Chapter 5
Electronics for Scanning Probe Microscopy

First we discuss some fundamental issues of electronics, such as voltage divider, low-pass filter, and operational amplifier. Then we continue to discuss topics more closely related to scanning probe microscopy such as the current amplifier in scanning tunneling microscopy and feedback electronics, which in SPM serves to stabilize the tip-sample distance. We close this chapter on electronics by discussing how digital-to-analog converters and analog-to-digital converters work in principle.

5.1 Voltage Divider

One of the simplest electronic circuits is the voltage divider, which is shown in Fig. 5.1a. Applying Kirchhoff's law and Ohm's law to this circuit results in the following equations

$$\begin{aligned} V_{in} &= V_1 + V_2 = I(R_1 + R_2) && \text{(Kirchhoff's voltage law)} \\ V_2 &= R_2 I = V_{out} && \text{(Ohm's law)} \end{aligned} \tag{5.1}$$

These equations can be solved for

$$\frac{V_{out}}{V_{in}} = H = \frac{R_2}{R_1 + R_2}. \tag{5.2}$$

The output voltage divided by the input voltage is the called transfer function H. We have assumed here that the output voltage is measured with an infinite inner resistance, i.e. no current flows at the output. The limiting cases for the transfer function are $H \approx 1$ for $R_1 \ll R_2$ and $H \approx R_2/R_1$ for $R_1 \gg R_2$.

B. Voigtländer, *Scanning Probe Microscopy*,
NanoScience and Technology, DOI 10.1007/978-3-662-45240-0_5

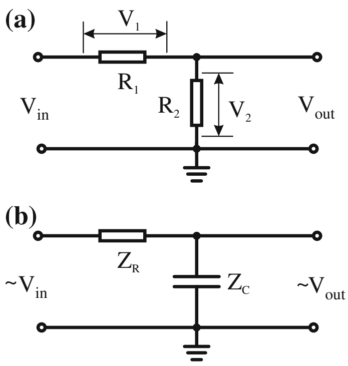

Fig. 5.1 **a** Circuit scheme of a voltage divider. The transfer function is given by $H = V_{\text{out}}/V_{\text{in}} = R_2/(R_1 + R_2)$. **b** This circuit is also a voltage divider, however, now R_2 is replaced by a capacitor and an AC input voltage is considered. Thus, we use the complex impedances Z_R and Z_C in order to obtain the transfer function

5.2 Impedance, Transfer Function, and Bode Plot

In the previous section, we considered DC voltages and currents. In the AC case, the voltages and currents can be written in the complex notation as

$$V = V_0 e^{i(\omega t + \varphi_V)}, \quad \text{and} \quad I = I_0 e^{i(\omega t + \varphi_I)}. \tag{5.3}$$

Of course, for ohmic resistors Ohm's law still reads as $V = RI$. For capacitances and inductors the concept of resistance can be extended to a complex impedance, which is defined as

$$\begin{aligned} Z_{\text{C}} &= \tfrac{1}{i\omega C} && \text{for a capacity } C, \text{ and} \\ Z_{\text{L}} &= i\omega L && \text{for a inductance } L, \text{ and of course} \\ Z_{\text{R}} &= R && \text{for a resistor } R. \end{aligned} \tag{5.4}$$

For the impedances, the equivalent of Ohm's law applies as $V = ZI$. For AC circuits, including several impedances Z, the usual Kirchhoff laws apply, and the rules for parallel and series resistors also hold for impedances, if the quantities are represented in a complex form.

As an example, we consider the circuit shown in Fig. 5.1b, which is similar to the voltage divider, except that one resistor is replaced by a capacitor and an AC input voltage is applied. Thus we consider the complex impedances Z_R and Z_C. The transfer function (now dependent on the frequency) can be calculated in analogy to (5.2) as

$$H(\omega) = \frac{V_{\text{out}}}{V_{\text{in}}} = \frac{Z_C I}{(Z_R + Z_C)I} = \frac{\frac{1}{i\omega C}}{R + \frac{1}{i\omega C}} = \frac{1}{1 + i\omega RC}. \tag{5.5}$$

The transfer function is a complex quantity. In the Bode diagram, the absolute value (modulus) of the complex transfer function and the phase difference between output voltage and input voltage are plotted, as shown in Fig. 5.2a. The corresponding equations are

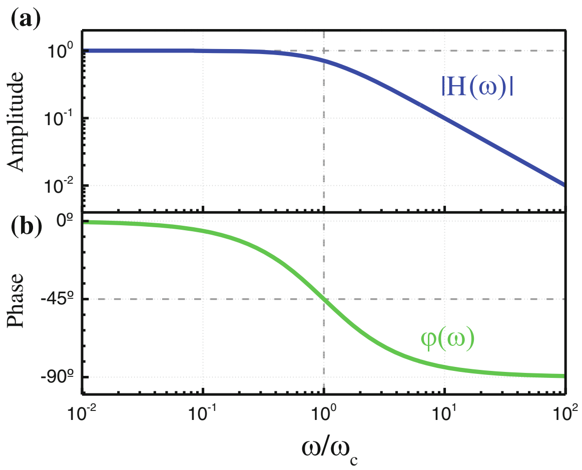

Fig. 5.2 The Bode plot shows the absolute value of the complex transfer function (gain) (**a**) and the phase shift of the output relative to the input signal (**b**). The figure shows the Bode plot of the circuit in Fig. 5.1b. The behavior of the absolute value of the transfer function (amplitude) approaches the value one for frequencies lower than the corner frequency, and decreases for higher frequencies, which is the characteristic of a low-pass filter

$$|H(\omega)| = \frac{|V_{\text{out}}|}{|V_{\text{in}}|} = \frac{1}{\sqrt{1+\omega^2 R^2 C^2}}, \quad \text{and} \quad \varphi_V = \arctan(-\omega RC). \tag{5.6}$$

For frequencies lower than the corner frequency $\omega_c = 1/(RC)$, the absolute value of the transfer function approaches unity, i.e. gain $|V_{\text{out}}| / |V_{\text{in}}|$ is one. For frequencies much larger than ω_c the absolute value of the transfer function decreases as $1/\omega$. At the corner frequency, the gain has the value $1/\sqrt{2}$ (which corresponds to -3 dB). In conclusion, the circuit shown in Fig. 5.1b is a low-pass filter, which transmits signals up to the frequency ω_c with gain one and suppresses signals with higher frequencies. Another way to express this is that this circuit corresponds to a low-pass filter with a bandwidth of $\omega_c = 1/(RC)$.

The phase behavior of this low-pass is shown in Fig. 5.2b. The phase shift is zero for frequencies much lower than the corner frequency and goes to $-90°$ for frequencies much larger than the corner frequency.

The analysis of the low-pass circuit was one simple example, another one is if the resistor and the capacitor in Fig. 5.1b are exchanged. This circuit corresponds to a high-pass filter. Also more complicated circuits can be analyzed using Kirchhoff's laws or the rules for impedances in parallel or in series. One requirement for the type of analysis described in this section is that the input signal V_{in} is a sinusoidal signal. If the transfer function for all frequencies is known this characterizes the behavior of the circuit at all frequencies. This is a basis to obtain the output signal for all periodic functions via Fourier methods.

5.3 Output Resistance/Input Resistance

In Fig. 5.3a we consider a device connected to a voltage source. Any kind of signal source can be replaced by an ideal voltage source with an resistor in series, which we call output resistance R_{output}, as shown in Fig. 5.3a. If the output of the signal source is connected to the input of a device, this can change the output voltage V_{out}, being no more identical to the ideal voltage source V_{signal}. The voltage V_{out} depends also on R_{input}, the input resistance of the device connected to the source. The circuit shown in Fig. 5.3a is (again) a voltage divider. Using (5.2) the output voltage V_{out} can be written as

$$V_{\text{out}} = V_{\text{signal}} \frac{R_{\text{device}}}{R_{\text{signal}} + R_{\text{device}}}, \tag{5.7}$$

and is shown in Fig. 5.3b. It can be seen that the output voltage approaches the signal voltage if $R_{\text{input}} \gg R_{\text{output}}$.

However, in relevant cases of small signal sources of sensors like photodiodes (in the case of atomic force microscopy), the inner resistance of the signal R_{output} is high. In such cases a so called impedance converter is used, which we discuss in Sect. 5.5.1 in order to convert the high output resistance of the signal source to a very low output resistance at the output of the impedance converter, which can be connected to devices with a modestly low input resistance, always maintaining the relation $R_{\text{input}} \gg R_{\text{output}}$.

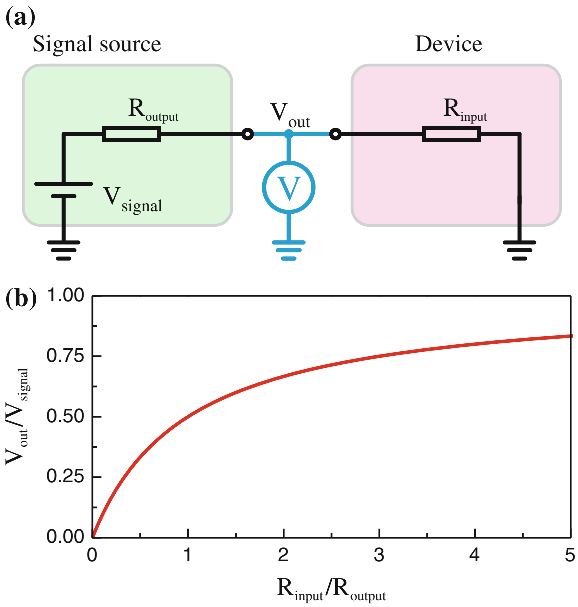

Fig. 5.3 **a** Signal source, consisting of an ideal voltaget source V_{signal} and an output resistance R_{output}, connected to a device with an input resistance, characterized by the resistance between the input and the ground R_{input}. **b** The output voltage for this circuit approaches V_{signal} only if $R_{\text{input}} \gg R_{\text{output}}$

The concept of output resistance and input resistance can be applied in sequence when connecting electronic circuits one after another. We can assign to each device in a sequence of devices an input resistance and an output resistance. In order to avoid the input of the next device modifying the output of the previous device, the relation the relation $R_{\text{input}} \gg R_{\text{output}}$ should be always maintained.

Here we considered the DC, however, the concept of output and input resistances can be extended to the AC case using the impedance replacing the resistance. Furthermore, this concept can also be used for active devices like circuits with operational amplifiers, discussed later.

5.4 Noise

If we consider a DC electric signal with some time-dependent fluctuations such as the current $I(t)$ or the voltage $V(t)$, it can be characterized by its average

$$\langle V\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T V(t)dt. \tag{5.8}$$

Fluctuations of the voltage around this average are called the noise as $\Delta V(t) = V(t) - \langle V\rangle$. This is still a time-dependent quantity and its average is zero. If the noise is due to random fluctuations, it is usually characterized by the following time independent quantity

$$\sqrt{\langle \Delta V^2\rangle} = \sqrt{\lim_{T\to\infty}\frac{1}{T}\int_0^T (V(t) - \langle V\rangle)^2 dt}. \tag{5.9}$$

also called root mean square (RMS) noise.

The above considerations about the noise were in the time domain, i.e. considering the time-dependent signal $V(t)$ and the time-dependent noise $\Delta V(t)$. In the following, we will consider the frequency dependence of the noise. The frequency dependence of the noise can be characterized by the power spectral density (PSD)[1] $N_V^2(\omega)$. An important property of the power spectral density of the noise is that it relates to the mean square noise as

$$\langle \Delta V^2\rangle = \int_0^\infty N_V^2(f)df = \frac{1}{2\pi}\int_0^\infty N_V^2(\omega)d\omega. \tag{5.10}$$

[1] The power spectral density of the noise $\Delta V(t)$ can be defined via the Fourier transform of the noise as $N_V^2(\omega) = \lim_{T\to\infty}\frac{1}{2\pi T}\left|\int_0^T \Delta V(t)e^{-i\omega t}dt\right|^2$.

If a detection scheme is used which measures the noise variable only within a certain (angular frequency) bandwidth $B_\omega = \omega_2 - \omega_1$ between ω_1 and ω_2, the mean square noise can be written as

$$\left\langle \Delta V^2 \right\rangle = \frac{1}{2\pi} \int_{\omega_1}^{\omega_2} N_V^2(\omega) d\omega. \tag{5.11}$$

This expression can also be considered as defining the noise power spectral density $N_V^2(\omega)$. The noise PSD indicates how much power the noise signal carries in a small region around ω. The noise amplitude spectral density is defined as $N_V = \sqrt{N_V^2}$. If the noise spectral density is constant between ω_1 and ω_2 and zero outside, (5.11) reduces to

$$\left\langle \Delta V^2 \right\rangle = \frac{1}{2\pi}(\omega_2 - \omega_1) N_V^2(\omega), \tag{5.12}$$

and we obtain

$$\sqrt{\left\langle \Delta V^2 \right\rangle} = N_V \frac{1}{\sqrt{2\pi}} \sqrt{B_\omega}. \tag{5.13}$$

The (constant) noise spectral amplitude density of the noise variable ΔV is expressed in the unit of the noise variable per $\sqrt{\text{rad} \cdot \text{Hz}}$, for instance volt/$\sqrt{\text{Hz}}$. The actual RMS value of the noise variable measured with a specific bandwidth is then given by the noise amplitude spectral density times the square root of the bandwidth. Note that the angular frequency bandwidth $B_\omega = \omega_2 - \omega_1$ is defined as angular frequency, i.e. in units of rad/s, not cycles/s. Similarly, the unit of the noise power spectral density $N_V^2(\omega)$ is volt^2/(rad $\cdot$ Hz). If the natural frequency f is considered, (5.13) reads

$$\sqrt{\left\langle \Delta V^2 \right\rangle} = N_V \sqrt{B}, \tag{5.14}$$

with $B = f_2 - f_1$ and N_V in volt/$\sqrt{\text{Hz}}$.

5.5 Operational Amplifiers

Since operational amplifiers are used in several parts of STM electronics a brief introduction to their operation is given. An operational amplifier can be considered as a "gain block" amplifying the difference between the input voltages (ideally possessing very high gain). The voltage at the output is the amplified voltage difference at the inputs. Outside of the gain block there is a feedback network (e.g. consisting of resistors), which controls the actual gain. Operational amplifiers operated close to DC have typically the following properties:

- Very high input resistance, with a typical input current of a few pA,
- Very low output resistance, typically a few ohm,
- Very large open-loop voltage gain G (10^4–10^6).

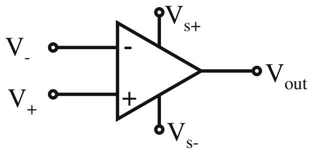

Fig. 5.4 Block diagram of an operational amplifier showing the supply voltages V_s, the input voltages $V_\pm$ and the output voltage V_{out}

We will show that if these properties of an operational amplifier are met the characteristics of the amplifier are determined by the feedback network only, not the gain block itself. We are not concerned with the inner working of the operational amplifier. A block diagram of an operational amplifier is shown in Fig. 5.4. The output voltage is the difference of the input voltages multiplied by the open loop gain G as

$$V_{\text{out}} = G(V_+ - V_-). \tag{5.15}$$

Due to the very high open loop gains of operational amplifies, they are usually not operated in an "open" configuration, because any voltage difference exceeding the sub-millivolt range will saturate the output voltage which is limited to the supply voltage V_s.

5.5.1 Voltage Follower/Impedance Converter

If we connect the output of an operational amplifier to its negative (inverting) input (Fig. 5.5) and apply a voltage signal to the non-inverting input, we will find that the output voltage of the op-amp closely follows that input voltage.

In order to find an expression for V_{out} for the circuit in Fig. 5.5 we start from (5.15) which states that the output voltage is the difference of the input voltages times the open loop gain. In our case the positive input voltage V_+ is V_{in} and the negative feedback voltage V_- is due to the negative feedback V_{out}. Thus (5.15) reads

$$V_{\text{out}} = G(V_{\text{in}} - V_{\text{out}}), \tag{5.16}$$

which leads to

$$V_{\text{out}} = V_{\text{in}} \frac{G}{1+G}. \tag{5.17}$$

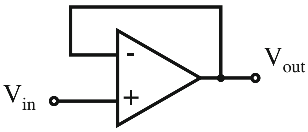

Fig. 5.5 Operational amplifier wired as a voltage follower. A negative feedback is realized by connecting the output to the negative (inverting) input

For a large open loop gain, the output voltage is approximately equal to the input voltage $V_{\text{out}} \sim V_{\text{in}}$.

Taking the output voltage of the operational amplifier and coupling it to the inverting input is a technique known as negative feedback. In this circuit the operational amplifier has the capacity to work in a linear mode, as opposed to merely being fully saturated (due to the high gain) with no feedback for voltage differences exceeding the mV range.

Here, as in the other operational amplifier circuits we will discuss, the actual gain (which is one here) is not determined by the open loop gain of the operational amplifier but by the outer feedback circuit (which is just a simple connection between V_{out} and $= V_{-}$). One could think that an amplifier with a gain of one is useless. However, this circuit acts as an impedance converter, since a high input resistance/impedance (being an intrinsic property of an op-amp) is converted to a low output resistance/impedance (being another intrinsic properties of an op-amp).

While having "only" a voltage gain of one, the voltage follower has a power (current) gain. The voltage follower is often used as "buffer" to interface a large impedance output signal to device with a low impedance (input) load. The voltage follower as impedance converter acts as "one-way" device for signals, drawing almost no current from the source supplying its input (because of its high input resistance), and it can supply a large amount of current to loads with low (input) impedance.

5.5.2 Voltage Amplifier

If we add a voltage divider to the feedback wiring (Fig. 5.6) only a fraction of the output voltage is fed back to the inverting input. In this case the output voltage is a multiple of the input voltage.

The gain of this circuit can be calculated taking the basic equation (5.15) into account. If the output is connected to the inverting input, via a voltage divider network, V_{-} can be written (using Ohm's and Kirchhoff's laws[2]) as $V_{-} = V_{\text{out}} \frac{R_1}{R_1+R_2} = V_{\text{out}} K$, and V_{in} is connected to the positive input V_{+}, then

$$V_{\text{out}} = G(V_{\text{in}} - KV_{\text{out}}). \tag{5.18}$$

Solving this equation for $V_{\text{out}}/V_{\text{in}}$, we find

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{G}{1+KG}. \tag{5.19}$$

[2] $V_{\text{out}} = V_1 + V_2 = I(R_1 + R_2) = (V_1/R_1)(R_1 + R_2) = V_{-} \frac{R_1+R_2}{R_1}$.

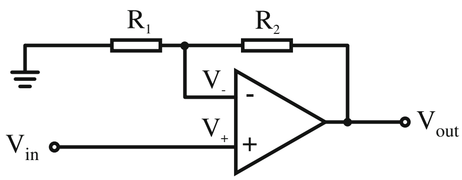

Fig. 5.6 Operation principle of non-inverting amplifier

If G is very large the gain becomes

$$\frac{V_{out}}{V_{in}} = \frac{1}{K} = 1 + \frac{R_2}{R_1}. \tag{5.20}$$

We can change the voltage gain of this circuit just by adjusting the values of R_1 and R_2 (changing the ratio of output voltage which is fed back to the inverting input).

While we have used in the basic equation for the operational amplifier (5.15) together with the analysis of the feedback circuit using Ohm's and Kirchhoff's laws, the analysis of operational amplifier circuits can be simplified using two simple rules. The rule that the input current of an operational amplifier vanishes we have already used in our analysis. In the previous two circuits the difference between the inputs V_+ and V_- approached zero. This is a general rule, leading to the following two "golden rules" which simplify the analysis of circuits with operational amplifiers.

- The input current to an operational amplifier vanishes (high input impedance).
- The difference between the inputs V_+ and V_- approaches zero.

In the following we calculate the output voltage for the circuit shown in Fig. 5.7 using above "golden rules" for operational amplifiers. In this circuit a negative feedback is provided through a voltage divider, but the input voltage is applied to the inverting input and the non-inverting input is grounded. The second "golden rule" tells us that the voltage at the inverting input is zero. Thus, the inverting input is referred to in this circuit as a *virtual ground*, being kept at ground potential (0 V) by the feedback, yet not directly connected to (electrically common with) ground. Since the input current to the operational amplifier is zero (first "golden rule"), the current through R_1 and R_2 are the same. By applying Ohm's law to the two resistors the gain can be calculated as

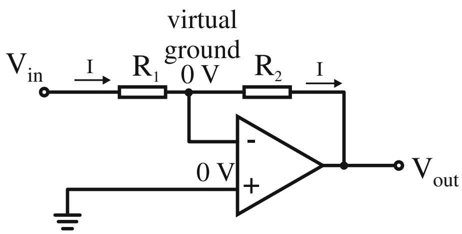

Fig. 5.7 Circuit of an inverting amplifier realized with an operational amplifier

$$\frac{V_{out}}{V_{in}} = \frac{-IR_2}{IR_1} = -\frac{R_2}{R_1}. \tag{5.21}$$

Note that the output voltage always has the opposite polarity of the input voltage. For this reason, this circuit is referred to as an inverting amplifier.

5.6 Current Amplifier

The tunneling current in STM has very small values, typically 0.01–10 nA. The current amplifier is an essential element of an STM since it amplifies the current and converts it to a voltage. Such amplifiers are called transimpedance amplifiers and already the circuit shown in Fig. 5.7 can serve as such a current-to-voltage converter. If we consider the voltage source plus the resistor R_1 as a current source, a current of $I_{in} = V_{in}/R_1$ flows to the virtual ground. Since the input current of the operational amplifier is practically zero (high input resistance), this current flows through the feedback resistor R_2. In the actual current amplifier shown in Fig. 5.8, the input current I_{in} has to flow through the resistor R_{FB}. Therefore, $I_{in} = I_{FB} = -V_{out}/R_{FB}$. Or

$$V_{out} = -I_{in}R_{FB}. \tag{5.22}$$

The input current is converted to an output voltage with R_{FB} as proportionality factor. As an example: If the feedback resistor has a value of $R = 1\,\mathrm{G\Omega}$, one nanoampere of input current results in an output voltage of 1 V. Due to the high input resistance of an operational amplifier and its low output resistance, a high input impedance is converted to a low impedance output which can be processed further.

Up to now we have considered the operational amplifier circuits as DC circuits. In the following, we consider the AC performance of the current amplifier shown in Fig. 5.8 and will show that its bandwidth is limited by the stray capacitance C_{stray} parallel to the feedback resistor. We use the complex impedance to analyze this AC circuit. The complex impedances for a resistor R and a capacity R are $Z_R = R$, and $Z_C = 1/(i\omega C)$, respectively. Since the two impedances in the feedback arm of

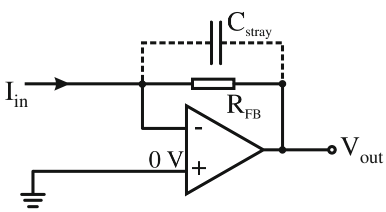

Fig. 5.8 Circuit used as current amplifier in STM. The gain (actually transconductance in V/A) is proportional to the resistance of the feedback resistor R_{FB}. The bandwidth of this current amplifier is limited by the stray capacitance C_{stray}

the operational amplifier are in parallel, the following expression results for the total (complex) impedance Z as

$$\frac{1}{Z} = \frac{1}{Z_R} + \frac{1}{Z_C} = \frac{1}{R} + i\omega C. \tag{5.23}$$

The absolute value of the complex impedance results as

$$|Z| = \frac{R}{\sqrt{1 + (\omega RC)^2}}. \tag{5.24}$$

Replacing according to (5.22) $V_{out} = -ZI_{in}$, and identifying R with R_{FB}, as well as $C = C_{stray}$ results in

$$V_{out} = \frac{-I_{in} R_{FB}}{\sqrt{1 + \left(\omega R_{FB} C_{stray}\right)^2}}. \tag{5.25}$$

This frequency dependence of the output voltage of the current amplifier is the same as that of a simple passive low-pass with a resistor and a capacitor. The corner frequency of such a low pass at which the output voltage drops by $1/\sqrt{2}$ is $f_{corner} = 1/\left(2\pi R_{FB} C_{stray}\right)$. As an example, if by careful design the stray capacitance can be reduced to 0.1 pF a bandwidth of 1.5 kHz is obtained for a feedback resistance of 1 GΩ. The bandwidth of the amplifier is the frequency range which is amplified without significant loss of the signal (i.e. from DC to $f_{corner} \sim 1/(2\pi R_{FB} C_{stray})$. It can be seen that the gain which is proportional to R_{FB} and the bandwidth proportional to $1/R_{FB}$ are opposing figures of merit. Increasing the amplification means decreasing the bandwidth and vice versa. Some numerical examples are given in Table 5.1.

Another figure of merit for amplifiers is the noise. The (RMS) noise induced by the thermal excitation of the electrons in a resistor R is called Johnson noise [12, 13] and can be calculated as

$$I_{noise} = \sqrt{\frac{4k_B T B}{R_{FB}}}. \tag{5.26}$$

with B being the bandwidth and k_B the Boltzmann constant. In Table 5.1 some numerical values are given.

Table 5.1 Gain, bandwidth and noise for a current amplifier with $R_{FB} = 100\,\text{M}\Omega$ and $R_{FB} = 1\,\text{G}\Omega$

$C_{stray} = 0.5\,\text{pF}$	$R_{FB} = 100\,\text{M}\Omega$	$R_{FB} = 1\,\text{G}\Omega$
Gain	10^8 V/A	10^9 V/A
Bandwidth	3 kHz	300 Hz
Noise	0.3 pA	0.1 pA

5.7 Feedback Controller

In scanning probe microscopy, a feedback controller is used to follow the surface topography. Before we come to the application of a feedback controller to SPM, we will consider feedback controllers in general. A general model for a feedback loop is shown in Fig. 5.9. In the control loop, the system output x is measured constantly by a sensor, and compared to the setpoint w by subtraction $w - x$. Depending on this error signal, the controller determines a system input (control signal) y, which is fed into the system in order to adjust the system output x to the set-point value w. This whole operation of the controller acts in a closed feedback loop as shown in Fig. 5.9. The control loop fulfills the task of adapting the system output to the setpoint in the presence of disturbing external noise.

Before we turn to the feedback controller of the STM, let us consider (as an example) a simpler system: the heating system of a house in winter. The simplest example of a feedback system is the on-off controller. On your thermostat you set a certain desired temperature (setpoint) w. If the measured temperature x is lower than w the controller gives a signal y to the system. For the case of the heating system of a house, y is the heating power which is turned on from zero to a certain power; thus the radiators heat the rooms until the set point temperature w is reached. Due to the inertia of the system (i.e. the time delays) the temperature in the rooms will continue to rise for some time after the heating has been switched off (temperature overshoot), because the radiators are still warm. You can easily imagine how this cycle continues. For instance, when the measured temperature x falls below the setpoint temperature w it will take some time before the radiators become warm. In conclusion, the actual temperature x fluctuates around the desired temperature w. What controller theory is all about is to find a smarter way to keep x as close as possible to w. There are two kinds of time delays in the feedback loop: First the time delay in the system itself (this delay is large for the case of the heating of the house and much smaller in the case of STM). For simplicity we will not consider this time delay in the following. Secondly, there is a time delay due to the controller, which we will consider in the following.

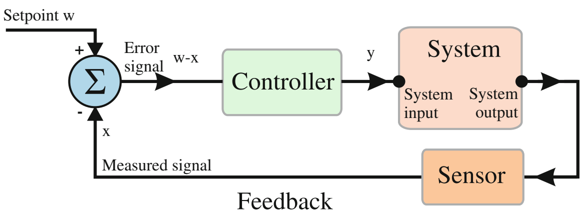

Fig. 5.9 A general model for a feedback loop

5.7.1 Proportional Controller

If in the example of the heating system of a house, a heater with a continuously variable heating power is available (not just on or off), a proportional controller (P controller) can be realized. For the P controller the output of the controller y is proportional to the error signal $w - x$, as

$$y = K_P(w - x). \tag{5.27}$$

The proportional constant K_P is called proportional gain. Since the heating power is now proportional to the error signal it is obvious that the temperature can be controlled much better with much less overshoot than for the on-off controller. (Actually, the on-off controller is a P controller with infinite gain K_P, which is only limited by maximum heating power of the heater). Since the output of the controller is instantaneously proportional to the error signal, the P controller is a fast reacting type of controller.

One problem with the proportional controller is that a pure proportional control will not settle at the set-point value w, but will retain a steady-state error, which is a function of the proportional gain. This can be qualitatively understood as follows. If in the example of our heating system we have continuous losses of heat (outside it is cooler than inside), therefore we need continuous heating power in order to maintain the setpoint temperature, even if the error signal is zero. However, the pure proportional controller does not provide this. According to (5.27) the actuating variable y is zero for zero error signal $w - x$. This means that the pure proportional controller cannot reach the setpoint w. The higher the load (i.e. the cooler it is outside) the greater is the deviation from the set-point value. Increasing the proportional gain can reduce the deviation but it never goes to zero and high gain can lead to instabilities (oscillations) in the feedback loop. The deviation between the output x and the setpoint w is proportional to the heat dissipation (load) and inversely proportional to the proportional gain K_P.

The time delay due to the controller is related to the proportional gain K_P. The greater K_P is, the shorter is the time delay of the controller, i.e. the controller can follow fast. However, a large value of K_P also leads to a larger overshoot.

An example of how a P controller can be implemented using an operational amplifier was shown in Fig. 5.7. The gain constant K_P can be modified by changing the resistances as $K_P = -R_2/R_1$.

In summary the advantage of the P-controller its fast reaction time, the controller output is instantaneously directly proportional to the error signal. The disadvantage of the P controller is the steady-state deviation of the system output from the desired set-point value.

5.7.1.1 Integral Controller

The integral controller provides a control signal proportional to the accumulated deviations from the setpoint. The contribution from the integral term is proportional to both the magnitude of the error and the duration of the error. Summing the instantaneous error over time (integrating the error) corresponds to an accumulated effect that should have been corrected previously. For the I controller the output of the controller y is written as

$$y(t) = K_I \int_0^t (w - x(\tau))d\tau. \quad (5.28)$$

The proportional constant K_I is called integral gain. The integral controller eliminates the residual steady-state error that occurs with a proportional controller. A disadvantage of this type of controller is the slow reaction to changes of the input signal, due to the integration. Of course also the I controller can be made faster (shorter time delay) by increasing K_I, however, this also increases the tendency towards overshooting and instable and oscillating behavior.

In a variant of the I controller, the integration is not performed from zero, but over a time interval Δt prior to the current time.

5.7.2 Proportional-Integral Controller

In a PI controller the P and the I control signals are added up, as shown in Fig. 5.10. In this controller, the advantages of both the P and I controllers are combined, while avoiding their individual disadvantages. Short-term deviations from the setpoint are compensated fast by the proportional controller and long-term deviations are compensated by the integral controller. This type of controller can regulate the error signal to zero in steady-state. The output signal can be written as

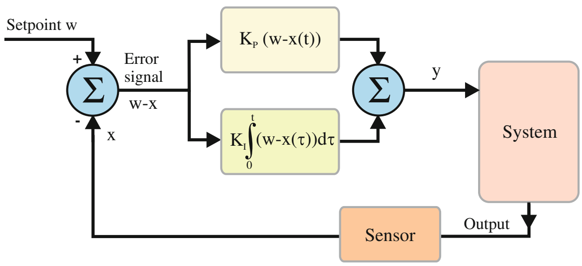

Fig. 5.10 Schematic of a PI controller

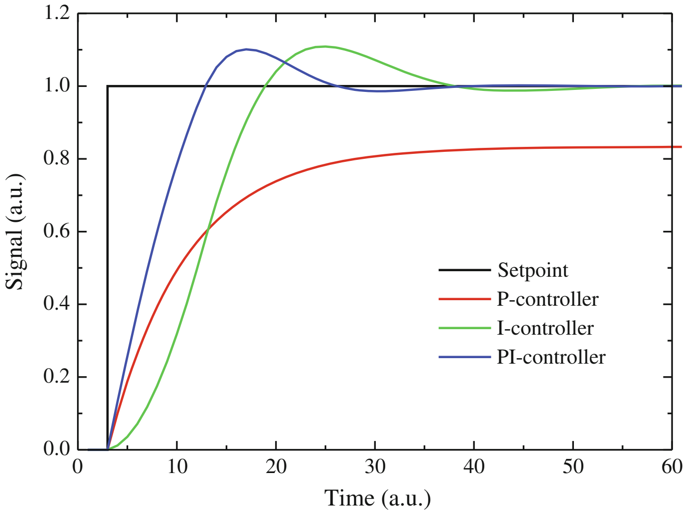

Fig. 5.11 Comparison of the step response of different controllers. The setpoint is a step function which changes from zero to one at time zero. Due to the steady-state error of the P controller the set point is never reached

$$y(t) = K_P(w - x(t)) + K_I \int_0^t (w - x(\tau))d\tau. \tag{5.29}$$

One way to show the performance of controllers is the step response. Step response means that the setpoint is changed instantaneously and the reaction of the controller (and the whole system) to reach the new setpoint is monitored. The step response of different controllers is compared in Fig. 5.11. The P controller does not reach the new set-point value, and the I controller alone is quite slow. The PI controller reaches the setpoint in a reasonable time for an appropriate choice of K_P, K_I.

5.8 Feedback Controller in STM

In STM or SPM in general the elements in the above-mentioned feedback loop have the following correspondence (Fig. 5.12).

- The system output x corresponds to the tunneling current, which is converted to a corresponding voltage by the current amplifier (sensor).
- The setpoint w corresponds to a voltage representing the desired tunneling current.
- The PI controller determines the system input (control variable) y, which is the voltage to be applied at the z-piezo element in order to change the tip-sample distance.

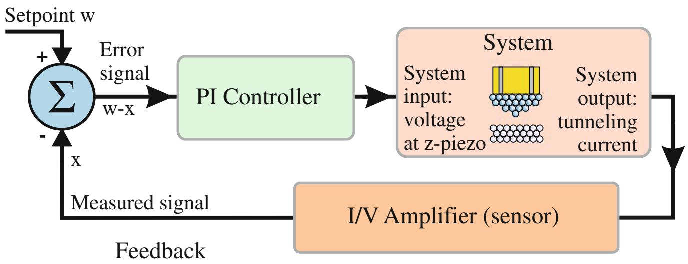

Fig. 5.12 Model of an STM feedback loop

- The most complex part of the feedback loop is the system itself. In the case of STM, it consists of DA converters, the high-voltage amplifiers (HVA) for the z-piezo voltage, the z-piezo element for the vertical positioning of the tip, and the tunneling contact.
- The noise of the z-signal arises due to external mechanical vibrations, the noise of the amplifiers, and the noise of DA and AD converters.

In STM, the P-part of the controller regulates fast deviations from the setpoint such as the atomic corrugation or atomic step edges (here the integrator helps to reach the final value, i.e. eliminates steady-state deviations).

In SPM, there is one effect which excretes the highest load to the feedback controller. Usually the sample is not oriented perfectly parallel to the xy-directions given by the scanner. This slope is usually the largest height signal in the original STM data and will be removed by appropriate background subtraction in the final image, as we will see later. However, the feedback has to follow this slope. As a quantitative example, if the xy-plane of the scanner and the sample surface are 3° off relative to the sample surface, this slope corresponds to a height of 500 Å for a 1 μm wide scan. This is by far the largest height signal compared to, for instance, a few atomic steps (3 Å high) in such an image.

The I controller has the advantage that it is less prone to noise. Depending on the conditions, the measured signal (tunneling current in STM) can be quite noisy. While the P controller reacts immediately to a noise spike of the measured signal, an I controller acts as a low-pass averaging out noise spikes.

Now we consider the problem that a feedback loop may become instable and start to oscillate. If the controller parameters (the gains of the proportional and integral terms) are chosen incorrectly, the feedback loop can become instable, i.e. its output starts to oscillate. An important reason for the instability of the feedback loop is the time delay (reaction) of the system. In our simple example of the heating system of a house, it takes some time after a deviation of temperature is detected before the radiators and the air in the house become hot. In the case of the STM, the time delay of the system is given by the time lag between a change of the z-voltage by the controller and a corresponding change of the tunneling current. Also the speed of the controller itself (given by the gains of the proportional and integral terms) is a source of time delay. It is intuitively clear that a large gain (heating power) and a long delay time of the system will give rise to large overshoots and result in an instability with oscillations of the controller system. A large part of controller theory is concerned with finding conditions for stability of a feedback loop. Here we will only provide a very qualitative intuitive discussion of the stability of a feedback loop.

A different way of characterizing the stability of a feedback loop than the analysis of the step response is to measure the output signal relative to a sinusoidal input signal (transfer function). The transfer function is the output signal divided by input signal. The knowledge of the (sinusoidal) output behavior as function of the sinusoidal input for all frequencies gives complete knowledge of the system response, since any input signal can be represented as a sum of the sinusoidal functions (Fourier theorem). The transfer function is a frequency dependent function and consists of an amplitude and a phase (complex number).

The transfer function of the whole feedback loop can be measured as shown schematically in Fig. 5.13. Initially the feedback loop is enabled and the STM is in tunneling operation. Then the (digital) feedback is switched off and the z-piezo voltage is modulated. The sinusoidal input signal is fed through all analogue components of the STM, HV amplifier, piezo actuator, tunneling junction, and current amplifier, as well as the controller. Then the output signal is measured (amplitude and phase), which results in the frequency dependent transfer function.

The measured transfer function (amplitude part) of the analogue components of a particular STM feedback loop is plotted in Fig. 5.14, as system output divided by system input amplitude as a function of frequency. The characteristics of this transfer function are the characteristics of a low-pass, and the amplitude drops significantly above 4 kHz. This corresponds to the bandwidth of the current amplifier, which is the bandwidth-limiting element of the analogue components in the system. The other elements of the system, HV amplifier (HVA), piezo actuator, and tunneling junction, do not limit the bandwidth of the system.

One very simplified condition for an instable feedback loop is the following: If for a certain frequency the output amplitude is larger than the input amplitude

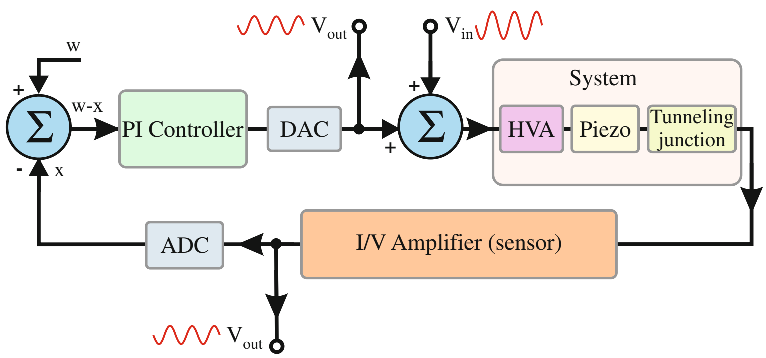

Fig. 5.13 Scheme of the measurement of the transfer function

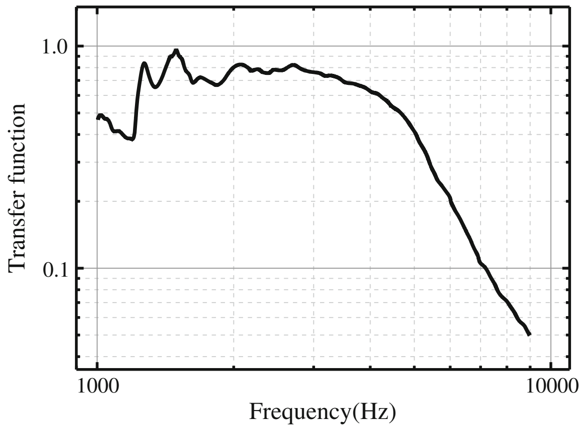

Fig. 5.14 Measured transfer function of the analogue components in the STM feedback loop: HV amplifier, piezo actuator, tunneling junction, and current amplifier

(amplitude of the transfer function ≥1) and the phase for this frequency is close to 0° the feedback loop will become instable. This means that small deviations from the setpoint will build up to an oscillation of large amplitude.

5.9 Implementation of an STM Feedback Controller

Feedback controllers are realized via a digital feedback loop nowadays. The tunneling current is measured by the current amplifier and then the corresponding voltage is digitized by analog digital converters (ADC), as shown in Fig. 5.15. These converters

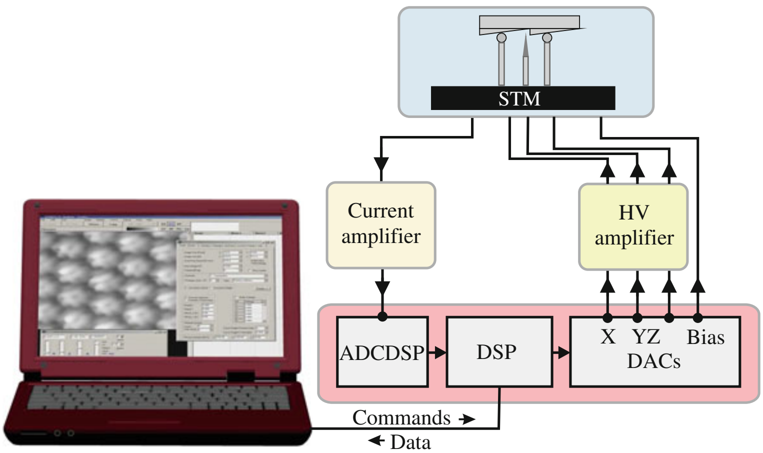

Fig. 5.15 Implementation of computer controlled STM electronics

can have, for instance, an accuracy of 20 bit in a range of ±10 V corresponding to a step width of 20 μV, which is usually far below the noise in the system and therefore sufficient for all practical purposes.

The actual feedback loop is often realized by a digital signal processor (DSP) (Fig. 5.15). A DSP is a own computer on which a single user single-task real-time program runs. From the measured (digitized) current and the current setpoint, the output, i.e. the actuator voltage for the z piezo motion, is calculated using a digitized version of a PI controller. Using a digital feedback loop has several advantages. First, it is very easy to stop the feedback and to perform spectroscopic measurements (i.e. to run a tunneling voltage ramp or a z-ramp), and also to measure the transfer function. Another advantage is that the feedback mode can be changed just by changing the software. The controller algorithm can be changed by a few lines in the DSP program. Furthermore, non-linear algorithms for noise reduction can be implemented.

An example for a pseudocode implementation of a PI controller is given in the following.

```
start
read measured_signal  x(t)
error_signal = set_point - measured_signal  w − x(t)
integral = integral + error_signal * dt  ∫_0^t (w − x(τ))dτ
controller_output = KP * error_signal + KI * integral  y(t)
goto start
```

Once the controller output (new z-voltage) is calculated, this number is converted into an actual voltage by (for instance) 20 bit digital analogue converters (DAC). This z-voltage (range: ±10 V) is then amplified by a high-voltage amplifier to a range of typically ±200 V (Fig. 5.15). This is enough to reach the necessary amplitude of the piezo actuators of a few micrometers. Regarding the resolution, the following reasoning can be applied: For a piezo constant of 60 Å/V and a high-voltage amplifier gain of 20 one DAC unit converts to a z-distance of 2 pm, which is usually more than enough. This means that with the high resolution DA and AD converters available today the digitization of the input and output quantities is no longer a problem since it is far below the usual noise limits. Also the tunneling bias voltage is supplied from the computer via a DAC in order to ramp this voltage in spectroscopic measurements.

When scanning an STM image, the DSP sends the xy-scan data to the DAC. The voltages for the x- and y-electrodes are finally amplified by the high-voltage amplifiers. The data about the height of the tip above the surface, i.e. the voltage applied to the z-piezo, generated by the feedback algorithm running on the DSP, is sent to the PC. The measurement program takes the height of the STM tip above the surface and displays it as an image, i.e. in gray scale as a function of x and y.

The digital control of the STM also allows an automated procedure to be used during the coarse approach of the tip towards the sample. A flow chart for an automated control could be as shown in (Fig. 5.16). After the automatic coarse approach a desired current setpoint is chosen and scanning can be started.

The bias voltage between tip and sample (usually between a few millivolts and a few of volts) can be applied to the sample (sample bias). In this case, the tunneling

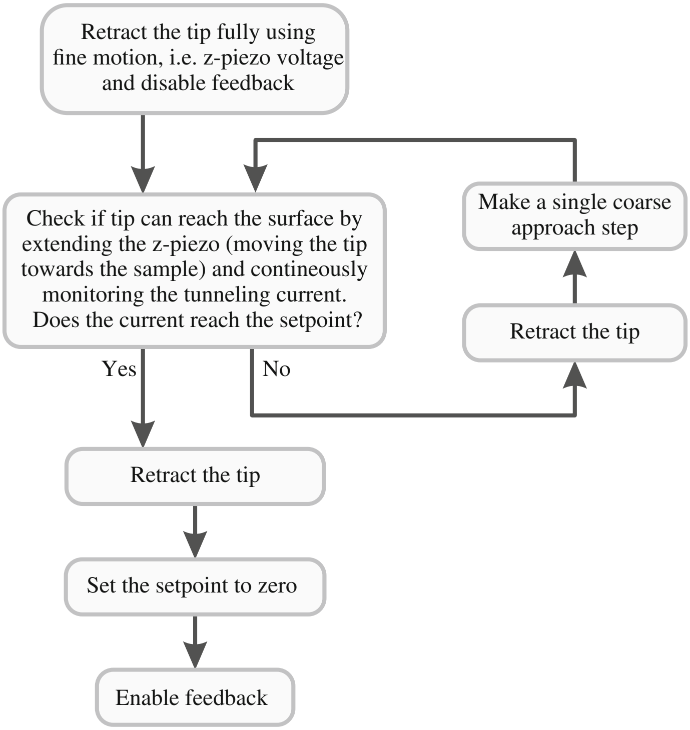

Fig. 5.16 Flow chart of the automatic approach procedure in scanning tunneling microscopy

current is measured relative to the ground. If the sample is grounded, the preamplifier has to float on a bias potential (tip bias) in order to apply a bias voltage between tip and sample.

5.10 Digital-to-Analog Converter

In a computer controlled data acquisition and control system, analog data have to be read to the computer and digital data generated by the computer have to be converted to analog signals. For instance, in scanning probe microscopy the xy-scan signals are generated by a computer program (digital values) and have to be converted to analog signal driving the piezo elements. For this task a digital-to-analog converter (DAC) is used. Here we describe the principle of how such a device can operate. However, actual digital-to-analog converters are more sophisticated than the basic idea explained here.

We assume that the digital signal is already present as voltages (high/low) at several wires of a connector. As an example, we will consider a four-bit signal in

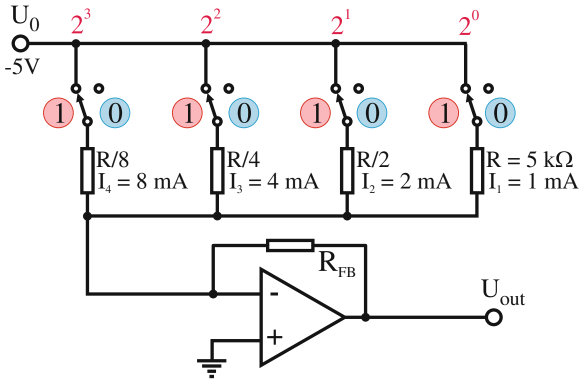

Fig. 5.17 Operating principle of a digital-to-analog converter

the following. In Fig. 5.17, the digital signal is represented by switches either open or closed (−5 V). Each of the lines (switches) has a different weight from 2^0 to 2^3 corresponding to the weight of the bit in the binary digital code. If all switches are open this corresponds to zero (0000), if all wires are connected to −5 V this corresponds to (binary 1111, i.e. 15). The task is now to convert the digitally coded voltage values present at the four connectors to 16 analog voltages relative to ground, ranging, for example, from 0 to 10 V. The resistor following each switch is chosen such that the current through it (when flowing to ground) corresponds to the weight of that bit. The least significant bit (2^0) has, for instance, a 5 kΩ resistor, corresponding to a current of 1 mA to ground, while the most significant bit (2^3) has an 8 times smaller resistor corresponding to an 8 times higher current of 8 mA in this line. All the lines are routed to the inverting input of an operational amplifier acting as a transimpedance amplifier. Since the positive input of the operational amplifier is on ground, the negative input is the virtual ground, as we have considered before. At the point where all these lines are brought together the sum of all the currents flows through R_{FB}. According to (5.22), the analog output voltage at the operational amplifier is

$$U_{\mathrm{out}} = -R_{\mathrm{FB}} U_0 \sum_{\mathrm{i=all\ closed\ switches}} \frac{1}{R_{\mathrm{i}}}. \tag{5.30}$$

The maximum output voltage can be chosen using a proper value for R_{FB}.

5.11 Analog-to-Digital Converter

In scanning tunneling microscopy, the analog voltage at the output of the current preamplifier has to be converted to a number (e.g. 16-bit value) proportional to the analog voltage (tunneling current). For this task, an analog-to-digital converter (ADC) is used. An ADC can be realized by the comparison of the analog signal

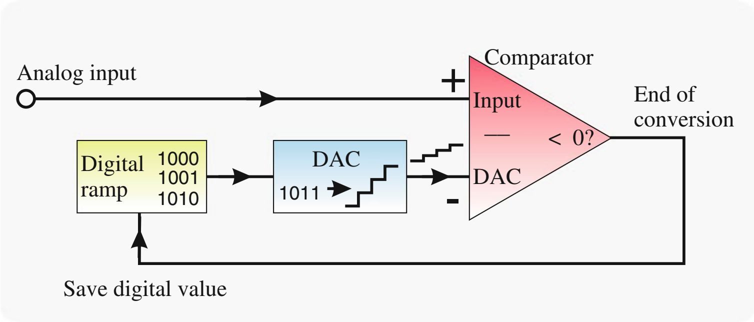

Fig. 5.18 Operating principle of an analog-to-digital converter

(to be digitized) to a voltage from a digitally generated voltage ramp. The principle of operation of one simple ADC is shown in Fig. 5.18. A digital voltage ramp is generated and converted to an analog voltage ramp using a DAC. The value of the generated voltage ramp is compared to the analog input signal to be digitized using a comparator. This comparator has a low digital signal as long as the voltage ramp has a lower voltage than the input voltage. A comparator can be realized by an operational amplifier without external feedback network. Due to its large open loop gain the output will always be maximally positive as long as the negative input voltage is smaller than the voltage at the positive input. The comparator signal changes to logically high if the voltage ramp exceeds the voltage to be measured (Fig. 5.18). This end of conversion signal is then fed to the ramp controller in order to stop the ramp and to read the actual (digital) ramp value. With this digital value of the ramp, a digital value of the analog input signal is saved and the conversion is stopped. Instead of ramping up all digital values from zero, also some interval-based algorithm can be also used in order to find the value closest to the analog input.

5.12 High-Voltage Amplifier

High-voltage amplifiers are needed to drive the piezo elements since the voltages supplied by the digital-to-analog converters are usually only in the range up to ± 10 V and are not high enough to generate sufficient extensions of the piezo elements of several micrometers. Therefore, the DAC voltages are amplified up to about 200 V, which generates the required piezo extensions. We assume here again piezo tubes as piezo elements. Much higher voltages are not advisable because they can lead to a depolarization of the piezo material. A reasonable upper limit for the required bandwidth of the high-voltage amplifiers is the resonance frequency of the piezo element. You cannot move a piezo element at a frequency higher than its resonance frequency. Therefore, 50 kHz is an upper limit for the required bandwidth. In practice, the feedback loop (actually the current amplifier) often has a much lower bandwidth in the range between 1 and 10 kHz. In this case, a low-pass filter at the output of the high-voltage amplifier can be used to reduce the noise. The output noise of the high-voltage amplifiers should be less than 1 mV. With a typical z-piezo constant of about 50 Å/V, this corresponds to a noise in the extension of the piezo in the z-direction of 0.05 Å, i.e. 5 pm.

The piezo motions during scanning are relatively slow. In order to move inertial sliders (Sect. 4.2), saw-tooth signals are applied to the piezo elements and the steepest possible slope of the piezo motion is required. This means a high slew rate (voltage change per time) of the high-voltage amplifier is required. The achievable slew rate depends on the capacitive load at the output of the amplifier, i.e. the capacity of the piezo elements. A high piezo capacity means that a lot of charge has to be pumped to or from the piezo element. If this has to be done in a short time, a high current has to flow. Therefore, high-voltage amplifiers driving piezo elements with a high capacity have to supply a high current in order to achieve a high slew rate. This can lead to problems of high power dissipation in the leads. This problem with the high capacitance occurs mostly for monolithic stacks of piezo elements. They have capacitances in the μF range, while piezo tubes, for instance, have only capacitances in the nF range.

5.13 Summary

- Operational amplifiers are characterized by a very large input resistance, a very low output resistance and a very large open loop gain.
- The actual gain of an operational amplifier including a feedback network is determined by the characteristics of the feedback network, not by the operational amplifier.
- Two golden rules can be applied when analyzing an op-amp circuit: (i) The input current vanishes. (ii) The voltage difference between the inputs is zero.
- A current amplifier converting the input current to an output voltage can be built using an operational amplifier. The output voltage depends on the feedback resistance as $V_{\text{out}} = -I_{\text{in}} R_{\text{FB}}$.
- In the proportional controller, the actuating variable is proportional to the error signal. In the integral controller the actuating variable is proportional to the time integral over to the error signal.
- The transfer function, output signal divided by the input signal (including amplitude and phase), is used to characterize the frequency response of electronic components.