

2 Number Systems

Student Group

First Name	Surname	Matrikel Nr.

Table of Contents

2. Number System	2
2.1 Number System	2
2.1.1 Addition Systems	2
2.1.2 Decimal and Binary	3
2.1.3 Other bases	5
Hexadecimal	5
Octal	6
2.1.4 From decimal to other bases	6
1. handling the integer decimal	6
2. handling the decimal places	7
2.1.5 Binary Coded Decimals	8
2.1.6 Marking the base of a numeral	8
When to use which base?	9
2.2 basic arithmetic operations in binary and hexadecimal	9
2.2.1 Addition	9
In Binary	10
In Hexadecimal	11
2.2.2 Subtraction	11
In Binary	11
In Hexadecimal	12
Note!	12
2.2.3 Multiplication and Division	13
further links	13

2. Number System

2.1 Number System

In the previous chapter we had a look onto the way a processor (and a computer) can deal with the digital values '0' and '1'. However, we haven't seen how the processor can handle larger numbers. In order to approach this, first a short historical outline is shown.

2.1.1 Addition Systems

The first used number system were **addition systems**. These are also still in use: when enjoying a German beer in the beer garden the waiter is counting the 'progress' by dropping dashes onto the coaster (see [figure 1](#))

Fig. 1: number of drinks on a German Bierdeckel (coaster)



In ancient Rome these systems were deeper elaborated. There were different symbols which represent numbers of various sizes:

- \$I = 1\$
- \$V = 5\$
- \$X = 10\$
- \$L = 50\$
- \$C = 100\$
- \$D = 500\$
- \$M = 1000\$

Besides this representation for quantities also the position of the symbol in the **numeral** were important:

- In general: the letters have to be arranged decreasing from left to right. For example $\text{MDCLXVI} = 1601$
- There are deviations of this rule: When up to three of the a lower symbol is written to the left, these have to be subtracted. Sounds complicated?
A kind of.... for example $\text{MCMXLIV} = 1944$.
Luckily, we do not have to learn this, but by this trick the length of the numeral could be shortened,

It becomes even more complicated, when trying to calculate with the numbers: what is the result of the multiplication $\text{CCMXXXVII} \cdot \text{DDIIX}$?

2.1.2 Decimal and Binary

Luckily, we have nowadays a better system for writing numbers: the [positional system](#).

We 'just know' what a number like 23 means. However, for understanding how the computer works we have to investigate this gut feeling and put some technical terms onto it.

1. We are accustomed to count with our fingers from 1 to 10 . For this we have 10 symbols to count: $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$. These group of distinguishable symbols are called **digits**.
2. The amount of the digits is called **base** B . We are used to the decimal base $B=10$, in logic we used binary (also called dual) $B=2$.
3. When we count beyond the maximum number we are used to 'enlarge the number to the left': after the 9 , we count 10 . But the digit 1 in 10 is more worth than the 1 in 31 (of course). It is on a different **position**, on the position of the tens.
4. Each position gets numbered: the ones count 0 , the tens 1 , the hundreds 2 , the thousands 3 and so on. This 'position number' is called **index** i .
5. Knowing the index, also the 'worth of the position' can be derived: the **place factor** p (like one, ten, hundred) can be calculated with the base and the index: $p=B^i$.
6. A **numeral** as a group of digits represents what is commonly known as a number.
7. A **code** or **encoding** means a way to translate one way to display information into another. E.g. A decimal numeral into a Binary, or an idea of an algorithm into a computer language.

In order to recapitulate this for $B=10$ we will calculate the amount of a decimal numeral here once in detail (click on the arrow to the right " $>$ " to see the next step, alternatively see [here](#)):

Ok, that was simple. But what about a binary number? Let's calculate the amount of a binary numeral:

So, what did we find out?

- The shown process is a relatively simple way to convert binary numerals to decimal.
- A 8 digit binary numeral is equal an up to 3 digit decimal numeral. --> Numerals in binary become lengthy

Therefore, it would be better to have a more structured way in presenting the numerals in the which are used in the processor. Internally, the processor just knows 0's and 1's. But investigating a huge bunch of these (e.g. when analyzing the internal memory or a file) is not really catchy in order to understand anything.

The first step is to group the bits:

- 4 bits are called a **nibble** (the name derives from 'to bit' and 'to nibble')
- 8 bits are called a **byte**
- 16 bits are (usually) called a **word**. In details, this depends on the processor.
- 32 bits are (usually) called a **double word**. Like the word this also depends on the processor.

By this, one can separate parts of information (e.g. in a file) better. It is also important to mark the order of the bits. For decimal numerals like \$42\$ the rightmost digit has always the lowest value. The technical term for the 'lowest value' in a binary numeral is called **lowest significant bit** or *LSB*. When the LSB is on position 0 (= rightmost) this order is called **LSB 0**. This was used in the calculation above and is commonly used. In some cases (some memory setup and communication protocols) the order is just the other way around. In this case the **most significant bit** is on position 0. This order is therefore called **MSB 0**. Example: $3_{10} = 0000_2$ (LSB 0), $3_{10} = 1100_2$ (MSB 0)

What is still missing are expressions for large amounts of data. We can describe these using prefixes and the powers of two. You may already know the prefixes such as kilo and mega, but it is worth brushing up on the powers of two. The easiest way to illustrate this is with a chessboard. Maybe you know the legend of the inventor of the chessboard, who was granted a wish by the king. His wish was that the king would give him one grain of rice on the first chessboard square and every time twice as much on each subsequent square. We'll play through this briefly, here to write down the powers of two:

- in the first square we enter two to the power of 0, which results in 1.
- In the second square, two to the power of 1, resulting in two.
- Then 4, 8, 16, 32, 64, 128, 256, 512 and then in the next line two to the power of ten, resulting in 1024.

You should definitely remember these sequence of the values for the power of 2! They are not only important for the exam, but also for the following semesters and computer science.

Fig. 2: chessboard with doubling numerals



1024 bit is also called **kbit** or **kilobit** in the semiconductor industry. You will notice here that the kilo is slightly more than 1000. To make it easier to distinguish, according to the ISO or ECE standard you should say **kibi** instead of "kilo" and write kibibit. The same applies to 2^{20} , i.e. 1024 to the power of two: **megabit** has become common there. However, **mebibit**, should be used to differentiate. The nomenclature continues in the same way for 2^{30} and 2^{40} : gibibit and tebibit.

2.1.3 Other bases

When looking onto multiple bases, it is important to clearly **mark the base** of the numerals. Already in the chapter before a numeral like \$110\$ could either be \$110\$ in decimal or \$110\$ in binary, which is \$6\$ in decimal.

In the following the base will be written as a subscript: \$110_2 = 6_{10}\$. In the end of this subchapter we will also see other ways to mark the base.

Hexadecimal

With the ideas of the previous subchapter we already can structure the bits. In order not to use the lengthy binary presentation it is common to use other bases. One important base is \$16\$ for hexadecimal numerals. There, we need 16 distinguishable symbols: \$0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F\$. With this digits it is possible to encode (=rewrite by other means) exactly 4 bit or one nibble. The encoding will be as follows:

Dual	Decimal	Hexadecimal
\$0000_2\$	\$0_{10}\$	\$0_{16}\$
\$0001_2\$	\$1_{10}\$	\$1_{16}\$
\$0010_2\$	\$2_{10}\$	\$2_{16}\$
\$0011_2\$	\$3_{10}\$	\$3_{16}\$
\$0100_2\$	\$4_{10}\$	\$4_{16}\$
\$0101_2\$	\$5_{10}\$	\$5_{16}\$
\$0110_2\$	\$6_{10}\$	\$6_{16}\$
\$0111_2\$	\$7_{10}\$	\$7_{16}\$
\$1000_2\$	\$8_{10}\$	\$8_{16}\$
\$1001_2\$	\$9_{10}\$	\$9_{16}\$
\$1010_2\$	\$10_{10}\$	\$A_{16}\$
\$1011_2\$	\$11_{10}\$	\$B_{16}\$
\$1100_2\$	\$12_{10}\$	\$C_{16}\$
\$1101_2\$	\$13_{10}\$	\$D_{16}\$
\$1110_2\$	\$14_{10}\$	\$E_{16}\$
\$1111_2\$	\$15_{10}\$	\$F_{16}\$

Tab. 1: hex encoding

This directly reduces the necessary amount of digits to show data. The hexadecimal representation is for example used in the file type *.hex. This is the output file of an embedded c compiler and contains code in an machine-readable representation. An example of a c code and its hex-file is shown in [figure 3](#). In the hex-file the bytes (= 2 nibbles = 2 digits) are visibly grouped.

Fig. 3: example of c code and hex-file  

The bytes in the first line are:

Byte0	Byte1	Byte2	Byte3	Byte4	Byte5	...
\$10_{16}\$	\$00_{16}\$	\$00_{16}\$	\$00_{16}\$	\$19_{16}\$	\$C0_{16}\$	...

But what the decimal value of this numerals?

This transfer is also possible the process shown in 2.1.1:

$Z_{10} = \sum_{i=-n}^m z_i \cdot B^i$, where z_i is the digit on position i .

As an example the Byte4 shall be written in decimal:

$Z_{10}(\text{Byte4}) = Z_{10}(19_{16}) = \sum_{i=0}^1 z_i \cdot 16^i = 1 \cdot 16^1 + 9 \cdot 16^0 = 16 + 9 = 25_{10}$

Octal

Another - much less common - base is 8. This number system is called **octal**. In this case only $0,1,2,3,4,5,6,7$ can be used for digits.

Similar to the number systems before, the following formula applies for the transfer in decimal: $Z_8(X) = \sum_{i=-n}^m z_i \cdot 8^i$

2.1.4 From decimal to other bases

Up to this point chapter, we only have converted other B -based numerals to decimal. Now we want to try to convert a decimal into hexadecimal, for example the numeral 315_{10} . One way is to first convert to binary by "try and subtract" and then convert the nibbles into hexadecimal:

For this example, we first try to subtract the highest power of 2, which not results in a negative number.

$$\begin{array}{rcl} 315_{10} & - & 256_{10} = 59_{10} \\ 59_{10} & - & 32_{10} = 27_{10} \\ 27_{10} & - & 16_{10} = 11_{10} \\ 11_{10} & - & 8_{10} = 3_{10} \\ 3_{10} & - & 2_{10} = 1_{10} \\ 1_{10} & - & 1_{10} = 0_{10} \end{array}$$

The number 307_{10} is similar to $2^8 + 2^5 + 2^4 + 2^3 + 2^1 + 2^0$. This is equal to 100111011_2 or $0001;0011;1011_2$. For the hexadecimal value these nibbles have to be converted, which leads to $13B_{16}$. This is often the fastest way, but is based on the constraint, that one remember the power of 2.

But how can we convert a decimal numeral like 452.12_{10} e.g. to hexadecimal?

The easiest way is to at first separate the value in **integer decimal z** and **decimal places f** :

$452.12 \rightarrow z = 452, f = 0.12$

1. handling the integer decimal

The integer decimal z can be converted the abovementioned method. A different way is via repeatedly dividing by the base (here $B=16$) and using the remainder:

$$\begin{array}{rcl} 452 & : & 16 = 28 \text{ R } 4 \\ 28 & : & 16 = 1 \text{ R } 12 \\ 1 & : & 16 = 0 \text{ R } 1 \\ 0 & : & 16 = 0 \text{ R } 0 \end{array}$$

For each of the shown steps, the integer of the division of the step before is used (28 , 1).

The last steps (and any following) result in a remainder of 0 .

For hexadecimal numeral we have to focus on the remainder from below to top and convert the value into hexadecimal digits

each position in decimal	1	12	4
each position in hexadecimal	1	C	4

The result is $1C4_{16}$

2. handling the decimal places

The decimal places 0.12 can be converted by repeatedly multiplying by the base (here $B=16$) and using the integer:

```
\begin{align*} 0.12 \cdot 16 &= \textcolor{blue}{1} \textcolor{blue}{.92} \quad \textcolor{blue}{.92} \cdot 16 = \textcolor{green}{14} \textcolor{green}{.72} \quad \textcolor{green}{.72} \cdot 16 = \textcolor{brown}{11} \textcolor{brown}{.52} \quad \textcolor{brown}{.52} \cdot 16 = \textcolor{red}{8} \textcolor{red}{.32} \quad \textcolor{red}{.32} \cdot 16 = \textcolor{grey}{5} \textcolor{grey}{.12} \quad \textcolor{grey}{.12} \cdot 16 = \textcolor{blue}{1} \textcolor{blue}{.92} \quad \textcolor{blue}{.92} \cdot 16 = \textcolor{green}{14} \textcolor{green}{.72} \quad \dots \quad \textcolor{blue}{.92} \cdot 16 = \textcolor{green}{14} \textcolor{green}{.72} \quad \dots \end{align*}
```

For each of the shown steps, the decimal places of the multiplication of the step before is used ($\textcolor{blue}{.92}$, $\textcolor{green}{.72}$, $\textcolor{brown}{.52}$, $\textcolor{red}{.32}$, $\textcolor{grey}{.12}$, $\textcolor{blue}{.92}$).

Decimal place in the second last line is equal to the first one. Therefore also any further places will also be equal. That leads to repeating decimals.

For hexadecimal numeral we have to focus on the integer from top to below and convert the value into hexadecimal digits.

each position in decimal	1	14	11	8	5	1	14	...
each position in hexadecimal	1	E	B	8	5	1	E	

The result is $0.\overline{1EB851E}_{16}$

Results:

1. A decimal numeral has to be separated into integer decimal and decimal places.
2. By dividing / multiplying with the base the integer decimal / decimal places can be converted to another base. This works for all other bases like 2 or 8 .
3. The results have to be converted (as least for base $B > 10$, like hexadecimal).
4. A small decimal places can lead to a longer (or even infinite) numerals in another base.

Especially the last result has a mayor impact for calulations on microcontrollers and computer: The internal logic is only based on binary, which also show this problem. However, the internal memory for a numeral is limited.

Even when stored in 32bit - it is not possible to exactly convert the 0.12_{10} to binary. In the following table the n -bit equivalent of 0.12_{10} in binary and hexadecimal system is shown.

Additionally, this value is also re-converted to a decimal numeral.

number of bits n	number system	numeral
\$8	binary	0.0001_2
	hex	$0.1F_{16}$
	equiv. dec	0.12109375_{10}
\$16	binary	$0.0001_2; 1110_2; 1011_2; 1000_2$
	hex	$0.1EB8_{16}$
	equiv. dec	0.11999511718_{10}
\$24	binary	$0.0001_2; 1110_2; 1011_2; 1000_2; 0101_2; 0010_2$
	hex	$0.1EB851_{16}$
	equiv. dec	$0.11999994516..._{10}$
\$32	binary	$0.0001_2; 1110_2; 1011_2; 1000_2; 0101_2; 0001_2; 1110_2; 1100_2$
	hex	$0.1EB851EC_{16}$
	equiv. dec	$0.12000000011..._{10}$

This might seem as a little issue. But there are a lot of areas, where exact decimal places are mandatory, like banking, or simulations.

2.1.5 Binary Coded Decimals

The first approach to this was the development of **Binary Coded Decimals**.

The encoding algorithm from a decimal numeral into BCD is simple: “don't use the division/multiplication method mentioned before - just take the same hexadecimal digit on each decimal position like the decimal one”.

This means the decimal numeral 391.21_{10} is encoded to 391.21_{BCD} . Inside the processor each digit is handled as hexadecimal number: $3_{16}; 9_{16}; 1_{16}; 2_{16}; 1_{16}$ equals $0011_2; 1001_2; 0001_2; 0010_2; 0001_2$.

The main disadvantage is the ineffective storage management and more complex algorithms for addition, subtraction and so on.

2.1.6 Marking the base of a numeral

Up to here, the marking was done by the subscript. Postfixes cannot be used in the software development environment, and sometimes are also not used in datasheets. Alternative ways for the marking are the following:

base	subscripted	prefixed (code)	postfixed
2 (binary)	0010_{10_2}	0b00101010	00101010B
10 (decimal)	1027_{10}	1027	1027D
8 (octal)	1027_{8}	01027	1027O
8 (hexadecimal)	$27D_{16}$	0x27D	27H (also \$27)

Be aware, that in the code 01027 is not equal to 1027!

In Binary

In the following some examples shall show the concept for binary:

```
\begin{align*} \begin{array}{lll} \{ a) \color{white}{+}0_2 \color{white}{+}0_2 \\ \color{white}{+}0_2 \color{white}{+}1_2 \\ \color{white}{+}1_2 \color{white}{+}0_2 \\ \color{white}{+}1_2 \color{white}{+}1_2 \end{array} \quad \begin{array}{l} \color{white}{+}0_2 \\ \color{white}{+}1_2 \\ \color{white}{+}0_2 \\ \color{white}{+}1_2 \end{array} \quad \begin{array}{l} \color{white}{+}0_2 \\ \color{white}{+}1_2 \\ \color{white}{+}0_2 \\ \color{white}{+}1_2 \end{array} \end{align*}
```

The four simplest examples show the carry only for $d) \quad 1_2 + 1_2$. Now we can also connect the number system with the logic gates! The addends are called A and B (e.g. $A=1_2$, $B=1_2$), the sum is the numeral CS_2 (e.g. $C=1$, $S=0 \rightarrow CS_2=10_2$).

When analysing the calculations above, we need one logic gate for S , which only results 1 , when either $A=1$ or $B=1$: this is the XOR gate.

For the carry C we only get 1 , when both of the inputs are one: Here a AND gate have to be used.

The [figure 4](#) show the resulting logic. This is also called a half adder.

Fig. 4: Simulation of a half adder

The next step shall be a bit more complicated - or better: some bit more complicated. We want to do the calculation $A + B = 0011_2 + 0111_2$.

```
\begin{align*} \color{white}{+}00\boldsymbol{11}_2 \color{white}{+}01\boldsymbol{11}_2 \\ \color{white}{+}00\boldsymbol{11}_2 \color{white}{+}01\boldsymbol{11}_2 \\ \color{white}{+}00\boldsymbol{11}_2 \color{white}{+}01\boldsymbol{11}_2 \\ \color{white}{+}00\boldsymbol{11}_2 \color{white}{+}01\boldsymbol{11}_2 \end{align*}
```

The rightmost, first step is easy, since we already had this in the examples before. The next step (marked in bold) is differing: not only do we have to consider the digits from A and B , but also the carry from the calculation before. The carry from before has to be added for all the next steps similarly.

For the calculation by hand, we did four individual additions from right to the left. The processor has to do the same. But first we have to expand the half adder. For a complete addition step we need a logic with the inputs A , B and carry from the step before C_{in} . The output will be still S and C .

Fig. 5: Simulation of a full adder

In [figure 5](#) the full adder is shown. It is an alternative representation from what we have done for calculation by hand:

1. first add one bit from A to one bit from B (half adder). The result is in S' and C' .
2. Then, take the result S' and add it with C_{in} (second half adder). The result is in S and C'' . C' is already the wanted output.
3. For the carry output we need to consider both carries C' and C'' . When two or all of the inputs A , B , C_{in} are high, then the carry has to be set to high. This can be implemented

`r{\red}{\boldsymbol{1}}}{\boldsymbol{1}}\overset{}{1_2}} \end{align*}`

In this example the bold column will be explained shortly:

```
\begin{align*} \color{white}{-}\,\,\boldsymbol{1}_2 \,\, - \boldsymbol{1}_2 \,\, \overline{\color{red}{1}} \,\, \color{white}{0} \,\, \overset{\color{red}{\boldsymbol{1}}}{\boldsymbol{1}}_2 \,\, \end{align*}
```

The calculation has to be executed as follows: $\boldsymbol{1}_2 - (\boldsymbol{1}_2 + \color{red}{\small{\boldsymbol{1}}}) = \boldsymbol{1}_2$. Additionally, another carry has to be taken from the next digit.

In Hexadecimal

The calculation in hexadecimal is conceptually again the same. Some examples, which we will discuss below:

```
\begin{align*} \begin{array}{lll} \{ a) \,\, \color{white}{-}15_{16} \,\, - \color{white}{1}3_{16} \,\, \\ \,\, \overline{\color{white}{-}\overset{}{12_{16}}}} \,\, \& \{ b) \,\, \color{white}{-}23_{16} \,\, - \\ \color{white}{B}6_{16} \,\, \overline{\color{white}{-}\overset{\color{red}{1}}{1}D_{16}}} \,\, \& \{ c) \,\, \\ \color{white}{-}3F_{16} \,\, - 1A_{16} \,\, \color{white}{-} \\ \overline{\color{white}{B}\overset{}{25_{16}}}} \,\, \& \{ d) \,\, \color{white}{+}\,\,38_{16} \,\, \\ + 1A_{16} \,\, \color{white}{+}\,\, \overline{\overset{\color{red}{1}}{1}\overset{}{B_{16}}}} \,\, \end{array} \end{align*}
```

a) This one is (again) simple: looks (again) like a decimal formula..

b) Here, the both hexadecimal numerals look like decimal numerals, but the result $3_{10} - 6_{10} = 7_{10} + C$ lead to an underflow, where we have to take a carry of $C = -10_{10}$ for a decimal calculation. In hexadecimal the carry differs. A better way to solve such a subtraction in hexadecimal is to add the carry before: $3_{16} - 6_{16} + \color{red}{10_{16}}$. Then the calculation can be converted to decimal: $3_{10} - 6_{10} + \color{red}{16_{10}} = 19_{10} - 6_{10} = 13_{10}$. This result has to be converted back to hexadecimal: $13_{10} = D_{16}$. For the next column the carry has to be considered.

c) For this formula the idea from b) helps, too: for each digit by digit subtraction, we will have a look, whether a transformation to decimal is beneficial. For the rightmost it is: $F_{16} - A_{16} = 15_{10} - 10_{10} = 5_{10}$. The rightmost result is now: $5_{10} = 5_{16}$.

d) We can do the same thing here, for $8_{16} - A_{16}$: consider the carry, convert to decimal, calculate, re-convert to hexadecimal. $8_{10} - 10_{10} + \color{red}{16_{10}} = 24_{10} - 10_{10} = 14_{10}$. The result is E_{16} .

Note!

The subtraction within the processor is done a bit differently, with the [two's complement](#). In short: the range of numerals will be divided in such a way that negative numbers can also be represented. By this the subtraction can be transformed into an addition. It is roughly like transforming $10_{10} - 4_{10}$ into $10_{10} + (-4)_{10}$.

