Heat treatment and quenching of gears is a complex process. Previous processing variables, as well as the heat treatment parameters of time, temperature, agitation, and quenchant must be carefully scrutinized for proper control of microstructure, distortion, and control of microstructure.

D. Scott MacKenzie**
Houghton International Inc.
Valley Forge, Pa.

The heat treatment of gears is a delicate balancing act. It is necessary to achieve specific mechanical properties at the case and core while achieving low part distortion. Excessive distortion can result in the need for excessive machining to achieve the desired dimensional properties and can cause too much noise in a gear train. Excessive residual stresses can also shorten the life of a gear because of mismatch or shortened fatigue life.

Quenching is defined as “the controlled extraction of heat.” The most important word in this definition is controlled. A quenchant is any medium that extracts heat from the part, and can be a liquid, solid, or gas.

Three stages of quenching (Fig. 1) when a hot part comes into contact with a liquid quenchant are:

- Vapor stage (Stage A or Vapor blanket)
- Boiling stage (Stage B or nucleate boiling)
- Convection stage (Stage C or convection cooling)

The vapor stage is encountered when the hot part surface initially comes into contact with the liquid quenchant, and the part becomes surrounded with a blanket of vapor. In this stage, heat transfer is very slow, and occurs primarily by radiation through the vapor blanket. Some conduction also occurs through the vapor phase. The blanket is very stable and its removal can only be enhanced by agitation or speed-improving additives. This stage is responsible for many of the surface soft spots encountered in quenching. High-pressure sprays and strong agitation can eliminate this stage, but undesirable microconstituents can form if they are allowed to persist.

The second stage encountered in quenching is nucleate boiling, where the vapor stage starts to collapse and all liquid in contact with the component surface erupts into boiling bubbles. This is the fastest stage of quenching; high heat extraction rates are due to carrying heat away from the hot surface and transferring it farther into the liquid quenchant, which allows cooled liquid to replace it at the surface. Many quenchants contain additives to enhance maximum cooling rates obtained by a given fluid. The boiling stage stops when the temperature of the component’s surface reaches a temperature below the boiling point of the liquid. For many distortion-prone components, high boiling temperature oils and liquid salts are used if the media is fast enough to harden the steel, but both of these quenchants see relatively little use in induction hardening.

The final stage of quenching is the convection cooling, which occurs when the component reaches a tem-

**Member of ASM International and member, ASM Heat Treating Society
temperature lower than the quenchant boiling temperature. Heat is removed by convection and is controlled by the quenchant’s specific heat and thermal conductivity and the temperature differential between the component’s temperature and that of the quenchant. The convection stage is usually the slowest stage, and is typically where most distortion occurs. Figure 2 shows the occurrence of the three phases of quenching on a cylindrical temperature probe.

The cooling characteristics of a quenchant can be represented as temperature as a function of time or temperature as a function of cooling rate (Fig. 3) for both normal speed and high speed quenching oils. However, cooling curves produced under laboratory conditions must be interpreted carefully and should not be considered in isolation. Results on used quenchants should be compared with reference curves for the same fluid.

The duration of the vapor phase and the temperature at which the maximum cooling rate occurs have a critical influence on the ability of the steel to harden fully. The rate of cooling in the convection phase is also important because it is generally within this temperature range that martensite transformation occurs and it can, therefore, influence residual stress, distortion, and cracking.

Obtaining the desired properties and low distortion is usually a balancing act. Often, optimal properties are obtained at the expense of high residual stresses or high distortion, and low distortion or residual stresses are usually obtained at a sacrifice in properties. Therefore, the optimum quench rate is one where properties are just met. This usually provides the minimum distortion.

**Oil Quenchants**

Many types of oils have been used in hardening ferrous metals including vegetable, fish, and animal oils, and certain sperm-whale oil. The first petroleum-based quenching oils were developed around 1880 by E.F. Houghton in Philadelphia. Since that time, much advancement has been made in quenching oil development to provide highly specialized products for use in specific applications. A wide range of quenching characteristics can be obtained through careful formulation and blending. High-quality quenching oils are formulated from refined base stocks of high thermal stability. Selected wetting agents and accelerators are added to achieve specific quenching characteristics. Complex anti-oxidant packages are included to maintain performance for long periods of continued use – particularly at elevated temperatures. Emulsifiers may be added to enable easy cleaning after quenching.

Petroleum-based quench oils can be divided into several categories, depending on the operational requirements including quenching speed, operating temperatures, and ease of removal.

Normal-speed quench oils have relatively low rates of heat extraction, and are used in applications where the material being quenched has a high hardenability. Medium-speed
Marquenching oils are a special case where the part is immersed into a quenchant at elevated temperature, typically 100 to 200°C (210 to 390°F). The workpiece is held in the quenchant until temperature equilibrium is established throughout the section, and then air-cooled to ambient temperature. During marquenching, components are quenched to an intermediate temperature close to the Ms temperature and held at this temperature. This eliminates temperature gradients across the surface, and consequently, during subsequent slow cooling after removal from the hot oil, transformation to martensite occurs uniformly throughout the section, which minimizes the generation of internal stresses and reduces distortion.

Because marquenching oils are used at relatively high temperatures, their formulation and physical properties are different from cold quenching oils. They are formulated from selected base stocks having high oxidation resistance and thermal stability. They have high flash points and viscosities, and contain complex antioxidant packages to provide long life. Selection of the marquenching oil is based on the operating temperature and quenching characteristics. A minimum of 50°C (120°F) should be maintained between the operating temperature of the oil and its flash point.

Heat Treatment

Stresses in parts prior to heat treatment are relieved during the treatment. Stress relaxation causes distortion as the part finds a stress-free equilibrium. Heat-up rates in the furnace can also cause distortion, as thermal gradients are formed and the thinner sections reach temperature quicker. Differences in thermal expansion also can cause sizable thermal strains to be developed within the part, which if large enough can result in plastic deformation and distortion. The use of a preheat stage to allow thicker sections to “catch-up” to the thinner section reduces distortion. The same problems can arise in a furnace that has nonuniform temperature in the workzone.

Racking of parts is a very important part of the heat-treating process.

**Fig. 4 — Fishbone diagram of the many causes of distortion.**
ically between 3-5% for carburized steel) that increases as the carbon content increases. The volume change causes differential transformation strains, which may cause distortion. While these strains may not cause distortion immediately after heat treat, they can appear immediately after any subsequent machining steps, as the part tries to achieve a new static equilibrium. These residual stresses can also manifest themselves by shortened fatigue life.

Proper atmosphere control is important because excessive soot can be carried into the quench oil, creating dirty parts and shortening the life of the quench oil. Proper atmosphere control can also reduce the amount of retained austenite, which can cause residual stresses and distortion.

Quenching
Figure 4 shows many sources of residual stress and distortion. In quenching, the primary source of distortion and residual stress is differential temperatures from the center of the part to the surface and from different locations on the surface. Large reductions in residual stresses and distortion can be achieved by reducing thermal gradients and differential temperatures. Factors that lead to the creation of large thermal gradients in parts during quenching are temperature, agitation, quenchant, and quenchant contamination.

Temperature. Increasing the oil temperature can reduce distortion and residual stresses in a heat-treated component. Temperature gradients in the part decrease as the temperature of oil increases (this is the basic principle of martempering). Increasing the temperature in cold oils up to the recommended use temperature (typically 180-200°F, or 80-95°C) can also reduce thermal gradients. The speed of cold oil can be increased by increasing the temperature of the oil to approximately 160°F (70°C).

Agitation. Distortion occurs due to differential temperature gradients in the part. Figure 6 shows that the three phases of cooling occur in the piece at different times. Therefore, some areas cool very slowly, while other parts cool rapidly, creating thermal gradients on the surface of the part, which can cause distortion. The purpose of agitation is to minimize surface temperature gradients. Quenching characteristics are influenced significantly by the degree of agitation. Increasing the degree of agitation reduces the stability of the vapor phase and increases the maximum rate of cooling. This also has the benefit of minimizing any vapor pockets that can occur, ensuring that the part has a uniform heat transfer across the surface of a part.

Quenchant. There are many types of petroleum-based quenchants. For most gear heat-treating applications, marquenching oils are used almost exclusively because of the benefits of reducing distortion. However, there are certain applications where cold oils are used, specifically in very large sections, or where press quenching is used.

Racking. Racking of gears is critical to minimize distortion. Parts must be positioned so the applied agitation will ensure uniform heat transfer on all surfaces of the gear. Uniformity of heat transfer minimizes the formation of thermal gradients on the surface of the parts. The positioning of parts also should not allow the creation of hot spots from adjacent parts or mechanical damage from part-to-part interactions.

Two primary methods for quenching gears are press quenching and placing on a grid or fixture for immersion. Press quenching is a specialized technique involving the physical restraint of distortion-prone parts on close-tolerance fixtures...
during the quenching operation. It minimizes distortion and movement, and is used mainly during the quenching of bearing rings and automotive transmission ring gears. It is a manually intensive operation, as each gear must be removed from the furnace manually, and placed on a quench fixture. The press is actuated, and a large flow of quenchant is passed through the fixture. Highly accurate and low distortion parts can be achieved in this manner.

There are several disadvantages to this technique. As mentioned previously, it is manually intensive, although some robotic applications have been implemented. Because hydraulic fluids are used to actuate the dies, contamination of the quenchant is a problem. This can cause a change in the cooling rate and quenching characteristics of the quenchant, which can cause cracking or fires. If fire-resistant hydraulic fluids are used, then some spots or cracking can occur on the part or the close-tolerance fixture. The quenchant must be routinely checked for contamination and water content. Close-tolerance fixtures used in press quenching must be designed for each gear configuration and they are expensive. A new fixture must be designed if gear dimensions change. Further, the die life is finite because of the thermal stresses experienced by the fixture. Fixture distortion and cracking also can lead to premature replacement. Generally, cold oils are used to harden the parts. This technique is generally limited to flat and symmetrical parts, such as ring gears.

Many gears are heat treated by placing them on a grid or fixture for immersion in the quench bath. The method greatly improves production rates. There are many ways to rack a gear that often depend on the type of furnace, quenchant, and the preference of the heat treater.

Typically, ring gears are either laid flat on a grid and stacked several grids high, either offset or directly on top of each other. Gears are often hung with supports under the gear. Each method has benefits that depend on the configuration of the gear. If gears are laid flat, they tend to bend (or “potato-chip”); gears on the bottom and top of the load are most prone to this type of distortion. This is due to differential cooling of the gears. In this case, the thermal mass of the grid retains heat, while the upper surface of the gear experiences the full quenching effect of the oil. The upper surface contracts due to thermal contraction, while the lower surface cools slower, and does not experience as much thermal contraction. As the upper surface cools to a point where the martensite transformation occurs, a volume change occurs, placing the upper surface in tension. As the lower surface cools to the martensite transformation temperature, a stress reversal occurs, placing the upper surface in tension, and the lower surface in compression. This is complicated by the round shape of the part, so that some areas bow up, while other areas bow down, resulting in the potato-chip shape. The degree of distortion is often dependent on how stiff the section is (polar moment of inertia). This can be overcome by the proper design of racking fixtures.

The grid must be flat when gears are laid onto it, because the heat treated part is hot and soft, and it will conform to the shape of the grid. Warped and badly cracked grids should not be used. Grids and racks also should be routinely stress relieved to relieve the build-up of quenching stresses over time; this minimizes cracking and extends service life.

When parts are hung, the weight of the gear often causes the gear to distort, with the gear becoming an oval shape. The degree of ovality often depends on the quality of support, and the weight of the part; smaller fully supported parts distort less. An advantage of hanging gears is that all sides experience similar heat transfer, assuming no hot spots or proximity of other parts (creating hot oil spots).

Pinion gears are racked vertically, racked either stem up or stem down. Often, the pinions are offset to allow uniform heat transfer and to minimize hot spots. Usually, spacers are used to maintain the pinions in the vertical position and to prevent movement of the parts.

Contamination and Oxidation. The condition of the quench oil can also contribute to gear distortion. Contamination with water must be avoided at all cost. As little as 0.05% water in quenching oil influences quenching characteristics significantly and may cause soft spots, dis-
properties. Understanding often is and parts on part distortion and understand the interaction of fluid flow treating process, it is difficult to un-
derstanding (FEA, or finite element analysis).

distortion (CFD, or computational fluid dynamics) and predicting part workload (CFD, or computational fluid dynamics) allow pre-

ing quenchants flow around a heat treated gear are quenchant ag-

multations over 0.5%, foaming during quenching is likely, which can cause fires and explosions. Other contam-

ates such as hydraulic oil and fire-resistant hydraulic fluids can also alter quenching characteristics, re-

sulting in increased distortion and residual stresses.

The level of oxidation of quenching oil is measured by the Precipita-

tion number or Total Acid number. As the oil oxidizes, it forms organic acids, and the oxidized constituents decreases the stability of the vapor phase and increases the maximum cooling rate, which can increase the risk of distortion and cracking. The use of stable, high-quality quench oils will reduce the possibility of this occurring. It is recommended to implement a proactive maintenance pro-

gram of monthly or quarterly checks for contamination and oxidation.

Modeling. Two of the most important variables affecting distortion of a heat treated gear are quenchant agita-
tion and racking. Improvements in computer modeling allow predicting quenchant flow around a workload (CFD, or computational fluid dynamics) and predicting part distortion (FEA, or finite element analysis).

Due to complexity of the heat-treating process, it is difficult to un-

erstand the interaction of fluid flow and parts on part distortion and properties. Understanding often is

achieved only by experience (that is, learning from mistakes). Today, there is less tolerance for trial and error and more emphasis on doing it right the first time. Unfortunately, there are few design rules that dictate racking of a part in a given furnace.

CFD allows one to investigate quenchant flow in a quench tank. CFD is a mathematical technique that models continuous fluids with partial differential equations, which are discretized into smaller algebraic equations that can then be solved using advanced mathematical techniques. The primary benefit of CFD is the ability to simulate physical behavior that is difficult to measure. The method allows the designer to simulate performance before building the part, avoiding costly build and test methodology previously used in quenching. It offers high accuracy, rapid results, and is cost-effective, and it has been widely used in a number of industries including metals, automotive, aerospace, chemical and petrochemical.

Four main steps carried out in any CFD model are:

- Model geometry description and meshing
- Material specification and process conditions (boundary conditions)
- Solution of the governing equations
- Post processing results

A detailed three-dimensional de-

scription of a water-tight model is needed in the first step. Usually, this is done by importing a CAD file that represents the quench tank and the carrying basket with the parts to be quenched. Any unnecessary details should be removed. The higher the resolution of the computational mesh, the more accurate the description of the model will be, but the higher the computational also will cost more.

After creating a computational mesh, the model for the CFD simula-
tion is set up by specifying the operating conditions (boundary conditions) and the material properties. After preprocessing, the CFD solver performs the calculations and produces the results.

The flow characteristics of the mesh are determined by solving the Navier-Stokes equations (a series of partial differential equations that describe the flow of incompressible fluids) at each node or corner of the mesh. These equations describe how velocity, pressure, temperature and viscosity of a moving fluid are related. The main process condition in simulating the flow inside the quench tank is the quenching liquid speed and direction as it leaves the agitator(s). Although CFD modeling is capable of modeling the flow inside the agitators, typically, only input and output conditions are defined.

Post-processing is the final step in CFD analysis, and involves organization and interpretation of the enormous data and images. CFD can be computationally intense. Previously, it required the use of specialized CRAY supercomputers or networked RISC workstations. However, the increase in computing capability and improved algorithms allows performing fairly complex CFD models on everyday office computers or laptops.

In most CFD analysis, a static analysis is performed; the effects of temperature are not considered, and that the flow is not time dependant. For most analysis, this is often adequate. The assumption is that if the flow is uniform, then heat transfer is uniform. If the heat transfer is uniform, then temperature differentials are small, and the resulting distor-

Fig. 8 — Typical grid and mesh used in CFD.
tion and residual stresses are mini-
mized. This is often a good first step
when looking at a quench tank for
agitation and flow uniformity.

The methodology can be illus-
trated by considering the heat treat-
ment of transmission pinions in
which distortion was a problem. The
distortion led to variability from load
to load and within the load. CFD was
used to modify the method of rack-
ing, and the flow of the agitation was
examined. The method of racking
was determined using CFD, and re-
sulted in a large reduction in dis-
tortion and distortion variability within
workloads. This translated into a
large cost savings in repair, scrap, re-
work, machining, and warranty re-
pairs. Example of the quench tank,
agitation system and results modeled,
are shown in Figs. 7, 8, and 9.

Determining part distortion
during heat-treating and predicting
part microstructure has been a long-
held (elusive) goal of the heat-
treating industry. The use of FEA has
been used extensively to solve struc-
tural and performance issues of com-
ponents for a long time. It has only
recently been used to predict part dis-
tortion and microstructure.

To accurately predict distortion or
the formation of residual stresses in
a part requires and understanding of
many factors including heat transfer,
elastic-plastic stress and strain be-
behavior, and microstructure. Heat
transfer is not a steady-state condi-
tion, and it requires the determina-
tion of heat-transfer coefficients as
function of fluid properties, geometry,
surface condition, and agitation.
It also is time and location dependant.

Elastic-plastic stress strain be-
behavior requires detailed constitutive
models of stress and strain as a func-
tion of strain rate, location, and tem-
perature. Knowledge of diffusion
transformations (pearlite and bainite)
occurring in the part and the nondif-
fusion transformations (austenite to
martensite transformation, recrystal-
lization, grain growth, etc.) is nec-
essary to accurately predict the mi-
crostructural development and its con-
tribution to distortion and resid-
ual stresses.

The use of FEA of microstructural
development and the development
of residual stresses and distortion is
in its infancy. With the availability of
better constitutive models, this tech-
nique offers great potential to solve
many distortion problems before the
part enters the furnace.

FEA requires data on material,
geometry, and boundary conditions.
Material information includes the
many relationships describing the
phase transformations and the
volume changes associated with the
phase transformations. These consti-
tutive equations are material depen-
dant and often are held proprietary
with the software, or are independ-
ently determined. CAD programs
define the geometry. Boundary con-
ditions depend on the quenchant,
and change as a function of time and
surface temperature.

Typically, only one part is exam-
ined, and boundary conditions,
residual stress, and distortion are
determined. It has only been recently
that CFD and FEA can be combined.
In a recent application, the flow con-
ditions were determined by CFD,
and these were translated into time-
dependent heat transfer coefficients.
The new boundary conditions were
then applied to an FEA model, and

<table>
<thead>
<tr>
<th>Position</th>
<th>Predicted, mm</th>
<th>Max</th>
<th>Avg.</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>A11</td>
<td>0.025</td>
<td>0.020</td>
<td>0.021</td>
<td>0.008</td>
</tr>
<tr>
<td>B11</td>
<td>0.015</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Fig. 9 — Typical results showing relative flow velocities in a quench tank through a rack of pinion gears racked vertically.

Fig. 10 — Predicted distortion of pinions quenched vertically for the flow conditions presented above.
comparative distortion within a workload was determined. This is one of the first practical cases where distortion has been predicted within a furnace workload. In this case, CFD results from the previous example were obtained, and the heat transfer coefficients as a function of temperature and velocity were determined. Using subroutines, the flow was translated into heat transfer coefficients and applied as boundary conditions to the selected parts. The distortion was then predicted as a function of the fluid flow conditions. Actual versus predicted results for pinion gears quenched vertically are shown in the table. Predicted distortion is shown in Fig. 10. The predicted results compare very closely with the actual results.

Bibliography


Pillar Induction Company

Innovative and reliable induction heat treating systems for over four decades.

Corporate Office: 21905 Gateway Rd. • Brookfield, WI 53045
Toll-Free: 1-800-558-7733 • 262-317-5300 • Fax: 262-317-5394
www.pillar.com • sales@pillar.com

ISO 9001: 2000 REGISTERED