Report: Hensel Lifting and Newton Iteration in Valutation Rings

Instructor: Amit Kumar Sinhababu and Sumanta Ghosh

Soham Chatterjee

sohamchatterjee999@gmail.com
Website: sohamch08.github.io

2024

Contents

Chapter 1

Introduction

Chapter 2

Hensel Lifting

The hensel method described here will lift an approximate factorization of a polynomial over a Hensel Ring R with valuation v where the factors are relatively prime. We will show a linear convergence and a quadratic convergence behavior for the liftings.

2.1 Hensel Ring

Definition 2.1.1: Hensel Ring

A ring with valuation $v : R \to \mathbb{R}_{\geq 0}$ is called a Hensel Ring if:

(i) $\forall\, a \in R, v(a) \leq 1$

(ii) $\forall\, a, b \in R, \forall\, \epsilon > 0, \exists\, c \in R$ such that $(v(a) \leq v(b) \implies v(a - bc) \leq \epsilon)$

In other words R is Hensel iff it is contained and dense in the valuation ring of its quotient field (with respect to the unique extension of v). We sometimes call such v a Hensel Valuation.

In condition (ii) we assume we can compute the c efficiently.

Theorem 2.1.1

Condition (i) of Hensel Ring $\implies$ v is Non-Archimedean.

Proof. Let $a, b \in R$. Now

$$\begin{aligned} v(a+b)^k &= v((a+b)^k) \\ &= v\left(\sum_{i=0}^{k} \binom{k}{i} a^{n-i} b^i\right) \leq \sum_{i=0}^{k} v\left(\binom{k}{i}\right) v(a)^{n-i} v(b)^i \\ &\leq \sum_{i=0}^{k} v(a)^{n-i} v(b)^i \leq \sum_{i=0}^{k} M^{n-i} M^i \qquad [m = \max v(a), v(b)] \\ &= \sum_{i=0}^{k} M^k = M^k(k+1) \end{aligned}$$

Hence

$$\left(\frac{v(a+b)}{M}\right)^k \leq (k+1) \iff \frac{v(a+b)}{M} \leq (1+k)^{\frac{1}{k}}$$

As $k \to \infty$ the RHS approaches 1 so $v(a+b) \leq M$. ■

Example 2.1 ($p-$adic Valuations)

- $\mathbb{Z}$ with $p-$adic valution v_p where $p \in \mathbb{N}$ is prime is a Hensel ring. Here $v_p(a) = p^{-n}$ where $n = \max\{k \geq 0 \mid p^k \mid a\}$
- $\mathbb{F}[y]$ with $p-$adic valutaion v_p where $p \in \mathbb{F}[y]$ is an irreducible polynomial is a Hensel Ring. Here $v_p(f) = 2^{-n \deg p}$ where $n = \max\{k \geq 0 \mid p^k \mid f\}$

Note:-

From the valuation v over R we naturally get a valuation v over the polynomial ring $R[x]$ by defining

$$\forall\, f \in R[x], \text{ let } f = \sum_{i=0}^{n} f_i x^i, \text{ then } v\left(\sum_{i=0}^{n} f_i x^i\right) = \max_i\{v(f_i)\}$$

2.2 Conditions related to Hensel's Lemma

We will define 5 conditinos. First suppose we have:

(1) $f \in R[x]$
(2) $f_0, \dots, f_m \in R[x] \quad \mathcal{F} = \{f_i : 0 \leq i \leq m\}$
(3) $f_0^*, \dots, f_m^* \in R[x] \quad \mathcal{F}^* = \{f_i^* : 0 \leq i \leq m\}$
(4) $s_0, \dots, s_m \in R[x] \quad \mathcal{S} = \{s_i : 0 \leq i \leq m\}$
(5) $s_0^*, \dots, s_m^* \in R[x] \quad \mathcal{S}^* = \{s_i^* : 0 \leq i \leq m\}$
(6) $z \in R$
(7) $\alpha, \delta, \epsilon \in \mathbb{R}$
(8) $\delta^* \in \mathbb{R}$
(9) $\gamma = \max\{\delta, \alpha\epsilon\}$

As you can see the set $\mathcal{F}^*$ basically represents the lift of $\mathcal{F}$ but here since we are saying the conditions in more generality we are not assuming any relations among them and we define some conditions involving them.

- $\boldsymbol{H_1(m, f, \mathcal{F}, \mathcal{S}, \epsilon)} := v\left(f - \prod_{i=0}^{m} f_i\right) \leq \epsilon < 1$
- $\boldsymbol{H_2(m, f, \mathcal{F}, \mathcal{S}, z, \delta)} := v\left(\sum_{i=0}^{m} s_i \prod_{j \neq i} f_i - z\right) < leq\delta < 1$
- $\boldsymbol{H_3(m, f, \mathcal{F}, S, z, \alpha, \delta, \epsilon)} :=$ (1) $f_1, \dots, f_m$ are monic
  (2) $\deg\left(\prod_{i=0}^{m} f_i\right) \leq \deg f$
  (3) $\deg s_i \leq \deg f_i \;\forall\, i \in [m]$
  (4) $\alpha\delta \leq 1, \alpha\epsilon^2 \leq 1$
  (5) $1 \leq \alpha v(z)$
- $\boldsymbol{H_4(m, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon)} :=$
  (1) $v(f_i^* - f_i) \leq \alpha\epsilon \quad \forall\, 0 \leq i \leq m$
  (2) $v(s_i^* - s_i) \leq \alpha\epsilon \quad \forall\, 0 \leq i \leq m$
  (3) $\deg f_i^* = \deg f_i \quad \forall\, i \in [m]$
  (4) $\deg s_i < \deg f_i \implies \deg s_i^* < \deg f_i^* \quad \forall\, i \in [m]$
- $\boldsymbol{H_5(m, f, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon, \delta^*)} :=$ Let $p \in [m]$. Then suppose
  - $\mathcal{I}_p = \{I_0, I_1, \dots, I_p\}$ be a partition of $\{0, \dots, m\}$ with $o \in I_0$.
  - $\overline{\mathcal{F}}_p^m = \{\overline{f}_i : i \in [p]\} \subseteq R[x]$ be a set of monic polynomials

Then define:

$$F_i = \prod_{j \in I_i} f_j, \qquad F_i^* = \prod_{j \in I_i} f_j^*, \qquad \mathfrak{s}_i^* = \sum_{j \in I_i} s_j \frac{F_i^*}{f_i^*}$$

So now we denote:

$$\mathscr{F} = \{F_i \colon 0 \le i \le p\}, \qquad \mathscr{F}^* = \{F_i \colon 0 \le^* i \le p\}, \qquad \mathscr{S} = \{\mathfrak{s}_i^* \colon 0 \le i \le p\}$$

Assume:

1. $v(\overline{f}_i - F_i) \le \alpha\epsilon \ \forall \ i \in [p]$
2. $\alpha v(s_i) \le 1 \ \forall \ 0 \le i \le m$
3. $\alpha\delta < 1$, $\alpha^2\delta \le 1$
4. $\alpha^2\epsilon < 1$, $\alpha^3\epsilon \le 1$

Then the following are equivalent:

(i) $\exists \ \overline{f}_0, \bar{s}_0, \ldots, \bar{s}_p \in R[x]$ denote

$$\overline{\mathcal{F}} = \{\overline{f}_i \colon 0 \le i \le p\}, \qquad \overline{\mathcal{S}} = \{\bar{s}_i \colon 0 \le i \le p\}$$

then the following conditions are true:

(a) $H_1(p, f, \overline{\mathcal{F}}, \overline{\mathcal{S}}, \epsilon^*)$

(b) $H_2(p, f, \overline{\mathcal{F}}, \overline{\mathcal{S}}, z, \delta^*)$

(c) $H_3(p, f, \overline{\mathcal{F}}, \overline{\mathcal{S}}, z, \alpha^*, \delta^*, \epsilon^*)$

(d) $H_4(p, f, \mathscr{F}, \overline{\mathcal{F}}, \mathscr{S}, \overline{\mathcal{S}}, z, \alpha^*, \delta^*, \epsilon^*)$

where $\alpha^* = \alpha$, $\epsilon^* = \alpha\epsilon\gamma$

(ii) $\exists \ \overline{f}_0 \in R[x]$ such that $H_1(p, f, \overline{\mathcal{F}}, \overline{\mathcal{S}}, \epsilon^*)$ is true

(iii) $\forall \ i \in [p]$ we have $v(\overline{f}_i - F_i^*) \le \epsilon^*$.

The first 3 conditions here togather imply that: From H_1 we get that $f_0 \cdots f_m$ is a good approximation of factorization of f with ϵ-precision, $H_2 \implies z$ plays a similar role to the gcd of $f_0, \ldots, f_m$ and it shows the generalized bezout's identity for gcd for multiple elements. In the usual treatment of Hensel's Lemma $f_0, \ldots, f_m$ are relatively prime (more precisely their images in the residue class field or R modulo the maximal ideal $\langle a \in R \mid v(a) < 1 \rangle$ satisfy the assumption then one can find $s_0, \ldots, s_m, \delta$ satisfying H_2 with $z = 1$. One can set $\alpha = 1$ or in general one can choose $\alpha = \frac{1}{v(z)}$. Thus H_2 staes that $f_0, \ldots, f_m$ are approximately pairwise relatively prime.

H_4 shows the connection between the lifts f_i^*, s_i^* and f_i, s_i.

H_5 basically states that the lifts are unique in the sense that one can group some of the $f_i' s$ to form $F_0, \ldots, F_p$ and change F_i to $\overline{f}_i$ with precision ϵ^* and still one will have the factorization of f with precision ϵ^*. H_5 is very important for the factorization algorithm in chapter 6.

Now we will state the Hensel's Lemma and will later give the algorithm to obtain the lifts.

2.3 Hensel's Lemma

First we will prove a helping lemma which will be very much usefull in the proof of Hensel's Lemma then we will state the actual theorem.

Theorem 2.3.1

(i) Let $a, f, p, s \in R[x]$ such that f is monic and $s = pf + a$ with $\deg a < \deg f$. Then we have $v(p) \le v(s)$ and $v(a) \le v(s)$

(ii) Let $h_0, \ldots, h_m \in R[x]$ and $h_o^*, \ldots, h_m^* \in R[x]$ such that we have $v(h_i^* - h_i) \le \epsilon$ for all $0 \le i \le m$. Then we have $v\left(\prod_{i=0}^{m} h_i - \prod_{i=0}^{m} h_i^*\right) \le \epsilon$

Theorem 2.3.2 Hensel's Lemma

Assume that we have $f \in R[x]$, $\mathcal{F} = \{f_0, \ldots, f_m\} \subseteq R[x]$, $\mathcal{S} = \{s_0, \ldots, s_m\} \subseteq R[x]$, $z \in R$ and $\alpha, \delta, \epsilon \in \mathbb{R}$ which satisfy:

1. $H_1(m, f, \mathcal{F}, \mathcal{S}, \epsilon)$
2. $H_2(m, f, \mathcal{F}, \mathcal{S}, z, \delta)$
3. $H_3(m, f, \mathcal{F}, S, z, \alpha, \delta, \epsilon)$

Then we can compute efficiently

$$\mathcal{F}^* = \{f_i^* \colon 0 \leq i \leq m\} \qquad \text{and} \qquad T = \{t_0, \ldots, t_m\}$$

such that

(i) **Linear Case**: $\mathcal{S}^* = \mathcal{S}$ and $\delta^* = \gamma$, $\epsilon^* = \alpha\gamma\epsilon$. Then we have the following conditions hold:

(a) $H_1(m, f, \mathcal{F}^*, \mathcal{S}^*, \epsilon^*)$

(b) $H_2(m, f, \mathcal{F}^*, \mathcal{S}^*, z, \delta^*)$

(c) $H_3(m, f, \mathcal{F}^*, \mathcal{S}^*, z, \alpha, \delta^*, \epsilon^*)$

(d) $H_4(m, f, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon)$

(e) $H_5(m, f, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon, \delta^*)$

(ii) **Quadratic Case**: $\mathcal{S}^* = T$ and $\delta^* = \alpha\gamma^2$, $\epsilon^* = \alpha\gamma\epsilon$. Assume that $\deg s_i > \deg f_i$ for $0 \leq i \leq m$. Then we have the following conditions hold:

(a) $H_1(m, f, \mathcal{F}^*, \mathcal{S}^*, \epsilon^*)$

(b) $H_2(m, f, \mathcal{F}^*, \mathcal{S}^*, z, \delta^*)$

(c) $H_3(m, f, \mathcal{F}^*, \mathcal{S}^*, z, \alpha, \delta^*, \epsilon^*)$

(d) $H_4(m, f, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon)$

(e) $H_5(m, f, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon, \delta^*)$

2.4 Hensel's Computation

Algorithm 1: Hensel's Computation

Input:

1. $f \in R[x]$, $\mathcal{F} = \{f_0, \ldots, f_m\} \subseteq R[x]$, $\mathcal{S} = \{s_0, \ldots, s_m\} \subseteq R[x]$
2. $z \in R$
3. $\alpha, \delta, \epsilon \in \mathbb{R}$

Output: $\mathcal{F}^* = \{f_0^*, \ldots, f_m^*\}$, $T = \{t_0, \ldots, t_m\}$

begin

Set $\gamma = \max\{\delta, \alpha\epsilon\}$, $\alpha^* = \alpha$, $\epsilon^* = \alpha\gamma\epsilon$ and $e = f - \prod_{i=0}^{m} f_i$

for $1 \leq i \leq m$ **do**

Compute $a_i, b_i, p_i \in R[x]$ such that

$$s_i e = p_i f_i + a_i, \qquad v(zb_i - a_i) \leq \epsilon\gamma, \qquad \deg b_i \leq \deg a_i < \deg f_i$$

Compute $a_0, b_0 \varnothing = inR[x]$ such that

$$a_0 = s_0 e + f_0 \sum_{i=1}^{m} p_i, \qquad v(zb_0 - a_0) \leq \epsilon\gamma, \qquad \deg b_0 \leq \deg f -^m f_i$$

for $0 \leq i \leq m$ **do**

$f_i^* = f_i + b_i$

for $1 \leq i \leq m$ **do**

Compute $c_i, d_i, g_i^* q_i \in R[x]$ such that

$$g_i^* = \prod_{j \neq i} f_j^*, \qquad s_i(s_i g_i^* - z) = q_i f_i^* + c_i \qquad v(zd_i - c_i) \leq \gamma^2 \qquad \deg d_i \leq \deg c_i < \deg f_i^*$$

Compute $g_0^* = \prod_{i=1}^{m} f_i^*$ and $c_0, d_0 \in R[x]$ such that

$$c_0 = s_o \left(\sum_{i=0}^{m} s_i g_i^* - z \right) + f_0^* \sum_{i=1}^{m} \left[q_i + s_i \left(\sum_{j \neq i} s_j \frac{g_j^*}{f_j^*} \right) \right], \quad v(zd_0 - c_0) \leq \gamma^2, \quad \deg d_0 \leq \deg f - \deg g_0$$

for $0 \leq i \leq m$ **do**

$t_i = s_i - d_i$

return $\mathcal{F}^* = \{f_i^* : 0 \leq i \leq m\}$, $T = \{t_i : 0 \leq i \leq m\}$

2.5 Proof of Hensel's Lemma

Chapter 3

Newton Iteration

Chapter 4

Solving Differential Equations

Chapter 5

Finding Short Vectors in Modules

Chapter 6

Factorization of Polynomials