

Soham Chatterjee Amit Kumar Sinhababu

CONTENTS

CHAPTER 1	Hensel Lifting in Valuation Rings	Page 3
1.1	Hensel Ring	3
1.2	Conditions related to Hensel's Lemma	4
1.3	Hensel's Lemma 1.3.1 Hensel's Computation 1.3.2 Proof of Hensel's Lemma	5 9 9
^HADTED 9	Part accus a party	D40F 10

Hensel Lifting in Valuation Rings

Now we will work with Valuation Rings. Till now we were assuming we can do a + b, ab, a^{-1} in a ring or field in O(1) time but for that we should have the whole table of addition, multiplication and inverse as in input which is not always possible if the ring or field is very big. That is why here we will describe Hensel Lifting for general Hensel Rings where we have an algorithm to efficiently find out a close enough element with respect to the given valuation. We will follow the paper from [vzG84].

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.

1.1 Hensel Ring

Definition 1.1.1: 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) $\forall a, b \in R, \forall \epsilon > 0, \exists c \in R \text{ such that } (v(a) \le v(b) \implies v(a bc) \le \epsilon)$

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 1.1.1

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

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

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

Hence

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

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

Example 1.1 (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 $n = \max\{k \ge 1\}$ $0 \mid p^k \mid a$
- $\mathbb{F}[y]$ with p-adic valutaion v_p where $p \in \mathbb{F}[y]$ is an irreducible polynomial is a Hensel Ring. Here $v_p(f) = 0$ $2^{-n \deg p}$ where $n = \max\{k \ge 0 \mid p^k \mid f\}$

Note:-

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

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

1.2 Conditions related to Hensel's Lemma

We will define 5 conditinos. First suppose we have:

- (1) $f \in R[x]$ $\mathcal{F} = \{f_i : 0 \le i \le m\}$ (2) $f_0, \dots, f_m \in R[x]$ $\mathcal{F}^* = \{f_i^* : 0 \le i \le m\}$ (3) $f_0^*, \dots, f_m^* \in R[x]$ $\mathcal{F}^* = \{f_i^* : 0 \le i \le m\}$ (4) $s_0, \dots, s_m \in R[x]$ $\mathcal{S} = \{s_i : 0 \le i \le m\}$ (5) $s_0^*, \dots, s_m^* \in R[x]$ $\mathcal{S}^* = \{s_i^* : 0 \le i \le m\}$

- (7) $\alpha, \delta, \epsilon \in \mathbb{R}$
- (8) $\delta^* \in \mathbb{R}$
- $\gamma = \max\{\delta, \alpha\epsilon\}$

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.

•
$$H_1(m, f, \mathcal{F}, \mathcal{S}, \epsilon) := v\left(f - \prod_{i=0}^m f_i\right) \le \epsilon < 1$$

•
$$H_2(m, f, \mathcal{F}, \mathcal{S}, z, \delta) := v\left(\sum\limits_{i=0}^m s_i\prod\limits_{j\neq i} f_i - z\right) \leq \delta < 1$$

- $H_3(m, f, \mathcal{F}, \mathcal{S}, z, \alpha, \delta, \epsilon) := (1)$ f_1, \dots, f_m are monic
 - (1) f_1, \dots, f_m are monte (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) $1 \leq \alpha v(z)$

•
$$H_4(m, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon) := (1)$$
 $v(f_i^* - f_i) \le \alpha \epsilon$ $\forall 0 \le i \le m$
 (2) $v(s_i^* - s_i) \le \alpha \epsilon$ $\forall 0 \le i \le m$
 (3) $\deg f_i^* = \deg f_i$ $\forall i \in [m]$
 (4) $\deg s_i < \deg f_i \implies \deg s_i^* < \deg f_i^*$ $\forall i \in [m]$

•
$$H_5(m, f, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon, \delta^*) := \text{Let } p \in [m]$$
. Then suppose

-
$$\mathcal{I}_p = \{I_0, I_1, \dots, I_p\}$$
 be a partition of $\{0, \dots, m\}$ with $o \in I_0$.
- $\overline{\mathcal{F}}_p^m = \{\overline{f}_i \colon i \in [p]\} \subseteq R[x]$ 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:

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

Assume:

1.
$$v(\overline{f}_i - F_i) \le \alpha \epsilon \ \forall \ i \in [p]$$

2.
$$\alpha v(s_i) \leq 1 \ \forall \ 0 \leq i \leq m$$

3.
$$\alpha \delta < 1, \alpha^2 \delta \leq 1$$

4.
$$\alpha^2 \epsilon < 1, \alpha^3 \epsilon \le 1$$

Then the following are equivalent:

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

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

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, $H_2 \Longrightarrow z$ 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, δ 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.

1.3 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 1.3.1

(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 $v(a) \le v(s)$

1.3: Hensel's Lemma Page 6

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

Proof. (i) Suppose $\deg s \geq \deg f \geq 0$ otherwise $\deg s < \deg f$ Then $s = 0 \times f + a = a$ so we get v(a) = v(s) and $v(p) = v(0) = 0 \leq v(s)$. Hence assume $l = \deg s - \deg f \geq 0$. Then $p = \sum\limits_{i=0}^{l} p_i x^i$ where $p_i \in R$ for all $0 \leq i \leq l$.

Now we will induct on l-i. Let $\deg f=n$ and $\deg s=m$. Hence assume $f=\sum\limits_{i=0}^n f_ix^i$ and $s=\sum\limits_{i=0}^m s_ix^i$ where for all $0\leq i\leq n$ and $0\leq j\leq m$, $f_i,s_j\in R$. Since f is monic $f_n=1$

Base Case (i = l):

$$v(p_l) = v(p_l)v(f_n) = v(p_lf_n) = v(s_m) \le v(s)$$

Inductive Step : $v(p_i) \le v(s)$ for all $l-k \le i \le l$. Now coefficient of $x^{(l-k-1)+n}$ in S is

$$s_{l-k-1+n} = \sum_{i=l-k-1}^{l} p_i f_{l-k-1+n-i}$$

Therefore $p_{l-k-1}f_n = s_{l-k-1+n} - \sum_{i=l-k}^{l} p_i f_{l-k-1+n-i}$. Therefore

$$\begin{split} v(p_{l-k-1}) &= v(p_{l-k-1}f_n) \leq \max\{v(s_{l-k-1+n}), v(p_if_{l-k-1+n-i}) \mid l-k \leq i \leq l\} \\ &\leq \max\{v(s_{l-k-1+n}), v(p_i) \mid l-k \leq i \leq l\} \\ &\leq v(s) \end{split}$$

Therefore by induction we have $v(p_i) \le v(s)$ for all $0 \le i \le l$. Then $v(p) \le v(s)$.

Hence

$$v(a) = v(s - pf) \le \max\{v(s), v(pf)\} \le \max\{v(s), v(p)\} \le v(s)$$

Therefore we have $v(a) \le v(s)$.

(ii) We will induct on $0 \le i \le m$.

Base Case : $v(h_0 - g_0^*) \le \epsilon$ Given

Inductive Step : Let this is true for i = k i.e.

$$v\left(\prod_{j=0}^k h_j - \prod_{j=0}^k h_j^*\right) \le \epsilon$$

Now

$$\prod_{j=0}^k h_j(h_{k+1} - h_{k+1}^*) + \prod_{j=0}^k h_j^*(h_{k+1} - h_{k+1}^*) = \left[\prod_{j=0}^{k+1} h_j - \prod_{j=0}^{k+1} h_j^*\right] - \left[h_{k+1}^* \prod_{j=0}^k h_j - h_{k+1} \prod_{j=0}^k h_j^*\right]$$

Therefore we have

$$\prod_{j=0}^{k+1} h_j - \prod_{j=0}^{k+1} h_j^* = \prod_{j=0}^{k} h_j (h_{k+1} - h_{k+1}^*) + \prod_{j=0}^{k} h_j^* (h_{k+1} - h_{k+1}^*) + \left[h_{k+1}^* \prod_{j=0}^k h_j - h_{k+1} \prod_{j=0}^k h_j^* \right]$$

Hence we have

$$\begin{split} v\left(\prod_{j=0}^{k+1}h_{j}-\prod_{j=0}^{k+1}h_{j}^{*}\right) &\leq \max\left\{v\left(\prod_{j=0}^{k}h_{j}\right)v(h_{k+1}-h_{k+1}^{*}),v\left(\prod_{j=0}^{k}h_{j}^{*}\right)v(h_{k+1}-h_{k+1}^{*}),v\left(h_{k+1}^{*}\prod_{j=0}^{k}h_{j}-h_{k+1}\prod_{j=0}^{k}h_{j}^{*}\right)\right\} \\ &\leq \max\left\{v(h_{k+1}-h_{k+1}^{*}),v\left(h_{k+1}^{*}\prod_{j=0}^{k}h_{j}-h_{k+1}\prod_{j=0}^{k}h_{j}^{*}\right)\right\} \\ &\leq \max\left\{\epsilon,v\left(h_{k+1}^{*}\prod_{j=0}^{k}h_{j}-h_{k+1}\prod_{j=0}^{k}h_{j}^{*}\right)\right\} \end{split}$$

Now we need to show that

$$v\left(h_{k+1}^*\prod_{j=0}^k h_j - h_{k+1}\prod_{j=0}^k h_j^*\right) \le \epsilon$$

Now

$$\begin{split} h_{k+1}^* \prod_{j=0}^k h_j - h_{k+1} \prod_{j=0}^k h_j^* &= \left(h_{k+1}^* \prod_{j=0}^k h_j - h_{k+1}^* h_k^* \prod_{j=0}^{k-1} h_j \right) + \left(\prod_{j=k-1}^{k+1} h_j^* \right) \left[\prod_{j=0}^{k-1} h_j \right] - \left[\prod_{j=k-1}^{k+1} h_j^* \right] \left[\prod_{j=0}^{k-2} h_j \right] \\ &+ \left(\prod_{j=k-1}^{k+1} h_j^* \right) \left[\prod_{j=0}^{k-2} h_j \right] - \left[\prod_{j=k-2}^{k+1} h_j^* \right] \left[\prod_{j=0}^{k-2} h_j \right] \right) \\ &\vdots \\ &+ \left(\prod_{j=t+1}^{k+1} h_j^* \right) \left[\prod_{j=0}^{t} h_j \right] - \left[\prod_{j=t}^{k+1} h_j^* \right] \left[\prod_{j=0}^{t-1} h_j \right] \right) \\ &\vdots \\ &+ \left(\left[\prod_{j=t+1}^{k+1} h_j^* \right] \left[\prod_{j=0}^{t} h_j^* \right] + \left(\prod_{j=0}^{k+1} h_j^* - \prod_{j=0}^{k} h_j^* \right) h_{k+1} \right) \\ &= h_{k+1}^* \prod_{j=0}^{k-1} h_j^* \left(h_k - h_k^* \right) + \left[\prod_{j=k}^{k+1} h_j^* \right] \left[\prod_{j=0}^{k-2} h_j \right] \left(h_{k-1} - h_{k-1}^* \right) \\ &+ \left[\prod_{j=k+1}^{k+1} h_j^* \right] \left[\prod_{j=0}^{k-3} h_j \right] \left(h_{k-2} - h_{k-2}^* \right) \\ &\vdots \\ &+ \left[\prod_{j=t+1}^{k+1} h_j^* \right] \left[\prod_{j=0}^{t-1} h_j \right] \left(h_t - h_t^* \right) \\ &\vdots \\ &+ \left[\prod_{j=t+1}^{k+1} h_j^* \right] \left(h_0 - h_0^* \right) + \left[\prod_{j=1}^{k} h_j^* \right] \left(h_{k+1}^* - h_{k+1} \right) \end{split}$$

Now for each $0 \le t \le k$ we have

$$v\left(\left[\prod_{j=t+1}^{k+1} h_{j}^{*}\right] \left[\prod_{j=0}^{t-1} h_{j}\right] (h_{t} - h_{t}^{*})\right) \leq v(h_{t} - h_{t}^{*}) \leq \varepsilon, \qquad v\left(\left[\prod_{j=1}^{k} h_{j}^{*}\right] (h_{k+1}^{*} - h_{k+1})\right) \leq v(h_{k+1}^{*} - h_{k+1}) \leq \varepsilon$$

Hence

$$v\left(h_{k+1}^*\prod_{j=0}^kh_j-h_{k+1}\prod_{j=0}^kh_j^*\right)\leq \max_{0\leq t\leq k}\left\{v\left(\left[\prod_{j=t+1}^{k+1}h_j^*\right]\left[\prod_{j=0}^{t-1}h_j\right]\left(h_t-h_t^*\right)\right),v\left(\left[\prod_{j=1}^kh_j^*\right]\left(h_{k+1}^*-h_{k+1}\right)\right)\right\}\leq \epsilon$$

Therefore we have

$$v\left(h_{k+1}^*\prod_{j=0}^k h_j - h_{k+1}\prod_{j=0}^k h_j^*\right) \le \epsilon \implies v\left(\prod_{j=0}^{k+1} h_j - \prod_{j=0}^{k+1} h_j^*\right) \le \epsilon$$

Hence by induction we have

$$v\left(\prod_{j=0}^{m}h_{j}-\prod_{j=0}^{m}h_{j}^{*}\right)\leq\varepsilon$$

1.3: Hensel's Lemma Page 8

Theorem 1.3.2 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

$$\mathcal{F}^* = \{f_i^* : 0 \le i \le m\}$$
 and $T = \{t_0, \dots, t_m\}$

such that

- (i) **Linear Case**: $S^* = S$ 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**: $S^* = T$ and $\delta^* = \alpha \gamma^2$, $\epsilon^* = \alpha \gamma \epsilon$. Assume that $\deg s_i > \deg f_i$ for $0 \le i \le m$. 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) $H_5(m, f, \mathcal{F}, \mathcal{F}^*, \mathcal{S}, \mathcal{S}^*, \alpha, \delta, \epsilon, \delta^*)$

1.3.1 Hensel's Computation

```
Algorithm 1: Hensel's Computation
  Input:
       1. f \in R[x], \mathcal{F} = \{f_0, \dots, f_m\} \subset R[x], \mathcal{S} = \{s_0, \dots, s_m\} \subset R[x]
       z \in R
       3. \alpha, \delta, \epsilon \in \mathbb{R}
  Output: \mathcal{F}^* = \{f_0^*, \dots, f_m^*\}, T = \{t_0, \dots, t_m\}
        Set \gamma = \max\{\delta, \alpha\epsilon\}, \alpha^* = \alpha, \epsilon^* = \alpha\gamma\epsilon \text{ and } e = f - \prod_{i=0}^m f_i
2
        for 1 \le i \le m do
3
              Compute a_i, b_i, p_i \in R[x] such that
4
                                          s_i e = p_i f_i + a_i, v(zb_i - a_i) \le \epsilon \gamma, \deg b_i \le \deg a_i < \deg f_i
        Compute a_0, b_0 \in R[x] such that
5
                                     a_0 = s_0 e + f_0 \sum_{i=1}^m p_i, \quad v(zb_0 - a_0) \le \epsilon \gamma, \quad \deg b_0 \le \deg f - f_i
        for 0 \le i \le m do
6
         \int_{i}^{*} f_{i} = f_{i} + b_{i}
        for 1 < i < m do
              Compute c_i, d_i, g_i^*q_i \in R[x] such that
                     g_i^* = \prod_{j \neq i} f_j^*, \quad s_i(s_i g_i^* - z) = q_i f_i^* + c_i \quad v(z d_i - c_i) \le \gamma^2 \quad \deg d_i \le \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_{i \neq i} s_j \frac{g_j^*}{f_i^*} \right) \right], \quad v(zd_0 - c_0) \leq \gamma^2, \quad \deg d_0 \leq \deg f - \deg g_0
        for 0 \le i \le m do
        t_i = s_i - d_i
```

1.3.2 Proof of Hensel's Lemma

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

CHAPTER 2

Bibliography

[vzG84] Joachim von zur Gathen. Hensel and newton methods in valuation rings. *Mathematics of Computation*, 42:637–661, 1984.