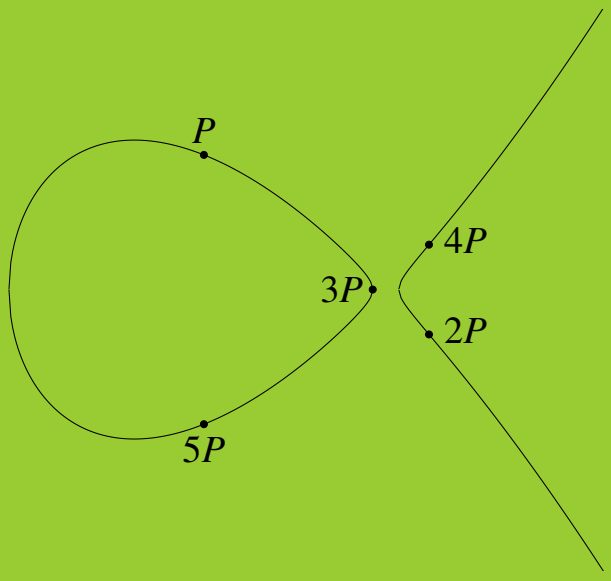
Elliptic Curves and Crypto Safety



Hans Montanus

Elliptic Curves and Crypto Safety

Hans Montanus

Preface

If somebody with some technical background is, for example, interested in the blockchain technology or a high school student, as another example, wants to write a practical assignment on cryptography, they will often face the following problem: already at the beginning of their investigation they read that it has something to do with the *multiplication of points on modular elliptic curves over finite fields*. They immediately ask themselves: what is an *elliptic curve*?, what is a *modular elliptic curve*? and what is the *multiplication of points on a modular elliptic curve*? This is already confusing, even more so since a modular elliptic curve does not look like a curve at all. Other questions which arise are: what is a *finite field* or even a *field*? A persistent student will find that a field has something to do with a *group* and that a group is something with properties like *associativity*, *commutativity* and *distributivity*. So, already after a few sentences they are drowning in concepts which are new and therefore difficult to them. At this point they may give up.

In the search for a less difficult book one might face the following problem: either one finds popular introductions with almost no mathematics or one arrives at university courses and books written by professors. The first are simple but do not satisfy the desire of the reader to understand things mathematically. The second are intended for university students. They are formal and technical, as they should. However, they are too difficult for readers with less mathematical experience in the field. A book which fills the gap should be mathematical on a very elementary level. The present book is intended to be a simple and informal introduction to the mathematics behind cryptography, cryptocurrency and blockchain technology. With simple is meant that a high school level of mathematics (together with the willingness to study) suffices to understand the contents. With informal is meant that the book is not organized as an enumeration of theorems and proofs. Instead it rather is a random walk through numbers en elliptic curves, some patterns are recognized and captured into relations. Proofs of these relations are omitted, except for a few obvious cases.

Since the contents in this book is very elementary and known for ages, citations are considered redundant. Citations were also omitted to avoid a technical and intimidating impression. However, it should be mentioned that I learned a lot from the book of Washington [1], the book of Koblitz [2] and the book of Silverman and Tate [3]. Of course I also obtained information from the internet. For this I wish to mention the following two references: An instructive explanation of the math behind the bitcoins is given by Rykwalder [4]. To understand the blockchain basics a 1blue3brown youtube video [5] was very helpful.
Together with what I already knew, I felt sufficiently equipped to write things in my own words. At every step I tried to put myself in the shoes of a layman. I also take sideways,

probably to show the reader the beauty of mathematics. The result is a somewhat unique presentation of the matter. The present book has just been written for educational purposes. It is intended for high school students with talent for mathematics and for readers with (a little more than) a high school level mathematical background.

Acknowledgement I wish to thank Ron Westdijk for his improvements and suggestions to the manuscript.

Februari 2025, Hans Montanus

ISBN 978-90-829889-5-6

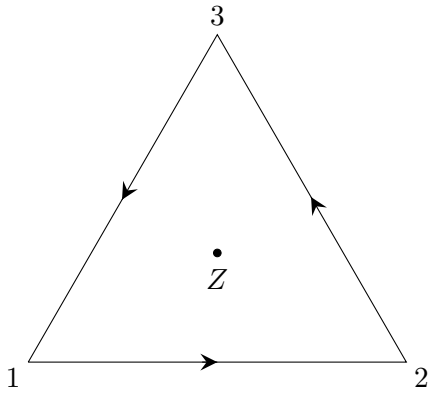
Contents

Chapter 1

Introduction to group theory

1.1 The group C_3

We start considering an equilateral triangle, see the figure. The arrows in the edges cause the triangle to have an 'orientation'.



The triangle is unaltered if it is rotated anti-clockwise over $2\pi/3$ around the barycenter Z, except that the figures at the corners have moved one position. Let us denote this rotation by r_1. The triangle also is unaltered if it is rotated anti-clockwise over $4\pi/3$. The figures at the corners then have moved two positions. This rotation is denoted as r_2. With a rotation angle of 2π both the triangle and the figures at the corners are rotated onto itself. With this full turn, which we could denote as r_3, nothing has changed. The result is the same as a rotation over 0 (no rotation at all). This is called a unit rotation (identity) and denoted as r_0 (sometimes also as e). First applying rotation r_1 and then rotation r_2 is denoted as r_2r_1. The result is the unit rotation: $r_2r_1 = r_0$. Similarly we have $r_1r_2 = r_0$, $r_1r_1 = (r_1)^2 = r_2$, $r_2r_2 = (r_2)^2 = r_1$, $r_1r_0 = r_1$, $r_0r_2 = r_2$, etc. One can also take longer sequences of rotations, for example $r_1r_1r_2$. Since $r_1r_2 = r_0$ we get $r_1(r_1r_2) = r_1r_0 = r_1$. We could also have chosen to replace r_1r_1 by r_2, then we get $(r_1r_1)r_2 = r_2r_2 = r_1$. The result does not depend on the order of the replacement: $r_1(r_1r_2) = (r_1r_1)r_2$. This property is known as *associativity*.

One can also rotate clockwise. It is the *inverse* (opposite) of rotating anti-clockwise. The inverse of r_1 is written as r_1^{-1}. Since a rotation followed by its inverse rotation is in effect no rotation at all we have $r_1^{-1}r_1 = r_0$. Since also $r_2 r_1 = r_0$ we obtain $r_1^{-1} = r_2$. Of course, we could have written the latter identity immediately just by looking at the action of r_1^{-1} and r_2 to the triangle. Similarly there holds $r_2^{-1} = r_1$ and $r_0^{-1} = r_0$.

The set $\{r_0, r_1, r_2\}$ is a GROUP because it satisfies the following demands:

1. the set contains a unit element, r_0 (in general e)
2. each element of the set has an inverse which is also an element of the set
3. associativity is satisfied, that is for each triple of elements a, b en c of the set there holds $(ab)c = a(bc)$.

A set is a group if all the three demands are satisfied. A group is called 'Abelian' if for each pair of elements a en b of the group there holds $ab = ba$. For instance, the triangle group $\{r_0, r_1, r_2\}$ is Abelian. The bookkeeping of the action of subsequent group elements is usually by means of a *Cayley table*. For the Abelian group $\{r_0, r_1, r_2\}$ it is as follows:

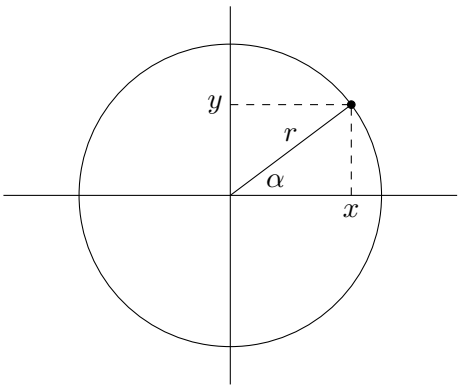
	r_0	r_1	r_2
r_0	r_0	r_1	r_2
r_1	r_1	r_2	r_0
r_2	r_2	r_0	r_1

The group $\{r_0, r_1, r_2\}$ can also be written as $\{(r_1)^0, r_1, (r_1)^2\}$. The element r_1 therefore is a *generator* of the group. The order of r_1 is 3 (it generates 3 group elements). The element r_2 also is a generator of the group $\{r_0, r_1, r_2\}$. The cyclic group $\{r_0, r_1, r_2\}$ is denoted as C_3. The number of elements in a group is the *order* of a group. In summary: the group C_3 is Abelian, it has order 3, and 1 generator (r_1 or r_2) is sufficient to generate the group.

Rotations can be described with matrices. For the coordinates (x, y) of a point on a circle with radius r and its centre at the origin we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} , \tag{1.1}$$

where α is the angle with respect to the x axis, see the next figure.



If the point (x, y) is rotated anti-clockwise over an angle θ, then the new coordinates are

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix} = \begin{pmatrix} r\cos\alpha\cos\theta - r\sin\alpha\sin\theta \\ r\sin\alpha\cos\theta + r\cos\alpha\sin\theta \end{pmatrix} \tag{1.2}$$

or

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.3}$$

The 2×2 matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{1.4}$$

is for an anti-clockwise rotation over an angle θ. The determinant of the matrices is 1. The matrices for rotations over 0, $2\pi/3$ en $4\pi/3$ are denoted as R_0, R_1 respectively R_2. Explicitly:

$$R_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R_1 = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}, \quad R_2 = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}. \tag{1.5}$$

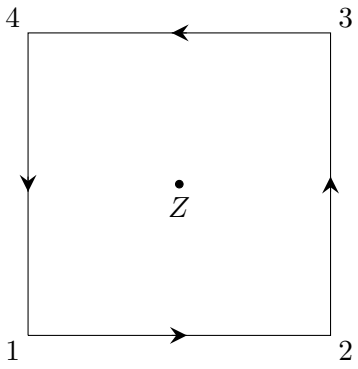
From matrix multiplication it follows $(R_1)^2 = R_2$, $(R_2)^2 = R_1$ and $R_1R_2 = R_2R_1 = R_0$. This means that the group of matrices $\{R_0, R_1, R_2\}$ is similar to $\{r_0, r_1, r_2\}$:

$$\begin{array}{l} r_0 \longleftrightarrow R_0 \\ r_1 \longleftrightarrow R_1 \\ r_2 \longleftrightarrow R_2 \end{array} \quad . \tag{1.6}$$

With this one to one relation the groups have the same group structure: the Cayley table is similar. The group $\{R_0, R_1, R_2\}$, which we will call M_3, therefore is *isomorphic* to the group C_3. The group M_3 is just another *representation* of the group C_3: the matrix representation. The isomorphism between C_3 and M_3 is expressed as $C_3 \cong M_3$.

1.2 The group C_4

Here we consider a square with arrows in the edges, see the figure.



The square has a fourfold rotational symmetry. The anti-clockwise rotations around the barycenter Z over 0, $\pi/2$, π en $3\pi/2$ are denoted as r_0, r_1, r_2 respectively r_3. The square also has point symmetry. That is, reflection in point Z leads to the same square. However, if you look what happens to the figures at the corners you will notice that the point reflection is actually the same as the rotation r_2. The group $\{r_0, r_1, r_2, r_3\}$ is Abelian with order 4. The group is generated by r_1 which has order 4. The group is denoted as C_4. The Cayley table is as follows:

	r_0	r_1	r_2	r_3
r_0	r_0	r_1	r_2	r_3
r_1	r_1	r_2	r_3	r_0
r_2	r_2	r_3	r_0	r_1
r_3	r_3	r_0	r_1	r_2

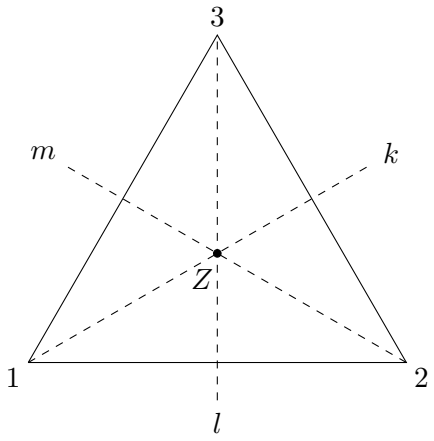
As for the triangle the rotations can be described with matrices. The matrices corresponding to a rotation over 0, $\pi/2$, π en $3\pi/2$ are denoted R_0, R_1, R_2 respectively R_3:

$$R_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad R_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad R_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (1.7)$$

The set matrices $\{R_0, R_1, R_2, R_3\}$ form a group which we will call M_4. As the reader may check $C_4 \cong M_4$.

1.3 The group D_3

Again we consider an equilateral triangle, but this time without the arrows in the edges, see the figure. The dashed lines are the medians.



The triangle has the same rotation symmetry as the triangle in section 1.1. Again, the rotations will be denoted as r_0,r_1 en r_2. Because of the absence of arrows the triangle also has mirror symmetry. For example, reflection in median k leads to the same triangle, except that the figures 2 and 3 at the corners are interchanged. This reflection will be denoted as s_0. The reflection in l and m is denoted as s_1 respectively s_2. The complete symmetry group is $\{r_0, r_1, r_2, s_0, s_1, s_2\}$ and has order 6. This so called dihedral group is denoted as D_3. The Cayley table is as follows:

	r_0	r_1	r_2	s_0	s_1	s_2
r_0	r_0	r_1	r_2	s_0	s_1	s_2
r_1	r_1	r_2	r_0	s_1	s_2	s_0
r_2	r_2	r_0	r_1	s_2	s_0	s_1
s_0	s_0	s_2	s_1	r_0	r_2	r_1
s_1	s_1	s_0	s_2	r_1	r_0	r_2
s_2	s_2	s_1	s_0	r_2	r_1	r_0

The combined actions $r_1 s_0$ (read: first reflection in k followed by a rotation over $2\pi/3$) has the same result as solely s_1, thus $r_1 s_0 = s_1$, see the table. Similarly we find $s_0 r_1 = s_2$. Since $r_1 s_0 \neq s_0 r_1$ the group D_3 is not Abelian. For the multiplication table it is not necessary to visualise all combinations of rotations and reflections. Instead one can explore the *algebra*

(rules of combined actions). For instance, from $r_1 s_0 = s_1$ it follows $r_2 r_1 s_0 = r_2 s_1$. Since $r_2 r_1 = r_0$ we find $r_2 r_1 s_0 = r_0 s_0 = s_0$ which results in $r_2 s_1 = s_0$. Convenient rules are:

$$r_i r_j = r_{i+j}\,,\ r_i s_j = s_{i+j}\,,\ s_i r_j = r_{i-j}\,,\ s_i s_j = r_{i-j}\,. \tag{1.8}$$

Since $i+j$ and $i-j$ always have to be 0, 1 or 2 one has to subtract 3 from $i+j$ if $i+j>3$ and add 3 to $i-j$ if $i-j<0$. That is, we count *modulo* 3.

Since $s_2 = s_0 r_1$ and $s_1 = s_0 r_2 = s_1 (r_1)^2$ the group D_3 is generated by 2 generators: r_1 and s_0. Explicitly: $D_3 = \{(r_1)^0, r_1, (r_1)^2, s_0, s_0 r_1, s_0 (r_1)^2\}$. The order of s_0 is 2. The group $C_2 = \{r_0, s_0\}$ (with generator s_0) is a subgroup of D_3. The group $C_3 = \{r_0, r_1, r_2\}$ (with generator r_1) is a subgroup of D_3. The order of D_3 is the product of the order of its 2 generators.

Reflections can also be expressed by matrices. The coordinates (x, y) of a point on a circle with radius r and its centre at the origin can be written as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} \tag{1.9}$$

If the point (x, y) is reflected in a line which forms an angle θ with the horizontal axis, then the new coordinates are

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} r\cos(2\theta-\alpha) \\ r\sin(2\theta-\alpha) \end{pmatrix} = \begin{pmatrix} r\cos\alpha\cos(2\theta) + r\sin\alpha\sin(2\theta) \\ r\cos\alpha\sin(2\theta) - r\sin\alpha\cos(2\theta) \end{pmatrix} \tag{1.10}$$

or

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) \\ \sin(2\theta) & \cos(2\theta) \end{pmatrix} \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.11}$$

The 2×2 matrix

$$\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \tag{1.12}$$

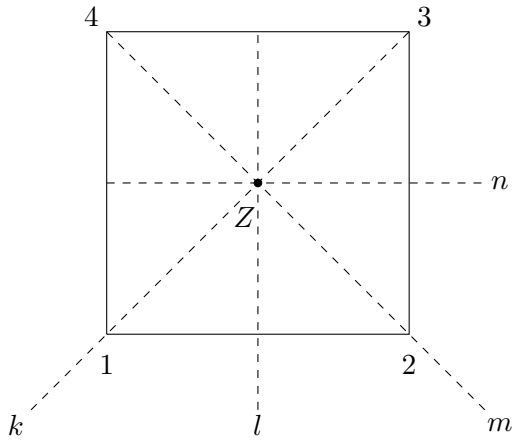
is the matrix for reflection in a line which forms an angle θ with the horizontal axis. For the lines k, l and m the angle θ is equal to 30°, 90° and 150°. The corresponding matrices, which we denote as S_0, S_1 respectively S_2, are:

$$S_0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}, \quad S_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S_2 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}, \tag{1.13}$$

With matrix multiplication it can be verified that for instance $(S_0)^2 = R_0$, $S_0 S_1 = R_1$ and $S_1 S_2 = R_2$ with R_0, R_1 and R_2 as given in the first section. The set $\{R_0, R_1, R_2, S_0, S_1, S_2\}$ is a group which we will denote as M_6. The Cayley table has the same structure as the table for D_3. thus $M_6 \cong D_3$.

1.4 The group D_4

We consider a square, but this time without the arrows in the edges, see the figure. The square has the same rotation symmetry as the square in C_3. Because of the absence of arrows the square also has mirror symmetry. The 4 lines of reflection are shown as dashed lines.



The rotations are denoted as r_0, r_1, r_2 en r_3 and the reflections in the lines k, l, m and n as s_0, s_1, s_2 respectively s_3. The complete symmetry group is $\{r_0, r_1, r_2, r_3, s_0, s_1, s_2, s_3\}$ and has order 8. This non-Abelian group is denoted as D_4. The Cayley table is as follows:

	r_0	r_1	r_2	r_3	s_0	s_1	s_2	s_3
r_0	r_0	r_1	r_2	r_3	s_0	s_1	s_2	s_3
r_1	r_1	r_2	r_3	r_0	s_1	s_2	s_3	s_0
r_2	r_2	r_3	r_0	r_1	s_2	s_3	s_0	s_1
r_3	r_3	r_0	r_1	r_2	s_3	s_0	s_1	s_2
s_0	s_0	s_3	s_2	s_1	r_0	r_3	r_2	r_1
s_1	s_1	s_0	s_3	s_2	r_1	r_0	r_3	r_2
s_2	s_2	s_1	s_0	s_3	r_2	r_1	r_0	r_3
s_3	s_3	s_2	s_1	s_0	r_3	r_2	r_1	r_0

The group D_4 is generated by 2 generators: r_1 and s_0. Explicitly:
$D_4 = \{(r_1)^0, r_1, (r_1)^2, (r_1)^3, s_0, s_0r_1, s_0(r_1)^2, s_0(r_1)^3\}$.
The rotation matrices are as in section 1.2. The matrices for reflection are:

$$S_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.14)$$

As can be verified, the group $\{R_0, R_1, R_2, R_3, S_0, S_1, S_2, S_3\}$ is isomorphic to D_4.

1.5 The group S_3

With the rotation r_1 of D_3 all corners move on one step: 1 moves to 2, 2 moves to 3 and 3 moves to 1. By means of cycles this is written as (1 2 3). For the rotation r_2 of D_3 corner 1 moves to 3, 3 moves to 2 and 2 moves to 1. This is expressed with the 3-cycle (1 3 2). The reflection s_0 of D_3 does interchange corners 2 and 3 while corner 1 is unaffected: 2 moves to 3, 3 moves to 2 and 1 'moves to' 1. This is the 3-cycle (1)(2 3). The latter is denoted more briefly with the 2-cycle (2 3), where it is understood that each missing numbers is in a 1-cycle. The unit element r_0 is in cycle notation (1)(2)(3) or shortly ( ). As can be verified, (1 2 3) = (2 3 1) = (3 1 2) and (2 3) = (3 2).

Each element of D_3 takes 1 to a, 2 to b and 3 to c, where a, b and c are 1, 2 or 3 such that $a \neq b$, $a \neq c$ and $b \neq c$. For a,b,c there are 6 possibilities: 1,2,3 and 1,3,2 and 2,1,3 and 2,3,1 and 3,1,2 and 3,2,1. Since a, b, c are a permutation of 3 different numbers, we have $3! = 6$ different permutations and thus 6 possibilities. The group S_3 is the permutation group for 3 different numbers. So, the group S_3 has order $3! = 6$. To each element of S_3 corresponds one element of D_3, see the next table.

a,b,c	element of S_3	element of D_3
1,2,3	( )	r_0
1,3,2	(23)	s_0
2,1,3	(12)	s_1
2,3,1	(123)	r_1
3,1,2	(132)	r_2
3,2,1	(13)	s_2

De cycle $(abcdef...xyz)$ has the same effect as $(ab)(bcdef...xyz)$. Indeed, $(bcdef...xyz)$ moves everything one position except that a 'moves to' a and z moves to b. Afterwards the cycle (ab) moves a to b and z to a. As a consequence the number positions are identical to the ones after $(abcdef...xyz)$. Therefore, each n-cycle ($n > 2$) can be written as a product of 2-cycles: $(abcdef...xyz) = (ab)(bc)(cd)(de)...(xy)(yz)$. Furthermore $(ab)(ab) = ( )$ since two reflections cancel each other. With these rules it follows for instance that (123) = (12)(23), or $r_1 = s_1 s_0$. Also (123)(132) = (312)(213) = (31)(12)(21)(13) = (31)(13) = ( ) or $r_1 r_2 = r_0$.

A consequence of the one-to-one correspondence between de cycles of S_3 and the elements of D_3 is that S_3 and D_3 are isomorphic: $S_3 \cong D_3$.

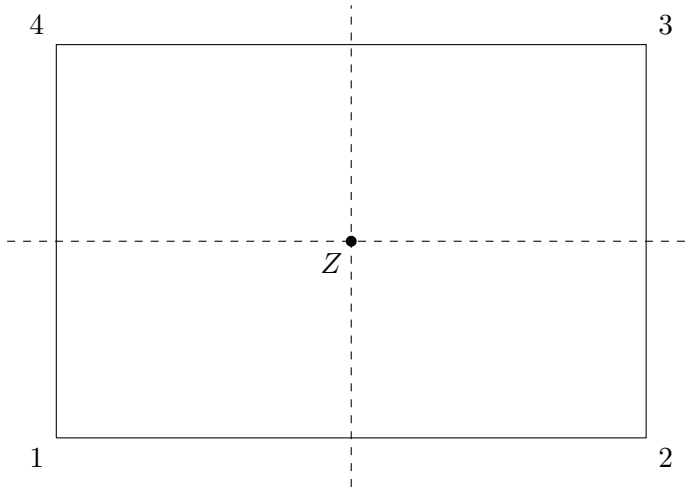
1.6 The group S_4

For $n = 4$ the permutation group has $4! = 24$ elementen, while D_4 has 8 elements. Therefore is D_4 one of the subgroups of S_4. The Cayley table for S_4 will not be shown since it is a 24 $\times$ 24 table. Instead, in the next table the 24 elements of S_4 are shown and, where applicable, the corresponding element of D_4.

a,b,c,d	element of S_4	element of D_4
1,2,3,4	( )	r_0
1,2,4,3	(34)	
1,3,2,4	(23)	
1,3,4,2	(234)	
1,4,2,3	(243)	
1,4,3,2	(24)	s_0
2,1,3,4	(12)	
2,1,4,3	(12)(34)	s_1
2,3,1,4	(123)	
2,3,4,1	(1234)	r_1
2,4,1,3	(1243)	
2,4,3,1	(124)	
3,1,2,4	(132)	
3,1,4,2	(1342)	
3,2,1,4	(13)	s_2
3,2,4,1	(134)	
3,4,1,2	(13)(24)	r_2
3,4,2,1	(1324)	
4,1,2,3	(1432)	r_3
4,1,3,2	(142)	
4,2,1,3	(143)	
4,2,3,1	(14)	
4,3,1,2	(1423)	
4,3,2,1	(14)(23)	s_3

1.7 Klein four-group V

We consider a rectangle without arrows, see the figure. The lines of reflection are dashed.



The rectangle is mapped onto itself by a rotation over 0, a rotation over π, a reflection in the horizontal axis and a reflection in the vertical axis. They are denoted as r_0, r_1, s_x and s_y respectively. The group $\{r_0, r_1, s_x, s_y\}$ has order 4 and is known as the Klein four-group, denoted as V. De Cayley table is as follows:

	r_0	s_x	s_y	r_1
r_0	r_0	s_x	s_y	r_1
s_x	s_x	r_0	r_1	s_y
s_y	s_y	r_1	r_0	s_x
r_1	r_1	s_y	s_x	r_0

We also consider the set $\{1, 3, 5, 7\}$. Multiplication is modulo 8 (that is, subtract multiples of 8 until the result is 0, 1, 2, 3, 4, 5, 6 or 7). The Cayley table is

	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

The two previous Cayley tables have the same structure. Although rotations and reflections may seem to have nothing to do with multiplications modulo 8, the Cayley tables learns that $\{r_0, r_1, s_x, s_y\}$ and $\{1, 3, 5, 7\}$ are isomorphic.

1.8 The group $\mathbb{Z}/n\mathbb{Z}$

The cyclic group $\mathbb{Z}/n\mathbb{Z}$ (also written as $\mathbb{Z}_n$) is the set $\{0, 1, ..., n-1\}$ where addition is modulo n. For example, $\mathbb{Z}/3\mathbb{Z} = \{0, 1, 2\}$ where addition is modulo 3. The Cayley table for $\mathbb{Z}/3\mathbb{Z}$ is

	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

The table has the same structure as for C_3, so $C_3 \cong \mathbb{Z}/3\mathbb{Z}$. In general $C_n \cong \mathbb{Z}/n\mathbb{Z}$ for all n.

The general linear group $GL(n, \mathbb{F})$ is a group of $n \times n$ invertible (non-zero determinant) matrices with matrix elements in $\mathbb{F}$. $\mathbb{F}$ can for instance be the complex numbers $\mathbb{C}$ or the reals $\mathbb{R}$. $\mathbb{F}$ can also be $\mathbb{Z}/n\mathbb{Z}$. The special linear group $SL(n, \mathbb{F})$ is a group of $n \times n$ matrices with determinant equal to 1 and with matrix elements in $\mathbb{F}$. Also here $\mathbb{F}$ can be $\mathbb{Z}/n\mathbb{Z}$. In words, modular counting can also be applied to matrix elements. For instance, $GL(2, \mathbb{Z}/3\mathbb{Z})$ is a group of 2×2 matrices with matrix elements in $\mathbb{Z}/3\mathbb{Z}$. Ignoring the determinant this would lead to $3^4 = 81$ possible matrices. A non-zero determinant, calculated modulo 3, reduces the number of possible matrices to 48. For the group $SL(2, \mathbb{Z}/3\mathbb{Z})$ this is further reduced to 24. As another example we consider the group $GL(2, \mathbb{Z}/2\mathbb{Z})$, which is identical to $SL(2, \mathbb{Z}/2\mathbb{Z})$. Ignoring the determinant this would lead to $2^4 = 16$ possible matrices. A non-zero determinant, calculated modulo 2, reduces the number of possible matrices to 6. These 6 different matrices are:

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \tag{1.15}$$

$$B_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{1.16}$$

The Cayley table for these matrices have the same structure as S_3: $GL(2, \mathbb{Z}/2\mathbb{Z}) \cong S_3$. Since $S_3 \cong D_3$ also $GL(2, \mathbb{Z}/2\mathbb{Z}) \cong D_3$.

1.9 Number of groups of order n

In all the tables shown an element never occurs more than once in a row (or in a column). The reason for this is as follows. Consider a group consisting of the different elements $\{a, b, c, d, ...\}$. Suppose that b followed by a has the same result as c followed by a. That would imply $ab = ac$. Since each element of a group has an inverse, we have $ab = ac \rightarrow a^{-1}ab = a^{-1}ac \rightarrow b = c$. The latter contradicts the initial assumption of b and c being different elements.

The question arises: how many groups of order n have a different, not isomorphic, Cayley table? Without specifying them we denote the elements as e, f, g, etc., where e is the unit element. For order 1 there is just 1 element: e. So, there is just 1 table possible:

	e
e	e

For order 2 we have: $\{e, f\}$. There is just 1 table possible, isomorphic to the table of $\mathbb{Z}/2\mathbb{Z}$:

	e	f
e	e	f
f	f	e

For order 3 there are 3 elements: $\{e, f, g\}$. To create the table $f^2 = e$ is not possible. The only possibility, $f^2 = g$, leads to 1 table, which is isomorphic to $\mathbb{Z}/3\mathbb{Z}$:

	e	f	g
e	e	f	g
f	f	g	e
g	g	e	f

For order 4 there are 4 elements: $\{e, f, g, h\}$. The table can be partly filled:

	e	f	g	h
e	e	f	g	h
f	f			
g	g			
h	h			

To complete the row for f we have three options. The first is $f^2 = g$. The requirement that each element occurs only once in a row or column limits the options for further filling the table to just one possibility:

	e	f	g	h
e	e	f	g	h
f	f	g	h	e
g	g	h	e	f
h	h	e	f	g

With $e \leftrightarrow 0$, $f \leftrightarrow 1$, $g \leftrightarrow 2$ and $h \leftrightarrow 3$ one sees the table is isomorphic to $\mathbb{Z}/4\mathbb{Z}$.

For the second option, $f^2 = h$, one is forced to the following table:

	e	f	g	h
e	e	f	g	h
f	f	h	e	g
g	g	e	h	f
h	h	g	f	e

That the latter table also is isomorphic to $\mathbb{Z}/4\mathbb{Z}$ can be seen by interchanging the rows and columns for g and h followed by replacing g for h and h for g. It can also be seen from the elements following cyclic from f: $f = f^1$, $h = f^2$, $g = fh = f^3$, $e = fg = f^4$.

For the third option, $f^2 = e$, it turns out we have two possibilities for further filling:

	e	f	g	h
e	e	f	g	h
f	f	e	h	g
g	g	h	f	e
h	h	g	e	f

	e	f	g	h
e	e	f	g	h
f	f	e	h	g
g	g	h	e	f
h	h	g	f	e

The left and right table are isomorphic to $\mathbb{Z}/4\mathbb{Z}$ respectively V. So, there are 2 groups of order 4: $\mathbb{Z}/4\mathbb{Z}$ and V. V is the smallest non-cyclic group. Cyclic groups are always Abelian. Non-cyclic groups are either Abelian or non-Abelian. The non-cyclic group V is Abelian while, for instance, the non-cyclic group D_3 is non-Abelian.

For order 5 one finds only one table and this table is isomorphic to $\mathbb{Z}/5\mathbb{Z}$. For order 6 one finds two tables: one isomorphic to $\mathbb{Z}/6\mathbb{Z}$ and one isomorphic to D_3. D_3 is the smallest non-Abelian group.

If the order of a group is a prime p, there is just one table and this table is isomorphic to $\mathbb{Z}/p\mathbb{Z}$.

1.10 Subgroups and classes

A subgroup is a subset of the group which satisfies the group conditions. In the Cayley tables for V and D_3 the members seem to be divided in blocks: 2×2 blocks for V and 3×3 blocks for D_3. In V is $\{e, f\}$ a subgroup of order 2. Also $\{e, g\}$ and $\{e, h\}$ are subgroups of order 2. The unit element e is a subgroup of order 1. The order of a subgroup is a divisor of the order of the group. A group whose order is a prime p can only have e as a subgroup. A group is always a subgroup of itself. We will illustrate subgroups and classes for D_3.

The group D_3 has one subgroup of order 1: r_0, three subgroups of order 2: $\{r_0, s_0\}$, $\{r_0, s_1\}$,$\{r_0, s_2\}$, one subgroup of order 3: $\{r_0, r_1, r_2\}$ and one subgroup of order 6: $\{r_0, r_1, r_2, s_0, s_1, s_2\}$. For D_3 we can calculate the result of gr_1g^{-1}, where g runs through all the elements of D_3. Thus $r_0r_1r_0^{-1}$, $r_1r_1r_1^{-1}$, $r_2r_1r_2^{-1}$, $s_0r_1s_0^{-1}$, $s_1r_1s_1^{-1}$ and $s_2r_1s_2^{-1}$. The result is either r_1 or r_2. For each g also gr_2g^{-1} is either r_1 or r_2. The set $\{r_1, r_2\}$ therefore is a *conjugacy class*. Similarly one finds that $\{s_0, s_1, s_2\}$ is a conjugacy class. The unit element, r_0, also is a conjugacy class. A subgroup consisting of complete conjugacy classes is called a *normal* subgroup. For D_3 is $\{r_0, r_1, r_2\}$ a normal subgroup. The subgroup $\{r_0, s_0\}$ is not a normal subgroup. The subgroup $\{r_0\}$ is a normal subgroup D_3; a unit element always is a normal subgroup. D_3 is a normal subgroup of itself; a group is always a normal subgroup of itself.

Chapter 2

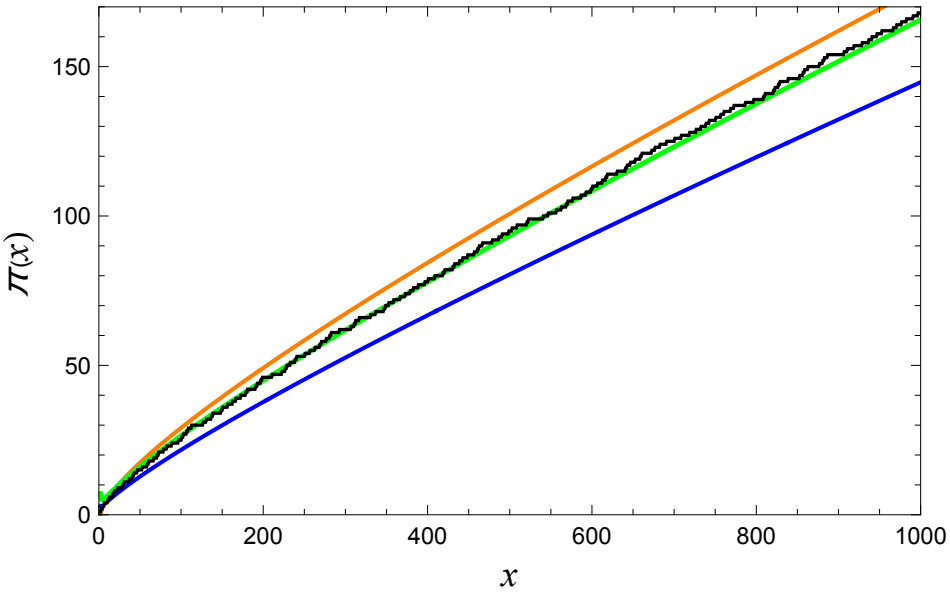
Modular Arithmetic

2.1 Some number theory

In number theory an important role is played by the prime numbers. The prime-counting function $\pi(x)$ counts the number of primes smaller than or equal to x. For instance, $\pi(11) = 5$ since there are 5 primes (2, 3, 5, 7 and 11) smaller than or equal to 11. A well known approximation for $\pi(x)$ is $\alpha(x) = \dfrac{x}{\ln x}$. A better approximation is $\mathrm{Li}(x) = \displaystyle\int_2^x \frac{1}{\ln t}\mathrm{d}t$, which requires a numerical evaluation. A convenient approximation is

$$\mu(x) = \frac{x}{\ln x}\left(1 + \frac{1}{\ln x}\right). \tag{2.1}$$

In the figure below all four functions are shown or $x \leq 1000$, $\alpha(x)$ in blue, $\mathrm{Li}(x)$ in orange, $\mu(x)$ in green and $\pi(x)$ in black.

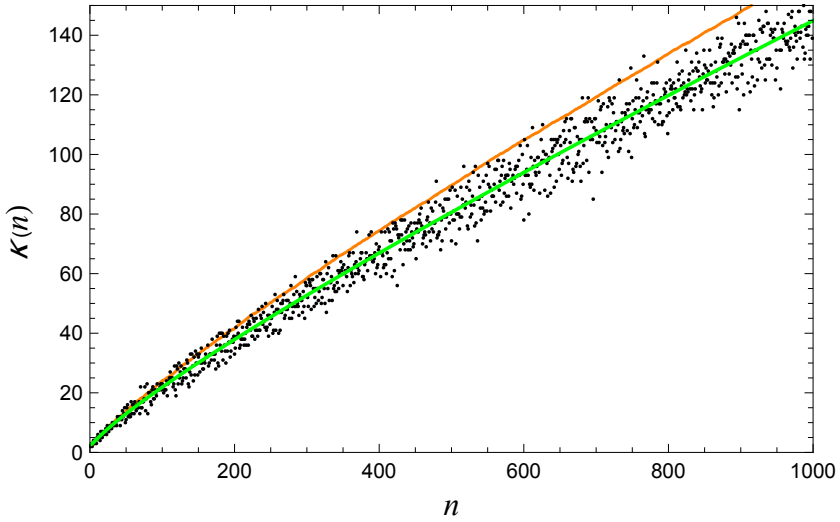


For $x > 8 \cdot 10^3$ the approximation $\mathrm{Li}(x)$ performs on average better than $\mu(x)$. For large x the performance of three approximations are tabulated:

10^n	$\alpha(x)/\pi(x)$	$\mathrm{Li}(x)/\pi(x)$	$\mu(x)/\pi(x)$
10^1	1.08574	1.28011	1.55727
10^2	0.86859	1.16324	1.05720
10^3	0.86170	1.05098	0.98644
10^4	0.88343	1.01309	0.97935
10^5	0.90553	1.00383	0.98419
10^6	0.92209	1.00164	0.98884
10^7	0.93355	1.00051	0.99147
10^8	0.94224	1.00003	0.99339
10^9	0.94901	0.99996	0.99481
10^{10}	0.95438	0.99995	0.99583
10^{11}	0.95874	0.99992	0.99659
10^{12}	0.96233	0.99993	0.99716
10^{13}	0.96535	0.99994	0.99759
10^{14}	0.96791	0.99994	0.99794
10^{15}	0.97013	0.99995	0.99821
10^{16}	0.97205	0.99992	0.99844
10^{17}	0.97374	0.99992	0.99862
10^{18}	0.97524	0.99993	0.99877
10^{19}	0.97658	0.99993	0.99890
10^{20}	0.97778	0.99994	0.99901
10^{21}	0.97886	0.99994	0.99911
10^{22}	0.97984	0.99995	0.99919
10^{23}	0.98074	0.99995	0.99926
10^{24}	0.98156	0.99995	0.99932
10^{25}	0.98231	0.99995	0.99937
10^{26}	0.98300	0.99996	0.99942

A well-known unsolved problem, one of the so called Landau's problems, is Legendre's conjecture: there always exist at least one prime between two consecutive perfect squares.
Let us denote the number of primes between two consecutive squares n^2 and $(n+1)^2$ as $\kappa(n)$. An estimate for $\kappa(n)$ is obtained as follows. Between the squares n^2 and $(n+1)^2$ there are $2n$ numbers. Half of it will be even and therefore not prime. This leaves $2n\left(1-\frac{1}{2}\right)$ odd numbers. Approximately a third of it will be a multiple of 3. This leaves $2n\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)$ possible primes. Repeating the argument for multiples of 5, 7, and so on, we obtain $\beta(n) = 2n\prod_{p_k \le n}\left(1-\frac{1}{p_k}\right)$ as an estimate for $\kappa(n)$.

From $\mu(x)$ another estimate is obtained: $\mu((n+1)^2) - \mu(n^2) \approx \ldots \approx \frac{n+1}{\ln(n+1)}$. We will denote it as $\xi(n)$, thus $\xi(n) = \frac{n+1}{\ln(n+1)}$. In the next figure we have plotted the function $\kappa(n)$ (black) and its estimates $\beta(n)$ (orange) and $\xi(n)$ (green).



The function $\beta(n)$ slightly overestimates. The function $\xi(n)$ follows accurately $\kappa(n)$, even for very large n. From $\xi(n)$ we obtain as an estimate for $\pi(n^2)$:

$$\pi(n^2) \approx \sum_{k=1}^{n-1} \xi(k) = \sum_{k=2}^{n} \frac{k}{\ln k} \tag{2.2}$$

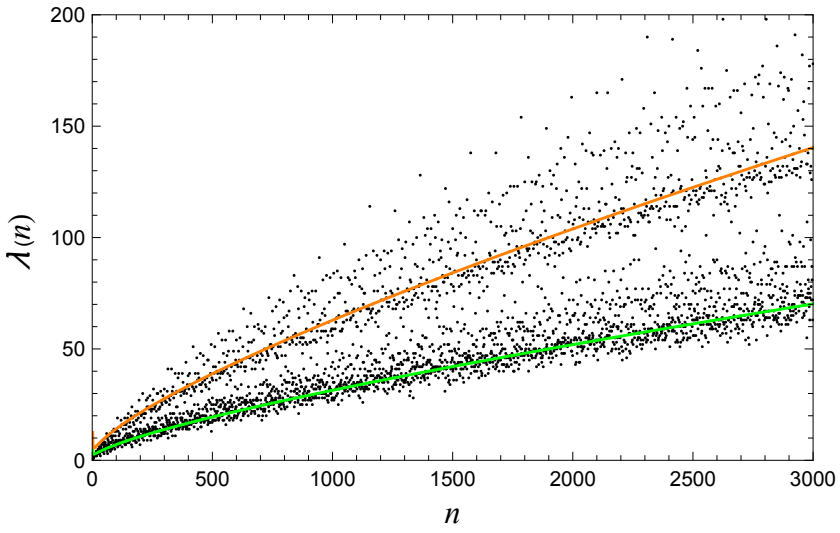
With the substitution of x for n^2 this is

$$\pi(x) \approx \sum_{k=2}^{\sqrt{x}} \frac{k}{\ln k} \approx \int_{\sqrt{2}}^{\sqrt{x}} \frac{v}{\ln v}\mathrm{d}v \approx \int_{\sqrt{2}}^{\sqrt{x}} \frac{1}{\ln v^2}\mathrm{d}v^2 = \int_2^x \frac{1}{\ln t}\mathrm{d}t. \tag{2.3}$$

This completes the circle since the latter is equal to $\mathrm{Li}(x)$. Although still not proven, the figure above suggests $\kappa(n) > 0$ for all $n > 0$. The conjecture might be stated a little stronger.

A numerical inspection suggests there always is a prime between n^2 and n^2+n and a prime between n^2+n and $(n+1)^2$, for $n>1$. If true, it implies $\kappa(n)>2$ for all $n>0$.

Another one of Landau's problems is the Goldbach conjecture: every even number larger than 2 can be written as the sum of two primes.
A lot of even numbers can be written as the sum of two primes in multiple ways. For instance, $20=3+17$ and $20=7+13$. Let us denote the number of ways an even number $2n$ can be written as a sum of primes as $\lambda(n)$. In the next figure we have plotted the function $\lambda(n)$ (black). The green and orange curves are $\dfrac{1.5n}{(\ln n)^2}$ respectively $\dfrac{3n}{(\ln n)^2}$.



The figure above clearly suggests $\lambda(n)>0$ for all $n>1$. Still, it is not proven.

Another one of Landau's problems is the twin prime conjecture: there exist infinitely many primes p such that $p+2$ is prime.
The number of twins smaller than or equal to x will be denoted as $\tau(x)$. An estimate for $\tau(x)$ is obtained as follows. From $\xi(n)$ it follows that the probability for a number x between n^2 and $(n+1)^2$ to be prime approximately is $\dfrac{1}{2n}\dfrac{n}{\ln n}=\dfrac{1}{2\ln n}$. Assuming the primes between n^2 and $(n+1)^2$ are randomly positioned the probability for a twin between n^2 and $(n+1)^2$ approximately is $\dfrac{1}{4(\ln n)^2}$. For the expected number of primes between n^2 and $(n+1)^2$ we then approximately have $2n\cdot\dfrac{1}{4(\ln n)^2}=\dfrac{n}{2(\ln n)^2}$. This leads to the following estimate:

$$\tau(x)\approx\sum_{k=2}^{\sqrt{x}}\frac{k}{2(\ln k)^2}\approx\int_{\sqrt{2}}^{\sqrt{x}}\frac{v}{2(\ln v)^2}\mathrm{d}v=\int_2^x\frac{1}{(\ln t)^2}\mathrm{d}t \tag{2.4}$$

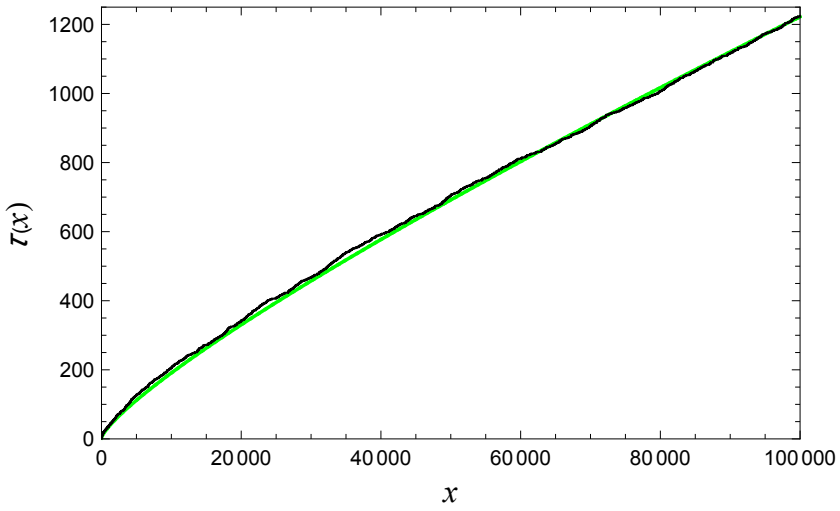
By means of partial integration we find

$$\int_2^x \frac{1}{(\ln t)^2}\mathrm{d}t = \int_2^x \frac{1}{\ln t}\mathrm{d}t - \left[\frac{t}{\ln t}\right]_2^x = \mathrm{Li}(x) - \frac{x}{\ln x} + \frac{2}{\ln 2} \tag{2.5}$$

Neglecting the $2/\ln 2$ and approximating $\mathrm{Li}(x)$ by $\mu(x)$ we obtain

$$\tau(x) \approx \frac{x}{\ln x}\left(1 + \frac{1}{\ln x}\right) - \frac{x}{\ln x} = \frac{x}{(\ln x)^2} \tag{2.6}$$

In the next figure $\tau(x)$ is plotted against x (black). The green curve is the estimate $\dfrac{1.63x}{(\ln x)^2}$ for $\tau(x)$. For increasing x a smaller value than 1.63 is required for a good approximation (ultimately to 1.32 for extremely large x).



The figure above suggests $\tau(x)$ is not limited. Still, it is not proven.

2.2 Some modular arithmetic

Modular arithmetic is a sort of cyclic counting; counting modulo a number. For instance, 49 mod 11 means: subtract from 49 a multiple of 11 such that the result is a number larger than or equal to zero and smaller than 11. For instance, $49 \mod 11 = 5$. We also say that 49 is *congruent* to 5 modulo 11 with the notation $49 \cong 5 \mod 11$.

Modular arithmetic can be very powerful. To verify that $67^{108} - 1$ is divisible by 165 we only have to check that $67^{108} \cong 1 \mod 165$. Since 165 is $3 \cdot 5 \cdot 11$ we proceed as follows:

$67 \cong 1 \mod 3 \rightarrow 67^{108} \cong 1^{108} \cong 1 \mod 3$.

$67 \cong 2 \mod 5 \rightarrow 67^4 \cong 2^4 \cong 1 \mod 5 \rightarrow 67^{108} \cong (67^4)^{27} \cong 1^{27} \cong 1 \mod 5$.

$67 \cong 1 \mod 11 \rightarrow 67^{108} \cong 1^{108} \cong 1 \mod 11$.

Now if a number is equal to 1 modulo 3, equal to 1 modulo 5 and equal to 1 modulo 11 it

must be 1 modulo the product of 3, 5 and 11 since 3, 5 and 11 have no factor in common. Hence, $67^{108} \cong 1 \mod 165$ and thus is $67^{108} - 1$ divisible by 165.

Another powerful result is Fermat's 'little theorem': if p is a prime number then
$a^p \cong a \mod p$ for any integer a.
One way to prove it is by means of induction. Firstly, $a^p \cong a \mod p$ is obviously true for $a = 0$ and for $a = 1$. Secondly, if p is a prime and $0 \leq k \leq p$ the numerator of $\binom{p}{k} = \frac{p!}{(p-k)!k!}$ contains a factor p not cancelled out by a number in the denominator. As a consequence, the identity $(a+1)^p = \sum_{k=0}^{p} \binom{p}{k} a^{p-k}$ is reduced to $(a+1)^p \cong a^p + 1 \mod p$. Then $a^p \cong a \mod p$ implies $(a+1)^p \cong (a+1) \mod p$. □

A little investigation learns that $a^5 \cong a \mod 30$ for all a. This can be understood with Fermat's little theorem. For example, modulo 3 we have $a^5 = a^3 \cdot a^2 \cong a \cdot a^2 = a^3 \cong a \mod 3$. Similarly, one finds $a^5 \cong a \mod 2$. Together with $a^5 \cong a \mod 5$ this implies $a^5 \cong a \mod 2 \cdot 3 \cdot 5$ since 2, 3 and 5 have no common factors. For each n we will search for the largest value m for which $a^n \cong a \mod m$. It is not necessary to look for values of m larger than $2^n - 2$ since they will violate $2^n \cong 2 \mod m$. So, m is a divisor of $2^n - 2$. Now if $a^p \cong a \mod p$ for some prime p than also $a^{p+k(p-1)} \cong a \mod p$. Thus in $2^n - 2$ occurs the factor 2 for $n = 2, 3, 4, ..., 2 + k, ...$, the factor 3 for $n = 3, 5, 7, ..., 3 + 2k, ...$, the factor 5 for $n = 5, 9, 13, ..., 5 + 4k, ...$,etc. That is, a prime factor p occurs in $p^n - p$ for $n \cong 1 \mod (p-1)$. In other words: a prime p is a factor of $2^n - 2$ if $p - 1$ divides $n - 1$. It quickly delivers the factors 2, 3, 5, 7 and 13 for $n = 13$. Since $2 \cdot 3 \cdot 5 \cdot 7 \cdot 13 = 2730$ we obtain $a^{13} \cong a \mod 2730$ for all a. In a similar way one finds for instance $a^{37} \cong a \mod 1\,919\,190$ or $a^{421} \cong a \mod 446\,617\,991\,732\,222\,310$. It is just a consequence of plain modular arithmetic.

For each n we denote the largest value m for which $a^n \cong a \mod m$ as $\nu(n)$, and the largest value m for which $a^{n-1} \cong 1 \mod m$ as $\eta(n)$. For the first 25 values of n the values of $\nu(n)$ and $\eta(n)$ are shown in the next table. Always is $\eta(n)$ a divisor of $\nu(n)$. The numbers $\nu(n)$ for successive n is known as the sequence A027760 of the OEIS [6].

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
$\nu(n)$	2	6	2	30	2	42	2	30	2	66	2	2730	2	6	2	510	2	798	2	330	2	138	2	2730
$\eta(n)$	1	3	1	5	1	7	1	5	1	11	1	13	1	3	1	17	1	19	1	11	1	23	1	13

The relation $a^p \cong a \mod p$ is always true if p is a prime and sometimes true when p is composite. For example, $a^{561} \cong a \mod 561$ for all a, while the number $561 = 3 \cdot 11 \cdot 17$ is composite. Such a number is a Carmichael number.

Since $p-1$ divides 560 for $p = 2$, 3, 5, 11, 17, 29, 41, 71, 113 and 281, and since $2 \cdot 3 \cdot 5 \cdot 11 \cdot 17 \cdot 29 \cdot 41 \cdot 71 \cdot 113 \cdot 281 = 15\,037\,922\,004\,270$ we obtain $a^{561} \cong a \mod 15\,037\,922\,004\,270$. Since the primes 3, 11 and 17 are factors of 15 037 922 004 270 we also have $a^{561} \cong a \mod 561$. Alternatively, n is a Carmichael number if it is a product of primes p for which $p-1$ divides $n-1$. Thus 561 is a Carmichael number because $3 \cdot 11 \cdot 17 = 561$ while 2, 10 and 16 are divisors of 560. In this way the next Carmichael numbers are easily found: $1105 = 5 \cdot 13 \cdot 17$, $1729 = 7 \cdot 13 \cdot 19$, etc. The smallest Carmichael number with 4 factors is $41041 = 7 \cdot 11 \cdot 13 \cdot 41$ and the smallest with 5 factors is $825265 = 5 \cdot 7 \cdot 17 \cdot 19 \cdot 73$, etc. For each number of factors there are infinitely many Carmichael numbers.

An equivalent form of Fermat's little theorem is:
if p is a prime number then $a^{p-1} \cong 1 \mod p$ for any integer a not divisible by p.

If $a^{p-1} \ncong 1 \mod p$ for some $a \ncong 0 \mod p$ it is certain that p is composite. However, if $a^{p-1} \cong 1 \mod p$ the number p is either prime or composite. Suppose we want to use Fermat's little theorem as a test for primality of 3281. If we try it for $a = 43$ we get $43^{3280} \cong 1 \mod 3281$. Let us try $a = 150$, then we get $150^{3280} \cong 1 \mod 3281$, still not conclusive. If we try $a = 2$ we get $2^{3280} \cong 3197 \ncong 1 \mod 3281$ and we finally know 3281 is composite: $3281 = 17 \cdot 193$. Among the values 0 through 3280 for a there are 256 values for which $a^{3280} \cong 1 \mod 3281$. For the Carmichael number 560 there even are 320 values for $a < 561$ for which $a^{560} \cong 1 \mod 561$. To know for sure that p is prime $a^{p-1} \cong 1 \mod p$ has to be tested for all numbers $a < p$. For large p this is time consuming. One can do better with Lehmer's theorem: if there exists an a such that $a^{p-1} \cong 1 \mod p$ and $a^{(p-1)/q} \ncong 1 \mod p$ for all primes q dividing $p-1$, then p is prime. Now one can stop testing as soon as an a has been found which satisfies Lehmer's theorem.

2.3 A small excursion

As a small side step we consider the value of $\rho(n) := \sum_{k=1}^{n-1} n \mod k$. For example, for $n = 41$ we have $\rho(41) = 297$, the values $n \mod k$ are shown in the next table

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	...	39	40
$41 \mod k$	0	1	2	1	1	5	6	1	5	1	8	5	2	13	11	9	7	5	3	1	20	19	18	...	2	1

If all the values of the bottom row run from 1 through 40 the sum would be $\frac{1}{2} \cdot 40 \cdot 41 = 820$. The value $\rho(41) = 297$ is a fraction 0.35357... of it. With the table for $n = 41$ at hand we can derive an estimation for $\rho(n)$ for large n. For $n = 41$ we see for $k = 21$ through 40 the

values of 41 mod k run from 1 through 20. It contributes to $\rho(41)$ with $\frac{1}{2}\cdot 20\cdot 21 = 210$. For large n this contribution to $\rho(n)$ would be approximately $\frac{1}{2}\left(\frac{n}{2}\right)^2$. For $k = 14$ through 20 the values of 41 mod k run from 1 through 13 with step size 2. For large n it would run from 1 through about $n/3$ with step size 2. If it would run from 1 through $n/3$ with step size 1 it would contribute to $\rho(n)$ approximately with $\frac{1}{2}\left(\frac{n}{3}\right)^2$. Since it runs with step size 2, the contribution is about the half of it: $\approx \frac{1}{2}\cdot\frac{1}{2}\left(\frac{n}{3}\right)^2$. For $k = 11$ through 13 the values of 41 mod k run from 2 through 8 with step size 3. Its contribution to $\rho(n)$ therefore approximately is: $\approx \frac{1}{3}\cdot\frac{1}{2}\left(\frac{n}{4}\right)^2$. Continuing the line of reasoning we obtain

$$\rho(n) \approx \frac{1}{2}n^2\left(\left(\frac{1}{2}\right)^2 + \frac{1}{2}\left(\frac{1}{3}\right)^2 + \frac{1}{3}\left(\frac{1}{4}\right)^2 + \frac{1}{4}\left(\frac{1}{5}\right)^2 + ...\right) = \frac{1}{2}n^2\sum_{k=1}^{\infty}\frac{1}{k(k+1)^2}. \qquad (2.7)$$

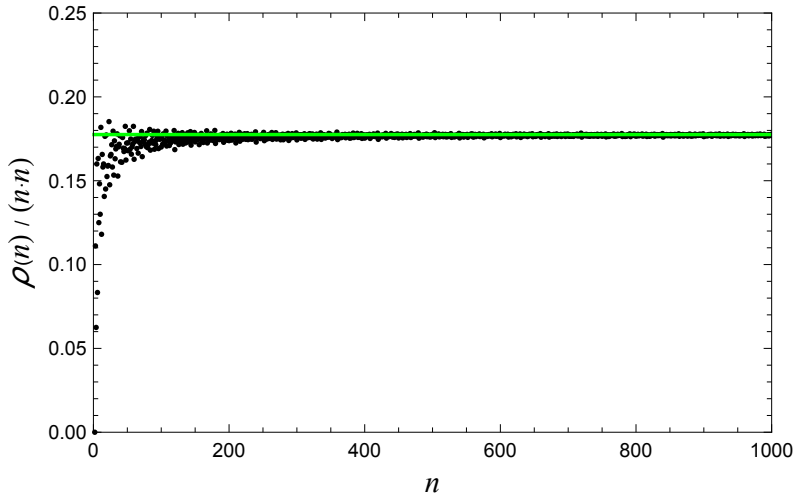
The latter can be elaborated to

$$\sum_{k=1}^{\infty}\frac{1}{k(k+1)^2} = \sum_{k=1}^{\infty}\frac{1}{k(k+1)} - \frac{1}{(k+1)^2} = \sum_{k=1}^{\infty}\frac{1}{k} - \frac{1}{k+1} - \frac{1}{(k+1)^2} = 2 - \sum_{k=1}^{\infty}\frac{1}{k^2}. \qquad (2.8)$$

The latter sum is the Riemann-zèta function $\zeta(2)$, and its value is $\pi^2/6$. Explicitly $\zeta(2) = \sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}$. As a result we therefore have

$$\lim_{n\to\infty}\frac{\rho(n)}{n^2} = \lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n-1} n \mod k \approx \frac{1}{2}\cdot\left(2 - \frac{\pi^2}{6}\right) = \left(1 - \frac{\pi^2}{12}\right). \qquad (2.9)$$

In the next diagram the ratio $\rho(n)/n^2$ is scattered against n. The limit value $1 - \pi^2/12$ is shown as a green line.



2.4 Euler's theorem

Two numbers m and n are 'relatively prime' if they have no common factors or, alternatively, if $\gcd(m,n)=1$. An important function in number theory is Euler's totient function φ. For a number n Euler's totient function counts the integers k ($1 \leq k \leq n$) which are relatively prime to n. For example, $\varphi(15)=8$ since there are 8 integers relatively prime to 15: 1, 2, 4 , 7, 8, 11, 13 and 14. Other examples: $\varphi(3)=2$ (1 and 2 are relatively prime to 3) and $\varphi(5)=4$ (1, 2, 3 and 4 are relatively prime to 5). In general $\varphi(p)=p-1$ if p is a prime. We see $\varphi(3)\cdot\varphi(5)=\varphi(15)$. In general $\varphi(m)\cdot\varphi(n)=\varphi(mn)$ if m and n are relatively prime. Another property is $\varphi(p^k)=p^k-p^{k-1}=p^k\left(1-\frac{1}{p}\right)$. Any number n can be written as a product of powers of primes (fundamental theorem of arithmetic): $n=p_1^{k_1}\cdots p_r^{k_r}$. From the latter is obtained Euler's product formula:

$$\varphi(n)=\varphi(p_1^{k_1})\varphi(p_2^{k_2})\cdots\varphi(p_r^{k_r})=n\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_r}\right). \tag{2.10}$$

Another property is: $\sum\varphi(d)=n$, where the summation is over all the divisors d of n.
Euler's theorem reads: $a^{\varphi(n)}\cong 1 \mod n$ for any integer a relatively prime to n. If n is a prime p it is reduced to $a^{(p-1)}\cong 1 \mod p$.

Writing x as the sum of y and a multiple of $\varphi(n)$ we have $a^x=a^{y+\varphi(n)k}=a^y\left(a^{\varphi(n)}\right)^k\cong a^y 1^k\cong a^y \mod n$. A consequence of Euler's theorem therefore is: if $x\cong y \mod \varphi(n)$, then $a^x\cong a^y \mod n$. If n is a prime p it is reduced to: if $x\cong y \mod (p-1)$, then $a^x\cong a^y \mod p$. The latter relation has been applied already in the second section when we searched for the largest value m for which $a^n\cong a \mod m$ for all a.
Here we will search for the smallest value m which for a given n satisfies $a^m\cong 1 \mod n$ for all a relatively prime to n. For each n these values of m is denoted as $\lambda(n)$. $\lambda(n)$ is known as the Carmichael function. For the first 28 values of n the Carmichael function and Euler's totient function are shown in the next table. See also A002322 and A000010 of the OEIS [6].

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
$\lambda(n)$	1	1	2	2	4	2	6	2	6	4	10	2	12	6	4	4	16	6	18	4	6	10	22	2	20	12	18	6
$\varphi(n)$	1	1	2	2	4	2	6	4	6	4	10	4	12	6	8	8	16	6	18	8	12	10	22	8	20	12	18	12

$\lambda(n)$ is equal to or a fraction of $\varphi(n)$. If n is a power of an odd prime or twice the power of an odd prime $\lambda(n)=\varphi(n)$. If n is 2 or 4 $\lambda(n)=\varphi(n)$. If n is a power of 2 larger than 4 $\lambda(n)=\frac{1}{2}\varphi(n)$. For other composite numbers n other fractions occur.
Carmichael's theorem reads: $a^{\lambda(n)}\cong 1 \mod n$ for any integer a relatively prime to n. If n is a prime p it is reduced to $a^{p-1}\cong 1 \mod p$.

2.5 Rings and fields

A set is a semigroup for a given operation ($+$ or $\cdot$ or whatever) if it satisfies associativity. A set is a monoid if it satisfies associativity and contains a neutral element. A set is a group if it satisfies associativity, contains a neutral element and each element has an inverse. To numbers we can apply addition and multiplication. For both they can be a group. For instance, the set of real numbers $\mathbb{R}$ is a group for addition:

1. $\mathbb{R}$ contains a neutral element, 0: $a + 0 = 0 + a = a$.
2. each element a of $(\mathbb{R}, +)$ has an inverse, $-a$: $a + (-a) = (-a) + a = 0$.
3. associativity is satisfied: $(a + b) + c = a + (b + c)$.

The set of real numbers $\mathbb{R}$ also is a group for multiplication:

1. $\mathbb{R}$ contains a neutral element, 1: $a \cdot 1 = 1 \cdot +a = a$.
2. each element a (except 0) of $(\mathbb{R}, \cdot)$ has an inverse, $1/a$: $a \cdot 1/a = 1/a \cdot a = 1$.
3. associativity is satisfied: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

The next two tables show the $\mathbb{Z}/5\mathbb{Z}$ structure for addition respectively multiplication.

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

We see $\mathbb{Z}/5\mathbb{Z}$ is a group for addition and, if we forget the 0, a group for multiplication . The situation changes for $\mathbb{Z}/6\mathbb{Z}$, see the next tables.

$+$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$\cdot$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

The set $\mathbb{Z}/6\mathbb{Z}$ is a group for addition. However, it is not a group for multiplication since 2, 3 and 4 have no inverse.

When both operations are considered together one obtains, depending on properties satisfied, rings or fields. To this end it is clarifying to enumerate properties (which should hold for every a, for every pair a, b and for every triple a, b, c) in the following order:

P1: associativity for $(+)$: $a + (b + c) = (a + b) + c$.

P2: neutral element for $(+)$: $a + 0 = 0 + a = a$.

P3: inverse for $(+)$: $a + (-a) = (-a) + a = 0$.

P4: commutative (Abelian) for $(+)$: $a + b = b + a$.

P5: associativity for $(\cdot)$: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

P6: distributivity: $a \cdot (b + c) = a \cdot b + a \cdot c$; $(a + b) \cdot c = a \cdot c + b \cdot c$.

P7: neutral element for $(\cdot)$: $a \cdot 1 = 1 \cdot a = a$.

P8: commutative (Abelian) for $(\cdot)$: $a \cdot b = b \cdot a$.

P9: no divisors of 0: if $a \cdot b = 0$ then $a = 0$ or $b = 0$.

P10: inverse (except for 0) for $(\cdot)$: $a \cdot (1/a) = (1/a) \cdot a = 1$.

Then we have the following nomenclature:

A set is a semigroup if **P1** is satisfied.

A set is a monoid if **P1** and **P2** are satisfied.

A set is a group if **P1**, **P2** and **P3** are satisfied.

A set is a commutative (Abelian) group if **P1**, **P2**, **P3** and **P4** are satisfied.

A set is a semiring, SR, if **P1**, **P2**, **P4**, **P5** and **P6** are satisfied.

A set is a ring, R, if **P1**, **P2**, **P3**, **P4**, **P5** and **P6** are satisfied.

A set is a unitary ring, UR, if **P1**, **P2**, **P3**, **P4**, **P5**, **P6** and **P7** are satisfied.

A set is a commutative unitary ring, CUR, if **P1** through **P8** are satisfied.

A set is an integral domain, ID, if **P1** through **P9** are satisfied.

A set is a field, F, if **P1** through **P10** are satisfied.

As a consequence: $F \subset ID \subset CUR \subset UR \subset R \subset SR$.

Some examples: the set of real numbers $\mathbb{R}$ satisfies **P1** through **P10** and therefore is a field. The same holds for the set of complex numbers $\mathbb{C}$. Also the set of rational numbers $\mathbb{Q}$ is a field. The set of integers $\mathbb{Z}$ is an integral domain (the inverse of for instance 3 is $\frac{1}{3} \notin \mathbb{Z}$). In general $\mathbb{Z}/n\mathbb{Z}$ is a ring. For instance, $\mathbb{Z}/6\mathbb{Z}$ is a ring. The subset $\{0, 2, 4\}$ of $\mathbb{Z}/6\mathbb{Z}$ also is a ring (with 4 as neutral element); it is a subring of the ring $\mathbb{Z}/6\mathbb{Z}$. We are more specific when we say that $\mathbb{Z}/6\mathbb{Z}$ is a CUR. Similarly, since $\{0, 2, 4\} \in \mathbb{Z}/6\mathbb{Z}$ satisfies **P9** we are more specific when we say that $\{0, 2, 4\} \in \mathbb{Z}/6\mathbb{Z}$ is an ID. $\mathbb{Z}/5\mathbb{Z}$ is a field. In general, the ring $\mathbb{Z}/p\mathbb{Z}$ is a field if p is a prime.

A 'unit' is an element of a ring which has a multiplicative inverse. For $\mathbb{R}$ every element, except 0, has an inverse; every element of $\mathbb{R}$ except 0 is a unit. $\mathbb{Z}$ has 1 and -1 as units. The 'unit group' of $\mathbb{Z}$ therefore is $\{1, -1\}$. For $\mathbb{Z}/6\mathbb{Z}$ is $\{1, 5\}$ the unit group. For $\{0, 2, 4\} \in \mathbb{Z}/6\mathbb{Z}$ is $\{4\}$ the unit group. For $\mathbb{Z}/p\mathbb{Z}$ with p a prime is every element except 0 a unit.

2.6 Polynomials

An expression of the form $\mathbb{K}[x] = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$ is a polynomial. In short $\sum_{i=0}^{n} a_i x^i$ is a polynomial (an infinite series such as a Taylor series of $\sin x$ is not a polynomial). The largest power of x, n, is the degree of the polynomial. If the largest power is 0, the polynomial is a constant: a_0. If the coefficients a_i are in a ring R, a UR, a CUR or an ID, then the polynomial also is a ring R, a UR, aCUR or a ID respectively. If the coefficients a_i are in a field F, then the polynomial is an ID; a polynomial in a field requires for the multiplicative inverse a fractional power of x which is outside the definition of a polynomial. Thus, although $\mathbb{R}$ is a field, $\mathbb{R}[x]$ is an ID.

A polynomial is reducible if it can be written as a product of factors, where a factor may not be a unit. Some examples:
The polynomial x^2+x+1 over $\mathbb{C}$ can be factored : $x^2+x+1 = (x+\frac{1}{2}+\frac{1}{2}i\sqrt{3})(x+\frac{1}{2}-\frac{1}{2}i\sqrt{3})$, while it can not be factored (is irreducible) over $\mathbb{R}$.
The polynomial x^2-x-1 over $\mathbb{R}$ can be factored : $x^2+x+1 = (x-\frac{1}{2}+\frac{1}{2}\sqrt{5})(x+\frac{1}{2}-\frac{1}{2}\sqrt{5})$, while it is irreducible over $\mathbb{Q}$.
The polynomial $x^2 + \frac{1}{6}x - \frac{1}{6}$ over $\mathbb{Q}$ can be factored: $x^2 + \frac{1}{6}x - \frac{1}{6} = (x + \frac{1}{2})(x - \frac{1}{3})$.
The polynomial $x^2 - 3x + 2$ over $\mathbb{Z}$ is reducible: $x^2 - 3x + 2 = (x - 1)(x - 2)$.
The polynomial $3x + 1$ over $\mathbb{Z}$ is irreducible, $3(x + \frac{1}{3})$ is not allowed since $\frac{1}{3} \notin \mathbb{Z}$. The polynomial $3x + 1$ also is irreducible over $\mathbb{Q}$, $3(x + \frac{1}{3})$ is not allowed since 3 is a unit of $\mathbb{Q}$.
If a polynomial is irreducible over $\mathbb{Z}$ it is irreducible over $\mathbb{Q}$. The reverse may not be true:
The polynomial $3x+3$ over $\mathbb{Z}$ is reducible: $3x+3 = 3(x+1)$, while it is irreducible over $\mathbb{Q}$; 3 is a unit (invertible) in $\mathbb{Q}$, while not a unit in $\mathbb{Z}$. The greatest common divisor of the coefficients of the latter polynomial is 3. Therefore 3 can be separated without causing a fraction in the other factor. Hence, if a polynomial is irreducible over $\mathbb{Q}$ and the greatest common divisor of the coefficients is equal to 1, then it is irreducible over $\mathbb{Z}$. A polynomial for which the greatest common divisor of the coefficients is equal to 1 is called a primitive polynomial.

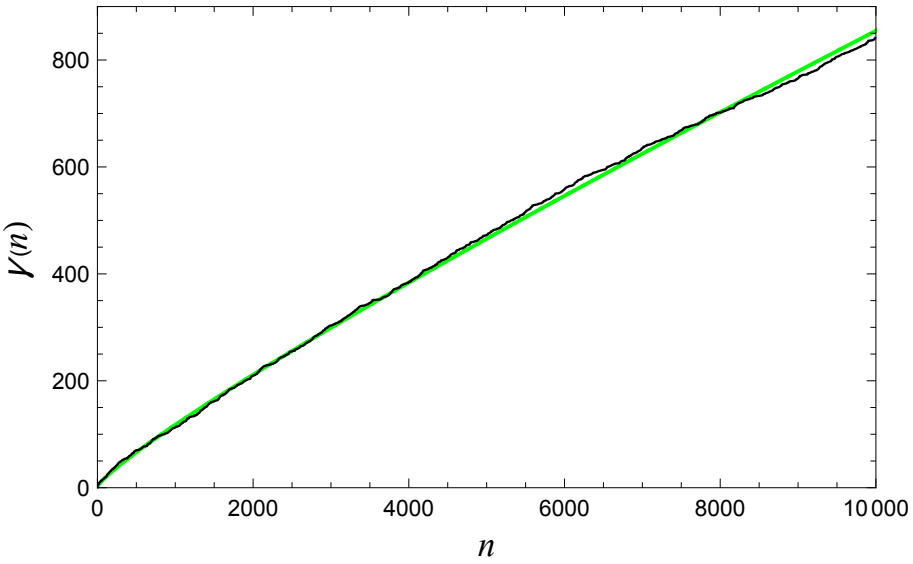
Some modular arithmetic examples: The polynomial $x^2 + x + 1$ over $\mathbb{Z}/2\mathbb{Z}$ is irreducible. Indeed, for $x = 0$ we have $0^2 + 0 + 1 \cong 1 \ncong 0 \mod 2$ and for $x = 1$ we have $1^2 + 1 + 1 \cong 1 \ncong 0 \mod 2$; there are no roots. For the polynomial $x^2 + x + 1$ over $\mathbb{Z}/3\mathbb{Z}$ we find for $x = 1$ that $1^2 + 1 + 1 \cong 0 \mod 3$. Hence, the polynomial $x^2 + x + 1$ over $\mathbb{Z}/3\mathbb{Z}$ is reducible: $(x - 1)^2 = x^2 - 2x + 1 \cong x^2 + x + 1 \mod 3$. The next value for n for which $x^2 + x + 1$ is reducible over $\mathbb{Z}/n\mathbb{Z}$ is $n = 7$: $(x - 2)(x - 4) = x^2 - 6x + 8 \cong x^2 + x + 1 \mod 7$. The list goes on for $n = 13, 19, 21, 31, ...$ The polynomial $x^2 + x + 1$ over $\mathbb{Z}/91\mathbb{Z}$ can be factored in two ways: $(x - 9)(x - 81) = x^2 - 90x + 729 \cong x^2 + x + 1 \mod 91$ and $(x - 16)(x - 74) = x^2 - 90x + 1184 \cong x^2 + x + 1 \mod 91$. Notice that 91 is not a prime

number. There are more examples for which the polynomial x^2+x+1 over $\mathbb{Z}/n\mathbb{Z}$ can be factored in multiple ways if n is not a prime. If for a prime p the polynomial x^2+x+1 over $\mathbb{Z}/p\mathbb{Z}$ is reducible, it can be factored in only one way.

As another example we consider the polynomial x^2+1 over $\mathbb{Z}/n\mathbb{Z}$. It is reducible for $n=2$: $(x-1)^2 = x^2-2x+1 \cong x^2+1 \mod 2$. Other values for n for which the polynomial x^2+1 over $\mathbb{Z}/n\mathbb{Z}$ is reducible are $5, 10, 13, 17, 25, ...$ For the composite number $n=65$ we have the first value for which the polynomial x^2+1 over $\mathbb{Z}/n\mathbb{Z}$ can be factored in two ways: $(x-8)(x-57) = x^2-65x+456 \cong x^2+1 \mod 65$ and $(x-18)(x-47) = x^2-65x+846 \cong x^2+1 \mod 65$. Again, for a prime p the polynomial x^2+1 over $\mathbb{Z}/p\mathbb{Z}$ can be factored, if it is reducible, in only one way. The reducibility of the polynomial x^2+1 over $\mathbb{Z}/p\mathbb{Z}$ for a prime p, thus $x^2+1 \cong 0 \mod p$ for some x, implies that x^2+1 is equal to p or a multiple of p for some x. This brings us to the fourth Landau problem: are there infinitely many primes of the form k^2+1 with $k \in \mathbb{N}$. Let us denote the number of such primes smaller than n^2+1 as $\gamma(n)$. An estimate for $\gamma(n)$ is obtained as follows. In the first section we saw the probability for a number between n^2 and $(n+1)^2$ to be prime approximately is $\dfrac{1}{2\ln n}$. This leads to the following estimate:

$$\gamma(n) \approx \sum_{k=2}^{n} \frac{1}{2\ln k} \approx \int_2^n \frac{1}{2\ln t}\mathrm{d}t = \frac{1}{2}\mathrm{Li}(n) \approx \frac{1}{2}\mu(n) \approx \frac{0.5n}{\ln n}\left(1+\frac{1}{\ln n}\right) \tag{2.11}$$

In the next figure we have plotted the function $\gamma(n)$ (black). The green curve is $\dfrac{0.71n}{\ln n}\left(1+\dfrac{1}{\ln n}\right)$.



The estimate suggests $\gamma(n)$ will not stop growing. Still, it is an open problem.

2.7 The Riemann zeta function

As another small excursion we consider the Taylor expansion of $\sin x/x$:

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \frac{x^8}{9!} - \frac{x^{10}}{11!}... \tag{2.12}$$

Since $\sin x$ has zero's for $x = n\pi$, $n \in \mathbb{Z}$, $(\sin x)/x$ also has zero's for $x = n\pi$ except for $n = 0$. Knowing all its roots the function $(\sin x)/x$ can be written as follows

$$\frac{\sin x}{x} = (1 - \frac{x}{\pi})(1 + \frac{x}{\pi})(1 - \frac{x}{2\pi})(1 + \frac{x}{2\pi})(1 - \frac{x}{3\pi})(1 + \frac{x}{3\pi})\cdots \tag{2.13}$$

or

$$\frac{\sin x}{x} = (1 - \frac{x^2}{\pi^2})(1 - \frac{x^2}{4\pi^2})(1 - \frac{x^2}{9\pi^2})\cdots \tag{2.14}$$

For our purpose we write the latter equation as

$$\frac{\sin x}{x} = (1 - y_1)(1 - y_2)(1 - y_3)\cdots = \prod_{k=1}^{\infty}(1 - y_k)\,, \tag{2.15}$$

where $y_k = \dfrac{x^2}{(k\pi)^2}$. Removing the brackets and grouping similar products, we obtain

$$\frac{\sin x}{x} = 1 - (y_1 + y_2 + ...) + (y_1y_2 + y_1y_3 + ... + y_2y_3 + ...) - (y_1y_2y_3 + ...), \tag{2.16}$$

which can be systematically denoted as

$$\frac{\sin x}{x} = S_0 - S_1 + S_2 - S_3 + ... = \sum_{n=0}^{\infty}(-1)^n S_n, \tag{2.17}$$

where $S_0 = 1$, $S_1 = \displaystyle\sum_{i=1}^{\infty} y_i$, $S_2 = \sum_{i=1}^{\infty}\sum_{j>i}^{\infty} y_iy_j$, $S_3 = \sum_{i=1}^{\infty}\sum_{j>i}^{\infty}\sum_{k>j}^{\infty} y_iy_jy_k$, etc.

The factors S_n can be systematically expressed as follows:

$$S_n = \frac{1}{n}\sum_{k=1}^{n}(-1)^{k+1} S_{n-k}T_k\,, \tag{2.18}$$

where $T_n = \displaystyle\sum_{k=1}^{\infty} y_k^n$. Solving for S_n we obtain

$S_0 = 1$,
$S_1 = T_1$,
$S_2 = \left(T_1^2 - T_2\right)/2!$,
$S_3 = \left(T_1^3 - 3T_1T_2 + 2T_3\right)/3!$,
$S_4 = \left(T_1^4 - 6T_1^2T_2 + 3T_2^2 + 8T_1T_3 - 6T_4\right)/4!$,
and so on.

Since $y_k = x^2/(k\pi)^2$ we have $T_n = \dfrac{x^{2n}}{\pi^{2n}} \displaystyle\sum_{k=1}^{\infty} \frac{1}{k^{2n}}$, and since $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^{2n}}$ is equal to the Riemann zeta function $\zeta(2n)$, we can write $T_n = \dfrac{x^{2n}}{\pi^{2n}}\zeta(2n)$. As a result we have

$$\frac{\sin x}{x} = 1 - \frac{x^2}{\pi^2}\zeta(2) + \frac{x^4}{2!\pi^4}\left(\zeta^2(2) - \zeta(4)\right) - \frac{x^6}{3!\pi^6}\left(\zeta^3(2) - 3\zeta(2)\zeta(4) + 2\zeta(6)\right) + \\ + \frac{x^8}{4!\pi^8}\left(\zeta^4(2) - 6\zeta^2(2)\zeta(4) + 3\zeta^2(4) + 8\zeta(2)\zeta(6) - 6\zeta(8)\right) - ...$$

Comparison with the series (2.12) gives

$$\frac{1}{3!} = \frac{1}{\pi^2}\zeta(2)$$
$$\frac{1}{5!} = \frac{1}{2!\pi^4}\left(\zeta^2(2) - \zeta(4)\right)$$
$$\frac{1}{7!} = \frac{1}{3!\pi^6}\left(\zeta^3(2) - 3\zeta(2)\zeta(4) + 2\zeta(6)\right)$$
$$\frac{1}{9!} = \frac{1}{4!\pi^8}\left(\zeta^4(2) - 6\zeta^2(2)\zeta(4) + 3\zeta^2(4) + 8\zeta(2)\zeta(6) - 6\zeta(8)\right),$$

and so on. Successively solving for $\zeta(2)$, $\zeta(4)$, $\zeta(6)$ and $\zeta(8)$, we obtain

$$\zeta(2) = \frac{\pi^2}{6}$$
$$\zeta(4) = \frac{\pi^4}{90}$$
$$\zeta(6) = \frac{\pi^6}{945}$$
$$\zeta(8) = \frac{\pi^8}{9450},$$

and so on.

We can also consider finite sums of positive powers of integers such as

$$\sum_{k=1}^{n} k = \frac{1}{2}n^2 + \frac{1}{2}n$$
$$\sum_{k=1}^{n} k^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$
$$\sum_{k=1}^{n} k^3 = \frac{1}{4}n^4 + \frac{1}{2}n^3 + \frac{1}{4}n^2$$
$$\sum_{k=1}^{n} k^4 = \frac{1}{5}n^5 + \frac{1}{2}n^4 + \frac{1}{3}n^3 - \frac{1}{30}n.$$

In general,

$$\sum_{k=1}^{n} k^m = \frac{1}{m+1} \sum_{j=0}^{m} \binom{m+1}{j} B_j \, n^{m+1-j} \,, \tag{2.19}$$

with B_j the j-th the Bernoulli number. The first Bernoulli numbers are shown in the next table.

j	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
B_j	1	$\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$	0	$\frac{1}{42}$	0	$-\frac{1}{30}$	0	$\frac{5}{66}$	0	$-\frac{691}{2730}$	0	$\frac{7}{6}$	0	$-\frac{3617}{510}$	0	$\frac{43867}{798}$

The Bernoulli numbers are related to the Riemann zeta functions. One of the relations is

$$\zeta(2n) = \frac{(-1)^{n+1} B_{2n} (2\pi)^{2n}}{2(2n)!} \quad , \qquad n \geq 0. \tag{2.20}$$

2.8 Divisor sum

The sum of the divisors of an integer n is denoted as $\sigma(n)$:

$$\sigma(n) = \sum_{d|n} d \,, \tag{2.21}$$

where $d|n$ means d is a divisor of n.
For instance 12 has 1, 2, 3, 4, 6 and 12 as divisors, so $\sigma(12) = 1 + 2 + 3 + 4 + 6 + 12 = 28$. Similarly, $\sigma(7) = 1 + 7 = 8$. If m and n have no common divisors then $\sigma(mn) = \sigma(m)\sigma(n)$. Thus $\sigma(84) = \sigma(12)\sigma(7) = 28 \cdot 8 = 224$. If p is prime then

$$\sigma(p^k) = 1 + p + p^2 + ... + p^k = (p^{k+1} - 1)/(p - 1) \,. \tag{2.22}$$

If $n = \prod_i p_i^{k_i}$ is the prime factorization of n, then

$$\sigma(n) = \prod_i \sigma(p_i^{k_i}) = \prod_i \frac{p_i^{k_i+1} - 1}{p_i - 1} \,. \tag{2.23}$$

Perfect numbers are numbers for which $\sigma(n) = 2n$. According to the Euclid-Euler theorem a number $n = 2^{p-1}(2^p - 1)$ is perfect if p is a prime and $2^p - 1$ is prime. Primes of the type $2^p - 1$ are known as Mersenne primes. It then follows that $\sigma(n) = \sigma(2^{p-1})\sigma(2^p - 1) = (2^p - 1)2^p = 2n$. The smallest perfect number is $2^1(2^2 - 1) = 6$, the second is $2^2(2^3 - 1) = 28$, the third is $2^4(2^5 - 1) = 496$, the fourth is $2^6(2^7 - 1) = 8128$, the fifth is $2^{12}(2^{13} - 1) = 33\,550\,336$. The number $2^{10}(2^{11} - 1) = 2\,096\,128$ is not perfect since $2^{11} - 1 = 2047 = 23 \cdot 89$ is not prime. A number n is multiperfect if $\sigma(n)$ is a multiple (larger than 2) of n. Examples, for which $\sigma(n) = 3n$ are 120, 672, 523 776, ... $\sigma(n) = 4n$ for 30 240, 32 760, ...

If we add $\sigma(1)$ through $\sigma(41)$ the result is 1384. In section 3 we saw the sum of 41 mod 1 through 41 mod 41 is equal to 297. It is no coincidence that $1384 + 297 = 1681 = 41^2$. In general, there holds the following identity

$$\sum_{k=1}^{n} n \mod k + \sum_{k=1}^{n} \sigma(k) = n^2. \tag{2.24}$$

As before, we denote the sum of the remainders as ρ. The identity then reads

$$\rho(n) + \sum_{k=1}^{n} \sigma(k) = n^2\,, \tag{2.25}$$

where $\rho(n) = \sum_{k=1}^{n} n \mod k$ and where $\sigma(k)$ is the sum of the divisors of k.

A proof of the relation is as follows. Since $n = 1 + ((n-1) \mod k)$ if k is not a divisor of n and $n = 0 = 1 + ((n-1) \mod k) - k$ if k is a divisor of n it follows that

$$\begin{aligned}
\sum_{k=1}^{n}(n \mod k) &= \sum_{k=1}^{n}(1 + ((n-1) \mod k))) - \sum_{k|n} k \\
&= \sum_{k=1}^{n} 1 + ((n-1) \mod n) + \sum_{k=1}^{n-1}((n-1) \mod k) - \sigma(n) \\
&= 2n - 1 + \sum_{k=1}^{n-1}((n-1) \mod k) - \sigma(n).
\end{aligned}$$

Hence

$$\rho(n) - \rho(n-1) = 2n - 1 - \sigma(n)\,. \tag{2.26}$$

A repetitive application of the latter leads to

$$\begin{aligned}
\rho(n) &= \rho(1) + \sum_{k=2}^{n}(\rho(k) - \rho(k-1)) = 0 + 2\sum_{k=2}^{n} k - \sum_{k=2}^{n} 1 - \sum_{k=2}^{n} \sigma(k) \\
&= \left(n^2 + n - 2\right) - (n-1) + \sigma(1) - \sum_{k=1}^{n} \sigma(k) = n^2 - \sum_{k=1}^{n} \sigma(k). \quad \square
\end{aligned} \tag{2.27}$$

For convenience a self explanatory scheme for $n = 9$ is given below.

$k \to$	1	2	3	4	5	6	7	8	9	sum of divisors
divisors of 1	1									$\sigma(1) = 1$
divisors of 2	1	2								$\sigma(2) = 3$
divisors of 3	1		3							$\sigma(3) = 4$
divisors of 4	1	2		4						$\sigma(4) = 7$
divisors of 5	1				5					$\sigma(5) = 6$
divisors of 6	1	2	3			6				$\sigma(6) = 12$
divisors of 7	1						7			$\sigma(7) = 8$
divisors of 8	1	2		4				8		$\sigma(8) = 15$
divisors of 9	1		3						9	$\sigma(9) = 13$
9 mod k	0	1	0	1	4	3	2	1	0	$\rho(9) = 12$
sum	9	9	9	9	9	9	9	9	9	81

From the relation $\lim_{n\to\infty} \frac{\rho(n)}{n^2} = 1 - \frac{\pi^2}{12}$ and the relation $\rho(n) + \sum_{k=1}^{n} \sigma(k) = n^2$ we obtain

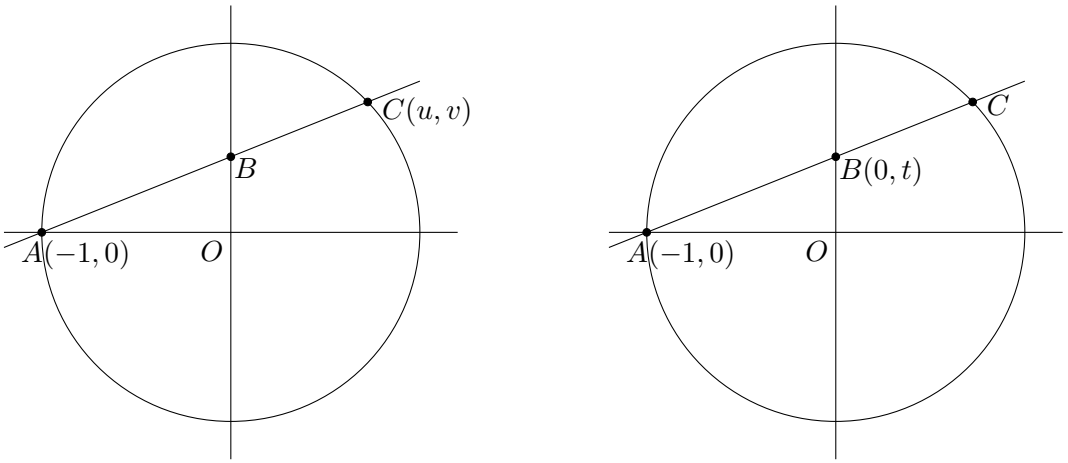
$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} \sigma(k) = \frac{\pi^2}{6} \,. \tag{2.28}$$

Chapter 3

Elliptic curves

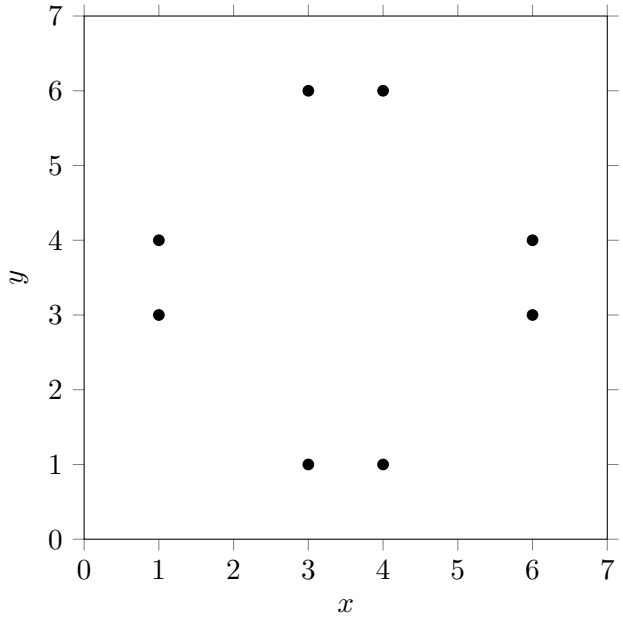
3.1 Rational points on a circle

As a start we consider rational points on a circle.



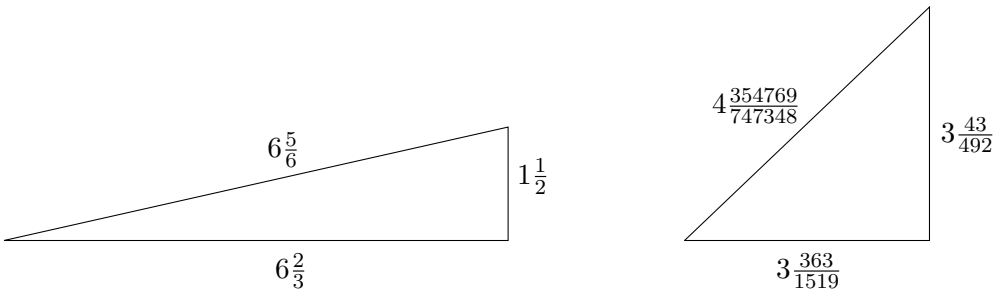
In the left side of the figure a line is drawn through $A(-1,0)$ and $C(u,v)$ where A and C are both on a unit circle. As can be calculated the line intersects the y axis in $B\left(0, \frac{v}{u+1}\right)$. This implies that the coordinates of B are rational if the coordinates u and v of C are rational. In the right side of the figure a line is drawn through $A(-1,0)$ and $B(0,t)$. The line intersects the circle at $C\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$. This implies that the coordinates of C are rational if the y coordinate t of B is rational. As a consequence, every rational point C on the circle is parameterized by a rational parameter t. If we write the rational number t as $t = \frac{m}{k}$, such that $\gcd(m,k) = 1$, then the coordinates of C read $u = \frac{k^2-m^2}{k^2+m^2}$ and $v = \frac{2mk}{k^2+m^2}$. Therefore we can find all right triangles with integer sides a, b and c (Pythagorean triples, satisfying $a^2 + b^2 = c^2$) by taking $a = k^2 - m^2$, $b = 2km$ and $c = k^2 + m^2$ and substituting integer values for k and m. The important conclusion is that there are rational points on the circle

$x^2+y^2=1$. This is not the case for the curves $x^n+y^n=1$ for $n=3,4,....$ (Fermat's theorem, proven by Wiles). There are no rational points on, for instance, the circle $x^2+y^2=3$. So, the occurrence of rational points on a circle $x^2+y^2=a$ depends on a. The possibilities are extended if we apply modular counting. For instance, $x^2+y^2 \cong 3 \mod 7$ is satisfied for $x \cong 1 \mod 7$ and $y \cong 4 \mod 7$. There are more x, y pairs satisfying $x^2+y^2 \cong 3 \mod 7$, see the next figure.

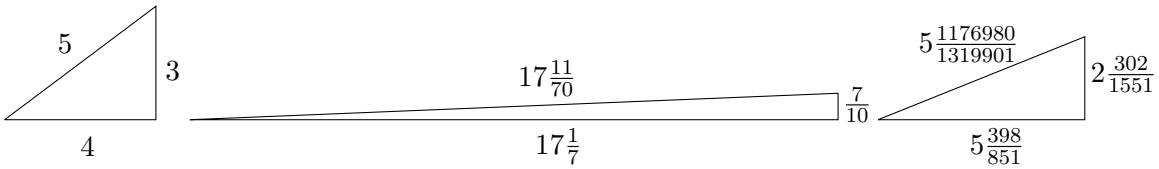


3.2 Right triangles with integer area

As a small excursion we consider right triangles with rational sides for which the area is an integer. This is always the case for Pythagorean triples. For instance, the Pythagorean $(3,4,5)$ triangle has area 6. The Pythagorean $(9,40,41)$ triangle has area 180. Since $180=5\cdot 6^2$ we can obtain a smaller integer area by dividing the sides by 6. Then the right triangle $(1\frac{1}{2}, 6\frac{2}{3}, 6\frac{5}{6})$ has area 5. Another example with area 5 is $(3\frac{43}{492}, 3\frac{363}{1519}, 4\frac{354769}{747348})$, see next figure.



For right triangles with rational sides the smallest integer area is 5. Examples of right triangles with area 6 are shown below.



The integer area n for right triangles with rational sides are known as 'integer congruent numbers'. The sequence of integer congruent numbers starts with $5, 6, 7, 13, 14, 15, 20, 21, 22, 23, 24, 28, 29, 30, 31, 34, ...$ It is sequence A003273 of the OEIS [6]. If we denote the rational sides of a right triangle as a, b and c, with c the hypothenuse, we have the Pythagorean relation $a^2 + b^2 = c^2$ and for the area n the relation $n = \frac{1}{2}ab$. Setting $x = \dfrac{nb}{c-a}$ and $y = \pm\dfrac{2n^2}{c-a}$ it follows that x and y satisfy the equation $y^2 = x^3 - n^2x$, which is an equation for an elliptic curve. If a, b, c and thus n are rational, then x and y are rational and (x, y) is a rational point on the elliptic curve.

Two $n = 5$ examples: for $(a, b, c) = (1\frac{1}{2}, 6\frac{2}{3}, 6\frac{5}{6})$ we obtain $(6\frac{1}{4}, 9\frac{3}{8})$ and $(6\frac{1}{4}, -9\frac{3}{8})$ as a rational point on the elliptic curve $y^2 = x^3 - 25x$. By changing roles of a and b we obtain for $(a, b, c) = (6\frac{2}{3}, 1\frac{1}{2}, 6\frac{5}{6})$ the rational point $(45, 300)$ and $(45, -300)$. Moreover, taking opposite sign for c leads to additional rational points: $(-4, -6)$, $(-4, 6)$, $(-\frac{5}{9}, -3\frac{19}{27})$ and $(-\frac{5}{9}, 3\frac{19}{27})$. In a similar way we find from the $(3\frac{43}{492}, 3\frac{363}{1519}, 4\frac{354769}{747348})$ right triangle the following rational points on the curve $y^2 = x^3 - 25x$: $(11\frac{97}{144}, 36\frac{71}{1728})$, $(11\frac{97}{144}, -36\frac{71}{1728})$, $(12\frac{473}{961}, 40\frac{13760}{29791})$, $(12\frac{473}{961}, -40\frac{13760}{29791})$, $(-2\frac{238}{1681}, -6\frac{42174}{68921})$, $(-2\frac{238}{1681}, 6\frac{42174}{68921})$, $(-2\frac{3}{2401}, -6\frac{56706}{117649})$ and $(-2\frac{3}{2401}, 6\frac{56706}{117649})$. These are not the only rational points on the curve $y^2 = x^3 - 25x$. Other rational points are, for instance, the zero's $(-5, 0)$, $0, 0)$ and $(5, 0)$.

Two $n = 6$ examples: from the $(3, 4, 5)$ triangle we obtain $(12, 36)$, $(12, -36)$, $(18, 72)$, $(18, -72)$, $(-3, -9)$, $(-3, 9)$, $(-2, -8)$ and $(-2, 8)$ as rational points on the curve $y^2 = x^3 - 36x$. From the $(17\frac{1}{7}, \frac{7}{10}, 17\frac{11}{70})$ triangle we obtain $(6\frac{1}{4}, 4\frac{3}{8})$, $(6\frac{1}{4}, -4\frac{3}{8})$, $(294, 5040)$, $(294, -5040)$, $(-5\frac{19}{25}, -4\frac{4}{125})$, $(-5\frac{19}{25}, 4\frac{4}{125})$, $(-\frac{6}{49}, -2\frac{34}{343})$ and $(-\frac{6}{49}, 2\frac{34}{343})$ as rational points on the curve $y^2 = x^3 - 36x$.

3.3 Elliptic curves

Third degree equations in two variables are in general given by

$$c_1y^3 + c_2y^2x + c_3yx^2 + c_4x^3 + c_5y^2 + c_6yx + c_7x^2 + c_8y + c_9x + c_{10} = 0,$$

where the coefficients c_i are elements of a field. If the equation is not singular, its curve is called an elliptic curve. For our purpose we restrict to the situation where the c_i are elements

of $\mathbb{Q}$, $\mathbb{Z}$ or $\mathbb{Z}/p\mathbb{Z}$. For these fields the elliptic curve can, by means of change of variables and coordinate transformations, be rewritten in the Weierstrass form: $y^2 = x^3 + ax + b$.

For $b = 0$ and $a = -25$ respectively $a = -36$ we obtain the elliptic curves from the previous section. They are shown, together with some of their rational points, in the next figure.

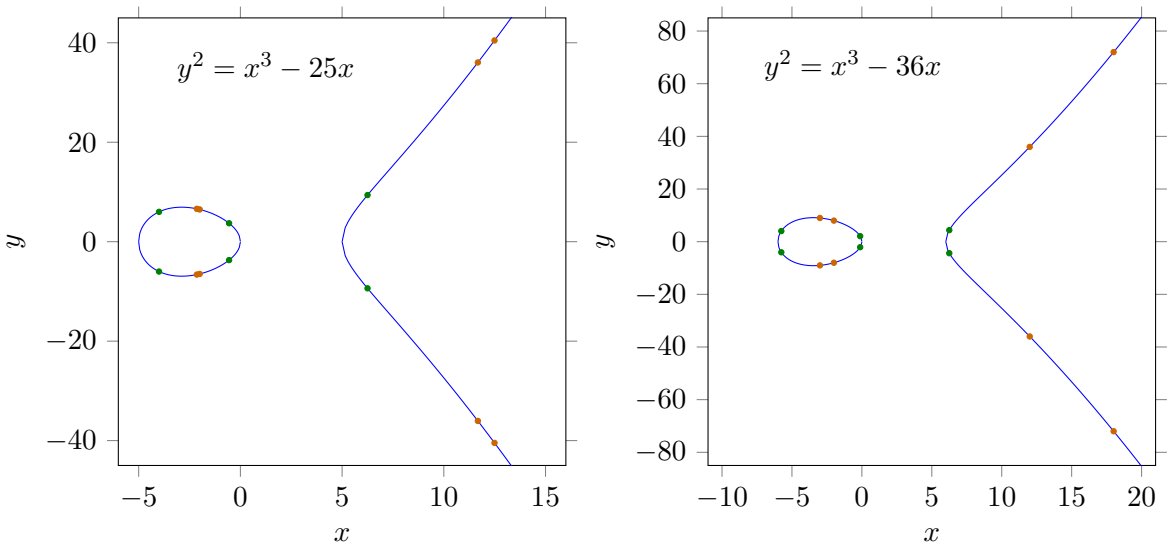
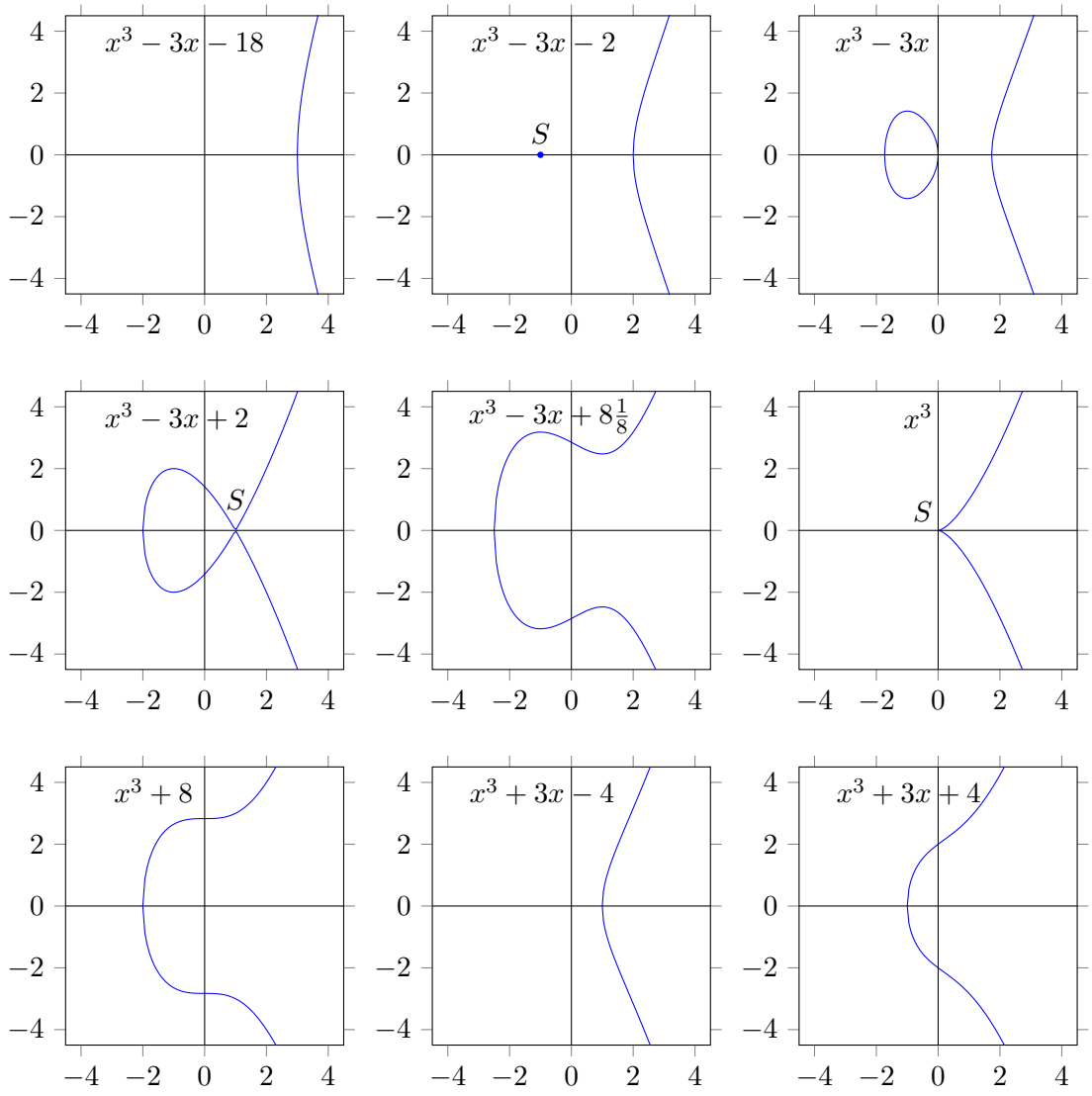


Figure 3.1: Left: the curve $y^2 = x^3 - 25x$ and some rational points found from the $(1\frac{1}{2}, 6\frac{2}{3}, 6\frac{5}{6})$ right triangle (green) and the $(3\frac{43}{492}, 3\frac{363}{1519}, 4\frac{354769}{747348})$ right triangle (orange). Right: the curve $y^2 = x^3 - 36x$ and some rational points found from the $(3, 4, 5)$ right triangle (green) and the $(17\frac{1}{7}, \frac{7}{10}, 17\frac{11}{70})$ right triangle (orange).

For the curve $y^2 = x^3 - n^2x$ the zeros are: $y = 0 \to x(x^2 - n^2) = 0 \to x = 0,\ x = -n,\ x = n$. For $-n \leq x \leq 0$ and $x \geq n$ the curve has a real value for y; outside these ranges the value of y is complex.

The shape of the elliptic curve depends on the coefficients a and b. This is illustrated in the next nine figures. For $y^2 = x^3 - 3x - 18$ there is a single real zero at $(3, 0)$ (upper left). If the value of b is increased the zero moves to the left. For $y^2 = x^3 - 3x - 2$ there are three real zero's: one at $(2, 0)$ and a twofold one at $(-1, 0)$ (upper middle). For $y^2 = x^3 - 3x$ there are three real zero's: $(-\sqrt{3}, 0)$, $(0, 0)$ and $(\sqrt{3}, 0)$ (upper right).
For $y^2 = x^3 - 3x + 2$ there are three real zero's: one at $(-2, 0)$ and a twofold one at $(1, 0)$ (middle left). For $y^2 = x^3 - 3x + 8\frac{1}{8}$ there is a single real zero at $(-2\frac{1}{2}, 0)$ (central figure). For $y^2 = x^3$ there is a threefold zero at $(0, 0)$ (middle right). For $y^2 = x^3 + 8$ there is a single zero at $(-2, 0)$ (lower left). For $y^2 = x^3 + 3x - 4$ there is a single zero at $(1, 0)$ (lower middle) and for $y^2 = x^3 + 3x + 4$ there is a single zero at $(-1, 0)$ (lower right).

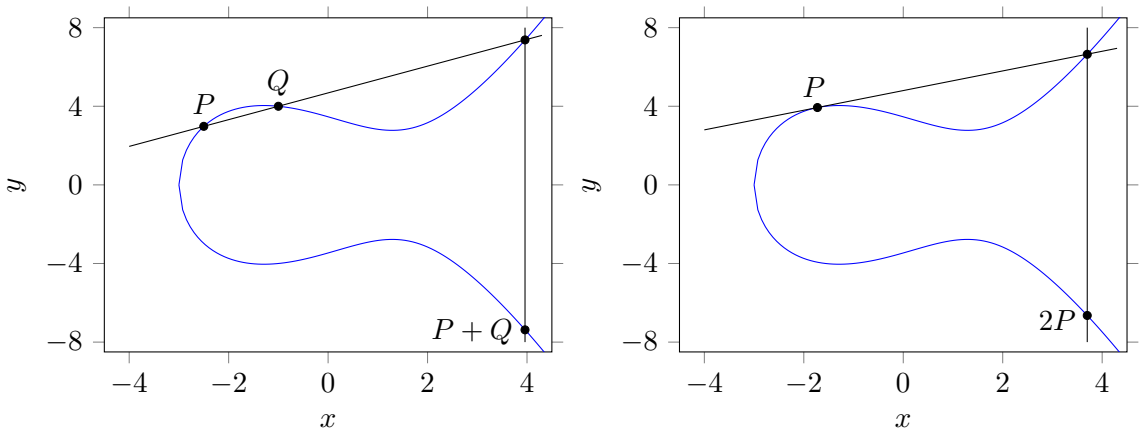
The curve for $y^2 = x^3 + ax + b$ is singular if a twofold or threefold zero is present. A singularity occurs if $4a^3 + 27b^2 = 0$. A twofold singularity is the case for $a = -3$, $b = -2$ (upper middle) and $a = -3$, $b = 2$ (middle left). A threefold singularity is the case for $a = 0$, $b = 0$ (middle right). The points of singularity are denoted as S.



3.4 Arithmetic on elliptic curves

A point P and a point Q on an elliptic curve can be composed ('added') to a point $P + Q$ as follows: draw a vertical line through the intersection point of the line through P and Q with the elliptic curve, the intersection point of the vertical line with the elliptic curve is $P + Q$. It is illustrated in the left diagram of the next figure. Now let Q approach P. In the limit that

$Q \to P$ the line through P and Q becomes the tangent line at P. This is illustrated in the right diagram of the next figure.



For the line through $P(x_P, y_P)$ and $Q(x_Q, y_Q)$ we have the equation $y = \lambda(x - x_P) + y_p$ where the slope is $\lambda = \frac{y_Q - y_P}{x_Q - x_P}$. Substituting it in the equation $y^2 = x^3 + ax + b$ for the curve, we obtain a third order equation for x:

$$x^3 - \lambda^2 x^2 + (a + 2\lambda^2 x_P - 2\lambda y_P)x + b - \lambda^2 x_P^2 + 2\lambda x_P y_P - y_P^2 = 0\,. \tag{3.1}$$

From the comparison with

$$(x - x_P)(x - x_Q)(x - x_{P+Q}) = x^3 - (x_P + x_Q + x_{P+Q})x^2 + ...x + ... = 0 \tag{3.2}$$

we see that $x_p + x_Q + x_{P+Q}$ has to be equal to λ^2. For the *addition* of P and Q we obtain:

$$x_{P+Q} = \lambda^2 - x_P - x_Q \quad , \qquad y_{P+Q} = \lambda(x_P - x_{P+Q}) - y_P \ , \ \text{with } \lambda = \frac{y_Q - y_P}{x_Q - x_P}\,. \tag{3.3}$$

For the tangent line through $P(x_P, y_P)$ we also have the equation $y = \lambda(x - x_P) + y_p$ while now the slope is the derivative in P: $\lambda = \dfrac{3x_P^2 + a}{2y_P}$. For the tangent line the point Q is equal to the point P. For the *doubling* of P we therefore obtain:

$$x_{2P} = \lambda^2 - 2x_P \quad , \qquad y_{2P} = \lambda(x_P - x_{2P}) - y_P \ ,, \ \text{with } \lambda = \frac{3x_P^2 + a}{2y_P}\,. \tag{3.4}$$

If P and Q have rational coordinates then $P + Q$ has rational coordinates and if P has rational coordinates then $2P$ has rational coordinates. So, starting with a rational point one can obtain a chain of other rational points.

For a point R on the curve where $y = 0$, a root, the tangent line is a vertical. As a consequence $2R$ is a point at infinity, denoted as $\mathcal{O}$, thus $2R = \mathcal{O}$. Since the elliptic curve is reflected in the $y = 0$ axis, to every point $P(x, y)$ corresponds a mirror point $-P(x, -y)$. In particular for a root R there holds $R = -R$, which is the same as saying that $2R = \mathcal{O}$.
Finally, since $P + Q = Q + P$ the group of points on an elliptic curve is abelian.

3.5 Torsion

For a root R of an elliptic curve we saw $2R = \mathcal{O}$. The group $\{R, \mathcal{O}\}$ is a cyclic group of order 2. Cyclic groups on elliptic curves are called *torsion* groups. Elliptic curves always have two points of inflection. The two points of inflection form, together with $\mathcal{O}$, a torsion group of order 3. Differentiating twice the equation $y^2 = x^3 + ax + b$ gives $yy'' = 3x - y'y'$. The inflection condition, $y'' = 0$, leads to $3x = (y')^2$. Substitution of $y' = (3x^2 + a)/2y$ gives $12xy^2 = (3x^2 + a)^2$. The substitution of $y^2 = x^3 + ax + b$ leads to $3x^4 + 6ax^2 + 12bx - a^2 = 0$. The four solutions of this equation are

$$x = \frac{1}{6}\sqrt{6}\left(\sqrt{-2a - \sqrt[3]{2D}} \;\pm\; \sqrt{-4a + \sqrt[3]{2D} \;\pm\; \frac{6\sqrt{6}b}{\sqrt{-2a - \sqrt[3]{2D}}}}\right), \tag{3.5}$$

where $D = -4a^3 - 27b^2$. Since the solution already looks complicated we restrict ourselves to the case where $b = 0$. Then the equation is reduced to $3x^4 + 6ax^2 - a^2 = 0$, which can be solved by hand:

$$x = \pm\sqrt{-a \pm \frac{2}{3}a\sqrt{3}}, \tag{3.6}$$

To visualize the result we take $a = 25$ and $a = -25$. For $a = 25$ the equation of the curve is $y^2 = x^3 + 25x$ and has one real root $R_0 = (0, 0)$. The equation for the point of inflection, $3x^4 + 150x^2 - 625 = 0$, has four solutions. The only solution for which both the x and y coordinate are real (in the sense of not complex) is for $x = 5\sqrt{-1 + \frac{2}{3}\sqrt{3}}$. If we denote the starting point as P then $P = \left(5\sqrt{-1 + \frac{2}{3}\sqrt{3}}, \frac{5}{3}\sqrt{30}\sqrt{\sqrt{-3 + 2\sqrt{3}}}\right)$. Numerically this is $P \approx (1.9666, 7.5346)$.

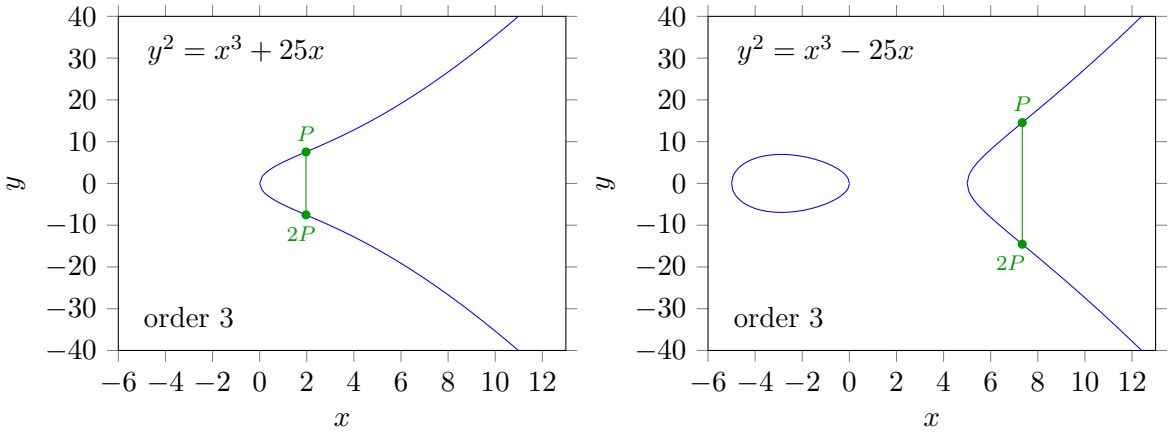
For $a = -25$ the equation of the curve is $y^2 = x^3 - 25x$ and has three real roots $R_- = (-5, 0)$, $R_0 = (0, 0)$ and $R_+ = (5, 0)$. The equation for the point of inflection, $3x^4 - 150x^2 - 625 = 0$, has four solutions. The only solution for which both the x and y coordinate are real is $P = \left(5\sqrt{1 + \frac{2}{3}\sqrt{3}}, \frac{5}{3}\sqrt{30}\sqrt{\sqrt{3 + 2\sqrt{3}}}\right)$. Numerically this is $P \approx (7.3394, 14.5558)$.

In the foregoing analysis we 'assumed' a point of inflection is part of a group of order 3. If one wants to be sure there is no other group of order 3, one can apply the arithmetic of the previous section in a straightforward manner. For $a = 25$ it goes as follows.
Start with a point P and double it to $2P$: $x_{2P} = \left(\frac{3x_P^2 + 25}{2y_P}\right)^2 - 2x_P$. If P has order 3, then $2P = -P$ and since $x_{-P} = x_P$ we obtain the condition $x_{2P} = x_P$. Hence

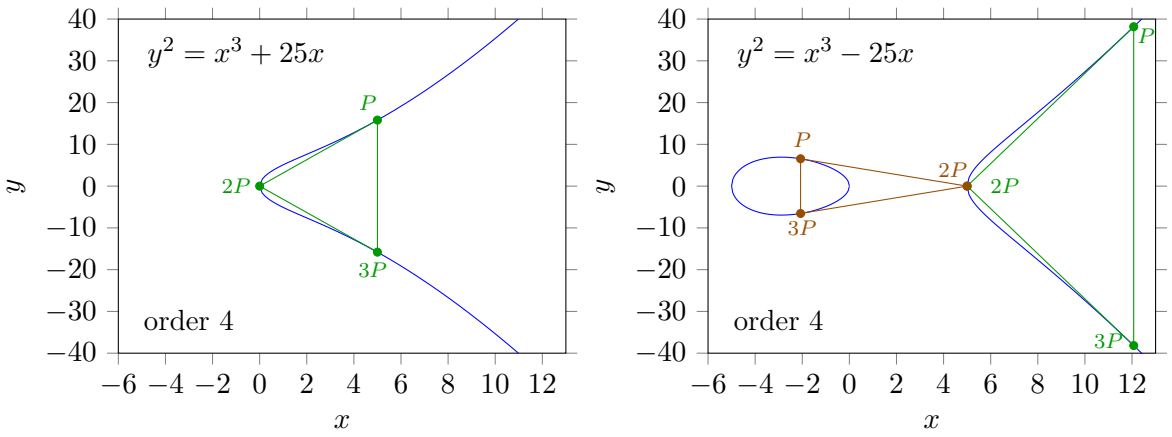
$$\left(\frac{3x_P^2 + 25}{2y_P}\right)^2 - 2x_P - x_P = 0 \quad \text{or} \qquad \left(\frac{3x_P^2 + 25}{2y_P}\right)^2 - 3x_P = 0$$

The y coordinate of P is eliminated by substitution of the equation for the elliptic curve: $y_P^2 = x_P^3 + 25x_P$. The elimination of y_P leads to the equation $3x_P^4 + 150x_P^2 - 625 = 0$ as

found above. It is a matter of inspection to identify the solution with an inflection point. For $y^2 = x^3 + 25x$ and $y^2 = x^3 - 25x$ the curves and the inflection points are shown in respectively the left and right diagram of the next figure.

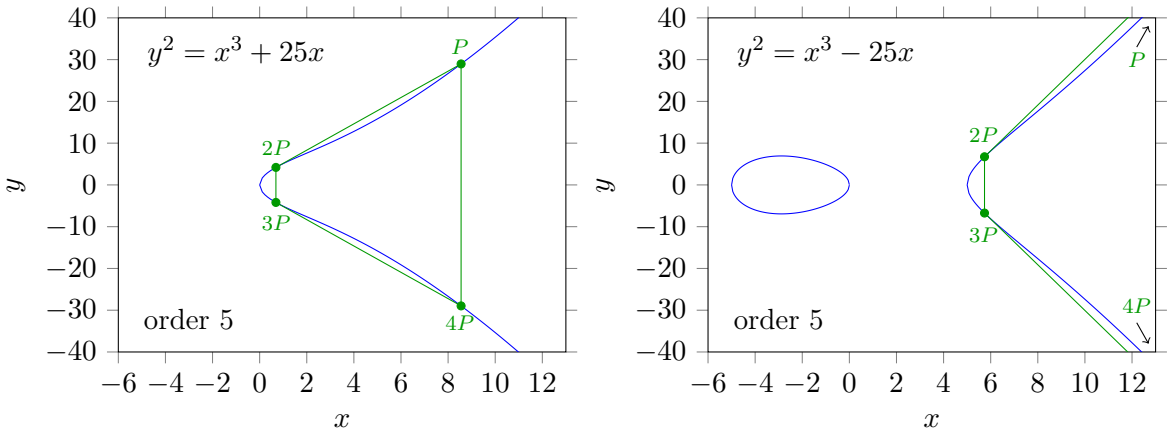


Hereafter we restrict to the situation with $a = 25$ and $a = -25$. A cyclic group of order 4 is found as follows: take a line through a root point R and tangent to the curve in a point P. This means that the slope of the tangent line has to equal the slope of the line through R and P. For $a = 25$ this means $\dfrac{3x^2 + 25}{2y} = \dfrac{y}{x}$. The latter can be elaborated to $3x^3 + 25x = 2y^2$. Substituting $y^2 = x^3 + 25x$ we obtain $x^3 - 25x = 0$, which factors in $x(x-5)(x+5) = 0$. From the three solutions only $x = 5$ is a valid x coordinate for P. The corresponding y coordinate is $5\sqrt{10}$. For $a = -25$ a similar analysis leads to the condition $x^3 - 15x^2 + 25x + 125 = 0$. There are two solutions. The first, which is on the 'egg' of the curve, is $\left(5(1-\sqrt{2}), 5\sqrt{5}(2-\sqrt{2})\right)$. The second, which is on the rounded cusp, is $\left(5(1+\sqrt{2}), -5\sqrt{5}(2+\sqrt{2})\right)$. Again, the solution could also have been obtained by equating x_{3P} with x_P and y_{3P} with $-y_P$ and determine the geometrical structure afterwards. The results are shown in the next figure.

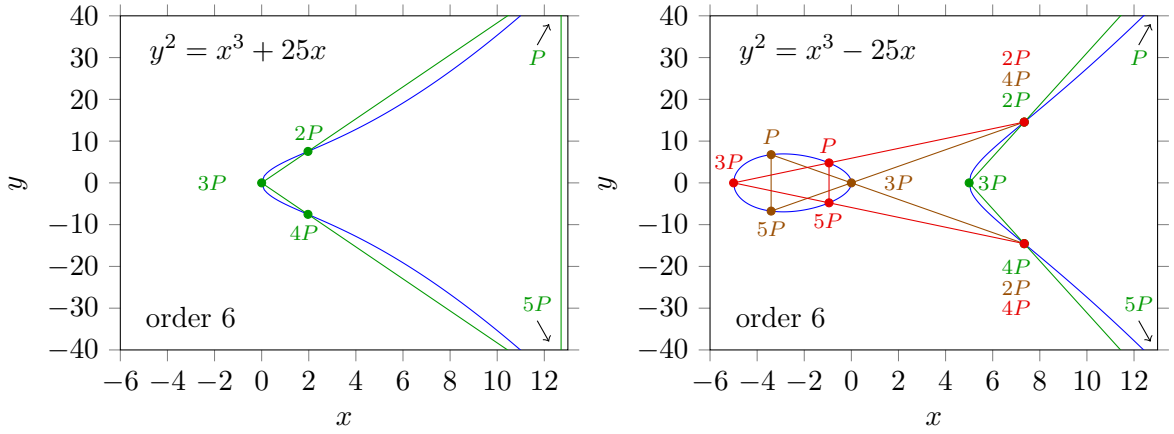


Notice that we could have started as well with $3P$. We could not have started with R_+ since $2R_+ = \mathcal{O}$. That is, $\{R_+, \mathcal{O}\}$ is a subgroup of $\{P, 2P, 3P, \mathcal{O}\}$, in the same way as C_2 is a subgroup of C_4.

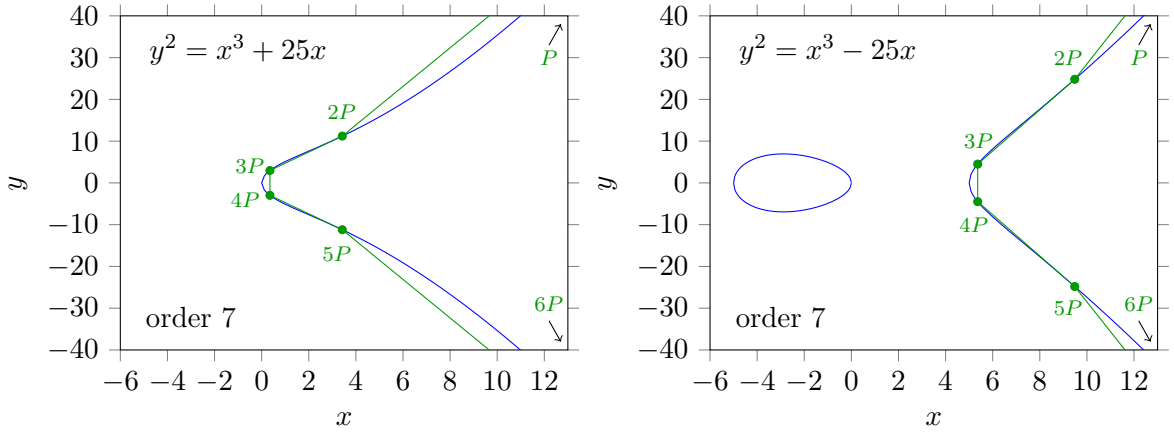
For $a = 25$ fivefold torsion leads to the equation $x^8 + 300x^6 - 16250x^4 - 812500x^2 + 390625 = 0$ with the solution
$x_P = 5\sqrt{-3 + 2\sqrt{5} - 2\sqrt{5 - 2\sqrt{5}}}$ and $y_P = 5\sqrt{10}\sqrt{\sqrt{-133 + 62\sqrt{5} + 6\sqrt{1025 - 458\sqrt{5}}}}$.
Numerically it is $P \approx (8.55164, 28.9685)$.
For $a = -25$ the equation is $x^8 - 300x^6 - 16250x^4 + 812500x^2 + 390625 = 0$ with the solution
$x_P = 5\sqrt{3 + 2\sqrt{5} + 2\sqrt{5 + 2\sqrt{5}}}$ and $y_P = 5\sqrt{10}\sqrt{\sqrt{133 + 62\sqrt{5} + 6\sqrt{1025 + 458\sqrt{5}}}}$.
Numerically it is $P \approx (18.4577, 76.334)$. Since 5 is a prime, one can start any of the four points.



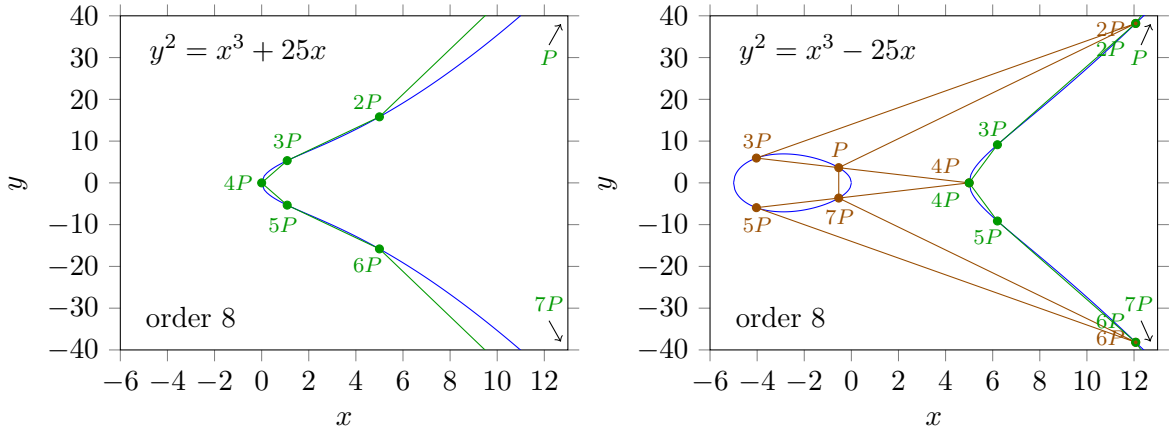
Sixfold torsion: for $a = 25$ we obtain the equation $x^4 - 150x^2 - 1875 = 0$ with the solution $x_P = 5\sqrt{3 + 2\sqrt{3}}$ and $y_P = 5\sqrt{10}\sqrt{\sqrt{45 + 26\sqrt{3}}}$. Numerically this is $P \approx (12.7123, 48.705)$. The points $2P$ and $4P$ are points of inflection. For $a = -25$ we obtain three equations. For the first equation, $x^4 - 20x^3 - 150x^2 - 500x + 625 = 0$, the solution is $x_P = 5\left(1 + \sqrt{3} + \sqrt{3 + 2\sqrt{3}}\right)$ and $y_P = 5\sqrt{10}\sqrt{18 + 10\sqrt{3} + \sqrt{627 + 362\sqrt{3}}}$. Numerically this is $P \approx (26.3726, 132.978)$. The group is shown in green in the right diagram of the next figure. For the second equation, $x^4 + 150x^2 - 1875 = 0$, the solution is $x_P = -5\sqrt{-3 + 2\sqrt{3}}$ and $y_P = 5\sqrt{10}\sqrt{\sqrt{-45 + 26\sqrt{3}}}$. Numerically this is $P \approx (-3.40625, 6.75538)$. The group is shown in brown. For the third equation, $x^4 + 20x^3 - 150x^2 + 500x + 625 = 0$, the solution is $x_P = -5\left(1 + \sqrt{3} - \sqrt{3 + 2\sqrt{3}}\right)$ and $y_P = 5\sqrt{10}\sqrt{-18 - 10\sqrt{3} + \sqrt{627 + 362\sqrt{3}}}$. Numerically this is $P \approx (-0.947955, 4.77986)$. The group is shown in red, see next figure. In all three cases the points $2P$ and $4P$ are points of inflection. Notice that the group $\{3P, \mathcal{O}\}$ and the group $\{2P, 4P, \mathcal{O}\}$ are subgroups of each sixfold torsion group (in the same way as C_2 and C_3 are subgroups of C_6).



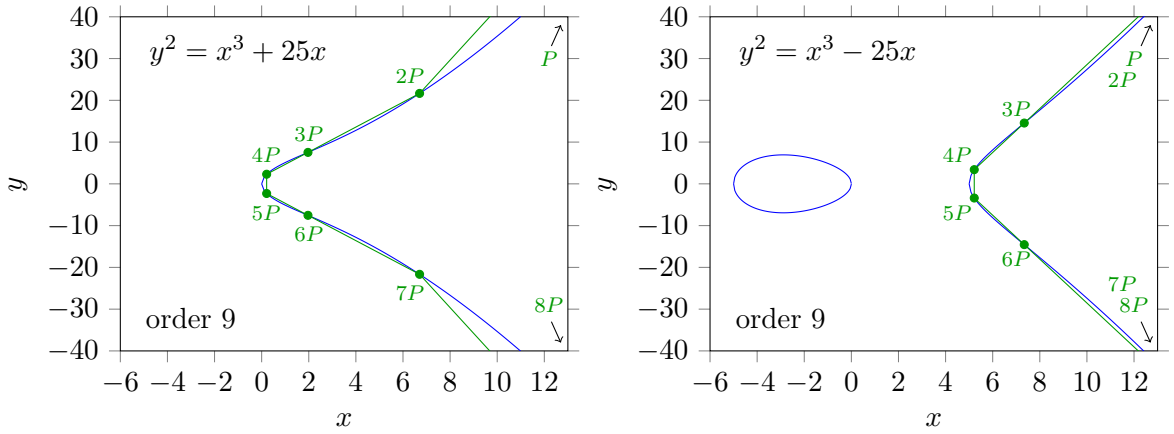
Sevenfold torsion: we obtain the equation $7X^{12} \pm 308X^{11} - 2954X^{10} \mp 19852X^9 - 35321X^8 \mp 82264X^7 - 111916X^6 \mp 42168X^5 + 15673X^4 \pm 14756X^3 + 1302X^2 \pm 196X^2 - 1 = 0$, where $X = \left(\frac{1}{5}x\right)^2$. The upper and lower part of the plus-minus symbols is for $a = 25$ and $a = -25$ respectively. For $a = 25$ the starting point is $P \approx (17.5386, 76.3763)$. For $a = -25$ the starting point is $P \approx (35.7759, 211.886)$. The results are shown in the next figure.



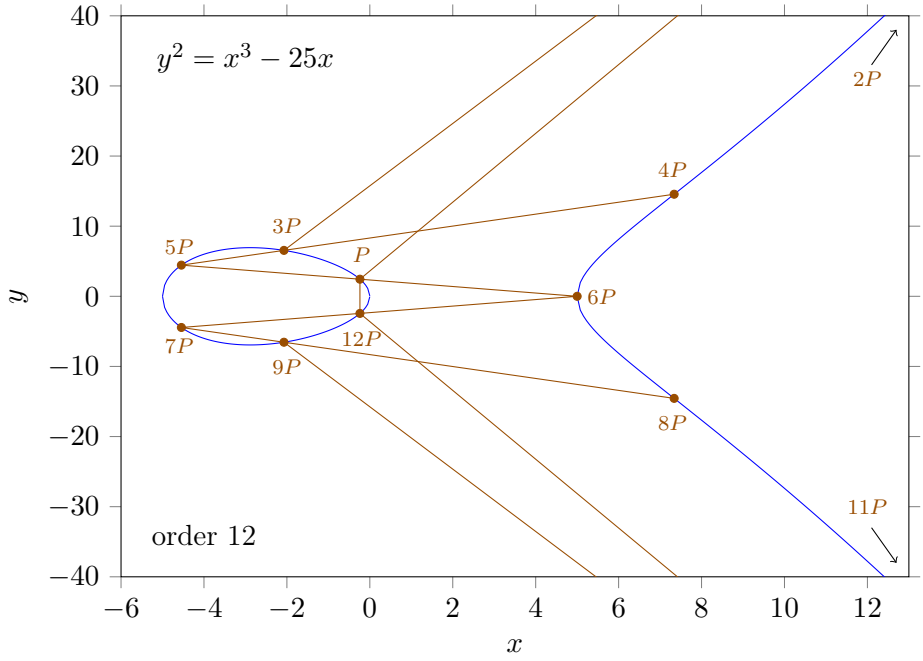
Eightfold torsion: for $a = 25$ the equation is $x^4 - 20x^3 - 50x^2 - 500x + 625 = 0$.
A solution is $x_P = 5\left(1 + \sqrt{2} + \sqrt{2 + 2\sqrt{2}}\right)$ and $y_P = 5\sqrt{10}\sqrt{13 + 9\sqrt{2} + 2\sqrt{82 + 58\sqrt{2}}}$. Numerically this is $(23.0579, 113.294)$. The result is drawn in the left diagram of the next figure. For $a = -25$ the equation is $x^8 - 40x^7 - 300x^6 - 1000x^5 + 23750x^4 + 25000x^3 - 187500x^2 + 625000x + 390625 = 0$. There are two solutions. The first starts with $P = (46.6517, 316.805)$ and is shown in green. The second starts with $P = (-0.535886, 3.63913)$ and is shown in brown. All solutions have $\{4P, \mathcal{O}\}$ as an order 2 subgroup and $\{2P, 4P, 6P, \mathcal{O}\}$ as an order 4 subgroup (as C_2 and C_4 are subgroups of C_8). If you follow the tangent lines you will see interesting geometrical properties which are not a priori obvious.



Ninefold torsion: we obtain the equation $-3814697265625 \pm 407409667968750X$ $+1666717529296875X^2 \pm 1046337890625000X^3 + 1406667480468750X^4 \pm 954676757812500X^5$ $+ 335182324218750X^6 \pm 108992109375000X^7 + 38843736328125X^8 \pm 8677938281250X^9$ $+777339140625X^{10} \mp 27707625000X^{11} - 13215656250X^{12} \mp 1252867500X^{13} - 170842500X^{14}$ $\mp 15174000X^{15} - 284625X^{16} \pm 1710X^{17} + 3X^{18} = 0$, where $X = \left(\frac{1}{5}x\right)^2$. The upper and lower part of the plus-minus symbols is for $a = 25$ and $a = -25$ respectively. For $a = 25$ the starting point is $P \approx (29.2843, 160.765)$. For $a = -25$ the starting point is $P \approx (58.9924, 451.47)$. The solutions have $\{3P, 6P, \mathcal{O}\}$ as an order 3 subgroup (as C_3 is a subgroup of C_9). The results are shown in the next figure.



The geometry of the green solutions becomes evident by now. Just to illustrate a more interesting geometry a torsion group of order 12 for $y^2 = x^3 - 25x$ is shown in the next figure. In this illustration as well as in the illustrations shown above there is no torsion point with both the x and the y coordinate rational. For rational points we should consider other elliptic curves.



3.6 Torsion lines

If we take the situation for sixfold torsion at hand we have six points: $P, 2P, 3P, 4P, 5P, \mathcal{O}$. For $y^2 = x^3 + 25x$ the line from P to $4P$ was tangent in P. Since P is twofold for its tangent we can denote the line as $(1, 1, 4)$. That is, it hits twice P and once $4P$. The line which intersects $3P$, $2P$ and P is $(3, 2, 1)$. The line tangent to the root hits twice the root $3P$ and once $\mathcal{O}$: $(3, 3, 0)$. The vertical line through the inflection points is $(2, 4, 0)$. The line intersecting $5P$, $4P$ and $3P$ is $(5, 4, 3)$. The tangent line in the point of inflection is threefold in $2P$: $(2, 2, 2)$. The line tangent to $5P$ is $(5, 5, 2)$. The line tangent to $5P$ can also be regarded as starting in $5P$, going to $2P$ and returning in $5P$: $(5, 2, 5)$. Alternatively, the order of the numbers do not matter. If we add the three numbers identifying a line we either obtain 0, 6 or 12. That is, for line (a, b, c), $c \cong (12 - a - b) \mod 6$. For n-fold torsion a line is $(a, b, (2n - a - b) \mod n)$. This can be seen as follows. If we add the points aP and bP we obtain the point $(a+b)P$. So, the line through aP and bP goes through $-(a+b)P = (-a-b)P$. Since the points are $\mod n$ we get for the line: $(a, b, (-a - b) \mod n)$, which is identical to $(a, b, (2n - a - b) \mod n)$. We can now generate all the lines: run a from 0 to $n - 1$ and b from 0 to $n - 1$ and calculate $c = 2n - a - b \mod n$. This leads to n^2 lines. However, the line (a, b, c) is the same line as (a, c, b). If a, b and c all three differ from each other we have 6 combinations for the same line. If two out of a, b and c are equal we have 3 combinations for the same line. This reduces the number of lines. Before we successively consider the situation for increasing order, we first will distinguish lines by their nature. The line (a, b, c) with $a \neq 0$, $b \neq 0$ and $c \neq 0$ all three different from each other is a line intersecting the elliptic curve in three different points. We

will denote it as type S. The line $(0, b, c)$ with $b \neq 0$ and $c \neq 0$ different from each other is a vertical line intersecting the elliptic curve in b, c and $\mathcal{O}$. We will denote it as type V. The line (a, b, b) with $a \neq 0$ and b different from a is a line tangent in b and intersecting the elliptic curve in a. We will denote it as type T. The line $(0, b, b)$ with $b \neq 0$ is a vertical line tangent in a root b. We will denote it as type R. The line (a, a, a) with $a \neq 0$ is a line tangent in a point of inflection a. We will denote it as type I. The line $(0, 0, 0)$ is a line through $\mathcal{O}$. We will denote it as type O. For order n the number of lines will be denoted as $L(n)$. The number of lines of type O, I, R, T, V, S will be denoted as $L_O(n), L_I(n), L_R(n), L_T(n), L_V(n), L_S(n)$. Their sum will be denoted as $L(n)$.

The results are tabulated.

type	$n = 1$	$L_{\text{type}}(1)$
O	(0,0,0)	1
I		0
R		0
T		0
V		0
S		0
sum		1

type	$n = 2$	$L_{\text{type}}(2)$
O	(0,0,0)	1
I		0
R	(0,1,1)	1
T		0
V		0
S		0
sum		2

type	$n = 3$	$L_{\text{type}}(3)$
O	(0,0,0)	1
I	(1,1,1) (2,2,2)	2
R		0
T		0
V	(0,1,2)	1
S		0
sum		4

type	$n = 4$	$L_{\text{type}}(4)$
O	(0,0,0)	1
I		0
R	(0,2,2)	1
T	(1,1,2) (2,3,3)	2
V	(0,1,3)	1
S		0
sum		5

type	$n = 5$	$L_{\text{type}}(5)$
O	(0,0,0)	1
I		0
R		0
T	(1,1,3) (3,3,4) (1,2,2) (2,4,4)	4
V	(0,1,4) (0,2,3)	2
S		0
sum		7

For instance, from the table for $n = 4$ we read of that 2 is root, that a line tangent in 1 intersects de curve in the root 2, that a line tangent in 3 intersects de curve in the root 2, that there is a vertical line through 1 and 3 and that there is a vertical line tangent to 2. This determines the geometry. In case of 3 roots one still has to find out for which root this is possible. One also has to find out if there is more than one possibility. Nevertheless, the tables can be of help for the understanding of the geometry of all the lines involved, in particular for increasing n.

type	$n = 6$	$L_{\text{type}}(6)$
O	(0,0,0)	1
I	(2,2,2) (4,4,4)	2
R	(0,3,3)	1
T	(1,1,4) (2,5,5)	2
V	(0,1,5) (0,2,4)	2
S	(1,2,3) (3,4,5)	2
sum		10

type	$n = 7$	$L_{\text{type}}(7)$
O	(0,0,0)	1
I		0
R		0
T	(1,1,5) (2,2,3) (4,4,6) (1,3,3) (2,6,6) (4,5,5)	6
V	(0,1,6) (0,2,5) (0,3,4)	3
S	(1,2,4) (3,5,6)	2
sum		12

type	$n = 8$	$L_{\text{type}}(8)$
O	(0,0,0)	1
I		0
R	(0,4,4)	1
T	(1,1,6) (2,2,4) (5,5,6) (2,3,3) (2,7,7) (4,6,6)	6
V	(0,1,7) (0,2,6) (0,3,5)	3
S	(1,2,5) (1,3,4) (3,6,7) (4,5,7)	4
sum		15

type	$n = 9$	$L_{\text{type}}(9)$
O	(0,0,0)	1
I	(3,3,3) (6,6,6)	2
R		0
T	(1,1,7) (2,2,5) (5,5,8) (1,4,4) (2,8,8) (4,7,7)	6
V	(0,1,8) (0,2,7) (0,3,6) (0,4,5)	4
S	(1,2,6) (1,3,5) (2,3,4) (3,7,8) (4,6,8) (5,6,7)	6
sum		19

type	$n = 10$	$L_{\text{type}}(10)$
O	(0,0,0)	1
I		0
R	(0,5,5)	1
T	(1,1,8) (2,2,6) (3,3,4) (6,6,8) (2,4,4) (2,9,9) (4,8,8) (6,7,7)	8
V	(0,1,9) (0,2,8) (0,3,7) (0,4,6)	4
S	(1,2,7) (1,3,6) (1,4,5) (2,3,5) (3,8,9) (4,7,9) (5,6,9) (5,7,8)	8
sum		22

type	$n = 11$	$L_{\text{type}}(11)$
O	(0,0,0)	1
I		0
R		0
T	(1,1,9) (2,2,7) (3,3,5) (6,6,10) (7,7,8) (1,5,5) (2,10,10) (3,4,4) (4,9,9) (6,8,8)	10
V	(0,1,10) (0,2,9) (0,3,8) (0,4,7) (0,5,6)	5
S	(1,2,8) (1,3,7) (1,4,6) (2,3,6) (2,4,5) (3,9,10) (4,8,10) (5,7,10) (5,8,9) (6,7,9)	10
sum		26

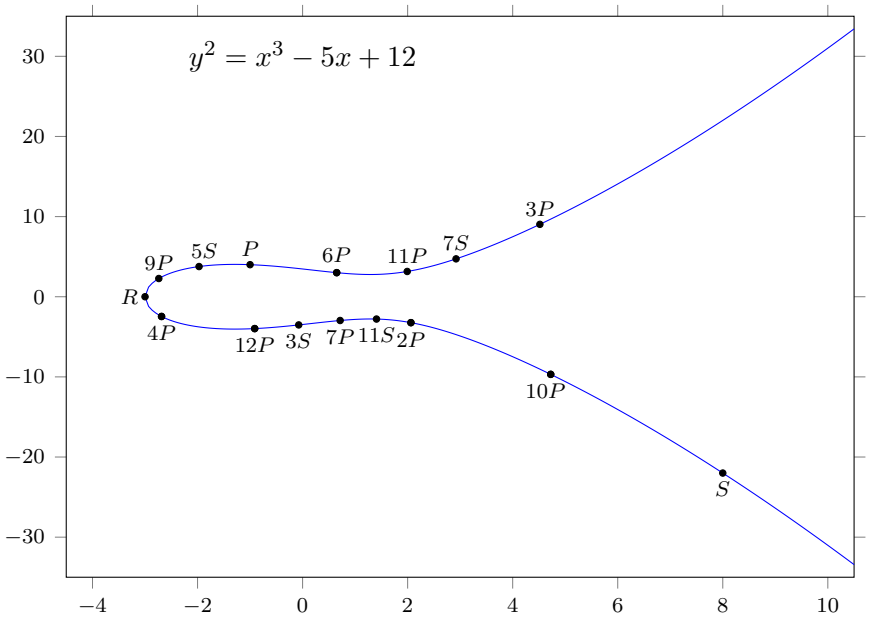
type	$n = 12$	$L_{\text{type}}(12)$
O	(0,0,0)	1
I	(4,4,4) (8,8,8)	2
R	(0,6,6)	1
T	(1,1,10) (2,2,8) (3,3,6) (7,7,10) (2,5,5) (2,11,11) (4,10,10) (6,9,9)	8
V	(0,1,11) (0,2,10) (0,3,9) (0,4,8) (0,5,7)	5
S	(1,2,9) (1,3,8) (1,4,7) (1,5,6) (2,3,7) (2,4,6) (3,4,5) (3,10,11) (4,9,11) (5,8,11) (5,9,10) (6,7,11) (6,8,10) (7,8,9)	14
sum		31

The number of lines of type O, I, R and T together will be denoted as L_{TRIO}. If we look at the sum of the number of lines of type O, I, R and T, then $L_{TRIO}(n) = n$. The number of lines of type V and S together will be denoted as L_{VS}. If we compare $L_{VS}(n)$ with $L(n)$ then $L_{VS}(n) = L(n-3)$. Since $L_{VS}(n) = L(n) - L_{TRIO}(n) = L(n) - n$ we obtain $L(n) = L(n-3)+n$. Starting with $n = 4$ we have $L(4) = L(1)+4 = 1+4$, $L(7) = L(4)+7 = 1+4+7$, $L(10) = L(7) + 10 = 1 + 4 + 7 + 10$ and so on. Hence, $L(n) = \frac{1}{6}n^2 + \frac{1}{2}n + \frac{1}{3}$ if $n \cong 1 \mod 3$. Starting with $n = 5$ we have $L(5) = L(2) + 5 = 2 + 5$, $L(8) = L(5) + 8 = 2 + 5 + 8$, $L(11) = L(8) + 11 = 2 + 5 + 8 + 11$ and so on. Hence, $L(n) = \frac{1}{6}n^2 + \frac{1}{2}n + \frac{1}{3}$ if $n \cong 2 \mod 3$. Starting with $n = 6$ we have $L(6) = L(3) + 6 = 4 + 6$, $L(9) = L(6) + 9 = 4 + 6 + 9$, $L(12) = L(9) + 12 = 4 + 6 + 9 + 12$ and so on. Hence, $L(n) = \frac{1}{6}n^2 + \frac{1}{2}n + 1$ if $n \cong 0 \mod 3$.

3.7 Generating rational points

As an example of an elliptic curve with rational points we consider the curve given by $y^2 = x^3 - 5x + 12$. The single root $R(-3, 0)$ is an integer torsion point of order 2. Next to the root $R(-3, 0)$ the curve has $P(-1, 4)$, $-P(-1, -4)$, $S(8, -22)$ and $S(8, 22)$ as integer points, where $S = R + P$. Starting with P we can calculate $2P$ with the 'doubling' formula. Thereafter we can calculate $3P = 2P + P$. We can calculate $4P$ either by regarding it as

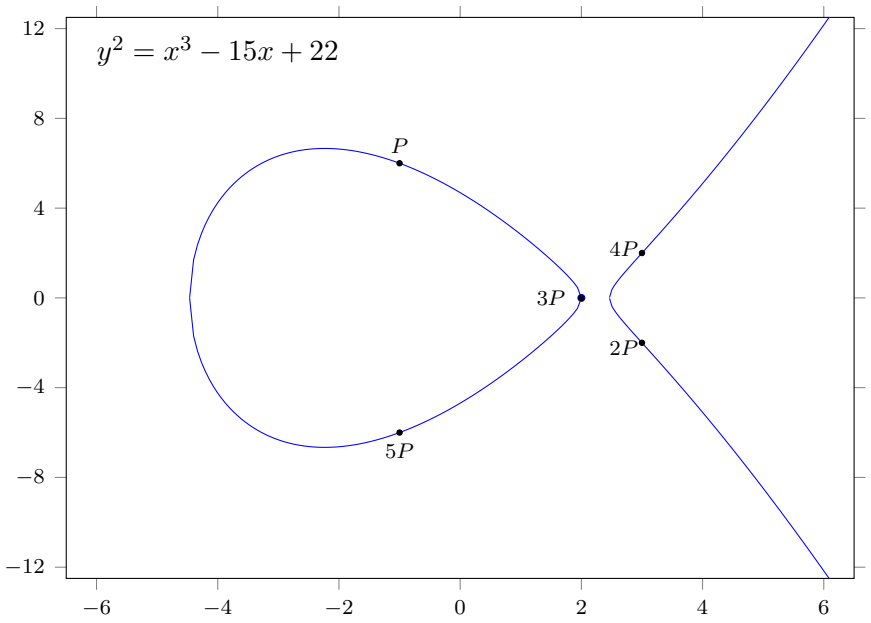
$4P = 2P + 2P$ and apply the doubling formula or by regarding it as $4P = P + 3P$ and apply the addition formula, etc. Either way, we obtain $2P = (\frac{33}{16}, -\frac{207}{64})$, for $3P = (\frac{10847}{2401}, \frac{1062860}{117649})$, $4P = (-\frac{7363967}{2742336}, -\frac{11182515137}{4541308416})$, etc. Notice that the denominator of the x coordinate is the square of a number while the denominator of the y coordinate is the cube of that number. The rational point P and some of its multiples are shown in the next figure.



From $S(8, -22)$ we obtain for $2S$ the coordinates $(\frac{33}{16}, -\frac{207}{64})$, for $3S$ $(-\frac{664}{9025}, -\frac{3015166}{857375})$, for $4S$ $(-\frac{7363967}{2742336}, -\frac{11182515137}{4541308416})$, etc. Since $S = R + P$ the points generated by S are not independent of the points generated by P. It also follows that $2S = 2R + 2P = \mathcal{O} + 2P = 2P$ and thus $4S = 4P$, $6S = 6P$, etc. The points S, P and their multiples are shown in the figure. For clarity, the mirror points $-P$, $-2P$,...,$-kP$, $-S$, $-2S$,..., $-kS$,.. are not shown.

The points on an elliptic curve form an abelian group, $E(\mathbb{R})$. The subgroup of rational points is denoted as $E(\mathbb{Q})$. The rank of an elliptic curve is the number of generators, 'starting points', needed to generate all the rational points. For instance, for the elliptic curve $y^2 = x^3 - 5x + 12$ the rational points are generated (whether or not with the help of the torsion point $R(-3, 0)$) by $P(-1, 4)$. Since there are no other rational points (just take it for granted because rank determination is complicated), the rank is 1. According to a theorem of Mordell the number of generators of rational points always is finite.

As another example we consider the curve $y^2 = x^3 - 15x + 22$. Next to the root $(2, 0)$ the curve has $(-1, 6)$, $(-1, -6)$, $(3, 2)$ and $(3, -2)$ as integer points. If we denote $(-1, 6)$ as P, then $2P = (3, -2)$, $3P = (2, 0)$, $4P = -2P = (3, 2)$, $5P = -P = (-1, -6)$ and $6P = \mathcal{O}$. The situation is shown in the next figure.



The order of P is 6: P is cyclic with cycle length 6. There is no rational point which generates an infinite number of rational points, so the rank is 0.

In general, if $nP = \mathcal{O}$ then n is the *order* of point P. For a point P with order n there holds $(n+1)P = P$. That is, the point P is cyclic with cycle length n. A cyclic point P is called a *torsion* point. An elliptic curve is denoted as E. The group of all points on the curve as $E(\mathbb{R})$. The group of torsion points E_{TORS} is a subgroup of $E(\mathbb{R})$.
The group of rational torsion points is called $E(\mathbb{Q})_{\text{TORS}}$. For the order n of a rational torsion point there holds $n \leq 12$ and $n \neq 11$; a theorem of Mazur. Only the neutral point $\mathcal{O}$ has order 1. Roots, points on the $y = 0$ axis, have order 2. There are 3 root points (of which 2 may be complex). Together with the neutral point we have 4 points of order 2. The points of inflection have order 3. The inflection equation, $y'' = 0$, is a fourth degree equation in x with 4 (of which 3 complex) solutions for x. For every solution x, y there also is a solution $x, -y$. So, we have 8 solutions. Together with the neutral point we have 9 points of order 3. In a similar way we have n^2 points (possibly complex) of order n. In general, torsion points are not rational. However, if the coefficients of the elliptic equation are integer, the torsion points also are integer. The group of torsion points E_{TORS} is infinite. However, the group of rational torsion points $E(\mathbb{Q})_{\text{TORS}}$ is finite. The group $E(\mathbb{Q})$ of rational points is generated by a finite number of generators. Every element Q of $E(\mathbb{Q})$ can be written as $Q = m_1P_1 + m_2P_2 + ... + m_rP_r + T$, where $P_1, ..., P_r$ are the generators, where $m_1, ..., m_r \in \mathbb{Z}$, where r is the rank and where $T \in E(\mathbb{Q})_{\text{TORS}}$. The group $E(\mathbb{Q})$ is isomorphic to $\mathbb{Z}^r \oplus E(\mathbb{Q})_{\text{TORS}}$.

3.8 Rational points on $y^2 = x^3 - 25x$

The elliptic curve $E:\ y^2 = x^3 - 25x$ has $(-5,0)$, $(0,0)$, $(5,0)$, $(-4,6)$, $(-4,-6)$, $(45,300)$ and $(45,-300)$ as integer points. The roots $(-5,0)$, $(0,0)$ and $(5,0)$ are three points each with order 2. Earlier we found that $(-4,6)$ corresponds to the $(1\frac{1}{2}, 6\frac{2}{3}, 6\frac{5}{6})$ right triangle with area 5. Denoting $(-4,6)$ as P we obtain $2P = (11\frac{97}{144}, -36\frac{71}{1728})$. Earlier we found it corresponds to the $(3\frac{43}{492}, 3\frac{363}{1519}, 4\frac{354769}{747348})$ right triangle. For the present purpose we write it as $a = \frac{12519}{492}$, $b = \frac{4920}{1519}$ and $c = \frac{3344161}{492\cdot 1519}$. Multiplication by 747348 leads to the Pythagorean triple: $(1519^2, 10\cdot 492^2, 3344161)$. By means of $3P = 2P + P$ we obtain $3P = (-\frac{2439844}{5094049}, 3\frac{5109762975}{11497268593})$. From the correspondence $x = \dfrac{nb}{c-a}$ and $y = \pm\dfrac{2n^2}{c-a}$ we obtain $a = \dfrac{x^2-n^2}{y}$, $b = \dfrac{2nx}{y}$ and $c = \dfrac{x^2+n^2}{y}$. For the coordinates of $3P$ it leads to a right triangle with rational sides $a = \frac{25353117}{3525434}$, $b = \frac{35254340}{25353117}$ and $c = \frac{654686219104361}{89380740677778}$. Multiplication by $89380740677778 = 3525434\cdot 25353117$ leads to the Pythagorean triple: $(25353117^2, 10\cdot 3525434^2, 654686219104361)$. For $4P$ we find the coordinates $(5\frac{12832131841}{2234116132416}, \frac{1791076534232245919}{3339324446657665536})$. It corresponds to a right triangle with rational sides $a = \frac{535583225279}{4998504070056}$, $b = \frac{49985040700560}{535583225279}$ and $c = \frac{24985059404727155836448064 1}{2677114931410801046145624}$. Multiplication by $267711493141080104614562 4 = 4998504070056\cdot 535583225279$ leads to the Pythagorean triple: $(535583225279^2, 10\cdot 4998504070056^2, 24985059404727155836448064 1)$.

coordinates (x,y) of nP	Pythagorean triangle A_n, B_n, C_n	generating (k_n, m_n)
$x_P = -2^2, \quad y_P = 2\cdot 3$	$A_1 = 3^2, \quad B_1 = 2^3\cdot 5, \quad C_1 = 41$	$k_1 = 5,\ m_1 = 2^2$
$x_{2P} = \dfrac{41^2}{2^4\cdot 3^2}$ $y_{2P} = -\dfrac{7^2\cdot 31\cdot 41}{2^6\cdot 3^3}$	$A_2 = 7^4\cdot 31^2,\ B_2 = 2^5\cdot 3^2\cdot 5\cdot 41^2$ $C_2 = 3344161$	$k_2 = 41^2$ $m_2 = 2^4\cdot 3^2\cdot 5$
$x_{3P} = -\dfrac{\alpha^2}{37^2\cdot 61^2}$ $y_{3P} = \dfrac{\alpha\cdot 3^2\cdot 587\cdot 4799}{37^3\cdot 61^3}$ where $\alpha := 2\cdot 11\cdot 71$	$A_3 = 3^4\cdot 587^2\cdot 4799^2$ $B_3 = 2^3\cdot 5\cdot 11^2\cdot 37^2\cdot 61^2\cdot 71^2$ $C_3 = 41\cdot 15967956563521$	$k_3 = 5\cdot 37^2\cdot 61^2$ $m_3 = 2^2\cdot 11^2\cdot 71^2$
$x_{4P} = \dfrac{\beta^2}{2^6\cdot 3^2\cdot 7^4\cdot 31^2\cdot 41^2}$ $y_{4P} = \dfrac{113279\cdot\beta\cdot 4728001}{2^9\cdot 3^3\cdot 7^6\cdot 31^3\cdot 41^3}$ where $\beta := 3344161$	$A_4 = 113279^2\cdot 4728001^2$ $B_4 = 2^7\cdot 3^2\cdot 5\cdot 7^4\cdot 31^2\cdot 41^2\cdot\beta^2$ $C_4 = 545834881\cdot 457740248460360961$	$k_4 = 3344161^2$ $m_4 = 2^6\cdot 3^3\cdot 5\cdot 7^4\cdot 31^2\cdot 41^2$

In the previous table coordinates for nP and the integer sides of the corresponding Pythagorean triangle and its generating numbers (k, m) ($a = k^2 - m^2, b = 2km, c = k^2 + m^2$ remember) are shown in factorized form. We see the generating (k_{2j}, m_{2j}) for $(2j)P$ follows from the (A_j, B_j, C_j) for jP via $k_{2j} = C_j^2$ and $m_{2j} = 2A_jB_j$. In conclusion, we can construct an infinite number of Pythagorean triangles for which $AB/2$ is 5 times a square.

Denoting the roots $(-5, 0)$, $(0, 0)$ and $(5, 0)$ respectively as R_-, R_0 and R_+ and denoting $(45, -300)$ as S we find $R_- + P = S$. Also here $2P = 2S$, so only odd multiples of S are new points. The first two of them are shown in the next table.

coordinates (x, y) of nS	Pyth. triangle A_n, B_n, C_n	generating k_n, m_n
$x_S = 3^2 \cdot 5, \quad y_S = 2^2 \cdot 3 \cdot 5^2$	$B_1 = 3^2,\ A_1 = 2^3 \cdot 5,\ C_1 = 41$	$k_1 = 5,\ m_1 = 2^2$
$x_{3S} = \dfrac{\alpha^2 \cdot 5}{4799^2}$ $y_{3S} = -\dfrac{\alpha \cdot 2^2 \cdot 5^2 \cdot 11 \cdot 37 \cdot 71}{4799^3}$ where $\alpha := 3^2 \cdot 587$	$B_3 = 3^4 \cdot 587^2 \cdot 4799^2$ $A_3 = 2^3 \cdot 5 \cdot 11^2 \cdot 37^2 \cdot 61^2 \cdot 71^2$ $C_3 = 41 \cdot 15967956563521$	$k_3 = 5 \cdot 37^2 \cdot 61^2$ $m_3 = 2^2 \cdot 11^2 \cdot 71^2$

We see that changing from $(2j + 1)P$ to $(2j + 1)S$ is a matter of changing roles of A_{2j+1} and B_{2j+1}. Not a surprise because this was the way we constructed S from P earlier.

Next we consider $V = R_0 + P$. We find for V the coordinates $(6\frac{1}{4}, 9\frac{3}{8})$, from which we can find new points $3V$, $5V$, etc. The first two are shown in the next table.

coordinates (x, y) of nV	Pyth. triangle A_n, B_n, C_n	generating k_n, m_n
$x_V = \dfrac{5^2}{2^2}, \quad y_V = \dfrac{3 \cdot 5^2}{2^3}$	$A_1 = 3^2,\ B_1 = 2^3 \cdot 5,\ C_1 = 41$	$k_1 = 5,\ m_1 = 2^2$
$x_{3V} = \dfrac{\alpha^2}{2^2 \cdot 11^2 \cdot 71^2}$ $y_{3V} = \dfrac{\alpha \cdot 3^2 \cdot 5 \cdot 587 \cdot 4799}{2^3 \cdot 11^3 \cdot 71^3}$ where $\alpha := 5 \cdot 37 \cdot 61$	$A_3 = 3^4 \cdot 587^2 \cdot 4799^2$ $B_3 = 2^3 \cdot 5 \cdot 11^2 \cdot 37^2 \cdot 61^2 \cdot 71^2$ $C_3 = 41 \cdot 15967956563521$	$k_3 = 5 \cdot 37^2 \cdot 61^2$ $m_3 = 2^2 \cdot 11^2 \cdot 71^2$

Finally we consider $W = R_+ + P$. We find for W the coordinates $(-\frac{5}{9}, -3\frac{19}{27})$, from which we can find new points $3W$, $5W$, etc. The first two are shown in the next table. The coordinates and the Pythagorean triangle sides of, for instance, $3P$ are governed by the prime numbers 2,3,5,11,37,61,71,587,4799 . For the coordinates of $3S$, $3V$ and $3W$ some of these prime numbers are moved from denominator to numerator and vice versa in comparison with $3P$.

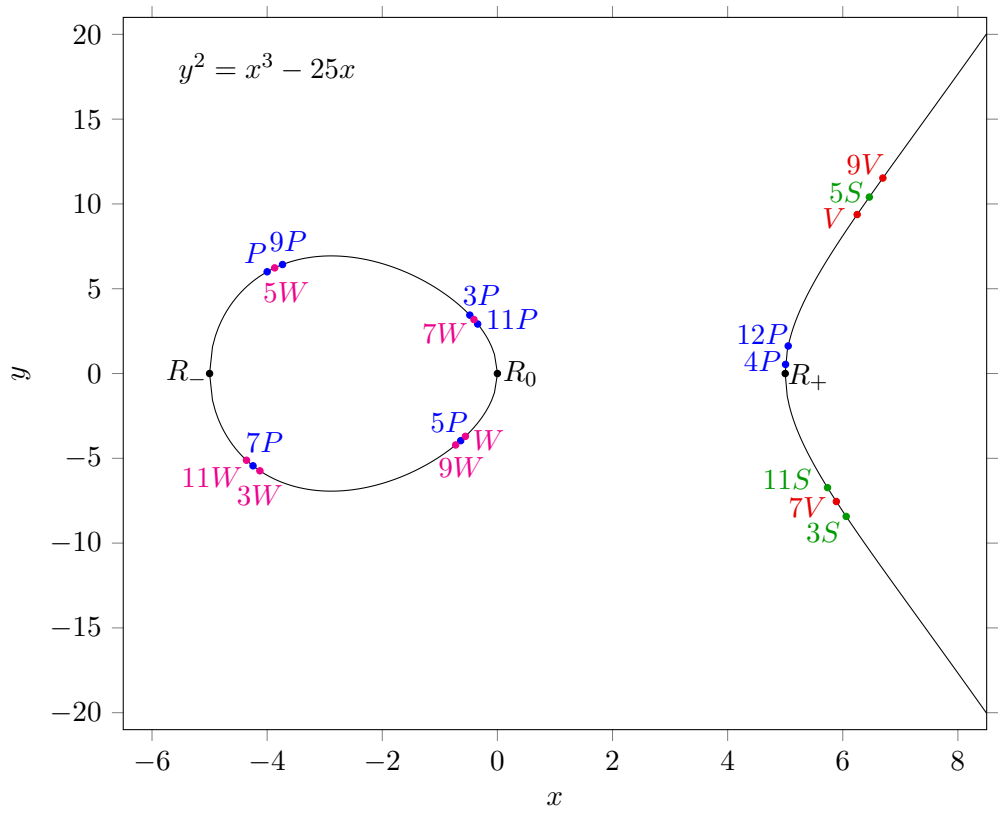
coordinates (x, y) of nW	Pyth. triangle A_n, B_n, C_n	generating k_n, m_n
$x_W = -\dfrac{5^2}{2^2}, \quad y_W = -\dfrac{3 \cdot 5^2}{2^3}$	$A_1 = 3^2,\ B_1 = 2^3 \cdot 5,\ C_1 = 41$	$k_1 = 5,\ m_1 = 2^2$
$x_{3W} = -\dfrac{\alpha^2}{3^4 \cdot 587^2}$	$B_3 = 3^4 \cdot 587^2 \cdot 4799^2$	$k_3 = 5 \cdot 37^2 \cdot 61^2$
$y_{3W} = -\dfrac{\alpha \cdot 2^2 \cdot 5 \cdot 11 \cdot 37 \cdot 61 \cdot 71}{3^6 \cdot 587^3}$	$A_3 = 2^3 \cdot 5 \cdot 11^2 \cdot 37^2 \cdot 61^2 \cdot 71^2$	$m_3 = 2^2 \cdot 11^2 \cdot 71^2$
where $\alpha := 5 \cdot 4799$	$C_3 = 41 \cdot 15967956563521$	

From the addition of two points P, S, V and W we obtain new points:

$P+S = V+W = \left(-\dfrac{5 \cdot 31^2}{7^4}, \dfrac{2 \cdot 3 \cdot 5^2 \cdot 31 \cdot 41}{7^6}\right)$, $P+W = S+V = \left(\dfrac{5 \cdot 7^4}{31^2}, \dfrac{2^3 \cdot 3 \cdot 5^2 \cdot 7^2 \cdot 41}{31^3}\right)$, $P + V = S + W = \left(-\dfrac{2^4 \cdot 3^2 \cdot 5^2}{41^2}, -\dfrac{2^2 \cdot 3 \cdot 5^2 \cdot 7^2 \cdot 31}{41^3}\right)$ and so on. Notice that W is not an 'independent' point since $W = S + V - P$. The roots R_-, R_0, R_+ are torsion points, each of order 2. Together with the rational point P all the other rational points are generated. So, the rank is 1.

One might wonder if there is a point, say H, such that its double is $P(-4, 6)$. Using the doubling formula in reversed order we obtain fourth degree equations for the coordinates of H. The four solutions are complex: $(2 + i, -1 + 7i)$, $(2 - i, -1 - 7i)$, $(-10 + 5i, 25 + 25i)$ and $(-10 - 5i, 25 - 25i)$. Often H is denoted as $\frac{1}{2}P$ for obvious reasons: $2 \cdot \frac{1}{2}P = P$. Application of the doubling formula in reversed order to $2P = (11\frac{97}{144}, -36\frac{71}{1728})$ leads to the following four solutions: $(-4, 6)$, $(45, -300)$, $(6\frac{1}{4}, 9\frac{3}{8})$ and $(-\frac{5}{9}, -3\frac{19}{27})$. That is, P, S, V and W as expected since $2S = 2V = 2W = 2P$.

In the next figure a number of multiples of P, S, V and W are shown (the mirror points are left).



The positions of rational points nP in the figure above are such that nP is close to $(n+8)P$. A similar observation can be made for the points nS, nV and nW. The reason for this is that the generator $(-4, 6)$ of the rational points on the curve $E : y^2 = x^3 - 25x$ is close to the non-rational torsion point $(-4.03198, 5.93736)$ (see the brown $3P$ in the eightfold torsion figure two sections earlier).

3.9 Modular counting on elliptic curves

As we did for a circle equation in the first section, we will apply modular counting on elliptic curves. As an example we will consider the curve $E : \ y^2 = x^3 + ax + b$ modulo a prime number. Alternatively, we consider $y^2 = x^3 + ax + b$ over the field $\mathbb{Z}/p\mathbb{Z}$ with p prime. The group of integer points on an elliptic curve E over $\mathbb{Z}/p\mathbb{Z}$ is usually denoted as $E(\mathbb{F}_p)$. The prime should not be 2 or 3 for reasons we will not go into. In addition we can only take prime numbers for which the curve does not become singular. A singularity occurs if the discriminant D becomes 0. The discriminant of an $n-$th degree equation is defined as the product of the squares of the distances between the roots times a_n^{2n-2}, where a_n is the leading coefficient, the coefficient of x^n. Thus

$$D = a_n^{2n-2} \prod_{i>j} (x_i - x_j)^2 \,. \tag{3.7}$$

For the quadratic equation $y = ax^2 + bx + c$ the leading coefficient is a. Since the two roots are $x_1 = \dfrac{-b - \sqrt{b^2 - 4ac}}{2a}$ and $x_1 = \dfrac{-b + \sqrt{b^2 - 4ac}}{2a}$, the distance between them is $x_2 - x_1 = \dfrac{\sqrt{b^2 - 4ac}}{a}$. Hence, $D = a^2(x_2 - x_1)^2 = b^2 - 4ac$.
For the cubic equation $y = x^3 + ax + b$ the leading coefficient is 1. If we denote the three roots as x_1, x_2 and x_3 then $D = 1^4(x_2 - x_1)^2(x_3 - x_1)^2(x_3 - x_2)^2$. The calculation of the roots of the cubic equation is standard in complex function theory. We just give the result: $x_1 = \dfrac{a}{T} + \dfrac{T}{3}$, $x_2 = \dfrac{1 + i\sqrt{3}}{2}\dfrac{a}{T} - \dfrac{1 - i\sqrt{3}}{2}\dfrac{T}{3}$ and $x_2 = \dfrac{1 - i\sqrt{3}}{2}\dfrac{a}{T} - \dfrac{1 + i\sqrt{3}}{2}\dfrac{T}{3}$, where $T = \sqrt[3]{\dfrac{-27b + 3\sqrt{3}\sqrt{4a^3 + 27b^2}}{2}}$. From these expressions one obtains the following expression for the discriminant: $D = -(4a^3 + 27b^2)$.

If we consider an elliptic equation modulo a prime p, then the curve is singular if the discriminant is $0 \mod p$. For example for $y^2 = x^3 - 5x + 12$ the discriminant is $D = -(4 \cdot -5^3 + 27 \cdot 12^3) = -3388$. Modulo 7 we have $D = -3388 \mod 7 = 0$. Since $3388 = 2^2 \cdot 7 \cdot 11^2$ the discriminant will also be 0 for $p = 11$: $D \mod 11 = 0$. Therefore $p = 7$ and $p = 11$ are not allowed for modulo counting on the curve $y^2 = x^3 - 5x + 12$. With this in mind we consider modular counting on some elliptic curves.

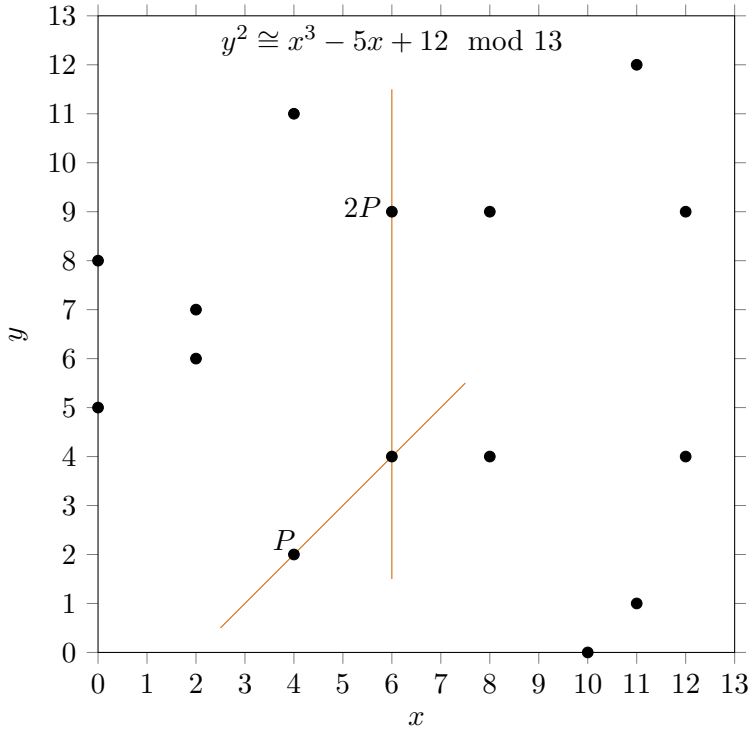
3.10 Modular counting on $y^2 = x^3 - 5x + 12$

In this section we will consider the curve $y^2 = x^3 - 5x + 12$ modulo a prime number p.
For instance, for $p = 13$ the integer points on 'the curve' are $(10, 0)$, $(11, 1)$, $(4, 2)$, $(6, 4)$, $(8, 4)$, $(12, 4)$, $(0, 5)$, $(2, 6)$, $(2, 7)$, $(0, 8)$, $(6, 9)$, $(8, 9)$, $(12, 9)$, $(4, 11)$ and $(11, 12)$. Together with $\mathcal{O}$ the group of 16 points is $E(\mathbb{F}_{13})$. If we denote $(4, 2)$ as P the doubling formula gives $2P = (6, 9)$. The calculation is as follows. First the slope of the tangent line in P:

$$\lambda = \frac{3x_P^2 + a}{2y_P} = \frac{3 \cdot 4^2 - 5}{2 \cdot 2} = \frac{43}{4} \cong \frac{4}{4} \cong 1 \mod 13\,. \tag{3.8}$$

Having obtained the slope we proceed: $x_{2P} = \lambda^2 - 2x_P = 1^2 - 2 \cdot 4 = -7 \cong 6 \mod 13$ and $y_{2P} = \lambda(x_P - x_{2P}) - y_P = 1(4 - 6) - 2 = -4 \cong 9 \mod 13$. Indeed $2P = (6, 9)$.

Since the tangent line through $P = (4, 2)$ has slope 1 it arrives in integer point: $(6, 4)$. The latter point is mirrored with respect to $y = 6\frac{1}{2}$, similar to the $y = 0$ mirror for continuous curves. The final point is $2P = (6, 9)$. The doubling of P is illustrated in the next figure.



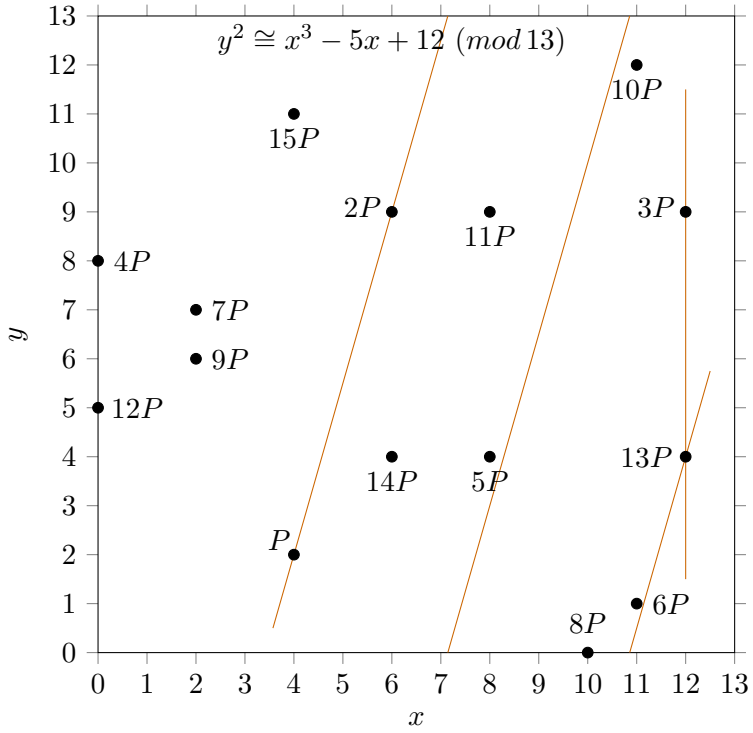
Next we apply the addition formula to obtain $3P$:

$$\lambda = \frac{y_{2P} - y_P}{x_{2P} - x_P} = \frac{9-2}{6-4} = \frac{7}{2} = 2^{-1} \cdot 7 \cong 7 \cdot 7 \cong 49 \cong 10 \mod 13\,. \tag{3.9}$$

Notice that 7 is the inverse of 2 since $7 \cdot 2 = 14 \cong 1 \mod 13$. This makes clear that unique inverses require the modulo counting with a prime number. Having obtained the slope of the line connecting P and $2P$ we proceed: $x_{3P} = \lambda^2 - x_P - x_{2P} = 10^2 - 4 - 6 = 90 \cong 12 \mod 13$ and $y_{2P} = \lambda(x_P - x_{3P}) - y_P = 10(4-12) - 2 = -82 \cong 9 \mod 13$. Hence $3P = (12, 9)$. Continuing the addition we obtain $4P = (0, 8)$, $5P = (8, 4)$, $6P = (11, 1)$, $7P = (2, 7)$, $8P = (10, 0)$, $9P = (2, 6)$, $10P = (11, 12)$, $11P = (8, 9)$, $12P = (0, 5)$, $13P = (12, 4)$, $14P = (6, 4)$, $15P = (4, 11)$ and $16P = \mathcal{O}$. So, P is of order 16. Since $1, 3, 5, 7, 9, 11, 13$ and 15 are relative prime to 16 (recall $\varphi(16) = 8$ with φ Euler's totient function), the 8 points $P, 3P, 5P, 7P, 9P, 11P, 13P, 15P$ have order 16. The point $2P$ has order 8, and since $1, 3, 5$ and 7 are relative prime to 8, the 4 points $2P, 6P, 10P, 14P$ have order 8. The point $4P$ has order 4, and since 1 and 3 are relative prime to 4, the 2 points $4P, 12P$ have order 4. The point $8P$ has order 2. The point $16P$ has order 1, $16P = \mathcal{O}$ is the single element with order 1. The full group $P, 2P, ..., 16P$ is isomorphic to the cyclic group C_{16}. Subgroups are C_8, C_4, C_2 and C_1.

The line through P and $2P$ goes through $(7\frac{1}{7}, 13)$ where it is wrapped to $(7\frac{1}{7}, 0)$. From there it goes to $(10\frac{6}{7}, 13)$ where it is wrapped to $(10\frac{6}{7}, 0)$ after which it arrives at the integer

point $(12, 4)$. The latter point is mirrored with respect to $y = 6\frac{1}{2}$ to $(12, 9)$. So, $3P = (12, 9)$. The addition of $P + 2P = 3P$ is illustrated in the next figure.

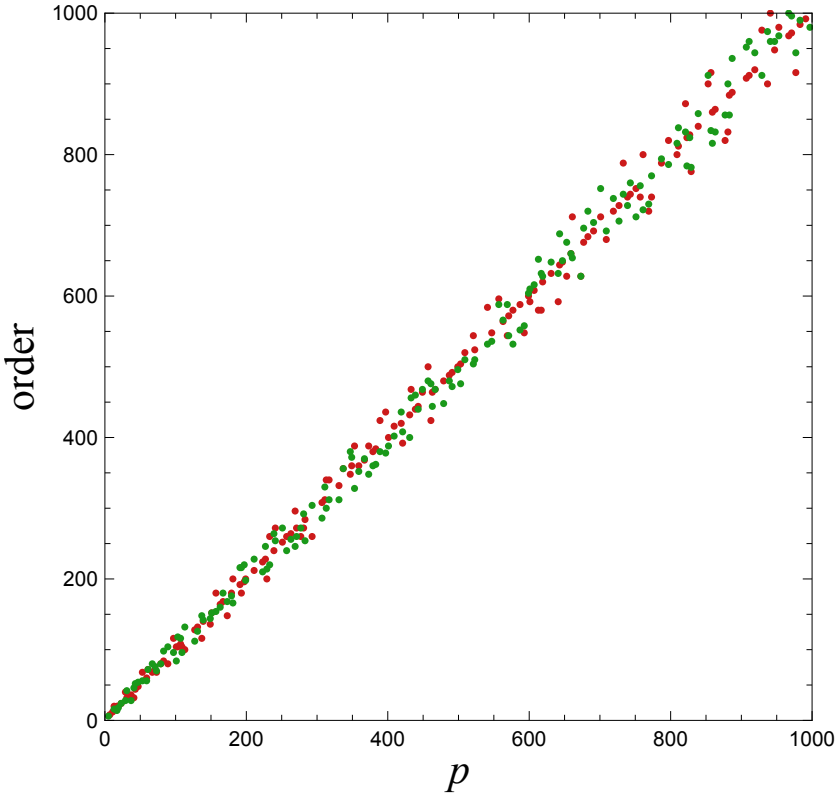


Since $15P = -P$ we see that going from P to $-P$ is a matter of reflection of the y coordinate with respect to the $y = 6\frac{1}{2}$ line. For an elliptic curve over the field $\mathbb{Z}/p\mathbb{Z}$ the points are mirrored in $y = p/2$.

For the next prime, $p = 17$, we find 13 integer points. Together with the neutral element $\mathcal{O}$ the points forms a group of order 14: $P, 2P, ..., 14P = \mathcal{O}$. The subgroup $2P, 4P, ..., \mathcal{O}$ has order 7, and the element $7P$ has order 2. The largest order of the elements is 14. The full group $P, 2P, ..., 14P$ is isomorphic to the cyclic group C_{14}. Subgroups are C_7, C_2 and C_1.

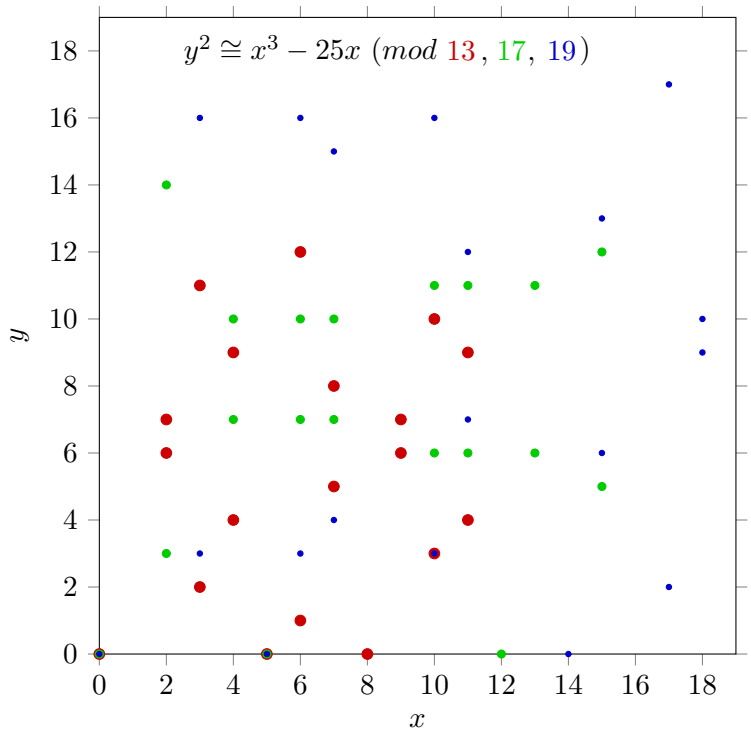
For $p = 19$ we obtain the following 17 points: $(16, 0)$, $(8, 3)$, $(7, 4)$, $(13, 4)$, $(18, 4)$, $(15, 5)$, $(5, 6)$, $(14, 8)$, $(3, 9)$, $(3, 10)$, $(14, 11)$, $(5, 13)$, $(15, 14)$, $(7, 15)$, $(13, 15)$, $(18, 15)$ and $(8, 16)$. Together with $\mathcal{O}$ we have 18 elements; the order of the group is 18. Denoting $(7, 4)$ as P we obtain $2P = (3, 10)$, $3P = (16, 0)$, $4P = (3, 9)$, $5P = (7, 15)$, $6P = \mathcal{O}$. Denoting $(8, 3)$ as Q, we obtain $2Q = (8, 16)$ and $3Q = \mathcal{O}$. The other points now are $P + Q = (5, 13)$, $2P + Q = (13, 4)$, $3P + Q = (18, 15)$, $4P + Q = (14, 8)$, $5P + Q = (15, 5)$, $P + 2Q = (15, 14)$, $2P + 2Q = (14, 11)$, $3P + 2Q = (18, 4)$, $4P + 2Q = (13, 15)$, $5P + 2Q = (5, 6)$. The group structure is $C_6 \times C_3$. The largest order of the elements of the group is 6.

The number of integer points $E(\mathbb{F}_p)$ including $\mathcal{O}$ is the order of $E(\mathbb{F}_p)$. The order of each element of $E(\mathbb{F}_p)$ is a divisor of the order of $E(\mathbb{F}_p)$. For a different elliptic curve such as $E:\ y^2 = x^3 - 25x$ we obtain 20 for the order of $E(\mathbb{F}_{13})$, $E(\mathbb{F}_{17})$ and $E(\mathbb{F}_{19})$. They are not far away from the corresponding values for the elliptic curve $E:\ y^2 = x^3 - 5x + 12$. They also are not far away from the prime p. That this is even more so for larger primes is illustrated in the next figure where the order of $E(\mathbb{F}_p)$ for the elliptic curves $E:\ y^2 = x^3 - 5x + 12$ (green dots) and $E:\ y^2 = x^3 - 25x$ (red dots) is plotted against p for $p < 1000$.



3.11 Modular counting on $y^2 = x^3 - 25x$

In this section we will consider the curve $y^2 = x^3 - 25x$ modulo a prime number.
For $p = 7$ there are 8 points (group structure $C_4 \times C_2$). For $p = 11$ there are 12 points. They are generated by two elements, one of order 6 and one of order 2 ($C_6 \times C_2$) . For $p = 13$, 17 and 19 there are 20 points. In all three cases the group structure is $C_{10} \times C_2$. For $p = 13$, 17 and 19 the corresponding points (except $\mathcal{O}$) are shown as respectively red, green and blue dots in the next figure.



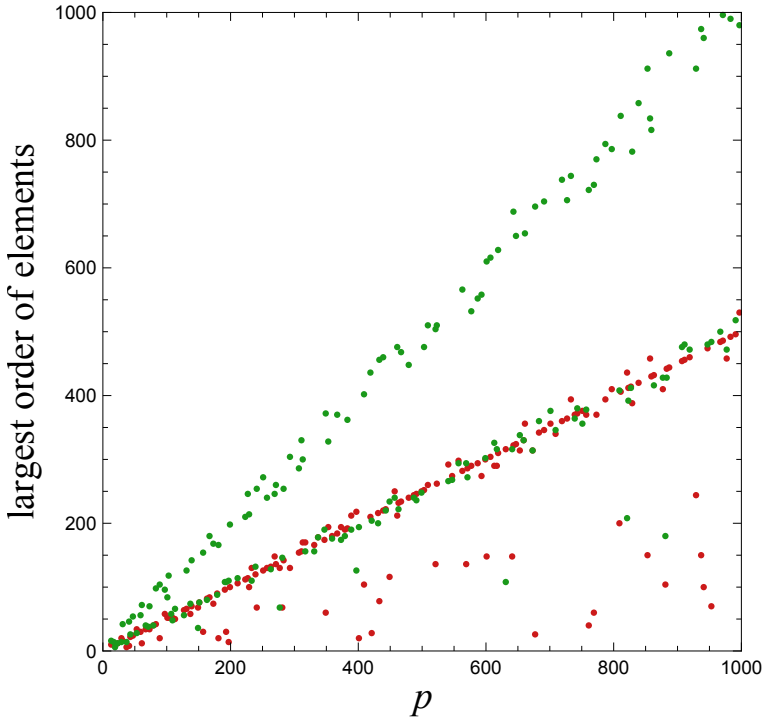
A common point for red and blue is $(10, 3)$. Common points for all three colours are the roots $(0, 0)$ and $(5, 0)$. For the third root, $(-5, 0)$, we find different values for different p: $(-5, 0) \cong (8, 0) \mod 13$, $(-5, 0) \cong (12, 0) \mod 17$ and $(-5, 0) \cong (14, 0) \mod 19$. The number of points with $y = 0$ is either 0 (if there is no integer root), 1 (in case of one integer root) or 3 (in case of three integer roots). For a group $E(\mathbb{F}_p)$ the sum of the y coordinates of points with the same x coordinate is p, because of the reflection with respect to the horizontal line $y = p/2$. If there are three points with the same y coordinate, the sum of the x coordinates is p or $2p$.

For $p = 13$ (red points) there are eight x values with 2 points. We will denote it as $X_2 = 8$. For $p = 13$ there are two x values with no points. We will denote it as $X_0 = 2$. In general we will denote the number of x values with k points as X_k and the number of y values with k points as Y_k. For $p = 13, 17$ and 19 the X_k and Y_k values are tabulated

p	X_1	X_2	Y_1	Y_2	Y_3	$\#E(\mathbb{F}_p)$
13	3	8	8	4	1	20
17	3	8	4	0	5	20
19	3	8	8	0	3	20

Taking $\mathcal{O}$ into account there holds $X_1 + 2X_2 + 1 = Y_1 + 2Y_2 + 3Y_3 + 1 = \#E(\mathbb{F}_p)$. The X_0 and Y_0 are not shown in the table since $X_0 = p - X_1 - X_2$ and $Y_0 = p - Y_1 - Y_2 - Y_3$.

For $E:\ y^2 = x^3 - 25x$ the largest order of the elements of $E(\mathbb{F}_{13})$, $E(\mathbb{F}_{17})$ and $E(\mathbb{F}_{19})$ is 10, while the order of these three groups is 20. In all three cases the largest order of the elements is half the order of the group. In the next figure the largest order of the elements of the group $E(\mathbb{F}_p)$ is plotted against p, $p < 1000$, for the elliptic curves $E:\ y^2 = x^3 - 5x + 12$ (green dots) and $E:\ y^2 = x^3 - 25x$ (red dots). We see that for the curve $E:\ y^2 = x^3 - 5x + 12$ the largest order of the elements of the group $E(\mathbb{F}_p)$ is in most cases as large as the order of the group $E(\mathbb{F}_p)$, while for the curve $E:\ y^2 = x^3 - 25x$ the largest order of the elements of the group $E(\mathbb{F}_p)$ is at most half the order of the group $E(\mathbb{F}_p)$.



3.12 A ratio in $E(\mathbb{F}_p)$

Let us define μ as the following ratio:

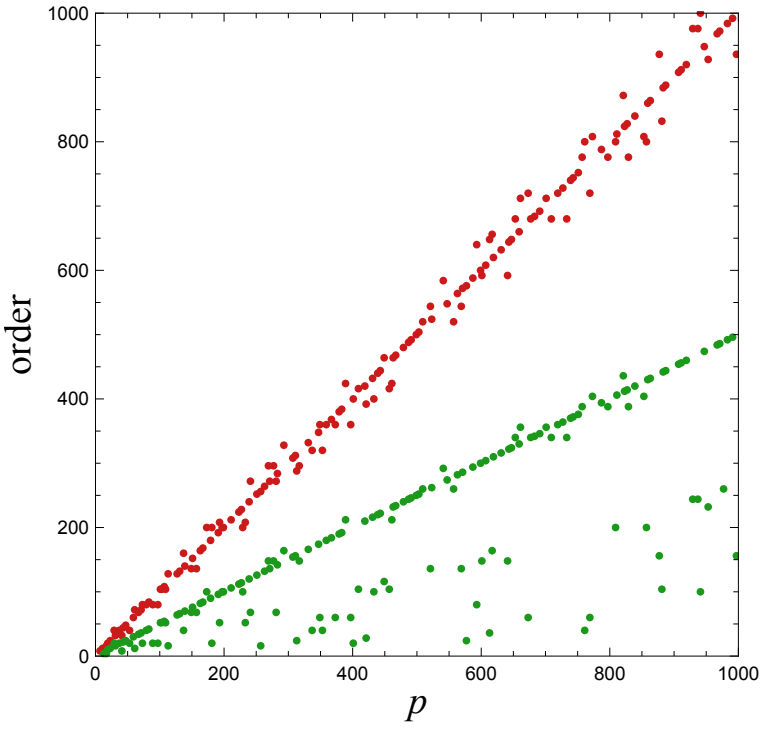
$$\mu(p; E) = \frac{\text{order of } E(\mathbb{F}_p)}{\text{largest order of the elements of } E(\mathbb{F}_p)}\,. \tag{3.10}$$

The set of different μ values depends on the elliptic curve and on p. For $p < 1000$ (and p not a divisor of the discriminant) the ratio μ takes on the values:
$1, 2, 3, 4, 5, 6$ for $E:\ y^2 = x^3 - 5x + 12$,
$1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 16$ for $E:\ y^2 = x^3 - 15x + 22$,
$1, 2, 3, 4, 6$ for $E:\ y^2 = x^3 - 3x + 18$,

$1, 2, 4, 5, 6, 7, 10, 12, 14, 23, 24$ for $E: y^2 = x^3 - 3x$,
$1, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 24, 28$ for $E: y^2 = x^3 + 8$,
$1, 2, 4, 6$ for $E: y^2 = x^3 + 3x - 4$,
$1, 2, 3, 4, 6$ for $E: y^2 = x^3 + 3x + 4$,
$2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24$ for $E: y^2 = x^3 - x$,
$2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24$ for $E: y^2 = x^3 - 16x$,
$2, 4, 6, 8, 10, 12, 14, 20, 26$ for $E: y^2 = x^3 - 25x$,
$2, 4, 6, 8, 10, 12, 14, 22, 24$ for $E: y^2 = x^3 - 36x$.

The set of μ's for $E: y^2 = x^3 - x$ and the set of μ's for $E: y^2 = x^3 - 16x$ are identical. In general the sets for $E: y^2 = x^3 - t^4 x$ are the same for any t. Moreover, for $E: y^2 = x^3 - t^4 x$ the order of $E(\mathbb{F}_p)$ as well as the largest order of the elements of $E(\mathbb{F}_p)$ only depend on p and not on t. In fact, for $E: y^2 = x^3 - t^4 x$ the group structures of $E(\mathbb{F}_p)$ only depend on p. This can be understood as follows. If we start with $y^2 = x^3 - x$ and perform the linear transformation $y' = \frac{y}{t^3}$, $x' = \frac{x}{t^3}$ we obtain $y'^2 = x'^3 - t^4 x'$. Since the group structure is not changed by a linear transformation, the group for $E: y'^2 = x'^3 - t^4 x'$ is identical to the one for $E: y^2 = x^3 - x$. In general, the group for $E: y^2 = x^3 + at^4 x + bt^6$ is identical to the one for $E: y^2 = x^3 + ax + b$.

For $E: y^2 = x^3 - x$ the order of the group $E(\mathbb{F}_p)$ against p is shown by the red dots and the largest order of the elements of $E(\mathbb{F}_p)$ against p is shown by the green dots in the next figure.



Chapter 4

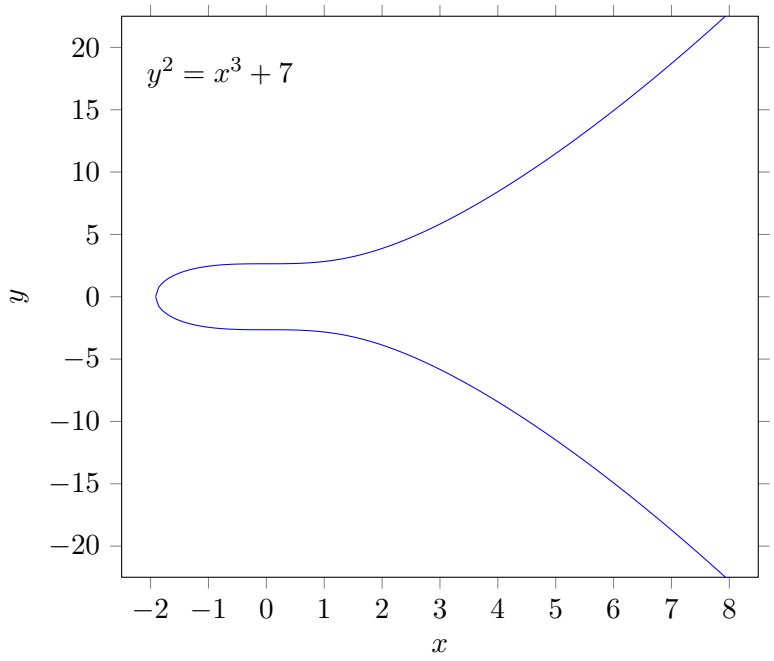
Modular elliptic curves

4.1 Modular counting on $y^2 = x^3 + 7$

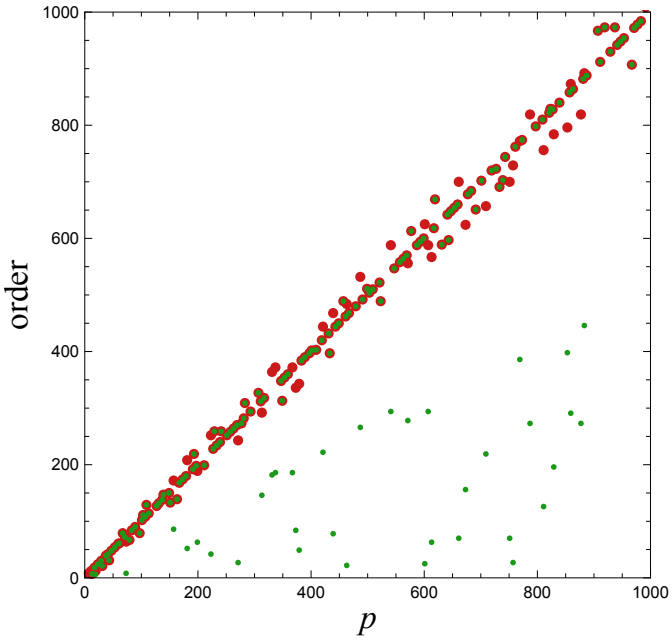
For elliptic curves with integer coefficients the integer torsion points can be systematically found by means of Nagel-Lutz theorem: if an elliptic curve with integer coefficients contains a torsion point the y coordinate of the point is either 0 or its square is a divisor of the discriminant: $y^2|D$. The reverse does not have to be true: an integer (x, y) for which $y^2|D$ it is not necessarily a torsion point. There also are integer points which are part of an infinite series of rational points generated by a generator. For each y satisfying $y^2|D$ one has to test if it belongs to a finite cyclic group or an infinite group.

For example, for $y^2 = x^3 - 25x$ we have for $y = 0$ the integer roots $x = -5$, $x = 0$ and $x = 5$; $(-5, 0)$, $(0, 0)$ and $(5, 0)$ are torsion points of order 2. The discriminant is $D = -(4 \cdot -25^3) = 62500$. The possible values for y^2 such that $y^2|D$ are $y^2 = 1, 5^2, 5^4, 5^6, 2^2 \cdot 5^2, 2^2 \cdot 5^4, 2^2 \cdot 5^6$. Since each of these values for y do not correspond to an integer value for x there are no further torsion points (except for the trivial $\mathcal{O}$). Integer points $(-4, 6)$ and $(45, 300)$ are part of an infinite series of rational points.
As another example, for $y^2 = x^3 - 5x + 12$ we have for $y = 0$ the integer root $x = -3$; $(-3, 0)$ is a torsion point of order 2. The discriminant is $D = -(4 \cdot -5^3 + 27 \cdot 12^2) = -3388$. The only possible values for y^2 such that $y^2|D$ are $y^2 = 1, 2^2, 11^2, 2^2 \cdot 11^2$. Only $y^2 = 22^2$ does correspond to an integer value for x, namely $x = 8$. However, the doubling of $(8, 22)$ does lead to a non-integer rational point. Therefore is $(8, 22)$ not a torsion point.
As a third example, for $y^2 = x^3 - 15x + 22$ we have for $y = 0$ the integer root $x = 2$; $(2, 0)$ is a torsion point of order 2. The discriminant is $D = -(4 \cdot -15^3 + 27 \cdot 22^2) = 432 = 2^4 \cdot 3^3$. The possible values for y^2 such that $y^2|D$ are $y^2 = 1, 2^2, 2^4, 3^2, 3^2 \cdot 2^2, 3^2 \cdot 2^4$. This leads to the following integer points: $(3, 2)$, $(3, -2)$, $(-1, 6)$, $(-1, -6)$. Starting with $P = (-1, 6)$ we obtain $2P = (3, -2)$, $3P = (2, 0)$, $4P = -2P = (3, 2)$, $5P = -P = (-1, -6)$ and $6P = \mathcal{O}$. Since the points are part of a finite cyclic group of integer points they are torsion points.

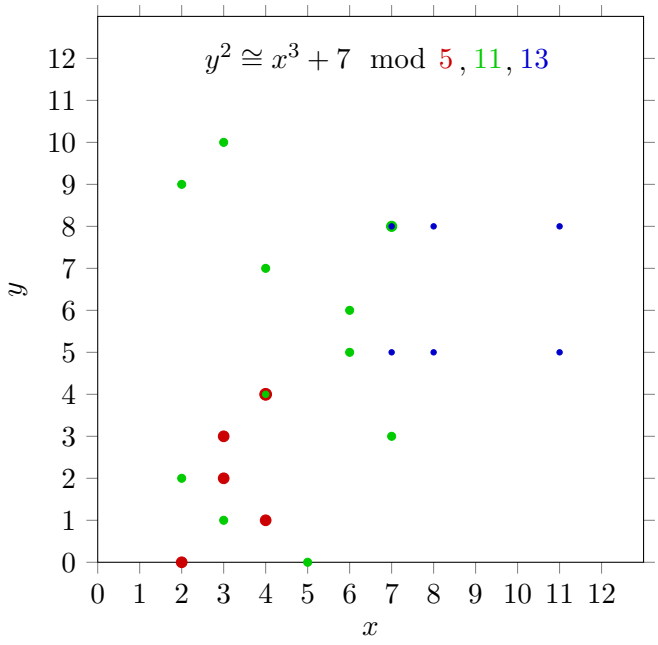
Now we consider the elliptic curve $E: y^2 = x^3 + 7$, which is used in the bitcoin blockchain. As for all curves of the type $y^2 = x^3 + b$ it has the property that $y' = 0$ and $y'' = 0$ for $x = 0$. For $x = 0$ we have $y(0) = \sqrt{7}$ which is not rational. The single root is for $x = \sqrt[3]{-7}$ which is not rational. The discriminant is $D = -1323 = -3^3 \cdot 7^2$. The $y^2|D$ are $y^2 = 1, 3, 7, 21$. None of these y values leads to an integer point. This means there are no torsion points (except for the trivial $\mathcal{O}$). In fact, there are no rational points at all; the rank is 0. The curve $E: y^2 = x^3 + 7$ is shown in the next figure.



Integer points come into existence if we apply modulo counting on $E: y^2 = x^3 + 7$. For the prime we can not take $p = 7$ because it would make the curve singular. For $p = 5$ we obtain $E(\mathbb{F}_5) = \{(4,1), (3,3), (2,0), (3,2), (4,4), \mathcal{O}\}$ with order 6. Denoting $(4,1)$ as P the successive elements are $P, 2P, 3P, 4P, 5P, 6P$. The group is C_6. There are two points of order 6, two of order 3, one of order 2 and one of order 1. The largest order of the elements is 6. Therefore, $\mu(5) = 1$. For $p = 11$ the group is C_{12}. So, the largest order of the elements is 12 and $\mu(11) = 1$. For $p = 13$ the group is C_7 and $\mu(13) = 1$. For $p = 17$ the group is C_{18} and $\mu(17) = 1$. For $p = 19$ the group is $C_6 \times C_2$ and $\mu(19) = 2$. For $p = 23$ the group is C_{24} and $\mu(23) = 1$. For $p = 29$ the group is C_{30} and $\mu(29) = 1$. For $p = 31$ the group is C_{21} and $\mu(31) = 1$. For $p = 37$ the group is C_{39} and $\mu(37) = 1$. For $p = 41$ the group is C_{42} and $\mu(41) = 1$. The second time where $\mu > 1$ is for $p = 73$. Then the group is $C_8 \times C_8$ and $\mu = 8$. We see the values of p for which $\mu \neq 1$ are sparse. In the next figure the order of the group $E(\mathbb{F}_p)$ (red dots) and the largest order of the elements of the group $E(\mathbb{F}_p)$ (green dots) are plotted against p for $p < 1000$.

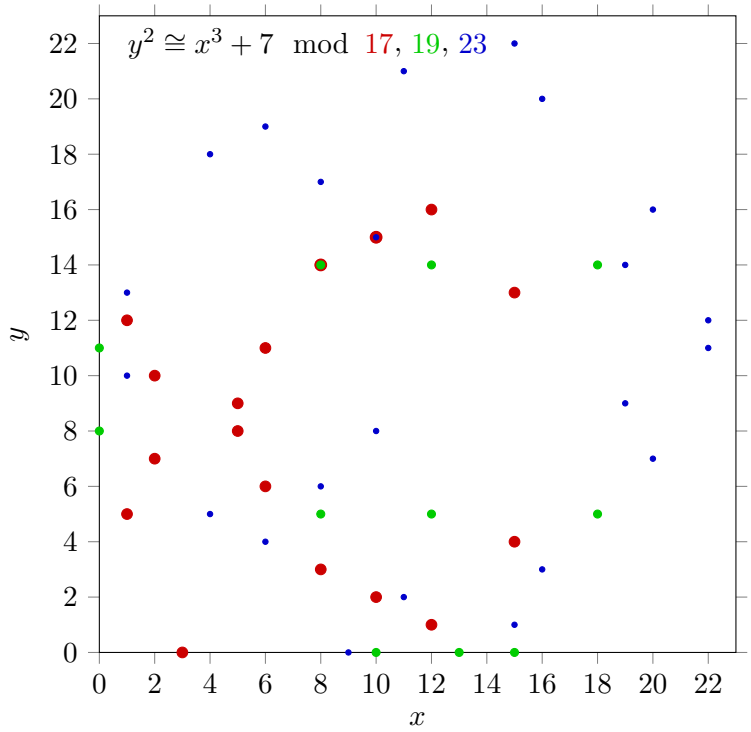


The order of $E(\mathbb{F}_p)$ is in approximately half the cases equal to $p+1$. For these cases the largest order of the elements of $E(\mathbb{F}_p)$ equals the order of $E(\mathbb{F}_p)$, see the green dots on top of the red dots. For the other cases the largest order of the elements of $E(\mathbb{F}_p)$ sometimes does not equal the order of $E(\mathbb{F}_p)$; $\mu \neq 1$. For $p < 1000$ the ratio μ takes on the values $1, 2, 3, 4, 6, 7, 8, 9, 10, 22, 25$ or 27. For $p = 5, 11$ and 13 the points of the group $E(\mathbb{F}_p)$ are shown as respectively red, green and blue dots in the next figure.



The red and green points have $(4, 4)$ in common and the green and blue points have $(7, 8)$ in common. For $p = 5$ there are 3 different x and 5 different y coordinates, 1 point for each y value. For $p = 11$ there are 6 different x and 11 different y coordinates, 1 point for each y value. For $p = 13$ there are 3 different x and 2 different y coordinates, 3 points for each occupied y value.

For $p = 17, 19$ and 23 the points are shown as respectively red, green and blue dots in the next figure.



Also for $p = 17, 19$ and 23 we see x values with 0, 1 or 2 points and y values with 0, 1 or 3 points. From inspection it is found for any $p < 1000$ that for every $0 \leq x \leq p - 1$ there are 0, 1 or 2 points and for every $0 \leq y \leq p - 1$ there are 0, 1 or 3 points. For $E : y^2 = x^3 + 7$ somehow y values with 2 points do not occur for $p < 1000$. As in the previous chapter we denote the number of x values with k points as X_k and the number of y values with k points as Y_k. The order of the group now is: $\#E(\mathbb{F}_p) = Y_1 + 3Y_3 + 1$ or $\#E(\mathbb{F}_p) = X_1 + 2X_2 + 1$. The addition with 1 is to account for the neutral element $\mathcal{O}$. In the next table we have tabulated for each prime p (first column) the value of X_1 (second column), X_2 (third column), Y_1 (fourth column), Y_3 (fifth column), the order of the group $E(\mathbb{F}_p)$ (sixth column) and $\mu(p)$ (seventh column). The table is for $p < 200$ and E; $y^2 = x^3 + 7$.

p	X_1	X_2	Y_1	Y_3	$\#E(\mathbb{F}_p)$	$\mu(p)$
5	1	2	5	0	6	1
11	1	5	11	0	12	1
13	0	3	0	2	7	1
17	1	8	17	0	18	1
19	3	4	2	3	12	2
23	1	11	23	0	24	1
29	1	14	29	0	30	1
31	0	10	2	6	21	1
37	0	19	2	12	39	1
41	1	20	41	0	42	1
43	0	15	0	10	31	1
47	1	23	47	0	48	1
53	1	26	53	0	54	1
59	1	29	59	0	60	1
61	0	30	0	20	61	1
67	0	39	0	26	79	1
71	1	35	71	0	72	1
73	3	30	0	21	64	8
79	0	33	0	22	67	1
83	1	41	83	0	84	1
89	1	44	89	0	90	1
97	0	39	0	26	79	1
101	1	50	101	0	102	1
103	0	55	2	36	111	1
107	1	53	107	0	108	1
109	0	64	2	42	129	1
113	1	56	113	0	114	1
127	0	63	0	42	127	1
131	1	65	131	0	132	1
137	1	68	137	0	138	1
139	0	73	2	48	147	1
149	1	74	149	0	150	1
151	0	66	0	44	133	1
157	3	84	0	57	172	2
163	0	69	0	46	139	1
167	1	83	167	0	168	1
173	1	86	173	0	174	1
179	1	89	179	0	180	1
181	3	102	0	69	208	4
191	1	95	191	0	192	1
193	0	109	2	72	219	1
197	1	98	197	0	198	1
199	0	94	2	62	189	3

In the table different categories can be distinguished. For instance, each time when $X_1 = 1$ then $Y_3 = 0$. The cases with $X_1 = 1$ and $Y_3 = 0$ belong to a categorie. For $E : y^2 = x^3 + 7$ over $\mathbb{F}_p$ we have the following five categories:

1. $X_1 = 1$ and $Y_3 = 0$
2. $X_1 = 3$ and $Y_1 = 2$
3. $X_1 = 3$ and $Y_1 = 0$
4. $X_1 = 0$ and $Y_1 = 2$
5. $X_1 = 0$ and $Y_1 = 0$.

4.2 Categories for $y^2 = x^3 + b \mod p$

For every b (integer of course) the elliptic curves $E : y^2 \cong x^3 + b \mod p$ can be divided in the same 5 categories as $E : y^2 \cong x^3 + 7 \mod p$. It should be noted that in general for $E : y^2 \cong x^3 + ax + b \mod p$ with $a \neq 0$ also categories do occur with $Y_2 \neq 0$. In the table below the category, the order $\#E(\mathbb{F}_p)$ and the ratio $\mu(p)$ are given for $b = 1, 2, ..., 8$ and $p = 5, 7, 11, 13, 17, 19, 23, 29, 31, 37$.

p \ b	1	2	3	4	5	6	7	8
5	1,6,1	1,6,1	1,6,1	1,6,1				
7	2,12,2	4,9,3	5,13,1	4,3,1	5,7,1	3,4,2		
11	1,12,1	1,12,1	1,12,1	1,12,1	1,12,1	1,12,1	1,12,1	1,12,1
13	2,12,2	5,19,1	4,9,3	4,21,1	3,16,4	5,7,1	5,7,1	3,16,4
17	1,18,1	1,18,1	1,18,1	1,18,1	1,18,1	1,18,1	1,18,1	1,18,1
19	2,12,2	5,13,1	5,13,1	4,21,1	4,27,3	4,21,1	2,12,2	3,28,2
23	1,24,1	1,24,1	1,24,1	1,24,1	1,24,1	1,24,1	1,24,1	1,24,1
29	1,30,1	1,30,1	1,30,1	1,30,1	1,30,1	1,30,1	1,30,1	1,30,1
31	2,36,6	2,36,6	5,43,1	2,36,6	4,39,1	5,43,1	4,21,1	2,36,6
37	2,48,4	5,49,1	4,39,1	4,39,1	5,37,1	3,28,2	4,39,1	3,28,2

For instance, for $p = 19$ and $b = 8$ we read of the numbers 3,28,2. This means that the points of $E : y^2 \cong x^3 + 8 \mod 19$ are in category 3, that $\#E_8(\mathbb{F}_{19}) = 28$ and that $\mu_8(19) = 2$. For $b \cong 0 \mod p$ the curve is singular and for $b > p$ the numbers can be read of at the column for $b \mod p$. This is the reason why for $p = 5$ and $p = 7$ the cells are left empty for $b \geq 5$ respectively $b \geq 7$.

We see that for primes of the type $p \cong 5 \mod 6$ the category is 1, $\#E(\mathbb{F}_p) = p + 1$ and $\mu = 1$. To save space we will confine to primes of the type $p \cong 1 \mod 6$, see the next table.

p \ b	1	2	3	4	5	6	7	8
7	2,12,2	4,9,3	5,13,1	4,3,1	5,7,1	3,4,2		
13	2,12,2	5,19,1	4,9,3	4,21,1	3,16,4	5,7,1	5,7,1	3,16,4
19	2,12,2	5,13,1	5,13,1	4,21,1	4,27,3	4,21,1	2,12,2	3,28,2
31	2,36,6	2,36,6	5,43,1	2,36,6	4,39,1	5,43,1	4,21,1	2,36,6
37	2,48,4	5,49,1	4,39,1	4,39,1	5,37,1	3,28,2	4,39,1	3,28,2
43	2,36,6	3,52,2	5,49,7	2,36,6	5,49,7	4,39,1	5,31,1	3,52,2
61	2,48,4	5,61,1	2,48,4	4,75,5	4,63,3	5,61,1	5,61,1	3,76,2
67	2,84,2	5,73,1	3,52,2	4,57,1	3,52,2	4,63,3	5,79,1	3,52,2
73	2,84,2	4,81,9	2,84,2	4,57,1	5,91,1	4,81,9	3,64,8	2,84,2
79	2,84,2	4,63,3	5,97,1	4,93,1	4,93,1	5,67,1	5,67,1	2,84,2
97	2,84,2	4,117,3	4,117,3	4,93,1	5,79,1	4,93,1	5,79,1	2,84,2
103	2,84,2	4,117,3	3,124,2	4,111,1	5,97,1	5,91,1	4,111,1	2,84,2
109	2,108,6	3,112,4	4,129,1	2,108,6	4,129,1	5,91,1	4,129,1	3,112,4
127	2,108,6	2,108,6	5,127,1	2,108,6	3,148,2	5,127,1	5,127,1	2,108,6
139	2,156,2	5,163,1	5,133,1	4,147,1	4,147,1	2,156,2	4,147,1	3,124,2
151	2,156,2	4,171,3	3,148,2	4,129,1	4,171,3	5,133,1	5,133,1	2,156,2
157	2,144,12	3,172,2	4,183,1	2,144,12	5,133,1	5,133,1	3,172,2	3,172,2
163	2,156,2	5,139,1	5,181,1	4,147,1	3,172,2	2,156,2	5,139,1	3,172,2
181	2,156,2	5,175,5	4,201,1	4,201,1	2,156,2	3,208,4	3,208,4	3,208,4
193	2,192,8	4,171,3	2,192,8	4,219,1	5,217,1	4,171,3	4,219,1	2,192,8
199	2,228,2	4,189,3	5,211,1	4,183,1	2,228,2	5,217,1	4,189,3	2,228,2
211	2,228,2	5,199,1	5,199,1	4,183,1	2,228,2	4,183,1	5,199,1	3,196,14
223	2,252,6	2,252,6	5,247,1	2,252,6	5,229,1	5,247,1	2,252,6	2,252,6
229	2,252,6	3,208,4	4,237,1	2,252,6	4,201,1	5,223,1	5,259,1	3,208,4
241	2,228,2	4,225,15	4,273,1	4,273,1	2,228,2	2,228,2	5,259,1	2,228,2
271	2,300,10	4,243,9	3,244,2	4,273,1	4,273,1	5,301,1	4,243,9	2,300,10
277	2,252,6	3,304,4	4,309,1	2,252,6	5,283,1	5,247,1	4,273,1	3,304,4
283	2,252,6	3,316,2	5,277,1	2,252,6	5,259,1	4,291,1	4,309,1	3,316,2
307	2,324,18	3,292,2	3,292,2	2,324,18	5,343,1	2,324,18	4,327,1	3,292,2
313	2,336,4	4,279,3	4,327,1	4,327,1	3,292,2	2,336,4	3,292,2	2,336,4
331	2,300,10	5,331,1	5,331,1	4,363,11	4,363,11	4,363,11	3,364,2	3,364,2
337	2,372,2	4,333,3	4,309,1	4,309,1	3,304,4	2,372,2	2,372,2	2,372,2
349	2,336,4	5,313,1	4,327,1	4,327,1	4,327,1	3,364,2	5,313,1	3,364,2
349	2,336,4	5,313,1	4,327,1	4,327,1	4,327,1	3,364,2	5,313,1	3,364,2
367	2,372,2	4,333,3	3,364,2	4,399,1	3,364,2	5,403,1	2,372,2	2,372,2
373	2,336,4	5,361,1	4,387,3	4,399,1	5,361,1	5,349,1	2,336,4	3,412,2
379	2,372,2	5,409,1	5,343,7	4,417,1	2,372,2	2,372,2	5,343,7	3,388,2

Of course the table is just a small part of what is intended to show. To the right the rows should be thought to run through $b = p-1$. An order may occur more than once. For example $\#E(\mathbb{F}_p) = 273$ for (p, b) equal to $(241, 3)$, $(241, 4)$, $(271, 4)$, $(271, 5)$ and $(277, 7)$. However, it is always accompanied by the same category and the same μ. This suggests that $\#E(\mathbb{F}_p)$ uniquely determines the category and μ.

By inspection of the tables we make the following observation: μ is a divisor of $p-1$. Since μ is a divisor of $\#E(\mathbb{F}_p)$, it also is a divisor of $\gcd(p-1, \#E(\mathbb{F}_p))$. This limits the values μ can possibly take on. If we apply it to, for instance, $p = 223$, then $\#E(\mathbb{F}_p) = 252$ for $E: y^2 = x^3 + 7$. Since $\gcd(222, 252) = 6$ the value of μ is 1,2,3 or 6. For this case $\mu = 6$.

4.3 Characteristics of categories

Category 1: $X_1 = 1$ and $Y_3 = 0$.
Characteristics category 1: $p \cong 5 \mod 6$), $\#E(\mathbb{F}_p) = p + 1$, $\mu = 1$, $X_2 \cong 2 \mod 3$) and $Y_1 \cong 5 \mod 6$. The set of all y values is $\{0, 1, 2, ..., p-1\}$ and ∞ (for point $\mathcal{O}$). Since $\mu = 1$ the points on $E(\mathbb{F}_p)$ are cyclic of order $p + 1$. Elliptic curves for which $\#E(\mathbb{F}_p) = p + 1$ are called *supersingular* (although it has nothing to do with a singularity). So, category 1 is the supersingular category.

In general, the order of a modular elliptic curve can be written as $\#E(\mathbb{F}_p) = p + 1 - d$. According to a theorem of Hasse $|d| \leq 2\sqrt{p}$. Thus $d = 0$ for category 1. For $E: y^2 = x^3 + b$ the cases with $d \neq 0$ occur for $p = 1 \mod 6$. This is the situation for categories 2 through 5.

Category 2: $X_1 = 3$ and $Y_1 = 2$.
Characteristics category 2: $p \cong 1 \mod 6$, $\#E(\mathbb{F}_p) \cong 0 \mod 12$, $\mu \cong 0 \mod 2$, $X_2 \cong 4 \mod 6$ and $Y_3 \cong 1 \mod 2$.

Category 3: $X_1 = 3$ and $Y_1 = 0$.
Characteristics category 3: $p \cong 1 \mod 6$, $\#E(\mathbb{F}_p) \cong 4 \mod 12$, $\mu = 0 \mod 2$, $X_2 \cong 0 \mod 6$ and $Y_3 \cong 1 \mod 2$.

Category 4: $X_1 = 0$ and $Y_1 = 2$.
Characteristics category 4: $p \cong 1 \mod 6$, $\#E(\mathbb{F}_p) \cong 3 \mod 6$, $\mu = 1 \mod 2$, $X_1 \cong 0 \mod 3$ and $Y_3 \cong 2 \mod 6$.

Category 5: $X_1 = 0$ and $Y_1 = 0$.
Characteristics category 5: $p \cong 1 \mod 6$, $\#E(\mathbb{F}_p) \cong 1 \mod 6$, $\mu = 1 \mod 2$, $X_2 \cong 0 \mod 3$ and $Y_3 \cong 0 \mod 2$.

4.4 Limitations for $E_{\mathbf{TORS}}(\mathbb{Q})$.

The elliptic curve $E: y^2 = x^3 + 1$ has 5 integer points: $(-1,0)$, $(0,1)$, $(0,-1)$, $(2,3)$ and $(2,-3)$ which form (together with $\mathcal{O}$) a torsion group of order 6: start with $P = (2,3)$ then $2P = (0,1)$ (a point of inflection), $3P = (-1,0)$ (the root), $4P = (0,-1)$, $5P = (2,-3)$ and $6P = \mathcal{O}$. There are no other rational points; the rank is 0.

If we start with $P = (2,3)$ on the modular curve $E: y^2 = x^3 + 1 \mod 5$ then $2P = (0,1)$, $3P = (4,0)$, $4P = (0,4)$, $5P = (2,2)$ and $6P = \mathcal{O}$. If we start with $P = (2,3)$ on the modular curve $E: y^2 = x^3 + 1 \mod 7$ then $2P = (0,1)$, $3P = (6,0)$, $4P = (0,6)$, $5P = (2,4)$ and $6P = \mathcal{O}$. The two examples illustrate that a torsion group on an elliptic curve E is present in $E(\mathbb{F}_p)$, except for a change of the coordinates because of the modulair counting. As a consequence the order of the torsion group is a divisor of the order of the modular group: $\#E_{\mathrm{TORS}}(\mathbb{Q})$ divides $\#E(\mathbb{F}_p)$. Since the latter holds for any p it holds for the greatest common divisor of different $\#E(\mathbb{F}_p)$. For $E: y^2 = x^3 + 1$ we have $\#E(\mathbb{F}_5) = 6$ and $\#E(\mathbb{F}_7) = 12$. Since $\gcd(6,12) = 6$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q})$ has to be a divisor of 6. There are no other rational points, so the rank is 0. Below follow some additional examples.

For the elliptic curve $E: y^2 = x^3 + 2$ we have $\#E(\mathbb{F}_7) = 9$ and $\#E(\mathbb{F}_{13}) = 19$. Since $\gcd(9,19) = 1$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q}) = 1$. The elliptic curve $E: y^2 = x^3 + 2$ has $(-1,1)$ and $(-1,-1)$ as integer points. Both generate an infinite sequence of rationals. In conclusion, $\mathcal{O}$ is the single torsion point and the rank is 1.

For the elliptic curve $E: y^2 = x^3 + 3$ we have $\#E(\mathbb{F}_7) = 13$ and $\#E(\mathbb{F}_{13}) = 9$. Since $\gcd(13,9) = 1$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q}) = 1$. The elliptic curve $E: y^2 = x^3 + 3$ has $(1,2)$ and $(1,-2)$ as integer points. Both generate an infinite sequence of rationals; the rank is 1.

For the elliptic curve $E: y^2 = x^3 + 4$ we have for $p \cong 5 \mod 6$: $\#E(\mathbb{F}_p) = p + 1 \cong 0 \mod 6$ and for $p \cong 1 \mod 6$ we see from the tables that $\#E(\mathbb{F}_p) \cong 0 \mod 3$. Since $\gcd(6,3) = 3$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q})$ divides 3. The elliptic curve $E: y^2 = x^3 + 4$ has $(0,2)$ and $(0,-2)$ as integer points. It are the points of inflection and form an integer torsion group of order 3. There are no other rationals; the rank is 0.

For the elliptic curve $E: y^2 = x^3 + 5$ we have $\#E(\mathbb{F}_{13}) = 16$ and $\#E(\mathbb{F}_{19}) = 27$. Since $\gcd(16,27) = 1$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q}) = 1$. The elliptic curve $E: y^2 = x^3 + 5$ has $(-1,2)$ and $(-1,-2)$ as integer points. Either one of them generates an infinite sequence of rationals; the rank is 1.

For the elliptic curve $E: y^2 = x^3 + 6$ we have $\#E(\mathbb{F}_7) = 4$ and $\#E(\mathbb{F}_{13}) = 7$. Since

$\gcd(4,7) = 1$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q}) = 1$. The elliptic curve $E: y^2 = x^3 + 6$ has no rational points; the rank is 0.

For the elliptic curve $E: y^2 = x^3 + 7$ we have $\#E(\mathbb{F}_{13}) = 7$ and $\#E(\mathbb{F}_{19}) = 12$. Since $\gcd(7,12) = 1$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q}) = 1$. As mentioned before, the elliptic curve $E: y^2 = x^3 + 7$ has no rational points; the rank is 0.

As a final example, for the elliptic curve $E: y^2 = x^3 + 8$ we have for $p \cong 5 \mod 6$: $\#E(\mathbb{F}_p) = p + 1 \cong 0 \mod 6$ and for $p \cong 1 \mod 6$ we see from the tables that $\#E(\mathbb{F}_p) \cong 0 \mod 2$. Since $\gcd(6,2) = 2$ it follows that $\#E_{\mathrm{TORS}}(\mathbb{Q})$ divides 2. The elliptic curve $E: y^2 = x^3 + 8$ has 1 integer point $(-2,0)$ of order 2 and 6 integer points, $(1,3)$, $(1,-3)$, $(2,4)$, $(2,-4)$, $(46,312)$ and $(46,-312)$, which generate infinite sequences of rationals. The latter 6 integer points follow from addition of the torsion point $(-2,0)$ and the integer point $(1,3)$. In conclusion, $\#E_{\mathrm{TORS}}(\mathbb{Q}) = 2$ and the rank is 1.

In the foregoing examples we took the $\gcd(\#E(\mathbb{F}_{p_1}), \#E(\mathbb{F}_{p_2}))$ for two different primes p_1 and p_2. In general, one should consider more primes to obtain the smallest gcd. For the elliptic curve $E: y^2 = x^3 + b$ the rank r, $\#E_{\mathrm{TORS}}(\mathbb{Q})$ and the smallest gcd are tabulated for $b = 1$ through 30, see next table, where t is an abbreviation for $\#E_{\mathrm{TORS}}(\mathbb{Q})$.

b	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
r	0	1	1	0	1	0	0	1	1	1	1	1	0	0	2	0	2	1	1	0	0	1	0	2	0	1	0	1	0	1
t	6	1	1	3	1	1	1	2	3	1	1	1	1	1	1	3	1	1	1	1	1	1	1	1	3	1	2	1	1	1
gcd	6	1	1	3	1	1	1	2	3	1	1	1	1	1	1	3	1	1	1	1	1	1	1	1	3	1	2	1	1	1

If one takes the gcd of the set $\{\#E(\mathbb{F}_p)\}$ for a sufficiently large number of primes p, then:

$$t = \gcd = \begin{cases} 6 & \text{if } b \text{ is a sixth power} \\ 3 & \text{if } b \text{ is a square} \\ 2 & \text{if } b \text{ is a cube} \\ 1 & \text{otherwise} \end{cases}$$

4.5 Polynomials for $y^2 = x^3 + ax + b$ over $\mathbb{F}_p$

The elliptic curve $y^2 = x^3 + ax \mod p$ can be investigated in a similar way as for $y^2 = x^3 + b \mod p$. One of the observations will be that $y^2 = x^3 + ax \mod p$ is supersingular if $p \cong 3$

mod 4. In this section our concern will be the more general equation $y^2 = x^3 + ax + b \mod p$. Usually it is written as $y^2 = x^3 + ax + b$ over $\mathbb{F}_p$ or as $E: y^2 = x^3 + ax + b$ over $\mathbb{F}_p$. For the present purpose it is convenient to use the following notation: $E(\mathbb{F}_p, a, b)$ is the elliptic curve $y^2 = x^3 + ax + b$ over $\mathbb{F}_p$. The group order, usually denoted as $\#E(\mathbb{F}_p)$, will be denoted as $N(p, a, b)$. The difference between $N(p, a, b)$ and $p + 1$ will be denoted as $d(p, a, b)$:

$$N(p, a, b) = p + 1 + d(p, a, b)\,. \tag{4.1}$$

The differences $d(p, a, b)$ are congruent (mod p) with a two dimensional polynomial P in a and b: $d(p, a, b) \cong P(p, a, b) \mod p$. Thus $N(p, a, b) \cong 1 + P(p, a, b) \mod p$. For the first few prime numbers the polynomials $P(p, a, b)$ are shown in the table.

p	$P(p, a, b)$
5	$3a$
7	$4b$
11	$2ab$
13	$6a^3 + 11b^2$
17	$15a^4 + 2ab^2$
19	$9a^3b + 11b^3$
23	$13a^4b + 14ab^3$
29	$19a^7 + 21a^4b^2 + 24ab^4$
31	$29a^6b + 30a^3b^3 + 4b^5$
37	$35a^9 + 31a^6b^2 + 8a^3b^4 + 10b^6$
41	$31a^{10} + 38a^7b^2 + 11a^4b^4 + 36ab^6$
43	$a^9b + 7a^6b^3 + 33a^3b^5 + 35b^7$
47	$10a^{10}b + 20a^7b^3 + 15a^4b^5 + 14ab^7$
53	$14a^{13} + 10a^{10}b^2 + 40a^7b^4 + 44a^4b^6 + 2ab^8$
59	$26a^{13}b + 37a^{10}b^3 + 2a^7b^5 + 53a^4b^7 + 3ab^9$
61	$51a^{15} + 39a^{12}b^2 + 55a^9b^4 + 31a^6b^6 + 9a^3b^8 + 47b^{10}$
67	$23a^{15}b + 64a^{12}b^3 + 46a^9b^5 + 41a^6b^7 + 66a^3b^9 + 16b^{11}$
71	$10a^{16}b + 7a^{13}b^3 + 35a^{10}b^5 + 59a^7b^7 + 21a^4b^9 + 68ab^{11}$
73	$6a^{18} + 20a^{15}b^2 + 45a^{12}b^4 + 63a^9b^6 + 49a^6b^8 + 24a^3b^{10} + 10b^{12}$

Table 4.1: Polynomials $P(p, a, b)$ for which $N(p, a, b) \cong 1 + P(p, a, b) \mod p$.

According to a theorem of Hasse there holds

$$|d(p,a,b)| \leq 2\sqrt{p}\,. \tag{4.2}$$

As a consequence, the polynomials uniquely determine the group order $N(p,a,b)$ if $p > 4\sqrt{p}$ $\rightarrow p > 16$. For example, if we want to know the group order for $y^2 = x^3 + 5x + 7$ over $\mathbb{F}_{19}$, we obtain $P(19,5,7) = 9{\cdot}5^3{\cdot}7{+}11{\cdot}7^3 = 11648$, $N(19,5,7) \cong 1{+}11648 \cong 2 \mod 19$. Of the numbers $\{..., -17, 2, 21, 40, 59, ...\}$ only 21 satisfies Hasse's theorem. Therefore $N(19,5,7) = 21$.

The powers n and m of the terms $a^n b^m$ in the polynomials have the property $4n+6m = p-1$. We can write the polynomials as

$$P(p,a,b) = \begin{cases} \sum_{k=0}^{m} c_k a^{\frac{p-1}{4}-3k} b^{2k} & \text{if } p \cong 1 \mod 4 \\ \sum_{k=0}^{m} c_k a^{\frac{p-3}{4}-3k-1} b^{2k+1} & \text{if } p \cong 3 \mod 4\,, \end{cases} \tag{4.3}$$

where $m = \dfrac{p - (p \mod 12)}{12} = \lfloor \dfrac{p}{12} \rfloor$.

For $b = 0$ the latter polynomials are reduced to

$$P(p,a,0) = \begin{cases} c_0 a^{\frac{p-1}{4}} & \text{if } p \cong 1 \mod 4 \\ 0 & \text{if } p \cong 3 \mod 4 \end{cases} \tag{4.4}$$

We can write the polynomials also as

$$P(p,a,b) = \begin{cases} \sum_{k=0}^{m} c_{m-k} a^{3k} b^{\frac{p-1}{6}-2k} & \text{if } p \cong 1 \mod 6 \\ \sum_{k=0}^{m} c_{m-k} a^{3k+1} b^{\frac{p-5}{6}-2k} & \text{if } p \cong 5 \mod 6\,. \end{cases} \tag{4.5}$$

Also here $m = \dfrac{p - (p \mod 12)}{12} = \lfloor \dfrac{p}{12} \rfloor$.

For $a = 0$ the latter polynomials are reduced to

$$P(p,0,b) = \begin{cases} c_m b^{\frac{p-1}{6}} & \text{if } p \cong 1 \mod 6 \\ 0 & \text{if } p \cong 5 \mod 6 \end{cases} \tag{4.6}$$

For the supersingular case $b \cong 0 \mod p$ and $p \cong 3 \mod 4$ we have $P(p,a,0) = 0$. For the supersingular case $a \cong 0 \mod p$ and $p \cong 5 \mod 6$ we have $P(p,0,b) = 0$. For the singular case $a \cong b \cong 0 \mod p$ we have $P(p,0,0) = 0$. For these cases the group order is equal to $p{+}1$.

For $b = a$ the polynomials reduce to

$$P(p, a, a) = \sum_{k=0}^{m} c_k a^{\lambda - k} \tag{4.7}$$

where $m = \dfrac{p - (p \mod 12)}{12} = \lfloor \dfrac{p}{12} \rfloor$ and $\lambda = \dfrac{p - (p \mod 4)}{4} = \lfloor \dfrac{p}{4} \rfloor$.

The coefficients c_k of the polynomials can be obtained by solving the following system of equations

$$N(p, a, a) \cong 1 + \sum_{k=0}^{m} c_k a^{\lambda - k} \mod p \,, \quad a = 1, 2, ..., m + 1 \,. \tag{4.8}$$

The group orders $N(p, a, a)$ in the system equations are calculated numerically.

Patterns are present in the coefficients c_k. Every prime p satisfying $p \cong 1 \mod 4$ can be written uniquely as $p = (\pm 2n)^2 + (\pm m)^2$ with n and m integers. Taking the signs such that $\pm(2n) \pm m \cong 1 \mod 4$ then the group order follows from (see page 115 of the book of Washington [1]):

$$N(p, a, 0) = \begin{cases} p + 1 - 2m, & \text{if } p - a \text{ is a fourth power mod } p \\ p + 1 + 2m, & \text{if } p - a \text{ is a square but not a fourth power mod } p \\ p \pm 4n, & \text{otherwise.} \end{cases} \tag{4.9}$$

For $p = 17$, for example, the squares $\mod 17$ are $\{1, 4, 9, 16, 8, 2, 15, 13\}$. The fourth powers mod 17 are $\{1, 16, 13, 4\}$. The squares which are not a fourth power are $\{9, 8, 2, 15\}$. Since $17 = 4^2 + 1^2$ we have

$$N(17, a, 0) = \begin{cases} 17 + 1 - 2 = 16, & \text{if } p - a \mod p \in \{1, 4, 13, 16\} \\ 17 + 1 + 2 = 20, & \text{if } p - a \mod p \in \{2, 8, 9, 15\} \\ 17 + 1 \pm 8, & \text{if } p - a \mod p \in \{3, 5, 6, 7, 10, 11, 12, 14\}. \end{cases} \tag{4.10}$$

For $b = 0$ the polynomial for $p = 17$ reduces to $P(17, a, 0) = 15a^4$, which also leads to $N(17, a, 0) = 16$ for $a = 1, 4, 13, 16$, to $N(17, a, 0) = 20$ for $a = 2, 8, 9, 15$, to $N(17, a, 0) = 26$ for $a = 3, 5, 12, 14$ and to $N(17, a, 0) = 10$ for $a = 6, 7, 10, 11$. A single value for a is sufficient to derive the coefficient c_0 from the group order theorem. For $a = 1$ it means that

$$c_0 = \begin{cases} -2m, & \text{if } p - 1 \text{ is a fourth power mod } p \\ +2m, & \text{if } p - 1 \text{ is a square but not a fourth power mod } p \\ \pm 4n, & \text{otherwise.} \end{cases} \tag{4.11}$$

If we apply the latter to for example $p = 29 = 2^2 + (-5)^2$, we find that $p - 1 = 28$ is a square: $12^2 \cong 28 \mod 29$ and that $p - 1 = 28$ is not a fourth power. Hence, $c_0 = 2m = 2 \cdot -5 =$

$-10 \cong 19 \mod p$. You can apply it yourself to the c_0 for $p = 5, 13, 37, 41$, etc.

As another example we consider a pattern in some of the coefficients guiding the terms $b^{\frac{p-1}{6}}$. If $p \cong 1 \mod 6$ then for the coefficient c_m of the term $c_m b^{\frac{p-1}{6}}$ there holds:
$c_m \cong 4 \mod p$ if the prime p has the form $12k^2 + 12k + 7$ with $k = 0, 1, 2, 3, ...$
$c_m \cong 10 \mod p$ if $p = 12k^2 + 25$ with $k = 0, 1, 2, 3, ...$
$c_m \cong 16 \mod p$ if $p = 12k^2 + 12k + 67$ with $k = 0, 1, 2, 3, ...$
$c_m \cong 22 \mod p$ if $p = 12k^2 + 121$ with $k = 0, 1, 2, 3, ...$
$c_m \cong 28 \mod p$ if $p = 12k^2 + 12k + 199$ with $k = 0, 1, 2, 3, ...$
$c_m \cong 34 \mod p$ if $p = 12k^2 + 289$ with $k = 0, 1, 2, 3, ...$
$c_m \cong 40 \mod p$ if $p = 12k^2 + 12k + 403$ with $k = 0, 1, 2, 3, ...$
$c_m \cong 46 \mod p$ if $p = 12k^2 + 529$ with $k = 0, 1, 2, 3, ...,$
and so on.
If $p \cong 1 \mod 12$ the pattern can be summarized as
$c_m \cong 10 + 12n \mod p$ if $p = 12k^2 + (6n + 5)^2$ with $k = 0, 1, 2, 3, ...,$
and if $p \cong 7 \mod 12$ the pattern can be summarized as
$c_m \cong 4 + 12n \mod p$ if $p = 12k^2 + 12k + (6n + 2)^2 + 3$ with $k = 0, 1, 2, 3, ...$
If we apply the latter to for example $p = 73 = 12 \cdot 2^2 + (6 \cdot 0 + 5)^2$, we find $c_m = c_6 = 10$. As a consequence, $N(73, 0, b) \cong 1 + 10b^{12} \mod 73$. Thus $N(73, 0, 1) \cong 1 + 10 \mod 73 = ..., 11, 84, 157, ...$ of which 84 is within the Hasse bounds; $N(73, 0, 1) = 84$. And $N(73, 0, 2) \cong 1 + 10 \cdot 2^1 2 \mod 73 = ..., 8, 81, 154, ...$ of which 81 is within the Hasse bounds; $N(73, 0, 2) = 81$. You can apply it yourself to the c_m for $p = 7, 31, 37, 67, 73$, etc.

4.6 Congruence relations for $N(p, a, 0)$ and $N(p, 0, b)$

Group orders are usually determined by numerical methods. A basic method consists in making a list of squares of the numbers 1 through $(p-1)/2$. Start with $n = 0$. Increase n by 2 for every $0 \leq x < p$ for which $x^3 + ax + b \mod p$ is an element of the list. Increase n by 1 for every $0 \leq x < p$ for which $x^3 + ax + b \cong 0 \mod p$. When you are finished $N(p, a, b) = n$. The method is slow. Faster methods are more complicated. A fast method is Schoof's algorithm. Very briefly, Schoof's algorithm consists in finding a point of which the order is larger than $4\sqrt{p}$. Then there is only one value for the group order (which is a multiple of the order of the point) satisfying Hasse's theorem. Since the evaluation of approximately $p/12$ expressions is time consuming the polynomials are not of practical use. An exception occurs for the case $b = 0$ and the case $a = 0$.

We first consider the case $b = 0$. From the equation (4.4) we obtain the following congruence relation:

$$d(p, a, 0) \cdot (a')^{\frac{p-1}{4}} \cong d(p, a', 0) \cdot a^{\frac{p-1}{4}} \mod p\,. \tag{4.12}$$

For p is a prime number the substitution of $a' = p - a$ in eq. (4.12) yields

$$d(p,a,0) = d(p,p-a,0)\,,\ \text{if } p \cong 1 \mod 8 \tag{4.13}$$

and

$$d(p,a,0) + d(p,p-a,0) = 0\,,\ \text{if } p \cong 5 \mod 8\,. \tag{4.14}$$

Alternatively,

$$N(p,a,0) = N(p,p-a,0)\,,\ \text{if } p \cong 1 \mod 8 \tag{4.15}$$

and

$$N(p,a,0) + N(p,p-a,0) = 2(p+1)\,,\ \text{if } p \cong 5 \mod 8\,. \tag{4.16}$$

From eq. (4.12) it follows

$$(N(p,a,0)-1)\cdot(a')^{\frac{p-1}{4}} \cong (N(p,1,0)-1)\cdot a^{\frac{p-1}{4}} \mod p\,. \tag{4.17}$$

That is, if $N(p,a',0)$ has been found numerically then one obtains $N(p,a,0)$ almost instantly. We only have to check the two smallest values larger than zero. As an example we consider $p = 53$ for which $N(53,1,0) = 68$. Suppose we want to know $N(53,11,0)$. From eq. (4.17) it follows $N(53,11,0)-1 \cong 67\cdot 11^{13} \cong 39 \mod 53$. From the two smallest values for $N(53,11,0)$, 40 and 93, the first satisfies Hasse's theorem. Hence, $N(53,11,0) = 40$.

Next we consider the case $a = 0$. From the equation (4.6) we obtain the following congruence relation:

$$d(p,0,b)\cdot(b')^{\frac{p-1}{6}} \cong d(p,0,b')\cdot b^{\frac{p-1}{6}} \mod p\,. \tag{4.18}$$

In particular for $b' = 1$ it is reduced to

$$d(p,0,b) \cong d(p,0,1)\cdot b^{\frac{p-1}{6}} \mod p\,. \tag{4.19}$$

Since p is a prime number the substitution of $b' = p - b$ in eq. (4.18) yields

$$d(p,0,b) = d(p,0,p-b)\,,\ \text{if } p \cong 1 \mod 12 \tag{4.20}$$

and

$$d(p,0,b) + d(p,0,p-b) = 0\,,\ \text{if } p \cong 7 \mod 12\,. \tag{4.21}$$

Alternatively,

$$N(p,0,b) = N(p,0,p-b)\,,\ \text{if } p \cong 1 \mod 12 \tag{4.22}$$

and

$$N(p,0,b) + N(p,0,p-b) = 2(p+1)\,,\ \text{if } p \cong 7 \mod 12\,. \tag{4.23}$$

From eq. (4.18) it follows

$$(N(p,0,b)-1)\cdot(b')^{\frac{p-1}{6}} \cong (N(p,0,1)-1)\cdot b^{\frac{p-1}{6}} \mod p\,. \tag{4.24}$$

If $N(p,0,b')$ has been found numerically then one obtains $N(p,0,b)$ almost instantly. We only have to check the two smallest values larger than zero. As an example we consider $p = 67$ for which $N(67,0,1) = 84$. Suppose we want to know $N(67,0,12)$. From eq. (4.24) it follows $N(67,0,12) - 1 \cong 83 \cdot 12^{11} \cong 11 \mod 67$. From the two smallest values for $N(67,0,12)$, 12 and 79, the second satisfies Hasse's theorem. Hence, $N(67,0,12) = 79$.

Congruence relations related to the ones given before are

$$N(p \cong 1 \mod 8, a, 0) \cong \begin{cases} 0 \mod 8 & \text{if } a \text{ is a fourth power} \\ 0 \mod 8 & \text{if } a \text{ is 4 times a fourth power} \\ 0,4 \mod 8 & \text{if } a \text{ is a square} \\ 0,4 \mod 8 & \text{if } a \text{ is 2 times a square} \\ 0,2,4 \mod 8 & \text{otherwise}, \end{cases}$$

$$N(p \cong 5 \mod 8, a, 0) \cong \begin{cases} 0 \mod 8 & \text{if } a \text{ is 4 times a fourth power} \\ 4 \mod 8 & \text{if } a \text{ is a fourth power} \\ 2 \mod 8 & \text{if } a \text{ is 2 times a square} \\ 0,4 \mod 8 & \text{if } a \text{ is a square} \\ 0,2,4 \mod 8 & \text{otherwise} \end{cases}$$

and

$$N(p \cong 1 \mod 6, 0, b) \cong \begin{cases} 0 \mod 12 & \text{if } b \text{ is a sixth power} \\ 0,4 \mod 12 & \text{if } b \text{ is a cube} \\ 0,3,9 \mod 12 & \text{if } b \text{ is a square} \\ 0,1,3,4,7,9 \mod 12 & \text{otherwise}. \end{cases}$$

4.7 Moments for $N(p,a,0)$ and $N(p,0,b)$

For $p \cong 5 \mod 8$ a consequence of eq. (4.16) is

$$\sum_{a=1}^{p-1} N(p,a,0) = (p-1)(p+1) = p^2 - 1 \tag{4.25}$$

By inspection it is found it also holds for $p \cong 1 \mod 8$. Therefore the identity

$$\sum_{a=1}^{p-1} N(p,a,0) = p^2 - 1 \tag{4.26}$$

holds for all $p \cong 1 \mod 4$.

The k-th moment $N(p,a,0)$ is defined as

$$\sum_{a=1}^{p-1} N^k(p,a,0) \tag{4.27}$$

For the supersingular case, $p \cong 3 \mod 4$, there holds for all k:

$$\sum_{a=1}^{p-1} N^k(p,a,0) = (p-1)(p+1)^k \,. \tag{4.28}$$

For $k \geq 0$ this can also be written as

$$\sum_{a=1}^{p-1} N^k(p,a,0) = (p-1)\sum_{j=0}^{j=k} \binom{k}{j} p^j \,. \tag{4.29}$$

For $p \cong 1 \mod 6$ we found by inspection for $0 \leq k \leq 3$:

$$\sum_{a=1}^{p-1} N^k(p,a,0) = (p-1)\sum_{j=0}^{j=k} \binom{k}{j}^2 p^j \tag{4.30}$$

For $k \geq 4$ the latter identity is violated. However, for $k \geq 4$ we have

$$\sum_{a=1}^{p-1} N^k(p,a,0) = (p-1)\left(\delta(k,p) + \sum_{j=0}^{j=k} \binom{k}{j}^2 p^j\right) \,, \tag{4.31}$$

where the integer $\delta(k,p)$ is the deviation. That is, $p-1$ still is a divisor of $\sum_{a=1}^{p-1} N^k(p,a,0)$.

For $p \cong 7 \mod 12$ a consequence of eq. (4.23) is

$$\sum_{b=1}^{p-1} N(p,0,b) = (p-1)(p+1) = p^2 - 1 \,. \tag{4.32}$$

From inspection it is found that it also holds for $p \cong 1 \mod 12$. Therefore, the identity

$$\sum_{b=1}^{p-1} N(p,0,b) = p^2 - 1 \,. \tag{4.33}$$

holds for all $p \cong 1 \mod 6$.

The k-th momentum of $N(p,0,b)$ is defined as

$$\sum_{b=1}^{p-1} N^k(p,0,b) \,. \tag{4.34}$$

For the supersingular case, $p \cong 5 \mod 6$, there holds for all k:

$$\sum_{b=1}^{p-1} N^k(p,0,b) = (p-1)(p+1)^k \,. \tag{4.35}$$

For $k \geq 0$ this can also be written as

$$\sum_{b=1}^{p-1} N^k(p,0,b) = (p-1)\sum_{j=0}^{j=k} \binom{k}{j} p^j \,. \tag{4.36}$$

For primes $p \cong 1 \mod 6$ we found by inspection for $0 \leq k \leq 5$:

$$\sum_{b=1}^{p-1} N^k(p,0,b) = (p-1)\sum_{j=0}^{j=k} \binom{k}{j}^2 p^j \tag{4.37}$$

For $k \geq 6$ the latter identity is violated. However, for $k \geq 6$ we have

$$\sum_{b=1}^{p-1} N^k(p,0,b) = (p-1)\left(\delta(k,p) + \sum_{j=0}^{j=k} \binom{k}{j}^2 p^j\right) \,, \tag{4.38}$$

where the integer $\delta(k,p)$ is the deviation. That is, $p-1$ is still a divisor of $\sum_{b=1}^{p-1} N^k(p,0,b)$.

4.8 Generating function

For the determination of the order of the elliptic curve $Y^2 - Y = X^3 - X^2$ over the field $\mathcal{F}_p$ with p a prime, one can apply the following infinite product:

$$F(q) = q\prod_{m=1}^{\infty} (1-q^m)^2 \left(1-q^{11m}\right)^2 \,. \tag{4.39}$$

Expansion of the product and elimination. of the brackets leads to

$$\begin{aligned} F(q) &= q\,(1-q)^2\left(1-q^{11}\right)^2\left(1-q^2\right)^2\left(1-q^{22}\right)^2\left(1-q^3\right)^2\left(1-q^{33}\right)^2 ... \\ &= q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7 - 2q^9 - 2q^{10} + q^{11} - 2q^{12} \\ &\quad +4q^{13} + 4q^{14} - q^{15} - 4q^{16} - 2q^{17} + 4^{18} + ... \end{aligned} \tag{4.40}$$

We can write

$$F(q) = \sum_{n=1}^{\infty} d_n q^n \,. \tag{4.41}$$

The sequence of successive coefficients d_n is the sequence A006571 of the OEIS [6]. The coefficient d_p is the deviation of the order from $p+1$:

$$\#E(\mathbb{F}_p) = p + 1 - b_p \,, \tag{4.42}$$

comparable with the equation (4.1).
Since it generates the series with coefficients d_p the product equation (4.39) is the generating function for the order of $E : Y^2 - Y = X^3 - X^2$ over $\mathcal{F}_p$. The equation (4.39) has been derived from the theory of *modular forms* [7].

We give three examples and compare things with the order determination on the basis of the polynomial $P(p, a, b)$.

Example 1
To find the order of $E : Y^2 - Y = X^3 - X^2$ over the field $\mathbb{F}_{13}$ one takes the coefficient $d_{13} = 4$ and subtract it from $p + 1$. The result is $\#E(\mathbb{F}_{13}) = 13 + 1 - 4 = 10$. We thus obtain 10 for the order of the elliptic curve $Y^2 - Y = X^3 - X^2$ over $\mathbb{F}_{13}$. The 10 points are $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(2, 6)$, $(2, 8)$, $(8, 3)$, $(8, 11)$, $(10, 7)$ and infinity $\mathcal{O}$.
To find the order we could also have used the polynomial $P(13, a, b) = 6a^3 + 11b^2$ from the table in section 4.5. To this end we have perform a linear transform of the elliptic curve $Y^2 - Y = X^3 - X^2$ to the Weierstrass form $y^2 = x^3 + ax + b$. By means of the substitution

$$X = \frac{x + 12}{36} \quad , \qquad Y = \frac{y + 108}{216} \tag{4.43}$$

the equation $Y^2 - Y = X^3 - X^2$ is transformed to the Weierstrass form

$$y^2 = x^3 - 432x + 8208\,. \tag{4.44}$$

For $a = -432$ and $b = 8208$ the discriminant is $D = -2^8 3^{12} 11$. Over $\mathbb{F}_{13}$ the constants a and b are reduced to $a = -432 \cong 10 \mod 13$ and $b = 8208 \cong 5 \mod 13$. Substitution of the latter in $P(13, a, b) = 6a^3 + 11b^2$ gives $P(13, a, b) \cong 9 \mod 13$. Of the set $\{..., -17, -4, 9, 22, 35, ...\}$ only 9 satisfies the Hasse bounds. In this way we also obtain 10 for the order of $E : y^2 = x^3 - 432x + 8208$ over $\mathbb{F}_{13}$. Now, the 10 points are $(1, 4)$, $(1, 9)$, $(3, 6)$, $(3, 7)$, $(8, 5)$, $(8, 8)$, $(10, 0)$, $(11, 4)$, $(11, 9)$ and infinity $\mathcal{O}$. The (x, y) coordinates on $E : y^2 = x^3 - 432x + 8208$ over $\mathbb{F}_{13}$ are related to the (X, Y) coordinates of $E : Y^2 - Y = X^3 - X^2$ over $\mathbb{F}_{13}$ via

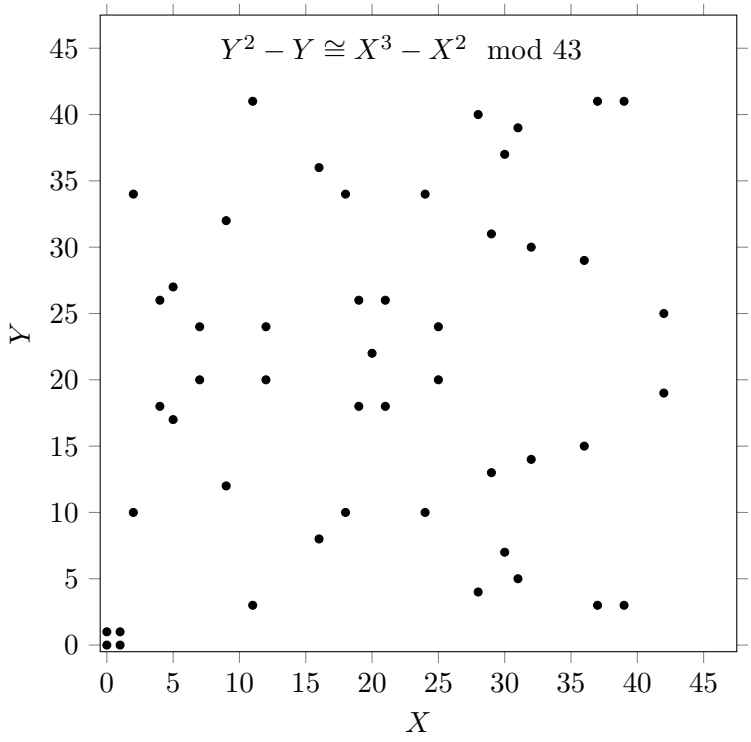
$$x \cong 36X - 12 \mod 13 \quad , \qquad y \cong 216Y - 108 \mod 13\,. \tag{4.45}$$

Example 2
For the order of $E : Y^2 - Y = X^3 - X^2$ over the field $\mathbb{F}_{17}$ one can take the coefficient $d_{17} = -2$ and subtract it from $p + 1$. The result is $\#E(\mathbb{F}_{17}) = 17 + 1 - -2 = 20$. We thus obtain 20 for the order of the elliptic curve $Y^2 - Y = X^3 - X^2$ over $\mathbb{F}_{17}$. The 20 points are $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(2, 9)$, $(7, 8)$, $(7, 10)$, $(8, 3)$, $(8, 13)$, $(9, 2)$, $(9, 16)$, $(11, 5)$, $(11, 13)$, $(12, 5)$, $(12, 13)$, $(13, 8)$, $(13, 10)$, $(15, 8)$, $(15, 10)$ and $\mathcal{O}$.
We could also have used the polynomial $P(17, a, b) = 15a^4 + 2ab^2$ from the table in section 4.5. For $y^2 = x^3 - 432x + 8208$ over $\mathbb{F}_{17}$ the constants a and b are reduced to $a = -432 \cong 10$

mod 17 and $b = 8208 \cong 14 \mod 17$. Substitution of the latter in $P(17, a, b) = 15a^4 + 2ab^2$ gives $P(17, a, b) \cong 2 \mod 17$. Of the set $\{..., -32, -15, 2, 19, 36, ...\}$ only 19 satisfies the Hasse bounds. In this way we also obtain 20 for the order.

Example 3

For the order of $E : Y^2 - Y = X^3 - X^2$ over the field $\mathbb{F}_{43}$ we take the coefficient $d_{43} = -6$ and subtract it from $p+1$. The result is $\#E(\mathbb{F}_{43}) = 43 + 1 - -6 = 50$. We thus obtain 50 for the order of the elliptic curve $Y^2 - Y = X^3 - X^2$ over $\mathbb{F}_{43}$. The 50 points are $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$, $(2,10)$, $(2,34)$, $(4,18)$, $(4,26)$, $(5,17)$, $(5,27)$, $(7,20)$, $(7,24)$, $(9,12)$, $(9,32)$, $(11,3)$, $(11,41)$, $(12,20)$, $(12,24)$, $(16,8)$, $(16,36)$, $(18,10)$, $(18,34)$, $(19,18)$, $(19,26)$, $(20,22)$, $(21,18)$, $(21,26)$, $(24,10)$, $(24,34)$, $(25,20)$, $(25,24)$, $(28,4)$, $(28,40)$, $(29,13)$, $(29,31)$, $(30,7)$, $(30,37)$, $(31,5)$, $(31,39)$, $(32,14)$, $(32,30)$, $(36,15)$, $(36,29)$, $(37,3)$, $(37,41)$, $(39,3)$, $(39,41)$, $(42,19)$, $(42,25)$ and $\mathcal{O}$. To get an impression the 49 finite integer points are shown in the next figure.



For the determination of the order we could also have used the polynomial $P(43, a, b) = a^9b + 7a^6b^3 + 33a^3b^5 + 35b^7$ from the table in section 4.5. For $y^2 = x^3 - 432x + 8208$ over $\mathbb{F}_{43}$ the constants a and b are reduced to $a = -432 \cong 41 \mod 43$ and $b = 8208 \cong 38 \mod 43$. Substitution of the latter in $P(43, a, b) = a^9b + 7a^6b^3 + 33a^3b^5 + 35b^7$ gives $P(43, a, b) \cong 6 \mod 43$. Of the set $\{..., -80, -37, 6, 49, 92, ...\}$ only 49 satisfies the Hasse bounds. In this way we also obtain 50 for the order.

Chapter 5

Cryptography

5.1 Introduction

Cryptography is the art of encrypting and decoding messages. The goal of the encryption is to keep a message secret. It is effective as long as others are not able to decode. Suppose Alice sends to Bob the following encrypted message: KRZEXJ JAFIWEMRCA. It is a bit difficult to decipher because the length of the message is very small. For a large message, say a page, you can count the frequency of characters and compare it with general frequencies. The leading character, the character which occurs most, probably is an encrypted E. The next to leading character probably is an encrypted T, etc. After a view trials you will obtain the original message. The encryption method for the given message is simple: Denote the characters A, B, ..., Z as $\alpha(1), \alpha(2), ..., \alpha(26)$. If $\alpha(m)$ is a character in the message, then the encrypted character is $\alpha(2m \mod 27)$. For example, E= $\alpha(5) \to \alpha(10)$ =J, and T= $\alpha(20) \to \alpha(40 \mod 27) = \alpha(13)$ =M. For the encryption with $2m \mod 27$ the decoding key is $\alpha(2^{-1}m \mod 27) = \alpha(14m \mod 27)$. For example, M=$\alpha(13) \to \alpha(13 \cdot 14 \mod 27) = \alpha(182 \mod 27) = \alpha(20)$ =T. Knowing the decryption key one easily finds the original message: SIMPLE ENCRYPTION. Alice and Bob could also have encrypted the message by $\alpha(m) \to \alpha(4m \mod 27)$: VIYJUT TBLRSJZIFB. Or by $\alpha(m) \to \alpha(5m \mod 27)$: NRKZFY YPOIQZSRUP. Or by $\alpha(m) \to \alpha(k \cdot m \mod 27)$ for any $1 < k < 27$ for which $\gcd(k, 27) = 1$. Of course, the method is extremely weak and in general keys are much larger numbers. Nevertheless, the example illustrates Alice and Bob somehow have to exchange the common key k. It makes the method vulnerable for eavedroppers. The exchange of a common key can be avoided by means of the Diffie-Hellman key exchange. It is based on modular elliptic curves and briefly works as follows: Alice and Bob use an elliptic curve $E(\mathbb{F}_p)$ and a point P on the curve. Both $E(\mathbb{F}_p)$ and P are not secret, it is the public key. Alice chooses a secret number a, Alice's private key, and Bob chooses a secret number b, Bob's private key. Alice computes the point aP and sends it to Bob. Bob computes the point bP and sends it to Alice. Alice computes the point abP and Bob computes baP. Both use $abP = baP$ for

their common key. They just have to convert the point abP to a number. For instance, by taking the x coordinate of abP as the common key. Anyway, Alice and Bob have established a common key without exchanging it.

Another property of the given encryption example is that the characters of the message are encrypted. Elliptic curves also are used for message encryption. A message can be represented as a point on an elliptic curve over $\mathbb{F}_p$ with p a large prime. A simple way is for instance to write A as 01, b as 02 though Z as 26, and a space delimiter as 00. Then the message SIMPLE ENCRYPTION is converted to a number m: $m = 19091316120500051403182516200915 14$. A famous method to send m is RSA (Rivest-Shamir-Adleman). Alice tells Bob she wants to sent him a secret message. Bob chooses two large primes, p and q, and computes the products: $n = pq$ and $k = (p-1)(q-1)$. Bob also chooses two integers d and e such that $de \cong 1 \mod k$. Bob sends n and e to Alice; n and e are public. In return Alice computes $c = m^e \mod n$ and sends it to Bob. With his secret number d Bob computes $c^d \mod n$. Since $c^d \cong (m^e)^d \cong m \mod n$ Bob recovers m. One should use very large primes to have $n > m$ and to make the factorization of n difficult. The message can also be sent by means of elliptic curves. Then one uses m as the x coordinate of a point on an elliptic curve. If there is no point on the curve for $x = 19091316120500051403182516200915 14$, on tries $x = 19091316120500051403182516200915 1401$, $x = 19091316120500051403182516200915 1402$, etc. until it is the x coordinate of a point M on the curve. Alice can send the message to Bob by means of Massey-Omura encryption. Alice chooses a secret number a, her private key, and Bob chooses a secret number b, his private key. Alice computes the point aM and sends it to Bob. Bob computes the point baM and sends it to Alice. Alice computes the point $a^{-1}baM$ and sends it to Bob. Bob computes $b^{-1}a^{-1}baM = M$, which he converts to characters to obtain the message.

Often the message or document itself is not confidential. Bob just wants to be sure the document is send by Alice. It requires an algorithm to verify the digital signature is valid and belongs to Alice. An algorithm based on elliptic curves is ECDSA (Elliptic Curve Digital Signature Algorithm). Alice and Bob use an elliptic curve $E(\mathbb{F}_p)$ with group order N. Alice chooses a secret number a, her private key. Alice chooses a point P on the curve of order N, and computes $Q = aP$. Alice and Bob also use a function f which converts a point (x, y) on the curve to a number. The function $f(x, y) = x$, as mentioned above, is an example. The set $(E(\mathbb{F}_p), N, P, Q)$ is the public key. For each message Alice chooses a random integer k and computes $R = kP$. The message or document is represented as an integer m by means of a hash function (hash functions will be considered further on). Alice computes $g = k^{-1}(m + ax_R) \mod N$, where x_R is the x coordinate of point R. Alice sends the three numbers m, R and g to Bob. Bob computes $u = g^{-1}m \mod N$, $v = g^{-1}x_R \mod N$ and verifies if $uP + vQ$ is equal to R. If the document is really signed by Alice, thus with the use

of a, then $uP + vQ = g^{-1}mP + g^{-1}x_RQ = g^{-1}(mP + x_RaP) = g^{-1}kgP = kP = R$. If it is signed by someone else with $g' = k^{-1}(m + a'x_R) \mod N$, then $uP + vQ \neq R$.

In the foregoing examples a lot of technical details are omitted for brevity and simplicity. For instance, one should use elliptic curves for which it is supposed to be difficult to determine the secret key a from the public points P and aP. To achieve the latter one should at least use large values for a and p. We also did not give a complete survey of existing methods. Instead, we will focus on the ECDSA as used in the bitcoin blockchain. Before we turn to the bitcoin ECDSA we first consider number bases.

5.2 Number bases

Our usual numbers are expressed in a decimal system by means of ten characters: 0,1,2,3,4,...,9. For example, 374 means $3\cdot10^2+7\cdot10^1+4\cdot10^0$ and 405.6 means $4\cdot10^2+0\cdot10^1+5\cdot10^0+6\cdot10^{-1}$. For the decimal system the base is 10. One can also use other bases. Suppose we want to count in base 7. Then we only use the characters 0,1,2,3,4,5 and 6. The number 532 in base 7 has the value $5 \cdot 7^2 + 3 \cdot 7^1 + 2 \cdot 7^0 = 268$. That is, 532 in base 7 equals 268 in base 10. We see a number obtains its proper value if you know its base. Numbers should therefore be expressed together with their base. For example, 532 in base 7 equals 268 in base 10 is written as $(532)_7 = (268)_{10}$. Suppose we want to count in base 12. Then we need two additional characters. Usually one takes 'a' for 10 and 'b' for 11. The number 5ab in base 12 has the value $5 \cdot 12^2 + 10 \cdot 12^1 + 11 \cdot 7^0 = 851$ in base 10. Thus $(5ab)_{12} = (851)_{10}$. In daily life one writes the numbers without the base since we just know the base is 10. Hereafter, a number in base 10 will be written without its base. Thus, $(532)_7 = 268$ and $(5ab)_{12} = 851$.

In computers a bit is either a 0 or a 1. A group of 8 bits make a byte. The first bit of the byte represents the sign of a number. With the other 7 bytes we can for instance make $(0001011)_2 = 2^3 + 2^1 + 2^0 = 11$ or $(1111111)_2 = 2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 = 127$. Since $(1111111)_2 + (0000001)_2 = (10000000)_2 = 2^7$, the latter can be briefly written as $(1111111)_2 = 2^7 - 1 = 127$. The base 2 system is known as the binary system.

For example, the number 447 in the bases 2 through 16 is: $(110111111)_2$, $(121120)_3$, $(12333)_4$, $(3242)_5$, $(2023)_6$, $(1206)_7$, $(677)_8$, $(546)_9$, $(447)_{10}$, $(377)_{11}$, $(313)_{12}$, $(285)_{13}$, $(23d)_{14}$, $(1ec)_{15}$ and $(1bf)_{16}$. The base 16 system is called the hexadecimal system.

If a number in base 10 ends on 0, 2, 4, 5, 6 or 8 then it is not a prime number, since 2 and 5 are divisors of 10. Similarly, if a number in base 12 ends on 0, 2, 3, 4, 6, 8, 9 or 'a' then it is not a prime number, since 2 and 3 are divisors of 12. For example, for the number 188321739 it is not immediately clear if it is a prime or composite. In base 12 it reads

$(5309a5a3)_{12}$. Since it ends on a 3 the number 188321739 must be divisible by 3 and is therefore composite. Similarly, for the number $(3199467)_{12}$ it is not immediately clear if it is a prime or composite. In base 10 it reads 9409615. Since it ends on a 5 we immediately see it is composite.

It may be illuminating to show the first 32 integers in different bases, see the next table.

(base 10) number \ base	2	4	8	12	16
1	1	1	1	1	1
2	10	2	2	2	2
3	11	3	3	3	3
4	100	10	4	4	4
5	101	11	5	5	5
6	110	12	6	6	6
7	111	13	7	7	7
8	1000	20	10	8	8
9	1001	21	11	9	9
10	1010	22	12	a	a
11	1011	23	13	b	b
12	1100	30	14	10	c
13	1101	31	15	11	d
14	1110	32	16	12	e
15	1111	33	17	13	f
16	10000	100	20	14	10
17	10001	101	21	15	11
18	10010	102	22	16	12
19	10011	103	23	17	13
20	10100	110	24	18	14
21	10101	111	25	19	15
22	10110	112	26	1a	16
23	10111	113	27	1b	17
24	11000	120	30	20	18
25	11001	121	31	21	19
26	11010	122	32	22	1a
27	11011	123	33	23	1b
28	11100	130	34	24	1c
29	11101	131	35	25	1d
30	11110	132	36	26	1e
31	11111	133	37	27	1f
32	100000	200	40	28	20

5.3 Bitcoin ECDSA

The bitcoin digital signatures uses the elliptic curve $E(\mathbb{F}_p): y^2 = x^3 + 7 \mod p$, where p is the prime

115792089237316195423570985008687907853269984665640564039457584007908834671663

Its value is equal to $p = 2^{256} - 2^{32} - 2^9 - 2^8 - 2^7 - 2^6 - 2^4 - 1$. In base 16 it reads $p =$ ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff fffffe ffffc2f. For this value of p the integer points on the curve $E(\mathbb{F}_p): y^2 = x^3 + 7 \mod p$ form a cyclic group of order $\#E(\mathbb{F}_p) = N =$ 115792089237316195423570985008687907852837564279074904382605163141518161494337.

In base 16 the order reads ffffffff ffffffff ffffffff fffffffe baaedce6 af48a03b bfd25e8c d0364141. The order, which is somewhat smaller than p, is a prime. The bitcoin base point is $P =$ (550662630222773436695787188951685343262506034537775941755001873603891167292 40, 32670510020758816978083085130507043184471273380659243275938904335757337482424).

In base 16 the base point reads $P =$
(79be667e f9dcbbac 55a06295 ce870b07 029bfcdb 2dce28d9 59f2815b 16f81798,
483ada77 26a3c465 5da4fbfc 0e1108a8 fd17b448 a6855419 9c47d08f fb10d4b8).

The base point P is an element of $E(\mathbb{F}_p)$ and its order is equal to the aforementioned group order N. The modular elliptic curve $E(\mathbb{F}_p): y^2 = x^3 + 7 \mod p$ with p and base point P as given above is called **secp256k1** by the Standards for Efficient Cryptography Group.

The signature process is as follows. Alice chooses (of course the computer software randomly chooses) a secret number a, her private key. Alice computes $Q = aP$, her public key. Since the order of P is a prime there is precisely one value for a for which $aP = Q$. Q can be calculated efficiently as will be shown below. However, the reconstruction of a from the public points Q and P is extremely time consuming; billions of years with the presently known algorithms. For each message Alice chooses a random integer k and computes $R = kP$. The message or document m is represented as an integer h by means of a hash function (hash functions will be considered further on). Alice computes $s = k^{-1}(h + r \cdot a) \mod N$, where r is the x coordinate of point R. The signature of the hashed message h is the pair (r, s). To check if the signature is legal one computes $u = s^{-1}h \mod N$ and $v = s^{-1}r \mod N$, and verifies if $uP + vQ$ is equal to R. If the document is really signed by Alice, thus with the use of a, then $uP + vQ = s^{-1}hP + s^{-1}rQ = s^{-1}(hP + r \cdot aP) = s^{-1}k \cdot sP = kP = R$. If it is signed by someone else with $s' = k^{-1}(h + a' \cdot r) \mod N$, then $uP + vQ \neq R$. That is, there is an extremely small probability (1 out of 10^{78}) that a' happens to equal a.

Since the numbers are extremely large it is very time consuming to do calculations explicitly. To show explicitly the calculation of aP for some $a < N$ we take a smaller prime: $p = 43$. For this value of p the integer points on the curve $E(\mathbb{F}_p) : y^2 = x^3 + 7 \mod p$ form a cyclic group of order $\#E(\mathbb{F}_p) = N = 31$. Let $P = (20, 3)$ be the base point and let 19 be the value for a. Thus, $Q = 19P$. It can be computed by successively calculating $2P$, $3P$, $4P$, ..., $19P$. That would take 18 steps. However, it is more efficient to calculate $2P = (13, 21)$, $4P = (12, 31)$, $8P = (42, 36)$ and $16P = (40, 25)$ with the doubling formula and than compute $16P + 2P + P$ with the addition formula. Then we obtain $19P = (38, 21)$ in 6 steps. Another way is to compute $9P = (37, 36)$ from $8P+P$, double it to $18P = (25, 25)$ and add P to obtain $19P = (38, 21)$. The latter also takes 6 steps. Either way one arrives at $Q = (38, 21)$. Suppose the hashed message is $h = 15$ and suppose further that $k = 26$ for the signature of the message.

The first computation is $R = kP = 26P = (34, 40)$. So, $r = 34$. The second computation is $k^{-1} \mod 31$ for $k = 26$. Since $6 \cdot 26 = 156 \cong 1 \mod 31$ we have $26^{-1} \cong 6 \mod 31$. The third computation is $s = k^{-1}(h + r \cdot a) \mod N$. Substituting the values we obtain $s = 6(15 + 34 \cdot 19) \cong 3966 \cong 29 \mod 31$. So, $s = 29$. The signature therefore is $(34, 29)$.

Next we consider the verification. The first computation is $s^{-1} \mod 31$. Since $15 \cdot 29 = 435 \cong 1 \mod 31$ we have $29^{-1} \cong 15 \mod 31$. The second computation is $u = s^{-1}h \cong 15 \cdot 15 \cong 225 \cong 8 \mod 31$. The third computation is $v = s^{-1}r \cong 15 \cdot 34 \cong 510 \cong 14 \mod 31$. The fourth computation is $uP + vQ = 8P + 14Q = 8(20, 3) + 14(38, 21) = (42, 36) + (25, 25) = (34, 40)$. Since $R = (34, 40)$ we have verified $uP + vQ$ indeed is equal to R and the signature is valid.

5.4 Hash function

In the foregoing example we just took $h = 15$ and did not worry about the hash function. What is denoted as h is actually a hash of a message m. That is, $h(m) \cong H(m) \mod N$, where H is the hash function. It is a repeated conversion of the input string by means of hash algorithms. For bitcoin signatures and keys the important hash function is SHA256. It is an abbreviation for 'Secure Hash Algorithm with an output length of 256 bits'. The output length is always 256 bits (32 bytes). Identical inputs lead to identical outputs. Slightly different inputs lead to completely different outputs. We will not show how the SHA256 works. We will just illustrate how it is applied. As a first example we consider the string: "message". After the application of SHA256 we obtain the following number in hexadecimal system:

ab530a13e45914982b79f9b7e3fba994cfd1f3fb22f71cea1afbf02b460c6d1d.

The number contains 32 bytes (64 characters). To consider the hashing of other messages

we add a counter; in the blockchain world it is called a nonce. Let us consider the string "message1". After the application of SHA256 we obtain the following number in hexadecimal system:

97d035e32036a670058f2be4e008a7c56355489750a5da6f2af342db4a968e99.

Next we increase the nonce: "message2". After SHA256 we obtain in hexadecimal system:

e09b16811444401b35c94081ee8c82a761bcd3cfd7260cf063e3fec520f5f5e9.

We see that a very small change of the string leads to a completely different number. In summary, no matter the length of the string the hash result of SHA256 always is a 32 byte number in hexagonal base. Furthermore, if two messages differ by just a single character the two hash results will be completely different. It therefore is practically impossible to find (with reverse engineering) an input string a given output string.

Now we make a jump in the nonce and consider the string "message15". Then we obtain that SHA256(message15) is equal to:

08ddaf8df28d5eee382f1b9ba191aec331260df321c5a715c5f39bc3a59c0cad.

We see the hash starts with a zero. For "message169" the hash even starts with two zero's:

009c5abcd3a674c926ec880886ab57f226e2cb981a2fc43e00b47b92a8e528b1.

There seems to be no structure or relation between the message string and the number of starting zero's of the hash. If the zero's and ones appear in a random order, one expects n starting zero's in a fraction 16^{-n} of the occasions. It will take on average 16^n attempts to find a SHA256 hash starting with n zero's. To find a hash starting with, say, 10 zero's requires about $16^{10} \approx 10^{12}$ attempts. That is a lot of work and it is an essential part of the blockchains technology.

5.5 Blockchain and mining

Transactions of cryptocurrencies are recorded. A list of transactions looks like:

0.01384267 bitcoin from Alice to Bob
0.31082855 bitcoin from Bob to Charlie
0.02581299 bitcoin from Alice to Charlie

and so on.

In reality, a transaction also contain the public key of the sender and the public key of the receiver, both Base58 encoded into a string of 34 alphanumeric characters from the set 1 through 9, 'a' through 'z' except the small L and A through Z except O and I(capital i). A public key is created as follows. One starts with, for instance, the hexadecimal word

023bb54d336d30a6fcb9cf17aa5bafefbdb6509c0465c0c13c6427f74a0fdce213.

The first byte, 02, is the parity of the y-coordinate and the other 32 bytes is the hexadecimal x-coordinate. First SHA256 is applied to the latter. The result is the hexadecimal word

d9847cd0e87fe3e6d9e6f2f8a600e6a04ad4468e2c1979436b1e5b59a9f0fd08.

Next the hash function RIPEMD160 is applied. The result is the hexadecimal word

61182fab45b8bb6141142f732ae9426fbdf5e409.

Two zero's are placed in front of it:

0061182fab45b8bb6141142f732ae9426fbdf5e409.

To the latter result SHA256 is applied:

2ac64aeb5c4a686f0c4d1ba110e4fa0552d693ecc86fe4d7c325f3b2818d8b84.

Then SHA256 is once more applied:

777097fe0c570f4ea8dc395cbcfb33a64fe86cd0e7528e9fe48fa0c5e0fd86cf.

The first four bytes, that is 777097fe, are concatenated to the end of the RIPEMD160 + front 00 result:

0061182fab45b8bb6141142f732ae9426fbdf5e409777097fe.

Finally, the latter hexadecimal word is Base58 encoded:

19rPX6VXBW8JjaR2QD8Hvd6VKu4sU1dp4u.

The latter is a the Base58 encoded public key as used in the blockchain.

All the transactions are grouped into blocks. If enough new transactions have appeared a new block will be created. Each block consists of a block header, a SHA256 hash of the block header and the list of transactions. Each block header consists of a version number (4 bytes), the previous block header hash (32 bytes), the Merkle root (32 bytes) which will be explained below, a UNIX timestamp (4 bytes), the difficulty target (4 bytes) which determines the number of starting zero's of the block header hash and the nonce (4 bytes).

Before we proceed we first consider the Merkle root. Each transaction is given a SHA256 hash: transaction 1 $\rightarrow$ H(1), transaction 2 $\rightarrow$ H(2), transaction 3 $\rightarrow$ H(3), etc. Then pairs of hashes are hashed: H(1)+H(2) $\rightarrow$ H(12), H(3)+H(4) $\rightarrow$ H(34), H(5)+H(6) $\rightarrow$ H(56), etc. Again pairs of hashes are hashed: H(12)+H(34) $\rightarrow$ H(1234), H(56)+H(78) $\rightarrow$ H(5678), etc. In the end we are left with a single hash, which is the Merkle root. Starting from the root the hashes form a binary tree. It is invented for time efficient verifications.

For the block header hash all the numbers in the block header are concatenated and the result is hashed with SHA256. If the latter hash does not start with the required number of zero's, the nonce is increased and a new hash is generated. The process of trying a large number of nonces until the hash starts with the required number of zero's is called *mining*. If a hash satisfies the requirements the block is accepted and the miner is rewarded with a given amount of bitcoins. A little later, about every 10 minutes, enough transactions will have appeared and a new block will be created. Part of the header of the new block is the hash of the previous block header. Therefore the blocks are connected: the block chain.

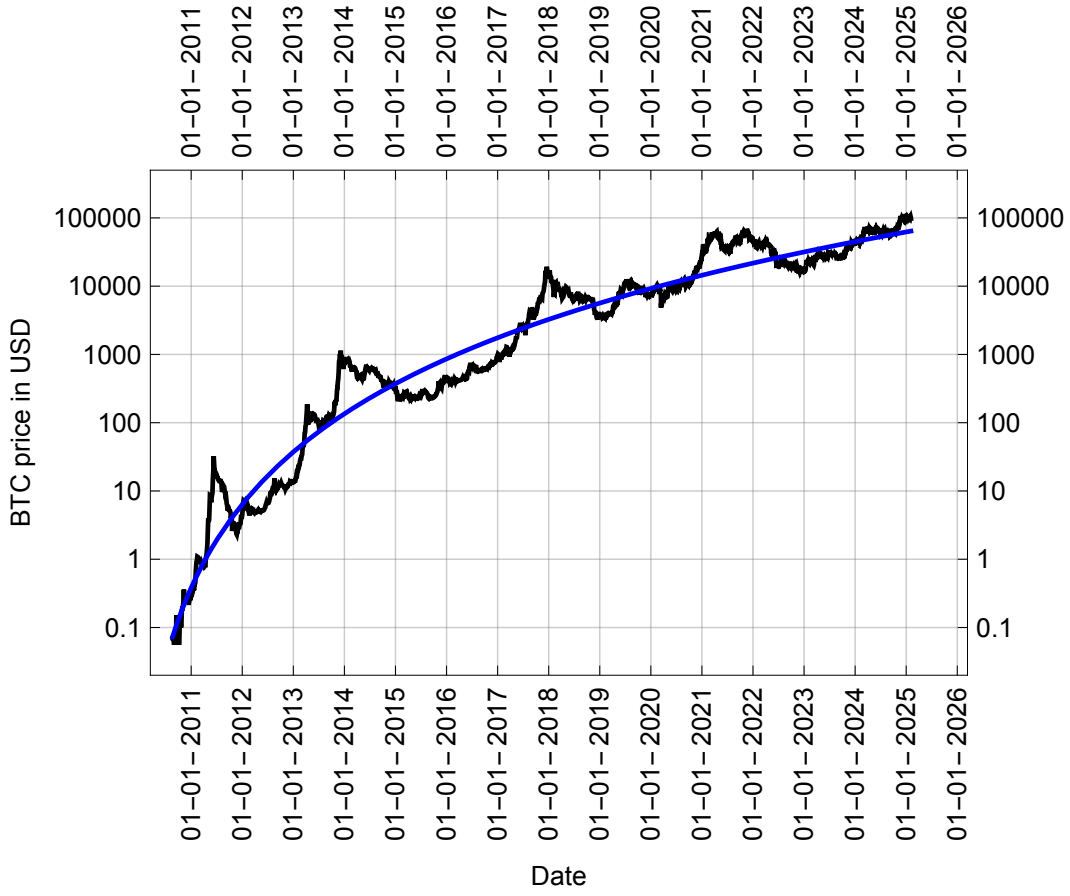
Suppose Alice buys a painting from art painter Bob for 0.1 bitcoin. Suppose after the transfer of the painting and the 'money' Alice decides to cheat Bob. Alice changes the transaction to, for instance, 0.01 bitcoin instead of 0.1 bitcoin. However, the small change, causes a different hash of the transaction and therefore to a different Merkle root. As a consequence the block header hash changes and Alice is forced to go to the long process of finding a new block header hash with the required number of starting zero's. By the time she has found new block header all the other miners are a few blocks ahead of her. Because of the chain her alternative block header hash will change the block header of the successor block which in turn will change the next successor block and so on. By the time Alice has found hashes for them too all the other miners are far ahead of her. In the end the longer list of blocks created by all the miners will be accepted and the single alternative block (or short list of alternative blocks) of Alice will be declined. The longer list of blocks has taken more work, delivered by all the miners. Because of this 'proof of work' it is regarded as the correct list of blocks.

It should be noted that not all the details are mentioned. For instance, the block header numbers being byte reversed is omitted. Such details are rather technicalities. In summary,

the validity of a transaction is achieved by means of ECDSA and the reliability of the chain of blocks of transactions is achieved by the hashing, with difficulty, of the block headers. Since we are more interested in the underlying mathematics we focussed on modular counting and elliptic curves in the preceding chapters. In the next and final section we will just consider the bitcoin rate.

5.6 Bitcoin rate

The price of a bitcoin (BTC) in US dollars (USD) is very volatile. The history of the BTC-USD rate is shown in the next diagram.



The historical trend curve (blue) approximately goes as $0.13\,t^{4.8}$, where t is the time in years from 1 october 2009. Of course, historical trends do say nothing about future developments.

Bibliography

[1] L.C. Washington, *Elliptic Curves, Number Theory and Cryptography*, Chapman & Hall/CRC, 2008.

[2] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, New York, 1993.

[3] J.H. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, New York, 1992.

[4] E. Rykwalder, *Explaining The Math Behind Bitcoin*, https://www.coindesk.com/math-behind-bitcoin.

[5] 1blue3brown, *Ever wonder how Bitcoin (and other cryptocurrencies) actually work?*, https://www.youtube.com/watch?v=bBC-nXj3Ng4

[6] N.J.A. Sloane, *The Online Encyclopedia of Integer Sequences*, https://oeis.org

[7] M. Eichler, Quaternäre quadratische Formen und die Riemannsche Vermutung für die Kongruenzzetafunktion, *Archiv für mathematische Logik und Grundlagenforschung*, **5**, 355 (1954).