

---

*Discrete Logs for  
Curves of Small Genus*

*Summer School on Elliptic and  
Hyperelliptic Curve Cryptography*

*Nicolas Thériault*

ntheria@fields.utoronto.ca

*Fields Institute*

---

With Enge–Gaudry, if  $g \geq v \log q$  ( $v \geq 1$ ) then we can compute the  $DL_a(b)$  in

$$L_{q^g} \left( 1/2, \sqrt{2} \left( \sqrt{1 + \frac{1}{2v}} + \sqrt{\frac{1}{2v}} \right) + o(1) \right)$$

Curves of large genus (compared to  $\log q$ ) should not be considered generic groups.

But what happens if  $g < \log q$ ?

If we look in detail at Enge–Gaudry, we chose

$$B \leq \left\lfloor \log_q \left( L_{q^g} \left( 1/2, 1/\sqrt{2} \right) \right) \right\rfloor$$

Clearly we can't choose  $B < 1$ , but we should still choose it as small as possible, so why not try  $B = 1$ ...

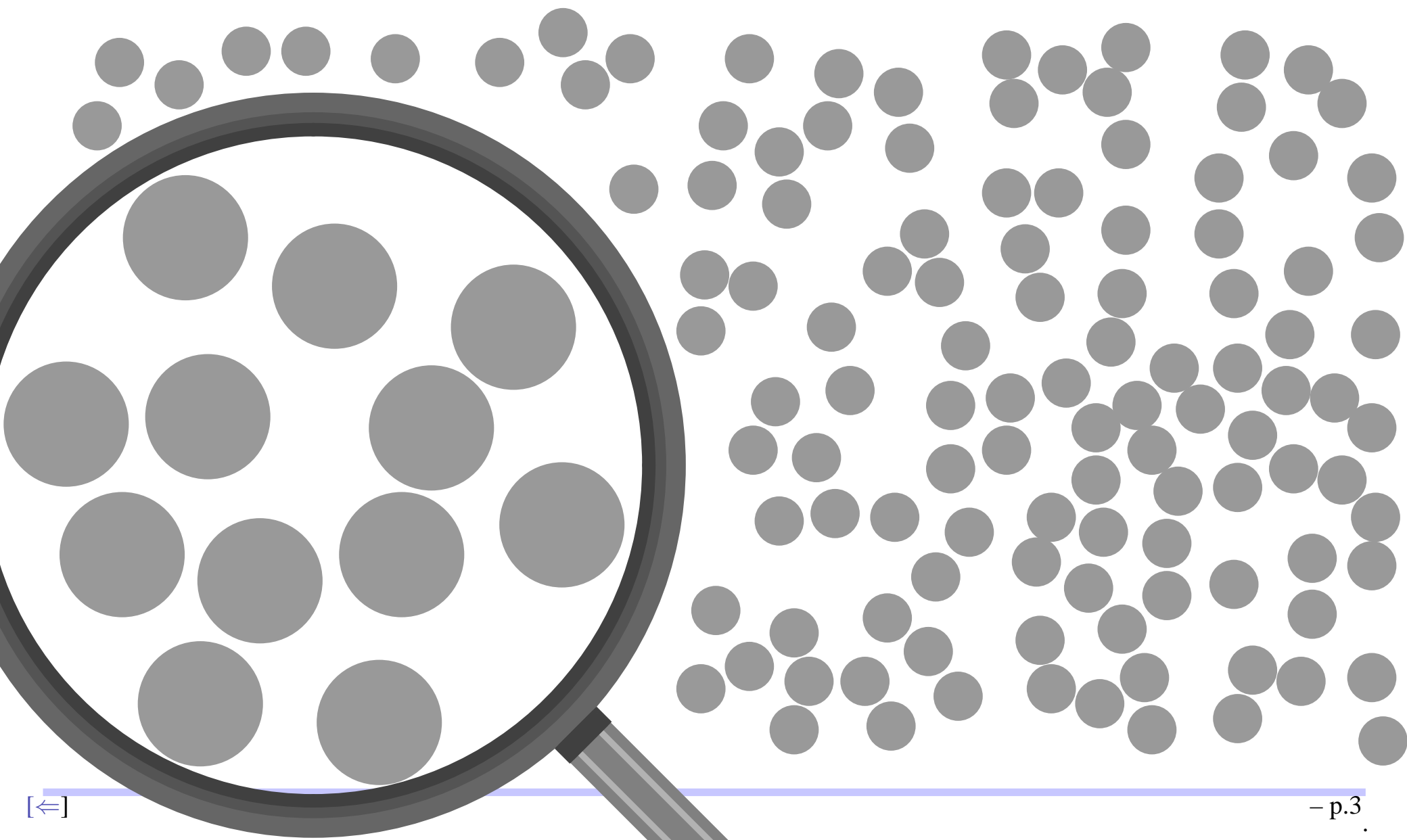
- $k_1 = |\{\text{prime divisors of degree 1}\}|$ , so  
 $k_1 = |C(\mathbb{F}_q) \setminus \{P_\infty\}| = q + O(g\sqrt{q})$ .
- $p_1 = Pr(\text{reduced divisor is 1-smooth}) \approx 1/g!$ .
- $\omega = g$ .

The index calculus algorithm will take

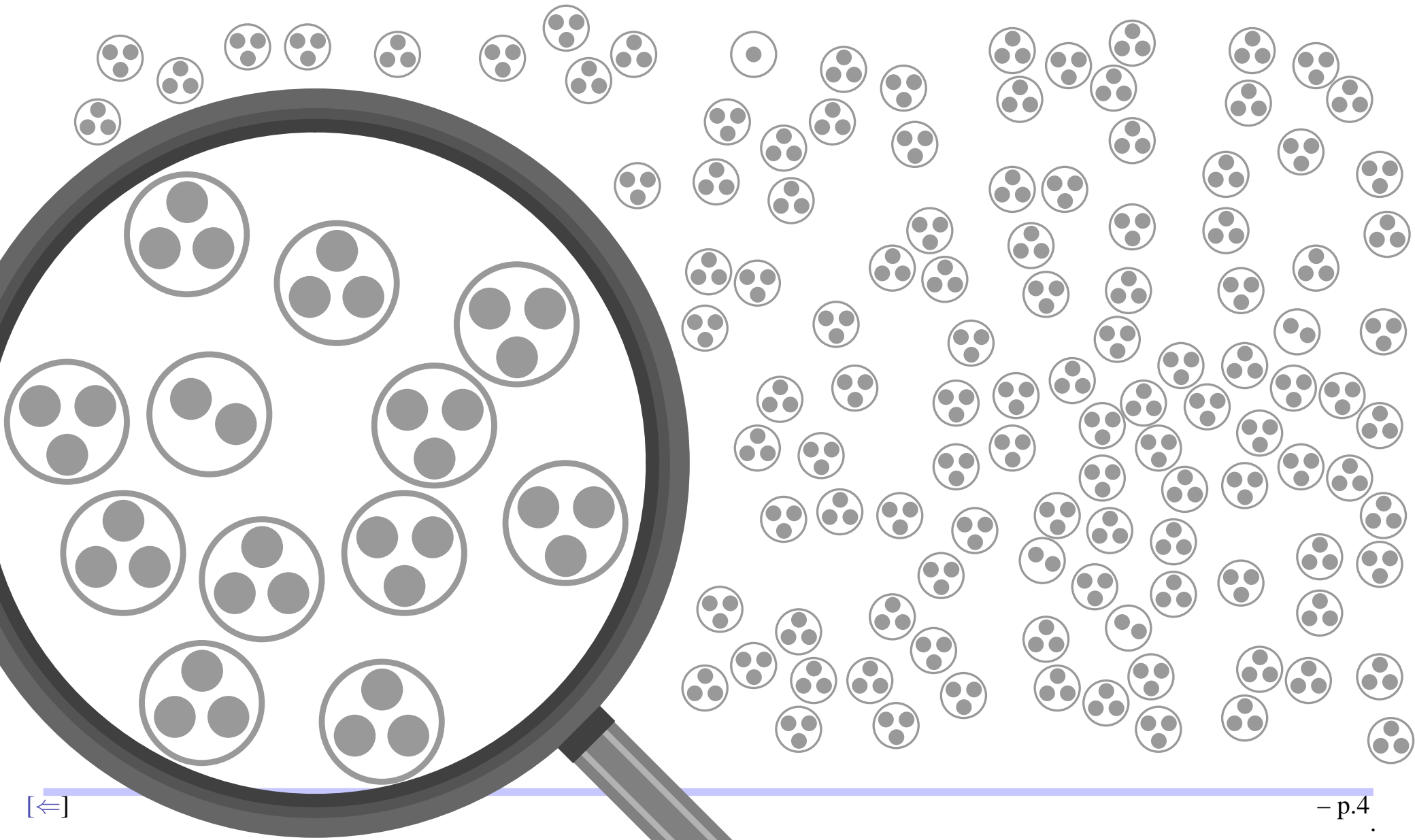
$$O(k_1/p_1) + O(\omega k_1^2) = O(g!q) + O(gq^2)$$

operations (ignoring the log factors).

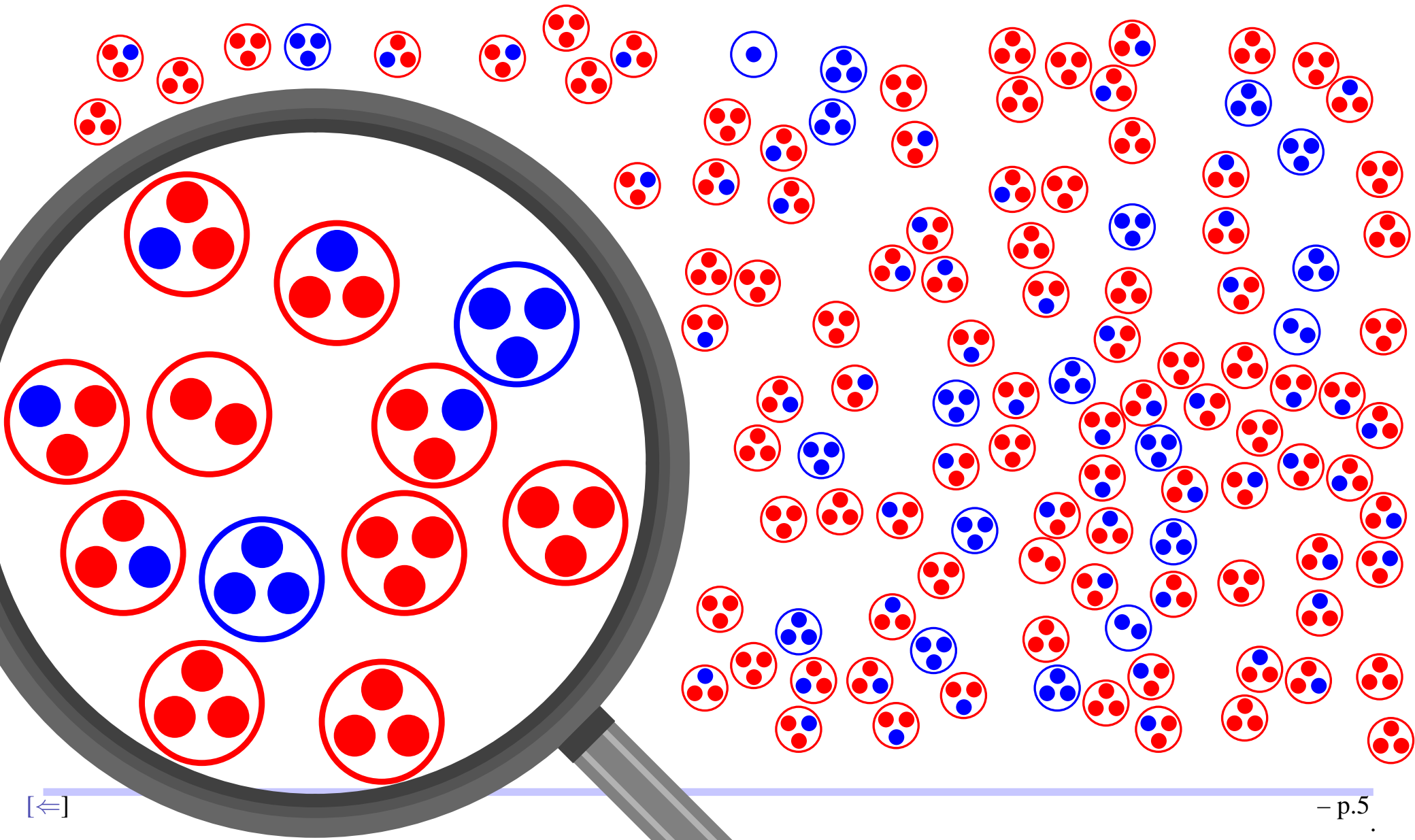
If  $g < \log q$ , the time is dominated by  $O(gq^2)$ .



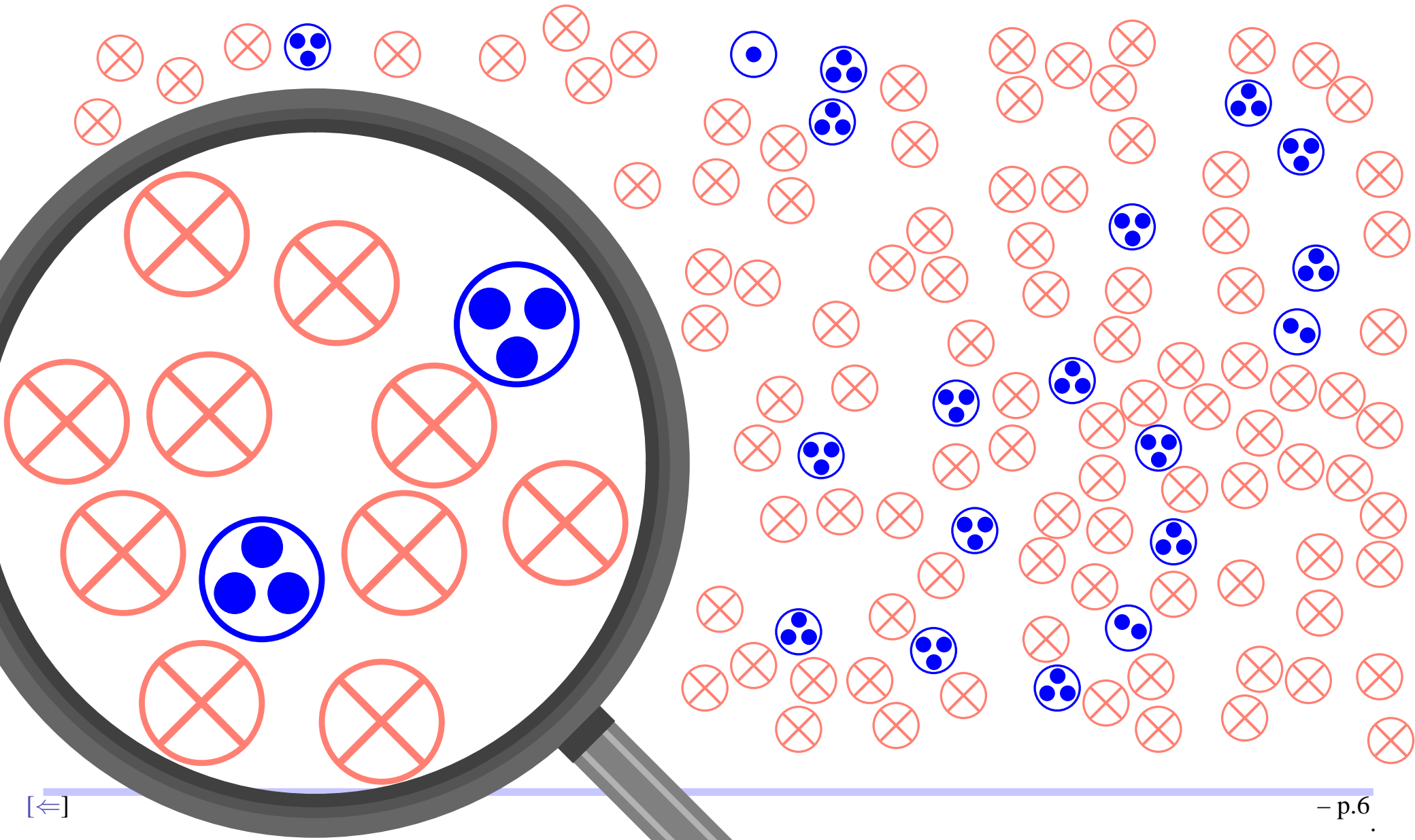
# *Index Calculus*



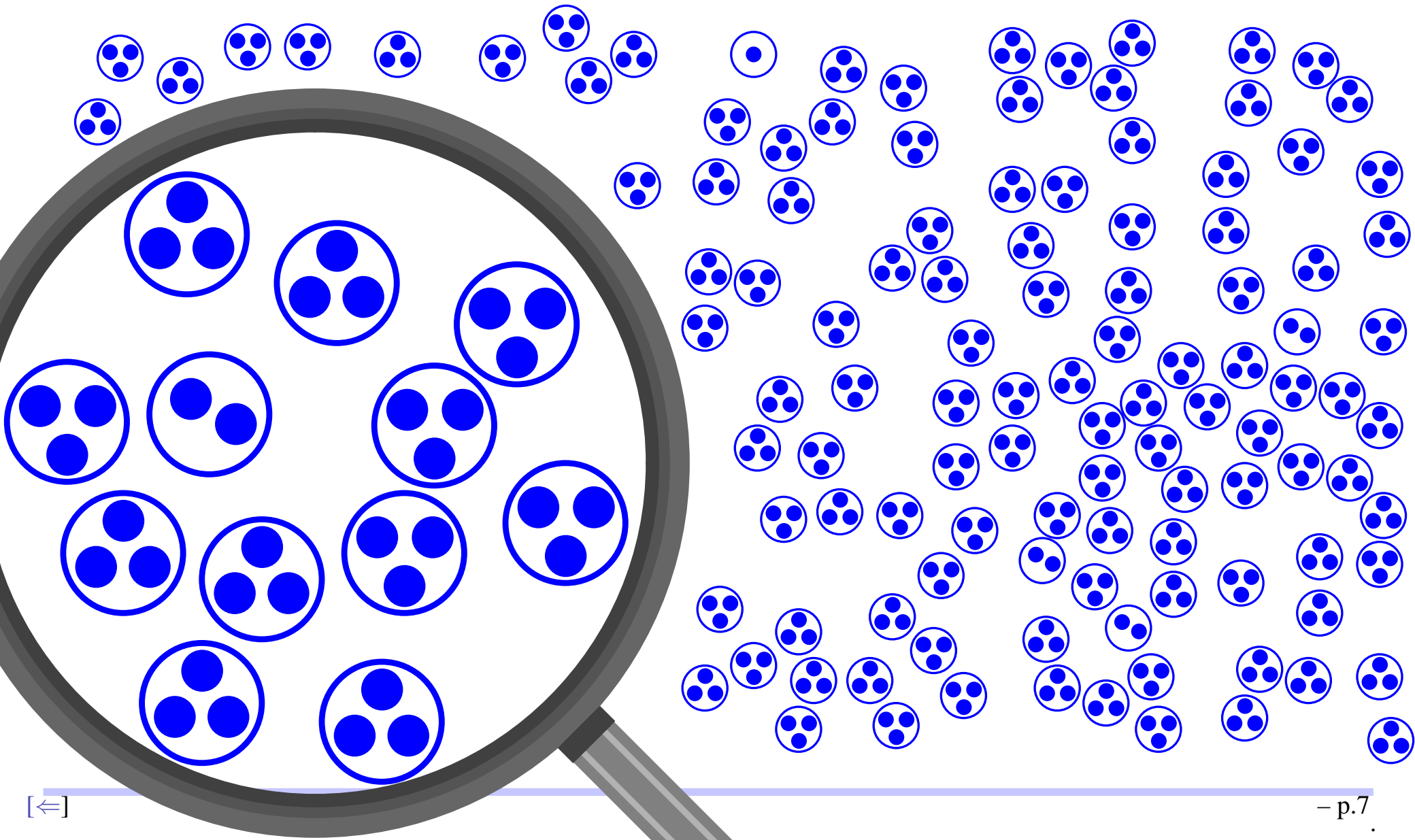
# Gaudry's Algorithm



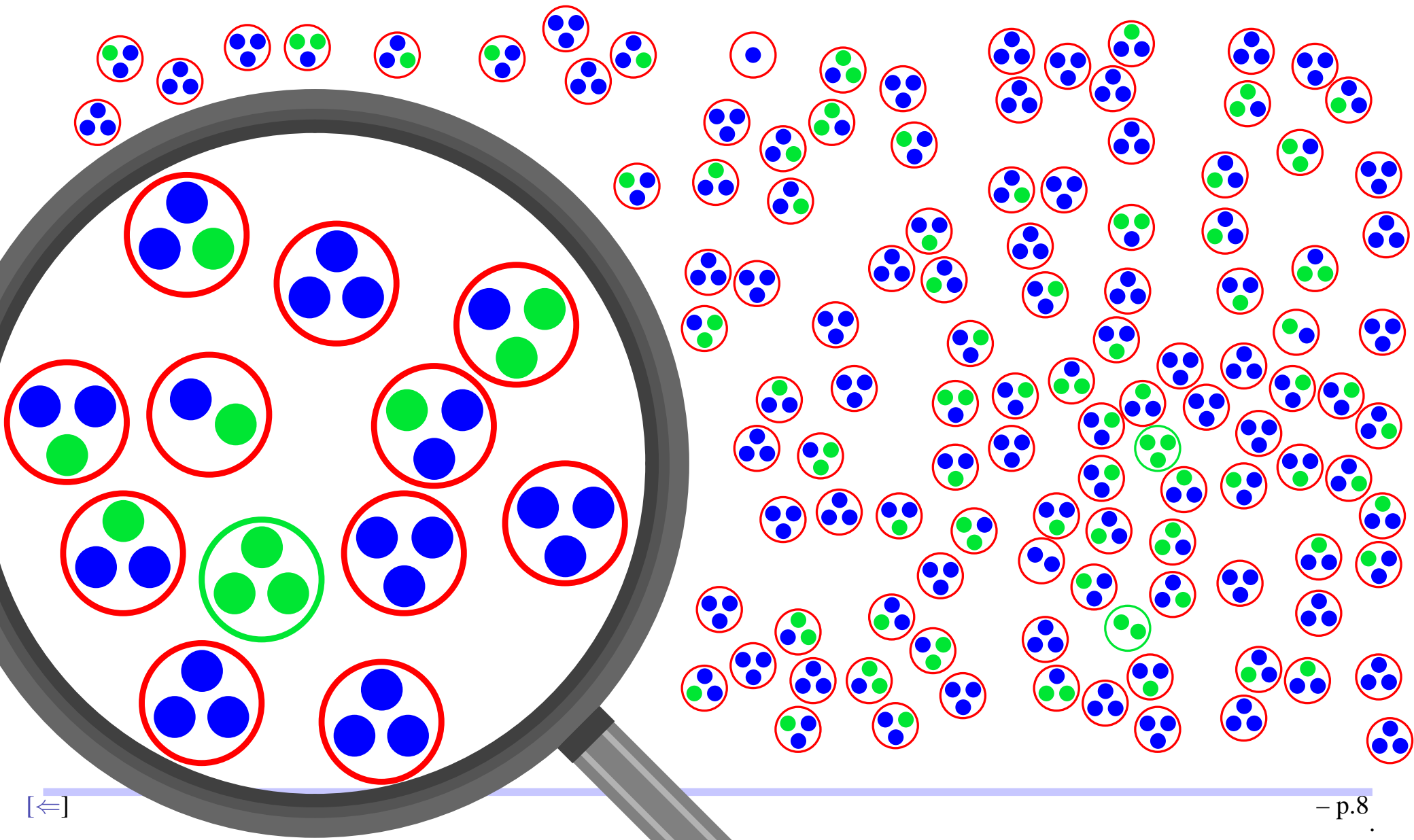
# Gaudry's Algorithm



# Gaudry's Algorithm



# Reduced Factor Base



Let  $\mathcal{P} = C(\mathbb{F}_q)$ , i.e.  $\mathcal{P}$  is the set of points of  $C$  over  $\mathbb{F}_q$ . Let  $\mathcal{B}$  be a subset of size (order)  $B$  of  $\mathcal{P}$ .

A divisor is  *$\mathcal{B}$ -smooth* if it is reduced and it can be written in the form

$$\sum_{i=1}^k P_i - kP_\infty$$

with the  $P_i$ 's in  $\mathcal{B}$  and  $k \leq g$ .

In this case,  $\mathcal{B}$  is called the *factor base*.

The probability that a reduced divisor is  $\mathcal{B}$ -smooth is

$$p_{\mathcal{B}} \approx \frac{1}{g!} \frac{B^g}{q^g}$$

The cost of the index calculus algorithm becomes

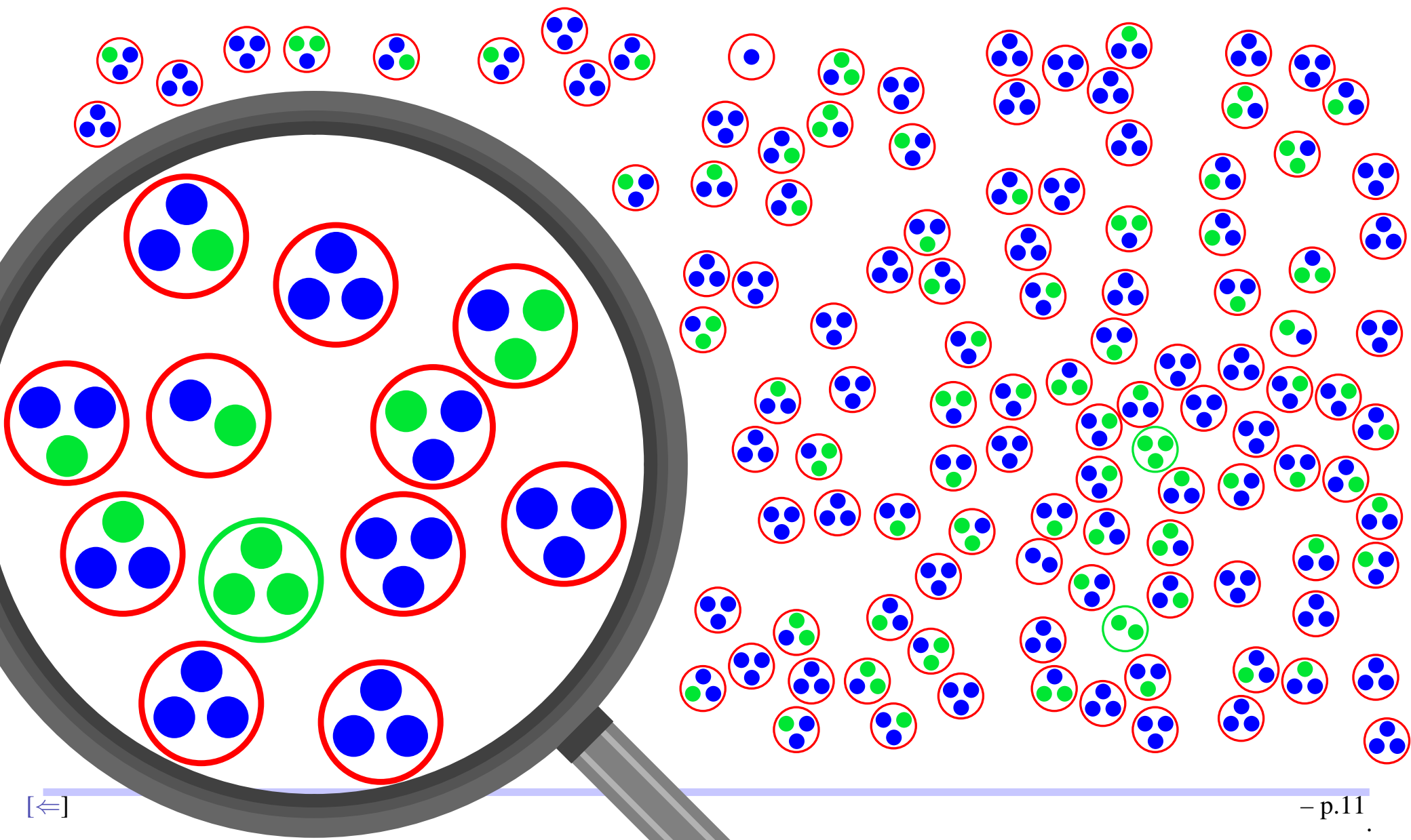
$$O(B/p_{\mathcal{B}}) + O(gB^2) = O(g!q^g/B^{g-1}) + O(gB^2) .$$

To minimize the cost, we choose  $B = O\left(gq^{1-\frac{1}{g+1}}\right)$ , and we get a total cost of

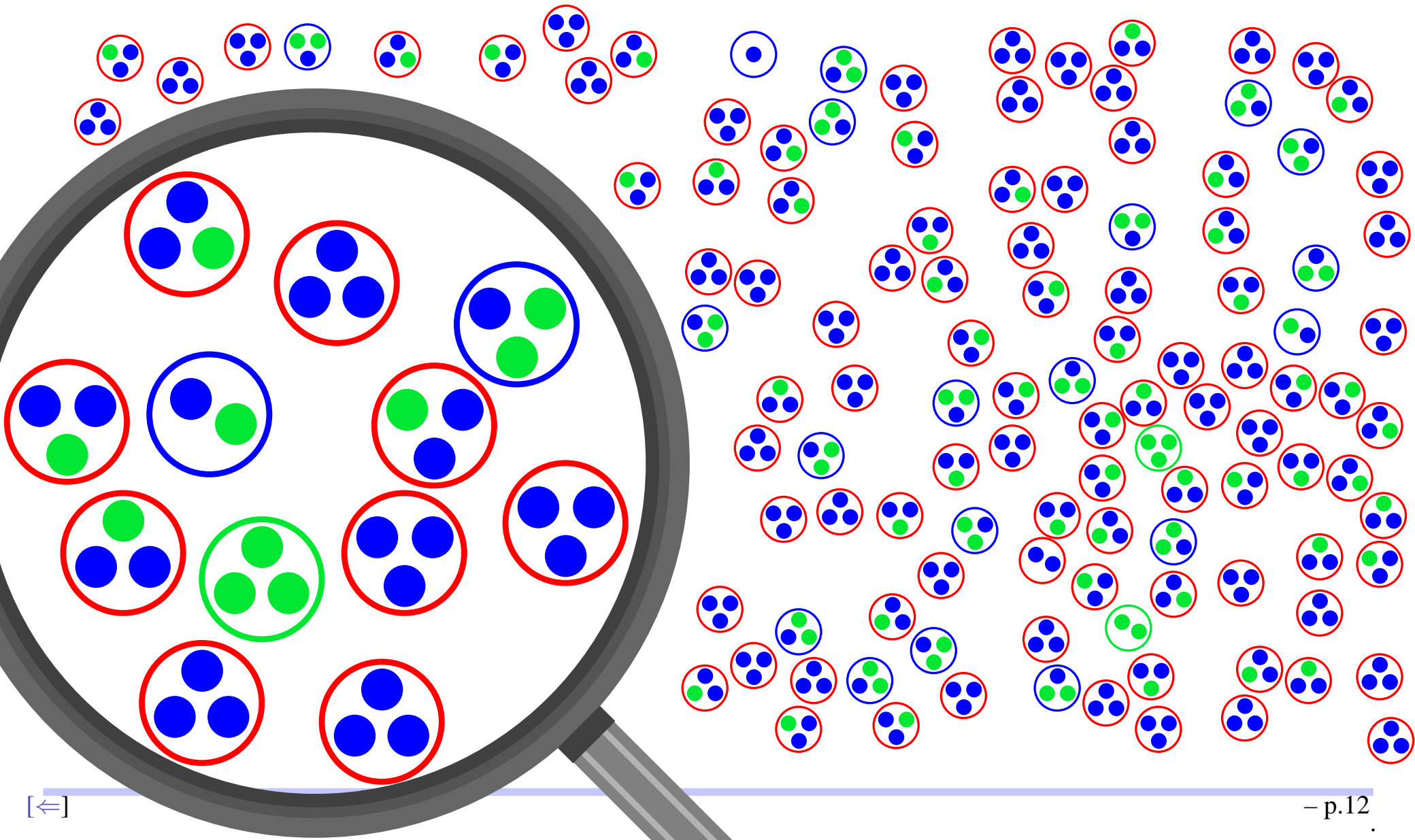
$$O\left(g^3 q^{2-\frac{2}{g+1}}\right)$$

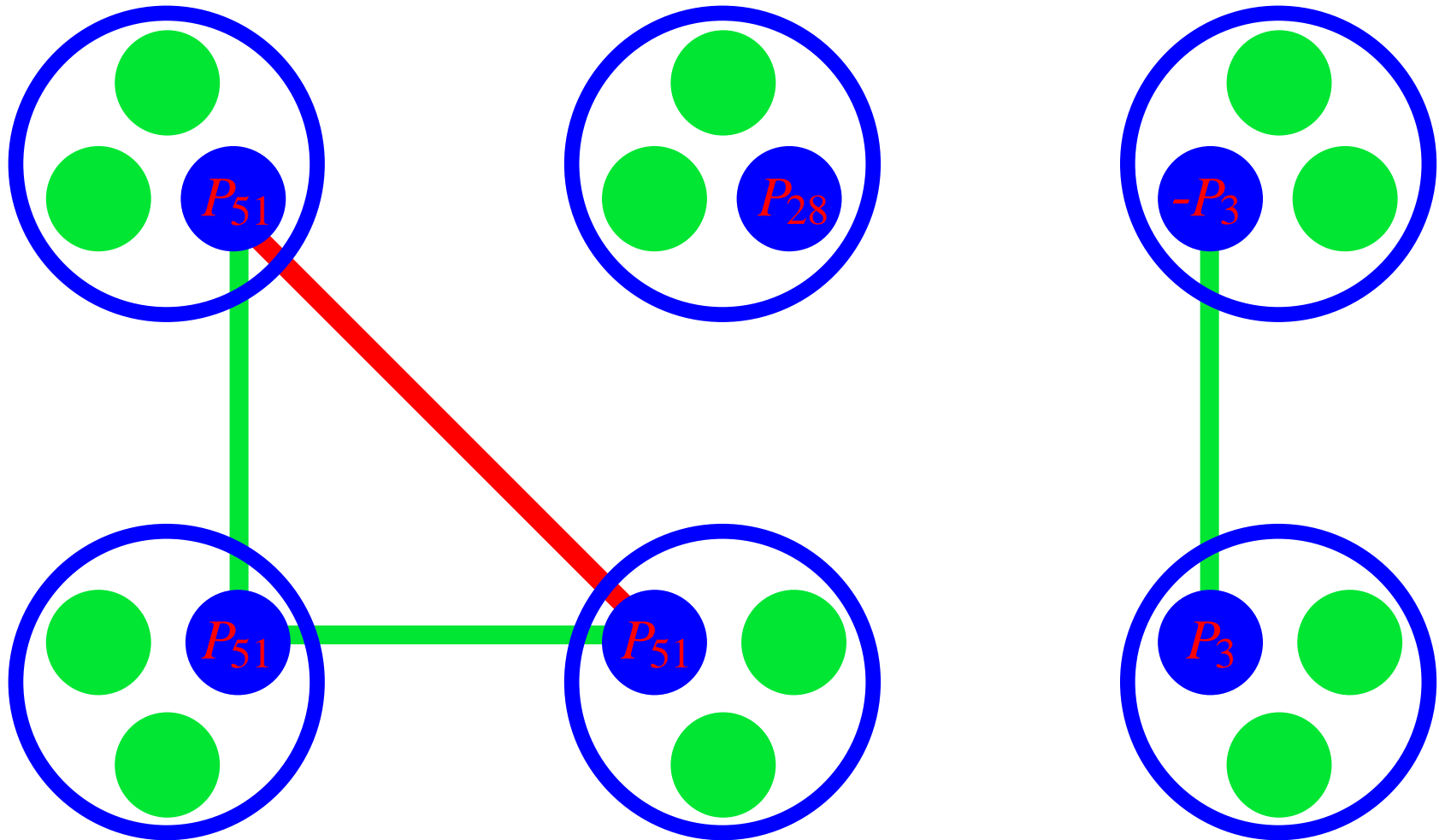
(group) operations.

# Large Primes



# Large Primes





Given a factor base  $\mathcal{B} \subset \mathcal{P}$ , a point  $P \in \mathcal{P}$  is called a *large prime* if  $P \notin \mathcal{B}$ .

A reduced divisor

$$D = \sum_{i=1}^k P_i - kP_\infty$$

is said to be *almost-smooth* if:

- all but one of the  $P_i$ 's are in  $\mathcal{B}$ ;
- the remaining  $P_i$  is a large prime.

- Let  $T_i$  be an almost-smooth divisor with the large prime  $P$ .
- $T_i$  is called an *intersection* if one of the previous almost-smooth divisor ( $T_j$ ) has large prime  $\pm P$ .
- We use the intersection of  $T_i$  with  $T_j$  to build a divisor that *factors over the factor base*.
- Intersections are used to decrease the time required to build the linear algebra system.
- $T_i$  is an intersection with *at most one* of the previous almost-smooth  $T_j$ 's.

Let  $R_{n,s}$  be the number of intersections after  $s$  draws from a set of  $n$  elements.

**Theorem:** If  $3 \leq s < n/2$ , then  $E[R_{n,s}]$  is between  $\frac{2s^2}{3n}$  and  $\frac{s^2}{n}$ .

Note:  $\text{Var}(R_{n,s}) = O(n)$ , so we can use Chebyshev...

For us,  $n \approx q$  and we want to choose  $s$  such that  $O(\frac{s^2}{n}) = B$ .

We need  $O(\sqrt{qB})$  almost-smooth divisors to get  $B$  intersections.

The probability that a reduced divisor is almost-smooth is

$$p_B \approx \frac{1}{(g-1)!} \frac{B^{g-1}}{q^{g-1}}$$

The cost of the index calculus algorithm becomes

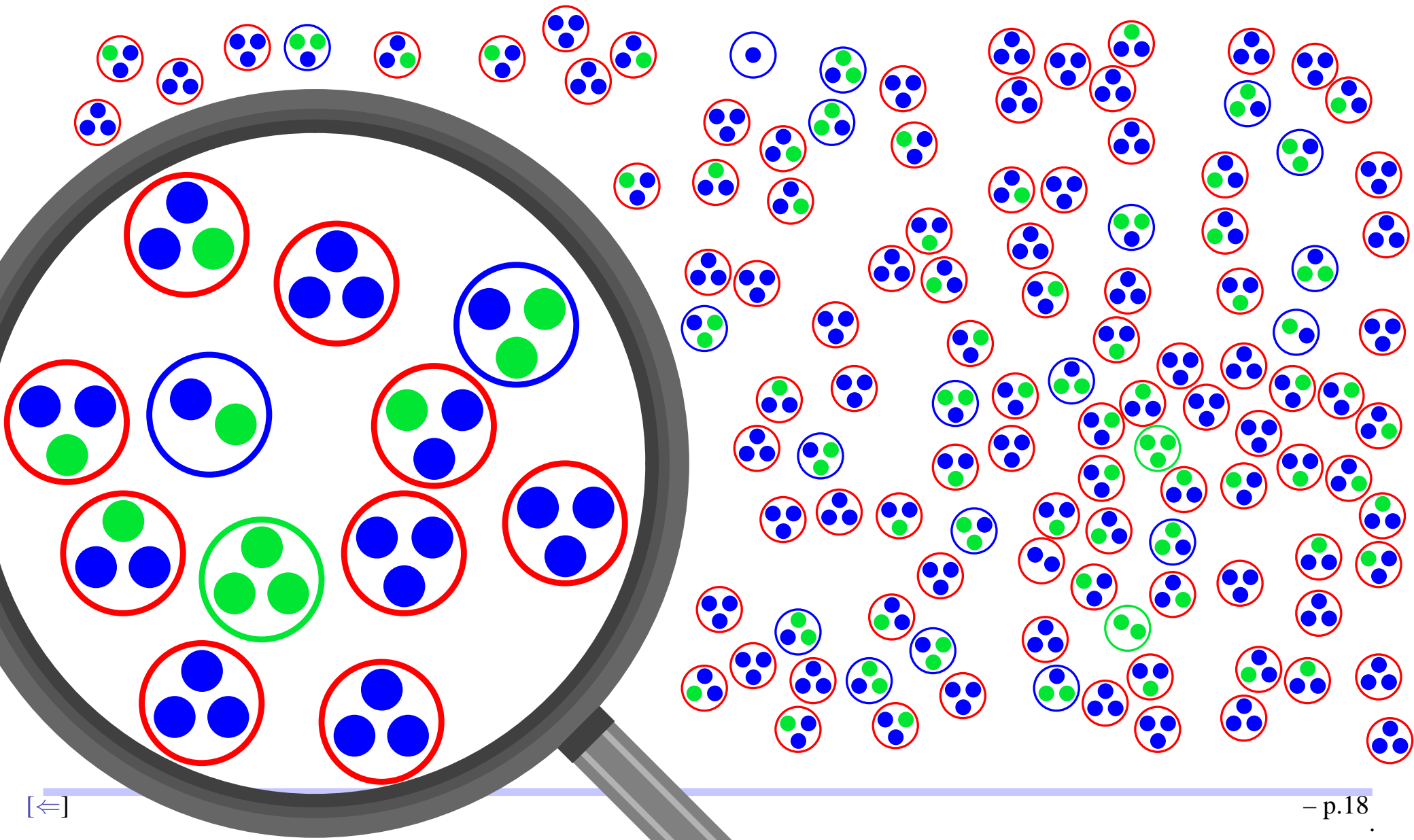
$$O(\sqrt{qB}/p_B) + O(gB^2) = O((g-1)!q^{g-1/2}/B^{g-3/2}) + O(gB^2) .$$

To minimize the cost, we choose  $B = O\left(gq^{1-\frac{1}{g+1/2}}\right)$ , and we get a total cost of

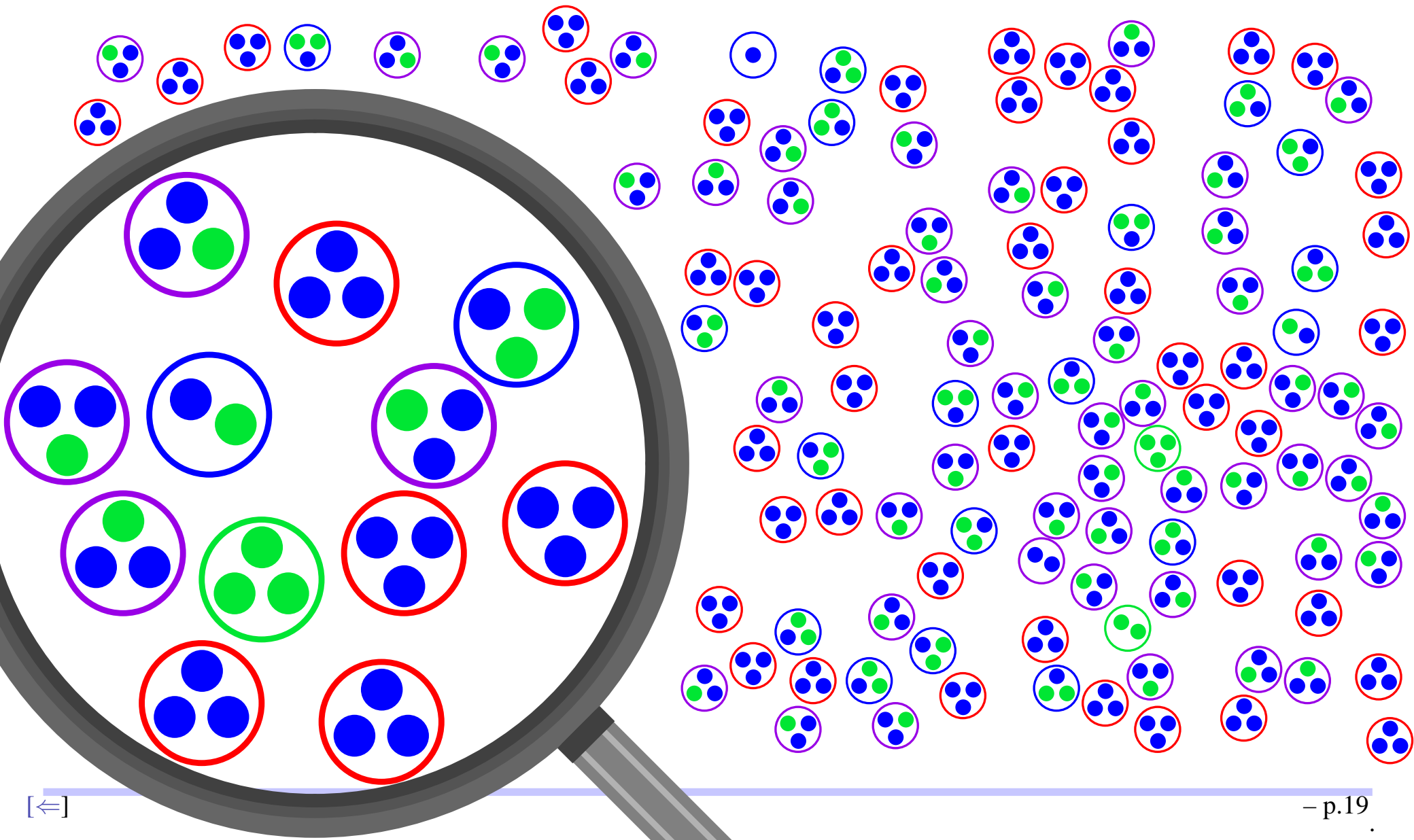
$$O\left(g^3 q^{2-\frac{2}{g+1/2}}\right)$$

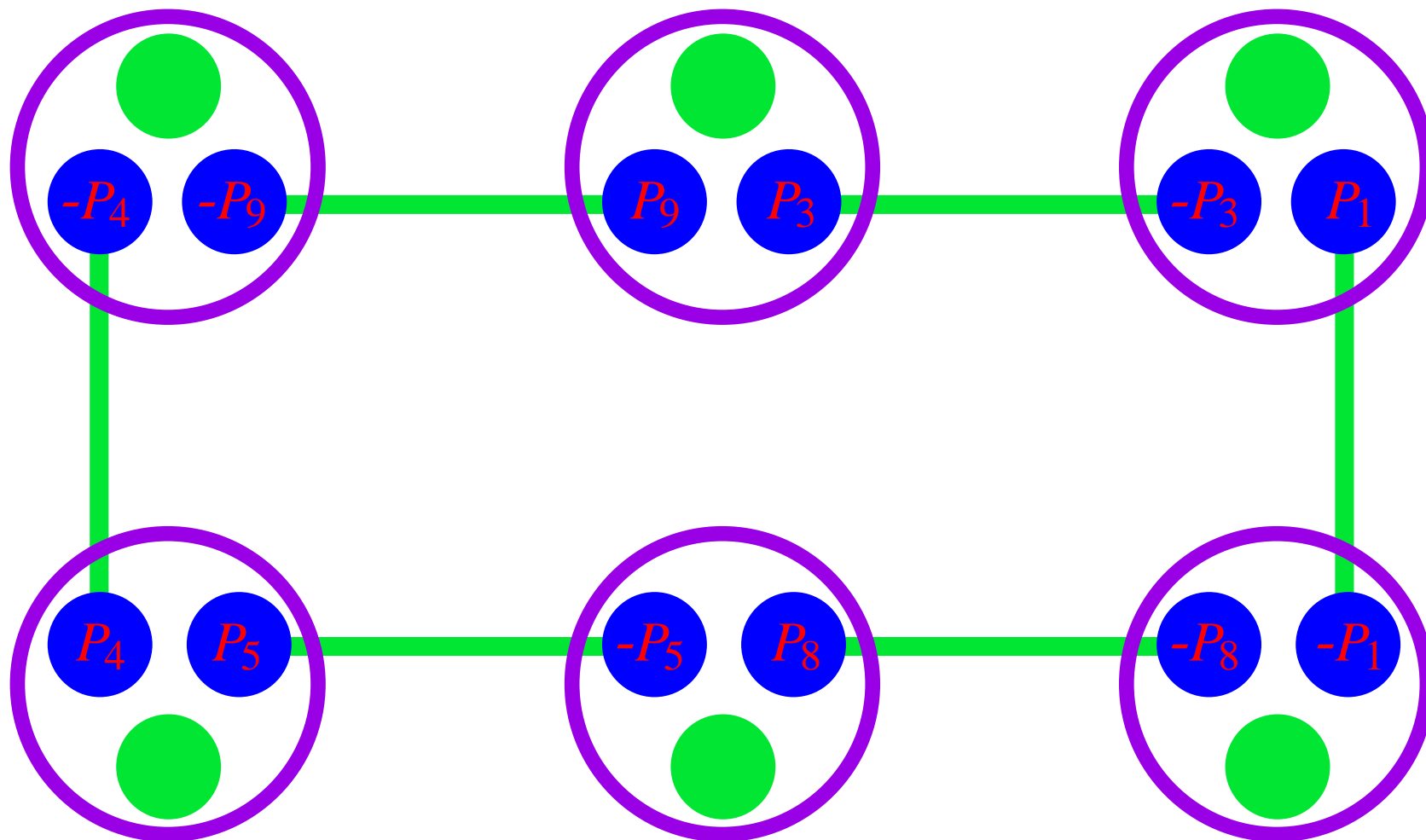
(group) operations.

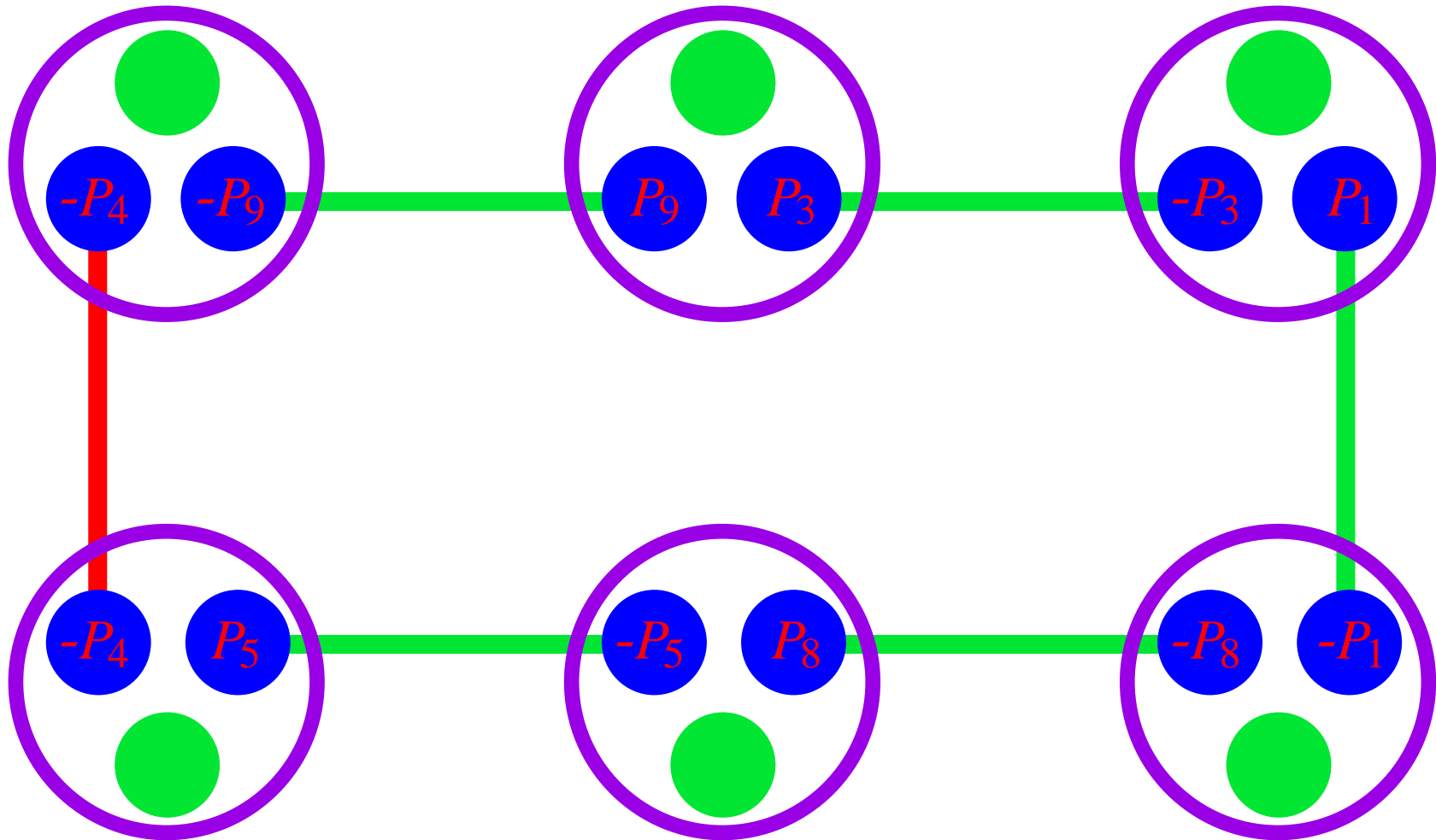
# Double Large Primes

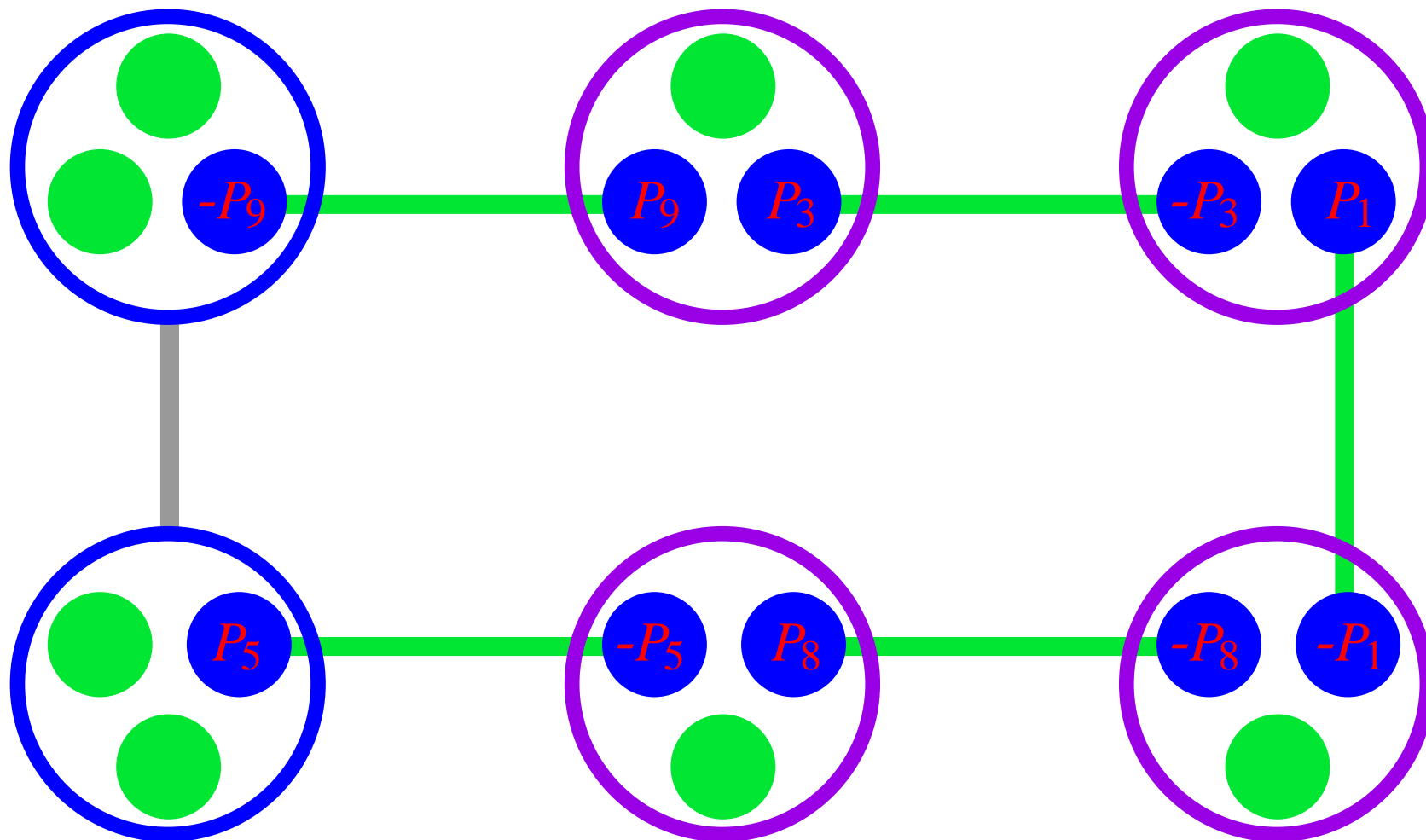


# Double Large Primes









A reduced divisor

$$D = \sum_{i=1}^k P_i - kP_\infty$$

is said to be *2-almost-smooth* if:

- all but two of the  $P_i$ 's are in  $\mathcal{B}$ ;
- the two remaining  $P_i$ 's are large primes.

- All almost-smooth and 2-almost-smooth divisors are used to define edges in a **graph of large primes** ( $\approx q$  vertices).
- The vertex **1** is used for the “second large prime” in almost-smooth divisors.
- Cycles in the graph correspond to smooth relations
  - If **1** is not in the cycle, only true 1/2 the time.
- First cycle obtained after including  $\sim q/2$  edges.
- $B$  cycles are found after  $\leq q + B$  edges are present (heuristically  $(1 + o(1))q/2$ ).
- **Problem: What is the average length of the cycles?**
- **Problem: Need to store  $O(q)$  divisors.**

# Simplified graph method

- Only consider edges when they are connected to **1**.
- Two phases:
  1. Build the graph (tree) until it has  $N_{max}$  edges.
  2. Build relations using the  $N_{max}$  large primes.
- For phase 1, The first edge is an almost-smooth divisor, all other edges are 2-almost-smooth divisors.
  - The asymptotic time is almost the same.
- For phase 2, we also use divisors with any number of large primes.
  - Decreases  $N_{max}$  to  $\sim q^{1-1/g+1/g^2}$  (from  $\sim q^{1-1/2g}$ ).

**Theorem:** *To get  $B$  cycles from the simplified graph we need  $O(q(\log q)/g)$  2-almost-smooth divisors. The cycles obtained have average length  $O((\log q)/g)$ .*

Note: With the full graph, we need  $O(q)$  2-almost-smooth divisors to get  $O(B)$  cycles, but we can't bound the average length of the cycles (for now).

# Double Large Primes

The probability that a reduced divisor is 2-almost-smooth is

$$p_B \approx \frac{1}{2!(g-2)!} \frac{B^{g-2}}{q^{g-2}}$$

The cost of the index calculus algorithm becomes

$$O\left(\frac{\log q}{g} \frac{q}{p_B}\right) + O((\log q)B^2)$$

To minimize the cost, we choose  $B = O\left(gq^{1-\frac{1}{g}}\right)$ , and we get a total cost of

$$O\left((\log q)g^2q^{2-\frac{2}{g}}\right)$$

(group) operations.

For small genus, we have:

| $g$ | square root attacks | original index calculus | smaller factor base | one large prime | two large prime |
|-----|---------------------|-------------------------|---------------------|-----------------|-----------------|
| 3   | $q^{3/2}$           | $q^2$                   | $q^{3/2}$           | $q^{10/7}$      | $q^{4/3}$       |
| 4   | $q^2$               | $q^2$                   | $q^{8/5}$           | $q^{14/9}$      | $q^{3/2}$       |
| 5   | $q^{5/2}$           | $q^2$                   | $q^{5/3}$           | $q^{18/11}$     | $q^{8/5}$       |
| 6   | $q^3$               | $q^2$                   | $q^{12/7}$          | $q^{22/13}$     | $q^{5/3}$       |

# Non-Hyperelliptic Curves

Most of what we said so far carries over nicely to non-hyperelliptic curves.

There are a few problems with the distributions of the primes and the large primes, but the difference is small and can be included in the error term.

But non-hyperelliptic curves seem to be even less generic than hyperelliptic curves.

To show this, Diem went back to the ideas of Adleman, DeMarrais and Huang.

If we intersect a line  $y = mx + b$  with a curve of degree  $d$ , we get  $d$  points of intersection. Their sum  $(-dP_\infty)$  is a divisor in the class  $\mathbf{0}$ .

# Non-Hyperelliptic Curves

In general, a non-hyperelliptic curve is given by an equation of degree  $d \leq g + 1$  where both  $x$  and  $y$  appear with degree greater than 2.

But to define a line we need two points, so if we select those in a smart way (in the factor base), we are left with  $d - 2$  “random” points. The “random” part acts like our randomly chosen reduced divisors.

Finding a smooth relation for  $a$  and one for  $b$ , and “hoping” they have an impact on the vector in the kernel, we can do everything we did before.

With the double large prime idea, we get  $O(q^{2 - \frac{2}{d-2}})$ , or  $O(q^{2 - \frac{2}{g-1}})$  in general.

Suppose we have a curve  $C$  defined over a field  $\mathbb{F}_{q^n} = \mathbb{F}_q[\alpha]/(r(\alpha))$  by an equation

$$F(X, Y) = 0 .$$

We can replace any  $a \in \mathbb{F}_{q^n}$  by  $\sum_{i=0}^{n-1} a_i \alpha^i$  where the  $a_i$  are in  $\mathbb{F}_q$ . Similarly, a variable  $X$  over  $\mathbb{F}_{q^n}$  can be replaced by a sum (in  $\mathbb{F}_{q^n}$ ) of variables  $\{x_0, x_1, \dots, x_{n-1}\}$  over  $\mathbb{F}_q$ . Substituting into  $F(X, Y)$ , we get

$$\sum_{i=0}^{n-1} F_i(x_0, x_1, \dots, x_{n-1}, y_0, y_1, \dots, y_{n-1}) \alpha^i = 0$$

where the  $F_i(\dots)$  are defined over  $\mathbb{F}_q$ .

Since the  $\alpha^i$  are linearly independent, we must have  $F_i(\dots) = 0$  for every  $i$ .

We get a  $n$  equations in  $2n$  variables. To obtain a curve, we intersect with  $n - 1$  hyperplanes, giving  $2n - 1$  equations in  $2n$  variables.

The result is a curve  $\tilde{C}$  of higher genus than  $C$ , but defined over  $\mathbb{F}_q$ , which could be easier to attack using index calculus.

We also get a map from  $Jac(C)(\mathbb{F}_{q^n})$  to  $Jac(\tilde{C})(\mathbb{F}_q)$ .

In some cases the genus increases by a factor of  $\approx n$ , and the discrete log problem may be easier in  $Jac(\tilde{C})(\mathbb{F}_q)$ .

**In general the genus increases exponentially, and the discrete log problem is harder in  $Jac(\tilde{C})(\mathbb{F}_q)$ .**

In practice, we use different methods to construct the curve  $\tilde{C}$  (so we have a better chance of obtaining a nice genus).

Some cases are more at risk:

- elliptic and hyperelliptic curves in characteristic 2 with  $n$  small, in particular  $n = 2, 3, 5, 7, 31$  (some)
- elliptic curves over  $\mathbb{F}_{2^{155}}$  (most of them)
- elliptic curves where  $n = 3, 4, 5, 6$  (all)
- hyperelliptic curves of genus 2 with  $n = 2, 3$  (all)
- hyperelliptic curves of genus  $> 2$  in odd characteristic with  $n = 2, 3$  (some)
- etc.