \operatorname{map}\ \widetilde{\text{tl}}\ as \overset{k[[X]][Y]}{*} bs\end{aligned}$$ (IH)

18

$$\begin{aligned}
&= 0_{K[[X]]} : [\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]}] \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ bs \\
&\overset{k[[X]][Y]}{+} 0_{K[[X]]} : [\widetilde{\mathrm{tl}}\, a] \overset{k[[X]][Y]}{*} bs \\
&\overset{k[[X]][Y]}{+} ((\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]}) \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b) : \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} [\widetilde{\mathrm{tl}}\, b] \\
&\overset{k[[X]][Y]}{+} (\widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b) : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as \overset{k[[X]][Y]}{*} [b]) \\
&\overset{k[[X]][Y]}{+} 0_{K[[X]]} : 0_{K[[X]]} : \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ zs : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ bs \\
&\overset{k[[X]][Y]}{+} 0_{K[[X]]} : 0_{K[[X]]} : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as \overset{k[[X]][Y]}{*} bs
\end{aligned}$$

def of $\overset{k[[X]][Y]}{+}$

19

$$\begin{aligned}
&= (0_{K[[X]]} : (\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]})) \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ bs \\
&\overset{k[[X]][Y]}{+} (0_{K[[X]]} : \widetilde{\mathrm{tl}}\, a) \overset{k[[X]][Y]}{*} bs \\
&\overset{k[[X]][Y]}{+} ((\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]}) \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b) : \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} [\widetilde{\mathrm{tl}}\, b] \\
&\overset{k[[X]][Y]}{+} (\widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b) : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as \overset{k[[X]][Y]}{*} [b] \\
&\overset{k[[X]][Y]}{+} (0_{K[[X]]} : 0_{K[[X]]} : \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as) \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ bs \\
&\overset{k[[X]][Y]}{+} (0_{K[[X]]} : 0_{K[[X]]} : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as) \overset{k[[X]][Y]}{*} bs
\end{aligned}$$

trivial

20

$$\begin{aligned}
&= ((\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]}) \overset{K[[X]]}{*} \widetilde{\mathrm{tl}}\, b) : \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as \overset{k[[X]][Y]}{*} [\widetilde{\mathrm{tl}}\, b] \\
&\overset{k[[X]][Y]}{+} (\widetilde{\mathrm{tl}}\, a \overset{K[[X]]}{*} b) : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as \overset{k[[X]][Y]}{*} [b] \\
&\overset{k[[X]][Y]}{+} (0_{K[[X]]} : (\widetilde{\mathrm{hd}}\, a : 0_{K[[X]]}) : \\
&\qquad \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ as) \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ bs \\
&\overset{k[[X]][Y]}{+} (0_{K[[X]]} : \widetilde{\mathrm{tl}}\, a : \mathrm{map}\ \widetilde{\mathrm{tl}}\ as) \overset{k[[X]][Y]}{*} bs
\end{aligned}$$

Def of $\overset{k[[X]][Y]}{+}$

21

$$\begin{aligned}
&= \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ (a : as) \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ [b] \\
&\overset{k[[X]][Y]}{+} \mathrm{map}\ \widetilde{\mathrm{tl}}\ (a : as) \overset{k[[X]][Y]}{*} [b] \\
&\overset{k[[X]][Y]}{+} (0_{K[[X]]} : \mathrm{map}\ (\lambda z \rightarrow \widetilde{\mathrm{hd}}\ z : 0_{K[[X]]})\ (a : as)) \overset{k[[X]][Y]}{*} \mathrm{map}\ \widetilde{\mathrm{tl}}\ bs \\
&\overset{k[[X]][Y]}{+} (0_{K[[X]]} : \mathrm{map}\ \widetilde{\mathrm{tl}}\ (a : as)) \overset{k[[X]][Y]}{*} bs
\end{aligned}$$

Def of map

22

$$\begin{aligned}
&= \mathrm{map}\ \widetilde{\mathrm{tl}}\ ((a : as) \overset{k[[X]][Y]}{*} [b]) \\
&\overset{k[[X]][Y]}{+} 0_{K[[X]]} : \mathrm{map}\ \widetilde{\mathrm{tl}}\ ((a : as) \overset{k[[X]][Y]}{*} bs)
\end{aligned}$$

(IH)

23

$$\begin{aligned}
&= \mathrm{map}\ \widetilde{\mathrm{tl}}\ ((a : as) \overset{k[[X]][Y]}{*} [b]) \\
&\overset{k[[X]][Y]}{+} \mathrm{map}\ \widetilde{\mathrm{tl}}\ (0_{K[[X]]} : (a : as) \overset{k[[X]][Y]}{*} bs)
\end{aligned}$$

$\widetilde{\mathrm{tl}}\ 0_{K[[X]]}$

24 $= \mathrm{map}\ \widetilde{\mathrm{tl}}\ ((a : as) \overset{k[[X]][Y]}{*} [b] \overset{k[[X]][Y]}{+} (0_{K[[X]]} : (a : as) \overset{k[[X]][Y]}{*} bs))$ 5.1.3

25 $= \mathrm{map}\ \widetilde{\mathrm{tl}}\ ((a : as) \overset{k[[X]][Y]}{*} [b] \overset{k[[X]][Y]}{+} (a : as) \overset{k[[X]][Y]}{*} (0_{K[[X]]} : bs))$ trivial

26 $= \mathrm{map}\ \widetilde{\mathrm{tl}}\ ((a : as) \overset{k[[X]][Y]}{*} (b : bs))$ 4.3.3