

Signals in digital systems

1 Dual-state signals

2 Three-state signals

3 Dynamic parameters of digital signals

Signals in digital systems

In digital systems, all the signals can achieve only a **finite** number of discrete values. This is different from analog systems in which the signals are continuous and may acquire in the defined range infinity number of values. In a digital system, **time** can also be a discrete quantity, i.e. signals can only change at certain moments. Such digital systems are then called **synchronous** - in contrast to asynchronous systems, in which signal changes can occur at **any** time. Synchronous systems are more common, because the existence of precisely determined moments of signal changes introduces "order" into **timing of signals** in the system and thus facilitates its design and operation. Accurate timing is ensured by clock pulses, which is a very important signal in the system.

Digital systems are divided into two groups - **combinational** and **sequential**. In combinational systems, the output signals depend only on the current input signals. In sequential systems, the output signals depend not only on the current input signals, but also on the past input signals. The system therefore has internal **memory**.

1 Dual state signals

As already mentioned, digital signals have only a finite number of discrete values. In the vast majority, these are just two values. Two-stage signals allow very modest tolerances in the manufacturing of circuits and hence their low cost. One signal value will be expressed by a lower voltage, the other by a higher voltage. We denote the two possible values of the signal as '0' and '1' (in accordance with the notation in Boolean algebra) . Their assignment to a lower or higher voltage is possible in two ways:

- for **positive** logic, lower voltage corresponds to '0' and higher voltage to '1'
- for **negative** logic, higher voltage corresponds to '0' and lower voltage to '1'

The same rule would apply with a negative supply voltage. For example, assigning a voltage of -10V to a value of '0' and -2V to a value of '1' would correspond to positive logic (which may be somewhat surprising). For the most systems, positive logic is used and will be automatically assumed in future texts.

The logic values are assigned to **voltage zones** (bands) - see Fig. 1. There is a zone of lower voltage L and a higher voltage H. All voltage values falling within the range L are identified as U_L and all voltage values within the zone H are identified as U_H . The limit values are U_{Lmax} and U_{Hmin} . The supply voltage is marked as U_{CC} .

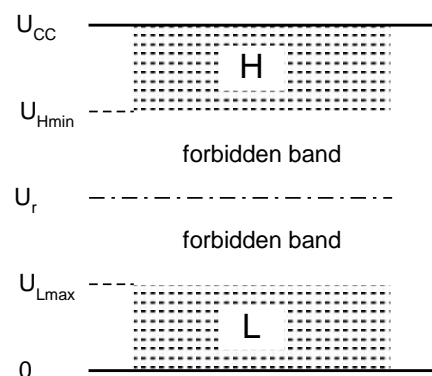


Fig. 1: Voltage bands L and H for digital signals

In positive logic, all voltage values $\leq U_{Lmax}$ would correspond to '0' and all voltage values $\geq U_{Hmin}$ would correspond to '1'. There is a **forbidden band** between these limits, through which the signal passes only when the status of the circuit changes, i.e. in **transients**. The long-time existence of voltages corresponding to the forbidden band can lead to unforeseen behavior of the circuit - possibly to its **oscillation**.

Between the extreme limits U_{Hmin} and U_{Lmax} lies the **reference voltage** U_{ref} - usually in the middle of the forbidden band. Some data on the dynamic parameters of the circuit relate to this voltage.

Under **ideal conditions**, i.e. at voltage values guaranteed within the permissible bands, and with infinitely fast transitions between them, it is possible to work with logical values '0' and '1' instead of voltage values. This greatly simplifies the design of the system, as the Boolean algebra apparatus can be used. In reality, however, other influences need to be taken into account:

- Transients cannot be neglected.
- Correct voltage levels can be disturbed by overloading digital components.
- Interfering signals from the surroundings may penetrate into the system, or the system may generate them by its own operation. Interfering signals are superimposed on the original signals.

Design using Boolean algebra is thus only part of the overall design.

If one signal can take two values, then a group of n signals can take 2^n values. Grouping of signals is very common in practice - e.g. addresses or data in a computer, etc. A group of signals with the same purpose is called a **vector** and a group of wires carrying these signals is called a **bus**. The vector has n **components**. A vector with $n+1$ signals will take $2^{n+1} = 2 \cdot 2^n$ values. Thus, adding one signal **doubles** the resolution. This is an important feature of digital systems.

Individual signals can be grouped in a **vector** that carries information in some code - for example, numbers in binary code. Then the individual components of the vector must be **marked** so that it is clear which signal corresponds to the **highest bit of the number** (data, addresses) and which to the **lowest bit**. The designations **MSB** for the most significant bit and **LSB** for the lowest bit are commonly used for this - see Fig. 2.

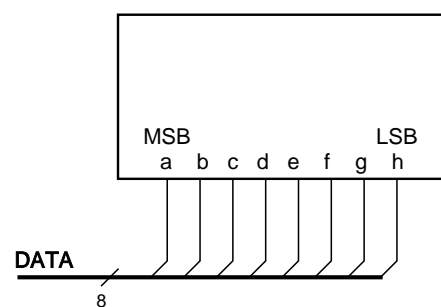


Fig. 2: Vector and its components

So far, we have noticed **voltage** signals, which is the most common case. **Current** signals are also used, even though less often. The current can flow in two directions and thus distinguish the logical value '0' and '1'. The advantage of current signals is that when the current changes, the voltage does not have to change (or only a very little) and thus the undesired effect of **parasitic capacitances** in the circuit is considerably reduced.

2 Three-state signals

In addition to the values '0' and '1', the three-state signals can also take on the value 'Z'. The value 'Z' corresponds to the state when the circuit that generates the signal does not supply any current (or only a negligible current) and is therefore effectively "disconnected". The 'Z' state is also called **high impedance**. A circuit with an output in the 'Z' state does not affect the voltage at its output in any way - if the circuit were separate, its output voltage would not be defined at all. However, the voltage is usually defined, for example, by a resistor connected to ground or to a supply voltage.

Components with three-state outputs can be connected in parallel. However, of the whole group of outputs connected in parallel, at most one can be in the '0' or '1' state, all others must be in the 'Z' state.

3 Dynamic parameters of digital signals

An approximately trapezoidal waveform is typical for digital signals. Therefore, the determination of dynamic parameters is relatively simple. The most important is to determine the **signal propagation delay** as the delay from the signal passage through the **reference level** at the input and the reference level at the output of the component. This period is referred to as t_{pd} .

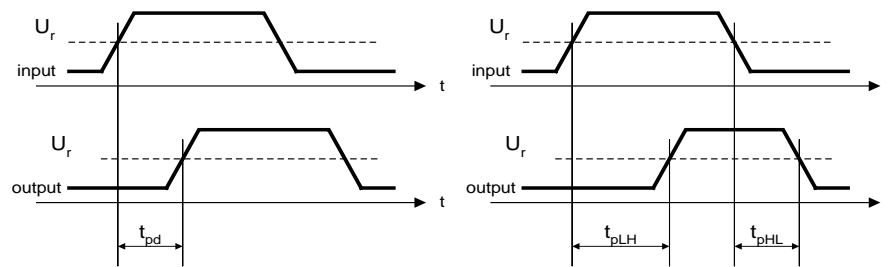


Fig. 3: Main dynamic parameters of a digital signal

In the simplified time diagram on the left side of Fig. 3, the delay is roughly the same when changing the input signal from L to H and from H to L. In fact, these delays may differ, as shown in the right figure. Then we distinguish the time t_{pLH} (when changing the input signal from L to H) and t_{pHL} (when changing in the opposite). For rough orientation, the parameter t_{pd} is calculated as an average of both.

In many cases, it is necessary to express only the moment of signal **change**, regardless of whether it is from L to H or vice versa. Figure 4 shows such an example where the delay t_{AD} between issuing an address and issuing data in a computer is marked. Both address and data are to be understood as vectors.

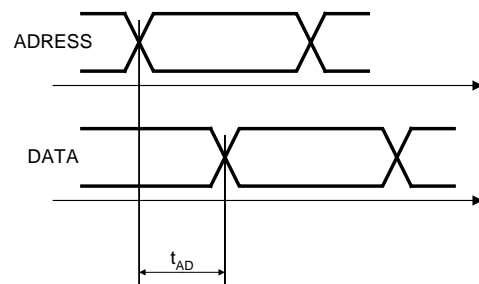


Fig. 4: Representation of points of signal changes

The delay times of common digital components (integrated circuits) on the printed board are of the order of **ns** (nanoseconds, 10^{-9} s), the delays inside the integrated circuits are of the order of units up to tens of **ps** (picoseconds, 10^{-12} s).

With such short times, the signal delay in the **connectors** must also be taken into account. On a normal printed circuit board, the delay is about 1 ns per 10 cm of path length. The signal delay on the chip is in units of ps per 0.1 mm. It can easily happen that the delay in the connection is greater than the delay in the component.

This can play an important role especially in the distribution of clock pulses. The timing of events at one end of a printed circuit board or chip is then different from that at its other end. The existence of delays in the connections substantially complicates the design of large-scale digital systems.

Other important parameters of digital signals are the **rise** and **fall** times of the pulse. These are generally measured at the signal intersections with the levels of 10% and 90% from the pulse amplitude - see Fig. 5. Definition of a **leading edge** and **trailing edge** is also apparent.

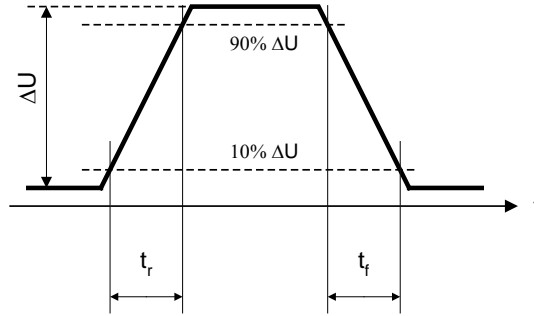


Fig. 5: Definition of rise and fall times

In the specification of the manufacturers of digital devices, rise and fall times might be defined otherwise.

The rise and fall time depend on the properties of the component, as well as on its load - mainly on the **capacitive component** of the load. Current required for charging and discharging load capacitance is approximately given as:

$$I_C = C \cdot \frac{\Delta U}{\Delta t} ,$$

where Δt is the whole time of the transient event, i.e. $1/(0.8 \cdot t_r \text{ or } t_f)$. However, the output current I_C of the component under test is limited, usually to a few mA. Therefore, the **slope** of the pulse $\Delta U/\Delta t$ will also be limited. For example: for a component with a voltage jump of 2 V, with an output current limited to 2 mA, a pulse slope of 2 V/ns in the unloaded state, and the capacitive load $C=10$ pF ($= 10^{-11}$ F), the transient time would extend to

$$\Delta t = C \cdot \frac{\Delta U}{I_o} = 10^{-11} \cdot \frac{2}{2 \cdot 10^{-3}} = 10^{-8} \text{ [s] , i.e. 10 ns.}$$

This would be probably a significant deterioration.

The capacitive loads of the components are both the **input capacitances** of the connected components (typically several pF per each input) and the parasitic capacitances of the **conductors**, which is roughly 1 pF per 1 cm of length. It depends, of course, on the design of the printed circuit board. The above numerical example is, therefore, quite realistic - it corresponds to the load by one component and about 5 cm of signal path.

The situation in an integrated circuit (on the chip) is basically similar, but with much lower delays and with shorter paths.

To maintain speed at higher capacitive loads, it is necessary to choose components with a **high output current** - usually tens of mA. Another method is to minimize the length of critical connections. However, it is to be expected that the system constructed of many interconnected circuits of smaller integration will always be slower than the system **integrated** in a single component, in which the length of the connections and parasitic capacitances is much lower.