Application Note PA–TS

Understanding Contact Temperature Sensors

An awareness of the principles involved in contact thermometry, knowledge of the various practical technologies available, and an understanding of their current relative cost and performance parameters are important to building or maintaining a competitive edge.

— **Irwin Bluestien**, RdF Corporation

Engineers need to maintain knowledge of all components used in the products that they design. While semiconductors have been evolving at blazing speed for more than 30 years with no letup in view, the evolution of sensor technology has been grinding out improvements at a comparative snail's pace. An awareness of the principles involved in contact thermometry, knowledge of the various practical technologies available, and an understanding of their current relative cost and performance parameters are important to building or maintaining a competitive edge. Some of the following may seem a bit obvious, but discussions with many engineers have convinced me that these fundamentals are often overlooked or misunderstood.

- The contact thermometer measures only its own temperature. It is brought into contact with or proximity to the substance whose temperature is to be measured, and that substance is supposed to bring the thermometer to the same temperature.
- To bring the thermometer to the temperature of the substance being measured requires the transfer of heat, a phenomenon that may involve several other substances. For example, if the sensing element is enclosed in a protective metal sheath, the heat must transfer from the process substance to the sheath, to the internal supporting structure, and then to the sensor itself.
- The amount of heat required to alter the temperature of the sensor may alter the temperature of the process.
- While the sensor is absorbing heat from the process, it is also yielding heat to its own environment.
- The measurement instrumentation measures the electrical effect that is related to the temperature of the thermometer.

Earlier, I said *"and the substance is supposed to bring the thermometer to the same temperature".* Lets examine the process qualitatively to improve our understanding of the factors involved in obtaining accurate data.

Limited Available Heat

Suppose we wanted to measure the temperature of a small quantity of water heated by a limited heat source.

1. To heat the water to the desired temperature using an external heater, it is necessary to heat the vessel, which in turn transfers heat to the water. The quantity of heat that must be added is related to the mass of the heater, plus the mass of the water, plus the mass of the vessel. The water is not heated directly by the heat source but by the vessel. Therefore we must include the specific heat of the vessel, the specific heat of the water, the mass of each, and the thermal conductivity between the vessel and the water. We should also consider the fact that while we are heating the vessel, it is losing heat to the atmosphere via conduction and radiation.

2. If we turn off the heat source when the water reaches the desired temperature it will begin to cool. The rate at which the temperature falls is related to:

A. The difference in temperature between the ambient air and the vessel's outer surface.

B. The thermal conductivity of the air, which is dependent on its moisture content.

C. The velocity of the ambient air surrounding the vessel: As the water cools, it heats the surrounding air. If the air is in motion, the rate of cooling is much faster than it is if the air is stagnant.

D. The thermal conductivity of the vessel: How well does it conduct the heat from the water?

E. The movement of the water in the vessel (is it being stirred?)

F. Convection within the vessel: External temperature gradients will conduct heat through the vessel walls at rates proportional to the temperature difference between the water and the external air. Such thermal gradients will set up convection currents inside the vessel.

3. If a thermometer is added through a port in the vessel, the water must give up the heat necessary to heat the thermometer. This is a function of the mass of the thermometer, its specific heat, and the thermal conductivity between the surface of the sensor and the sensing element within.

4. Additional heat must be supplied to maintain the thermometer at the water temperature. This is dependent on the amount of additional surface area the sensor presents to the ambient air and its thermal conductivity.

These considerations are mentioned not to complicate the discussion, but to point out how difficult it may be to answer simple-sounding questions relating to accuracy or speed of response. When faced with a limited heat source, it's important to restrict the mass and thermal conductivity of the sensor to minimize its effect on the temperature of the system.

In many practical applications it is necessary to measure the temperature of a small quantity of fluid or gas. Not only is the amount of available heat very small, but the temperatures within the system may be subject to rapid change. The guidelines are therefore:

- Minimize the mass of the sensor.
- Maximize the thermal conductivity of the tip of the sensor where the sensing element is closest to the process.
- Minimize the thermal conductivity of the sensor structure where it is not needed to conduct the heat from the process to the sensing element.

Large Systems

If the system has a lot of heat, then it may not be subject to rapid temperature changes; it may, however, contain pressures and flow rates that mandate increasing the sensor assembly mass to withstand the forces applied by the process fluid. In many real–world cases, the increased mass makes it difficult to achieve the response speed necessary for process control.

To minimize the loss of response speed, the thermal probe is usually spring loaded for positive contact at the bottom of the thermowell bore. In some cases, thermal grease is added inside the well tip to improve thermal conductivity.

With the above considerations in mind, what other factors determine the sensor technology most suited to the task? Let's consider the four most popular types of electrical contact thermometers in use today:

- **Thermocouple**
- **Thermistor**
- **Semiconductor**
- **Platinum RTD**

Thermocouples

The oldest and still the most common industrial thermometer is the thermocouple. Its origin can be traced back to Thomas Seebeck's 1822 report of the Seebeck effect. The correct definition of the Seebeck effect is that *a difference of potential will occur if a homogeneous material having mobile charges has a different temperature at each measurement contact.*

This definition is different from the one you find in most publications. Seebeck's original experiment was with a single homogeneous conductor. He noted that a voltage difference appeared when the wire was heated at one end. Regardless of temperature, if both ends were at the same temperature there was no voltage. If the circuit were completed with wire made of the same material there was no current flow. (see **Figure 1**)

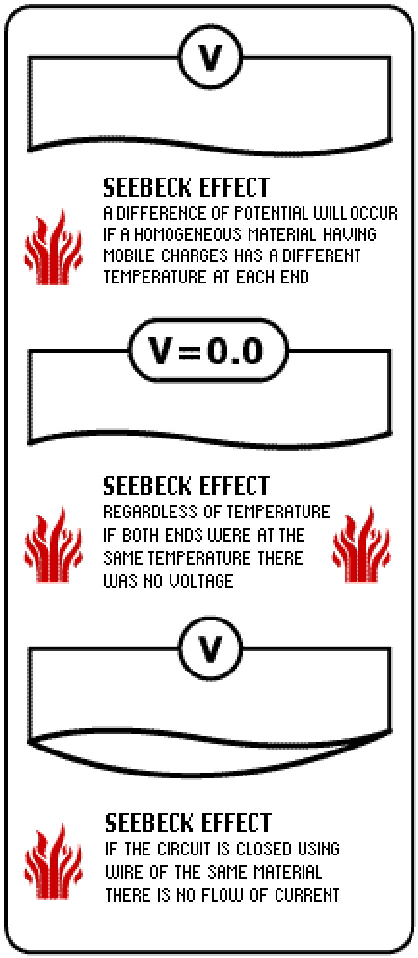
The thermoelectric voltage resulting from the temperature difference from one end of the wire to the other is actually the sum of all the voltage differences along the wire from end to end:

$$V = \int_0^L S\left(dt/dx\right)dx \qquad \textbf{(1)}$$

where:
S = Seebeck coefficient in µV/°C

As long as the thermocouple wires remain homogeneous, the temperature difference between the endpoints determines the net effect of the summation of all the infinitesimal voltages developed across the infinitesimal distances resulting from the temperature gradient. The Seebeck coefficient of a particular wire material is the first derivative of the thermoelectric voltage as a function of temperature—in other words, the difference in potential that results from a difference in temperature. Unfortunately, the Seebeck coefficient is not a constant but instead varies with temperature. The Seebeck coefficient for Type K thermocouples, for example, drops rapidly below -40°C, where it is ~36 µV/°C, to 17.8 µV/°C at -190°C. It peaks at ~630°C at 42 µV/°C.

Thermocouples can be made from a variety of metals and cover a temperature range of -200°C to 1800°C. Advances in metallurgy have made the production of high–purity thermocouple material more reliable than in the past. Comparisons of thermocouples to other types of sensors should be made in terms of the tolerance given in ASTM E 230 or product data.



SEEBECK EFFECT
A DIFFERENCE OF POTENTIAL WILL OCCUR IF A HOMOGENEOUS MATERIAL HAVING MOBILE CHARGES HAS A DIFFERENT TEMPERATURE AT EACH END

SEEBECK EFFECT
REGARDLESS OF TEMPERATURE IF BOTH ENDS WERE AT THE SAME TEMPERATURE THERE WAS NO VOLTAGE

SEEBECK EFFECT
IF THE CIRCUIT IS CLOSED USING WIRE OF THE SAME MATERIAL THERE IS NO FLOW OF CURRENT

Figure 1

Thermocouples have three really significant advantages:

- They are capable of being used to directly measure temperatures as high as 1800°C.
- The thermocouple junction may be grounded and brought into direct contact with the material being measured.
- They can be very rugged. Thermocouples are frequently made of swaged thermocouple wire by a process that entails drawing the steel sheath and the thermocouple wires while at the same time crushing and compacting the internal ceramic insulation. They are also made by inserting an assembly consisting of thermocouple wires threaded through ceramic beads into a steel, Inconel, or ceramic sheath. Beaded construction is most common for the higher temperatures.

That's about the end of the advantages. Here are a few disadvantages:

- Every temperature measurement made with a thermocouple requires that two temperatures be measured, the junction at the work end (usually called the hot junction) and the junction formed when the thermocouple wires meet the instrumentation (copper) wires (cold junction). There are actually two cold junctions formed at the measuring instrument. The positive thermocouple lead and one of the copper leads form one junction and the negative thermocouple lead and the other copper lead form the second. Both junctions are considered to be at the same temperature, and, if that is the case, there is no additional error introduced by the copper wires of the measuring instrument. In the past, the cold junction was held at 0°C by actually placing the wires in an ice bath or by the use of an electronic thermocouple reference junction. Today, most electronic instruments measure the temperature at the terminal block using a semiconductor thermistor, or RTD and digitally adding in the additional voltage that would have existed at the terminal block if it had been at 0°C.
- Thermocouples may appear simple but their operation is complex, with many potential sources of error. The fact that the thermoelectric voltage is developed along the entire length of the thermocouple wires explains the reason that loss of homogeneity, which may result from corrosion occurring anywhere along the length of the thermocouple wire, will introduce error.
- The materials of which thermocouple wires are made are not inert. They are subject to instability resulting from a variety of factors such as the atmosphere to which they are exposed and the rearrangement of their molecular structure resulting from temperature exposure.

- Thermocouple measuring circuits must not draw current. The wire has much higher resistivity than copper, and despite the fact that a short circuit exists at the hot junction, by the time it reaches the measuring instrument it can become an effective antenna for radiated interference (EMI).

- If there is significant distance between the point in the process where the temperature measurement must be made and the electronic measuring instrument, the cost of the thermocouple wire connecting the two may become an important consideration.

- The relationship between the thermocouple millivolt signal and the process temperature is not linear. Because of the availability of microprocessor based measuring instrumentation, this is much less a factor today than it was in the past. The equations representing thermocouples are polynomials consisting of constants to 10 decimal places with temperatures raised to the -23rd power.

- When time response requirements require that the thermocouple junction be grounded, an isolated signal conditioner is essential to prevent the ground loop that would result from a direct connection.

For many industrial measurements the thermocouple is the only cost–effective solution but if accuracy and stability are required, the calibration of the installed sensors must be checked frequently. This calibration should be carried out while the thermocouple is in use by comparing it to a nearby comparison thermocouple. If the thermocouple is removed and placed in a calibration bath, the output integrated over the length is not reproduced exactly and it must be assumed that the thermocouple is still homogeneous. Since this is an unlikely situation, it must be concluded that this type of calibration is flawed and if accuracy is important, the thermocouples should be replaced routinely.

Thermistors

It seems that Seebeck almost discovered semiconductors. When we study semiconductor physics we learn about three energy levels referred to as bands: the valence band, the conduction band, and the gap between them (depletion band). Thermal excitation of a pure semiconductor (the addition of heat energy) causes more carriers to be available for conduction, a phenomenon referred to as intrinsic conduction. At low temperatures the conductivity is very lowbecause of the shortage of intrinsic carriers, so the material is usually dopedwith additional carriers. Their contribution is known as extrinsic conduction.

Seebeck was working with metal conductors and would have seen only the wire resistance increasing with temperature, although he also saw an increase in voltage resulting from carriers moving from the valence to the conduction bands. The characteristics of the metal conductors in conjunction with the observed changes in thermoelectric voltages so thoroughly complicated the situation as to result in his noting the effect without explanation. The effect bearing his name is much more pronounced with semiconductors than with conductors.

A thermistor is a semiconductor with all its material at the same temperature. As a consequence, no difference of potential exists. Thermistors are mixtures of metal oxide semiconductor materials formulated to exhibit specified characteristics. The relationship between the resistance of the thermistor and its temperature is a combination of its intrinsic and extrinsic characteristics.
Intrinsic conduction can be expressed as:

$$R_t = R_{t_0} \cdot e^{\left[\beta\left(1/T - 1/T_0\right)\right]} \qquad \textbf{(2)}$$

where:

T = temperature in kelvin

T_0 = reference temperature in kelvin

β = a constant of the material that represents the change in lnR_t as a function of temperature.

Attempts to modify Equation (2) for extrinsic conduction resulted in several empirical equations that serve to further complicate calculations and are therefore rarely used. Current practice is to determine polynomial equations to fit the characteristics of the particular thermistor as follows:

$$lnR_t = A_0 + A_1/T \ldots + A_N/T^N \qquad \textbf{(3)}$$

$$1/T = a_0 + a_1\left(lnR_t\right) \ldots + a_n\left(lnR_t\right)^n \qquad \textbf{(4)}$$

An accepted practical version, the Steinhart–Hart equation **(5)**, is used by thermistor producers.

$$1/T = \mathrm{A} + \mathrm{B}\left(lnR_t\right) + \mathrm{C}\left(lnR_t\right)^3 \qquad \textbf{(5)}$$

With the introduction of the transistor, temperature compensation of electronic circuits became a much more important task than it had been with vacuum tubes. The latter required heaters to function and tended to become stable after a short warm-up period. Transistors require nowarm-up and their dynamic characteristics are very much affected by temperature. Thermocouples are not well suited for this purpose because the measurement must be relative to a fixed cold-junction temperature, making the implementation impractical. Into the 1970s RTDs were made of wire. Platinum wire RTDs were often considered too costly and some RTDs were made of nickel or copper wire or foil instead. Today the RTD cost comparison is reversed because platinum is available in high–performance, low cost thin film and efficient wire designs.

The leakage current associated with semiconductors was often used in temperature stabilization circuits. It is a consequence of the Seebeck effect on a reverse–biased semiconductor junction, and the development of the thermistor was a natural offshoot of that technology. Using homogeneous material (no junction) controlled to yield specific resistance vs. temperature characteristics, thermistors were developed that could provide accurate temperature measurement. Early thermistors produced in 1960 cost as much as $40.00 each, but found use in many applications because of their great sensitivity and small size.

Thermistors are restricted to measuring temperature over only a portion of their limited operating temperatures because of the dynamic range required if they are connected via an amplifier. A typical thermistor having a resistance of 14,000Ω at 0°C will have <500Ω at 100°C (dynamic range of 28:1). The excitation current must be kept low to avoid self-heating; if the thermistor is made larger to increase its mass and thus decrease self-heating concerns, response speed is sacrificed. At the time thermistors were invented, RTDs were made by winding small-diameter wire on a mandrel. The thermistors were smaller, faster, and more rugged, and quickly became much lower in cost.

Semiconductor Circuits

There are families of IC temperature sensors available today that produce linear current or voltage signals related to temperature over the -40°C to 150°C range. Such circuits could have been produced using discrete semiconductor components but that would have been of little advantage. Why place the entire circuit in the environment to be measured when you could put a thermistor there and keep the balance of the circuitry in a better regulated environment? The development of the first linear semiconductor IC to produce a linear current–to–temperature relationship (the Analog Devices AD590) was an achievement that was quickly recognized and copied. Contrary to the general practice of digitizing everything, these are analog circuits that linearize the Seebeck characteristics of thermistors over a limited temperature range.

Platinum Resistance Thermometers

RTDs are based on the fact that the resistance of metal wire increases with temperature. Practical devices are made of metals that exhibit relatively large temperature coefficients of resistance. The obvious and best choice is platinum because it has a relatively high resistivity, is very stable, and does not alloy into another material, regardless of the atmospheric conditions to which it is exposed.

When RTDs first appeared, the only practical way to make them was to wind a fine metal wire around a nonconducting mandrel. Platinum is expensive, and the ability to draw this soft metal into wire was not nearly so refined as it is today. As a consequence, RTDs have been made from a variety of metals, the most popular of which were copper, nickel, and nickel iron alloys. Although all these metals have positive temperature coefficients of resistance, each has different characteristics. Their temperature coefficients of resistance differ in magnitude and linearity, and the resistivity of each is different.

Copper, the most linear but with the lowest resistivity, was widely used by the manufacturers of electrical power generators and transformers. Its low resistivity led to the 10Ω copper RTD. Nickel and nickel–iron alloys were used because they exhibit higher temperature coefficients and the wire is both much easier to work with and less expensive than platinum. The nickel–iron alloy became popular because the material is made especially for temperature sensor use and its high resistivity supported higher resistance RTDs. The definition of pure nickel has changed over the years as the ability to produce higher purity grades improved. While nickel and the nickel alloy are relatively stable, they do not rival platinum. The materials are not interchangeable batch to batch and operation is restricted to 260°C. The added processing required to attain reasonable interchangeability adds more cost than can be saved by using the lower cost resistor material. The use of the nickel materials in new sensor designs persists because it may be formed by chemical etching into useful configurations. The bottom line is that platinum is much more stable, and when it comes to long–term measurement of temperature the others are not even close.

The ability to purify platinum and knowledge of its characteristics led to the early development of very stable, albeit very expensive devices. Today, the lower cost of manufacturing platinum wire eliminates any price justification for use of the other metals. In fact, platinum now costs less in most cases. The development of thin film technology has resulted in the production of reliable platinum RTDs at costs so low that wire-wound platinum devices are used in a very small minority of new situations. The standard of 100Ω at 0°C arose quickly and a few variations based on the achievable temperature coefficient of resistance became popular.

The technology of sputtering thin film platinum devices resulted in the addition of controlled amounts of impurities and/or stabilization exposures in the range of 900°C to control the temperature coefficient of resistance. Platinum RTDs are controlled worldwide to a universally accepted standard incorporated in DIN 43-760 and IEC 751, which are essentially identical.

Progress in semiconductor manufacturing technology has been applied to the manufacture of thin film RTDs and has resulted in very inexpensive and reliable components. Today's thin film platinum RTDs are well matched to the supporting substrate. Their failure rates are comparable to those achieved by the most reliable wire–wound units. The mean time between failures achieved in practice appears to have more to do with the packaging of the devices than with the process used in their manufacture.

The development of thin film RTDs has progressed to incorporate practical capabilities necessary for unrestricted use:

- Strong lead attachment
- Tightly controlled temperature coefficients including the effects of material bonding on alumina supports
- Automated trimming to resistance tolerances that are the same as those for manually trimmed wire RTDs
- Thin film RTDs packaged in standard packages similar to semiconductors to support automated manufacturing

Thin film RTDs are inexpensive, stable, reliable, small, fast, accurate, and manufactured to the same internationally recognized standards by every manufacturer in the world who makes these devices.

Summary

Semiconductor based devices, thermistors, and ICs are well suited for use as component parts of electronic apparatus. The practice of packaging them in metal housings with threaded fittings would never have become established if today's platinum devices had existed 20 years ago. RTDs and thermocouples are packaged with threaded fittings and connection heads and thermowells in order to facilitate their replacement upon failure.

Thermocouples are exposed to the harshest environments and highest temperatures and are therefore prone to failure at rates sufficiently high as to continue to justify the common industrial configurations.

When transistors first appeared, our vacuum tube mentality told us to expect failures and mount them in sockets. The failures proved to be mostly failures of the sockets and equipment reliability soared as the sockets were eliminated.

RTDs have advanced so far in the past decade that the practice of packaging them like thermocouples will gradually disappear. Most failures associated with current RTD installations are failures in the interconnecting hardware. As manufacturers learn to trust the modern RTD they will learn to mount them permanently in locations that will facilitate optimum operating efficiency at the expense of easy maintenance.

Considering that available, low–cost ICs can accurately digitize, in engineering units, the resistance of platinum RTDs with more than sufficient resolution to achieve all the accuracy available, it no longer makes any sense to sacrifice the universal interchangeability and stability of the RTD for the increased sensitivity of the semiconductor based devices. Manufacturers who are slow to recognize the significance of this will find themselves at a serious competitive disadvantage. Their equipment will not only cost more, but it will also be less reliable.

While there will always be certain applications that require the fastest or most sensitive component regardless of cost, it is clear that in the temperature range from -200°C to 600°C the use of any contact temperature sensor other than a platinum RTD will be increasingly difficult to justify.

Irwin Bluestein is National Sales Manager,
RdF Corporation. ©RdF, 2003, edited.
Original Publishing: SENSORS, January 1999,
AN ADVANSTAR PUBLICATION