THE SCOPE OF THE STRAIN GAGE PRINCIPLE

Prof. Dr.-Ing. Klaus Bethe
Institut für Elektrische Meßtechnik
und Grundlagen der Elektrotechnik
Technische Universität Braunschweig
Hans-Sommer-Straße 66
D-3300 Braunschweig

Abstract
The 50 years old strain gage principle has found widespread application in the measurement of dynamometric quantities, i.e. sensors for force/mass, pressure, torque or acceleration. It further plays an important role in the measurement of geometric quantities like displacement or level. In total, about 50% of all sensors (by value) rely on the strain gage principle. A very important second application of strain gages is the "experimental stress analysis" of structural parts.

The piezoresistive effect, i.e. the change of electrical resistance upon mechanical (elastic) deformation can be described by a "gage factor" k resulting from a geometric effect plus a stress-induced modulation of the electric field charge transport. Depending on the relative direction of stress and electric field, a "longitudinal" and a "transverse k-factor are defined. These gage factors ranging near a value of 2 (in metals), can reach more than 100 in semiconductors. On this basis a broad spectrum of sensors can be realized, reaching from the especially tailored super-precision transfer load cell (max. error 2*10^{-3}/diameter: 25cm/price: 2500$) down to a massproduced tiny monolithic silicon pressure sensor (±1%/6mm/$5). Finally, a short look at the potential of digital microelectronics in connection with these sensors shows up three areas of interest:

1. The improvement of the imperfect sensor (repectively, the possibility to make the naked sensor less perfect)
2. Digital sensor signal electronics avoiding aging and drift

 Which of these concepts will be realized and where they will be allocated (in the sensor housing or somewhere else) is mainly a question of environmental conditions, reliability and yield/price.

Introduction
In industry and trade and, more recently, in the private household numerous sensors (= transducers) are needed for monitoring, control and accounting purposes. (See table 1)

Table 1: Fields of application of sensors
- Process engineering (basic industry: petrochemistry, iron, concrete; chemistry)
- Mechanical production (Robots, CIM; Product quality)
- Power industry
- Traffic, transport, mining industry
- Structural safety
- Trade
- Heating plants, room climate
- Consumer goods (automobile, domestic equipment)

These sensors, i.e. the eyes, ears, noses or fingers of modern electronic circuitry take part in the blooming of electronic's market resulting in a turnover of 17*10^9 DM in 1986 (world-wide). One third of this turnover is being achieved in Western Europe. Among these 5,4*10^9 DM more than 50% were due to dynamometric sensors, i.e. sensors for force and force-related quantities like mass, pressure, momentum etc. That's why it makes sense to have a closer look at this predominant sensor family, particularly because it often forms the bottleneck for a realization of a proposed...
system, either because of availability, price or environmental robustness.

Principle of dynamometric sensors

The usual principle of measuring a force is represented by the classical spring balance: The force to be measured deforms some elastic element, the deformation of which (in a second step) is transformed into an electrically useful signal. Among these deformation sensors by far the most important is the strain gage with its numerous forms of technical realization. (Besides it should be mentioned, that in our days capacitive sensing experiences a remarkable renaissance.)

Piezoresistivity

Application of a tensile load upon a metal wire results in a slight change of its electrical resistance. This effect, discovered by William Thompson (later called Lord Kelvin) in 1846, is called the piezoresistive effect. Its technical realization is the strain gage, on which principle about 50% of all technical sensors/transducers are based. The strain gage principle is applied for the measuring of the following quantities:
- force
- mass
- pressure
- differential-pressure
- torque
- mass flow
- acceleration
- height
- level
- displacement/deformation
- creep
- fatigue
- cracks

Finally, it is an important factor in modern thin film resistive temperature sensors.

All these sensors differ not only in the measurand, but also in measuring range, in their accuracy and in their environmental durability. We further have to discern between mass products for consumer applications with a typical price of 6$ or less and, on the other hand, special precision transducers, selling in small numbers only; the latter may cost more than 200 times as much as a consumer sensor. Consequently, for those divergent tasks one has to make use of different sensing materials, different technologies and different designs of the strain gage.

These problems will be the topic of the first part of my presentation demonstrating the whole spectrum of strain gage sensors.

In a second part, however, we shall have a short look at the mutual interactions between modern strain gage sensor's design and microelectronic circuitry.

Strain gage materials

Given a cuboid made of a homogenous, electrically conducting material (Fig. 1 a). Application of an axial tensile force \( F_x \) leads to an elongation \( \Delta l \) of this body, together with a decrease of its transverse dimensions \( -\Delta w; -\Delta t \). Measuring the electrical resistance between the two end faces (1 and 1') of this cuboid one observes a change \( \Delta R \) which is more less proportional to the relative elongation, i.e. the strain:

\[
\frac{\Delta R}{R} = k \cdot \varepsilon
\]  

(1)

A simple analysis of the situation leads to the expression
\[ k_i = 2 - \frac{A}{E} \frac{d(n\cdot\mu)}{n\cdot\mu} \]  \tag{2}

Equation (3) describes the change of the number of free charged carriers \( n \) in the body and the mobility thereof \( \mu \) due to the mechanically induced deformation of the lattice. Thus, the "longitudinal strain gage effect", equation (1), consists of the constant "geometry factor" of 2 plus a generally nonlinear term caused by a modulation of the mechanism of electric charge transport, i.e. a typical solid state effect. In Table 2 a few experimental gage factors \( k_i \) of polycrystalline metals are compiled:

**Table 2:**

**Longitudinal gage factors of some metals**

<table>
<thead>
<tr>
<th>Pure Metals</th>
<th>( k_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al, Ag, Au</td>
<td>3.1...3.8</td>
</tr>
<tr>
<td>Cu</td>
<td>2.9</td>
</tr>
<tr>
<td>Pt</td>
<td>4.4</td>
</tr>
<tr>
<td>Ta, W, Ni</td>
<td>2.9...3.2</td>
</tr>
<tr>
<td>Ni</td>
<td>-8.5...-11(l)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alloys</th>
<th>( k_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Konstantan, NiCr, CrSi</td>
<td>1.9...2.1</td>
</tr>
<tr>
<td>Zeranin</td>
<td>1.05</td>
</tr>
<tr>
<td>Manganin</td>
<td>0.56</td>
</tr>
<tr>
<td>Pt W 10</td>
<td>4.5...5.5</td>
</tr>
</tbody>
</table>

Those materials, revealing a gage factor different from the purely geometric value of 2 are liable for some degree of nonlinearity, they further are sensitive to any metallurgical change (structure, texture, chemical composition).

The above "longitudinal piezoresistive effect" is characterized by a parallel position of electric field and strain.

If, instead of faces 1 and 1' (planes of constant \( x \) in Fig. 1) the two opposing faces 2 and 2' (at constant values of \( y \)) are used as electrodes, we speak of a "transverse piezoresistive effect" in so far as now electric field \( E \) and strain \( \varepsilon \) are in an orthogonal position.

Theory in this case gives for the transversal sensitivity of the free wire:

\[ k_t = 2 - \frac{A}{E_t} \frac{d(n\cdot\mu)}{n\cdot\mu} \]  \tag{4}

The above Table 2 only contains (technical) metals, which are isotropic in that they consist of randomly oriented microcrystals. The situation becomes much more complicated in the case of semiconducting single crystals where, in addition, the crystallographic direction plays its role. Tables 3 and 4 show that in a crystalline semiconductor one can have extremely high positive or negative (or nearly vanishing) piezo-coefficients, depending on crystallographic direction and doping type. This is true for the longitudinal as well as for the transverse effect \( k_1 \) and \( k_t \) respectively. High doping levels reduce the gage-factors \( k_1 \); \( k_t \) considerably (see Table 3; Si).

**Table 3:**

**Longitudinal gage factor \( k_1 \) in semiconducting single crystals for the three fundamental crystallographic directions**

<table>
<thead>
<tr>
<th>( \langle 100 \rangle )</th>
<th>( \langle 111 \rangle )</th>
<th>( \langle 110 \rangle )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si n ( p )</td>
<td>-132</td>
<td>-14</td>
</tr>
<tr>
<td>( \langle 100 \rangle )</td>
<td>(-12)x</td>
<td>(-24)x</td>
</tr>
<tr>
<td>( \langle 110 \rangle )</td>
<td>(+8)x</td>
<td>(+175)x</td>
</tr>
<tr>
<td>( \langle 111 \rangle )</td>
<td>(105)x</td>
<td></td>
</tr>
<tr>
<td>Ge n ( p )</td>
<td>-5.3</td>
<td>+102</td>
</tr>
<tr>
<td>( \langle 100 \rangle )</td>
<td>(-10.9)</td>
<td>(-157)</td>
</tr>
<tr>
<td>( \langle 110 \rangle )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GaAs n ( p )</td>
<td>-3.2</td>
<td>-8.9</td>
</tr>
<tr>
<td>( \langle 110 \rangle )</td>
<td>(-12)</td>
<td>(+36.2)</td>
</tr>
<tr>
<td>InSb n ( p )</td>
<td>-48</td>
<td>-61.5</td>
</tr>
<tr>
<td>( \langle 111 \rangle )</td>
<td>(-0)</td>
<td>(+30)</td>
</tr>
</tbody>
</table>

3-33
Table 4: Transverse gage factor $k_t$ in semiconducting single crystals

<table>
<thead>
<tr>
<th></th>
<th>${100}$</th>
<th>${111}$</th>
<th>${110}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si n</td>
<td>69</td>
<td>-5,6</td>
<td>-28</td>
</tr>
<tr>
<td>(95Ω•cm)p</td>
<td>-1.4</td>
<td>-83</td>
<td>-110</td>
</tr>
<tr>
<td>Ge n</td>
<td>-5,5</td>
<td>63,5</td>
<td>90</td>
</tr>
<tr>
<td>(95Ω•cm)p</td>
<td>5</td>
<td>-51</td>
<td>73</td>
</tr>
</tbody>
</table>

Calculated from data published by Hollander, Vick & Hock (65).

Although being considered for some advanced technical sensors, the transverse piezoresistive effect heretofore is not used deliberately. Nevertheless it plays an important parasitic role in many sensor designs, i.e., those, where a 2-dimensional strain field exists, e.g., in an elastic plate, being used in pressure and many force transducers.

### Practical strain gages

In practice, the strain gage is not a bulky cuboid as shown in Fig. 1, but it is a slim bar, wire or foil stripe, the length $l$ of which measuring at least 5 times its width $w$. In a typical design, the strain gage forms a metal foil meander, being backed up by a thin plastic sheet. This strain sensing device, like a tiny stamp, is cemented to the massive body to be measured. In this conventional technique the strain gage is only sensitive to those components of the strain within the underlying body, which are parallel to the gage length $l$. Transverse strains are not effectively transmitted from the body to the slim sensing resistive stripe.

As an alternative to the just described conventional bonded strain gage the following 3 technologies are quickly expanding their share:

- **b) Thin film**
- **c) Thick (printed) film**
- **d) Integral monolithic Silicon sensor.**

A short description:

In case b) a sequence of thin films is deposited onto the elastic body consecutively: An isolating layer ($\text{SiO}_2$, $\text{Al}_2\text{O}_3$, ...), a strain-sensing resistive layer (NiCr, CrSi, Si or Ge) and a conducting layer, e.g., Au or Al. Vacuum processes like Vapor Deposition, Sputtering or Plasma-CVD are applied to form those above films of typically 1 μm thickness. Contouring of the strain gage resistors is done either by photolithography or by the use of a shadow-mask.

c) Thick film strain gages are based on the screen printing of special pastes, containing glasses ($\rightarrow$ isolation) or metals/metaloxides like RuO$_2$ as strain sensor or e.g., Au for conducting interconnections. The elastic body preferably is made of non-conducting ceramic or glass.

d) The "integral monolithic silicon sensor" starts with an elastic body made of lightly n-doped Si. As a typical example Fig. 2 shows such a Si-membrane, surrounded by a rigid circular rim ($n^+\text{Si}$). In the center region of this membrane — just where the applied pressure $p$ causes maximum strain — a narrow stripe of a $p^+$-layer is formed by a local boron counter-dope. Either a diffusion or an ion-implantation-process is used for this purpose. Now this "diffused piezo-resistor" acts as a strain gage, being isolated by a $p/n$-junction.

![Integrated monolithic Si pressure sensor (principle)](image)

So, the actual strain gage repertoire is the following:

1. Conventional metal foil strain gage
   
   $k_1 = 2; k_2 = 0$

2. Thin film strain gage
   
   a) Metal (NiCr,CrSi): $k_1 = 2; k_2 \leq 0.1$
   
   (PtW): $k_1 = 4.5; k_2 = 2$

3 - 34
b) Poly-Si or -Ge: \( k_1 = 25; k_{\text{int}} = 0.05...0.4 \)
(depending on texture, grain size...)

3. Thick film strain gage
\( k_1 = 9...14; \frac{\kappa}{\kappa_{\text{int}}} = 0.6...0.7 \)

4. Integrated monolithic Si-Sensor
\( k_1, k_2 = -100...+100 \)
\( (k_1 \leq k_2, \text{see tables 3 \\& 4} \)

Table 5:
Comparison of 4 alternative strain gages (with respect to their use in dynamometric sensors)

<table>
<thead>
<tr>
<th>Flexibility in sensor design (measurand, range, size)</th>
<th>Precision</th>
<th>Environmental ruggedness</th>
<th>Output signal (full bridge)</th>
<th>Optimum lot size</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal: any measurand, any size, except miniature sensors</td>
<td>a) low b) good c) very good</td>
<td>a) +130°C (+760°C) b) poor c) medium</td>
<td>2 mV/V</td>
<td>small/medium</td>
<td>medium...high</td>
</tr>
<tr>
<td>2a) any measurand: only medium &amp; small sensors; miniaturization possible</td>
<td>a) very low b) very good c) very good</td>
<td>a) 200°C b) 350°C (Ti N or Pt) c) medium</td>
<td>2 mV/V</td>
<td>medium/high</td>
<td>medium...low</td>
</tr>
<tr>
<td>2b) any measurand: only medium sized sensors; no miniaturization</td>
<td>a) low b) good c) very good</td>
<td>a) 150°C b) medium c) good</td>
<td>20 mV/V</td>
<td>medium/high</td>
<td>medium...low</td>
</tr>
<tr>
<td>3 any measurand: only medium sized sensors; no miniaturization</td>
<td>a) medium b) medium/fair c) good</td>
<td>a) +150°C b) good c) good</td>
<td>6 mV/V</td>
<td>medium/high</td>
<td>medium</td>
</tr>
<tr>
<td>4 only pressure and acceleration; very small sensors only</td>
<td>a) medium b) good c) medium</td>
<td>a) +120°C b) poor c) very good</td>
<td>25 mV/V</td>
<td>high</td>
<td>medium...low</td>
</tr>
</tbody>
</table>

Technical sensors

As already mentioned, a great many of sensors rely on these above four types of strain gages. (Besides, bonded conventional metal foil strain gages are widely used for experimental stress analysis of structural parts.)

In the following photographs a few examples of strain gage sensors for force/mass, pressure and acceleration are shown.
Fig. 5: Compact, low price load cells (Profiled circular plate-type). Thin-film-strain-gage systems made of NiCr. At the right: 20 mm diameter, 120 kg; left rear: 2800 kg. [emg/Philips]

Fig. 6: Early polysilicon strain gages (LP-CVD-Process; substrate: quartz-glass) [Philips, 1976]

Fig. 7: Very low-load force sensor (0,1N). The planar elastic element (21x23x0,2 mm) is made of CuBe2. Four strain gages of poly-Ge at the left resp. right edge (dark rectangles). [emg]
Fig. 8: Commercial pressure-/force-sensor (Poly-Ge on CuBe/AgPd). [Valvo]

Fig. 9: Thick-film pressure sensor. Membrane is made of alumina ceramic (diameter: 34mm) [Transbar]

Fig. 10: Monolithic silicon pressure sensors:
Foreground: 2 single chips
Left: Chips mounted on glass pedestal
Rear: Si-Wafer containing about 300 pressure sensor chips
At the right: Single Si strain gage [Philips]

Fig. 11: Pressure-loaded monolithic Si sensor with radial and tangential strain gages. [Keller]

Fig. 12: Housed Si pressure sensors [Siemens]

Fig. 13: Monolithic Si pressure sensor-chip with rectangular membrane. Two electronic circuits are integrated on the rigid frame offering a frequency-analog- and a voltage/current-analog output signal. [Toyota]
Digital Electronic plus Sensor

If digital microelectronics gets into sensors - may it be in an integrated or hybrid form - what can be done?

a) Improving the imperfect sensor by signal correction
   - Storing of output signal at zero and maximal load
   - Storing of the whole (nonlinear) sensor characteristic/computing true values
   - Correction of environmental influences (e.g. temperature)
   - Computing of creep corrections
   - Self-check/Self-calibration
   (All these measures may allow to make sensors less perfect in the future.)

b) Sensor signal electronics
   - Digital compensator for bridge readout
   - Driving a bridge circuit with binary noise instead of (e.g.) a sinusoidal 5kHz-signal
   - Direct frequency output
   - Digital output/Bussystem

c) Solving special measuring problems
   - Motional weighing
   - Weighing without gravitation
   - Use of a highly nonlinear (progressive) elastic element for range compression and improvement of dynamics.

Which of these ideas (and other similar ones) one day may enter technical sensors depends on questions like
- environmental conditions (temperature, reliability, acceleration, EMC)
- price.