Systems of linear equations

and

quadratic equations

Thure Dührsen

September 25, 2012 20:08

Contents

Things waiting to get done

Part I.
Preliminary stuff

What we need to know before we can do stuff.

This is heavily a work in progress.

1. Super-quick overview of number sets

What is a number?

1.1. The natural numbers

When counting things, such as playing cards on a table, you start out with the number one and, one by one, you advance to the higher numbers. Note the wording "one by one", because it implies that when you have reached a certain number, going to the *next* number is all you need to reach numbers arbitrarily large so that, while in practice you cannot keep counting indefinitely, in theory you can. [1] So the set of counting numbers is infinite. We'll call it the **natural numbers** and denote it by the symbol $\mathbb{N}$.

1.1.1. Definition of the natural numbers

So what have we done so far? We have seen that

(1) 1 is a natural number and

(2) whenever n is a natural number, the *next* number is a natural number as well.

For every natural number n, we call the next number the **successor** of n and denote it by n^*.

We also call n the **predecessor** of n^*. We'll denote the predecessor of a natural number n by n_*.

Suppose a few cards are lying on the table and we pick them up, again one by one. As long as we haven't picked up all of them, the number of cards on the table is a natural number. When we pick up the last card, there are no more cards left on the table. So the

[1] From a mathematical point of view, you don't even need names for the numbers.

number of cards on the table should be the predecessor of 1, which does not exist in $\mathbb{N}$, as we started out with the number 1 when defining the natural numbers. Another way of saying this is

(3) 1 is not the successor of any natural number.

So far, we do not have a symbol for the concept of nothing being there. But fortunately, someone came up with a new symbol, 0, to represent this. The set of natural numbers, united with the set that contains the new symbol 0, is written $\mathbb{N}_0$.

Now we get the usual names of the natural numbers as follows:

$$\begin{array}{rcccc} & & 1 & \in & \mathbb{N} \\ 2 & := & 1^* & \in & \mathbb{N} \\ 3 & := & 2^* & \in & \mathbb{N} \\ 4 & := & 3^* & \in & \mathbb{N} \\ \text{and so on.} & & & & \end{array}$$

1.1.2. Adding natural numbers

Now, using the definitions above, we can define the addition of natural numbers as follows:

Definition 1 (Addition of natural numbers)**.**

($\mathrm{Add}_{\mathbb{N}} - 1$) For all natural numbers n, adding 1 to n is defined as advancing to the successor of n, i. e.

$$n+1 \quad := \quad n^* \qquad \text{for all } n \in \mathbb{N} \qquad .$$

($\mathrm{Add}_{\mathbb{N}} - 2$) For all natural numbers n and for all natural numbers k, adding the successor of k to n is defined as adding k to n and advancing to the successor of the sum, i. e.

$$n+k^* \quad := \quad (n+k)^* \qquad \text{for all } k \in \mathbb{N} \text{ and for all } n \in \mathbb{N} \qquad .$$

How does adding numbers this way work? Consider the following example:

$$
\begin{aligned}
3+4 &= 3+3^* && \text{(Definition of the number 4)} \\
&= (3+3)^* && (\mathrm{Add}_{\mathbb{N}}-2) \\
&= (3+2^*)^* && \text{(Definition of the number 3)} \\
&= \big((3+2)^*\big)^* && (\mathrm{Add}_{\mathbb{N}}-2) \\
&= \big((3+1^*)^*\big)^* && \text{(Definition of the number 2)} \\
&= \Big(\big((3+1)^*\big)^*\Big)^* && (\mathrm{Add}_{\mathbb{N}}-2) \\
&= \Big(\big((3^*)^*\big)^*\Big)^* && (\mathrm{Add}_{\mathbb{N}}-1) \\
&= \Big(\big((4)^*\big)^*\Big)^* && \text{(The successor of 3 is 4)} \\
&= \big((4^*)^*\big)^* && \text{(Leave off the innermost parentheses)} \\
&= \big((5)^*\big)^* && \text{(The successor of 4 is 5)} \\
&= (5^*)^* && \text{(Leave off the innermost parentheses)} \\
&= (6)^* && \text{(The successor of 5 is 6)} \\
&= 6^* && \text{(Leave off the innermost parentheses)} \\
&= 7 && \text{(The successor of 6 is 7)}
\end{aligned}
$$

1.1.3. Is 1 really the first natural number?

Earlier we defined the natural numbers as beginning with the number 1. It is quite clear, however, that we could just as well have taken the number 0 to be the first natural number. Then we have to alter the three laws from that section as follows:

(1') 1 is a natural number and

(2') whenever n is a natural number, the *next* number is a natural number as well.

(3') 1 is not the successor of any natural number.

Once we have done that, we define 1 to be the successor of 0. We also have to redefine addition to include the law $n+0 := n$ for all natural numbers n. But that's pretty much all we have to do. So in effect, it does not really matter what number we choose to be the first natural number – but having it be anything other than 0 or 1 would be highly "unnatural", if you will forgive the pun.

1.1.4. Comparing natural numbers

The next thing we'll do is compare natural numbers.

When looking at the two natural numbers 4 and 9, what does it mean to say that 4 is smaller than 9?

Well, one answer to this question is obvious: When counting from 1, we reach 4 before we reach 9. When looking at the definition of addition, we see that this means there is a natural number, 5 in this case, such that $4+5=9$. Here it is important that the augend is a natural number! Consider the equation $7+n=5$. Is there any natural number n such that this equation is true? I can't find any. I can even say why: Adding a natural number n to 7 means repeatedly advancing to the next number, exactly n times. But no matter how many times we do this, we cannot ever reach 5, as 5 occurred already *before* we reached 7 when counting from 1.

Definition 2. [Comparison of natural numbers] Let a and b be natural numbers.

($\mathrm{Comp}_{\mathbb{N}}-1$) a is **smaller than** b, by definition, if and only if there exists a natural number m such that $a+m=b$. We then write $a<b$.

($\mathrm{Comp}_{\mathbb{N}}-2$) a is **greater than** b, by definition, if and only if b is smaller than a. We then write $a>b$.

($\mathrm{Comp}_{\mathbb{N}}-3$) a is **equal to** b, by definition, if and only if a is neither smaller nor greater than b. We then write $a=b$.

($\mathrm{Comp}_{\mathbb{N}}-4$) a is **greater than or equal to** b, by definition, if and only if a is not smaller than b. We then write $a \geqslant b$.

($\mathrm{Comp}_{\mathbb{N}}-5$) a is **smaller than or equal to** b, by definition, if and only if a is not greater than b. We then write $a \leqslant b$.

1.1.5. Subtracting natural numbers

Now that we can compare natural numbers, we can easily introduce subtraction:

Definition 3 (Subtracting natural numbers)**.** Let a and b be natural numbers such that a is greater than b.

By definition (2), this means that b is smaller than a, i. e. there exists a natural number m such that $b+m=a$. The number m is called the **difference** of a and b, and the operation of finding it is called **subtraction**.

1.1.6. The commutative law of addition

One of the laws governing the natural numbers we use literally every day is the commutative law of addition: For all natural numbers m, n it is true that $m + n = n + m$. Can we prove this?

Yes, indeed – but not without doing some work beforehand.

1.1.7. The associative law of addition

So far, we've used parentheses whenever there were more than two summands in a sum, and when evaluating a sum, all we needed to know was how to add exactly two numbers, as only that kind of sum occurred inside the innermost pair of parentheses. The reason we did this is that up to now, we have only defined sums of two terms. Whenever the augend was greater than 1, we used the property ($\mathrm{Add}_{\mathbb{N}} - 2$) to break it down until we reached 1, where we used property ($\mathrm{Add}_{\mathbb{N}} - 1$) to advance to the successor.

Now what if we want to add more than two natural numbers? If $a_1, a_2, \ldots, a_n$ are natural numbers, we would need to write

$$\Big(\Big(\ldots((a_1 + a_2) + a_3) + \ldots\Big) + a_{n_*}\Big) + a_n$$

to denote their sum, which is quite cumbersome, if you ask me. So it would be nice if we could do away with the parentheses and simply write

$$a_1 + a_2 + a_3 + \ldots + a_{n_*} + a_n \qquad .$$

1.1.8. Semigroups

1.1.9. Multiplying

1.2. The integers

1.2.1. Definition of the integers

1.2.2. Adding and subtracting

1.2.3. Groups

1.2.4. Multiplying

1.2.5. Rings

1.3. The rational numbers

1.3.1. Adding and subtracting

1.3.2. Comparing

1.3.3. Multiplying

1.3.4. Dividing

1.3.5. Fields

1.4. The real numbers

Part II. The real stuff

2. Systems of linear equations

Definition 4. A **system of linear equations**, or **linear system** for short, is a collection of *linear equations* all involving the same set of *variables* or *unknowns*. Note that there is no requirement for the number of equations to be the same as the number of variables.

A linear system of m equations in the n variables $x_1, \ldots, x_n$ can always be written in the standard form

$$\left\{ \begin{array}{rcl} a_{1,1}\cdot x_1 + a_{1,2}\cdot x_2 + \ldots + a_{1,n}\cdot x_n & = & b_1 \\ a_{2,1}\cdot x_1 + a_{2,2}\cdot x_2 + \ldots + a_{2,n}\cdot x_n & = & b_2 \\ & \vdots & \\ a_{m,1}\cdot x_1 + a_{m,2}\cdot x_2 + \ldots + a_{m,n}\cdot x_n & = & b_m \end{array} \right.$$

where the numbers $a_{i,j}$ and b_i are known real numbers.[2] The numbers $a_{i,j}$ are called the system's *coefficients*. The numbers b_i form the *right-hand side (RHS)*. If the RHS is zero in all the equations, the system is called *homogenous*, otherwise we say it is *inhomogenous*.

The brace to the left of the equations is not eye candy. It is intended to signify that they form a set that is to be solved simultaneously; simply writing an equation beneath

[2] At least that's the way things are in high school. Generally, $\mathbb{R}$ is not the only field those numbers can be taken from. However, in any given linear system, only one field is involved. — What does *field* mean in this context? The short answer is that a field F is a set of numbers endowed with a certain structure that allows one to take any sum, difference, product and quotient of two numbers $x, y \in F$ and know that the result will also be a member of F. In the case of division, however, the divisor may not be 0. In particular, this means that for every $x \in F \setminus \{0\}$ there exists $y \in F$ such that $x \cdot y = 1$. In that case, y is denoted x^{-1} or $\frac{1}{x}$ and called the *multiplicative inverse of* x.

One of the most important characteristics of fields is that they do not have any *zero divisors*, that is, for any two $x, y \in F$ that satisfy $x \cdot y = 0$, it follows that at least one of x and y is itself zero. Why is that so? Well, let x and y belong to F such that their product is zero. If $x = 0$, we are done. If, however, $x \neq 0$, then all we have to show is that $y = 0$. Because $x \neq 0$, x^{-1} exists in F. Consequently,

$$0 \;=\; x^{-1}\cdot 0 \;=\; x^{-1}\cdot(x\cdot y) \;=\; (x^{-1}\cdot x)\cdot y \;=\; 1\cdot y \;=\; y,$$

i. e. $y = 0$. $\square$

If you are turning this over in your head and asking yourself what magic property the reals have that the rationals lack, the answer is in the concept of *completeness* as explained in calculus. As an example, consider the well-known case of $\sqrt{2}$ not belonging to $\mathbb{Q}$. – If one restricts the set of numbers even further – to the integers, say –, things get rather boring, as the simplest of equations, such as $2x = 3$, are unsolvable.

another means that the second one follows from the first.[3]

Solving a linear system means finding numbers $x_1, \ldots, x_n \in F$ such that all of the equations are satisfied simultaneously. That's why **simultaneous equations** is another term for a collection of equations (not necessarily linear ones).

Writing a system down is one thing, solving it is quite another. Over the reals, consider the system

$$\begin{cases} 2\cdot x_1 + 3\cdot x_2 &= 13 \\ 2\cdot x_1 + 3\cdot x_2 &= 12 \end{cases} \quad ,$$

which is unsolvable because no matter what we set x_1 and x_2 to be, $2x_1 + 3x_2$ cannot be 12 and 13 at the same time, or even (again over the reals)

$$\begin{cases} 0\cdot x_1 = 7 \end{cases} \quad .$$

Even when a system has a solution, that solution does not at all have to be unique. Over the reals, let S be the set of solutions to the system

$$\begin{cases} x+y &= 2 \\ 3x+3y &= 6 \end{cases} \quad .$$

We can mangle the first equation to read $y = 2 - x$; plugging this into the second equation yields $3x + 3\cdot(2-x) = 6$, which simplifies to $6 = 6$. This last statement is true for any x and y we care to choose and only serves to indicate that the system is solvable at all. We have thus "used up" both of the given equations, but obtained only one real constraint on the unknowns, namely $y = 2 - x$. Therefore we are free to choose one of the variables, say x, to write down the solution set as follows

$$S = \{(t, 2-t) \mid t \in \mathbb{R}\}$$

and read it as "all those pairs of numbers where [the first component is a real number and] the second component is two minus the first component".

But we can just as well choose y as the free variable and state the solution set as

$$S = \{(2-t, t) \mid t \in \mathbb{R}\} \quad .$$

[3] It does *not* mean that the lines are equivalent! If one wishes to express that, the equivalence sign $\Longleftrightarrow$ is needed in front of the second line.

2.1. Einsetzungsverfahren

There are many methods of solving linear systems, and we've seen one of them above. It is called „Einsetzungsverfahren“ in German, which might be called “substitution method” in English. The idea is to pick one of the equations and solve it for one of the unknowns in terms of the remaining unknowns. After that, plug the expression for the single unknown into the other equations. That way, one variable has been eliminated from all the equations save the one we used to derive the expression. Now we repeat this process until we have reached a point where we can solve one equation completely, and back-substitute until the entire solution is found.

An example will probably make things clear. Over the real numbers, consider the system

$$\begin{cases} x+3y-2z &= 5 \\ 3x+5y+6z &= 7 \\ 2x+4y+3z &= 8 \end{cases} .$$

We'll take the first equation and rewrite it as $x=5-3y+2z$. Plugging this into the second and third equation yields

$$\begin{cases} x &= 5-3y+2z \\ 3\cdot(5-3y+2z)+5y+6z &= 7 \\ 2\cdot(5-3y+2z)+4y+3z &= 8 \end{cases} ,$$

which simplifies to

$$\begin{cases} x &= 5-3y+2z \\ -4y+12z &= -8 \\ -2y+7z &= -2 \end{cases} ,$$

and so we have eliminated the variable x from the second and third equation. Now we can apply the same technique to those equations. Manipulate the second equation into the form $y=2+3z$, so one of the variables is isolated, and plug this into the third equation and simplify. Now the system is

$$(*) \quad \begin{cases} x &= 5-3y+2z \\ y &= 2+3z \\ z &= 2 \end{cases} ,$$

so one of the unknowns has a numeric value that we can simply read off. What is more, we can also plug this value into the second equation to get $y=8$, which we can plug into the first equation and obtain $x=-15$. So the system's unique solution is $(-15,8,2)$.

Also note that the system (∗) resembles a triangle in that, when reading it from bottom to top, each equation has one more variable than the previous one. This might have given rise to the term “triangular form” (German: „Dreiecksform“) to describe such systems.

2.2. Additionsverfahren

Another way of solving linear systems is called „Additionsverfahren“. This time, I’ll risk an attempt at translating it as “addition method”. As the name suggests, the idea is adding two equations in such a way that the resulting equation contains fewer variables than the summands, and repeating the process. Care should be taken to treat the variables one by one.

Over the reals, consider the system

$$\begin{cases} \text{(I)} & 2x + 3y + 4z = 20 \\ \text{(II)} & 3x + 2y + 5z = 22 \\ \text{(III)} & 4x + 5y + z = 17 \end{cases} .$$

This looks like a lot of work using the Einsetzungsverfahren, so let’s do it as follows: We’ll multiply the first equation by 3 and the second one by (-2), then add them:

$$\begin{array}{r|rcrcrcr} 3\cdot\text{(I)} & 6x & + & 9y & + & 12z & = & 60 \\ \oplus\quad -2\cdot\text{(II)} & -6x & - & 4y & - & 10z & = & -44 \\ \hline \text{(IV)} & & & 5y & + & 2z & = & 16 \end{array}$$

In this way, x has vanished from the result. We’ll now multiply the first equation by 4 and the third one by (-2), then add them:

$$\begin{array}{r|rcrcrcr} 4\cdot\text{(I)} & 8x & + & 12y & + & 16z & = & 80 \\ \oplus\quad -2\cdot\text{(III)} & -8x & - & 10y & - & 2z & = & -34 \\ \hline \text{(V)} & & & 2y & + & 14z & = & 46 \end{array}$$

So x has also vanished from this sum. It should be clear now that we’ll next multiply equation (IV) by (-7), then add it to equation (V):

$$\begin{array}{r|rcrcr} -7\cdot\text{(IV)} & -35y & - & 14z & = & -112 \\ \oplus\quad \text{(V)} & 2y & + & 14z & = & 46 \\ \hline \text{(VI)} & -33y & & & = & -66 \end{array}$$

We can now read off $y = 2$. Back-substituting as above yields $z = 3$ and, finally, $x = 1$.

2.3. Matrix representation of linear systems

When looking at a linear system

$$\begin{cases} a_{1,1}\cdot x_1 + a_{1,2}\cdot x_2 + \ldots + a_{1,n}\cdot x_n &= b_1 \\ a_{2,1}\cdot x_1 + a_{2,2}\cdot x_2 + \ldots + a_{2,n}\cdot x_n &= b_2 \\ &\vdots \\ a_{m,1}\cdot x_1 + a_{m,2}\cdot x_2 + \ldots + a_{m,n}\cdot x_n &= b_m \end{cases}$$

we see that we need to write down quite a lot of boilerplate in every step: the set $\{x_1, \ldots, x_n\}$ of variables does not change, and we know that the LHS is a sum of *monomials* (that is, products of coefficients and their corresponding variables) so that we can also do away with the plus and minus signs. In this way, we can encode all the system's information using the *augmented coefficient matrix*[4]

$$\left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \ldots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} & b_m \end{array}\right)$$

where every row contains the information of an equation in the system, and every column except for the rightmost one corresponds to an unknown; the rightmost column contains the system's right-hand side.

2.4. Gaussian Elimination

The basic idea of Gaussian elimination is adding equations (that is, rows in the matrix) in such a way that the resulting system is in triangular form so that we can obtain the entire solution recursively. This is done in a specified order – in the beginning, only the first column is treated, and all its entries below the first row are set to zero, essentially by using the addition method. Then the other columns are treated in order until – if the system has a unique solution – the system is in triangular form.

[4] We call it the *augmented* coefficient matrix because it contains the system's right-hand side as its last column. Therefore, if we put only the system's coefficients into a matrix, we call it the coefficient matrix. This is quite useful if the same coefficents, but different right-hand sides occur across a number of systems.

2.5. First example

Over the real numbers, look at the system

$$\left\{\begin{array}{lrcrcrcr} \text{(I)} & x_1 & + & 2x_2 & + & 3x_3 & = & 6 \\ \text{(II)} & x_1 & + & 3x_2 & + & 7x_3 & = & 16 \\ \text{(III)} & 3x_1 & + & 3x_2 & - & 2x_3 & = & -9 \end{array}\right. .$$

If it has a unique solution, then the system can be rewritten as

$$\left\{\begin{array}{lcccc} \text{(I)} & x_1 & & & = \alpha_1 \\ \text{(II)} & & x_2 & & = \alpha_2 \\ \text{(III)} & & & x_3 & = \alpha_3 \end{array}\right. ,$$

where $(\alpha_1, \alpha_2, \alpha_3)$ is the system's unique solution. So, in matrix-speak, we want to get the matrix

$$A := \left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 1 & 3 & 7 & 16 \\ 3 & 3 & -2 & -9 \end{array}\right)$$

into the form

$$U := \left(\begin{array}{rrr|r} 1 & \widetilde{a}_{1,2} & \widetilde{a}_{1,3} & \alpha_1 \\ 0 & 1 & \widetilde{a}_{2,3} & \alpha_2 \\ 0 & 0 & 1 & \alpha_3 \end{array}\right)$$

so that we can read off the solution for the variable x_3 and back-substitute to generate the entire solution. Here I use the notation $\widetilde{a}_{i,j}$ to indicate that the value of $a_{i,j}$ will most likely have changed.

How do we go about this? Well, we have already seen the addition method in action. So if we take the second row of A and subtract the first row from it (that is, multiply the first row by (-1) and add it to the second row, in the terminology of the addition method), we get

$$(1,3,7,16) \quad - \quad (1,2,3,6) \quad = \quad (0,1,4,10) \quad ,$$

we see that the resulting row has a zero in its first component, so we've done the first step in getting to the matrix U. Now we take the third row and again add some multiple of the first row to it so that the first component of the sum is zero. How do we know which multiple to add? Obviously, all we need to do is look at the first component of the first row, which is 1 in our case. So if we multiply the first row by (-3), we get (-3), which we can add to the first component of the third row to get zero. So in effect, we multiply the first row by $\frac{a_{3,1}}{a_{1,1}}$ and subtract the result from the third row, which we'll denote as follows.[5]

$$\text{III} \quad \mapsto \quad \text{III} - \frac{a_{3,1}}{a_{1,1}} \cdot \text{I}$$

It is worth noting that for this to work, $a_{1,1}$ had better not be zero!

Now the matrix is

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 0 & 1 & 4 & 10 \\ 0 & -3 & -11 & -27 \end{array}\right) \quad ,$$

and we'll now treat the second column.

All we do here is produce a zero entry at the position $(3,2)$, that is, the third row should get a zero in the second column. Can we use our familiar recipe here, that is, can we subtract some multiple of the first row from the third row? Well, if we do

$$\text{III} \quad \mapsto \quad \text{III} - \frac{a_{3,2}}{a_{1,2}} \cdot \text{I}$$

that is,

$$\text{III} \quad \mapsto \quad \text{III} + \frac{3}{2} \cdot \text{I} \quad ,$$

we end up with

$$\text{III} \quad \mapsto \quad \text{III} + \frac{3}{2} \cdot \text{I} \quad = \quad (0,-3,-11,-27) + \frac{3}{2} \cdot (1,2,3,6) \quad = \quad \left(\tfrac{3}{2}, 0, -\tfrac{13}{2}, -18\right) \quad ,$$

so there is a zero in the third row of the second column, all right, but *we have destroyed the zero entry in the third row of the first column!*

So are we out of luck? No.

What we can do is subtract some multiple from the *second* row instead of subtracting it from the first row. Let's do

$$\text{III} \quad \mapsto \quad \text{III} - \frac{a_{3,2}}{a_{2,2}} \cdot \text{II} \quad .$$

The result is

$$\text{III} \quad \mapsto \quad \text{III} - \frac{-3}{1} \cdot \text{II} \quad = \quad (0,-3,-11,-27) + 3 \cdot (0,1,4,10) \quad = \quad (0,0,1,3) \quad .$$

Presto! We produced a zero entry at $(3,2)$ and at the same time we retained the zero entry at $(3,1)$.

[5] Technically, the third row has changed as a result of the transformation, so we should introduce a new symbol, like III', III", etc. after each step, but I can't be bothered.

The matrix is now

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 4 & 10 \\ 0 & 0 & 1 & 3 \end{array}\right) \quad ,$$

so we can read off $x_3 = 3$. (So the system has a unique solution.) As usual, we plug this into the second equation to get $x_2 = 10 - 4x_3 = 10 - 4 \cdot 3 = -2$, and, plugging these two values into the first equation, we get $x_1 = 6 - 2x_2 - 3x_3 = 6 - 2 \cdot (-2) - 3 \cdot 3 = 1$. So the system's solution is $(1, -2, 3)$.

2.6. Second example: No solution

Why don't we do a more complex example? Over the real numbers, let's look at the system

Dieses Beispiel fertig stellen.

2.7. Third example: More than one solution

Now what does the resulting matrix look like when the system has more than one solution?

Dieses Beispiel fertig stellen.

2.8. What to do if you can't divide

We noted earlier that for the row transformations

$$\mathrm{III} \quad \mapsto \quad \mathrm{III} - \frac{a_{i,j}}{a_{j,j}} \cdot \mathrm{I}$$

to be meaningful, the entry $a_{j,j}$ must not be zero.

We'll now do a bigger example using a system that has a unique solution. This will also show how to avoid fractions.

$$\left\{\begin{array}{rrcrcrcrcrcrcrcrcr}
\text{(I)} & & + & x_2 & + & 4x_3 & & & + & 2x_5 & - & 3x_6 & + & 3x_7 & = & -7 \\
\text{(II)} & 4x_1 & & & + & 3x_3 & - & 3x_4 & + & x_5 & - & x_6 & + & 2x_7 & = & 5 \\
\text{(III)} & -2x_1 & & & + & x_3 & + & 2x_4 & & & & & + & 3x_7 & = & -5 \\
\text{(IV)} & 2x_1 & - & x_2 & & & + & 3x_4 & & & - & 2x_6 & + & x_7 & = & 13 \\
\text{(V)} & -3x_1 & & & + & 2x_3 & + & x_4 & + & 4x_5 & + & 3x_6 & - & x_7 & = & -21 \\
\text{(VI)} & x_1 & - & 2x_2 & - & x_3 & & & & & + & 2x_6 & - & 3x_7 & = & 6 \\
\text{(VII)} & -x_1 & - & 3x_2 & + & 2x_3 & + & x_4 & & & + & 3x_6 & - & 2x_7 & = & -1
\end{array}\right. .$$

The matrix form of this system is

$$\left(\begin{array}{rrrrrrr|r}
0 & 1 & 4 & 0 & 2 & -3 & 3 & -7 \\
4 & 0 & 3 & -3 & 1 & -1 & 2 & 5 \\
-2 & 0 & 1 & 2 & 0 & 0 & 3 & -5 \\
2 & -1 & 0 & 3 & 0 & -2 & 1 & 13 \\
-3 & 0 & 2 & 1 & 4 & 3 & -1 & -21 \\
1 & -2 & -1 & 0 & 0 & 2 & -3 & 6 \\
-1 & -3 & 2 & 1 & 0 & 3 & -2 & -1
\end{array}\right)$$

Dieses Beispiel fertig stellen.

2.9. Gauss-Jordan elimination

There is a special twist to Gaussian elimination particularly useful for systems that have a unique solution.

Let's return to the system

$$\left\{\begin{array}{rrcrcrcr}
\text{(I)} & x_1 & + & 2x_2 & + & 3x_3 & = & 6 \\
\text{(II)} & x_1 & + & 3x_2 & + & 7x_3 & = & 16 \\
\text{(III)} & 3x_1 & + & 3x_2 & - & 2x_3 & = & -9
\end{array}\right. .$$

with its matrix form

$$A = \left(\begin{array}{rrr|r}
1 & 2 & 3 & 6 \\
1 & 3 & 7 & 16 \\
3 & 3 & -2 & -9
\end{array}\right) .$$

We've already seen that we can manipulate the matrix A into triangular form

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 4 & 10 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

using Gaussian elimination. We then had to do back-substitution to find the entire solution. As the system has a unique solution,

$$\begin{cases} \text{(I)} & x_1 & & & = & 1 \\ \text{(II)} & & x_2 & & = & -2 \\ \text{(III)} & & & x_3 & = & 3 \end{cases} \quad ,$$

it is possible to get the matrix into the form

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \quad .$$

Can this be done using Gaussian-elimination-like manipulations? Indeed it can. First do

$$\mathrm{II} \mapsto \mathrm{II} - 4 \cdot \mathrm{III}$$

to get

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \quad ,$$

then do

$$\mathrm{I} \mapsto \mathrm{I} - 3 \cdot \mathrm{III}$$

to get

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \quad ,$$

finally do

$$\mathrm{I} \mapsto \mathrm{I} - 2 \cdot \mathrm{II}$$

to get

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \quad ,$$

so you can read off the entire solution. Of course this only works if the system has a unique solution. So whenever, at the end of Gaussian elimination, you find that a system has a unique solution, you can do Gauss-Jordan elimination to find the entire solution more quickly than by doing back-substitution.

Gauss-Jordan elimination is also useful if a system does not have a unique solution.

Genauer ausführen.

2.10. Other methods

Other methods of solving linear systems include the „Gleichsetzungsverfahren“ which consists of solving two equations for the same variable and equating the two expressions obtained. Naturally, those expressions no longer contain the variable in question so that, once again, a finite number of steps leaves an equation with only one unknown, and back-substituting generates the entire solution.

3. Quadratic equations

Definition 5. A **quadratic equation** is an equation of the form

$$ax^2 + bx + c = 0 \qquad . \tag{1}$$

Here $a \neq 0, b$ and c are known real numbers[6], the *coefficients*. The left-hand side of this equation is a second-degree polynomial. ax^2 is called the *quadratic term*, bx is called the *linear term*, c is called the *constant term*. If $a = 1$, then the quadratic equation is said to be in *monic form*, otherwise to be in *general form*.

Over the reals, a given quadratic equation does not have to be solvable, and if it is solvable, there are at most two solutions.[7]

So how do we solve these?

3.1. Special cases

Missing linear term. If the linear term is missing (that is, $b = 0$), equation (1) reads $ax^2 + c = 0$, which is equivalent to $x^2 = -\frac{c}{a}$. Therefore, if $-\frac{c}{a}$ is positive or zero, we can take the square root and state the solution(s) as $x_{1,2} = \pm\sqrt{-\frac{c}{a}}$.

Missing constant term. Now equation (1) is $ax^2 + bx = 0$, that is $x \cdot (ax + b) = 0$, which we'll write in the form

$$x = 0 \quad \vee \quad ax + b = 0$$

where the symbol $\vee$ is a fancy way of saying "or". So $x = 0$ is one solution, and the other one is $x = -\frac{b}{a}$.

[6] At least in high school.

[7] There is a field, called the *complex numbers*, in which every quadratic equation is solvable, and there are other sets of numbers where a quadratic equation can have more than two solutions. But those definitely won't be on the exam. ◡̈

3.2. General case.

A number of other ways of solving quadratic equations exist.

Vieta's formulas. For a moment, let's only consider quadratic equations in monic form, that is,

$$x^2 + px + q = 0 \qquad .$$

If we call the two solutions (assuming for a moment that they exist) s_1 and s_2, we may write

$$(x - s_1) \cdot (x - s_2) = 0$$

or, equivalently,

$$x^2 - (s_1 + s_2) \cdot x + s_1 \cdot s_2 = 0$$

from which we see, by *equating coefficients*, that

$$s_1 + s_2 = -p \qquad \text{and} \qquad s_1 \cdot s_2 = q \qquad .$$

So given an equation like $x^2 - 7x - 44 = 0$, what do you do? You look for two numbers s_1 and s_2 such that $s_1 + s_2 = 7$ and $s_1 \cdot s_2 = -44$. (This is a system of equations, albeit not a linear system.) Do such numbers exist? Yes, they do, and all you need is mental arithmetic: $s_1 = 11$, $s_2 = -4$.

This also works in reverse: Given the numbers $s_1 = -7$ and $s_2 = 5$, can you state the [8] quadratic equation with those numbers as its solutions? Yes, you can, without clumsy four-term binomial expansions: $s_1 + s_2 = -2$, $s_1 \cdot s_2 = -35$, and so the equation is $x^2 + 2x - 35 = 0$.

As a side note, Vieta's formulas easily generalise to higher-degree polynomials. By way of example, if the real cubic equation $x^3 + a_2x^2 + a_1x + a_0 = 0$ has the three real solutions s_1, s_2 and s_3, we expand the product, collect like powers, and write

$$\begin{aligned}
x^3 + a_2x^2 + a_1x + a_0 &= (x - s_1) \cdot (x - s_2) \cdot (x - s_3) \\
&= (x^2 - s_1x - s_2x + s_1s_2) \cdot (x - s_3) \\
&= x^3 - s_1x^2 - s_2x^2 + s_1s_2x - s_3x^2 + s_1s_3x + s_2s_3x - s_1s_2s_3 \\
&= x^3 - (s_1 + s_2 + s_3) \cdot x^2 + (s_1s_2 + s_1s_3 + s_2s_3) \cdot x - s_1s_2s_3
\end{aligned}$$

[8] Over the real numbers this problem has, of course, a unique solution, so that we can speak of *the* equation instead of *an* equation.

to conclude, again by equating coefficients, that

$$\begin{aligned} a_2 &= -(s_1+s_2+s_3) \\ a_1 &= s_1s_2+s_1s_3+s_2s_3 \\ a_0 &= -s_1s_2s_3 \qquad . \end{aligned}$$

This is, again, a non-linear system of equations. Already in the case of three equations with three variables, it is rather more difficult to solve offhand than the system with two equations and two variables, but it (and its higher-degree analogues) does have its applications. For example, it can be used to show that any *integer* solution of a polynomial equation with *integer* coefficients divides the polynomial's constant term – a fact that comes in very handy when guessing zeros of high-degree polynomials.

Completing the square. Another method, and perhaps the most intuitive one, is called *completing the square*. We'll first do two examples, then derive a formula for solving quadratic equations from it.

Looking, over the reals, at the equation

$$2x^2 - 12x = 32 \qquad ,$$

we express it in monic form,

$$x^2 - 6x = 16 \qquad ,$$

and see that the left-hand side would be a perfect square if we added 9 to it, because $x^2 - 6x + 9 = (x-3)^2$. So we add 9 to both sides,

$$x^2 - 6x + 9 = 16 + 9 \qquad ,$$

and simplify

$$\begin{aligned} (x-3)^2 &= 25 \\ \iff x-3=5 &\vee x-3=-5 \\ \iff x=8 &\vee x=-2 \end{aligned}$$

to arrive at the solution set $\{-2, 8\}$.

Now consider the quadratic equation

$$-3x^2 - 12x - 15 = 0 \qquad ,$$

again over the real numbers. As before, we put it into monic form,

$$x^2 + 4x + 5 = 0 \qquad ,$$

shift the constant term to the right-hand side,

$$x^2 + 4x = -5 \qquad ,$$

and see that the left-hand side would be a perfect square if we added 4 to it, because $x^2 + 4x + 4 = (x+2)^2$. So we add 4 to both sides,

$$x^2 + 4x + 4 = -5 + 4 \qquad ,$$

and simplify

$$(x+2)^2 = -1$$

to arrive at an equation that is clearly unsolvable, as the square of a real number is always greater than or equal to zero. Therefore, the original equation, $-3x^2 - 12x - 15 = 0$, has no solution over the reals.

It should be noted that completing the square is a much simpler way to show that a quadratic equation is unsolvable than using Vieta's formulas. Revisiting the equation $x^2 + 4x + 5 = 0$, to use Vieta's formulas we are asked to find two real numbers s_1 and s_2 such that $s_1 + s_2 = -4$ and $s_1 \cdot s_2 = 5$. Here, all I can come up with is guessing, because the methods from the previous section, such as the substitution method,

$$s_1 + s_2 = -4 \iff s_1 = -s_2 - 4$$

$$s_1 \cdot s_2 = 5 \iff (-s_2 - 4) \cdot s_2 = 5 \iff -s_2^2 - 4s_2 = 5 \iff s_2^2 + 4s_2 = -5$$

yield the very same quadratic equation we started out with! So one needs to use some other method – completing the square, for example.

We can of course also use this method to factor polynomials. As we can split off a linear factor $(t-a)$ from a polynomial $p \in F[t]$ [a] if and only if p has a *zero* at a – that is $p(a) = 0$ – [b], we solve the equation $p(x) = 0$ to find the factors of p.

As an example, consider the polynomial $t^2 + t - 1 \in \mathbb{R}[t]$. We have to solve the equation $x^2 + x - 1 = 0$. We'll first look at the term $x^2 + x$ only. What must be added to it in order to make it a perfect square? The answer is $\frac{1}{4}$, as $x^2 + x + \frac{1}{4} = \left(x + \frac{1}{2}\right)^2$. Of course we have to subtract that $\frac{1}{4}$ again in order to leave the equation intact, but as we have produced a perfect square we can take the square root and simplify. In this way, we end up with the following decomposition:

$$t^2 + t - 1 = \left(t - \tfrac{-1+\sqrt{5}}{2}\right) \cdot \left(t - \tfrac{-1-\sqrt{5}}{2}\right)$$

$$\begin{aligned}
x^2 + x - 1 &= 0 \\
x^2 + x + \tfrac{1}{4} - \tfrac{5}{4} &= 0 \\
\left(x + \tfrac{1}{2}\right)^2 - \tfrac{5}{4} &= 0 \\
\left(x + \tfrac{1}{2}\right)^2 &= \tfrac{5}{4} \\
x + \tfrac{1}{2} &= \pm\sqrt{\tfrac{5}{4}} \\
x + \tfrac{1}{2} &= \pm\tfrac{\sqrt{5}}{2} \\
x &= -\tfrac{1}{2} \pm \tfrac{\sqrt{5}}{2} \\
x &= \tfrac{-1\pm\sqrt{5}}{2}
\end{aligned}$$

---

[a] $F[t]$ is shorthand for "the set of polynomials in the variable t with coefficients from F", where F, at least in high school, is any field. (See the first footnote on page 9.)

[b] Let's quickly prove that. — For all $p \in F[t]$, let $\deg(p)$ denote the degree of p.

„$\Longrightarrow$" If a is a zero of p, we do polynomial long division to obtain polynomials $r, s \in F[t]$ such that

$$p = (t-a) \cdot s + r \quad , \qquad \deg(r) < 1 = \deg(t-a).$$

Therefore $\deg(r) = 0$, that is, r is a constant polynomial. Observing

$$0 = p(a) = (a-a) \cdot s + r = 0 \cdot s + r = 0 + r = r \quad ,$$

we find $r = 0$, that is, the division leaves no remainder.

„$\Longleftarrow$" If $(t-a)$ divides p with no remainder, we have $p = (t-a) \cdot s$ for some $s \in F[t]$. Plugging in a for t, we get $p(a) = 0 \cdot s(a)$. As $s(a)$ belongs to F and as F does not have any zero divisors, we conclude $p(a) = 0$. Therefore a is a zero of p. $\square$

So what happens here?

$$
\begin{aligned}
ax^2+bx+c &= 0\\
ax^2+bx &= -c\\
x^2+\tfrac{b}{a}\cdot x &= -\tfrac{c}{a}\\
x^2+2\cdot\tfrac{b}{2a}\cdot x &= -\tfrac{c}{a}\\
x^2+2\cdot\tfrac{b}{2a}\cdot x+\left(\tfrac{b}{2a}\right)^2 &= \left(\tfrac{b}{2a}\right)^2-\tfrac{c}{a}\\
\left(x+\tfrac{b}{2a}\right)^2 &= \left(\tfrac{b}{2a}\right)^2-\tfrac{c}{a}\\
\left(x+\tfrac{b}{2a}\right)^2 &= \tfrac{b^2}{4a^2}-\tfrac{c}{a}\\
\left(x+\tfrac{b}{2a}\right)^2 &= \tfrac{b^2}{4a^2}-\tfrac{4ac}{4a^2}\\
x+\tfrac{b}{2a} &= \pm\sqrt{\tfrac{b^2-4ac}{4a^2}}\\
x_{1,2} &= -\tfrac{b}{2a}\pm\tfrac{\sqrt{b^2-4ac}}{2a}\\
x_{1,2} &= \tfrac{-b\pm\sqrt{b^2-4ac}}{2a}
\end{aligned}
$$

We start out with the quadratic equation in the general form, bring it into monic form and manipulate the LHS in order to apply the binomial theorem. The term $\frac{b}{2a}$ is the „quadratische Ergänzung“ that is used to complete the square.
Then take the square root and simplify, and there you have it: a simple formula you can stick values into and read off the solutions. But don’t neglect the other methods if they can save you time!

References

[1] Wikipedia contributors, *"System of linear equations"*; Wikipedia, the free encyclopedia; URL: `http://en.wikipedia.org/w/index.php?title=System_of_linear_equations&oldid=448970441#Elimination_of_variables`, accessed 14 September 2011.

[2] Teschl, Gerald/Teschl, Susanne: *Mathematik für Informatiker. Band 1: Diskrete Mathematik und Lineare Algebra*, Springer, Berlin/Heidelberg, 2. Auflage 2006–2007, S. 307–313, S. 331.

[3] Roolfs, Günter: *Gleichungssysteme mit drei Variablen*, URL: `http://nibis.ni.schule.de/~lbs-gym/klasse9pdf/GleichungssytemedreiVariablen.pdf`, accessed 14 September 2011.

[4] Autoren der Wikipedia: *„Quadratische Gleichung"*; Wikipedia, die freie Enzyklopädie; URL: `http://de.wikipedia.org/w/index.php?title=Quadratische_Gleichung&oldid=93562987#a-b-c-Formel_.28Mitternachtsformel.29`, abgerufen am 18. September 2011.

[5] Autoren der Wikipedia: *„Quadratische Ergänzung"*; Wikipedia, die freie Enzyklopädie; URL: `http://de.wikipedia.org/w/index.php?title=Quadratische_Erg%C3%A4nzung&oldid=100191628#L.C3.B6sung_einer_quadratischen_Gleichung`, abgerufen am 26. März 2012.

[6] Brill, Manfred: *Mathematik für Informatiker. Einführung an praktischen Beispielen aus der Welt der Computer*, Hanser, München, 2. Auflage 2005, S. 214. Online bei Google Books, URL: `http://books.google.de/books?id=R3SzmBJBjJ4C&pg=214`, abgerufen am 27. November 2011.

[7] Furlan, Peter: *Das gelbe Rechenbuch 1. Lineare Algebra, Differentialrechnung*, Verlag Martina Furlan, Dortmund, Druck im Jahr 2010, S. 68–78. URL: `http://www.das-gelbe-rechenbuch.de/band1.shtml`, abgerufen am 16. Mai 2012.

[8] Autoren der Wikipedia: *„Peano-Axiome"*; Wikipedia, die freie Enzyklopädie; URL: `http://de.wikipedia.org/w/index.php?title=Peano-Axiome&oldid=101246838#Axiome`, abgerufen am 8. Juni 2012.

[9] DaMenge (Bewohner von Matroids Matheplanet): *„Beweis einfacher Rechengesetze"*; Matroids Matheplanet; URL: `http://www.matheplanet.com/matheplanet/nuke/html/article.php?sid=316`, abgerufen am 8. Juni 2012.