Foundations of Computer Algebra: Symbolic Newton Iteration

Marc Moreno Maza
University of Western Ontario
CS 424 / CS 556 – Winter 2005

January 7, 2008

1 P-adic Expansions and Approximations

The goal of this section is to formalize in an algebraic point of view the notions of expansions and approximations well known in numerical analysis.

Definition 1 *Let R be an Euclidean domain and let $p \in R$ be a prime element. Let $a \in E$ be an element. We say that a has* degree zero w.r.t. p *and we write* $\deg(a, p) = 0$ *if for every positive integer $k \in \mathbb{N}$ the quotient of a w.r.t. p^k is zero. Assume from now on that a does not have degree zero w.r.t. p. If for every positive integer ℓ there exists a positive integer $k > \ell$ such that the quotient of a w.r.t. p^k is not zero, then we say a has* infinite degree w.r.t. p *and we write* $\deg(a, p) = +\infty$*, otherwise we say that a has* finite degree w.r.t. p *and we write* $\deg(a, p) < +\infty$*. Assume furthermore that* $\deg(a, p) < +\infty$ *holds. The largest positive integer $k \in \mathbb{N}$ such that the quotient of a w.r.t. p^k is not zero is called the* degree of a w.r.t. p *and is denoted by* $\deg(a, p)$.

Definition 2 *The Euclidean size δ of an Euclidean domain R is said* regular *if for every $a, b \in R$ with $\delta(b) \neq 0$ there exists a couple $(q, r) \in E \times E$ such that*

1. $a = bq + r$,
2. $r \neq 0 \quad \Rightarrow \quad \delta(r) < \delta(b)$,

3. $q \neq 0 \quad \Rightarrow \quad \delta(b) \leq \delta(a)$,

4. $q \neq 0$ and $r \neq 0 \quad \Rightarrow \quad \delta(q) < \delta(a)$.

Example 1 The absolute value in the ring $\mathbb{Z}$ of integers and the degree in the ring $k[x]$ of univariate polynomials over a field k are regular Euclidean sizes.

Proposition 1 *Let R be an Euclidean domain with a regular Euclidean size δ. Let p be a prime element. Then for every non-zero $a \in E$ there exists an integer $d \in N$ and elements $a_0, a_1, \ldots, a_d \in E$ such that*

1. $a = a_d p^d + \cdots + a_1 p + a_0$,

2. $a_d \neq 0$,

3. for every $i = 0 \cdots d$ we have $\deg(a_i, p) = 0$.

*The sequence $(a_0, a_1, \ldots, a_d)$ is called a p-*adic *expansion of a w.r.t. p.*

PROOF $\rhd$ Let (q_0, a_0) be a couple quotient-remainder given as in Definition 2 when dividing a w.r.t. p.

First we prove that a_0 must have degree 0 w.r.t. p. This is obvious if $a_0 = 0$. So let us assume that $a_0 \neq 0$ holds. This implies $\delta(p) > \delta(a_0)$. If a_0 would have positive degree k w.r.t. p, then, by the properties of a regular Euclidean size, we would have $\delta(a_0) \geq \delta(p^k)$. Since $\delta(p^k) \geq \delta(p)$ holds (by definition of an Euclidean domain) we would have a contradiction with $\delta(p) > \delta(a_0)$. Therefore in any case we have $\deg(a_0, p) = 0$.

Next, let us prove the proposition for those elements a which have quotient 0 w.r.t. p. Observe that if one quotient of a w.r.t. p is zero then we have $\delta(a) < \delta(p)$. Hence every quotient of a w.r.t. p is null. (Otherwise we would have $\delta(p) \leq \delta(a)$ by the properties of a regular Euclidean size.) Since $a = a_0$ and since a_0 has degree 0 w.r.t. p, the element a has degree 0 w.r.t. p and (a_0) is a p-adic expansion of a w.r.t. p. Therefore the proposition is proved in the case of the elements a which have quotient 0 w.r.t. p.

We can assume now that a does not have quotient 0 w.r.t. p. Then, two cases arise:

- Either both $q_0 \neq 0$ and $a_0 \neq 0$ hold, and this case we have $\delta(q_0) < \delta(a)$; then, we consider (q_1, a_1) a couple quotient-remainder given as in Definition 2 when dividing q_0 w.r.t. p.

- Or $a_0 = 0$. Then, there exists a smallest integer $e_0 > 1$ such that the remainder in an division of a w.r.t. p^{e_0} is not zero. (This follows from the fact that R is a Unique Factorization Domain and p is a prime.) Let (q_0, a_0) be a couple quotient-remainder in an division of a w.r.t. p^{e_0} with $a_0 \neq 0$.
  - If $q_0 = 0$, then we stop.
  - If $q_0 \neq 0$, then $\delta(q_0) < \delta(a)$; then, we consider (q_1, a_1) a couple quotient-remainder given as in Definition 2 when dividing q_0 w.r.t. p.

In each case where we consider a couple quotient-remainder (q_1, a_1), we can repeat a similar discussion as we did for (q_0, a_0).

Continuing in this manner we obtain a finite sequence $a_0, a_1, \ldots, a_d \in R$, a finite sequence $e_0, e_1, \ldots, e_d \in \mathbb{N}$ and a finite sequence $q_0, q_1, \ldots, q_d \in R$ such that we have

$$\begin{array}{rcllrcl}
a & = & q_0 p^{e_0} + a_0 & \text{with} & q_0 & \neq & 0 \\
q_0 & = & q_1 p^{e_1} + a_1 & \text{with} & q_1 & \neq & 0 \\
q_1 & = & q_2 p^{e_2} + a_2 & \text{with} & q_2 & \neq & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
q_{\ell_2} & = & q_{\ell-1} p^{e_{\ell-1}} + a_{\ell-1} & \text{with} & q_{\ell-1} & \neq & 0 \\
q_{\ell-1} & = & q_\ell p^{e_\ell} + a_\ell & \text{with} & q_\ell & = & 0
\end{array}$$

This process stops since we have

$$\delta(a) > \delta(q_0) > \delta(q_1) > \delta(q_2) > \cdots > \delta(q_{\ell-1}) \tag{1}$$

and since the Euclidean size of an element has value in $\mathbb{N} \cup \{-\infty\}$. Clearly, there exitst $d \in \mathbb{N}$ such that a writes

$$a = a_d p^d + \cdots + a_1 p + a_0.$$

where $a_d, \ldots, a_1, a_0$ are null or have degree 0 w.r.t. p, by the first part of this proof. Moreover $a_d \neq 0$ holds. $\triangleleft$

Remark 1 Proposition 1 applies in the ring $\mathbb{Z}$ of integers and the ring $\mathbf{k}[x]$ of univariate polynomials over a field $\mathbf{k}$. It extends also for the ring $R[y]$ of univariate polynomials over an Euclidean domain R with a regular Euclidean size δ and a prime element $p \in R$. Indeed, one can apply Proposition 1 to each coefficient of a polynomial $\phi \in R[y]$.

As we shall see now, it extends to a ring $R[y]$ of univariate polynomials over a commutative ring with identity element and to the element $p = y - g$ where g is an element of R.

Proposition 2 *Let R be a commutative ring with identity element. Let $\phi \in R[y]$ be a polynomial of degree n and let $g \in R$ be an element. We define $p = y - g$. There exists a unique sequence $\phi_0, \phi_1, \ldots, \phi_n \in R$ such that we have*

$$\phi = \phi_0 + \phi_1 p + \cdots + \phi_n p^n. \tag{2}$$

The above expression is called the Taylor expansion *of ϕ at g,*

PROOF $\triangleright$ By induction on n. If $n = 0$, the result is clear. Otherwise let q and r be the quotient and the remainder in the division of ϕ by p. We have

$$\phi = qp + r \quad \text{and} \quad r \in R.$$

Since $r = \phi(g)$ the conclusion (existence and unicity) follows by the induction hypothesis. $\triangleleft$

Remark 2 We can characterize the Taylor expansion of ϕ at g by means of the notion of a formal derivative.

Definition 3 *Let R be a commutative ring with identity element. For*

$$\phi = \phi_0 + \phi_1 y + \phi_2 y^2 + \cdots + \phi_n y^n$$

we define the formal derivative *of ϕ by*

$$\phi' = \phi_1 + 2\phi_2 y + \cdots + n\phi_n y^{n-1}. \tag{3}$$

Remark 3 The terminology formal derivative comes from the fact that we do not make use of the notion of a limit to define this derivative. Indeed, if R is a finite ring, then this would not make clear sense. However we retrieve the usual properties of derivatives.

Proposition 3 *Let R be a commutative ring with identity element. Let $\phi, \psi \in R[y]$ and let $a, b \in R$. Then the following properties hold*

1. $(a\phi + b\psi)' = a\phi' + b\phi'$,
2. $(\phi\psi)' = \phi'\psi + \phi\psi'$,
3. $(\phi(\psi))' = \psi'\phi'(\psi)$.

Proof $\triangleright$ See Lemma 9.19 in [GG99]. $\triangleleft$

Proposition 4 *Let R be a commutative ring with identity element. Let $\phi \in R[y]$ and let $g \in R$. There exists a polynomial $\psi \in R[y]$ such that we have*

$$\phi(y) = \phi(g) + \phi'(g)(y-g) + \psi(y)(y-g)^2. \tag{4}$$

Proof $\triangleright$ Follows Proposition 2 we have

$$\phi = \phi_0 + \phi_1 p + \cdots + \phi_n p^n \tag{5}$$

where $p = y - g$. The division of $\phi(y)$ by p shows that $\phi(g) = \phi_0$. Similarly, from (Equation 3), we have $\phi'(g) = \phi_1$. Plugging $\phi(g) = \phi_0$ and $\phi'(g) = \phi_1$ in (Equation 5) leads to

$$\phi = \phi(g) + \phi'(g)(y-g) + \psi(y)(y-g)^2 \tag{6}$$

for a polynomial ψ. $\triangleleft$

Remark 4 Proposition 2 admits several generalizations. First one can use a monic polynomial p of degree $d > 1$. In this case the coefficients $\phi_0, \phi_1, \ldots, \phi_n$ are polynomials of degree less than d. Another generalization, given by Proposition 5, is for multivariate polynomials. To understand this generalization, it is important to make the following three observations.

- First, in Proposition 4, one can consider g as a new indeterminate, and then the result holds for a polynomial $\phi \in R[y, g]$.
- Secondly, the notion of a formal derivative of a univariate polynomial (over a commutative ring with identity element) leads naturally to the notion of a partial derivative of a multivariate polynomial w.r.t. a variable. We will not detail this here, since this is quite straightforward
- The quantity $\psi(y-g)^2$ in Relation (4) lies in the ideal $\langle y-g \rangle^2$.

Proposition 5 *Let R be a commutative ring with identity element. We consider the ring of multivariate polynomials $P = R[x_1, \ldots, x_r][y_1, \ldots, y_r]$. We define $\mathbf{x} = (x_1, \ldots, x_r)$ and $\mathbf{y} = (y_1, \ldots, y_r)$. For every polynomial $\phi \in R[\mathbf{x}]$ there exists a polynomial $\psi \in P$ lying in the ideal $\langle y_1, \ldots, y_r \rangle^2$ such that*

$$\phi(x_1 + y_1, \ldots, x_r + y_r) = \phi(\mathbf{x}) + \Sigma_{j=1}^{j=r} \frac{\partial \phi}{\partial x_j}(\mathbf{x}) y_j + \psi. \tag{7}$$

Remark 5 Using the vocabulary of numerical analysis, Propositions 4 and 5 provides a Taylor expansion of ϕ of order 2. This leads to the following definition.

Definition 4 *Let R be a commutative ring with identity element and let $\mathcal{I}$ be an ideal of R. Let $a, b \in R$ be elements and k be a positive integer. We say that a is an $\mathcal{I}$*-adic approximation *of b at order k if we have*

$$a \equiv b \mod \mathcal{I}^k. \tag{8}$$

Remark 6 We conclude by two useful properties of ideal-adic approximations.

Proposition 6 *Let R be an Euclidean domain with a regular Euclidean size δ. Let $p \in R$ be a prime element. Let $a \in R$ be an element and let $(a_0, a_1, \ldots, a_d)$ be a p-adic expansion of a w.r.t. p. For every positive integer $k \leq d+1$ we define*

$$a^{(k)} = a_0 + a_1 p + \cdots a_{k-1} p^{k-1} \tag{9}$$

Then $a^{(k)}$ is a p-adic approximation *of a at order k, that is*

$$a^{(k)} \equiv a \mod p^k. \tag{10}$$

Proof $\triangleright$ Indeed, we have

$$a - a^{(k)} = a_k p^k + a_{k+1} p^{k+1} + \cdots + a_d p^d = (a_k + a_{k+1} p + \cdots + a_d p^{d-k}) p^k$$

$\triangleleft$

Proposition 7 *Let R be a commutative ring with identity element and let $\mathcal{I}$ be an ideal of R. Let $a, b \in R$ be elements, let k be a positive integer and let $\phi \in R[y]$ be a polynomial. Then we have*

$$a \equiv b \mod \mathcal{I}^k \quad \Rightarrow \quad \phi(a) \equiv \phi(b) \mod \mathcal{I}^k \tag{11}$$

That is, if a is an $\mathcal{I}$-adic approximation of b at order k, then $\phi(a)$ is an $\mathcal{I}$-adic approximation of $\phi(b)$ at order k.

PROOF $\triangleright$ This is simply because of the ring-homomorphism between R and $R/\mathcal{I}$. $\triangleleft$

2 Symbolic Newton Iteration

2.1 Linear Newton Iteration

Theorem 1 *Let R be an Euclidean domain with a regular Euclidean size δ. Let $p \in R$ be a prime element. Let $\phi \in R[y]$ and let $a, a_0 \in R$ such that*

1. $\phi(a) = 0$,
2. $a \equiv a_0 \mod p$,
3. $\phi'(a_0) \neq 0 \mod p$.

If a is unknown, but ϕ and a_0 are known then we can compute a p-adic expansion of a. Therefore we can compute a **exactly**.

PROOF $\triangleright$ Let $(a_0, a_1, \ldots, a_d)$ be a p-adic expansion of a w.r.t. p. Let $k < d+1$ be a positive integer. By Proposition 6, the element

$$a^{(k)} = a_0 + a_1 p + \cdots a_{k-1} p^{k-1} \tag{12}$$

is a p-adic approximation of a at order k. By Proposition 4 there exists a polynomial $\psi \in R[y]$ such that

$$\phi(a^{(k)} + a_k p^k) = \phi(a^{(k)}) + \phi'(a^{(k)}) a_k p^k + \psi(a^{(k)} + a_k p^k)(a_k p^k)^2. \tag{13}$$

Since we have

$$a \equiv a^{(k)} + a_k p^k \mod p^{k+1}$$

we deduce from Proposition 7

$$\phi(a) \equiv \phi(a^{(k)} + a_k p^k) \mod p^{k+1} \tag{14}$$

Since $\phi(a) = 0$ this shows that $\phi(a^{(k)} + a_k p^k)$ is in the ideal generated by p^{k+1}. Similarly, $\phi(a^{(k)})$ is in the ideal generated by p^k. Therefore we can divide $\phi(a^{(k)} + a_k p^k)$ and $\phi(a^{(k)})$ by p^k, leading to

$$\frac{\phi(a^{(k)} + a_k p^k)}{p^k} = \frac{\phi(a^{(k)})}{p^k} + \phi'(a^{(k)}) a_k + \psi(a^{(k)} + a_k p^k)(a_k)^2 p^k. \tag{15}$$

Now observe that

$$\frac{\phi(a^{(k)} + a_k p^k)}{p^k} \equiv 0 \equiv \psi(a^{(k)} + a_k p^k)(a_k)^2 p^k \mod p. \tag{16}$$

Let us denote by Φ_p the canonical homomorphism from R to $R/\langle p \rangle$. Then we obtain

$$\Phi_p(\frac{\phi(a^{(k)})}{p^k}) = -\Phi_p(\phi'(a^{(k)}))\Phi_p(a_k) \tag{17}$$

Now, since $a \equiv a_0 \mod p$ holds we have

$$\Phi_p(\phi'(a^{(k)})) \equiv \Phi_p(\phi'(a_0)) \mod p \tag{18}$$

Since $\phi'(a_0) \not\equiv 0 \mod p$ holds we can solve Equatiion 17 for $\Phi_p(a_k)$. Finally, from Theorem 1, we have $\deg(a_k, p) = 0$ such that we can *view* $\Phi_p(a_k)$ as a_k. This is straightforward if $R = \mathbb{Z}$ or if $R = \mathbf{k}[x]$ where $\mathbf{k}$ is a field, since we can choose for $\Phi_p(x)$ the remainder of x modulo p. $\triangleleft$

Theorem 2 *Let R be a commutative ring with identity element and let $\mathcal{I}$ be a finitely generated ideal of R. Let r be a positive integer. Let $f_1, \ldots, f_n \in R[x_1, \ldots, x_r]$ be n multivariate polynomials in the r variables $x_1, \ldots, x_r$. Let $a_1, \ldots, a_r \in R$ be elements. Let U be the Jacobian matrix of $f_1, \ldots, f_n$ evaluated at $(a_1, \ldots, a_r)$. That is, U is the $n \times r$ matrix defined by*

$$U = (u_{ij}) \quad \text{where} \quad u_{ij} = \frac{\partial f_i}{\partial x_j}(a_1, \ldots, a_r) \tag{19}$$

We assume that the following properties hold

1. *for every $i = 1 \cdots n$ we have $f_i(a_1, \ldots, a_r) \equiv 0 \mod \mathcal{I}$.*
2. *the Jacobian matrix U is left-invertible modulo $\mathcal{I}$.*

Then, for every positive integer ℓ we can compute $a_1^{(\ell)}, \ldots, a_r^{(\ell)} \in R$ such that

1. *for every $i = 1, \ldots, n$ we have $f_i(a_1^{(\ell)}, \ldots, a_r^{(\ell)}) \equiv 0 \mod \mathcal{I}^\ell$,*
2. *for every $j = 1, \ldots, r$ we have $a_j^{(\ell)} \equiv a_j \mod \mathcal{I}$.*

Proof $\triangleright$ We proceed by induction on $\ell \geq 1$. For $\ell = 1$ the claim follows from the hypothesis of the theorem. So let $\ell \geq 1$ be such that the claim is true. Hence there exist $a_1^{(\ell)}, \ldots, a_r^{(\ell)} \in R$ such that

$$f_i(a_1^{(\ell)}, \ldots, a_r^{(\ell)}) \equiv 0 \mod \mathcal{I}^\ell, \quad i = 1, \ldots, n \tag{20}$$

and

$$a_j^{(\ell)} \equiv a_j \mod \mathcal{I}, \quad i = 1, \ldots, r \tag{21}$$

Since $\mathcal{I}$ is finitely generated, then so is $\mathcal{I}^\ell$ and let $g_1, \ldots, g_s \in R$ such that

$$\mathcal{I}^\ell = \langle g_1, \ldots, g_s \rangle \tag{22}$$

Therefore, for every $i = 1, \ldots, r$, there exist $q_{i1}, \ldots q_{is} \in R$ such that

$$f_i(a_1^{(\ell)}, \ldots, a_r^{(\ell)}) \; = \; \Sigma_{k=1}^{k=s} q_{ik} g_k \tag{23}$$

For each $j = 1, \ldots, r$ we want to compute $B_j \in R$ such that

$$a_j^{(\ell+1)} = a_j^{(\ell)} + B_j \tag{24}$$

is the desired *next approximation*. We impose $B_j \in \mathcal{I}^\ell$ so let $b_{j1}, \ldots, b_{js} \in R$ be such that

$$B_j \; = \; \Sigma_{k=1}^{k=s} b_{jk} g_k \tag{25}$$

Using Proposition 5 we obtain

$$\begin{array}{rcl} f_i(a_1^{(\ell+1)}, \ldots, a_r^{(\ell+1)}) & = & f_i(a_1^{(\ell)} + B_1, \ldots, a_r^{(\ell)} + B_r) \\ & \equiv & f_i(a_1^{(\ell)}, \ldots, a_r^{(\ell)}) + \Sigma_{j=1}^{j=r} u_{ij}^{(\ell)} B_j \mod \mathcal{I}^{\ell+1} \\ & \equiv & \Sigma_{k=1}^{k=s} q_{ik} g_k + \Sigma_{j=1}^{j=r} u_{ij}^{(\ell)} B_j \mod \mathcal{I}^{\ell+1} \\ & \equiv & \Sigma_{k=1}^{k=s} q_{ik} g_k + \Sigma_{j=1}^{j=r} u_{ij}^{(\ell)} \left(\Sigma_{k=1}^{k=s} b_{jk} g_k \right) \mod \mathcal{I}^{\ell+1} \\ & \equiv & \Sigma_{k=1}^{k=s} \left(q_{ik} + \Sigma_{j=1}^{j=r} u_{ij}^{(\ell)} b_{jk} \right) g_k \mod \mathcal{I}^{\ell+1} \end{array} \tag{26}$$

where $\left(u_{ij}^{(\ell)} \right)$ is the Jacobian matrix of $(f_1, \ldots, f_n)$ at $(a_1^{(\ell)}, \ldots, a_r^{(\ell)})$.

Hence, solving for $a_1^{(\ell+1)}, \ldots, a_r^{(\ell+1)} \in R$ such that

$$f_i(a_1^{(\ell+1)}, \ldots, a_r^{(\ell+1)}) \equiv 0 \mod \mathcal{I}^{\ell+1}$$

leads to solving the system of linear equations:

$$q_{ik} + \Sigma_{j=1}^{j=r} u_{ij}^{(\ell)} b_{jk} \equiv 0 \mod \mathcal{I} \tag{27}$$

for $k = 1, \ldots, s$ and $i = 1, \ldots, n$.

Now using $a_j^{(\ell)} \equiv a_j \mod \mathcal{I}$ for $j = 1, \ldots, r$ we obtain

$$u_{ij}^{(\ell)} \equiv u_{ij} \mod \mathcal{I} \tag{28}$$

Therefore the linear system equations given by Relation (27) has solutions. $\triangleleft$

2.2 Quadratic Newton Iteration

Let R be a commutative ring with units. Let $p \in R$ and $\phi \in R[y]$. We are concerned with solving the equation

$$\phi(g) \ = \ 0. \tag{29}$$

The following Proposition 8 states that Relation (31) computes better and better approximations of solutions of Relation (29) that is, approximations modulo higher and higher powers of p. Comparing to the results of the previous section, this proposition offers a significant improvement: the accuracy doubles at each step, leading to the so-called *Quadratic Symbolic Newton Iteration*.

Proposition 8 *Let $\phi \in R[y]$ and let $g \in R$ be such that*

$$\phi(g) \equiv 0 \mod p \tag{30}$$

Assume that $\phi'(g)$ is invertible modulo p and let denote by $\phi'(g)^{-1}$ its inverse. Consider an element $h \in R$ such that

$$h \equiv g - \phi(g)\,\phi'(g)^{-1} \mod p^2 \tag{31}$$

Then the following assertions hold:

1. $\phi(h) \equiv 0 \mod p^2$,
2. $h \equiv g \mod p$,
3. $\phi'(h)$ *is invertible modulo p, and thus modulo p^2.*

Proof $\triangleright$ Since $\phi'(g)$ is invertible modulo p then $\phi'(g)$ is invertible modulo p^2. (This is easy to check: Exercise!) Hence Relation 31 makes sense. In fact, Algorithm **??** computes $\phi'(g)^{-1} \mod p^2$ given $\phi'(g)^{-1} \mod p$.

Since p divides p^2, the following congruence holds

$$h \equiv g - \phi(g)\, \phi'(g)^{-1} \mod p \tag{32}$$

Since $\phi(g) \equiv 0 \mod p$, this leads to

$$h \equiv g \mod p \tag{33}$$

and we have proved the second claim. Let us prove the first one. Using the Taylor expansion of ϕ at h and the fact that p^2 divides $(h-g)^2$ we obtain

$$\begin{array}{rcl} \phi(h) & \equiv & \phi(g) + \phi'(g)(h-g) \mod p^2 \\ & \equiv & \phi(g) - \phi'(g)\phi(g)\, \phi'(g)^{-1} \mod p^2 \\ & \equiv & 0 \mod p^2 \end{array} \tag{34}$$

proving the first claim. Finally, observe that $h \equiv g \mod p$ implies $\phi'(h) \equiv \phi'(g) \mod p$. Indeed ϕ' is a polynomial in $R[x]$. Therefore, $\phi'(h)$ is invertible modulo p, since $\phi'(g)$ is invertible modulo p, by hypothesis. $\triangleleft$

Remark 7 Intuitively, Proposition 8 says that if g is a good approximation of a zero of ϕ then h is a better approximation, at least *twice as accurate*.

Remark 8 Proposition 8 leads to an algorithm provided that we can

- check whether $\phi'(h)$ is invertible modulo p and
- compute its inverse modulo p in case.

Then for computing the inverse of $\phi'(g)$ modulo p the idea is to apply the algorithm to itself! That is, to the particular case of

$$\phi(g) = 1/g - f \tag{35}$$

This leads to Algorithm 1. Finally recall that for a prime element p in an Euclidean domain R, the computation of $\phi'(g)^{-1}$ modulo p can be done by the Extended Euclidean Algorithm.

Algorithm 1

Input: $\ell \in \mathbb{N}$, $\phi \in R[y]$, $p \in R$ and $g_0, s_0 \in R$ such that $\phi(g_0) \equiv 0 \mod p$ and $s_0\, \phi'(g_0) \equiv 1 \mod p$.

Output: $g \in R$ such that $\phi(g) \equiv 0 \mod p^\ell$ and $g \equiv g_0 \mod p$.

$r := \lceil \log_2(\ell) \rceil$
for $i = 1 \cdots r-1$ **repeat**
$g_i \equiv g_{i-1} - \phi(g_{i-1}) s_{i-1} \mod p^{2^i}$
$s_i \equiv (2 s_{i-1} - \phi'(g_i)\, {s_{i-1}}^2) \mod p^{2^i}$
$g := g_{r-1} - \phi(g_{r-1}) s_{r-1} \mod p^\ell$
return g

Theorem 3 *Algorithm 1 is correct.*

Proof $\triangleright$ Let $g_r \equiv g_{r-1} - \phi(g_{r-1}) s_{r-1} \mod p^{2^r}$. Then we have

$$g \equiv g_r \mod p^\ell \tag{36}$$

In other words g_r is a *solution of higher precision.* Then proving the theorem turns into proving the invariants for $i = 0 \cdots r$

1. $g_i \equiv g_0 \mod p$,
2. $\phi(g_i) \equiv 0 \mod p^{2^i}$,
3. if $i < r$ then $s_i\, \phi'(g_i) \equiv 1 \mod p^{2^i}$.

The case $i = 0$ is clear and we assume that $i > 0$ holds. Then by induction hypothesis $p^{2^{i-1}}$ divides $\phi(g_{i-1})$ and $s_{i-1} - \phi'(g_{i-1})^{-1}$. Then p^{2^i} divides their product, that is

$$\phi(g_{i-1}) s_{i-1} \equiv \phi(g_{i-1}) \phi'(g_{i-1})^{-1} \mod p^{2^i}. \tag{37}$$

Now we calculate

$$\begin{aligned} g_i &\equiv g_{i-1} - \phi(g_{i-1}) s_{i-1} \mod p^{2^i} \\ &\equiv g_{i-1} - \phi(g_{i-1}) \phi'(g_{i-1})^{-1} \mod p^{2^i} \end{aligned} \tag{38}$$

Now applying Proposition 8 with $g = g_{i-1}$, $h = g_i$ and $m = p^{2^{i-1}}$ we obtain the first two invariants. In addition, this leads to

$$g_i \equiv g_{i-1} \mod p^{2^{i-1}} \tag{39}$$

for $i < r$. This implies

$$\phi'(g_i)^{-1} \equiv \phi'(g_{i-1})^{-1} \equiv s_{i-1} \mod p^{2^{i-1}} \tag{40}$$

Now, s_i is obtained by one Newton step for inversion (see Algorithm **??**). Therefore the third invariant follows. $\lhd$

References

[BL82] Bruno Buchberger and Rüdiger Georg Konrad Loos. *Algebraic Simplification*, pages 11–43. Springer Verlag, Wien–New York, 1982.

[DST93] J. Davenport, Y. Siret, and E. Tournier. *Calcul Formel*. Masson, 1993. 2^e^ édition révisée.

[GCL92] K.O. Geddes, S.R. Czapor, and G. Labahn. *Algorithms for Computer Algebra*. Kluwer Academic Publishers, 1992.

[GG99] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, 1999.