Computer Algebra

Summer semester 2008

Personal notes of
Yves Radunz

Contents

Lecture on 2008-04-14

Lecture on 2008-04-21

Chapter 1

Elementary big integer arithmetic (Chapt. 3 in Pethö)

Proposition 1.1. *For any base* $1 < \beta \in \mathbb{N}$ *and* $0 < a \in \mathbb{N}$ *there exists a unique* $l = l(a)$ *and coefficients* $a_j \in \mathcal{B} = \{0, ..., \beta\}$ *for* $j = 0, ..., l-1$ *such that* $a = \sum_{j=0}^{l-1} a_j \beta^j$ *with* $a_{l-1} \neq 0$.
Notation: $a = (a_0, a_1, ..., a_{l-1})_\beta$.

Proof:
Based on mappings from $(a_0, ..., a_{l-1}) \in \mathcal{B}^l$ into $\mathbb{N}$ being strictly monotonic with respect to lexicographic ordering in $\mathcal{B}^l$. □

Remark
Most popular choices are $\beta \in \{2, 10\}$ oriented towards hardware and I/O respectively. For computer algebra other word choices, e.g. $\beta = 2^{16}$, $\beta = 2^{32}$ or $\beta = 2^{15}, 2^{31} + sign$, may be advantadgeous, because such words, being elements of $\mathcal{B}$ can be efficiently transfered in memory and arithmetically manipulated.

Corollary 1.1. (1.2) *The length* $l(a)$ *satisfies uniquely*

$$\beta^{l(a)-1} \leq a_{l(a)-1}\beta^{l(a)-1} \leq a \leq \beta^{l(a)} - 1 < \beta^{l(a)}.$$

Proof:
No Proof. □

Algorithm (1.1 (Developement base $\beta \rightarrow$ number))
number alg1.1($\beta > 1, l, a_0, ..., a_{l-1} \in \mathbb{N}$){
    $a = a_{l-1}$;
    for ($i = l - 2$; $i \geq 0$; i--)
        $a = a \cdot b + a_i$;
    return a;
}

Inversion follows with modular arithmetic:
$a = \beta \cdot q + a_0 \Leftrightarrow a_0 = a \mod \beta$
$a - a_0 = \beta^2 \tilde{q} + a_1 \Rightarrow a_1 = \frac{(a - a_0) \mod \beta^2}{\beta}$
...

Algorithm (1.2 (number → Development base β))

number alg1.2($0 < a \in \mathbb{N}, 1 < \beta \in \mathbb{N}$){
  $i = 0$;
  while ($a > 0$) {
    $a_i = a \mod \beta$;
    $++i$;
    $a = a$ div β;
  }
}

Generally, we have the identity $a = \beta(a \text{ div } \beta) + a \mod \beta$ (Division by β with remainder).

Remark

From now on we assume that all integers are represented by there development with respect to β, i.e. $a = (\varepsilon, a_0, ..., a_{l-1})_\beta$ with $\varepsilon \in \{0, \pm 1\}$ being the sign bit. We assume that arithmetic operations and data transfers on coefficients in $\mathcal{B} \equiv \{0, ..., \beta - 1\}$ can be performed with unit complexity. The count of such operations is denoted by $\mathcal{B} - OPS_{ALG}(I)$, where ALG is a particular algorithm and I a set of input instances.

Addition of big integers

We have $\sum_{j=0}^{m-1} a_j\beta^j + \sum_{j=0}^{n-1} b_j\beta^j = \sum_{j=0}^{\max\{m,n\}-1}(a_j + b_j)\beta^j = \sum_{j=0}^{\max\{m,n\}} c_j\beta^j$, where missing a_j or b_j are padded with zeros.

When $a_j + b_j \geq \beta$ we need to carry one bit into the next power β^{j+1}. If that happens for the highest term, the length of $a + b$ is $l(a + b) = \max\{l(a), l(b)\} + 1$.

Algorithm (1.2 (Integer Addition ISUM))

integer[] isum($a = (a_0, ..., a_{m-1})_\beta, b = (b_0, ..., b_{m-1})_\beta$){
  $r = 0$;
  for($i = 0$; $i \leq m - 1$; i++) {
    $s = a_i + b_i + r$;
    if($s < \beta$) {
      $c_i = s$
      $r = 0$;
    } else {
      $c_i = s - \beta$;
      $r = 1$;
    }
  }
  $c_m = r$;
  return c;
}

Complexity: Bound(WorstCase)

$\mathcal{B} - OPS_{ISUM}(I) \leq O(m)$, where $I = \{(a, b) \in \mathbb{N} | l(a) \leq m \geq l(b)\}$.

Algorithm (1.5 (Naive multiplication - IMULT))

$c = \left(\sum_{j=0}^{m-1} a_j\beta^j\right)\left(\sum_{j=0}^{n-1} b_j\beta^j\right) = \sum_{k=0}^{m+n-2}\left(\sum_{j=0}^{k} a_j b_{k-j}\right)\beta^k = \sum_{j=0}^{m+n-1} c_j\beta^j$

W.l.o.g. we have $n \leq m$.

integer[] imult($a = (a_0, ..., a_{m-1})_\beta, b = (b_0, ..., b_{m-1})_\beta$){
  $r = 0$;
  for ($k = 0$; $k \leq m + n - 2$, k++) {
    for ($i = \max\{0, k - n + 1\}$; $i \leq \min\{k, m - 1\}$, i++)
      r+=$a_i b_{k-1}$;
    $c_k = r \mod \beta$;

```
        r = r div β;
    }
    c_{n+m-1} = r;
    return c;
}
```

Remark
The above algorithm may generate intermediate values r of size $\max(m,n)\cdot(\beta-1)^2 \gg \beta-1$. To fix that we perform a carry as soon as it becomes possible.

Algorithm (1.6. (Alternative Accumulation of c_k))

```
integer[] IMULT2(a = (a_0, ..., a_{m-1})_β, b = (b_0, ..., b_{m-1})_β){
    for (j = 0; j ≤ n − 1; j++){
        r = 0;
        for (i = 0; i ≤ m − 1; i++){
            r+=a_j b_i + c_{i+j};
            c_{i+j} = r  mod  β;
            r = r div β;
        }
        c_{j+m} = r;
    return c;
}
```

Lemma 1.1. *By induction follows that $r \le \beta-1, t \le \beta^2-1$.*

Proof:
Exercise. □

Lemma 1.2. (Complexity of Algorithm 1.6)
$\mathcal{B}-OPS_{IMULT}(I) = O(m\cdot n)$, *where* $I = \{(a,b)\in\mathbb{N}^2 | l(a)\le m, l(b)\le n\}$

(no proof)

□

Faster Alternative: Karatsuba method (Divide & Conquer)
Idea: Consider a,b with $l(a)=l(b)=2^p$ by padding. Take $n=l$ and $m=\frac{l}{2}$.

$$\left(\sum_{j=0}^{n-1} a_j\beta^j\right)\left(\sum_{j=0}^{n-1} b_j\beta^j\right) = \left(\underbrace{\sum_{j=0}^{m-1} a_j\beta^j}_{=A_0} + \beta^m \underbrace{\sum_{j=m}^{n-1} a_j\beta^{j-m}}_{=A_1}\right)\left(\underbrace{\sum_{j=0}^{m-1} b_j\beta^j}_{=B_0} + \beta^m \underbrace{\sum_{j=m}^{n-1} b_j\beta^{j-m}}_{=B_1}\right)$$
$$= (A_0+\beta^m A_1)(B_0+\beta^m B_1)$$
$$= A_0B_0 + \beta^{2m}A_1B_1 + \beta^m\left((A_0+A_1)(B_0+B_1) - A_0B_0 - A_1B_1\right)$$

Key observation:
The calculation of $a\cdot b$ requires the calculation of three products $A_0B_0, A_1B_1, (A_0+A_1)(B_0+B_1)$ of half the length $m=l/2$ plus some $O(n)$ housekeeping operations.

Complexity of Karatsuba
$M_n = \mathcal{B}-OPS_{IMULT}(I)$, where $I=\{(a,b)\in\mathbb{N}^2 | l(a)\le n=2^p\ge l(b)\}$ satisfies for some constant $r\in\mathbb{N}$ the recurrence $M_n \le 3M_{\frac{n}{2}} + \gamma n$.
$\Rightarrow \underbrace{M_n + 2\gamma n}_{=\tilde{M}_n} \le 3M_{\frac{n}{2}} + 3\gamma n = 3(M_{\frac{n}{2}} + 2\gamma(\frac{n}{2}))$
$\Rightarrow \tilde{M}_n \le 3\tilde{M}_{\frac{n}{2}} \le 3^p\tilde{M}_1 = 3^p(1+\gamma)$
$\Rightarrow M_n \le (1+\gamma)3^{\log_2 n} - 2\gamma n \le (1+\gamma)n^{\log_2 3}$

Lecture on 2008-04-28

Algorithm (Karatsuba)

integer karatsuba$(a,b)\{$ $\qquad a \in \mathbb{N} \ni b$
  $n = \max\{l(a), l(b)\}$; $\qquad$ base $\beta > 1, n \neq 2^p$ possible
  if $(n < n_0)$
    return basemultiply(a,b);
  else$\{$
    $m = \lceil \frac{n}{2} \rceil$;
    $A0 = a \mod \beta^m$; $A1 = a \text{ div } \beta^m$;
    $B0 = a \mod \beta^m$; $B1 = a \text{ div } \beta^m$;
    $C0 = \text{karatsuba}(A_0, B_0)$; $\qquad l(C0) \leq n+1$
    $C1 = \text{karatsuba}(A_1, B_1)$; $\qquad l(C1) \leq n+1$
    $C2 = \text{karatsuba}(A_0 + A_1, B_0 + B_1)$; $\qquad l(A0+A1) \leq m+1 \geq l(B_0+B_1)$
    $l(C2) \leq n+3$
    return $C0 + C1 \cdot \beta^{2m} + \beta^m(C2 - C0 - C1)$; $\qquad l(a \cdot b) \leq 2n+1$
  $\}$
$\}$

Remark

Alternative to this "additive" version there is a "subtractive" version where the middle term is formed with $A_0 - A_1$ and $B_0 - B_1$ avoiding carries. This requires some management of signs (see exercise).

According to Brad and Zimmerman the efficiency threshhold is $n_0 \approx 100$.

1.1 Division with Remainder

Lemma 1.3. (Basic Identity) *Let be $a, b \in \mathbb{N}, 1 < b \leq a$.*
Then, we have $a = bq + r$ with $q \equiv a$ div $b \in \mathbb{N}, r = a \mod b < b$.

Proof:

Proof and computation can be based on recursion. $a \text{ div } b = \begin{cases} 0 & \text{if } a < b \\ (a-b) \text{ div } b + 1 & \text{otherwise} \end{cases}$

Recursion could be implemented in base β arithmetic using roughly $a \text{ div } b + 1 \approx \frac{a}{b}$ subtractions $\approx \beta^{l(a)-l(b)} \cdot l(a)$ operations in $\mathcal{B} = \{0, ..., \beta - 1\}$. □

Long Division Algorithm base β

For $m \geq 0$ there is the equality $\underbrace{\sum_{j=0}^{n+m} a_j \beta^j}_{=a} = \underbrace{\left(\sum_{j=0}^{n-1} b_j \beta^j\right)}_{=b<a} \underbrace{\left(\sum_{j=0}^{m} q_j \beta^j\right)}_{=q} + \underbrace{\sum_{j=0}^{n-1} r_j \beta^j}_{=r}$.

The first "β-digit" $q_m \in \{1, ..., \beta - 1\}$ is defined by the subtask with $m = 0$
$\sum_{j=m}^{n+m} a_j \beta^{j-m} = \sum_{j=0}^{n} a_{m+j} \beta^j = \left(\sum_{j=0}^{n-1} b_j \beta^j\right) q_m + \sum_{j=0}^{n-1} \tilde{r}_j \beta^j$
Once the subtasks are solved, multiplication by β^m and subtraction from original yields
$\sum_{j=0}^{m-1} a_j \beta^j + \sum_{j=0}^{n-1} \tilde{r}_j \eta^{m+j} = \sum_{j=0}^{n+m-1} \tilde{a}_j \eta^j = \left(\sum_{j=0}^{n-1} b_j \beta^j\right) \left(\sum_{j=0}^{m-1} q_j \beta^j\right) + \sum_{j=0}^{n-1} r_j \beta^j$

In the subtask consider two highest terms
LHS: $a_n \beta^n + a_{n-1} \beta^{n-1} + O(\beta^{n-2})$ $\qquad$ RHS: $b_{n-1} q_m \beta^{n-1} + O(\beta^{n-2})$
$\Rightarrow a_n \beta^n + a_{n-1} \beta^{n-1} + \beta^{n-1} > b_{n-1} q_m \beta^{n-1}$
$\Rightarrow a_n \beta + a_{n-1} + 1 > b_{n-1} q_m$
This motivates the following result:

Proposition 1.2. (1.5) *For the subtask we have with $q = q_m$:*
$q \leq q_* = \min\{\beta - 1, \lfloor \frac{a_n \beta + a_{n-1}}{b_{n-1}} \rfloor\}$

Moreover if b is normalized such that $b_{n-1} \geq \lfloor \frac{\beta}{2} \rfloor$ (if $\beta = 2$ this means that the leading digit is 1), we get $q \geq q_* - 2$, so we have at worst three tries.

Proof:
See "Satz 3.5" in Pethö or "Theorem A" in 4.3.1 of Knuth. □

Lemma 1.4. (1.6) *Unless for $b_{n-1} \geq \lfloor \frac{\beta}{2} \rfloor$ already that condition holds, after multiplication of a, b and (implicitly) r by $d = \lfloor \frac{\beta}{b_{n-1}+1} \rfloor$ we get $a = bq + r \Leftrightarrow ad = (bd)q + (rd)$.*

Proof:
$l(b) = n \Rightarrow b < (b_{n-1} + 1)\beta^{n-1} \Rightarrow db < \lfloor \frac{\beta}{b_{n-1}+1} \rfloor (b_{n-1} + 1)\beta^{n-1} \leq \beta^{n-1}\beta = \beta^n \Rightarrow l(db) = n$
$\Rightarrow$ The new b_{n-1} is less or equal to $\beta - 1$ and there does not appear a carry.
Supposing $b_{n-1}\lfloor \frac{\beta}{b_{n-1}+1} \rfloor < \lfloor \frac{\beta}{2} \rfloor$ implies $\lfloor \frac{\beta}{b_{n-1}+1} \rfloor < \frac{1}{b_{n-1}} \lfloor \frac{\beta}{2} \rfloor \leq \lfloor \frac{\beta}{2b_{n-1}} \rfloor$.
$\Rightarrow \frac{\beta}{b_{n-1}+1} < \frac{\beta}{2b_{n-1}} \Rightarrow b_{n-1} + 1 > 2b_{n-1} \Rightarrow b_{n-1} < 1 \Rightarrow$ contradiction □

Remark
The normalization of b is required only once, not for every subtask.
Consequence: If $r^* = a - q^*b < 0$, we need to add b only once or twice to each $r \geq 0$ with $q_* + 2 \leq q \leq q_*$.
The total complexity $\mathcal{B}{-}OPS_{DIV}(a, b) \sim (l(a){-}l(b){+}1)l(b)$ is achieved by the following algorithm.

Algorithm (DIV)
integer[] div(a, b)\{
  $d = \lfloor \frac{\beta}{1+b_{n-1}} \rfloor$;
  $a = a \cdot d$; $b = b \cdot d$;
  for $(j = 0; j < m - n; j{+}{+})$\{
    $q^* = \min\{\beta - 1, \lfloor \frac{a_{m-j-1}+\beta a_{m-j}}{b_{n-1}} \rfloor\}$;
    while$((a_{m-j-1}...a_{m-1})_\beta - q^*b < 0)$
      q^*--;
    $q = \beta q + q^*$; $a = a - q^*\beta^{m-n-j}b$;
  \}
  return $(q, q/d)$;
\}

Assumed costs: $O(1)$ $\mathcal{B} - OPS$

Explanation: $(a_{m-j-1}...a_{m-1})_\beta = \sum_{k=0}^{n+j-1} a_{m-n-j+k}\beta^k$

Remark
The Implementation is a nontrivial project.

Lecture on 2008-05-05

Chapter 2

Euclidean and factorial rings or domains

2.1 General stuff

Throughout R is a commutative ring with 1.

Definition (2.1)

1. $a \in R$ is called *divisor* of $b \in R$ (short $a|b$), if there exists $c \in R$ such that $b = a \cdot c$.
2. $a, b \in R$ are called *zero divisors* if $a \cdot b = 0$, but $a \neq 0 \neq b$.
3. a is called *unit* if $a \cdot b = 1$ for some $b = a^{-1} \in R$.

Lemma 2.1. (2.2)

1. *$a|b$ is a partial ordering on R which is reflexive, but generally not anti-symmetric.*
2. *The units form a multiplicative group $U = U(R)$ so that $a \sim b \Leftrightarrow ac = b$ for some $c \in U$ is an equivalence relation. $a \sim b$ are said to be associated.*
3. *In a finite ring R every nonzero element is either a unit or zero-divisor.*

Proof:

1. In $\mathbb{Z}$ we have $-a|a$ and $a|-a$ but $a \neq -a$ if $a \neq 0$. $\Rightarrow$ not anti-symmetric.
2. $ab = 1 = a'b' \Rightarrow (aa')(bb') = 1 \Rightarrow (aa')^{-1} = bb'$
3. Look at the sequence of powers $(a^j)_j = 1, \ldots$. Due to assumed finiteness we must have $a^j = a^k$ for two indices $1 \leq j < k$.
   $\Rightarrow a^{k-j}(a^j - 1) = 0 \Rightarrow a(\underbrace{a^{k-j-1}(a^j - 1)}_{=\tilde{a}}) = 0$
   $\Rightarrow a$ and $\tilde{a}$ are zero divisors or $a^j - 1 = 0 \Rightarrow a^{j-1} = a^{-1}$. □

Definition (integral domain (2.3))
A ring R without zero-divisors is called *integral domain*.

Lemma 2.2. (2.4) *Let be R an integral domain.*

1. *$a|b \wedge b|a \Rightarrow a \sim b$*

2. *R has a (minimal) extension to field of quotients.*

3. *If $|R| < \infty$, then R is itself a field.*

Proof:

1. $a = bc \wedge b = da \Rightarrow b = dbc \Rightarrow b(1 - dc) = 0 \Rightarrow dc = 1 \Rightarrow d, c \in U \Rightarrow a \sim b$

2. —

3. —

Example

1. The integral domain $\mathbb{Z}$ is contained in its field of quotients $\mathbb{Q}$. There are further extensions of $\mathbb{Q}$ to $\mathbb{R}$ and $\mathbb{C}$.

2. $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z} = \mathbb{Z} \mod m, m > 1$ is integral domain iff m is prime.

   ($m = ab$ implies that $ab \mod m = 0 \Rightarrow a, b$ are zero divisors)

   Otherwise by Fermat: $a^{m-1} = 1 \Rightarrow a^{m-2} = a^{-1}$.

Question
For which integral domains exist "proper" greatest common divisors (gcd), i.e. in some sense unique maximal common divisors of a given pair a, b? (Maximality: $\forall e : e|a, e|b \Rightarrow e|gcd(a, b)$)
Answer: Not always. In some rings, e.g. in the quadratic extension $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} | a, b \in \mathbb{Z}\}$ there may be several maximal common divisors that are not even associated.

Definition (MAX (2.5))
Denoting by $MAX(S)$ the set of maximal elements with respect to the division ordering ("|") for $S \subseteq R$, we define the set $\gcd(a, b) = MAX\{c \in R \big| c|a \wedge c|b\}$.
The recursive extension to several arguments is unique due to subsequient lemma (1) and (4) $\gcd(a, bc) = \gcd(\gcd(a, b), c) = \gcd(a, \gcd(b, c))$.

Lemma 2.3. (2.5)

0. *$c|a \wedge c|b \Rightarrow c|\tilde{c}$ for some $\tilde{c} \in \gcd(a, b)$*

1. $\gcd(a, b) = \gcd(b, a)$ *by commutativity*

2. $\gcd(ca, cb) = c \gcd(a, b)$

3. $\gcd(a, b + a) = \gcd(a, b)$ *Basis for Euclidean Algorithm*

4. $\gcd(\gcd(a, b), c) = \gcd(a, \gcd(b, c))$ *Associativity*

5. $e \in U(R) \Rightarrow e \gcd(a, b) = \gcd(a, b)$ gcd*'s contain all associative elements*

Proof:
Exercise. □

Definition (2.7)

1. R is *gcd domain* if each $\gcd(a,b)$ for $a,b \in R$ consists exactly of one coset of U so that $c, \tilde{c} \in \gcd(a,b) \Rightarrow c \sim \tilde{c}$. (Consequence: $c|a \wedge c|b \Rightarrow c|\tilde{c}$ for all $\tilde{c} \in \gcd(a,b)$)
2. $a \in R$ is *prime* iff $a|(b \cdot c) \Rightarrow a|b \vee a|c$.
3. $a \in R$ is irreducible, iff $a = bc \Rightarrow b \in U \vee c \in U \Rightarrow a \sim c \vee a \sim b$.

Proposition 2.1. (2.8) *In an integral domain R primality implies irreducibility and the converse holds iff R is gcd domain.*

Proof:

1. "$\Rightarrow$" (by contradiction, i.e. we prove the contrapositive)
   Let be a irreducible.
   $\Rightarrow a = bc \wedge b \notin U \wedge c \notin U$
   $\Rightarrow a|bc \wedge a \not| b \wedge a \not| c$
   This is a contradiction to the primality of a.
2. "$\Leftarrow$"
   later.

Definition (Eucledian ring / Eucledian domain (2.9))
An integral domain K is called *Euclidean* (domain or ring) iff there exists a norm function $\varphi: R \to \mathbb{N}$ such that

1. $a \neq 0 \Rightarrow \varphi(a) > 0$
2. $\varphi(ab) = \varphi(a)\varphi(b)$
3. $a, b \in R \wedge b \neq 0 \Rightarrow \exists q, r \in R : a = bq + r$ with $\varphi(r) < \varphi(b)$

Remark
In literature one often uses instead of φ its logarithm $\deg(a) = \log_2 \varphi(a)$ (possibly scaled). Then, multiplicativity turns into $\deg(a \cdot b) = \deg(a) + \deg(b)$.

Lemma 2.4. (2.10)

1. $\varphi(0) = 0$ $\quad (\deg(0) = -\infty)$
2. $\varphi(a) = 1 \Leftrightarrow a \in U$
3. $b|a \wedge b \not\sim a \Rightarrow \varphi(b) < \varphi(a)$

Proof:
Left as exercise. □

Example

1. $R = \mathbb{Z}$ with $\varphi(a) = |a|$
2. $R = F[X]$ (set of polynomials with coefficients in F) with $\varphi(p) = 2^{\deg(p)}$.
3. Counterexample: $R = \mathbb{Z}[X]$, or other sets of polynomials with coefficients in rings that are not a field, are not Euclidean.

Proposition 2.2. (Existence of representation of gcd (2.11)) *A Euclidean domain is a gcd domain and the up to units unique $g = gcd(a,b)$ has the representation $g = sa + tb$ for $s, t \in R$.*

Proof:
Sketch: $S(a,b) = \{as + bt | s, t \in R\}$ is an ideal and Euclidean
$\Rightarrow S(a,b) = \{gt | t \in R\} \Rightarrow g$ is greatest common divisor and unique up to units. □

Algorithm (gcd)
The computation of gcd can be based on the observation $a = bq + r \Rightarrow \gcd(a, b) = \gcd(b, r)$.
If $\varphi(a) \geq \varphi(b)$ we get $\varphi(a) \geq \varphi(b) > \varphi(r)$.
Iterative notation from $a_0 = a, a_1 = b$ with $\varphi(a_0) \geq \varphi(a_1)$ generates a sequence a_k for $k = 2, ...$ such that $a_{k+1} = a_{k-1} \mod a_k = a_{k-1} - q_k a_k$ until for the first time $a_{n+1} = 0$.
Then $0 \neq a_n \wedge a_n | a_{n+1} \Rightarrow a_n = \gcd(a_{n-1}, a_n) = ... = \gcd(a_0, a_1) = \gcd(a, b)$.

Example
$R = \mathbb{Z}, a = 612, b = 228$

k	0	1	2	3	4	5
a_k	612	228	156	72	12	0
q_k		2	1	2	6	
s_k	1	0	1	-1	3	
t_k	0	1	-2	3	-8	

$a_4 = 12 = \gcd(612, 228) = s_4 a + t_4 b = 3 \cdot 612 - 8 \cdot 228$
To get s, t update identity $a_k = s_k a + t_k b$ starting from
$a_0 = 1a + 0b = s_0 a + t_0 b$
$a_1 = 0 \cdot a + 1 \cdot b = s_1 a + t_1 b$
$a_2 = (s_0 - q s_1) a + (t_0 - q t_2) b$
$\vdots$
$a_{k+1} = s_{k+1} a + t_{k+1} b$ with $s_{k+1} = s_{k-1} - q_k s_k$, $t_{k+1} = t_{k-1} - q_k t_k$
We get $a_{k-1} = a_{k+1} + q_k a_k \geq a_{k+1} + a_k$, hence a_k is bounded below by the Fibonacci sequence.

Lecture on 2008-05-19

Remark (Representation with matrices)
We have $\begin{pmatrix} a_k & s_k & t_k \\ a_{k+1} & s_{k+1} & t_{k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & -q_k \end{pmatrix}}_{=Q_k} \begin{pmatrix} a_{k-1} & s_{k-1} & t_{k-1} \\ a_k & s_k & t_k \end{pmatrix}$.
$\Rightarrow \begin{pmatrix} a_k & s_k & t_k \\ a_{k+1} & s_{k+1} & t_{k+1} \end{pmatrix} = \prod_{j=k,...,1} Q_j \begin{pmatrix} a & 1 & 0 \\ b & 0 & 1 \end{pmatrix}$.

Example
We consider $R = \mathbb{Z}_5[x]$ with $\mathbb{Z}_5 = \mathbb{Z}/5\mathbb{Z} = \mathbb{F}_5$ what is a finite field.
Then, we get:

k	0	1	2	3
a	$x^3 + 3$	$2x^2 + 2$	$4x + 3$	0
q	$-$	$3x$	$2x + 4$	0
s	1	0	1	$2x + 1$
t	0	1	$-3x$	$4x^2 + 2x + 1$

$\Rightarrow \gcd(x^3 + 3, 2x^2 + 2) = 4x + 3 \sim 16x + 12 \sim x + 2$

Observation
Here as for polynomials in general the degree of $a_k = a_k(x)$ typically decreases by 1 at each stage. Hence the number of divisions in Euclid is bounded by $m \leq \deg(a_0) - \deg(a_1) = \log_2(\frac{\varphi(a_0)}{\varphi(a_1)})$.

Lemma 2.5. (2.12) *In $R = \mathbb{Z}$ the number of divisions m needed for $a > b > 0$ is bounded by $m = O(\log a)$.*

Proof:
It follows from $a_m > 0 < a_{m-1}$ and $a_{k-1} = a_k q_k + a_{k+1} \geq a_k + a_{k+1}$ that $a_0 \geq F_{m+1}$ (which is the $(m+1)$-st Fibonaccy number.

It is well known that F_{m+1} has the explicit representation
$F_{m+1} = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^{m+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{m+1}}{\sqrt{5}} \geq \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{m}$
$\Rightarrow a_0 \geq \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{m} \Rightarrow m \leq \frac{\log a + \log\sqrt{5}}{\log(1+\sqrt{5}) - \log 2} = O(\log a)$ since $a \geq 2$ □

Complexity of gcd
Consequence:
Since $a_k \geq \gcd(a_0, a_1)$ and $\mathcal{B} - OPS_{DIV}(a_k, a_{k+1}) = O(l(a_{k+1})(l(a_k) - l(a_{k+1}) - 1))$ for each k, we have overall the simple bound $\mathcal{B} - OPS_{GCD}(a, b) = O(l(a)l(b)(l(a) - l(gcd(a, b)) + 1))$.
In other words we have a complexity bound in terms of word operations, which still grows like the third power of $\min\{l(a), l(b)\}$. A quadratic rather than cubic bound is obtained as follows.

Proposition 2.3. *For $a > b > 0$ in $R = \mathbb{Z}$ we have $\mathcal{B} - OPS_{GCD}(a, b) = O(l(a)l(b))$.*

Proof:
Since $l(a_{k-1}) \leq l(a_k) + l(q_k) + 2$ and $l(q_k) \geq 1$, $k \leq m$ we get $l(a_{k-1}) - l(a_k) + 1 \leq l(q_k) + 3 = O(l(q_k))$.
$\Rightarrow \sum_{k=0}^{m-1} O(l(a_k) - l(a_{k+1}) + 1) = \sum_{k=0}^{m-1} O(l(q_{k+1}))$
Furthermore, we have

$$\mathcal{B} - OPS_{DIV}(a_k, a_{k+1}) = O(l(a_{k+1})(l(a_k) - l(a_{k+1}) - 1)) = l(b)O(l(a_k) - l(a_{k+1}) + 1).$$

Since $a_k = a_{k+1}q_{k+1} + a_{k+2} \geq a_{k+1}q_{k+1}$ we know $q_{k+1} \leq \frac{a_k}{a_{k+1}}$.
Hence $\prod_{k=0}^{m-1} q_{k+1} \leq \frac{a_0}{a_m} = \frac{a}{\gcd(a,b)} \leq a$.
Since $q_{k+1} \geq \beta^{l(q_{k+1})-1}$ we get $x = \sum_{j=0}^{l(x)-1} x_j\beta^j \geq x_{l(x)-1}\beta^{l(x)-1} \geq \beta^{l(x)-1}$ and $\prod_{k=0}^{m-1} \beta^{l(q_{k+1})-1} \leq \beta^{l(a)}$.
$\Rightarrow \sum_{k=0}^{m-1} l(q_{k+1}) - 1 \leq l(a)$
$\Rightarrow \sum_{k=0}^{m-1} l(q_{k+1}) \leq l(a) + m = l(a) + O(\log(a)) = O(l(a))$
Hence we conclude that
$\sum_{k=0}^{m-1} O(l(a_{k+1})(l(a_k) - l(a_{k+1}) + 1)) = O(l(a_1)\sum_{k=0}^{m-1} O(q_{k+1})) = O(l(b)O(l(a))) = O(l(b)l(a))$
□

Remark
Pethö establishes the sligthtly stronger bound $O(l(b)(l(a) - l(\gcd(a, b)) + 1))$ at some considerable expense in terms of proof complexity. D.H. Lehmer obxerved that the quotient sequence $q_1, ..., q_m$ is typically for $k \in [1, \overline{k}]$ identical to the one obtained for a pair of leading terms $a' = a \text{ div } \beta^n$ and $b' = b \text{ div } \beta^m$ for some $m < \min\{l(a), l(b)\}$.

Example
Consider the pair $(880458, 307373) \rightsquigarrow (880, 308), (881, 307)$.
$\Rightarrow \frac{880}{308} < \frac{880458}{307373} < \frac{881}{307}$
It can be shown (Exercise 2.4c, see also Pethö) that as long as the quotients obtained for lower and upper bound agree with each other they are also identical to the ones for the fraction in the middle. Then apply $Q'_k = \begin{pmatrix} 0 & 1 \\ 1 & q_k \end{pmatrix}$ to original term and continue from remaining (a_k, a_{k+1}) with same scheme.

Remark (Final complexity of gcd **on** $R = \mathbb{Z}$**)**
Using the above idea of Lehmer (plus FFT based multiplication and division) Knuth and Schönhage designed versions of Euclid with complexity $O(l(a)(\log a)^p \log\log a)$ with $p = 5$ and $p = 2$, respectively. This algorithms are essentially linear in $l(a) = \max\{l(a), l(b)\}$.

Application of Euclid for $\bmod$ **inverses**
Given $a, m \in R$ (R Euclidean) we get $ms + at = \gcd(m, a) = g$.
This means that in $\bmod m$ arithmetic, i.e. in $R_m = R/mR$, we have $ta \bmod m = (ms + ta) \bmod m = \gcd(m, a) \bmod m = g \bmod m$.
Now, if $g \in \mathcal{U}(R)$ is a unit (i.e. $g = 1$) then t is exactly the inverse of a in $\bmod m$ arithmetic.

Example (Inverses (1))
Let's consider $m = 21$ (not prime!) and $a = 16$. Obviously, we have $\gcd(m, a) = 1$.
Using Euclid, we get:

a	21	16	5	1	0
q	$-$	1	3	5	0
t	0	1	-1	4	

$\Rightarrow t_3 = 4 = 16^{-1} \bmod 21$

Example (Inverses (2))
Let be $m(x) = x^5 + x^2 + 1 \in \mathbb{Z}_2[x]$ and $a(x) = x^2 + 1$.

a	$x^5 + x^2 + 1$	$x^2 + 1$	x	1
q	$-$	$x^3 + x + 1$	x	x
t	0	1	$x^3 + x + 1$	$x^4 + x^2 + 1$

Hence $t(x) = x^4 + x^2 + x + 1 = (x^2 + 1)^{-1} \bmod x^5 + x^2 + 1$.

Lecture on 2008-05-26

Lecture on 2008-06-02

Goal
We want to apply newton's Method to polynomials.

Definition (Newton's method)
Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a C^1-function. Let $x^* \in \mathbb{R}^d$ such that $f(x^*) = 0$ and $x^{(1)} \in \mathbb{R}^d$ somehow "near" to x^*. Then, the sequence given by

$$x^{(k+1)} = x^{(k)} - Df(x^{(k)})^{-1} f(x^{(k)}), \text{ i.e. } f(x^{(k)} + Df(x^{(k)})(x^{(k+1)} - x^{(k)}) = 0,$$

converges. (We have $x^{(k)} \to x^*$.)

Definition (formal and partial derivatives (3.7))
Let R be a commutative ring with 1 and $f = \sum_{k=0}^{n} a_k x^k \in R[x], a_n \neq 0$.
The *formal derivative* of f is given by $f' = \sum_{k=1}^{n} a_k k x^{k-1} = \sum_{k=0}^{n-1} a_{k+1}(k+1)x^k$.
Attention: The variable k in the term $a_k k x^{k-1}$ has the meaning of $\underbrace{1 + 1 + \ldots + 1}_{k \text{ times}}, 1 \in R$.
Analogously, we define the *partial derivative* $\partial_j f(x_1, \ldots, x_d)$ of some $f \in R[x_1, \ldots, x_d]$ and the Jacobian $Df(x_1, \ldots, x_d) \in R^{n,d}$ of $f = (f_1, \ldots, f_n) \in R[x_1, \ldots, x_d]^n$.

Lemma 2.6. (Taylor Approximation (3.8)) *Let R be a commutative ring with 1, $y_1, \ldots, y_d$ some unknowns and $f \in R[x_1, \ldots, x_d]$.*
$\Rightarrow f(x_1 + y_1, \ldots, x_d + y_d) \equiv f(x_1, \ldots, x_d) + \sum_{k=1}^{d} \partial_k f(x_1, \ldots, x_d) y_k f(x_1, \ldots, x_d) + Df(x_1, \ldots, x_d) \cdot y$
$\bmod (y_1, \ldots, y_d)^2$ *with* $y = \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix}$.
(The ideal $(y_1, \ldots, y_d)^2$ is used in the sense of definition 3.2.)

Proof:
Since derivation is a linear operation we may assume w.l.o.g. $f = \prod_{j=1}^{d} x_j^{n_j}$.
We define $I = (y_1, \ldots, y_d)$.
$\Rightarrow \partial_k f(x_1, \ldots, x_d) = n_k x^{n_k - 1} \prod_{j \neq k} x_j^{n_j}$
$\Rightarrow f(x_1 + y_1, \ldots, x_d + y_d) = \prod_{j=1}^{d} (x_j + y_j)^{n_j} = \prod_{j=1}^{d} (x_j^{n_j} + n_j y_j x^{n_j - 1} + I^2(\ldots))$
$\equiv_{\mod I^2} \prod_{j=1}^{d} (x_j^{n_j} + n_j y_j x_j^{n_j - 1})$
$= \prod_{j=1}^{d} x_j^{n_j} + \sum_k n_k y_k x_k^{n_k - 1} \prod_{j \neq k} x_j^{n_j} + I^2(\ldots)$
$\equiv_{\mod I^2} f(x_1, \ldots, x_d) + Df(x_1, \ldots, x_d) y$ □

Proposition 2.4. (Linear Hensel Lifting (3.9)) *Let R be a commutative ring with 1, I a finitely generated ideal in R, $f = (f_1, \ldots, f_d) \in R[x_1, \ldots, x_d]^n$ and $z \in R^d$ such that $f(z) \equiv 0 \mod I$. Further let be $W \in R^{d,h}$ the right inverse of $Df(z) \mod I$, i.e. $Df(z)W \equiv 1_n \mod I$. Then one can construct a sequence $(z^{(k)})_{k \in \mathbb{N}} \subseteq R^d$ with $z^{(k)} \equiv z \mod I$ and $f(z^{(k)}) \equiv 0 \mod I^k$.*

Proof:
Let be $z^{(1)} = z$.
Now, we use induction: Assume $z^{(k)}$ is given and $f(z^{(k)}) \equiv 0 \mod I^k$.
Ansatz (inspired by Newton): $z^{(k+1)} = z^{(k)} - \underbrace{Wf(z^{(k)})}_{= \Delta z \in I^k, \text{ componentwise}}$
$\Rightarrow f(z^{(k+1)}) = f(z^{(k)} + \Delta z) \equiv_{(Delta z)^2 \subseteq I^{2k} \subseteq I^{k+1}} f(z^{(k)}) - Df(z^{(k)}) W f(z^{(k)})$ (Taylor)
$\Rightarrow f(z^{k+1}) \equiv (1_n - Df(z^{(k)}) W) \cdot \underbrace{f(z^{(k)})}_{\in I^k}$ (*)
Since $z^{(k)} \equiv_I z$, we get $Df(z^{(k)}) \equiv_I Df(z)$
$\Rightarrow 1_n - Df(z^{(k)})W \equiv_I 1_n - Df(z)W \in I$
$\Rightarrow$ The first factor of the equality (*) is contained in I.
$\Rightarrow f(z^{k+1}) \in I^{k+1}$ □

Remark
We used a fixed W in each step. If a $W^{(k)} \in R^{d,n}$ with $Df(z^{(k)})W^{(k)} \equiv 1 \mod I^{2^k}$ is available, then "quadratic convergence" is possible, i.e. $f(z^{(k)}) \in I^{2^k}$.
Proof: $\Delta z = -W^{(k)} f(z^{(k)}) \in I^{2^k} \Rightarrow f(z^{(k+1)}) \equiv_{I^{2^k}} (1 - Df \cdot W) f(z^{(k)}) \in (I^{2^k})^2 = I^{2^{k+1}}$ □
Such $W^{(k)}$'s really exist!
(Apply proposition 3.9 to $F_j^{(k)}(x_1, \ldots, x_d) = Df(z^{(k)})x - e_k, e_k = (0, \ldots, 0, 1, 0, \ldots, 0)^t$. The j-th column of $W^{(k)}$ is a root of $F_j^{(k)} \mod I^k$.)

Example ($R = \mathbb{Z}$)
Consider $f(x_1, x_2) = x_1 x_2 - x_2^2 - 10$ and $I = (2)$.
$\Rightarrow Df(x_1, x_2) = (x_2, x_1 - 2x_2) \Rightarrow f \equiv x_1 x_2 - x_2^2 \mod I$
$z = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = z^{(1)} \Rightarrow f(z^{(1)}) = -10 \equiv_2 0$
$\Rightarrow Df(z) = (0, 1)$

1. We take $W = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

   $\Rightarrow z^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, f(z^{(1)}) = -10, \Delta z^{(k)} = -W f(z^{(1)}) = \begin{pmatrix} 10 \\ -10 \end{pmatrix}$

   $z^{(2)} = \begin{pmatrix} 11 \\ 10 \end{pmatrix}, f(z^{(2)}) = 0$ (exact solution!)

2. With $W = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ instead of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, we get

$z^{(1)} = \begin{pmatrix}1\\0\end{pmatrix}, f(z^{(1)}) = -10, \Delta z^{(2)} = \begin{pmatrix}0\\10\end{pmatrix}$

$z^{(2)} = \begin{pmatrix}1\\10\end{pmatrix}, f(z^{(2)}) = -100, \Delta z^{(3)} = \begin{pmatrix}0\\100\end{pmatrix}$

$z^{(3)} = \begin{pmatrix}1\\110\end{pmatrix}, f = -12000 \equiv 0 \mod 2^3 = 8$

$z^{(3)} + 8\begin{pmatrix}n\\m\end{pmatrix}$ is also a solution $\mod 8$ for all $n, m \in \mathbb{Z}$. Hence, we could try to find an exact solution with smaller values in the first and/or second component.

Remark (Convergence)
For a fixed prime p we define $ord_p : \mathbb{Z} \to \mathbb{N}_0 \cup \{\infty\}$ by $0 \neq x = \pm\prod_{q\in\mathbb{P}} q^{ord_q(x)}$, e. g.
$ord_3(20) = ord_3(2^2 \cdot 5) = 0$,
$ord_2(20) = 2$.
For $x = 0$ we define $ord_p(0) = \infty$.
(One could say that $ord_p(x)$ is the highest power of p such that x is divisible by this power of p. Since 0 always divisible by every power of a prime number p, we get $ord_p(0) = \infty$.)
We define the p-absolute value $|\cdot|_p$ on $\mathbb{Z}$ by $|x|_p = p^{-ord_p(x)}$. This is an absolute value: It is multiplicative, positive definite and fullfilles the triangle inequalities.
$\Rightarrow d_p(x, y) = |x - y|_p$ is a metric.
Applying Hensel with $R = \mathbb{Z}, n = 1, I = (p)$ for the sequence $(z^{(k)})_k$ we get:
$z^{(k+1)} - z^{(k)} = \Delta z^{(k+1)} \in I^k = (p^k)$
$\Leftrightarrow |z^{(k+1)} - z^{(k)}|_p \leq p^{-k} \overset{k\to\infty}{\longrightarrow} 0$
$\Rightarrow (z^{(k)})$ is a Cauchy sequence.
$\Rightarrow (z^{(k)})$ converges in some larger space.
(Note: That space is the set of p-adic numbers $\mathbb{Q}_{(p)}$. But $f(z^{(k)}) = 0$ for some $k \in \mathbb{N}$ does usually not happen!)

Lecture on 2008-06-09

Chapter 3

Factorization in Non Euclidean Rings

3.1 There is no suitable headline yet.

According to Magnotte "modern" algorithms for factorization of $F \in \mathbb{Z}[x]$ proceed as follows:

0. Reduce $F(x)$ to square free form
1. Chose suitable prime and factor $f \equiv F \mod p$ in $\mathbb{Z}_p$.
2. Refine previous factorization modulo p to modulo p^n with n sufficiently large (Hensel lifting)
3. Check if factorization is correct in that all factors are truly divisors of F (may be too fine)

Review of relations between integral domain properties
In short, we have:
Euclidian ⇒ factorial ("unique" factorization) ⇒ gcd ⇒ irreducible ⇒ primal

Proposition 3.1. (4.1) *Let be k an arbitrary field. Then the ring $k[x]$ is Euclidean and thus factorial, i.e. $0 \neq P \in k[x]$ admits the factorization $P(x) = c\prod_{i=1}^{m} P_i(x)^{\alpha_i}$ where $c \in k$, $P_i(x)$ suitable and pairwise non-associated and $\alpha_i > 0$. This decomposition is unique up to permutation of the index i and rescaling P_i and $c \in k$.*

Observation
For $R[x]$ over a nonfield R division with remainder may not work, hence the ring $R[x]$ is not Euclidean.
Standard example: Take $x^2 + 1, 2x \in \mathbb{Z}[x]$.

Definition (4.2)
If R is factorial, and thus gcd, then $P(x) = \sum_{i=0}^{n} a_i x^i$ is called *primitive* if the socalled *content* $cont(P) = \gcd(a_n, a_{n-1}, ..., a_0)$ is associated to 1, i.e. if it is a unit.
In any case the *primitive part* $pp(P) = \frac{P(x)}{cont(P)}$ is primitive.

Lemma 3.1. (Gaußlemma) *The product of two (or more) primitive polynomials is also primitive.*

Proof:
Let be $P = \sum_{i=0}^{n} a_i x^i, Q = \sum_{j=0}^{n} b_j x^j$ and $R = PQ = \sum_{i=0}^{n+m} c_i x^i$.
Suppose $cont(R) = \gcd(c_0, ..., c_{n+m}) = d = pd', p$ prime. Such p and d' exist, since $d \neq 1$ (i.e. R is not primitive).
⇒ $c_{n+m} = a_n b_m$ is a multiple of d and thus of p.

$\Rightarrow$ W.l.o.g. $p|a_n$ (otherwise interchange P and Q if necessary).
$\Rightarrow$ There exists a maximal $k < n$ with $p \not| a_k$ because otherwise $p \in cont(P) \in \mathcal{U}$.
$\Rightarrow$ $\underbrace{c_{k+m}}_{\text{divisible by } p} = \underbrace{a_n b_{k+m-n} + a_{n-1} b_{k+m-n+1} + ... + a_{k+1} b_{m-1}}_{\text{divisible by } p} + a_k b_m$
$\Rightarrow p|a_k b_m \Rightarrow p|b_m$ (p can not divide a_k by definition of a_k)
$\Rightarrow$ There exists a maximal $j < m$ such that $p \not| b_j$.
$\Rightarrow c_{k+j} = a_k b_j + \sum_{i \neq 0} \underbrace{a_{k-i} b_{j+i}}_{\text{divisible by } p}$
$\Rightarrow p|a_k b_j$ what is a contradiction. □

Proposition 3.2. (4.4) *If R is factorial so is $R[x]$ and $R[x_1, ..., x_n]$ with $n \in \mathbb{N}$.*

Proof:
Let Q_R be the field of quotients generated by R.
Then any $P(x) \in R[x] \subseteq Q_R[x]$ has a factorization of the form $P(x) = c\tilde{P}_1(x) \cdot ... \cdot \tilde{P}_m(x)$ with $\tilde{P}_i \in Q_R[x]$ and $c \in Q_R$. Hence each $\tilde{P}_i$ can be written as $\tilde{P}_i(x) = \sum_{j=0}^{n_i} \frac{a_{i,j}}{b_{i,j}}, \gcd(a_{i,j}, bi_j) = 1$ (otherwise cancel the gcd).
Now, we define the factors $d_i = \begin{cases} \gcd(a_{i,0}, ..., a_{i,n_i}) = cont(\tilde{P}), & \text{if} b_{i,j} = 1, 0 \leq j \leq n_j \\ \frac{1}{\text{lcm}(b_{i,0}, ..., b_{i,n_i})}, & \text{if otherwise} \end{cases}$.
We set $P_i(x) = \frac{1}{d_i} \tilde{P}_i(x) \in R[x], i = 1, ..., m$. All the P_i are primitive.
By the GaußLemma, $\prod_{i=1}^m P_i(x)$ is also primitive.
Now, consider the product $\left(\prod_{i=1}^m P_i(x)\right) \left(c \prod_{i=1}^n d_i\right) = P(x)$.
$c \prod_{i=1}^n d_i \in R \Rightarrow c \prod_{i=1}^n d_i = \prod_{i=1}^r p_i$ with p_i prime in R.
$\Rightarrow p_1 ... p_r P_1(x) ... P_m(x) = P(x)$ represents the factorization into irreducible factors in $R[x]$. □
The last assumption follows since $P(x_1, ..., x_n) \in R[x_1, ..., x_n]$ can be interpreted as an element of the polynomial ring $R[x_1, ..., x_{n-1}][x_n]$, i.e. the ring of polynomials in x_n whose coefficients belong to $R[x_1, ..., x_{n-1}]$.

Observation
The Euclidean division works in $R[x]$, provided the leading coefficient of the divisor is a unit. In particular:
$\underbrace{(a_n x^n + ... + a_1 x + a_0)}_{=P(x)} / (x - a) = \underbrace{(a_n x^{n-1} + ... + \tilde{a}_0)}_{=Q(x)} + \frac{b}{x-a}$
$\Rightarrow P(x) = Q(x)(x - a) + b$

Lemma 3.2. (4.5) *a is a root of $P \in R[x]$ (i.e. $P(a) = 0$) $\Leftrightarrow P(x) = Q(x)(x - a)\alpha, Q(a) \neq 0$*
The exponent $\alpha \in \mathbb{N}$ is called the multiplicity *of the root a.*

Definition
$P(x) \in R[x]$ is called *square free*, if $Q(x)|P(x) \wedge \deg Q > 0 \Rightarrow Q(x)^2 \not| P(x)$.

Goal
In step 0 of the above factorization we want to reduce P in a way that it is square free.

Proposition 3.3. (4.6) *If $P(x) = \prod_{i=1}^m (x - a_i)^{\alpha_i}$ then we obtain with $P'(x)$ (the algebraic derivative satisfying $(cx^n)' = ncx^{n-1}$) the square free reductions $Q(x) = \frac{P(x)}{\gcd P(x), P'(x)} = \prod_{i=1}^m (x - a_i)^1$.*
Furthermore, we have $\frac{Q(x)}{\gcd(Q(x), P'(x))} = \prod_{\alpha_i = 1} (x - a_i)$.

Lecture on 2008-06-16

Consider a factorial domain with $a_i, i = 0, ..., m$. Then, we have $a_i = \prod_{j \in J} p_j^{\alpha_{i,j}}$ where the p_j are irreducible nonassociated elements.
$\Rightarrow \gcd(a_0, ..., a_m) = \prod_{j \in J} p_j^{\alpha_j}$,

$\underline{\alpha_j} = \min_{0\le i\le m}\{\alpha_{i,j}\}$ *and*
$\operatorname{lcm}(a_0,...,a_m) = \prod_{j\in J} p_j^{\overline{\alpha_j}}$,
$\overline{\alpha_j} = \max_{0\le i\le m}\{\alpha_{i,j}\}$.

Consequence:
Rescale of $P(x) = \sum_{i=0}^{m} \frac{a_i}{b_i} x^i \in Q_R[x]$, $\gcd(a_i, b_i) = 1$ *to obtain* $P(x) = dP'(x) \in R[x]$ *(*P' *primitive) with* $d = \frac{\operatorname{lcm}(b_0,...,b_m)}{\gcd(b_0,...,b_m)}$.

Proof:
Factorize up to units uniquely: $a_i = \prod_{j\in J} p_j^{\alpha_{i,j}}, b_j = \prod_{j\in J} p_j^{\beta_{i,j}}$.
Divide by $\gcd(a_0,...,a_m) = \prod_{j\in J} p_j^{\underline{\alpha_j}}$ beforehand so that afterwards without loss of generality $\gcd(a_0,...,a_m) = 1$. $\operatorname{lcm}(b_0,...,b_m) = \prod_{j\in J} p_j^{\overline{\beta_{i,j}}}$ yields

$$P(x) = dP'(x) = \sum_{i=0}^{m} \frac{a_i}{b_i} dx^i = \sum_{i=0}^{m} x^i \prod_{j\in J} p_j^{\alpha_{i,j}+\overline{\beta_j}-\beta_{i,j}} \in R[x].$$

It remains to be shown that $cont(P(x)) = \gcd_{i=0}^{m}(\prod_{j\in J} p_j^{\alpha_{i,j}}) = 1$.
Suppose otherwise: For some j and all i we have $\alpha_{i,j} + \overline{\beta_j} - \beta_{i,j} > 0$, $\overline{\beta_j} = \beta_{k,j}$ for some k.
$\Rightarrow \alpha_{k,j} > 0 \Rightarrow \beta_{k,j} = 0$ (since $\gcd(a_k, b_k) = 1$)
$\Rightarrow \overline{\beta_j} = 0$
$\Rightarrow \beta_{i,j} = 0$ for all i
$\Rightarrow \alpha_{i,j} > 0$ for all i, what contradicts $\gcd(a_0,...,a_m) = 1$ □

3.2 GCD computation in factorial domains

TODO
subsection 4.1 machen

Corollary 3.1. (from the Gauss Lemma (4.3)) *Let be* R *factorial and* $P, Q \in R[x]$. *Then, we have:*

1. $pp(PQ) \sim pp(P) \cdot pp(Q)$
2. $cont(PQ) \sim cont(P) \cdot cont(Q)$
3. $cont(\gcd(P,Q)) \sim \gcd(cont(P), cont(Q))$
4. $pp(\gcd(P,Q)) \sim \gcd(pp(P), pp(Q))$
5. $gcd(P,Q) = \gcd(P, \frac{1}{q}Q)$ *if* P *is primitive and* $q|cont(Q)$

Algorithm (simple gcd**-algorithm)**

1. $d = \gcd(P,Q)$ in $Q_R[x]$ (Euclidean)
2. $g = \gcd(pp(P), pp(Q)$ (in $Q_R[x]$)
3. renomalize g to become primitive in $R[x]$ (as above)
4. return $dg \sim \gcd(P,Q)$

Remark
Disadvantage: The detour through the field of quotients can be very costly (Knuth).
An alternative is the generalized Euclidean algorithm which is based on pseudo division.

Lemma 3.3. (4.8) *Given* $a(x) = \sum_{j=0}^{m} a_j x^j$ *and* $b(x) = \sum_{j=0}^{m} b_j x^j$ *there exist* $q(x) = \sum_{j=0}^{m-n} q_j x^j$ *and* $r(x) = \sum_{j=0}^{n-1} r_j x^j$ *(all in* $R[x]$*) such that* $b_n^{m-n+1} a(x) = b(x)q(x) + r(x)$.
$r(x) = pres(a(x), b(x))$ *is called* pseudo residual.

Proof:
We prove this lemma by induction on $m - n$.
In the case of $m - n = 1$, we have $b^{m-n+1}a(x) = a(x) = r(x)$ and $deg(r(x)) < n = deg(b(x))$.
Now, let's work on the reduction of $m-n$ to $m-n-1$, achieved by calculation: Leading coefficiens of $q(x)$ are set to remainders residual $q_{m-n} = b_n^{m-n} a_m$, $\tilde{a}(x) = b_n a(x) - a_m b(x) x^{m-n}$.
$\Rightarrow$ The $(m-n)$-th power of x in $\tilde{a}()x)$ has the coefficient $b_n a_m - a_m b_n = 0$.
$\Rightarrow deg(\tilde{a}(x)) < deg(a) \Rightarrow\} \tilde{m} = m - 1$.
$\Rightarrow$ This process can be repeated until $m = n - 1$.
By induction hypothesis there exist $\tilde{q}(x)$ with $deg(\tilde{q}) = m - n - 1 = \tilde{m} - n$ and $r(x) \in R[x]$ with $deg(r(x)) < n$ such that $b_n^{m-n}\tilde{a}(x) = b(x)\tilde{q}(x) + r(x)$.
Substituting the definition of $\tilde{a}(x)$ implies $b_n^{m-n}(b_n a(x) - b_n^{m-n} a_m b(x) x^{m-n}) = b(x)\tilde{q} + r(x)$.
$\Rightarrow b_n^{m-n+1} a(x) = b(x)(b_n^{m-n} a_m x^{m-n} + \tilde{q}(x)) + r(x)$, where $b_n^{m-n} a_m x^{m-n}$ is a monomial $q(x)$ of degree less or equal $m - n$. □

Example (Knuth)
Consider $a(x) = x^8 + x^6 - 3x^4 + 3x^3 + 8x^2 + 2x - 5$, i.e. $m = 8$, $a_8 = 1$, and
$b(x) = 3x^6 + 5x^4 - 4x^2 - 9x + 21$, i.e. $n = 6, b_6 = 3$.

1. $k = m - n = 2$

   $q_k = q_2 = 9 = 1 \cdot 3^2$

   $\tilde{a}(x) = 3x^8 + 3x^6 - 9x^4 - 9x^3 + 24x^2 + 6x - 15 - 3x^8 - 5x^6 + 4x^4 + 9x^3 - 21x^2$
   $= 0 \cdot x^7 - 2x^6 - 5x^4 + 3x^2 + 6x - 15$

2. $k = m - n = 1$

   $q_1 = \tilde{a}_7 \cdot 3^1 = 0$

   $\tilde{\tilde{a}} = -6x^6 - 15x^4 + 9x^2 + 18x - 5$

3. $k = m - n = 0$

   $q_0 = 3^0 \cdot \tilde{\tilde{a}}_6 = -6$

   $r(x) = \tilde{\tilde{\tilde{a}}}(x) = -18x^6 - 45x^4 + 27x^2 + 54x - 135 + 18x^6 + 30x^4 - 24x^2 - 54x + 126 = -15x^4 + 3x^2 - 9$

   $\Rightarrow 27(x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5) = (3x^6 + 5x^4 - 4x^2 - 9x + 21)(9x^2 - 6) + (-15x^4 + 3x^2 - 9)$

Proposition 3.4. (4.9) *Suppose:* $a, b \in R[x]$ *are primitive and* R *factorial.*
$r \equiv pres(a, b)$ *implies* $\gcd(a, b) = \gcd(b, r) = \gcd(b, pp(r))$.

Proof:
We have $b_n^{m-n-1} a(x) = q(x)b(x) + r(x)$ by definition of r.
$c|a \wedge c|b \Rightarrow c|r \Rightarrow \gcd(a, b)|\gcd(b, r)$
$c|b \wedge c|r \Rightarrow c$ is primitive $\Rightarrow c|b_n^{m-n-1} a(x) \Rightarrow x|a \Rightarrow \gcd(b, r)|\gcd(a, b)$
The last assertion holds since $cont(\gcd(a, b)) = \gcd(cont(a), cont(b)) = \gcd(1, 1) = 1$ and
$cont(\gcd(b, r)) = \gcd(cont(b), cont(r)) = \gcd(1, cont(r)) = \gcd(1, 1) = 1$.

Algorithm (Euclidean algorithm)
$d = \gcd(cont(a), cont(b))$;
$a = pp(a)$; $b = pp(b)$;
while $(r = pres(a, b) \neq 0)\{$
    $a = b$;
    $b = pp(r)$;
}
return $g = db$;

Remark
For the above example we get $pp(r) = -5x^4 + x^2 - 3$.

Example
In the following example of the Euclidean Algorithm we will only write down the coefficients of the polynomials:

a	b
$1, 0, 1, 0, -3, -3, 8, 2, -5$	
$3, 0, 5, 0, -4, -9, 21$	$-15, 0, 3, 0, -9$
$5, 0, -1, 0, 3$	$-585, 1125, 2205$
$13, 25, -4$	$-233310, 307500$
$4663, -6150$	$143193869 = \gcd(a, b)$

Remark
Repeated $cont(r) = \gcd$ calculations can be avoided by scaling pseudoresiduals differently.
In this way one can achieve (as has been shown by Henrick) that with $d > \max\{deg(a(x)), deg(b(x))\}$ for $R = \mathbb{Z}[x]$ and $H = \max\left\{\sqrt{\sum_{j=0}^{n} b_j^2}, \sqrt{\sum_{j=0}^{n} b_j}\right\}$ the $\gcd(a, b)$ can be calculated with a number of integer operations of $O(d^4 \ln^2 H)$.

Lecture on 2008-06-23

Remark (Back to Factorization)

- Step 0: Eliminate repeated factors by division by $\gcd(a, a')$.
- Step 1: Factorize $a(x) \mod p$, i.e., in $\mathbb{F}_p = \mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$.

  Justification: $a(x) = b(x)c(x) \equiv b(x) | a(x)$ in $\mathbb{Z}[x]$
  $\Rightarrow a_p(x) = b_p(x)c_p(x) \equiv b_p(x) | a_p(x)$ in $\mathbb{Z}_p[x]$
  where a_p, b_p and c_p are obtained by taking the coefficients modulo p.

  Warning: The converse is not true, as for example $(x^2 + 2) \equiv (x + 1)(x + 2) \mod 3$, but $X^2 + 2$ is irreducible in $\mathbb{Z}[x]$. In Fact $(x^2 + 2) \equiv (x^2 - 1) \equiv (x + 1)(x - 1) \mod 3$.

- Hence, factors obtained in $\mathbb{Z}_p$ must be checked with respect to validity in $\mathbb{Z}[x]$ afterwards. (Step 2)

Observation
The factorization in $\mathbb{Z}_p[x]$ (more generally in $\mathbb{F}_q, q = p^n, n \in \mathbb{N}, p \in \mathbb{P}$) is simpler than the factorization in $\mathbb{C}$, because there are at most $p^{\frac{n}{2}+1}$ distinct polynomials $b(x)$ with $\deg(b(x)) \leq \frac{n}{2}$ which could possibly divide given $a(x) \in \mathbb{Z}_p[x]$ of degree $n = \deg(a(x))$.
$\Rightarrow$ Exhaustive search is finite (but costly and dimwitted).

Lemma 3.4. (Basic relations in $\mathbb{Z}_p$ (4.11))

1. $\forall \alpha \in \mathbb{Z}_p : \alpha^p = \alpha$, $\forall \alpha \in \mathbb{Z}_p, \alpha \neq 0 : \alpha^{p-1} = 1$
2. $x^p - x = x(x-1)(x-2) \cdot \ldots \cdot (x-(p-1)) \mod p$
3. $(a(x) + b(x))^p = a(x)^p + b(x)^p$ *(Raising to the p-th power is linear mapping on $\mathbb{Z}_p[x]$.)*
4. $(a(x))^p = a(x^p)$

Proof:

1. Standard/Exercise
2. Standard/Exercise
3. $(a(x)+b(x))^p = \sum_{j=0}^{p} \binom{p}{j} a(x)^j b(x)^{p-j} = a(x)^0 b(x)^{p-0} + a(x)^p b(x)^{p-p} = a(x)^p + b(x)^p$, since $\binom{p}{j}$ is divisible by p for $j = 1, \ldots, p-1$.
4. $(a(x))^p) = (\sum_{j=0}^{n} \alpha_j x^j)^p = \sum_{j=0}^{n} (\alpha_j x^j)^p = \sum_{j=0}^{n} \alpha_j^p x^{jp} = \sum_{j=0}^{n} \alpha_j (x^p)^j = a(x^p)$ □

Derivation of Bertehamp (Factorization by Linear Algebra)
By CRT (=Chinese Remaineder Theorem in $\mathbb{Z}_p[x]$):
Let be $a(x) = \prod_{j=1}^{m} a_j(x), a_j(x)$ irreducible (i.e. prime), monic (i.e. the leading coefficient is 1) and distinct.
$\Rightarrow b(x) \equiv \gamma_j \mod a_j(x), j = 1, \ldots, m$ has a solution $b \in \mathbb{Z}_p[x]$ which is unique up to multiples of $a(x) = \prod_{j=1}^{m} a_j(x)$.
In particular, there is exactly one representer $b(x)$ with $\deg(b(x)) < n = \deg(a(x))$.

Proposition 3.5. (4.12)

1. *The set of such $b(x) \in \mathbb{Z}_p[x]$ with $\deg(b(x)) < n$ forms an m-dimansional vector space $\mathcal{B}_a$, shich is isomorphic to $(\mathbb{Z}_p)^m$.*
2. $b(x) \in \mathcal{B}_a \equiv b(x)^p - b(x) = 0 \mod a(x), \deg(b) < n$
3. $b(x) \in \mathcal{B}_a \Rightarrow a(x) = \prod_{0 \leq \gamma < p} \gcd(b(x) - \gamma, a(x))$

Proof:

1. This follows the isomorphy with $(\mathbb{Z}_p)^m$ by mapping $b \mapsto (b \mod (a_j))_{j=1}^{m}$.
2. (a) "⇒"
   $b(x) \in \mathcal{B}_a \Rightarrow b(x) = \gamma_j \mod a_j \Rightarrow b(x)^p = \gamma_j^p \mod a_j$
   $\Rightarrow b(x)^p - b(x) = 0 \mod a_j, j = 1, \ldots, m$
   $\Rightarrow (b(x))^p - b(x) = 0 \mod a = \prod_{j=1}^{m} a_j$
   (b) "⇐"
   $a(x) = \prod_{j=1}^{m} a_j(x) | (b(x))^p - b(x) = b(x)(b(x)-1)(b(x)-2) \cdot \ldots \cdot (b(x)-p+1)$
   (The factors are pairwise relatively prime.)
   $\Rightarrow \forall j \exists r_j : a_j(x))(b(x) - r_j) \Rightarrow b(x) = \gamma_j \mod a_j(x)$
3. This follows from the last argument as γ is unique. □

Corollary 3.2. (4.13)
A square free monic polynomial $a(x) \in \mathbb{Z}_p[x]$ is irreducible, iff $\mathcal{B}_a \equiv \{0, 1, \ldots, \gamma, \ldots, p-1\} \equiv \mathbb{Z}_p$.

Proof:

1. "⇒"
   If $b(x) \in \mathcal{B}_a, \deg(b) > 0$, then $b(x) \neq \gamma$ for all $\gamma = 0, \ldots, p-1$ and thus $\forall \gamma = 0, \ldots, p-1 :$ $\gcd(b(x) - \gamma, a(x)) \neq a(x)$.

2. "⇐"

$b(x) \equiv r \Rightarrow b(x)^p = \gamma^p$ with resprect to any factor $a_j(x)$. □

Proposition 3.6. (4.14) $b(x) = \sum_{j=0}^{m} \beta_j x^j \in \mathcal{B}_a \Leftrightarrow (\beta_0, \ldots, \beta_{n-1}) \in \mathbb{Z}_:p^n$ *is a left eigenvektor associated with the eigenvalue* 1 *of the matrix* $Q = \begin{pmatrix} q_{0,0} & \cdots & q_{0,n-1} \\ \vdots & \ddots & \vdots \\ q_{n-1,0} & \cdots & q_{n-1,n-1} \end{pmatrix} \in \mathbb{Z}_p^{(n-1)\times(n-1)}$ *where the* $q_{i,j}$ *are uniquely defined by the equations* $x^{ip} = q_{i,0} - q_{i,1}x + \ldots + q_{i,n-1}x^{n-1} \mod a(x)$, $i = 0, \ldots, n-1$.

Proof:
$b(x) = \sum_{j=0}^{n} x^j \beta_j = \sum_{j=0}^{n-1} x^j \sum_{i=0}^{n-1} \beta_i q_{i,j} = \sum_{i=0}^{n-1} \beta_i \sum_{j=0}^{n-1} q_{i,j} x^j = \sum_{i=0}^{n-1} \beta_i x^{pi} = b(x^p) = (b(x))^p$
The second equality follows from the eigenvector property and vice versa. □

Remark
The sucessive calculation of q_{ij} as subset of $c_{k,j}$'s is given by:

$$\begin{aligned} x^k &= c_{k,0} + c_{k,1}x + \ldots + c_{k,n-1}x^{n-1} \mod a(x) \\ x^{k+1} &= c_{k,0}x + c_{k,1}x^2 + \ldots + c_{k,n-1}x^{n-1} \mod a(x) \\ &= -(\alpha_0 + \alpha_1 x + \ldots + \alpha_{n-1}x^{n-1}) + c_{k,0}x + c_{k,1}x^2 + \ldots + c_{k,n-1}x^{n-1} \mod a(x) \\ &= c_{k+1,0} + c_{k+1,1}x + \ldots + c_{k+1,n-1}x^{n-1} \end{aligned}$$

(We have $a(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n-1}x^{n-1} + x^n$ (since a is monic).
Hence, we get $x^n = -(\alpha_0 + \alpha_1 x + \ldots + \alpha_{n-1}x^{n-1})$.)

Example
Let be $a(x) = x^4 + x^3 + x - 1 \in \mathbb{Z}_3[x]$.
Then, we get:

j	0	1	2	3
α_j	-1	1	0	1
x^0	1	0	0	0
x^1	0	1	0	0
x^2	0	0	1	0
x^3	0	0	0	1
x^4	1	-1	0	-1
x^5	-1	$2 = -1$	-1	1
x^6	1	1	-1	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x^9	0	1	0	0

$$\Rightarrow Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow Q - I = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

Since left null vectors stay unchanged under elementary column operations, we get the matrix $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ which has rank 2. Hence there is a 2 dimensional nullspace.
One non trivial left nullvector is $(\beta_0, \beta_1, \beta_2, \beta_3) = (0, 1, 0, 1) \Leftrightarrow b(x) = x + x^3 \Rightarrow a(x)$ is reducible.

Lecture on 2008-06-30

Algorithm (4.15: Factorization)
Factorization in $\mathbb{Z}[x]$

1. Given squarefree $P(x) \in \mathbb{Z}_p[x]$ of degree n, we form the matrix $Q = (q_{i,j})_{i,j=0,...,n-1} \in \mathbb{Z}_p^{n\times n}$ by appropriate method.

2. Using alementary column operations and permutations, we reduce $Q - I$ to trapezoidal form $\begin{pmatrix} 0 & R \\ 0 & U \end{pmatrix}$ with $\det U \neq 0$ for $n \in \mathbb{Z}_p^{(n-m)\times(n-m)}$.

   $\Rightarrow rank(U - I) = n - m$

   $\Rightarrow$ The number of irreducible factors is exactly m.

3. If $m = 1$ stop as P is irreducible. Otherwise, consider the $m-1$ polynomials $b^j(x), j = 2, ..., m$ whose coefficients form the j-th row of the matrix.

   $B \equiv [I, -RU^{-1}] \in \mathbb{Z}_p^{m\times n} \Rightarrow B \cdot Q = B \Leftrightarrow B(Q - I) = 0$

   Then determine all non trivial factors (irreducible): $\gcd(b^j(x) - \gamma, P(x))$ for $j = 0, ..., p - 1$ and $j = 2, ...m$ until m distinct factors have been found.

Proposition 3.7. (Satz 4.5 in Pethö (4.16)) *The worst case complexity of above algorithm is* $O(\max(mpn^2, n^3 \log p)$ *operations in* $\mathbb{Z}_p$.

Remark (Average complexity)
The improvement due to Cantor-Zassenhaus yields expected complexity $O(\max(n^3, n^2 \log p))$ where $n = \deg(p)$ and m is the number of irreducible factors.

Step 2: Hensel lifting
Now, we consider the Hensel lifting from $\mod p$ to $\mod p^k$.
Recall that $\mathbb{Z}_{p^k}$ has zero divisors and thus no extension of quotients. Hence, we must leave the arithmetic in fields.

Bound on the number of coefficients
Is there a bound on the size and thus the number of coefficients of a factor $G(x) \in \mathbb{Z}[x]$ of given $P(x) \in \mathbb{Z}[x]$?
Answer: Yes, see the following.

Lemma 3.5. (Mignotte Bound (4.17)) *With* $H(P)$ *the maximal modulus of any coefficient of* $P(x) \in \mathbb{Z}[x]$ *the maximal coefficient of any divisor* G *of* P *with* $\deg(G) \leq \frac{1}{2}\deg(P) = \frac{n}{2}$ *is bounded by* $B = 2^{\frac{n}{2}}\sqrt{n+1}H(P)$. $H(P)$ *is also called the* height *of* P.
Hence we can select for any prime p *a power* k *such that* $c \in \left[-\frac{p^k-1}{2}, \frac{p^k-1}{2}\right] \Rightarrow |c| \leq B$.

Proof:
No proof. □

Corollary 3.3. (4.18) $P(x), G(x), H(x) \in \mathbb{Z}[x], P(x) = G(x)H(x)$ *and* $p^k, \deg(G) \leq \frac{1}{2}\deg P$ *implies* $G(x)|P(x)$ *in* $\mathbb{Z}[x]$.

$\mod p^k$ **factoriziation**
How to obtain $\mod p^k$ factorization?
Answer: Hensel lifting.
Setting: We have $f(G, H) = G \cdot H - P \mod I$ where $G, H, P \in \mathbb{Z}[x]$ and $I = (p) = p\mathbb{Z}[x]$.
Jacobian of f with respect to G and H is $f'(G, H) = \nabla f(G, H) = (\frac{\partial f}{\partial G}, \frac{\partial f}{\partial H}) = (H, G) \in \mathbb{Z}[x]^{1\times 2}$.
The right Inverse $W = (G^*, H^*)^T \in \mathbb{Z}[x]^{2\times 1}$ should satisfy $f'(x) \cdot W(x) = H \cdot G^* + G \cdot H^* = 1 \mod p$.

Assuming that H and G are coprime, which holds in particular if G is irreducible, $\gcd(H,G) \in \mathbb{Z}_p[x]$ equals 1 and the G^*, H^* can be computet by the extended Euclidean Algorithm.
The by proposition 2.4 (Linear Hensel Lifting - 3.9) obtained

$G_{k+1} = G_k - (G_k \cdot H_k - P)G^* = G_k - f(G_k, H_k)G^*$

$H_{k+1} = H_k - (G_k \cdot H_k - P)H^* = H_k - f(G_k, H_k)H^*$

starting from $G_1 = G, H_1 = H$ yields $G_k = G_1 \mod p$, $H_k = H_1 \mod p$ and $f(G_k, H_k) = 0 \mod p^k$.
This is equivalent to $P(x) = G_k(x)H_k(x) \mod p^k$.

Example
Let be $G_1(x) = G(x) = x^2 + 1$, $H_1(x) = H(x) = x^5 + x^2 + 1$ and $P(x) = x^7 + x^5 + x^4 + 1 \in \mathbb{Z}[x]$.
$f(G,H) = (x^2+1)(x^5+x^2+1) - x^7 - x^5 - x^4 - 1 = x^7 + x^4 + x^2 + x^5 + x^2 + 1 - x^7 - x^5 - x^4 - 1 = 2x^2$
$\Rightarrow f(G,H) = 0 \mod p$ with $p = 2$
The Extended Euclidean Algorithm in $\mathbb{Z}_2$ yields $(x^5 + x^2 + 1)x + (x^2 + 1)(x^4 + x^2 + x + 1) = 1 \mod 2$.
$G_2(x) = G_1(x) - f(G_1, H_1)G^*(x) = x^2 + 1 - 2x^2 \cdot x = x^2 + 1 - 2x^3$
$H_2(x) = H_1(x) - f(G_1, H_1)H^*(x) = x^5 + x^2 + 1 - 2x^2(x^4 + x^2 + x + 1)$
$\quad = x^5 + x^2 + 1 - 2x^6 - 2x^4 - 2x^3 - 2x^2$

$$\begin{aligned} G_2(x) \cdot H_2(x) \mod 4 &= (x^2 + 1 - 2x^2 \cdot x)(x^5 + x^2 + 1 - 2x^2(x^4 + x^2 + x + 1)) \mod 4 \\ &= (x^2 + 1)(x^5 + x^2 + 1) - 2x^2(x(x^5 + x^2 + 1) + (x^2 + 1)(x^4 + x^2 + x + 1)) \\ &\quad + \underline{4}x^2(\ldots) \mod 4 \\ &= P(x) + 2x^2 - 2x^2(x(x^5 + x^2 + 1) + (x^2 + 1)(x^4 + x^2 + x + 1)) \mod 4 \\ &= P(x) - 2x^2(x^6 + x^3 + x + x^6 + x^4 + x^4 + x^2 + x^3 + x + x^2 + 1) \mod 4 \\ &= P(x) - 4x^2(x^6 + x^4 + x^3 + x^2 + x) \mod 4 \end{aligned}$$

$\Rightarrow G_2(x) \cdot H_2(x) = P(x) \mod 4$
$\vdots$

Algorithm (Berlecamp factorization (4.19))
The Berlecamp factorization in $\mathbb{Z}[x]$ is:

0. Reduce $F(x)$ to square free form
1. Factorize P in $\mathbb{Z}_p[x]$ into $P(x) = P_1(x) = \cdot \ldots \cdot P_m(x)$.
2. If $m = 1$ then $P(x)$ is irreducible, hence stop.
3. Otherwise choose partition $I \cup J = \{1, \ldots, m\}$ with $I \cap J = \emptyset$.
   Put $G_1 = \prod_{j \in I} P_j$ and $H_1 = \prod_{j \in J} P_j$.
   $\Rightarrow P = G_1 H_1 \mod p, \gcd(G_1, H_1) = 1$
4. Apply Hensel to obtain G_k, H_k with $P = G_k \cdot H_k \mod p^k$ and $p^k > 2B$, with B as above.
5. (a) If $G_k | P$ in $\mathbb{Z}[x]$ apply the procedure recursively to G_k and P/G_k.
   (b) If $G_k \not| P$ and some partitions in 3. have not been checked try the next one.
   (c) Otherwise stop factorization.

Remark
Only one p is necessary for correctness though efficiency may depend on it. The smaller p the more factors m can be expected and thus the more partitions need to be considered.

Lecture on 2008-07-07

Chapter 4

Gröbner Basis and Buchberger Algorithm

4.1

TODO
Find a name for this subsection

Aim
Solve Systems of Polynomial Equations (e.g. Robotics and Wavelet Construction):
$P_i(x) = 0$, $i = 1, ..., s$ with $P_i \in F[x_1, ..., x_n]$

Definition (5.1)
Consider an ideal $I = \langle P_1, .., P_s \rangle \equiv \{\sum_{i=1}^{s} Q_i(x)P_i(x) | Q_i \in F[x_1, ..., x_n]\}$.
We define the *variety* $V(I) = \{x \in F^n | \forall i = 1, ..., s : P_i(x) = 0\} = \{x \in F^n | \forall P \in I : P(x) = 0\}$.

Hilbert's Result

1. Nullstellensatz:

   If F is agebraically complete, e.g. $\mathbb{C}$ or $\mathbb{Z}_p$, then $V(I) = \emptyset \Leftrightarrow I \ni 1 = \sum_{i=1}^{s} Q_i P_i$.

2. Basissatz: For any field F, e.g. $F = \mathbb{Q}$ or $F = \mathbb{R}$, any ideal is finitely generated. (Equivalently: Any ascending chain of ideals terminates, i.e. the ring is Noethersch.)

Questions to be answered

1. $V(I) \neq \emptyset$?
2. $V(I) =$? (characterize solutions)
3. $I \ni P$? (Membership)
4. $I = F[x_1, ..., x_n]$ (Triviality)

Approach
Generate structured basis (=generateing system) by successive eliminations of "higher order" terms. Elemination means linear combinations in this ideal.

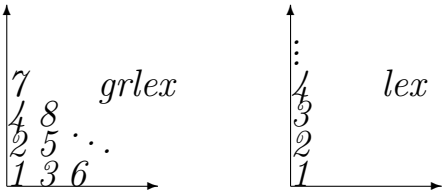
Definition (5.2)

1. $x^\alpha = x_1^{\alpha_1} x_2 \alpha_2 ... x_n^{\alpha_n}$ for $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$
2. $x^\alpha \prec x^\beta \Leftrightarrow \alpha \prec \beta$ if
   - the first nonzero component of $\beta - \alpha \in \mathbb{Z}^n$ is positive — lex: lexicographic ordering
   - $\sum_{i=1}^s \alpha_i < \sum_{i=1}^n \beta_i$ or $(\sum_{i=1} s\alpha_i = \sum_{i=1}^s \beta_i$ and $\alpha \prec_{lex} \beta)$ — grlex: graded lexicographic ordering

Remark
A third ordering grevlex is obtained like grlex but with numbering reversed (grevlex: graded reverse lexicographic ordering).

Lemma 4.1. (5.3) *The above orderings satisfy*

1. *The (usual, partial) componentwise ordering on $\mathbb{Z}^n$ ($\alpha \leq \beta$) implies $\alpha \preceq \beta$. (monotonicity)*
2. *$\alpha \preceq \beta, \gamma \in \mathbb{N}^n \Rightarrow \alpha + \gamma \preceq \beta + \gamma$ (additive invariance)*
3. *$\alpha \prec \beta \vee \alpha = \beta \vee \alpha \succ \beta$ for any (α, β) (total ordering)*
4. *Any $S \subseteq \mathbb{N}^n$ has a unique minimal element.*



Definition (5.4)
Let be $P \in F[x_1, ..., x_n]$.

1. $P(x) = \sum_{\alpha \in \mathbb{N}^n} c_\alpha(P) x^\alpha$ with only finitely many $c_\alpha(P) \neq 0$
2. $le(P) = \max_\prec \{\alpha | c_\alpha(P) \neq 0\}$ (leading exponent)
3. $lc(P) = c_{le}(P) \in \mathbb{N}^n$ (leading coefficient)
4. $lt(P) = c_{le}(P) x^{le(P)}$ (leading term)

Lemma 4.2. (5.5) *Let be $P, Q \in F[x_1, ..., x_n] \setminus \{0\}$.*

1. $le(P \cdot Q) = le(P) + le(Q)$
   $(\Leftrightarrow lt(P \cdot Q) = lt(P) \cdot lt(Q))$
2. $le(P + Q) \preceq \max_\prec \{le(P), le(Q)\}$
   where equality holds if $le(P) \neq le(Q)$.
3. $le(P_1) \leq le(P) \Leftrightarrow lt(P_1) | lt(P)$
   $\Rightarrow le\left(P - \frac{lt(P)}{lt(P_1)} P_1\right) \prec le(P)$

Proof:

1. and 2. Can be easily checked.
3. Clearly, the old leading term of P is eliminiated by subtraction. The remaining terms of P and nonleading terms of P_1 have exponents $\prec le(P)$ and $\prec le(P_1) + le(P) - le(P_1) = le(P)$.
   $\Rightarrow$ Assertion follows from 2. □

Example

	$xy+1$	$y+1$
xy^2+1	y	
$-xy^2-y$		
$-y+1$		-1
$-y-1$		
2		

$\Rightarrow (xy^2+1) = y(\underbrace{xy+1}_{=P_1}) - (\underbrace{y+1}_{=P}) + 2$

	$y+1$	$xy+1$
xy^2+1	xy	
$-xy^2-xy$		
$-xy+1$	$-x$	
$xy+x$		
$1+x$		

$\Rightarrow (xy^2+1) = (xy-x)(y+1) + x + 1 + 0 \cdot (xy+1)$

Observation
The representation $P = Q_1P_1 + Q_2P_2 + R$ where no term of R is divisible by $lt(P_1)$ or $lt(P_2)$ is not unique.

Algorithm (Multidivision with Remainder)
$F[x]^n$ multidivision $(P, P_1, ..., P_s)$ {
  $R = 0$;
  $\tilde{P} = P$;
  for $(i = 1, ..., s)$
    $Q_i = 0$;
  while $(\tilde{P} \neq 0)$
    if $(\exists i \in \{1, ..., s\} : lt(P_i) | lt(\tilde{P}))$ {
      $Q_i = Q_i + \frac{lt(\tilde{P})}{lt(P_i)}$;
      $\tilde{P} = \tilde{P} - \frac{lt(\tilde{P})}{lt(P_i)} P_i$;
    } else {
      $R = R + lt(\tilde{P})$;
      $\tilde{P} = \tilde{P} - lt(\tilde{P})$;
    }
  return $Q_1, ..., Q_s, R$;
}

Proposition 4.1. (5.6) *The above algorithm yields*

- *throughout* $P = \tilde{P} \sum_{i=1}^{n} Q_i P_i + R$ *with* $le(Q_i) \preceq le(P)$
- *at the and* $P = \sum_{i=1}^{s} Q_i P_i + R$ *with* $c_\alpha(R) \neq 0 \Rightarrow \alpha \not\geq le(P_i)$ *for* $i = 1, ..., s$ *(*$\Leftrightarrow lt(P_i) \not| x^\alpha$ *for* $i = 1, ..., s$*).*

Proof:
By Exercise. □

Lecture on 2008-07-14

Remark
Division by $\{P_1, ..., P_s\}$ yields $P(x) = R(x) + \sum_{j=1}^{n} Q_j(x)P_j(x)$ with $le(P_iQ_i) \preceq le(P)$ and $c_\alpha(R) \neq 0 \Rightarrow \forall i : \alpha \not\geq le(P_i)$.

Lemma 4.3. (5.7) *For any ideal $I \subseteq F[x_1, ..., x_n]$ there exist exponents $\alpha_i \in \mathbb{N}^n$ for $i = 1, ..., s$ such that:*
$le(I) = \{le(P)|P \in I\} = \bigcup_{i=1}^{s}\{\alpha_i + \mathbb{N}^n\}$

Proof:
Let be $\alpha = le(P)$ with $P \in I$ and $\beta \in \mathbb{N}^n$.
$\Rightarrow le(x^\beta P(x)) = \alpha + \eta \in le(I)$
$\Rightarrow le(I) = \bigcup_{\alpha \in le(I)}\{\alpha + \mathbb{N}^n\} = \bigcup_{\alpha \in A}\{\alpha + \mathbb{N}^n\}$
where A is the set of minimal elements with respect to the partial ordering $\leq$. Its finiteness can be checked by induction on the dimension n. □

Proposition 4.2. (5.8) *For $A = \{\alpha_i\}_{i=1}^{s}$ as obtained in lemma 4.3 any selection $\{P_i\}_{i=1}^{s} \subseteq I$ with $le(P_i) = \alpha_i$ for $i = 1, ..., s$ generates I. It is then called a* Gröbner basis.

Proof:
By the lemma there exists for any $P \in I$ a representation $P = R + \sum_{i=1}^{s} Q_iP_i$ (with $R \in I$).
$\Rightarrow R = 0 \vee le(R) \geq le(P_i) = \alpha_i$ for some i
$\Rightarrow R = 0$ (by construction of R, i.e. by division)
$\Rightarrow \{P_i\}_{i=1}^{s}$ generates I. □

Corollary 4.1. (5.9) *For any $P \in F[x_1, ..., x_n]$ the residual R with respect to a given Gröbner basis $G = \{P_1, ..., P_s\}$ is unique and thus denoted by $R = P \operatorname{rem} G$.*

Proof:
$P = R + \sum_{i=1}^{s} Q_iP_i = R' \sum_{i=1}^{s} Q'_iP_i$
$\Rightarrow \Delta R = R' = R = \sum_{i=1}^{s}(Q_i - Q'_i)P_i \in I$
$\Rightarrow \Delta R = 0 \vee le(\Delta R) \preceq \max_\prec\{le(R'), le(R)\} \not\geq le(P_i)$ for all $i = 1, ..., s$
$\Rightarrow \Delta R = 0$ (otherwise the leading exponent of ΔR would not be contained in $le(I)$) □

Consequence
The Membersihp Problem $P \in I$ can be answered by computing $R = P \operatorname{rem} G$: $P \in I \Leftrightarrow R = 0$.

Example
$P = x^3 - 2xy, Q = x^2y - 2y^2 + x$
The grlex ordering yields $lt(P) = x^3 \succ x^2y = lt(Q)$.
$I = \{RP + SQ|R, S \in F[x_1, ..., x_n]\}$
Is $\{P, Q\}$ already a Gröbner Basis? If not, how can we construct one?
$S(P, Q) = x^{\max\{0, le(Q) - le(P)\}} \frac{P(x)}{lc(P)} - x^{\max\{0, \leq(P) - \leq(Q)\}} \frac{Q(x)}{lc(Q)}$, where max is used componentwise
$le(P) = (3, 0) \succ (2, 1) = le(Q)$
$\Rightarrow \max\{le(P), le(Q)\} = (3, 1)$
$R = x^0y^1(x^3 - 2xy) - x^1y^0(x^2y - 2y^2 + 1) = yx^3 - 2xy^2 - x^3y + 2xy^2 - x^2 = -x^2$
$\Rightarrow le(R) = (2, 0) \leq (3, 0) = le(P)$
$U = P + xR = (x^3 - 2xy) - x^3 = -2xy \Rightarrow le(U) = (1, 1)$
$V = Q + yR = -2y^2 + x \Rightarrow le(V) = (0, 2)$
$\Rightarrow$ Unless x or y or 1 is contained in I, we find that V, U, R form a Gröbner basis.

Definition (S-polynomial (5.10))
The polynomial $S(P, Q) = x^{\max\{0, le(Q) - le(P)\}} \frac{P(x)}{lc(P)} - x^{\max\{0, le(P) - le(Q)\}} \frac{Q(s)}{lc(Q)}$ is called *S-polynomial*.

Lemma 4.4. *The polynomial R from the definition above satisfies $le(R) \prec \max\{le(P), le(Q)\}$.*

Proof:
The powers have been chosen such that the highest terms in multiples of P and Q cancel. □

Lemma 4.5. (5.11) *For $P_1, ..., P_s \in I$ with $lc(P_i) = 1$ for $i = 1, ..., s$ holds:*
If $Q = \sum_{i=1}^{s} c_i x^{\max_{j=1,...,s}\{le(P_j)-le(P_i)\}} P_i$ has $le(Q) \prec \max_{j=1,...,s}\{le(P_j)\}$, then there exist coefficients $\tilde{c}_i$ for $i = 1, ..., s-1$ such that $Q = \sum_{i=1}^{s-1} \tilde{c}_i S(P_i, P_{i+1}) x^{\max\{P_j - \max_{j=1,...,s}\{P_i, P_{i+1}\}\}}$.

Proof:
By induction on $s = 2, 3, ...$:
For $s = 2$ the only combination of P_1 and P_2 that yields a leading exponent $\prec \max\{le(P_1), le(P_2)\}$ is the S-polynomial $S(P_1, P_2)$ multiplied by a constant $c_1 = \tilde{c}_1$.
Induction step:
Take the first term
$c_1 x^{\max_{j=1,...,s}\{le(P_j)-\max(le(P_1,P_2))\}} x^{\max\{0,le(P_2)-le(P_1)\}} P_1 = c_1 P_1 x^{\max_{j=1,...,s}\{le(P_j)-le(P_1)\}}$ (read opposite order).
$\Leftrightarrow \max\{le(P_j) - \max_{j=1,...,s}\{le(P_1), le(P_2)\}\} + \max\{0, le(P_2) - le(P_1)\} = \max_{j=1,...,s}\{le(P_j) - le(P_1)\}$
The first term is equal to
$c_1 x^{\max_{j=1,...,s}\{le(P_j)-\max(le(P_1),le(P_2))\}} \left(S(P_1, P_s) + x^{\max\{0,le(P_1)-le(P_2)\}} P_2\right)$
$= \tilde{c}_1 x^{\max_{j=1,...,s}\{le(P_j)-\max(le(P_1),le(P_2))\}} S(P_1, P_s)$ as claimed with $\tilde{c}_1 = c_1$
$+ c_1 P_2 x^{\max\{0,le(P_1)-le(P_2)\}}$(can be added to second term in original expansion with coefficient $c_1 + c_2$)
The Remaining sum of $s-1$ terms has by induction hypothesis a decomposition into a sum of $S(P_i, P_{i+1})$, with $i = 1, ..., s-1$. □

Corollary 4.2. (5.12) *$G = \{P_1, ..., P_s\}$ is Gröbner basis of the ideal generated by itself if and only if $S(P_i, P_j) \operatorname{rem} G = 0$ for any $0 < i < j \leq s$.*

Algorithm (Simplified Buchberger Algorithm)
GröbnerBasis buchberger($G = \{P_1, ..., P_s\}$) {
  while (true) {
    $S = \emptyset$;
    for ($\{P_i, P_j\} \subset G$) {
      $R = S(P_i, P_j) \operatorname{rem} G$;
      if ($R \neq 0$)
        $S = S \cup \{R\}$;
    }
    if ($S = \emptyset$)
      return G;
    else
      $G = G \cup S$;
  }
}

Lemma 4.6. (5.13) *For any ideal I there is a unique "reduced" Gröbner basis $G = \{P_1, ..., P_s\}$ with the following property:*
$\forall i = 1, ..., n : c_\alpha(P_i) \neq 0 \Rightarrow \alpha \not\geq le(P_i)$

Proof:
No Proof. □

Remark (Good News)
If $V(P_1, ..., P_s) \equiv \{x \in \mathbb{R}^n | \forall P \in I : P(x) = 0\}$ consists of finitely many isolated solution points (well determined system in Numerical Analysis Sense) then the Gröbner basis with respect to lexicographic ordering can be ordered in "triangular" fashion such that $lt(P_i) = x_i^{\nu_i}$ for some $\nu_i \geq 1$ and $i = 1, ..., n$. Hence the solution components can be computed one by one as in Backsubstitution in Linear Algebra.

Example
On the above example, we get $x = 0$ by $-x^2 = 0$ and hence $-2y^2 + x = 0 \Rightarrow y = 0$.

Remark (Bad News)
Computing a Gröbner basis by Buchberger is EXPSPACE-complete, i.e. the space requirement grows in the worst case (and typically) like 2^{2^n} where n measures the input size.

Index