Chapter 8
Process Design and Process Monitoring

8.1 Finding Economical Cutting Parameters

8.1.1 Cutting Parameter Limits

The further development of workpiece materials and cutting tool materials and the increasing automation of machine tools demand an occasional correction of machining standard values.

Standard values should be determined under consideration of machinability and provide reference data concerning depth of cut, cutting speed, feed and the expected tool life parameters for the given tool life criteria. Variation in the machinability of a material of the same standard designation but different batch can be larger than three to one in extreme cases. This ratio is the more unfavourable the larger the permissible wear and the higher the cutting speed [Kluf83]. For an optimal cutting value determination, the tool life parameters and wear properties of the tool are the most important criteria for evaluating the tool life behaviour of a material/cutting tool material combination [Köni71].

Generally, widths of flank wear land of $\mathbf{VB} = 200-800\,\mu\mathrm{m}$ are permitted for flank face wear when cemented carbides are used. The tool life criterion for crater wear should be determined as a function of the feed f with the following equation:

$$\mathrm{KT} = 0.06 + 0.3 \cdot f \tag{8.1}$$

In addition to tool wear, the other main evaluation parameters of machinability – the cutting force, surface quality and chip form – also serve, depending on the case at hand, as tool life criteria. When determining the cutting parameters, the limits of the system workpiece/cutting medium/tool/machine must be taken into consideration. Specification of these limit values refers to the description of the cross-section of undeformed chip using the theoretical parameters uncut chip thickness h and undeformed chip width b and must be converted in accordance with geometrical relations that are true for the respective cutting process into the technical parameters depth of cut a_p and feed f (Chaps. 9 and 10).

In order to obtain the largest possible volume removal rate within a given tool life, first a depth of cut a_p should be selected that is as large as possible provided it is not

F. Klocke, *Manufacturing Processes 1*, RWTH edition,
DOI 10.1007/978-3-642-11979-8_8,

Table 8.1 Maximum of undeformed chip width

Tool included angle ε_r/°	Chip width b
80–90	2/3 · cutting edge length
60	1/2 · cutting edge length
35–55	1/4 · cutting edge length

already defined by the machining allowance to be lifted in one cut. A large depth of cut reduces the number of required steps, so that besides the primary processing time (Sect. 8.1.2), the secondary processing time for lifting, returning and reengagement is economised.

The maximum depth of cut a_{pmax} is limited by the stability of the tool's cutting edge. For roughing, tools with a large tool included angle ε_r are used. The rules of thumb given in Table 8.1 serve as the standard values.

Larger depths of cut or undeformed chip width can lead to cutting edge fractures. After determining the depth of cut, the feed should be set as large as possible in order to minimize the primary processing time. The smaller influence of feed on wear formation compared with cutting speed and the fact that specific cutting force is reduced with increased feed both speak for this sequence in determining the cutting parameters. In this way, the machine's power can be exploited more fully.

In addition to the maximum and minimum permissible feed of the machine, tool limits must again be considered in selecting the feed. In order to avoid unmachined areas on the workpiece, the following inequality can serve as a guide:

$$f_{max} \leq r_\varepsilon \tag{8.2}$$

The lower feed limit is determined using the minimum chip thickness h_{min}, which is a function of the stiffness of the system workpiece/cutting medium/tool/machine and is intended to guarantee chip removal especially in the case of rounded off or chamfered cutting edges. Following SOKOLOWSKI [Soko55], the following can serve as a standard value, for turning for example:

$$0.25 < \frac{h_{min}}{r_n} < 1.125 \tag{8.3}$$

This minimum chip thickness can be limited both by the material (e.g. in the case of austenitic steels, which have a propensity to strain hardening) and by the tool cutting edge form. For example, coated indexable inserts always have a cutting edge rounding due to manufacturing conditions, the radius of which is between 20 and 60 µm. In the case of roughing and when interrupted cuts require a large amount of cutting tool material toughness, these roundings stabilize the cutting edges just like the chamfer of the rake face. In the case of very small chip thicknesses or chip widths, the elastic portion is greatly increased during the separation process and chip removal can no longer be guaranteed.

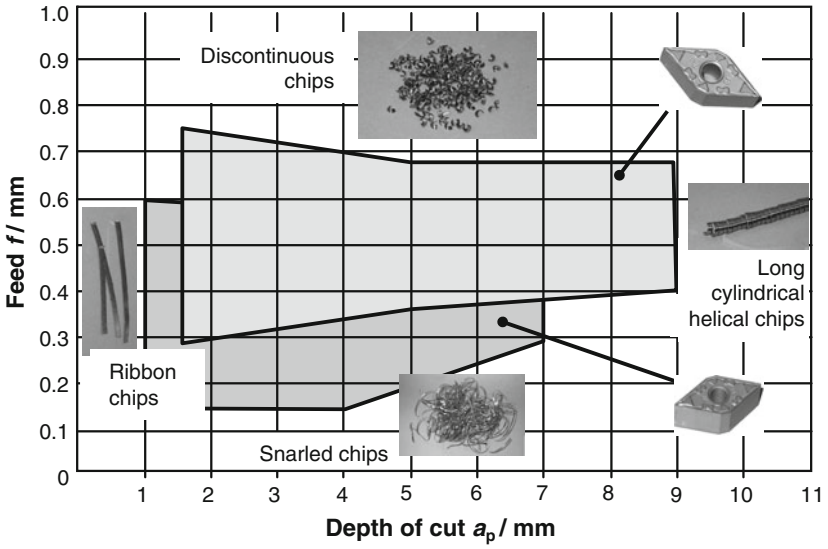

Fig. 8.1 Chip form recommendation for inserts with chip breaker

Another parameter which limits feed is the predefined surface roughness. This depends primarily on the feed, so the feed should not be set too high in order to maintain a certain level of surface quality (Sect. 3.5).

The chip form is of great importance for an undisturbed manufacturing process (especially in the case of automated manufacture with NC and CNC machines) and for the protection of the machine operator. Figure 8.1 shows solution fields of favourable chip forms in feed as a function of the depth of cut for two indexable inserts of the same dimensions with different chip breakers. The images on the edges of the respective solution field show what chip forms can be expected.

Such solution fields are applicable only:

- for the applied manufacturing method,
- for the applied process kinematics,
- for the applied tools,
- for the applied workpiece material and
- for the applied cutting tool material

The cutting speed is determined in accordance with the previously defined tool life parameter on the basis of a preset tool life criterion by a tool life parameter function (Chap. 7) or with standard value tables.

The cutting speeds given in the standard value recommendations are usually applied to stable cutting conditions on pre-machined workpieces. When machining workpieces with rim zone structures that are especially difficult to cut (e.g. forging, rolling or cast iron skin) the given cutting speeds should be multiplied by a factor of 0.65–0.8.

8.1.2 Optimizing the Cutting Parameters

Due to labour costs, machine investment costs, the constantly dropping costs for tool cutting edges and improving wear properties, the cutting parameter recommendations of cutting tool material manufacturers currently refer to a tool life of 15 min. The necessity of shortening the tool life in the case of capital-intensive machine tools is understandable when one considers the costs. Moderate cutting conditions result in a long tool life, little tool change and low tool costs. On the other hand, long machining times also result, which lead to high labour and machine costs depending on the volume machined. Figure 8.2 illustrates the relation between cutting parameters and labour costs.

Since labour and machine costs have increased considerably while tool and tool change costs have risen much more slowly (e.g. by the use of automated tool changers), tool life reduction by increasing the cutting conditions leads to lower manufacturing costs. Improving the cutting tool materials also results in improved wear resistance. This allows for higher potential cutting speeds.

Depending on the machining task, a target value must be selected. In roughing, two optimization targets stand in the foreground, proceeding from the aims of management policy:

- minimal manufacturing costs K_{Fmin}
- minimal allowed time for a process t_{emin}

In the case of finishing, other optimization targets are required. Here, lower workpiece tolerances, predefined surface qualities or other parameters that are important for the functional reliability of the component must be observed. The following is oriented formally and content-wise towards the guideline published by VDI regarding cutting parameter optimization [VDI3321]. When optimizing individual machining processes as well, marginal conditions stemming from the production process such as machine availability or cycle times of linked production plants must be considered when determining the machining parameters. Corresponding to the required optimization target, the optimal value function is derived from the

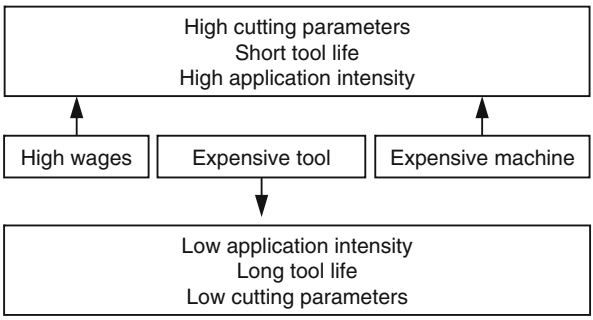

Fig. 8.2 Impact of cutting parameters by values of labour costs, acc. to VDI 3321

manufacturing time or manufacturing cost equation. *Time per unit* t_e is composed of the *basic time* t_g, *additional time* t_v and *recovery time* t_{er}:

$$t_e = t_g + t_v + t_{er} \tag{8.4}$$

The *additional time* t_v is the total time required for all irregular events such as procuring necessary resources.

The *recovery time* t_{er} takes all pauses into consideration during which the machine tools are not in operation.

The basic time t_g is the sum of the *main process time* t_h and *auxiliary process time* t_n*:*

$$t_g = t_h + t_n \tag{8.5}$$

The *main process time* t_h is the time in which direct progress in the sense of the production order is made via machining [VDI 3321].

The following is valid for the main process time:

$$t_h = \sum_i \frac{L_i}{v_{fi}} + \sum_j \frac{S_j}{v_{fj}} \tag{8.6}$$

To calculate the main process time, the feed paths L_i and cut lengths S_j are considered in conjunction with the respective feed velocity v_f. The number of processes required until completion is taken into account by the number of feed paths i and the cut length j.

The *auxiliary process time* t_n is the time during which all indirect processes arising during the machining operation (e.g. tightening, measuring, adjusting, pro rata tool change and workpiece change) are executed [VDI3321]. The following is valid for the auxiliary process time:

$$t_n = \frac{t_r}{t_{WM}} + \frac{t_W}{n_{WT}} + t_{WST} \tag{8.7}$$

Tool change time t_W is the time that passes until a tool is changed, and both the position correction and positioning for re-entry have taken place. This time is partially contained in the auxiliary process time t_{ap} [VDI3321].

The *workpiece change time* t_{WST} is the time that passed until a workpiece is changed.

Since the addends of the auxiliary process times sometimes do not accrue for every workpiece or every process, they are considered proportionately. Set-up time is based on the batch time m:

$$\text{number of workpieces per machine: } n_{WM} = m \tag{8.8}$$

The tool change time is based on the tool life T or the tool operating life n_{WT}:

$$\text{number of workpieces per tool life: } n_{WT} = \frac{T}{t_c} \tag{8.9}$$

Here, the cutting time t_c is the time in which the tool is actually cutting.

Finally, the following is true for the time per unit t_e per workpiece or process:

$$t_e = t_h + \frac{1}{m} \cdot t_r + \frac{t_c}{T} \cdot t_{WZ} + t_{WST} \tag{8.10}$$

8.1.2.1 Optimal-Cost Cutting Speed

The manufacturing costs per workpiece K_F comprise labour and non-wage labour costs K_L, pure machine costs K_M, tool costs and residual factory overheads K_x.

$$K_F = K_L + K_M + K_W + K_r \tag{8.11}$$

Labour costs are calculated as follows:

$$K_L = K_{LH} \cdot t_e = L \cdot (1 + p_L) \cdot t_e \tag{8.12}$$

with K_{LH} as labour and non-wage labour costs per hour, L as hourly wage and p_L as the amount of non-wage costs.

Machine costs are derived from the machine-hour rate and the time per unit t_e:

$$K_M = K_{MH} \cdot t_e \tag{8.13}$$

The proportionate tool costs are defined as

$$K_W = \frac{K_{WT}}{n_{WT}} = \frac{t_c}{T} \cdot K_{WT} \tag{8.14}$$

with K_{WT} as tool costs for the tool operating life n_{WT} including pre-adjustment costs, number of workpieces per tool life T and t_c as cutting time. For the tool costs per tool life for re-grindable tools we have:

$$K_{WT} = \frac{K_{Wa} - K_{Wu} + n_s \cdot K_{Ws}}{n_s + 1} + K_{Wv1} + K_{Wv2} \tag{8.15}$$

with K_{Wa} as the tool acquisition value, K_{Wu} as the tool residual value, n_s the number of potential regrinds, K_{Ws} costs per regrind and K_{Wv1} and K_{Wv2} as pre-adjustment costs outside and inside the machine.

The following is true for tools with indexable inserts:

$$K_{WT} = \frac{K_{WP}}{n_{TP}} + \frac{K_{WH}}{n_{TH}} + \frac{K_{WE}}{n_{TE}} + K_{Wv1} + K_{Wv2} \tag{8.16}$$

with K_{WP} as the indexable insert acquisition value, K_{WH} as the tool holder acquisition value and K_{Wv1} and K_{Wv2} as pre-adjustment costs outside and inside the machine.

The tool operating life n_T with $i = P, H, E$ usually has highly differing values.

The residual factory overheads are calculated in the following way:

$$K_x = \frac{K_{xH}}{60} \cdot t_{eB} \tag{8.17}$$

with K_{xH} as the residual factory overheads per hour and t_{eB} as the occupation time on the machine.

The machine and labour cost hourly rate K_{ML} is defined as:

$$K_{ML} = \frac{t_e}{t_{eB}} \cdot K_{LH} + K_{MH} + K_x \tag{8.18}$$

with K_{LH} as the labour and non-wage labour costs, K_{MH} as the machine hour-rate and K_x as the residual factory overhead. By using the machine and labour cost hourly rate, manufacturing costs can be summarized in the following manner:

$$K_F = K_L + K_M + K_x + K_W = K_{ML} \cdot t_{eB} + K_W \tag{8.19}$$

Figure 8.3 shows graphs of the manufacturing costs, tool costs and costs contingent on the main processing time as a function of cutting speed.

Steadily increasing the cutting speed cannot lower the manufacturing costs further. Due to the shortening of the tool life with increasing speeds, more frequent tool change is necessary, raising the tool costs. Thus at very high cutting speeds the proportionate tool costs can become the largest addend of the manufacturing costs.

The cutting parameters depth of cut a_p, feed f and cutting speed v_c should be evaluated differently with respect to their optimization. The influence of the depth of cut on tool wear is minimal. In the first step in cost-optimization, one can first select the optimal-cost depth of cut a_{pok} at the maximum.

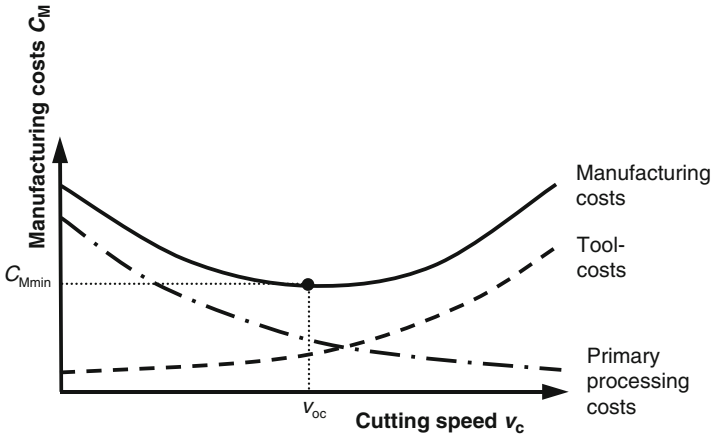

Fig. 8.3 Manufacturing costs as function of cutting speed

Although manufacturing costs have an absolute minimum as a function of feed and cutting speed, the feed is of only secondary importance as a freely selectable optimization parameter since the optimal feed is above the technically possible feed for most practical cases. In due consideration of the cutting parameter limits (Sect. 8.1.1), the second step is to select the maximum technically possible feed as the optimal-cost feed f_{ok}. As opposed to the parameters a_p and f, the cutting speed v_c is freely selectable across a wide range utilizable for optimization. Cutting speed has a considerable influence on the wear behaviour of the tool, which is designated by the tool life parameter, which in turn influences the tool change costs.

Rising machine and labour costs as well as shorter auxiliary process times cause a shift in the economical cutting speeds towards the lower range of tool life parameters. The description of tool life behaviour is limited in general to a range linearized in a double logarithmic system in which this behaviour can be described with sufficient accuracy (Sect. 7.2.1 and Fig. 8.4). Mathematically, a regression analysis is executed applying a general exponential function.

Figure 8.4 clarifies the dependence of the slope of the tool life straight lines on the cutting tool material when machining iron materials. Steeply falling tool life straight lines are characteristic of temperature-sensitive cutting tool materials (e.g. high speed steel), levelly falling tool life straight lines on the other hand for more high-temperature resistant cutting tool materials (e.g. cutting ceramics).

The expanded tool life function from a formulation made by TAYLOR contains all three parameters – depth of cut, feed and cutting speed:

$$T = C_v \cdot v_c^k \cdot C_f \cdot f_z^{k_{fz}} \cdot C_a \cdot a_p^{k_a} \tag{8.20}$$

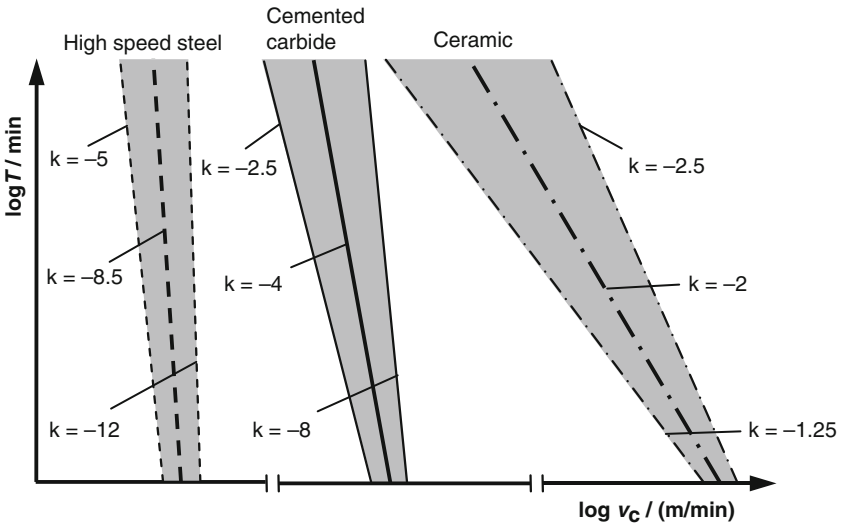

Fig. 8.4 Tool life straight lines for cylindrical turning, acc. to VDI 3321

Since both the optimal-cost depth of cut a_{pok} and the optimal-cost feed f_{ok} have already been selected, it is sufficient to use the TAYLOR's simplified tool life equation to derive the optimal value function of the costs.

$$T = C_{\mathrm{v}} \cdot v_{\mathrm{c}}^{k} \tag{8.21}$$

The optimal value function f_{ow} is set up by inserting the tool life Eq. (8.21) and the main process time Eq. (8.4), taking into consideration the process being used, into Eq. (8.19).

$$K_{\mathrm{F}} = K_{\mathrm{ML}} \cdot (t_{\mathrm{h}} + t_{\mathrm{n}} + t_{\mathrm{v}} + t_{\mathrm{er}}) + \frac{t_{\mathrm{c}}}{T} \cdot K_{\mathrm{WT}} \tag{8.22}$$

Using the known structure of the auxiliary process time we obtain:

$$K_{\mathrm{F}} = K_{\mathrm{ML}} \cdot \left(t_{\mathrm{h}} + \frac{t_{\mathrm{w}}}{n_{\mathrm{WT}}} + \frac{t_{\mathrm{r}}}{n_{\mathrm{WM}}} + t_{\mathrm{WST}} + t_{\mathrm{v}} + t_{\mathrm{er}} \right) + \frac{t_{\mathrm{c}}}{T} \cdot K_{\mathrm{WT}} \tag{8.23}$$

Here it is advisable to consider the dependencies of the individual amounts of cutting speed. All time parameters that are not a function of the cutting speed are viewed as constants and substituted in the following way:

$$C_1 = \frac{t_{\mathrm{r}}}{n_{\mathrm{WM}}} + t_{\mathrm{WST}} + t_{\mathrm{v}} + t_{\mathrm{er}} \tag{8.24}$$

Inserted, we obtain:

$$K_{\mathrm{Fo}}(v_{\mathrm{c}}) = K_{\mathrm{ML}} \cdot \left(t_{\mathrm{h}} + \frac{t_{\mathrm{c}}}{T} \cdot t_{\mathrm{w}} + C_1 \right) + \frac{t_{\mathrm{c}}}{T} \cdot K_{\mathrm{WT}} \tag{8.25}$$

Since the optimal value function should be a function of cutting speed, the time amounts have to be substituted by the cutting speed. In the case of a machining method using a *rotational main motion* with constant cutting parameters, the following equations are valid for the main process time and for the cutting time using Eq. (8.6) and substituting the constant items with the constant C_3:

$$t_{\mathrm{h}} = \sum_i \frac{C_i}{v_{\mathrm{f}i}} + \sum_j \frac{s_j}{v_{\mathrm{f}j}} = C_2 + t_{\mathrm{c}} \tag{8.26}$$

$$t_{\mathrm{n}} = \frac{t_{\mathrm{r}}}{n_{\mathrm{WM}}} + \frac{t_{\mathrm{w}}}{n_{\mathrm{WT}}} + t_{\mathrm{WST}} \tag{8.27}$$

$$t_{\mathrm{c}} = j \cdot \frac{\pi \cdot D \cdot s_j}{v_{\mathrm{c}} \cdot f_{\mathrm{ok}}} = C_3 \cdot \frac{1}{v_{\mathrm{c}}} \tag{8.28}$$

The tool life equation must also be valid, so for the optimal value function we obtain:

$$\begin{aligned} K_{\mathrm{Fo}}(v_{\mathrm{c}}) &= K_{\mathrm{ML}} \left(C_2 + C_3 \cdot \frac{1}{v_{\mathrm{c}}} + C_3 \cdot \frac{t_{\mathrm{w}}}{v_{\mathrm{c}} T} + C_1 \right) + C_3 \cdot \frac{1}{v_{\mathrm{c}} \cdot T} \cdot K_{\mathrm{WT}} \\ &= K_{\mathrm{ML}} \cdot \left(C_2 + C_3 \cdot \frac{1}{v_{\mathrm{c}}} + \frac{C_3}{C_{\mathrm{v}}} \cdot \frac{t_{\mathrm{w}}}{{v_{\mathrm{c}}}^{k+1}} + C_1 \right) + \frac{C_3}{C_{\mathrm{v}} \cdot {v_{\mathrm{c}}}^{k+1}} \cdot K_{\mathrm{WT}} \end{aligned} \tag{8.29}$$

Now the methods of a general extreme value task can be used so that the necessary and sufficient condition are met. For the first derivation according to cutting speed we first obtain:

$$\frac{\mathrm{d}K_{\mathrm{F0}}}{\mathrm{d}v_{\mathrm{c}}} = K_{\mathrm{ML}} \left(-C_3 \cdot \frac{1}{{v_{\mathrm{c}}}^2} - (k+1) \cdot \frac{C_3}{C_{\mathrm{v}}} \cdot \frac{t_{\mathrm{w}}}{{v_{\mathrm{c}}}^{(k+1)}} \right) - (k+1) \cdot \frac{C_3 \cdot K_{\mathrm{WT}}}{C_{\mathrm{v}} {v_{\mathrm{c}}}^{(k+2)}} \tag{8.30}$$

This expression can be simplified as follows:

$$\frac{\mathrm{d}K_{\mathrm{F0}}}{\mathrm{d}v_{\mathrm{c}}} = -\frac{K_{\mathrm{ML}} \cdot C_3}{{v_{\mathrm{c}}}^2} - (k+1) \cdot \frac{C_3 \cdot (K_{\mathrm{ML}} \cdot t_{\mathrm{W}} + K_{\mathrm{WT}})}{C_{\mathrm{v}} \cdot {v_{\mathrm{c}}}^{(k+2)}} \tag{8.31}$$

The necessary condition for extreme value tasks demands that the first derivation becomes zero so that the following is true for the optimal cutting speed:

$$\frac{\mathrm{d}K_{\mathrm{F0}}}{\mathrm{d}v_{\mathrm{c}}} = 0 \rightarrow v_{\mathrm{c}} = \sqrt[k]{-\frac{(k+1) \cdot C_3 \cdot (K_{\mathrm{ML}} \cdot t_{\mathrm{w}} + K_{\mathrm{WT}})}{K_{\mathrm{ML}} \cdot C_3 \cdot C_{\mathrm{v}}}} \tag{8.32}$$

To make the cutting speed minimal-cost, the second derivation must be greater than zero. At this location however, we shall dispense with the proof. After re-substitution, the following is valid for the optimal-cost cutting speed:

$$v_{\mathrm{ok}} = \sqrt[k]{\frac{-(k+1) \cdot \left(t_{\mathrm{w}} + \frac{K_{\mathrm{WT}}}{K_{\mathrm{ML}}} \right)}{C_{\mathrm{v}}}} \tag{8.33}$$

After inserting the optimal-cost cutting speed, Eq. (8.33) into the tool life equation, Eq. (8.21), we obtain the following for the optimal-cost tool life:

$$T_{\mathrm{ok}} = -(k+1) \cdot \left(t_{\mathrm{w}} + \frac{K_{\mathrm{WT}}}{K_{\mathrm{ML}}} \right) \tag{8.34}$$

8.1.2.2 Optimal-Time Cutting Speed

To determine the optimal-time cutting speed, the manufacturing time must be optimized with respect to the cutting speed. Figure 8.5 clarifies the difference between cost and time optimization.

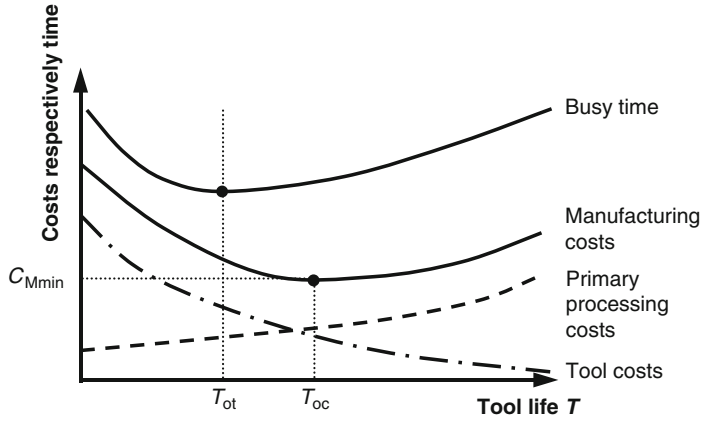

Fig. 8.5 Difference between time and cost optimum, acc. to VDI 3321

This means that a cutting speed should be selected which allows for a short main process time and a long tool life, and thus the ratio of cutting time to tool life only comprises a small amount of the auxiliary process time.

Analogously to the determination of the optimal-cost cutting speed (Eqs. (8.26) and (8.33)), this can be carried out for the optimal-time cutting speed, obtained via:

$$v_{ot} = \sqrt[k]{-(k+1) \cdot \frac{t_w}{C_v}} \tag{8.35}$$

After inserting the optimal-time cutting speed (Eq. (8.35)) into the tool life Eq. (8.21), we obtain for the optimal-cutting time tool life:

$$T_{ot} = -(k+1) \cdot t_w \tag{8.36}$$

After determining and optimizing the cutting parameters, the cutting parameters a_p, f and v_c must be evaluated for their realizability with respect to available spindle power and spindle moment. In so doing, the following demands must be met:

$$P_{spindel} > P_c = F_c \cdot v_c \tag{8.37}$$

$$M_{spindel} > M_c = F_c \cdot \frac{D}{2} \tag{8.38}$$

The cutting force can be determined by the calculation basis given by SALOMON [Salo24], which is used here in KIENZLE's notation [Kien52] in expanded form with K_{vk} as the correction factor for tool wear [Lang72, VDI3206, Degn00].

$$F_c = k_{c1.1} \cdot b \cdot h^{1-m_c} \cdot K_{vk} \tag{8.39}$$

One can assume 5% per 100 µm width of flank wear land as a standard value for the correction factor K_{vk}.

If the spindle power is insufficient, the depth of cut or feed should be reduced accordingly.

8.1.3 Calculating the Machine Hour-Rate

The machine hour-rate describes the costs to be calculated of a machine tool per hour. When determining the machine hour-rate, the

- imputed amortizations $K_{A,}$
- imputed interest $K_{Z,}$
- maintenance costs $K_{I,}$
- room costs $K_{R,}$
- and energy costs K_E

are taken into consideration. The machine hour-rate K_{MH} can be calculated as follows:

$$K_{MH} = \frac{K_A + K_Z + K_I + K_R + K_E}{T_N} \tag{8.40}$$

The yearly machine runtime T_N amounts for example to 1600–1800 h/a for single-shift operation. In the case of multi-shift operation, the runtime is increased proportionately (e.g. two-shift operation ca. 3200 h/a or three-shift operation ca. 4800 h/a).

The imputed amortizations K_A are comprised of the acquisition value of the machine K_{AW}, the amortization duration t and the sales revenue of the machine K_{VE} after expiration of the amortization duration:

$$K_A = \frac{K_{AW} - K_{VE}}{t} \tag{8.41}$$

The duration of amortization t is relevant for determining the imputed amortization K_A. The duration of amortization is calculated from the minimum of the expected technical or economical service life. The decreasing value of capital assets is designated in fiscal terms as *allowance for amortization* (AfD). The financial basis for the duration of amortization can be found in such AfD lists, since they turn out differently for different assets.

Figure 8.6 clarifies the calculation of the imputed interest for the amortization term. The sales revenue K_{VE} can be depreciated over the entire term. The acquisition value K_{AW} does not remain constant over the entire term however, but gradually decreases. The imputed interest over the amortization term K_{Zges} can be interpreted geometrically as surface area under the cost function. It should be taken into consideration that the symbol I always indicates the rate of interest per year. The following is valid:

$$K_{Zges} = \left(K_{VE} + \frac{K_{AW} - K_{VE}}{2}\right) \cdot i \cdot t = \frac{K_{AW} + K_{VE}}{2} \cdot i \cdot t \tag{8.42}$$

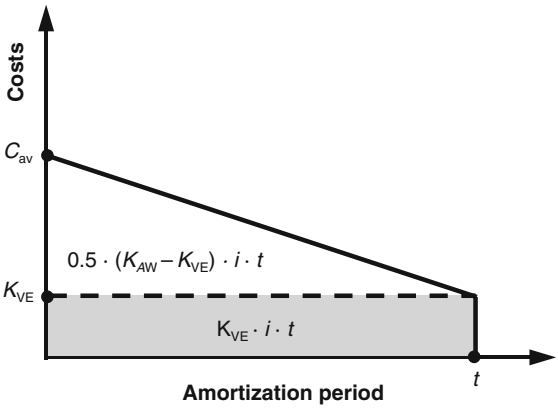

Fig. 8.6 Imputed interest calculation

Since the costs of Eq. (8.39) relate to 1 year, the imputed interest for 1 year K_Z must be inserted in this equation, resulting in:

$$K_Z = \frac{K_{AW} + K_{VE}}{2} \cdot i \tag{8.43}$$

Amortization, interest, room costs and fixed maintenance costs are all fixed quantities. On the other hand, energy costs, tool consumption and the variable maintenance costs are variable quantities. If the manufacturing labour is dependant on the machine run-time, it can be included in the rate of payment. In this case, we refer to it as a *cost per man-hour*.

8.1.4 Planning Methods and Tools

According to the type of data determination, we distinguish between manual (external) and computer-aided cutting data determination. The boundaries here, depending on the computer use, are fluid (Fig. 8.7).

In the case of manual cutting parameter determination and optimization, documents from various sources are utilized:

- guidelines of different standardization committees
- catalogues and cutting parameter recommendations of the tool manufacturers
- trial results from research papers
- in-house data banks for cutting parameters

Machining standards, such as are published by the Verein Deutscher Ingenieure (VDI) (English: Association of German Engineers) for various machining processes as VDI guidelines, give the process planner only very rough reference values for the target cutting parameters [VDI3206, VDI3208, VDI3209]. Usually, cutting speeds are recommended for broadly associated material groups independently of chip cross-section, tool life behaviour, machine type and machine-hour rate that often

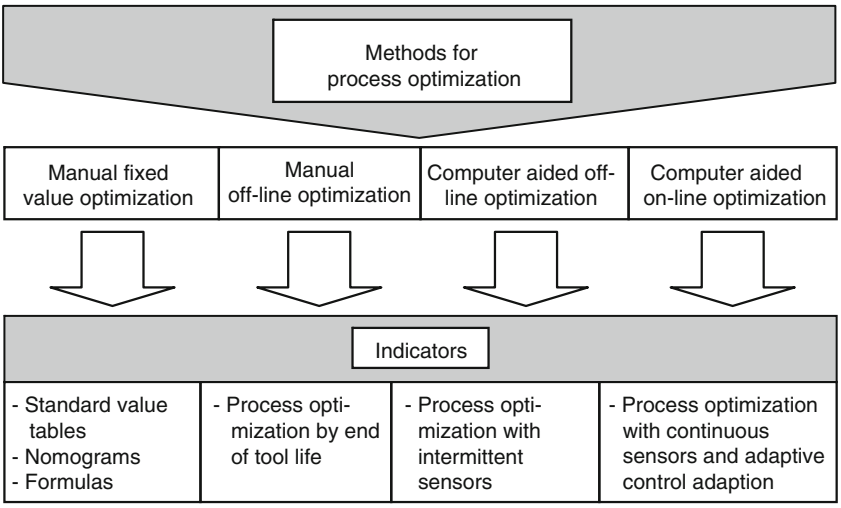

Fig. 8.7 Methods of process optimization, acc. to GEBAUER [Geba80]

lead to an uneconomically high tool life. Many tool manufacturers today offer software systems (e.g. CIMSOURCE, CoroGuide™, WinTool, etc.), with the help of which the determination of the cutting parameters can be combined directly with the selection of a suitable tool. More exact results can be obtained from test reports, which however are valid in most cases for special machining conditions and can only be applied to other machining cases with reservations.

Based on these sources, internal cutting data collections are often made as part of a company's process preparation. These collections make reference to the accumulated practical machining cases in manufacture and also take into consideration the long-term experiences of specialists and skilled workers. Information centres for cutting data, which collect cutting data from industry and research and make these available to companies in a form specifically tailored to them, are another modern phenomenon. Figure 8.8 shows the input mask of a data bank as an example of such a type of cutting parameter selection. These data bank systems contain cutting data for set tool life parameters for a comparable workpiece material/cutting tool material combination and under consideration of geometrical influencing variables. Such a data bank makes it possible to process large collections of previously determined machining data under consideration of many influencing parameters and to provide improved target values.

8.2 Process Monitoring

Secure process monitoring has a crucial role in manufacturing technology. Uninterrupted, fully or partially automated manufacturing processes guarantee high productivity and optimal capacity utilization. Process disturbances are, however, unavoidable. Stochastically occurring tool failure, deviations in the composition of

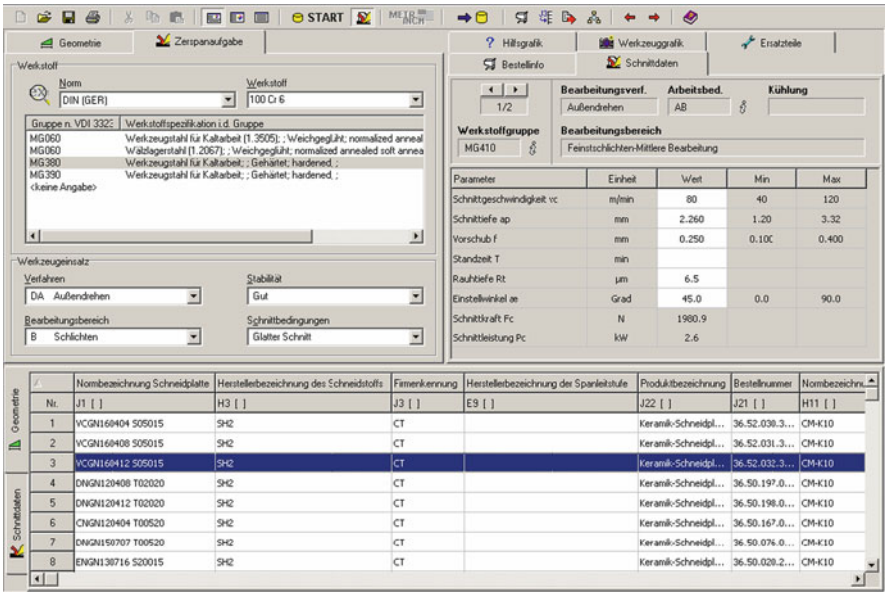

Fig. 8.8 Example for technology data organization (Source: CIM GmbH)

the machined material or faulty planning/programming prior to machining are all potential causes for such interruptions.

The goal of process monitoring in manufacturing technology is to recognize these process disturbances in a timely fashion and to introduce proper countermeasures. For example, extensive tool, workpiece or machine tool damage can result from tool fractures if the proper countermeasures are not promptly applied, be it manually or automatically (i.e. with the help of monitoring systems). This means, among other things, that the machining process must be monitored in order to minimize consequential costs of process disturbances.

Process monitoring systems make it possible to monitor cutting tools. The tools are thereby monitored for sudden tool fractures or overloads and for increasing tool wear. Tool monitoring has been developed to varying extents depending on the various machining processes with defined cutting edge geometries. Monitoring of the drilling process has been widely researched and finds extensive use in industry. There are also different approaches from research and practice for realizing a reliable tool fracture and wear-recognition monitoring system for turning processes. In the case of milling, fracture and wear monitoring is much more costly than in the abovementioned processes due to the more complicated process kinematics. Accordingly, there are only a few monitoring solutions that exist for milling processes.

Common to all monitoring systems is that process-relevant information is measured with the help of specialized sensors. The most common sensors and sensor principles today will be introduced in the next chapter.

8.2.1 Sensors for Process Monitoring

A number of sensors are available on the market for monitoring manufacturing processes, of which most can be classified in six physical functional principles. Table 8.2 shows an allocation of typical measurement parameters and the physical functional principles that primarily underlie them.

In order to standardize usage and to lend semantic clarity to the terms used, the essential fundamental concepts of metrology will first be defined [Pfei01]:

- The *measurement parameter* is the physical parameter that is to be measured.
- *Measurement* is the execution of activities planned for the quantitative comparison of measurement parameters with the unit.
- The *measurement result* is the value taken from the measurements estimated as the true value of a measurement parameter.
- The *measurement principle* is the physical basis of the measurement.
- The *measurement method* is defined as a special kind of procedure used in the measurement independent of the measurement principle.
- A *measurement process* is the practical application of a measurement principle and a measurement method.

A further distinction is drawn between direct and indirect measurement methods. Direct measurement methods record the measurement parameter of interest by means of a comparison with a normal of the same physical parameter. Indirect measurement methods, in order to determine the measurement parameter, capture an auxiliary parameter which stands in a known and describable relation to the measurement parameter [Pfei01].

While for direct methods, it is often difficult to reach the measurement location with suitable sensors, the problem with indirect measurement methods is in the determination of an assignment rule, the validity of which should not be affected

Table 8.2 Relation of measurement parameters to physical functional principles

Physical functional principles for process description	Associated measurement parameters
Mechanical (motion, displacement, stiffness)	Position, acceleration, velocity, force, torque, longitudinal/torsional stress, longitudinal/torsional strain, pressure etc.
Thermal (kinetic energy of atoms and molecules)	Temperature, heat flow, specific heat, thermal conductivity etc.
Electric (electric field)	Voltage, current, charge, conductivity etc.
Magnetic (magnetic field)	Permeability, magnetic current etc.
Radiation (electromagnetic radiation)	Energy, intensity, emission, reflection, permeability etc.
Chemical (forces between atomic nuclei and electrons, bond energies of molecules)	Chemical components, concentrations etc.

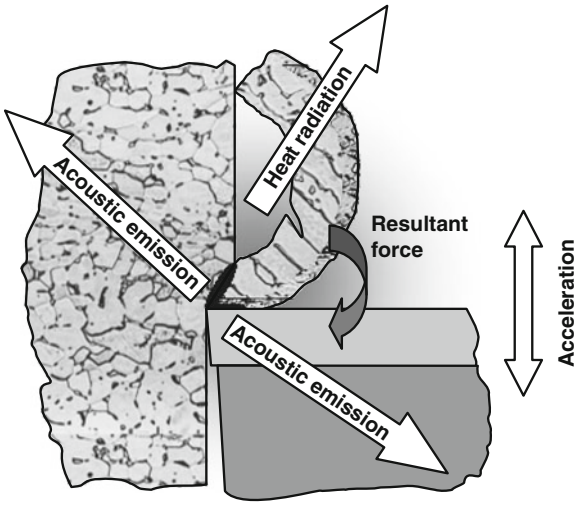

Fig. 8.9 Process emissions for monitoring of cutting process

by changes to the process parameters. Yet the advantages and disadvantages of the direct and indirect measurement methods cannot be generalized and must be weighed for every particular case.

The following is concentrated on the monitoring of cutting processes. Figure 8.9 shows the process emissions relevant for this area.

Numerous sensors use the strain of an object as an auxiliary function for measuring force, torque and acceleration. For this reason, the following sections will examine the workings of strain gauge strips and piezoelectric sensors.

8.2.1.1 Strain Gauges

A strain gauge converts a mechanical strain into an electrical resistance change. The OHM resistance R of a metallic conductor is a function of the material (ρ_e), length (L) and cross-section (A) according to the following relation:

$$R = \rho_e \cdot \frac{L}{A} \tag{8.44}$$

In the simplest case, a strain gauge consists of a resistance wire, which is frequently incorporated into a supporting matrix. The strain gauge is fastened to the shifting position on the measuring body. The adhesive is of great importance, because the strain or compression of the measuring body must be transferred unaltered to the strain gauge via the adhesive connection. Assuming that the material properties of the strain gauge are not changed by the strain (ρ_e = constant), the resistance change is dependent only on the length and cross-section change. In general, the relation between strain and relative resistance change can be described by Eq. (8.45). Strain sensitivity k, for example for a constantan strain gauge, amounts

to $k = 2.044$. The normal operating range of strain gauges extends up to a relative strain/compression of about 1%.

$$\Delta R/R = k \cdot \varepsilon \tag{8.45}$$

In order to reduce the dimensions in comparison to the single-wire form, usually meander-formed designs are utilized (Fig. 8.10).

However, the deflection areas cause a sensitivity of up to 3% transverse to the measurement direction. Strain gauges are not limited to the use of resistance wires. Semiconductor, foil and thin-film materials are also employed.

The resistance changes arising during the use of strain gauges are very small and are only a fraction of its total resistance. With the help of a bridge circuit, it is possible to measure the arising resistance changes in the form of a voltage variation U_M. For a balanced bridge circuit with $R_1 = R_2 = R_3 = R_4 = R$ and a feed voltage U_0, $U_M = 0$. The WHEATSTONE bridge circuit is shown in Fig. 8.11.

Also valid is the relation given in Eq. (8.46)

$$U_M = U_0 \cdot \frac{(R + \Delta R_1)(R + \Delta R_4) - (R + \Delta R_2)(R + \Delta R_3)}{(2R + \Delta R_1 + \Delta R_2)(2R + \Delta R_3 + \Delta R_4)} \tag{8.46}$$

Depending on whether two or four active strain gauges are used, this relation can be simplified in the form of a quarter, half or full bridge. For a quarter bridge with

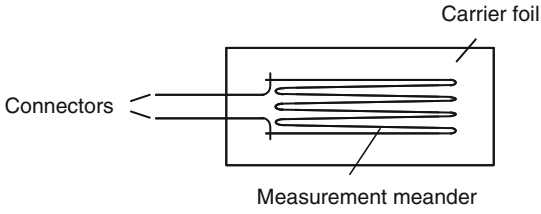

Fig. 8.10 Sketch of a strain gauge strip with measurement meander

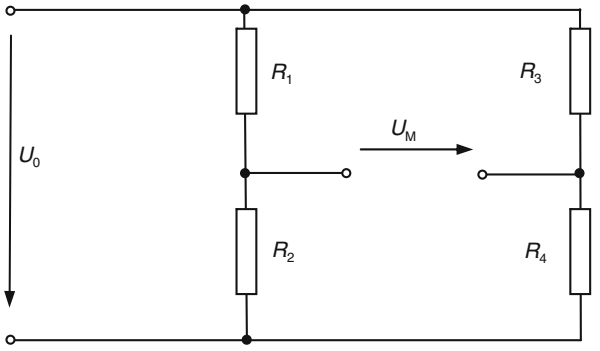

Fig. 8.11 Wiring diagram of a WHEATSTONE bridge

$\Delta R_1 = \Delta R$ *and* $\Delta R_2 = \Delta R_3 = \Delta R_4 = 0$ and $\Delta R << R$ the Eq. (8.47) is valid. It is used for simple measurements with a single strain gauge.

$$U_M = U_0 \cdot \frac{\Delta R}{4R} \tag{8.47}$$

A half bridge can be used to increase sensitivity by applying two identical strain gauges so that the amount of strain of one is equal to the amount of compression of the other, i.e. $\Delta R_1 = \Delta R, \Delta R_2 = -\Delta R$ and $\Delta R_3 = \Delta R_4 = 0$. Eq. (8.48) is true in this case.

$$U_M = U_0 \cdot \frac{\Delta R}{2R} \tag{8.48}$$

In the case of the full bridge, all four arms of the bridge circuit are fit with strain gauges. This makes it possible to raise sensitivity ever further and, in case all the strain gauges are exposed to the same temperature, a temperature compensation is also possible. Equation (8.49) is valid here with $\Delta R_1 = \Delta R_4 = -\Delta R$ and $\Delta R_2 = \Delta R_3 = -\Delta R$.

$$U_M = U_0 \cdot \frac{\Delta R}{R} \tag{8.49}$$

8.2.1.2 Piezoelectric Sensors

Piezoelectricity is caused by the electromagnetic interaction between the mechanical and electrical state of crystals that have no centre of symmetry, e.g. Quartz (SiO_2), tourmaline and ferroelectrical ceramics. By deforming the crystal lattice along the polar axes, the positive and negative lattice elements are displaced relative to one another, generating an electrical dipole moment. This charge transfer can be converted into a voltage signal with the help of a charge amplifier. In the case of piezo-actuators, the inverse piezoelectric effect is utilized, in which a mechanical change in length is produced by applying voltage.

Piezoelectric materials have one or more polar axes. Their piezoelectric coefficient is direction-dependent and indicates sensitivity in the direction of the corresponding axis. Piezoelectric coefficients range from 2 to a present maximum of 1500 pC/N. This should be taken into consideration when cutting crystals (Fig. 8.12).

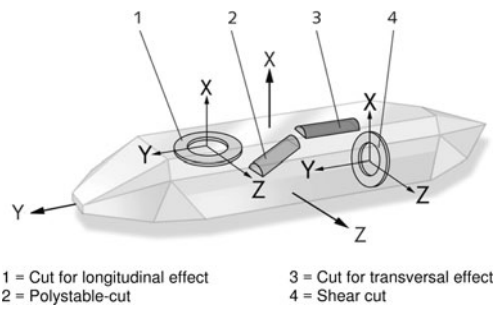

Fig. 8.12 Possible cuts through a synthetic crystal (Source: Kistler)

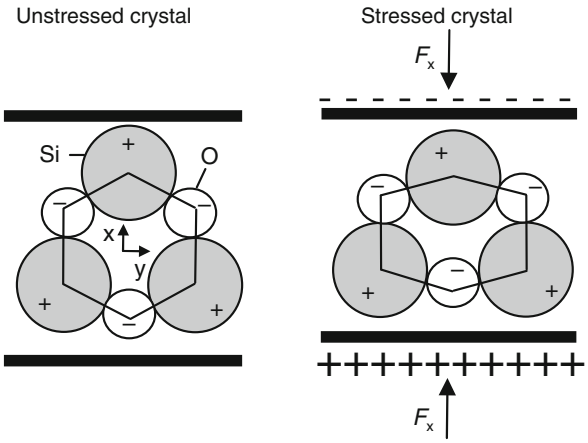

Fig. 8.13 Longitudinal effect in piezoelectric materials

In the case of the longitudinal effect, a charge transfer occurs on the force contact surface as a result of a mechanical deformation along a polar axis of the crystal perpendicular to the pickup of the charge (Fig. 8.13).

The charge difference should not be increased with the geometrical dimensions of the crystal plates, but only by mechanical series connection and electrical parallel connection of several plates. The amount of charge can be calculated for the longitudinal effect with Eq. (8.50), with d_{11} as the piezoelectric coefficient in the x-direction, F_x as force in the x-direction and n as the number of crystal plates.

$$Q_x = d_{11} \cdot F_x \cdot n \tag{8.50}$$

In the case of the shear effect, stress occurs as in the longitudinal effect along the polar axis of the crystal. However, in this case a shear force arises in the direction of the charge pickup (Fig. 8.14).

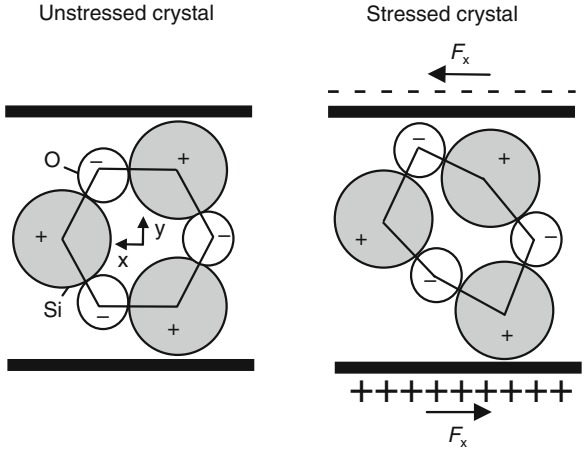

Fig. 8.14 Shear effect in piezoelectric materials

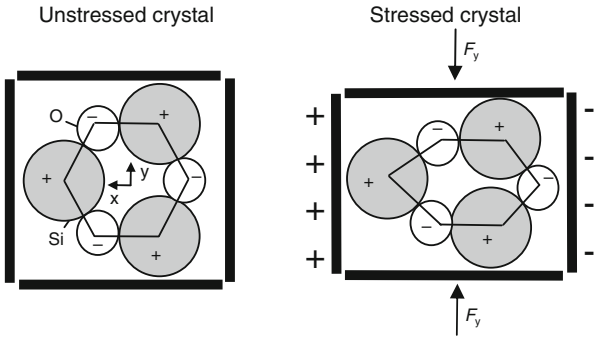

Fig. 8.15 Transversal effect in piezoelectric materials

The charge thereby released can be computed using Eq. (8.51), with d_{11} as the piezoelectric coefficient in the x-direction, F_x as force in the x-direction and n as the number of crystal plates.

$$Q_x = 2 \cdot d_{11} \cdot F_x \cdot n \tag{8.51}$$

In the case of the transversal effect, stress is applied in the direction of a neutral axis of the crystal, and the charge occurs orthogonally to it on a polar axis (Fig. 8.15).

In this case, the amount of charge should be influenced by the dimensions of the piezo-element and can be calculated with Eq. (8.52), with d_{11} as the piezoelectric coefficient in the x-direction, F_y as force in the y-direction and b/a as the geometric ratio.

$$Q_y = -d_{11} \cdot F_y \cdot \frac{b}{a} \tag{8.52}$$

Another essential component of piezoelectric sensors are the isolators. Isolators are subject to extremely high requirements in order to prevent small, piezoelectrically generated charge transfers from flowing off rapidly and distorting the reading, especially in the case of static and quasistatic measurements. The isolation effect must also be guaranteed during use under high temperatures and under the influence of moisture. Furthermore, the isolation materials must have a high level of stiffness for transferring mechanical stresses in order to support the piezo-element against the housing.

8.2.1.3 Force Sensors

For measuring forces and torques in manufacturing technology, piezoelectric sensors and, occasionally, inductive and capacitive position transducers are utilized. In general, all systems convert a stress into an electrically evaluable magnitude.

Fundamentally, one can measure either in the main force connection or in the secondary force connection. Especially when the sensor is arranged in the main force

connection, one must be sure that, on the one hand, the stiffness properties of the entire system are not negatively affected by the integration of the sensor and, on the other hand, that the sensor itself is not damaged by an overload. The high stiffness and high overload stability are advantageous in this regard. However, the sensors can sustain only extremely low tensile stresses, so that they must be sufficiently prestressed with a compressive stress in the case of tensile or tumescent stress.

Figure 8.16 shows the structure of a piezoelectric sensor used for measuring force components. Two longitudinally sensitive quartz plates serve as the sensor elements.

In order to measure torque, two basic principles can be used. When longitudinally sensitive force sensors are used, it is possible to make inferences about the torque from non-parallel sensors under consideration of their reference coordinates. If shear-sensitive force sensors are used, the torque applied can be directly measured (Fig. 8.17). The plates are arranged mechanically in a series and connected electrically parallel so that the charge is proportional to the torque applied. Since the shear stresses are non-positively transferred, the sensors require a high level of pre-stressing.

In the case of sensors designed to capture several components, a corresponding number of additional quartz plate pairs of varying orientation are used. Figure 8.18 shows a four-component dynamometer used for measuring three force components and torque.

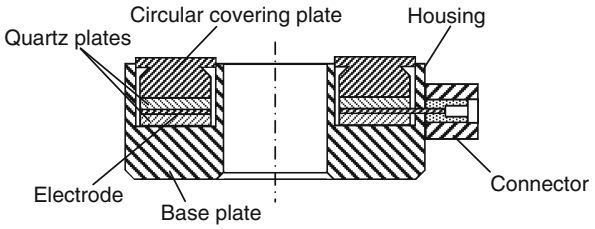

Fig. 8.16 Schematic view of a piezoelectric dynamometer

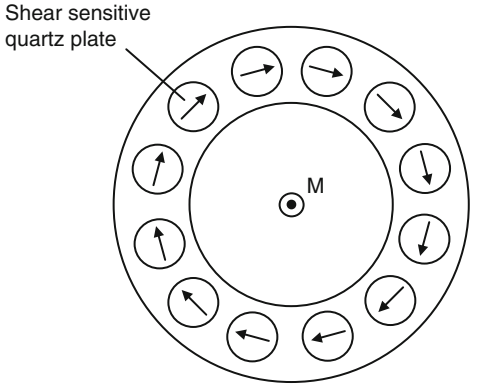

Fig. 8.17 Schematic view of assembly of shear sensitive quartz plates in a torque sensor

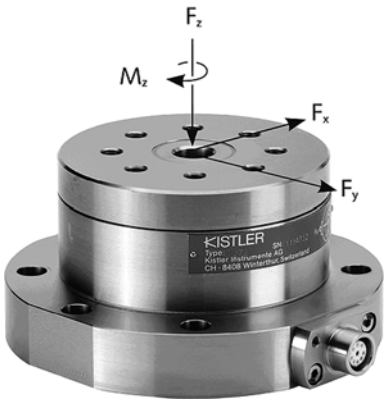

Fig. 8.18 Dynamometer for drilling processes (Source: Kistler)

Similar systems exist for turning and milling processes as well. Figure 8.19 shows a rotating four-component dynamometer for milling.

Piezoelectric measuring systems are characterized by their high levels of stiffness. If beyond this the oscillating mass located in front of the sensor element is small, the measurement section has a high characteristic frequency. In this way it is possible to analyze reliably dynamic process parameters even in high-frequency areas. In common usage, this is often designated as the capacity of piezoelectric measuring systems for highly dynamic measurement. Figure 8.20 shows the schematic of tool holder for turning, in which a small piezoelectric force sensor for measuring cutting force has been fit beneath the cutting edge.

This principle can also be applied to rotating tools. In this case, a signal transmission between the rotating and stationary parts of the measuring system is required. Such systems are currently still in prototype stage.

If however high dynamics are unnecessary in the measurement, measuring in the secondary force connection has as a rule the advantage of simpler sensor integration into the machine's structure. Besides strain gauges, above all piezoelectric force

Fig. 8.19 Rotating dynamometer for milling processes (Source: Kistler)

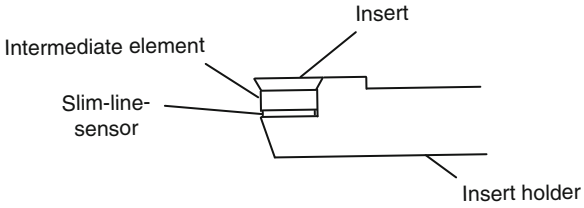

Fig. 8.20 Turning tool with integrated dynamometer

Fig. 8.21 Quartz transverse measuring pin
(Source: Kistler)

measurement pins are suitable as sensors in such cases (Fig. 8.21). These indirectly measure force via the strain of the structure into which they are integrated.

A general disadvantage of measuring in the secondary force connection is that it is necessary to analyze the machine structure in order to determine the optimal assembly location. If this is chosen correctly, it is then possible to monitor tool fracture and overload.

Force measurement rings – integrated into the main or feed spindle bearings of drilling, turning and milling machines – are another possible way to measure resultant force. Figure 8.22 shows a piezoelectric ring sensor. In some cases, it is also possible to apply strain gauges on the outer bearing ring or in a special bearing sleeve.

Fig. 8.22 Multicomponent force sensor for integration into spindle bearings (Source: Kistler)

8.2.1.4 Accelerometers

For measuring mechanical process dynamics in machining, usually piezoelectric accelerometers are used. The basic functional principle is based on the proportionate relation between force and acceleration. The factor of proportionality is the inert mass of the body upon which the force is acting. The main components of an accelerometer are therefore the seismic mass, sensor element and housing (Fig. 8.23).

In detail, we differentiate between three different accelerometer designs, in which a charge transfer is produced by normal forces, shear forces or flexural forces. A sensor element acting on normal forces is shown in Fig. 8.24. As a result of the design, the large surface contact between the sensor element and the base plate leads to the stresses of the measured object being transferred via the base plate to the sensor element. This brings about a measurement error. Furthermore, the pre-stressing is altered by temperature changes, so that thermally caused errors in measurement can also occur. The susceptibility of the sensor type to these sources of disturbance is much lower in the case of shear-sensitive accelerometers (Fig. 8.24). Besides their lower basic strain sensitivity, shear-sensitive piezoelectric ceramics exhibit no thermally caused charge transfer (pyroelectricity).

Every mass additionally mounted on the measured object changes its vibrational properties (characteristic frequency) and thus in certain conditions call into question the result of a measurement. It is especially important to take this fact into consideration in the case of experimental modal analysis. The mass of the accelerometer must be much smaller than that of the structural mass to be analyzed. One example

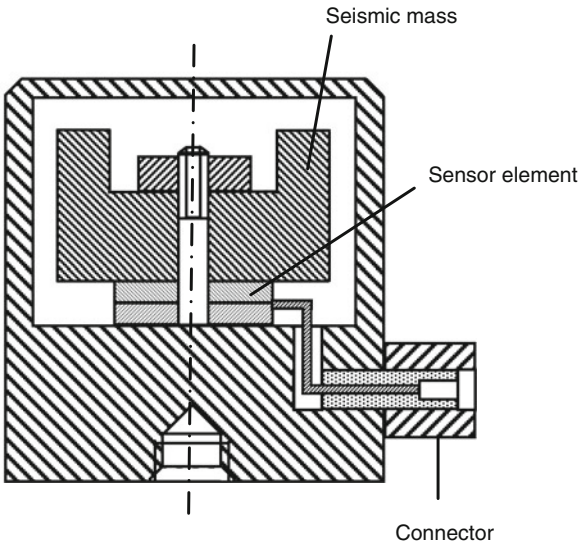

Fig. 8.23 Schematic assembly of an accelerometer

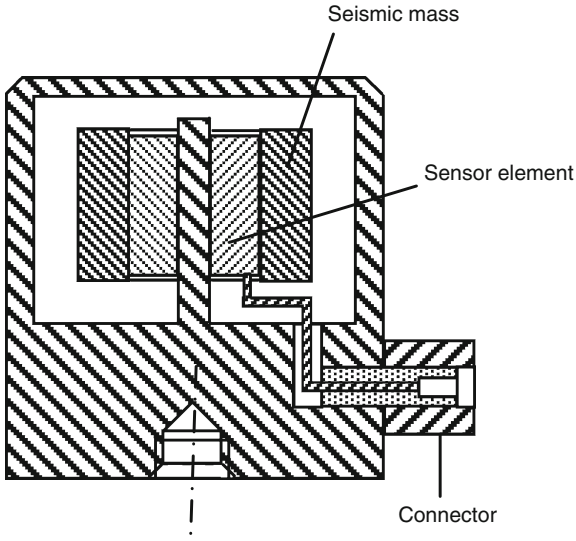

Fig. 8.24 Accelerometer with shear sensitive sensor element

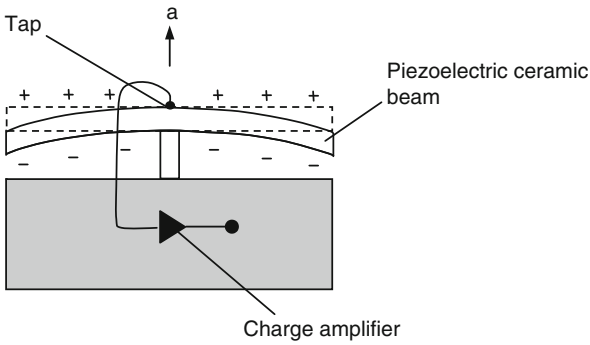

Fig. 8.25 Accelerometer with a piezoelectric beam

of a mass-reduced accelerometer is shown in Fig. 8.25. The active component of the accelerometer consists here of a piezo-ceramic bending beam. Upon displacement of the beam due to acceleration, a charge transfer is caused. In this case, the sensor element and the seismic mass are identical. Besides the low characteristic frequency and the pyroelectric attributes of the sensor, improper use results in marked sensitivity to mechanical destruction due to beam fracture.

The sensor is selected primarily in accordance with the measurement range, sensitivity, the mass and characteristic frequency. Simultaneous fulfilment of all goals represents an optimization problem due to physical limitations. For example, increasing sensitivity over a larger seismic mass results in a lower characteristic frequency and thus to a more limited available frequency range. This can been seen in Eq. (8.53), in which f_0 is the characteristic frequency, c the stiffness of the sensor element and m the seismic mass.

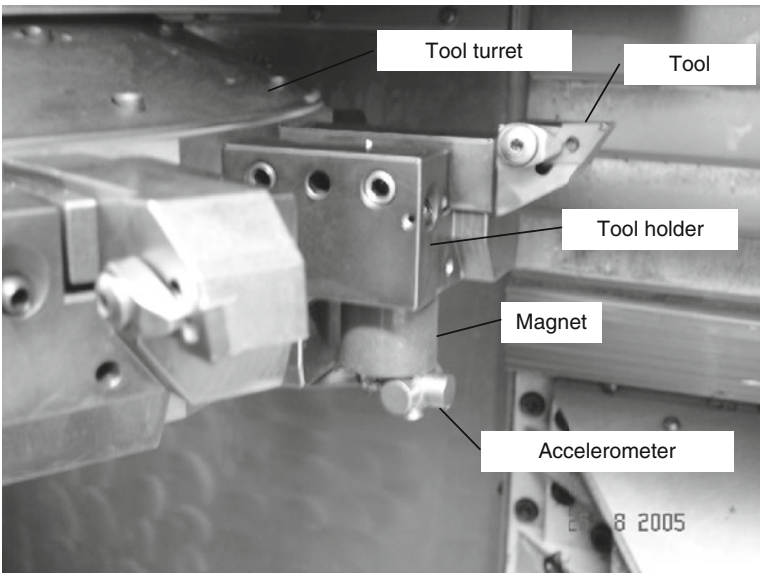

Fig. 8.26 Application of an accelerometer at the tool holder

$$f_0 = \frac{1}{2\pi} \cdot \sqrt{\frac{c}{m}} \tag{8.53}$$

In comparison to quartz, some piezoelectric ceramics provide charge differences that are 100 times higher under otherwise identical conditions. This property can be utilized in order to reduce the seismic mass by a factor of 100 at equal sensor sensitivity. However, ceramic elements are generally not as stiff, so this approach leads to losses with respect to the characteristic frequency.

In order to capture the dynamics of manufacturing process as freely as possible of disturbances in practice, a triaxially measuring accelerometer should be applied close to the action point, e.g. directly under the turning tool. This ideal solution is not always realizable within the machining space of a machine tool. In Fig. 8.26, the sensor was thus fastened with a magnet on the tool holder.

8.2.1.5 Acoustic Emission Sensors

There are also dynamics in the province of acoustic emissions [Eise88]. Depending on the type and magnitude of the emission source, the frequency range extends from audible sound (ca. 16 kHz) to the high ultrasound range (ca. 30 MHz). Figure 8.27 shows a diagram of possible sources and causes of the development of acoustic emission (AE) in machining processes.

Figure 8.28 shows an AE-sensor in cross-section. The surface waves arrive via a thin membrane with integrated docking element at the sensor element, which is surrounded by a damping mass. No additional seismic mass is required to detect

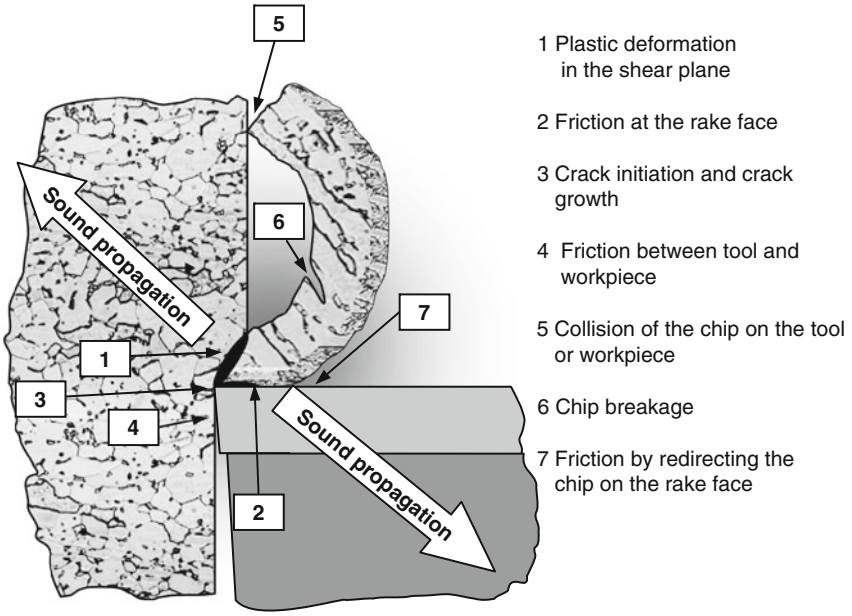

Fig. 8.27 Sources of acoustic emission

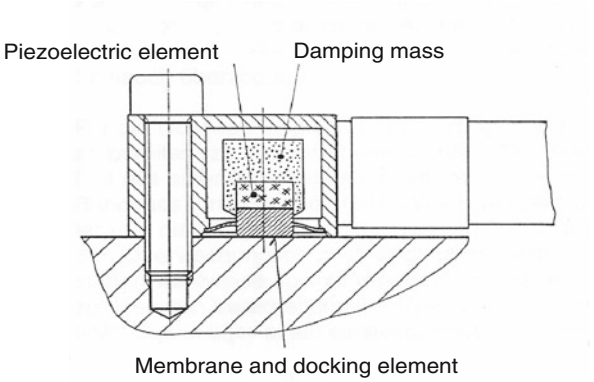

Fig. 8.28 Schematic assembly of an AE-sensor

the AE signal. The inertia of the sensor element itself is sufficient to produce a measurable charge transfer at typical frequencies between 50 kHz and 2 MHz.

Piezoelectric ceramics are especially suited to AE-measurement because of their high sensitivity. AE-sensors are constructed such that the mechanical oscillation direction is either perpendicular or parallel to the electric polarisation. In practice however, still other piezoelectric or elastic coupling possibilities exist that can lead to undesired resonances during the vibration measurement [Saxl97]. Thus, the area

of use of sensors is limited to suitable frequency ranges. Typical frequency ranges are, for example, 50–400 kHz or 100–900 kHz.

Surface sensors are attached to the machine structure and register the surface waves produced by the AE-signals. It is important that the optimal position for sensor assembly is determined. Sensor selection and positioning are decisive criteria for a successful process monitoring. AE-signals are attenuated at joining points and material inhomogeneities in an order of about 11 dB per point of intersection [Dorn93, Kett96]. The extent of this signal attenuation is also strongly dependent on frequency. High-frequency signal components are more strongly attenuated than low-frequency ones. When applying AE-sensors therefore, one is faced with the basic conflict of goals between attachment near the process in order to receive signals that are as little attenuated as possible and a positioning which protects the sensor from interference from hot chips or cutting fluid and does not obstruct tool or workpiece change. In order to guarantee constant and reproducible coupling conditions, the surface making contact with the sensor element must be machined to a high level of surface quality.

Fluid acoustic sensors measure the AE-signal via a fluid jet (e.g. cutting fluid), which is directed either at the workpiece, the tool holder or the tool.

With AE-sensors, fracture and wear can be recognized in turning as well as in milling and drilling operations [Diei87, Dorn89, Kett96, Köni89b, Köni92b, Mori80, Reub00]. Another area of application is in the recognition of unfavourable chip forms in automated turning processes [Kutz91, Köni96, Kloc05a]. One example from the sphere of grinding operations (see Manufacturing Processes, Volume 2) is the recognition of burn by means of acoustic emission analysis [Saxl97]. Monitoring solutions based on acoustic emission are suitable for finishing processes as well, since the use of other sensor principles has remained problematic due to the small cross-sections of undeformed chip and low resultant forces.

8.2.1.6 Effective-Power Sensors

Changes to the resultant force components lead not only to dislocations or deformations but also to a change in current and power consumption of the main and feed drives. The power consumed by the motor of a machine tool is composed of an effective and an idle component. Because of its proportionality to the torque emitted by the motor, the effective power is often used as a signal within the control system for quantifying the motor load. When monitoring cutting processes, usually external effective-power measurement tools are used, often with associated evaluation software and visualization unit.

The principle of effective-power measurement is based on determining the voltage U, current I and phase shift ϕ between both quantities. The effective power is obtained from these using Eq. (8.54). We make a distinction between one, two and three-phase systems, whereby three-phase systems have the highest resolution because they execute the measurement in all three phases.

$$P_{\mathrm{w}} = \sqrt{3} \cdot U \cdot I \cdot \cos\phi \tag{8.54}$$

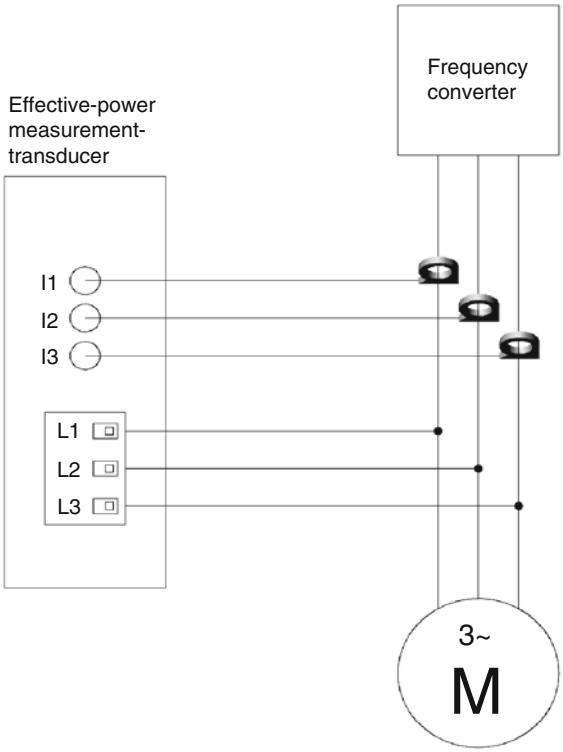

Fig. 8.29 Wiring diagram for the assembly of effective-power measurement system

External effective-power modules are installed in machine tools between the frequency converter and the motor. Figure 8.29 shows a typical wiring diagram for a three-phase system. To read the current, hall sensors are used that measure the current along with its phasing via the magnetic field surrounding the conductor in a circular shape. During integration into the machine tool, signal adaptation is very simple. Hall sensors are offered in various performance classes. The initial adaptation can be achieved by varying the number of conductor loops that are led through the hall sensors. Current systems also have electronic signal conditioners.

Effect-power measurement systems are characterized above all by the fact that they do not affect the mechanical properties of the machine tool. The machining torque can be measured during the operation without the integration of external sensors into the electric flux of the machine. The primary area of application is the recognition of tool fractures and collisions in the workspace. Sufficiently large force changes are necessary for efficient wear monitoring. Furthermore, the systems are inexpensive and easily retrofittable. Their suitability as an efficient process monitoring tool depends to a decisive extent on the ratio of the power input generated by the machining process and the total power of the drive. Small machining torques can thus no longer be reliably measured in the case of spindles with a collectively large power input. Reliable monitoring is particularly difficult in the case of small

cross-sections of undeformed chip or small process forces and torques, e.g. in the case of finishing or of drilling with small diameters. The method is hardly suitable for measuring dynamic machining torques due to the high inertia (of the motor armature, clamping mechanism and so on) in front of the sensors. The inert masses of the power train represent a low-pass filter that considerably impedes the swift measurement of dynamic magnitudes because of the low cutoff frequency.

8.2.1.7 Temperature Sensors

The cutting temperature is an important process-characteristic quantity for evaluating thermal stress on the workpiece and cutting tool material surfaces in the contact zone. In addition to thermoelements, resistance thermometers and thermocameras, quotient pyrometers are also used for measurement purposes, which provide highly dynamic, precise and absolute information on temperature.

Thermoelements make use of the thermoelectrical effect between two metals. Besides the single-cutter system and the twin-cutter system [Gott25, Vier70], it is also possible to integrate a thermoelement into the tool [Küst54] or the workpiece [Hopp03].

Resistance thermometers utilize the temperature-dependence of the resistance of a conductor/semiconductor for temperature measurement. Usually metallic materials are used, particularly platinum and nickel, the resistance of which increase with temperature in an easily reproducible manner. Some semiconductors have negative temperature coefficients, i.e. their resistance decreases with increasing temperature.

Thermocameras allow for a non-contact, extensive measurement of temperature. They function according to the principle of thermography. Every object emits a band of infrared radiation, the intensity of which is a function of temperature. The wavelength range is between 0.7 and 1000 μm. The majority of commercial infrared cameras use however only the spectral range of medium and long-wave infrared from 3.5 to 14 μm.

In the case of a quotient pyrometer – also referred to as a ratio pyrometer or 2-colour pyrometer – intensity is not only measured with one wavelength but by the ratio of the intensities of two different wavelengths. Infrared radiation is captured fibre-optically and sent to an evaluation unit. The glass fibre is sheathed with a protective tubing. In the pyrometer housing are the optics, filter, detectors and a measuring amplifier (Fig. 8.30). The principle of the quotient pyrometer is extensively discussed in the dissertation of Müller [Müll04].

The electronic housing contains the control of the Peltier element cooling and the power supply. Data acquisition and evaluation can be automatically executed with an attached laptop.

Variations in intensity that are not caused by temperature, but rather, for example, by partial impurification of the optics, have no influence on the measurement result. Furthermore, measurement by quotient formation is for almost completely independent of the degree of emission of the material so long as there is no significant dependence on wavelength.

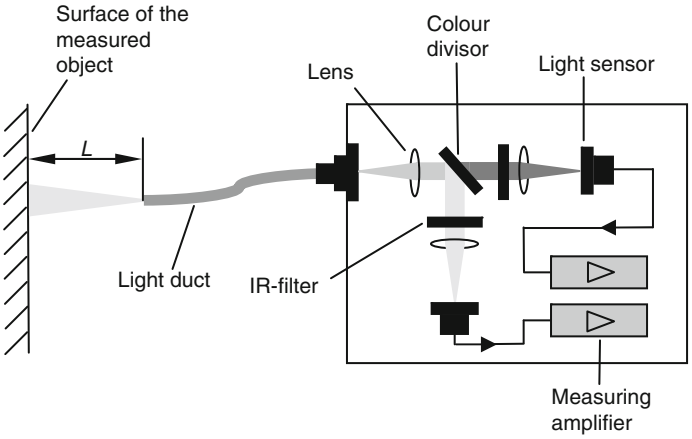

Fig. 8.30 Functional principle of a quotient pyrometer (Source: WSA RWTH Aachen)

Metallically sheer surfaces have small degrees of emission. In order to measure at optically difficult to access and thermally/mechanically heavily stressed areas, it is possible to position a fibre-optic cable very close to the contact zone between the tool and the workpiece. To do this, the measurement location must as a rule be prepared with a hole, into which the fibre can be inserted. Figure 8.31 shows a typical application on an indexable insert for measuring temperature on the machined surface of the workpiece.

Measurement data acquisition and evaluation can be carried out with a PC. The result of the measurement is a temporally and spatially high-resolution temperature signal such as is shown in Fig. 8.32 using an example from drilling.

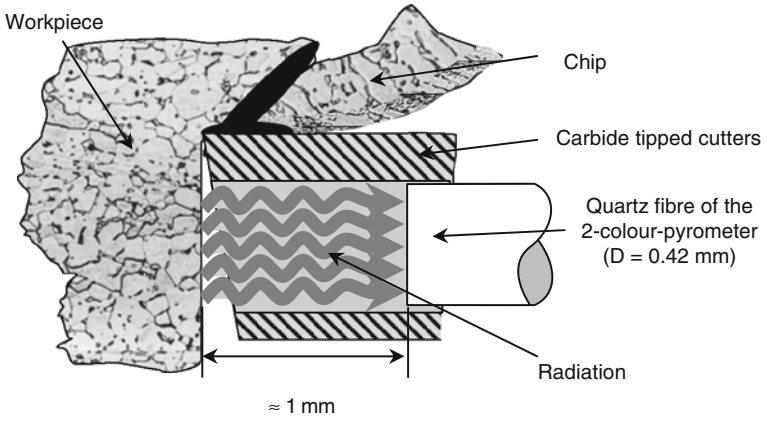

Fig. 8.31 Temperature measurement of machined workpiece surface

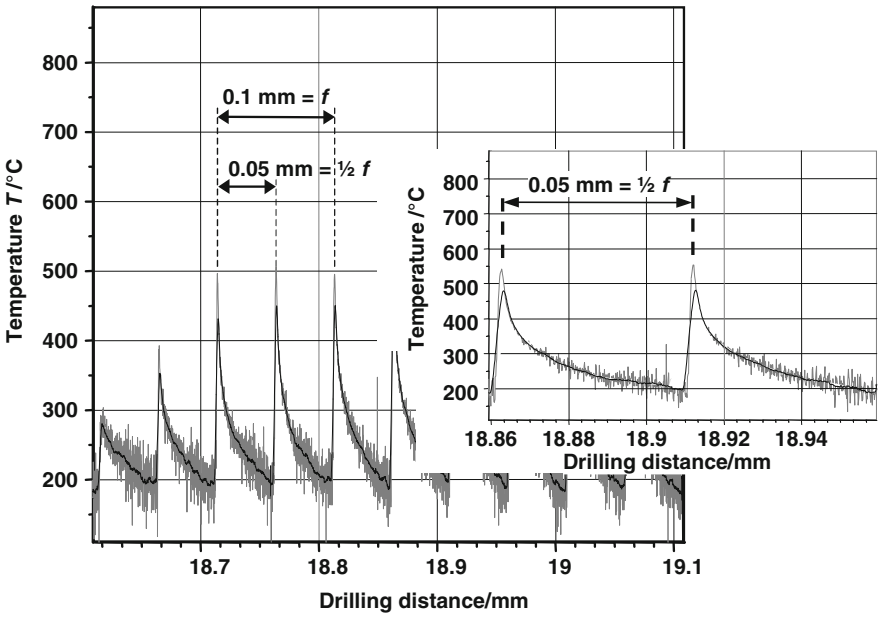

Fig. 8.32 Example of a highly detailed temperature signal

In the case of the system described here, the lower limit of measurable temperature is about 200°C. The potential temperature resolution depends on the selection of the amplification factor. A higher amplification factor results in a higher resolution but also reduces the measurable temperature range and the measurable maximum temperature. While losses of intensity by partial impurification of the optics do not distort the measurement result, measurement errors do occur in the case of differing emission degrees for both of the wavelengths on which the measurement is based. When cutting fluids are used, measurements are extremely unstable due to the different optical properties of the cutting fluid and the fibre material if the optic is wetted.

8.2.2 Signal Processing and Monitoring Strategies

In order to develop process monitoring strategies, signal processing is first required. Independently of whether the signals are obtained from external sensors or by machine tool control readouts, these signals should exhibit a close correlation to the process.

Depending on the required system reaction speed, the methods used either accompany the process or are intermittent. A continuous, process-concomitant measurement offers the most rapid possibility of detecting events and introducing a suitable reaction. The intermittent measurement permits reaction after reinstatement

of the measurement cycle at the earliest, so it is possible that disturbances arising outside the measurement cycle are not reflected in the measured signals.

No new information is generated during signal processing. Signal processing has the exclusive function of extracting from the total received information that information which relates to the process parameters to be monitored. A basic distinction is drawn between online and offline signal processing systems. In the case of online processing, signal reception and signal processing occur simultaneously. In offline processing, signal reception and processing are physically and temporally separated.

The properties and characteristics of the systems used determine to quite a significant extent the basic properties of the signals that are picked up. Therefore, it is necessary to determine signal processing strategies with an eye to the signal properties. Figure 8.33 shows an example of a measurement path for the acquisition of AE-signals in a machining process. This measurement path is characteristic of many other manufacturing processes. The signals generated by the sensors generally have low energy contents. The signal-to-noise ratio (SNR) is thus low, making disturbance-free signal transmission and evaluation more difficult. In order to improve the SNR therefore, the signal is subject to preamplification. Often the signal data is then filtered in accordance with the frequency range that is relevant for the given monitoring task. Depending on the monitoring task and the type of sensor signals, either highpass, lowpass, or bandpass filtering takes place.

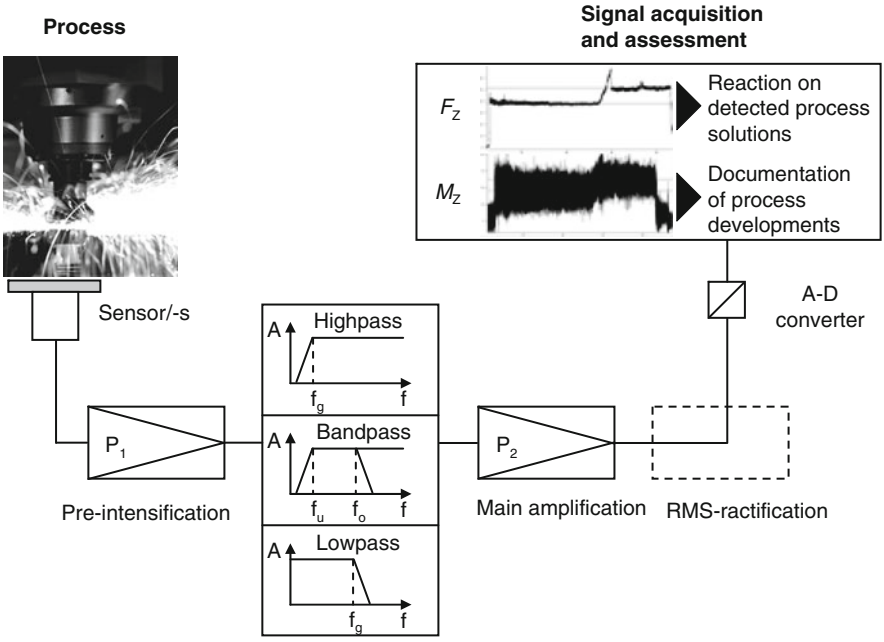

Fig. 8.33 Measurement path for acquisition of acoustic emission

The essential function of signal filtering is to filter out those signal components from the raw signal that do not correlate with the process parameters to be monitored. The irrelevant signal components might contain information assigned to the process but which is not relevant for the application at hand, or they might be extra disturbance signals that were introduced into the measurement chain from outside. Examples of disturbances include irrelevant machine vibrations or external electromagnetic disturbance fields. After signal filtering, amplification takes place in the main amplifier. Optionally, especially in the case of AE-signals, the effective value of the sensor signals can be rectified following Eq. (8.55).

$$s_{\text{eff}} = \sqrt{\frac{1}{T}\int_0^T s^2(t)\text{dt}} \tag{8.55}$$

For the sake of simplicity, preprocessing is applied to a signal with constant frequency (Fig. 8.34). In the effective value rectification, a time-dependent, sliding, quadratic average – the RMS (root-mean-square) – is formed from an output signal [Saxl97, Reub00].

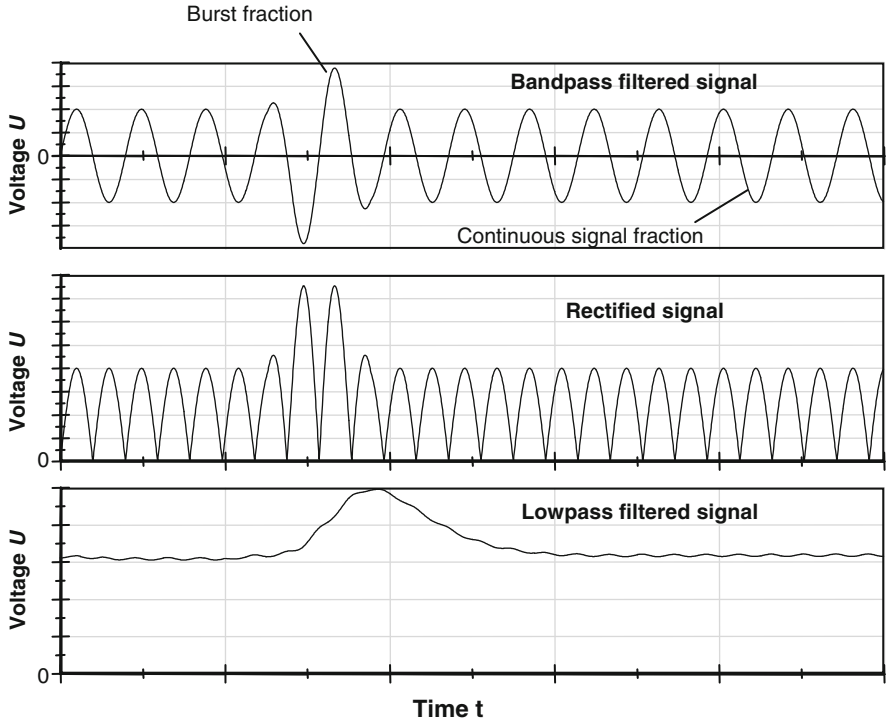

Fig. 8.34 Filtering and RMS-rectification

RMS signals have only positive signal components and are used to determine monitoring parameters. After the effective value is rectified, further filtering or amplifying of the signals may be necessary depending on the case at hand. In order to make information-technological processing of the signals possible, the analogue measurement values are digitalized in an analogue/digital (A/D) converter. The measurement values can then be acquired, evaluated with the help of software and further processed in digital form.

Principally, it is possible to consider signals in the time range or frequency range. Wave analysis makes it possible to combine information coming from both ranges. Table 8.3 compares general attributes of the three methods [Reub00].

In the case of time-consuming monitoring processes, signal evaluation in the time range has the advantage that it does not require a time-consuming transformation into the frequency range that demands a large amount of computer capacity. In the simplest case, process disturbances lead to significant signal changes that can be determined directly from the temporal signal profile.

Action limits and tolerance zones are useful for identifying disturbances. Figure 8.35 shows some classic examples. Besides a static limit value, a trend, a tolerance zone and revolving thresholds, it is also possible to identify characteristic signal profiles.

Monitoring by means of static thresholds is the simplest type of signal-based monitoring. The signal level for the undisturbed process profile is determined in preliminary tests. This signal level is then used for monitoring in the form of a stable limit. In the case of monitoring with dynamic thresholds, a sliding average is calculated from the recorded signal over a defined time period. This average is

Table 8.3 Characteristics of signal analysis in the time and frequency ranges

Time range	Frequency range
Temporal signal profile	Transformation of the signal into the frequency range
Continuous signal analysis or in temporally limited sections	Spectral signal composition
Almost no information regarding signal frequencies	Discontinuous signal evaluation in temporally limited signal sections
No time-consuming transformation	No information regarding temporal signal changes in the transformed signal section
	Limited information regarding the temporal signal profile by analysis of several consecutive signal sections

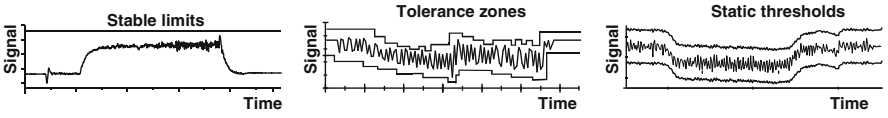

Fig. 8.35 Action limits and tolerance zones

associated with a percentage addition or deduction, making it possible to monitor the signal profile dynamically. The advantage of this strategy compared with a stable limit is that it is also possible to monitor signals that do not proceed uniformly or whose level constantly changes even in an undisturbed process.

Saxler provides an example of signal analysis in the time range in his dissertation on the structure of a system for recognizing grinding burn via acoustic emission analysis [Saxl97]. Grinding burn is a form of undesirable thermal rim zone damage and is described more thoroughly in volume 2 of this series. When profile grinding tooth flanks, a structure-borne sound sensor is affixed near the workpiece. The profiles of the RMS values of the signals are shown in Fig. 8.36 for the roughing and finishing phases of the machining of a total of 1650 gearwheels. Of particular interest is that both graphs have asynchronous profiles until the 900th gearwheel. Beyond that point, the graphs are almost parallel and differ by a signal difference of about 1 V. After reaching the end of tool life (i.e. after grinding 1550 gearwheels) are found the highest RMS values of the acoustic emission within the tool life with the exception of the first two evaluation points.

Irrespective of the fact that the appearance of grinding burn at the first two evaluation points can not be detected by exceeding the threshold value, the total amplitude differences are not large enough to lead to reliable information about the development of grinding burn. This practical example will be taken up again when dealing with signal analysis in the frequency range.

Besides the pure detection of deviations, breach of the action limits can be used to produce an automatic reaction of the machine by altering the process parameters (Fig. 8.37).

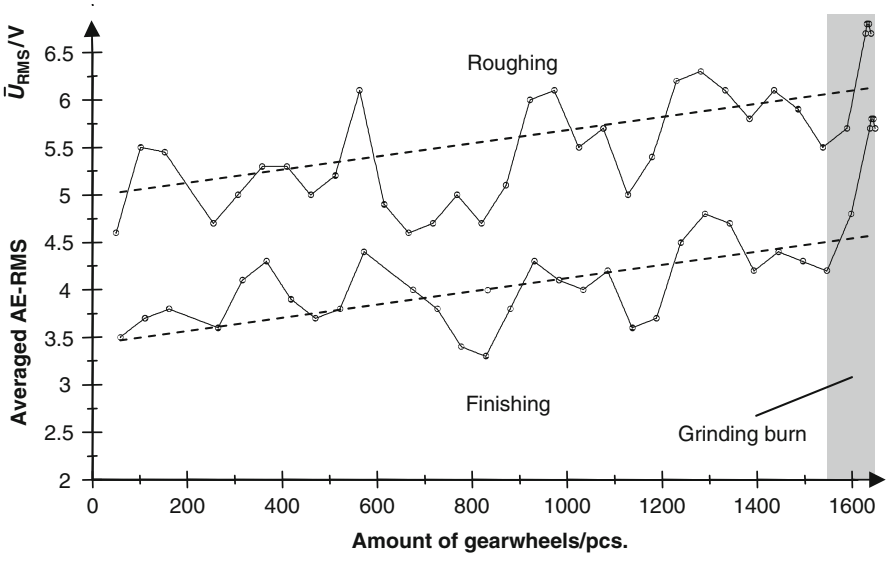

Fig. 8.36 Analysis of acoustic emission in time interval

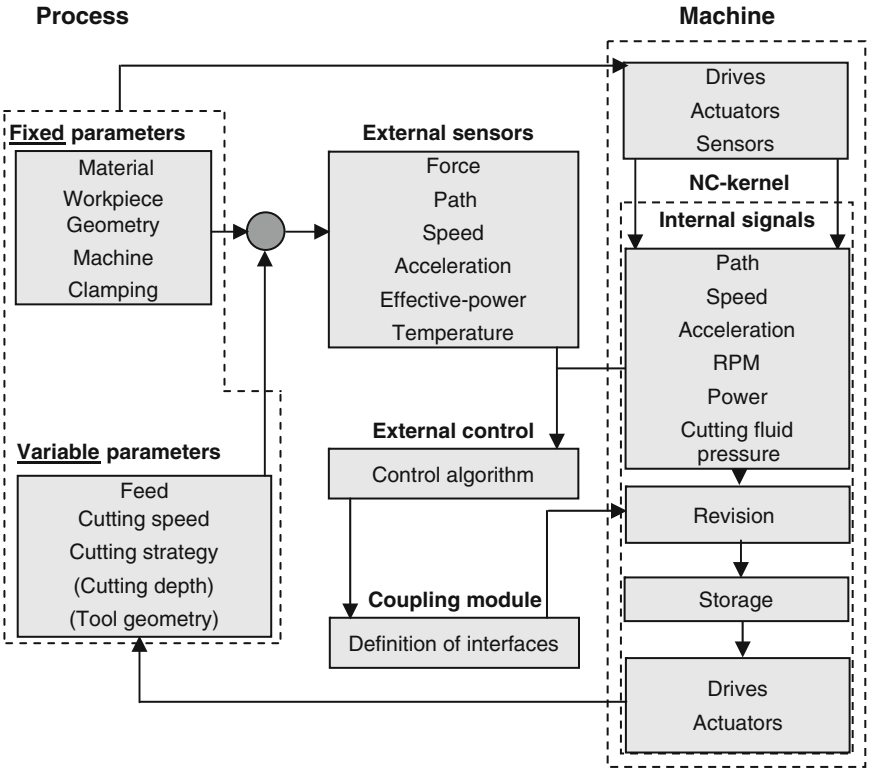

Fig. 8.37 Concept for intervention in the manufacturing process

For this purpose, there are two basic control strategies, adaptive control constraint [Gies73, Häns74, Müll76, Gath77] and adaptive control optimization [Esse72, Otto76]. In the case of adaptive control constraint (ACC), the variable of the control loop is varied such that the control variable reaches a level that is as constant as possible. One classic example of an adaptive control constraint in machining is the maintenance of a constant level of torque by adjusting the feed in the case of an alternating overmeasure. Adaptive control optimization involves varying one or more variables so that the control variable follows a learning curve. The learning curve could represent, for example, the profile of spindle performance over the feed path and be stored for a specific machining situation as an optimal process sequence. The parameters feed and speed could then be freely selected within defined limits so that the performance follows the learning curve despite wear.

New developments in process monitoring are focused on model-based monitoring methods or simultaneous evaluation of different sensor signals in order to obtain a maximum amount of security from interruption. This is necessary because frequent false alarms are a major problem in the industrial use of process monitoring systems in the case of the devices used today. For adaptive control constraint and

adaptive control optimization, often control algorithms are therefore implemented that combine and evaluate one or more pieces of sensor information. The output of the system can be characterized by several variables.

Signal evaluation in the frequency range is especially sensible when different periodic signal components are superpositioned. This is often the case in machining operations since dominant process frequencies such as turning or tooth engagement frequencies as well as their overtones are superpositioned by the characteristic frequencies of the machine and tool. From the representation of amplitude performance spectra over frequency, information can be obtained regarding the appearance of signal components at different frequencies and across frequency shifts. Often, frequency analysis also makes it possible to identify process disturbances, for example the appearance of chatter vibrations or significant changes on selected cutting edges in milling operations. Process disturbances of this kind are easier to detect in the frequency range than in the analysis of process signals in the time range.

The most well-known method of frequency analysis is based on the FOURIER transformation. The method most used in industrial practice today is the fast Fourier transformation (FFT), which is characterized by its fast and efficient transformation algorithm.

Generally, when analyzing process information in the frequency range, one must bear in mind that the time information of when a certain signal event occurs is lost. For this reason, often analyses in the time and frequency ranges are combined. This will be dealt with later (wavelet analysis).

Saxler's system for recognizing grinding burn via acoustic emission analysis again provides a practical example for signal analysis in the frequency range. Roughing and finishing signals during tooth flank profile grinding undergo an FFT within a frequency range of 200–400 kHz. Throughout the tool life of the grinding wheel, the profiles show an increase of about 1 dBV in the averaged acoustic emission amplitude. Analysis of the measurement data gathered from the finishing phase demonstrates a clear relation between the averaged acoustic emission amplitudes and the appearance of grinding burn (Fig. 8.38).

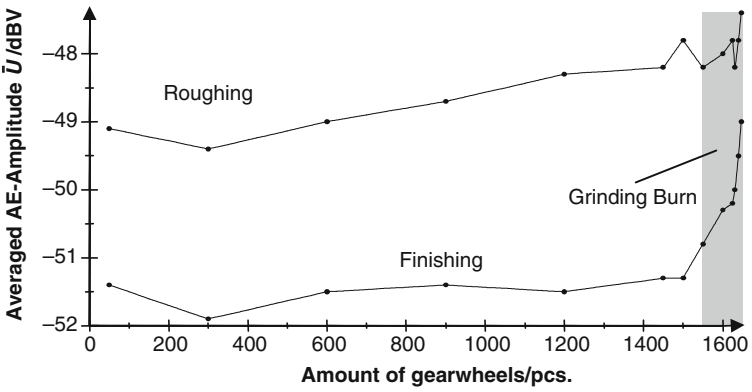

Fig. 8.38 Analysis of acoustic emission in frequency range

Both in the time range and in the frequency range, it is possible to form significant signal characteristic values and to set them in relation with each other. The characteristic values are calculated either from very narrow frequency bands or from the amplitude spectrum in one relevant frequency band. In the first case, the values of the amplitude peaks are usually directly drawn upon for evaluation, in the second often integrating characteristic values are ascertained from the profile of the amplitude spectrum that correlate with the energy content of the signal. Figure 8.39 compares example signals from the time and frequency ranges and illustrates the corresponding characteristic values [Reub00].

The Fourier transformation's usefulness is limited in the case of time-critical signal analyses and highly dynamic processes. One weak point is in the delayed preparation of the evaluation result because information about included signal characteristics can only be made after complete transformation of the signal section. The maximum delay results in case a disturbance-related signal characteristic appears at the beginning of the signal section under consideration and is only recognized after the end of that section. No information about the time or temporal sequence of the appearance of the frequencies contained in the signal exist within the analyzed signal section.

One possibility of increasing the time resolution consists in selecting the width of the signal section so narrow that every signal segment can be considered as quasi-stationary and then to consider the sequential progression of several successive signal sections. This is the functional principle of the short-time FOURIER analysis (STFT). In comparison with FFT, short-term changes in amplitude are not as level. The result for the shortened signal section is represented analogously to FFT. The faster the relevant signal changes arise and the shorter the permissible delay time for their recognition, the smaller the signal sections should be selected. However, a

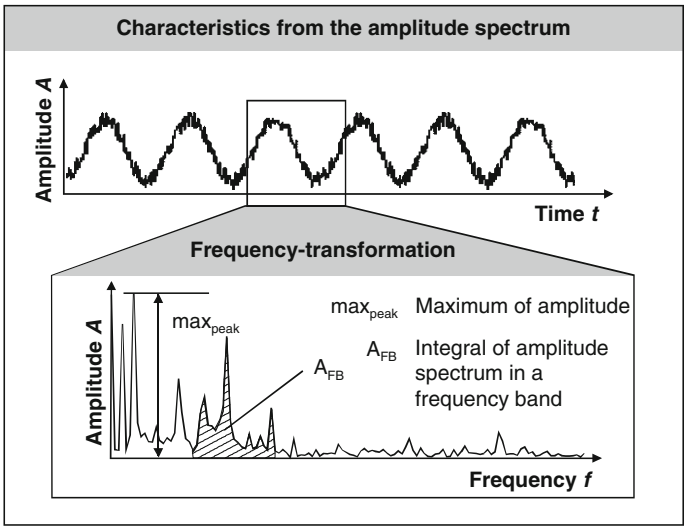

Fig. 8.39 Transformation from time into frequency range and characteristic values

wide signal section is necessary for a high frequency resolution capacity in order to obtain information about the low-frequency signal components that is as complete as possible. One approach to solving this goal conflict is raising the sample rate when digitalizing the signal. This makes a larger amount of discrete values available for the transformation, but requires a larger amount of computing time and memory requirements.

Wavelet analysis is the logical development of the notion of analyzing a signal in real time and completely in its contained frequencies. Different frequency ranges of a signal can be investigated with different temporal resolution. While in the case of the FOURIER transformation only a constant signal section can be selected for the entire frequency range of the signal, the wavelet algorithm adjusts the size of the section to the respective frequency band under consideration. So in the case of high-frequency signal components, there is a very good temporal resolution and at low frequencies very good spectral resolution. Another advantage is that the transformed signal can be fully reconstructed by means of an inverse wavelet transformation.

We differentiate between continuous wavelet transformation (CWT) and discrete wavelet transformation (DWT). In CWT, the output signal is multiplied by a wavelet function. The wavelet function ψ has a constant number of vibrations and is thus a wave packet. It has a shifting parameter τ and a scaling parameter s (Eq. (8.56)).

$$\psi_{\tau,s} = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \tag{8.56}$$

The shifting parameter contains the time information in the transformation range, while the variation of the scaling parameter correlates with the frequence-related evaluation (Fig. 8.40).

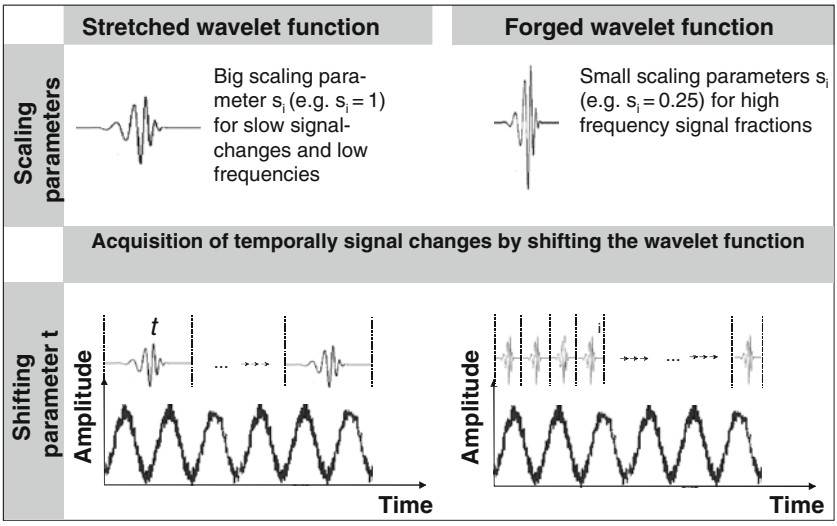

Fig. 8.40 Shifting and scaling of the wavelet function

The signal is temporally resolved by shifting the wave packet along the time axis, as the shifting parameter contains the current location along the time axis. By varying the scaling parameter, the wave packet is expanded or compressed so that its length becomes a measure for the analyzed frequency range.

In accordance with the principle of multiple solutions, the temporal signal profile is transformed several times while varying both parameters. The product of the respective signal section $x(t)$ with the wave packet creates a function whose integral corresponds with the wavelet coefficient c (Eq. (8.57)).

$$c(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \; \psi\left(\frac{t-\tau}{s}\right) \mathrm{d}t \tag{8.57}$$

The calculated coefficient is a measure for the similarity between the analyzed signal section and the shifted or scaled wave packet. Pattern recognition is an important strength of the algorithm. In order to exploit this property, it is necessary to adjust the wave packet to the current signal characteristics by selecting a suitable basic form.

For practical applications in signal transmission, often the discrete wavelet transformation (DWT) is used since they require less computing time and provide extensive possibilities in signal profile evaluation and presentation. The key difference between CWT and DWT is that in the case of discrete wavelet transformations the signal is broken down into individual frequency ranges by repeated highpass and lowpass filtering. The individual layers of analysis are called decomposition layers. Only the highpass-filtered signal components are coded with wavelet coefficients. The output signal of the lowpass filtering is prepared for the next evaluation level (Fig. 8.41) [Reub00]. "Downsampling" compresses the signal by purging the number of discrete individual signal values of information represented in the wavelet coefficient (making it redundant) after highpass/lowpass filtering.

In the case of DWT, a wavelet coefficient represents, simply considered, the difference of two individual signal values. In every decomposition level, a vector of wavelet coefficients is thereby formed which has half as many inputs as the decomposed signal. It represents exactly that signal information that is contained in the frequency band resulting from highpass filtering of the respective decomposition level and can thus be assigned to a certain frequency band. The signal vector resulting from lowpass filtering forms the basis for the next evaluation level. It is thereby reduced by half of the individual signal values, which is possible, in agreement with the Shannon sampling theorem, for the half-band frequency range considered in the next evaluation level without loss of information.

The advantage of this procedure is that the amount of individual signal values used in every transformation level for the computation algorithm is adjusted to the decomposed frequency band. While transformation into the frequency range is bound to a constant number of signal values as a function of the spectral resolution, in the case of "downsampling" during wavelet transformation the number of discrete signal values in each processing level is clearly reduced. This saves not only

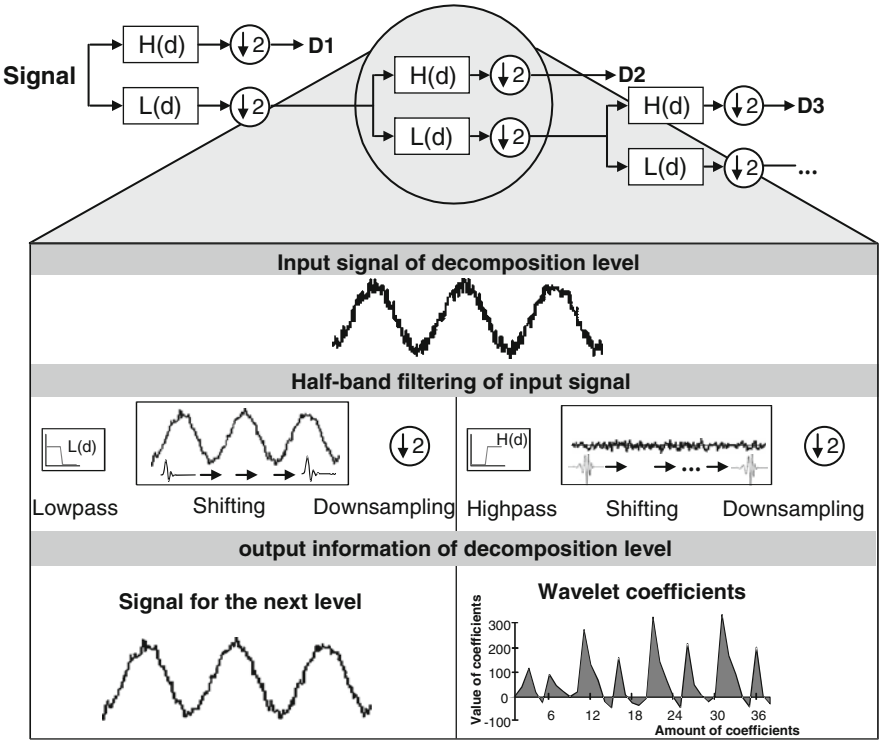

Fig. 8.41 Signal decomposition by discrete wavelet transformations

memory capacity but also computation time, which is advantageous particularly in the context of online monitoring.

The dissertation of PLAPPER provides a practical application of wavelet analysis [Plap04]. In it, the good pattern recognition properties of the method is utilized to detect local defects on the guideways of ball screw guides. Some defects on the guideways made their presence known in such an unsteady fashion in the internal control signals of the machine tool that they could not be traced back to the overrun frequencies in the time signal by means of a Fourier transformation. A clear retrace was possible with wavelet analysis.

A further practical example is given by REUBER in his dissertation on process monitoring finish milling operations on free formed surfaces [Reub00]. In it, the distinct sensitivity of the algorithm with respect to deviations from the constant signal patterns of the disturbance-free process is exploited to generate dynamic, wear-dependent characteristic values. In the investigations, a wavelet function of the DAUBECHIES family [Daub93] was used with the regularity 8 (degree of differentiability). Due to its steep flanks, the function approximates the profile of the resultant force signals very closely and also makes it possible to separate the individual frequency ranges effectively. Three characteristic values were investigated

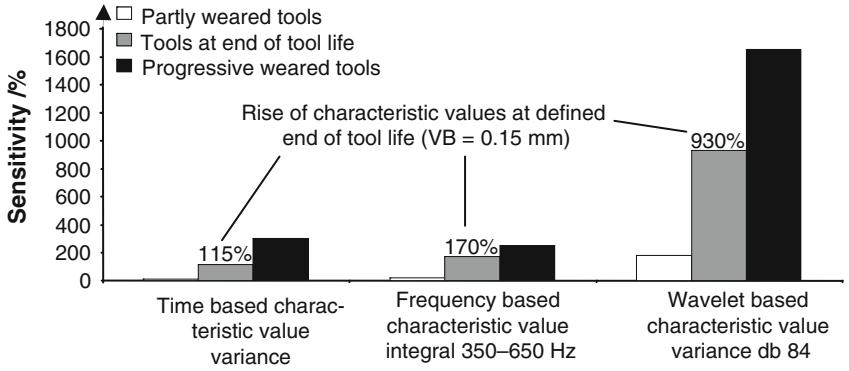

Fig. 8.42 Comparison of wear sensitivity of dynamic wear characteristic values

with respect to their sensitivity to wear: variance of the temporal signal profile, the integral of the amplitude spectrum in the frequency range of 350–650 Hz and variance of the wavelet coefficient in the respective wear-sensitive decomposition level (Fig. 8.42).

Comparison of these characteristic values shows that variance of the wavelet coefficients exhibits a significantly higher level of sensitivity than the characteristic values from the time signal and the amplitude spectrum. The pattern recognition effect represents wear-relevant signal developments much better than they can be recognized with the increased intensity of signal amplitudes in the frequency range or from the temporal signal profile.