Part I: Index Calculus in Finite Fields
–
Part II:Hyperelliptic Curves

Tanja Lange

Department of Mathematics

Technical University of Denmark

t.lange@mat.dtu.dk

30.08.2006

Overview

- Index calculus attacks in prime fields.
- Index calculus attacks in binary fields.
- More efficient index calculus attacks in binary fields.
- Definition of hyperelliptic curves.
- Divisor class group.
- Elliptic curves (reminder).
- Pictorial description of group law.
- Mumford representation.
- Cantor's algorithm.
- Index calculus attacks on HECC.

Index calculus attacks in prime fields

- Index calculus is a method to compute discrete logarithms, also called indices.
- p prime, elements of $\mathbb{F}_p$ represented by numbers in $\{0, 1, \ldots, p-1\}$; g generator of multiplicative group.
- If $h \in \mathbb{F}_p$ factors as $h = h_1 \cdot h_2 \cdots h_n$ then

$$h = g^{a_1} \cdot g^{a_2} \cdots g^{a_n} = g^{a_1 + a_2 + \cdots + a_n}$$

  with $h_i = g^{a_i}$.
- Knowledge of the a_i, i.e. the discrete logarithms of h_i to base g gives knowledge of the discrete logarithm of h to base g.
- If h factors appropriately ...

Index calculus attacks in prime fields

- If h factors appropriately?!
- This point can be taken care of by having $h \cdot h'$ and h' factor over the h_i. This is possible since in $h \cdot h'$ one reduces modulo p.
- So far: instead of finding one DL we have to find many DLs and they have to fit to h and we have to find a suitable h' and factor numbers.
- Will show now that this works indeed.
- Note that we are working in two different settings – the integers modulo p and the integers themselves. The factorization takes place over the integers while the left hand side is always reduced modulo p.

Terminology

- Select some set $\{g_1, g_2, \ldots, g_m\}$ so that a number $\bar{h} < p$ is likely to factor into powers of the g_i.
- The set of g_i is called the factor base.
- An equation of the form $\bar{h} = g_1^{n_1} \cdot g_2^{n_2} \cdots g_m^{n_m}$ with $n_i \in \mathbb{Z}$ is called a relation.
- Choose the factor base to consist of the small primes e.g. $g_1 = 2, g_2 = 3, g_3 = 5, g_4 = 7, \ldots$ to get a large chance that a random $\bar{h}$ factors over this factor base.
- Generate many relations with known DL of $\tilde{h}_j = g^{k_j}$

$$\tilde{h}_j = g_1^{n_{j1}} \cdot g_2^{n_{j2}} \cdots g_m^{n_{jm}} = g^{k_j}$$

  by doing a random walk modulo p starting from some g^{k_1} and keeping track of the k_j.

Matrix of relations

For each relation

$$\tilde{h}_j = g_1^{n_{j1}} \cdot g_2^{n_{j2}} \cdots g_m^{n_{jm}} = g^{k_j}$$

enter the row

$$(n_{j1} n_{j2} \ldots n_{jm} | k_j)$$

into a matrix

$$M = \left(\begin{array}{cccccc|c} n_{11} & n_{12} & \ldots & n_{1i} & \ldots & n_{m1} & k_1 \\ n_{21} & n_{22} & \ldots & n_{2i} & \ldots & n_{m2} & k_2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ n_{l1} & n_{l2} & \ldots & n_{li} & \ldots & n_{lm} & k_l \end{array} \right)$$

The i-th column corresponds to the unknown a_i so that $g_i = g^{a_i}$.

Computing DLPs

- Use linear algebra to solve matrix M for the unknowns a_i.
- This step needs to be done only once per field and generator, it does not depend on the target DLP $h = g^a$.
- A single relation involving h factoring over the factorbase gives the DLP in one step.
- Other approach is to include h in the factor base. Then matrix cannot be reused. Advantage in proving running times (but if an attack works then it just works . . .)
- Running time with much more clever way of finding relations is

$$O(\exp(c \log p^{1/3} \log(\log p)^{2/3}))$$

for some c. This is subexponential in $\log p$!

Translation to $\mathbb{F}_{2^n}$

- Elements of $\mathbb{F}_{2^n}$ are represented as

$$\mathbb{F}_{2^n} = \{\sum_{i=0}^{n-1} c_i x^i | c_i \in \mathbb{F}_2, 0 \le i < n\},$$

  i.e. polynomials of degree less than n modulo an irreducible polynomial $f(x) \in \mathbb{F}_2[x]$.
- Factoring into powers of small primes is replaced by factoring into irreducible polynomials of small degree.
- Same approach works, same problem of balancing size of factorbase (and thus complexity of the matrix step) and the likelyhood of splitting completely over the factorbase.

Arbitrary finite fields $\mathbb{F}_q = \mathbb{F}_{p^n}$

- Same approach (or better mixture of both) works for general finite finite fields.
- Easy to get running time

$$O(\exp(c' \log q^{1/2} \log(\log q)^{1/2}))$$

  for some c'.
- For all finite fields the DLP has subexponential complexity.
- Result at Crypto 2006:
  For all p, n one can obtain running time

$$O(\exp(c'' \log q^{1/3} \log(\log q)^{2/3}))$$

  for some constant c''.

Very small factorbase

- Restrict factorbase to linear polynomials. So there are p elements in the factor base.
- Number of polynomials of degree $< n$ splitting into linear factors is approximately $\sim \frac{1}{n!}p^n$.
- Total number of polynomials of degree $< n$ is $q = p^n$.
- Probability of splitting in reduced factor base is $\sim \frac{1}{n!}$.
- Need $O(n!p)$ tries to find p relations, $O(p^2)$ for sparse matrix.
- For n fixed, p growing the running time $O(n!p + p^2)$ translates to
$$O(p^2)$$
- Very fast – beware of constants!

Even faster

- Tiny factorbase:
  Restrict factorbase to subset of linear polynomials to have p^r elements in the factor base, for some $r \in (0, 1)$.
  Gives
  $$\tilde{O}(p^{2-\frac{2}{n+1}}).$$
- Use large prime variation, i.e. have a further set of elements for which relations are accepted. Then for each of them linear algebra is used to cancel them out (slightly more entries per row).
- Use double large prime variation, i.e. use single large prime variation as above and allow some further elements for which at least two relations must exist – so they cancel out immediately.


And now for
something
completely
different...

Hyperelliptic curves

Affine equation of hyperelliptic curve of genus g (with $\mathbb{F}_q$-rational Weierstraß-point at infinity)

$$C : y^2 + h(x)y = f(x)$$

$h(x), f(x) \in \mathbb{F}_q[x]$, f monic, $\deg f = 2g + 1, \deg h \leq g$
non singular, i. e. in no point $(a, b) \in C/\overline{\mathbb{F}_q}$ over the algebraic closure $\overline{\mathbb{F}_q}$ we have both partial derivatives vanishing, i.e.

$2b + h(a) = 0$ and $h'(a)b - f'(a) = 0$.

Otherwise (a, b) is called a singular point.

Try to mimic approach in elliptic curves, but first some examples.

Examples

Concerning the arithmetic properties one can consider elliptic curves as hyperelliptic curves, i. e.

$$y^2 + (a_1x + a_3)y = x^3 + a_2x^2 + a_4x + a_6$$

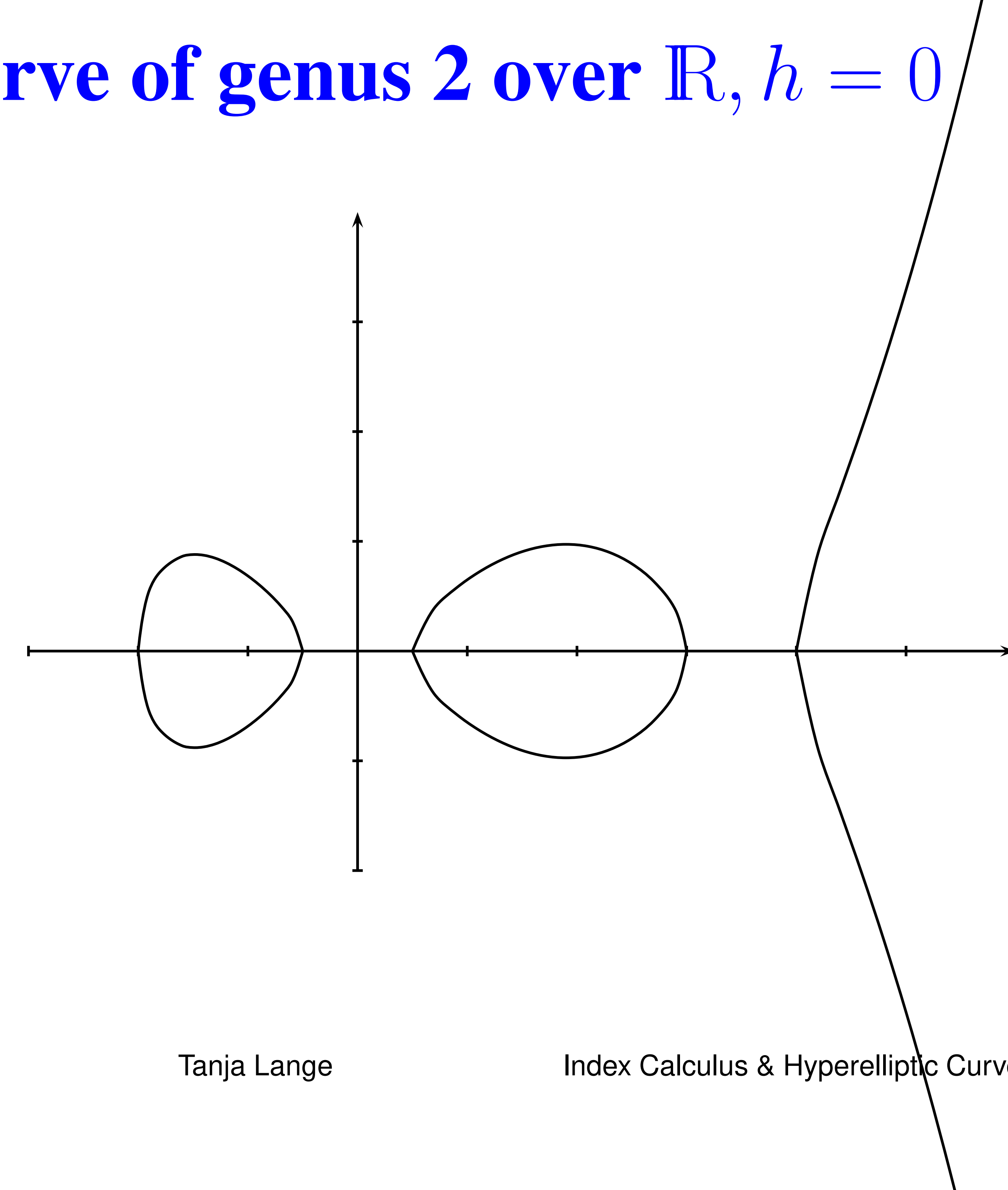
is considered as curve of genus one.

Curve of genus 2 over field of odd characteristic

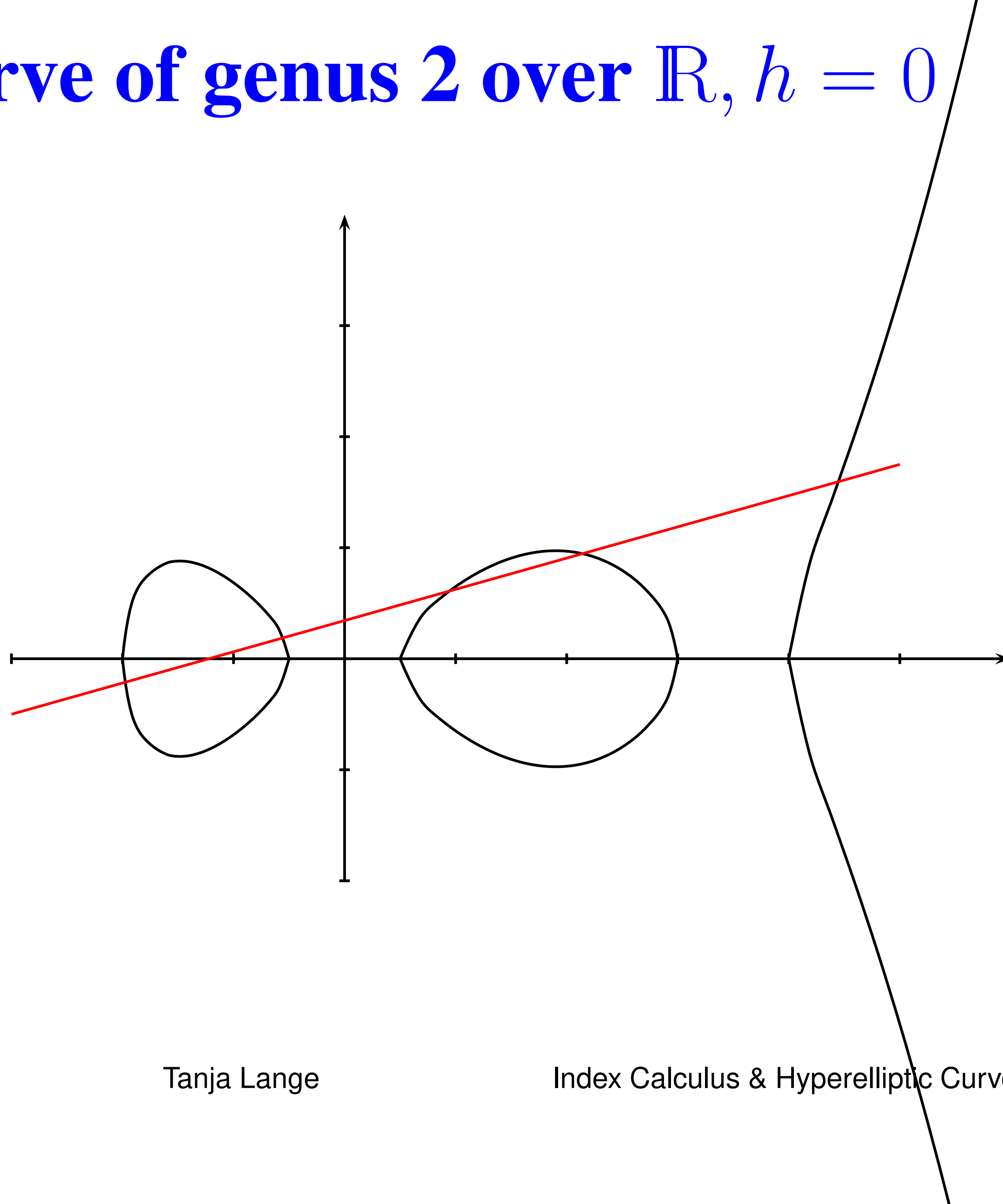
$$y^2 = x^5 + f_3x^3 + f_2x^2 + f_1x + f_0,$$

provided $f(x)$ has no multiple roots.

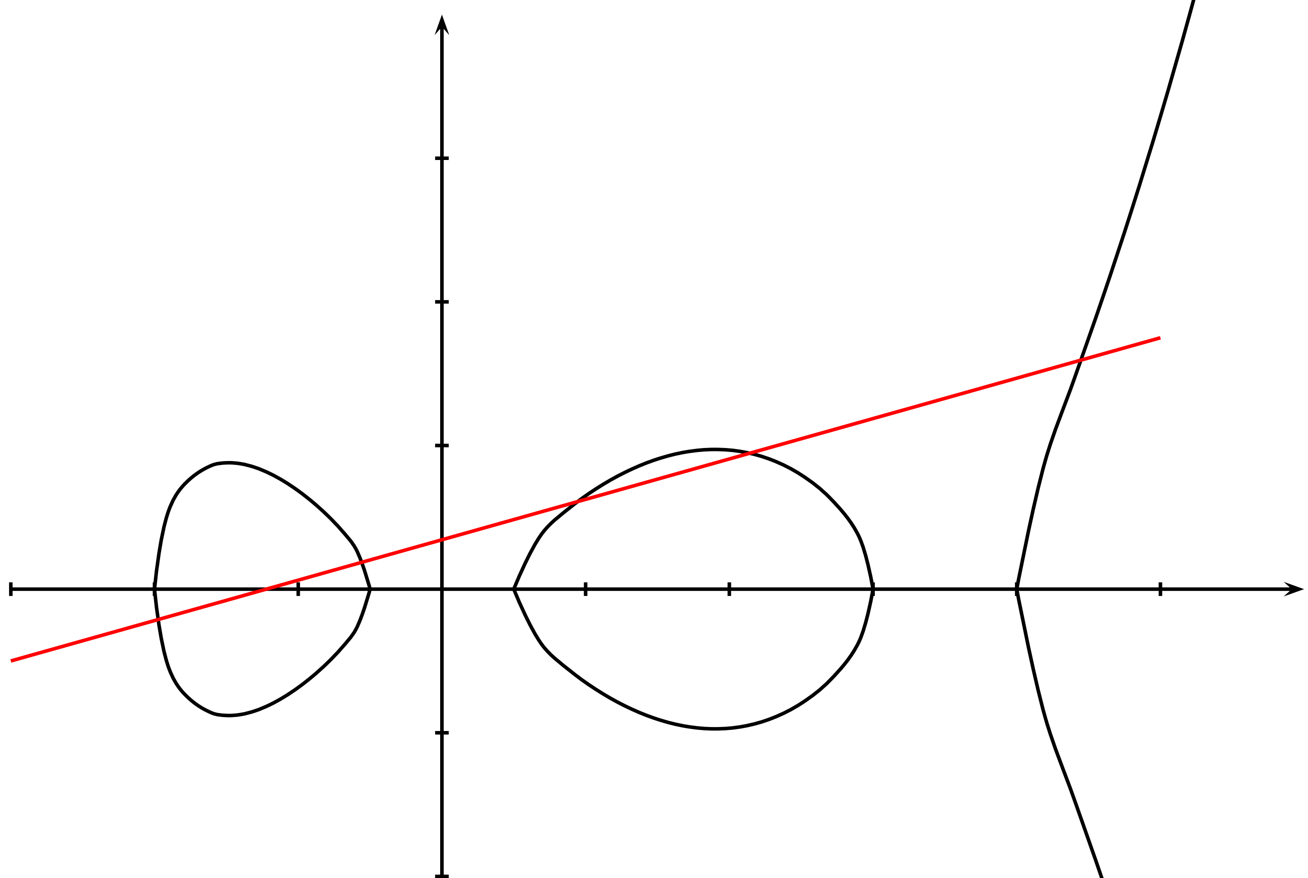
Curve of genus 2 over $\mathbb{R}, h = 0$



Curve of genus 2 over $\mathbb{R}, h = 0$



Curve of genus 2 over $\mathbb{R}, h = 0$



points do **not** form a group!

Group of Divisors

- Construct group from points on curve. Free abelian groups are in particular groups, and so associativity etc. follow immediately.
- Construction uses Divisors, i. e. finite sums of points (elements of the free abelian group),

$$\sum_{P \in C(\overline{\mathbb{F}_q})} n_P P,\ n_P \in \mathbb{Z}$$

  with $n_P = 0$ for almost all P.
- Addition works component-wise:

$$(P_1 + 2P_2 - P_3) + (P_1 + P_2 + P_4) = 2P_1 + 3P_2 - P_3 + P_4.$$

Divisors

- Effective divisors are divisors

$$D = \sum_{P \in C(\overline{\mathbb{F}_q})} n_P P,\ n_P \in \mathbb{Z}$$

  for which each $n_P \geq 0$.
- The degree of a divisor is

$$\deg(D) = \sum_{P \in C(\overline{\mathbb{F}_q})} n_P.$$

  $\deg(P_1 + 2P_2 - P_3) = 1 + 2 - 1 = 2, \deg(P_1 + P_2 + P_4) = 3,$
  $\deg(2P_1 + 3P_2 - P_3 + P_4) = 5$.
- Divisors of degree zero form a group Div_C^0 with component-wise addition.

Principal divisors

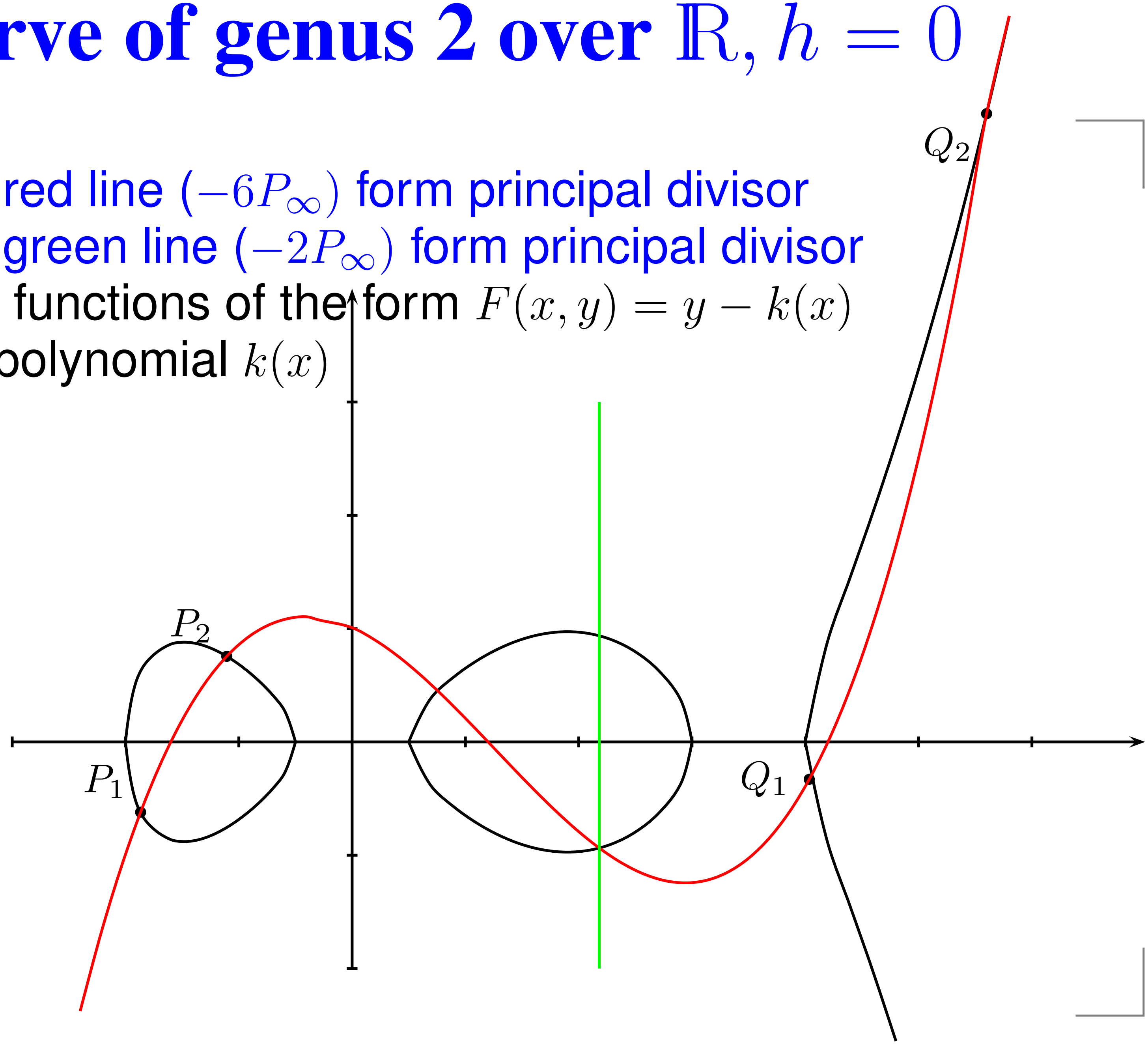
- For any function $F(x,y)$ the graph $F(x,y)=0$ intersects curve in some points of $C(\overline{\mathbb{F}_q})$.
- Let v_P be normalized valuation $P \in C(\overline{\mathbb{F}_q})$, thus $v_P(F) = n \geq 0$ iff F has intersection of multiplicity n with curve at P (simple intersection has $n = 1$ while tangent has multiplicity $n \geq 2$).
- Negative values for multiplicity of poles.
- Associate divisor

$$\mathrm{div}(F) = \sum_{P \in C(\overline{\mathbb{F}_q})} v_P(F) P$$

  to function $F \in \mathbb{F}_q(C)$.
- Such divisors are called principal divisors Princ_C. One can show that they have degree zero.

Curve of genus 2 over $\mathbb{R}, h = 0$

points on red line $(-6P_\infty)$ form principal divisor
points on green line $(-2P_\infty)$ form principal divisor
Here only functions of the form $F(x, y) = y - k(x)$
for some polynomial $k(x)$



Divisor class group

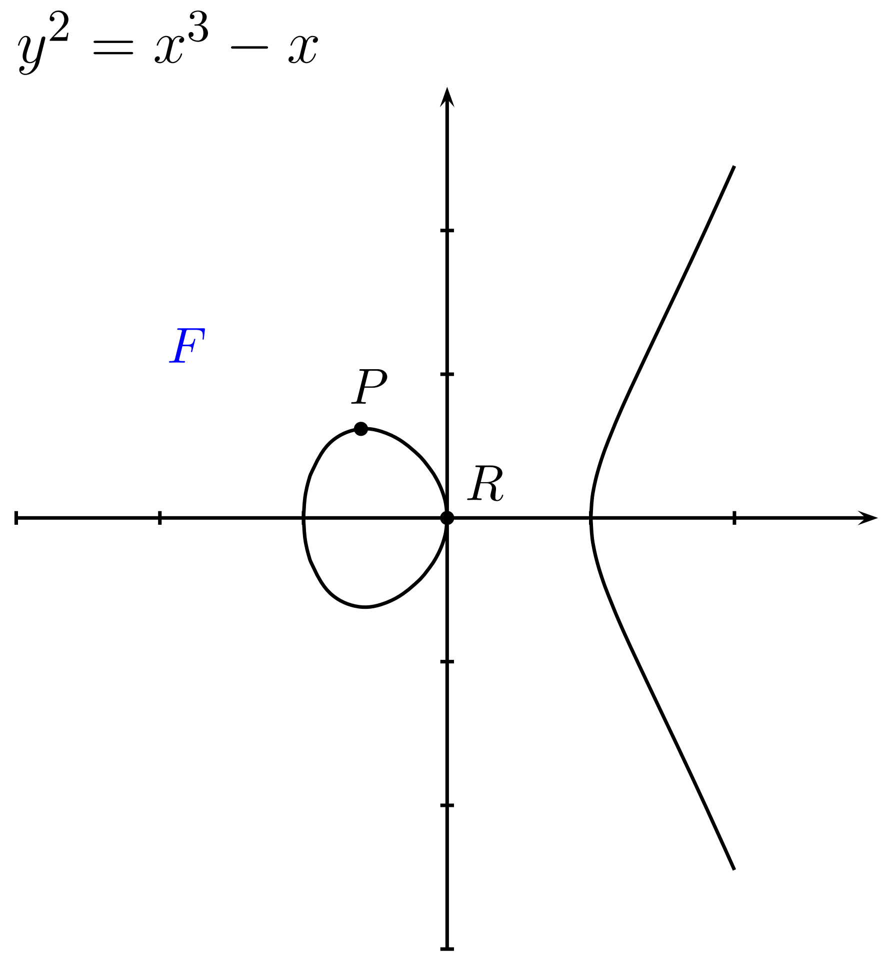
Consider the factor group of the group of divisors of degree zero Div_C^0 modulo the principal divisors. This way one constructs the divisor class group of degree zero.

$$\mathrm{Pic}_C^0 = \mathrm{Div}_C^0 / \mathrm{Princ}_C.$$

Meaning will become clear soon. First example ECC.

So far working over algebraic closure.
First definition: The $\mathbb{F}_q$-rational group elements $\mathrm{Pic}_C^0(\mathbb{F}_q)$ are those which remain fixed under applying the Frobenius endomorphisms, i.e. computing q-th powers of all coordinates. Note that not each point needs to remain fixed for that (sum can be rearranged).

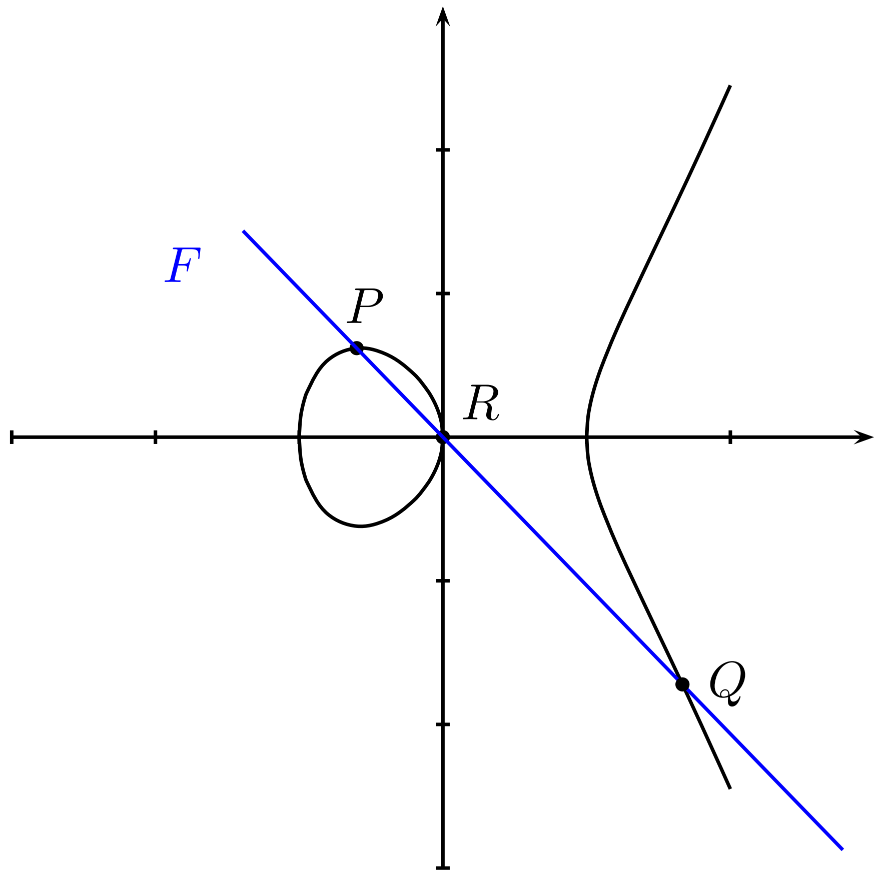
Example: $E(\mathbb{R}), h = 0$



$\mathrm{div}(F(x,y)) = P + Q + R - 3P_\infty$

Example: $E(\mathbb{R}), h = 0$

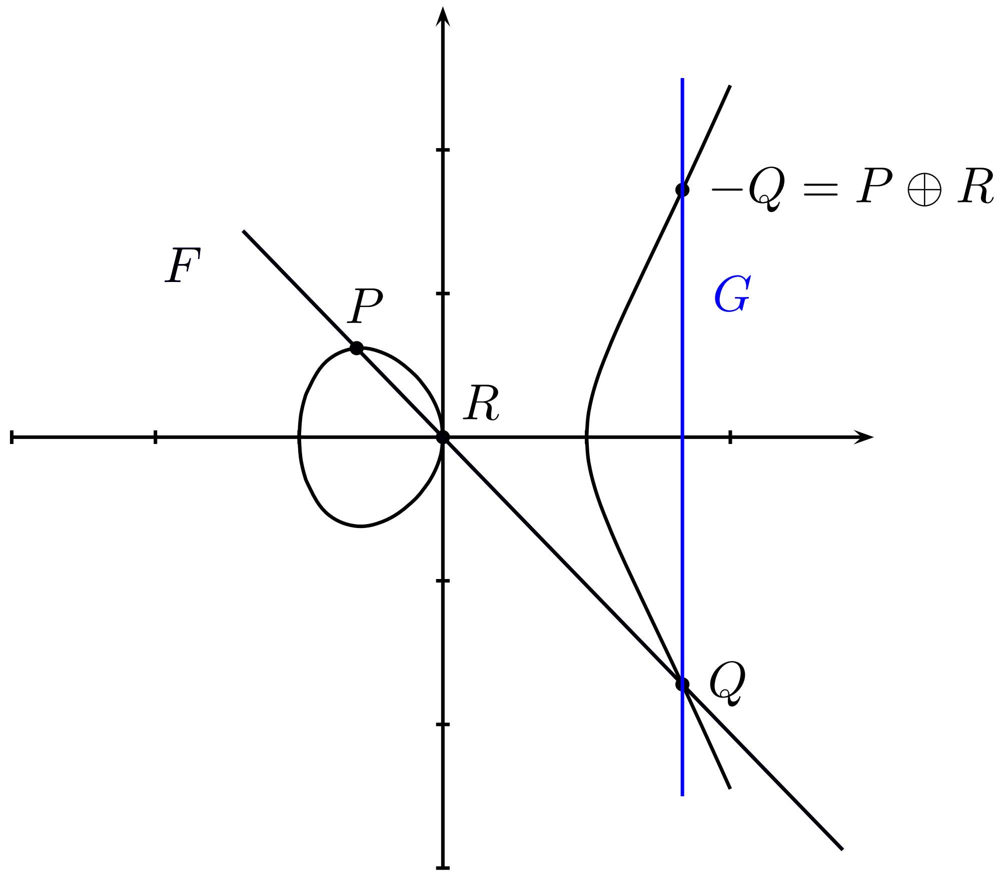
$y^2 = x^3 - x$



$\mathrm{div}(F(x,y)) = P + Q + R - 3P_\infty$

Example: $E(\mathbb{R}), h = 0$

$y^2 = x^3 - x$



$\mathrm{div}(F(x,y)) = P + Q + R - 3P_\infty$ $\qquad$ $\mathrm{div}(G(x,y)) = Q + (-Q) - 2P_\infty$

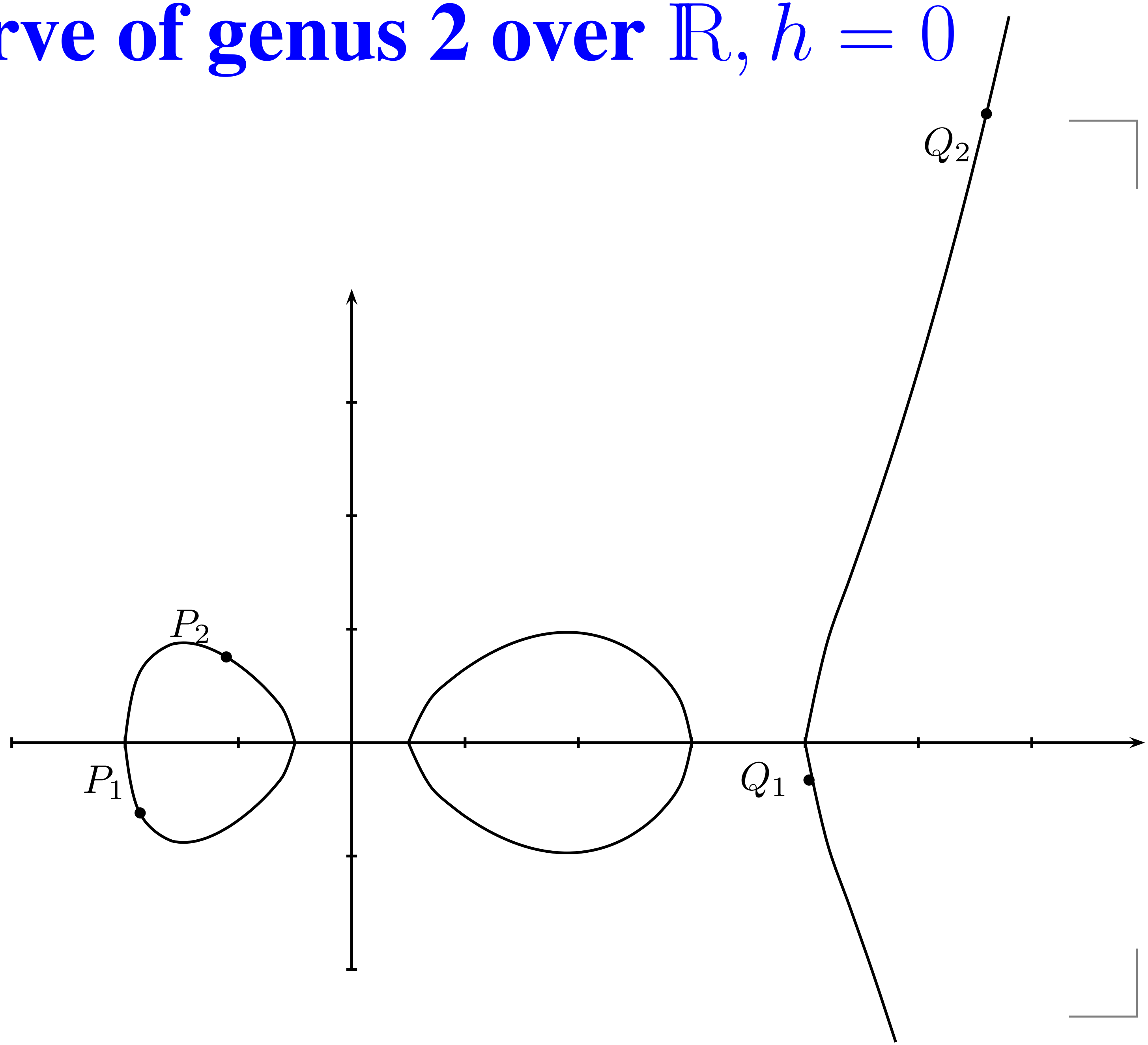
Reduced divisors

- Divisor D is semi-reduced if

$$D = \sum_{\substack{i=1 \\ P_i \in C(\overline{\mathbb{F}_q}) \setminus \{P_\infty\}}}^{m} P_i - mP_\infty$$

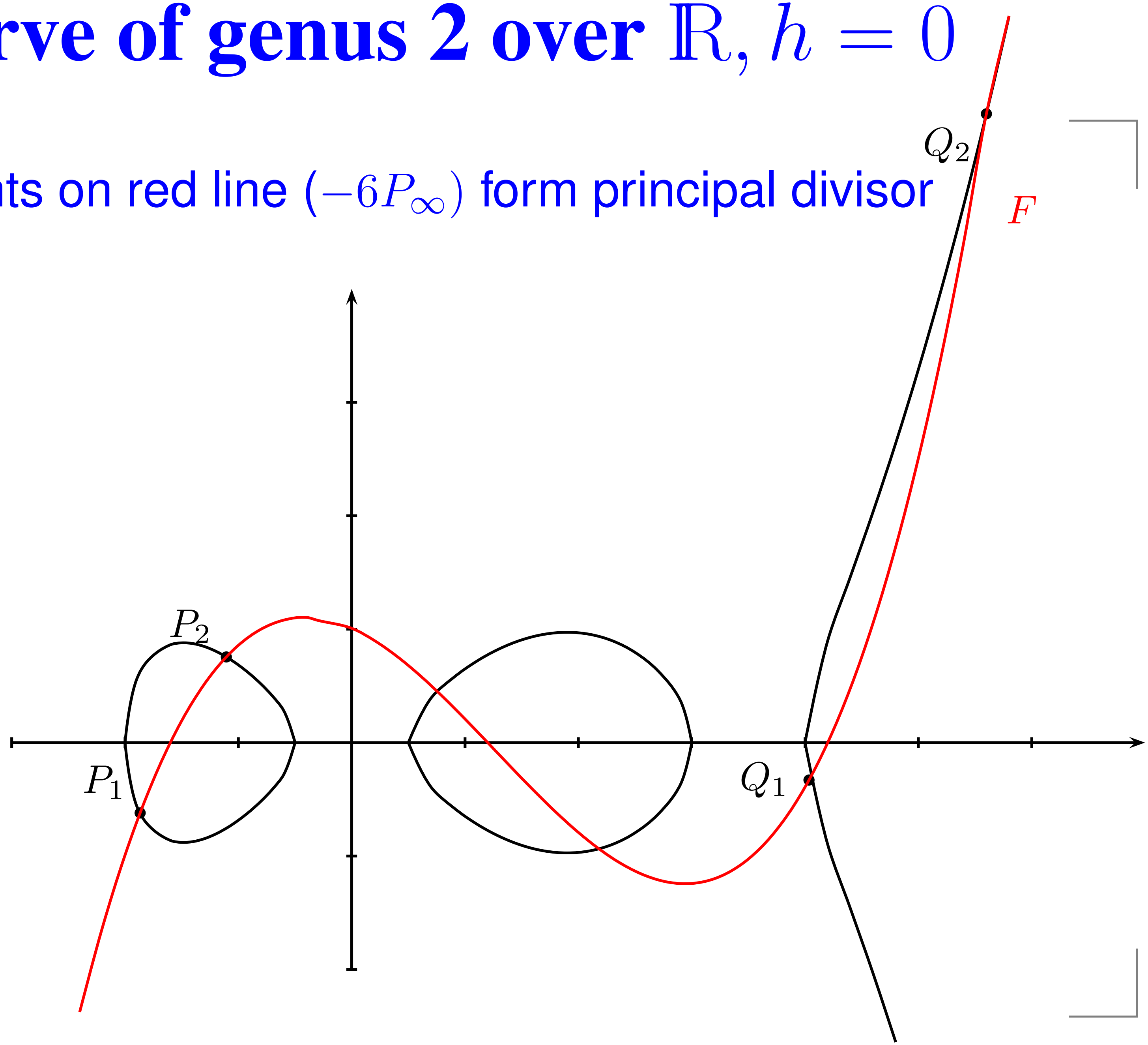
  and $P_i \neq -P_j$ for $i \neq j$ (no restriction on the number of points)
- Divisor D is reduced if it is semi-reduced and $m \leq g$.
- Important result on representation:
  Each divisor class has a unique reduced representative.

Curve of genus 2 over $\mathbb{R}$, $h = 0$



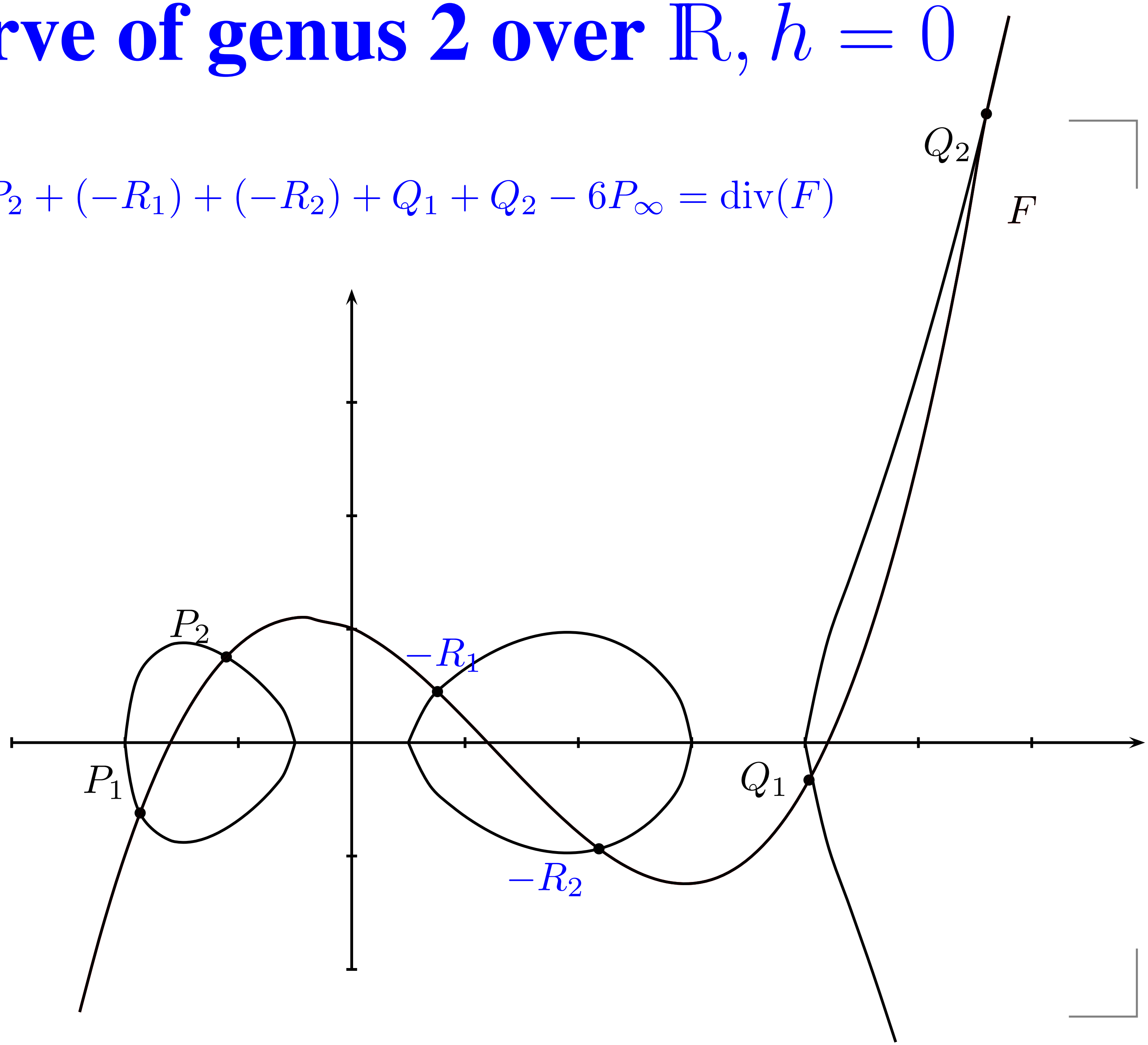
Curve of genus 2 over $\mathbb{R}, h = 0$

points on red line $(-6P_\infty)$ form principal divisor



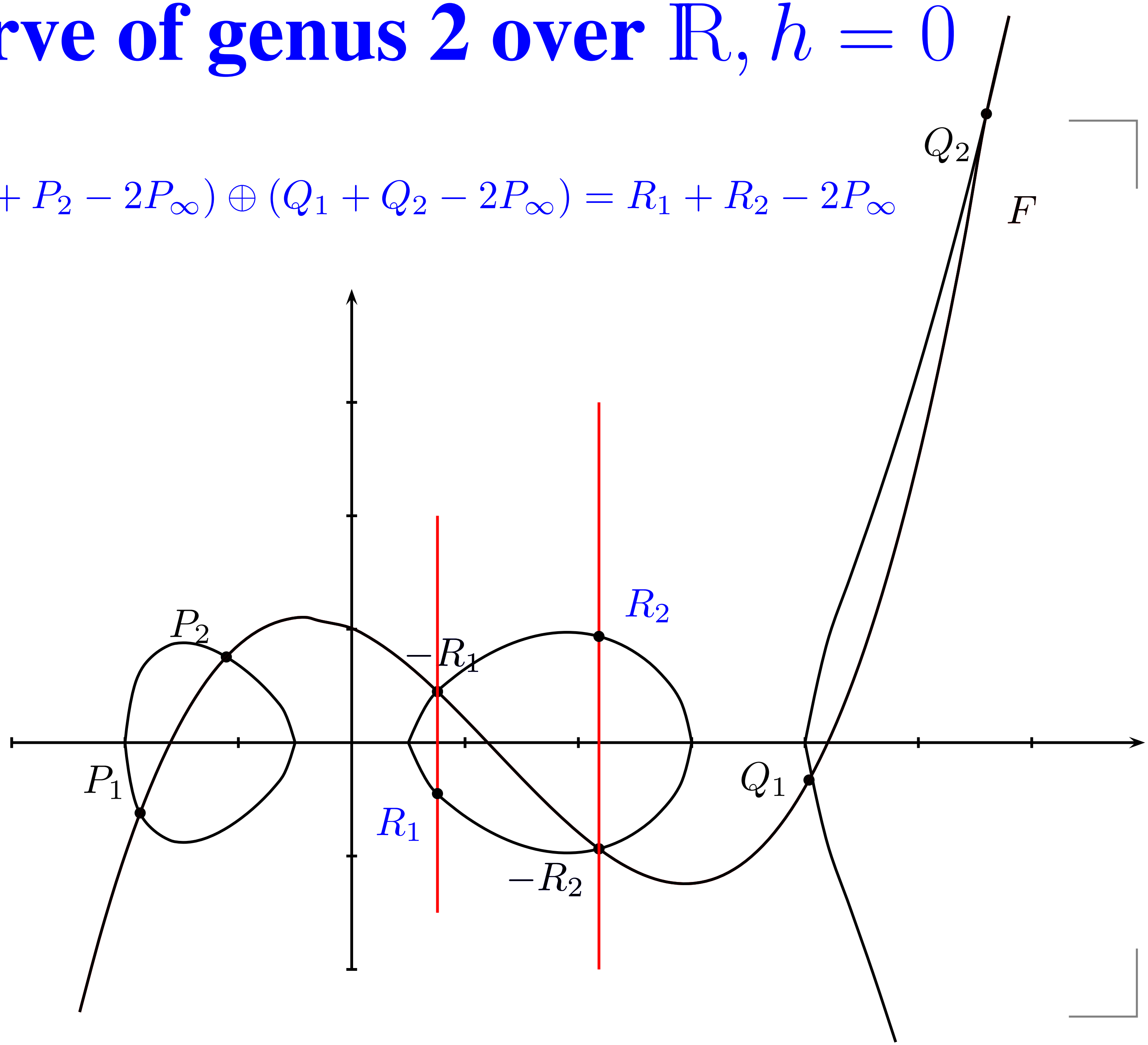
Curve of genus 2 over $\mathbb{R}$, $h = 0$

$$P_1 + P_2 + (-R_1) + (-R_2) + Q_1 + Q_2 - 6P_\infty = \mathrm{div}(F)$$



Curve of genus 2 over $\mathbb{R}$, $h = 0$

$$(P_1 + P_2 - 2P_\infty) \oplus (Q_1 + Q_2 - 2P_\infty) = R_1 + R_2 - 2P_\infty$$



Representation

- Still need a compact representation.
- Idea: use polynomials to represent divisors, ignore P_∞ – multiplicity dictated by affine part.
- Let semi-reduced $D = \sum_{i=1}^{m} P_i - mP_\infty$ with $P_i = (x_i, y_i)$. Put

  $$u(x) = \prod_{i=1}^{m}(x - x_i) \text{ and define } v \text{ by } v(x_i) = y_i \text{ with multiplicity}$$

  (latter gives conditions on derivative of v).
- $\deg(v) < \deg(u) = m$.
- Reduced divisors have $\deg(u) \leq g$.

Mumford Representation

- Easy characterization for field of definition!
  Divisor classes defined over $\mathbb{F}_q$ have $u, v \in \mathbb{F}_q[x]$.
- Divisor classes can be represented by reduced divisors $\Rightarrow$ each class can be represented by two polynomials

$$[u(x), v(x)]; u, v \in \mathbb{F}_q[x],$$

  u monic, $\deg v < \deg u \leq g, \quad u | v^2 + vh - f.$

  $\Rightarrow$ compact representation.
- Alternative viewpoint:

  We define group on elements $[u(x), v(x)]$ with conditions as above according to algorithm on next slide.
- That's all we really need to do crypto on hyperelliptic curves.

Arithmetic on hyperelliptic curves

Composition & Reduction (Cantor/Koblitz)

`IN:` $D_1 = [u_1, v_1], D_2 = [u_2, v_2], C : y^2 + h(x)y = f(x)$
`OUT:` $D = [u, v]$ reduced with $D \sim D_1 + D_2$

1. compute $d_1 = \gcd(u_1, u_2) = e_1 u_1 + e_2 u_2$
2. compute $d = \gcd(d_1, v_1 + v_2 + h) = c_1 d_1 + c_2(v_1 + v_2 + h)$
3. let $s_1 = c_1 e_1, s_2 = c_1 e_2, s_3 = c_2$
4. $u = \frac{u_1 u_2}{d^2} \quad v = \frac{s_1 u_1 v_2 + s_2 u_2 v_1 + s_3(v_1 v_2 + f)}{d} \bmod u$

This result $[u, v]$ corresponds to a semireduced divisor.

Arithmetic on hyperelliptic curves

Composition & Reduction (Cantor/Koblitz)

`IN:` $D_1 = [u_1, v_1], D_2 = [u_2, v_2], C : y^2 + h(x)y = f(x)$
`OUT:` $D = [u, v]$ reduced with $D \sim D_1 + D_2$

1. compute $d_1 = \gcd(u_1, u_2) = e_1 u_1 + e_2 u_2$
2. compute $d = \gcd(d_1, v_1 + v_2 + h) = c_1 d_1 + c_2(v_1 + v_2 + h)$
3. let $s_1 = c_1 e_1, s_2 = c_1 e_2, s_3 = c_2$
4. $u = \frac{u_1 u_2}{d^2} \quad v = \frac{s_1 u_1 v_2 + s_2 u_2 v_1 + s_3(v_1 v_2 + f)}{d} \bmod u$
5. let $u' = \frac{f - vh - v^2}{u} \quad v' = (-h - v) \bmod u'$
6. if $\deg u' > g$ put $u := u', v := v'$ goto step 5
7. make u monic

This result $[u, v]$ corresponds to a reduced divisor.

Index calculus attacks on HECC

- First polynomial u captures x-coordinates of points.
- If u splits over $\mathbb{F}_q$ then the points all have x-coordinates in $\mathbb{F}_q$ (and y in $\mathbb{F}_q$ or $\mathbb{F}_{q^2}$).
- Use factor base only the irreducible polynomials u of small degree.
- Factorization of u tells whether element can be composed from factor base elements. This rules out genus $g \geq 10$.
- Gaudry's idea: use only polynomials of degree 1 for factorbase. This rules out genus $g \geq 5$.
- Harley and Thériault's idea: use only a subset of these polynomials and "large primes". This rules out genus $g \geq 4$ and scratches $g = 3$.

More HECC tomorrow!