Journal of Heat Transfer

Published Monthly by ASME

RESEARCH PAPERS

Micro/Nanoscale Heat Transfer

Bubbles, Particles and Droplets

Heat Transfer in Manufacturing

Evaporation, Boiling, and Condensation

Forced Convection

Conduction

This journal is printed on acid-free paper, which exceeds the ANSI Z39.48-1992 specification for permanence of paper and library materials. ⊗™
☉ 85% recycled content, including 10% post-consumer fibers.

(**Contents continued**)

TECHNICAL BRIEFS

The ASME Journal of Heat Transfer is abstracted and indexed in the following:

Applied Science and Technology Index, Chemical Abstracts, Chemical Engineering and Biotechnology Abstracts (Electronic equivalent of Process and Chemical Engineering), Civil Engineering Abstracts, Compendex (The electronic equivalent of Engineering Index), Corrosion Abstracts, Current Contents, E & P Health, Safety, and Environment, Ei EncompassLit, Engineered Materials Abstracts, Engineering Index, Enviroline (The electronic equivalent of Environment Abstracts), Environment Abstracts, Environmental Engineering Abstracts, Environmental Science and Pollution Management, Fluidex, Fuel and Energy Abstracts, Index to Scientific Reviews, INSPEC, International Building Services Abstracts, Mechanical & Transportation Engineering Abstracts, Mechanical Engineering Abstracts, METADEX (The electronic equivalent of Metals Abstracts and Alloys Index), Petroleum Abstracts, Process and Chemical Engineering, Referativnyi Zhurnal, Science Citation Index, SciSearch (The electronic equivalent of Science Citation Index), Theoretical Chemical Engineering

Local Heat Transfer Measurements in Microchannels Using Liquid Crystal Thermography: Methodology Development and Validation

R. Muwanga
e-mail: rmuwanga@alcor.concordia.ca

I. Hassan[1]
e-mail: ibrahimH@alcor.concordia.ca

Department of Mechanical and Industrial
Engineering,
Concordia University,
1515 St. Catherine W.,
Montreal, Quebec,
Canada, H3G 2W1

Microchannel heat transfer governs the performance of the microchannel heat sink, which is a recent technology aimed at managing the stringent thermal requirements of today's high-end electronics. The microencapsulated form of liquid crystals has been well established for use in surface temperature mapping, while limited studies are available on the use of the un-encapsulated form. This latter form is advantageous since it offers the potential for high spatial resolution, which is necessary for microgeometries. A technique for using un-encapsulated thermochromic liquid crystals (TLCs) in order to measure the local heat transfer coefficient in microchannel geometries is shown in the present study. Measurements were made in a closed loop facility combined with a microscopic imaging system and automated data acquisition. A localized TLC calibration was used to account for a non-uniform coating and variation of lighting conditions. Three test section configurations were investigated with each subsequent configuration arising due to a shortfall in the previous. Two of these configurations are comprised of single wall heated rectangular channels, while the third is a circular tube channel. Validation results are also presented; overall, the methods developed and utilized in this study have been shown to provide the local heat transfer coefficient in microchannels.
[DOI: 10.1115/1.2193541]

Keywords: liquid crystal, heat transfer, microchannel, experimental method

1 Introduction

In recent years the advent of very large scale integrated circuit technology has propelled the electronics industry into the fabrication of electronic circuits with feature sizes on the order of microns and submicrons [1]. This has allowed designers to pack these components with more features at much smaller tolerances and in so doing improve the speed and performance of these electronics. However, as a result, there has been an increase in the heat produced per unit area, which in turn has generated the requirement for effective, miniaturized cooling technologies. One such device is the microchannel heat sink which was pioneered in the early 1980s by Tuckerman and Pease [2]. These heat sinks demonstrated heat transfer rates about two orders of magnitude greater than those of commercial technologies for cooling arrays of integrated circuits [3]. The micro heat sink/exchanger is characterized by a set of microchannels conventionally machined or micromachined into a conducting block.

A number of studies have been conducted in recent years in order to understand and improve the heat transfer performance of these devices and recent summaries may be found in [3–7]. Some of the prevalent characteristics and remarks related to single phase continuum flows follows. There are discrepancies reported between microchannel results and the accepted correlations and trends for their macrocounterparts. For example, the critical Rey-

nolds number for transition to turbulent flow in microchannels has been reported to be as low as 200 [6] compared to the commonly known approximate value of 2300. In addition, significant variations between the results of different investigators have been reported for the heat transfer coefficients and friction factor with no clear relationship on the differences in some instances [4]. Finally, the channel size at which deviations will occur from conventionally sized channels remains undetermined. With the exception of the confinement number for boiling flows, no physics based classification of channel sizes has been proposed [8]. Overall, a review of the literature indicates the need for additional and improved studies of heat transfer in mini and microchannels [6].

Most work regarding experimental heat transfer measurements in microchannels uses thermocouples (point sensors) for wall temperature measurements. One problem with such sensors is their large size compared to the channel dimensions and hence their potential influence on the flow characteristics. This may be alleviated by placing the sensors at a distance from the fluid-wall interface and assume one-dimensional heat transfer, or by using a fin analysis method [1]. These approximations, however, introduce inaccuracies in the internal wall surface temperature measurement. Also, the relative size of these sensors, if discreetly placed at the wall, gives only an area averaged temperature measurement. Another inconvenience of such point sensors is the need for an array of sensors to obtain full surface temperature measurements.

Non-intrusive techniques such as infrared (IR) thermography, laser induced fluorescence and liquid crystal measurements have traditionally been employed [9]. There are a variety of challenges associated with IR thermography, which include accurately characterizing its performance and calibration, determining the measured body's surface emmisivity, identifying the measurement

point, and accounting for radiation losses and additional optical equipment [10]. In addition, for the measurement of microobjects, accurately accounting for the background radiation is necessary. Hetsroni et al. [10], while measuring the surface temperature of a capillary tube, utilized a technique whereby the background temperature was controlled to within that of the tube wall temperature. The aim of this was to minimize the uncertainty in measurement due to background radiation during IR measurement.

Thermochromic liquid crystal (TLC) thermometry is a non-intrusive technique based on the changing pitch size of the helical molecular structure of the material during temperature change. Additional information about the characteristics of liquid crystal and their use in thermometry and heat transfer may be found in Ireland and Jones [11] and Hallcrest [12]. The TLC technique has been used by many researchers since the late 1960s when it was first applied to non destructive testing and flow visualization [13].

The successful use of TLCs on a microgeometry has been demonstrated by Höhmann and Stephan [14] while investigating evaporation from a liquid meniscus. In their work, they utilized the un-encapsulated form of TLCs applied to the backside of a 20 μm stainless steel foil. They were able to obtain spatial resolutions of 0.83 μm, where the uncertainty in temperature was 0.51°C or 7.5%. Their use and limitations in boiling heat transfer have been investigated and discussed by Kenning et al. [15] and Klausner et al. [16]. These studies determined that in many instances, due to the thermal lag between the liquid crystal and the inner wall surface, the response time of the liquid crystal is not sufficient to capture all the unsteady phenomena occurring during nucleation boiling.

A first glimpse at the use of TLCs in minichannels with boiling has been recently presented by Chin [17,18] during an investigation of boiling incipience and convective heat transfer in narrow vertical channels. These authors used the TLCs in their common microencapsulated (as opposed to un-encapsulated) form, which, due to capsule diameters of 10–25 μm, restricts the spatial resolution capability of the TLC material. The use of un-encapsulated TLCs for quantitative, high resolution heat transfer measurements in small channels has, to a limited extent, been investigated [19]. This is likely due to the difficulty in handling un-encapsulated TLCs, since they are easily contaminated by solvents or dust, and prolonged exposure to ultraviolet light can disturb their calibration and even destroy their response [12].

In the work of Aligoodarz et al. [19], rectangular channel geometries 2×1 mm and 2×2 mm with three walls made of 75-μm-thick stainless steel were studied for boiling flows. The fourth wall was made of glass, which was used for visualization while the opposing stainless steel wall was coated with un-encapsulated TLC. The authors document difficulties in obtaining a continuous coating of the TLC on the surface, as well as difficulties in accounting for its stringent requirements for a contaminant free area. Their choice of the un-encapsulated form, however, was motivated by the ability to apply it in thinner layers than the microencapsulated form, hence obtaining an improved response time, which was approximately 4 ms. Another benefit of the un-encapsulated form is that the spatial resolution is now restricted to the optical configuration [20]. Thus, submicron resolution is possible. TLCs therefore offer the potential for improved temperature measurements in microchannels due to their nonintrusive character and possibility for high spatial resolution measurements. The combination of the un-encapsulated TLC's capability and potential, as well as the need for increased measurements in microchannels, was therefore the motivation for the present work. This paper describes the development of a local heat transfer measurement system for mini and microchannels using un-encapsulated thermochromic liquid crystals.

2 Experimental Facility Methodology

2.1 Flow Loop. A schematic of the main components employed in the closed loop test facility are shown in Fig. 1. The
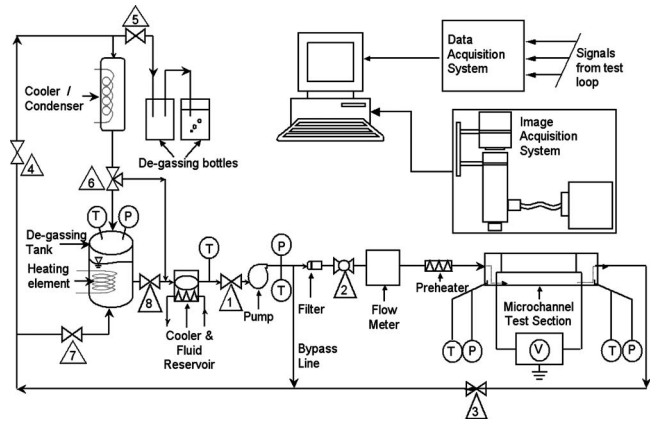
Fig. 1 Schematic of the experimental test facility

flow loop concept was designed to accommodate boiling experiments that will be presented in a later study. Only the components relevant to the single phase experiments are discussed. The fluid used was distilled water, however the facility is set up to handle a variety of refrigerants and engineered fluids.

Flow enters the loop from the main tank and is continuously circulated by a magnetically coupled gear pump (Fig. 1). The pump runs at a constant speed and supplies a flow rate of 290 ml/min with a maximum pressure of 517 kPa (75 psi). A nutating digital output flowmeter provided by DEA Engineering was used to monitor the flow rate. This meter outputs a 5 V square wave signal at a frequency proportional to the time for the nutator to complete one cycle. The flowmeter was initially calibrated using the weighing method, and has a reliable range from 10 to 250 ml/min. Upstream of the flowmeter is a 25 μm filter, which is used to remove any accumulating particles. A preheater is located at the test section for additional flow temperature control, while the exit of the test section has a cooler, which is used to restore the temperature of the fluid.

2.2 Test Module. As mentioned above, heat transfer in mini and microchannels is the main motivation for the experimental methods developed in the present study. Such channels will come in many shapes, which are predominantly dictated by the manufacturing processes available. An additional constraint will be the measurement technique utilized in experiments. To design a heated channel test module utilizing TLCs for measurement, a primary criterion is optical access to the measured surface. In the most basic case, a flat surface is utilized with the camera at its normal. Three test module designs were investigated in this research program with each subsequent design being a consequence of shortcomings in the previous configuration. They are described hereafter and results are presented for the final design.

The initial test concept was manufactured from G-10 fiberglass with three of the four channels produced from a milled out slot. The material's low conductivity allowed for approximately adiabatic conditions. The fourth wall was provided via a thin foil supported on a polycarbonate cover. A schematic of this setup is shown in Fig. 2. The TLC material was applied on the backside of this foil before it was placed on the cover. Several difficulties were encountered with this approach. First, the use of un-encapsulated TLCs directly in the channel caused the coating to erode over time. Unlike the microencapsulated form, the un-encapsulated TLC material does not solidify, but rather remains in an oily state. Therefore, any air pockets present on the backside of the foil where the TLC coating is located will draw fluid and/or will allow for the TLC material to escape over the course of an experiment. A second drawback of this design was the use of a thin foil, which was very difficult to handle, to produce accurately, and to align. A third drawback of this design was the use of an O
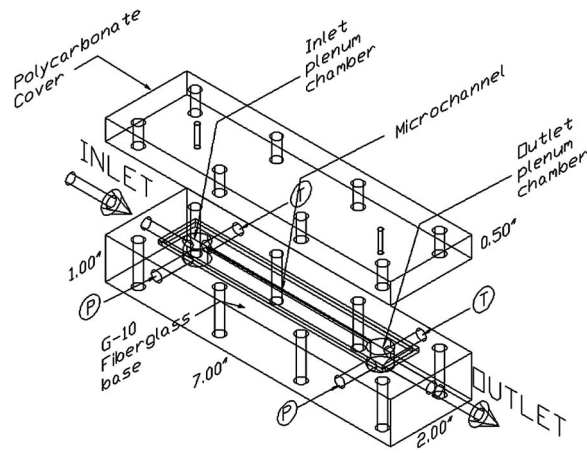
(a)

(b)

Fig. 2 Configuration for Test Module I
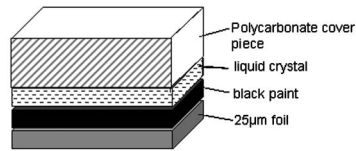
Fig. 4 Configuration for Test Module III—circular tube

ring for sealing, as it introduced additional uncertainty into the height of the channel, as well produced secondary paths for the fluid flow. It is noted that these difficulties were likely not encountered in similar previous works [18,21], due to the larger channels and their use of encapsulated TLCs. The primary conclusion from the above design was that if the TLC coating could be isolated from the flow path it could be an effective measurement tool.

A second test module design was investigated to address the above drawbacks. In this concept, a wide foil was used rather than a narrow foil. The wider foil was easier to handle and minimized local faults produced during its slitting to size. The foil was permanently secured to the base using an epoxy and thus eliminated the need for an O-ring seal. The foil required cooling in the section not exposed to the channel and so a cooling sink of copper blocks with water running in a copper pipe over top was incorporated. The use of a metallic heat sink required electrical isolation from the foil, while minimizing thermal resistance and so an alumina based epoxy that fit this requirement was utilized. A schematic of the channel is shown in Fig. 3.

Measurements utilizing this test setup were found to be high in

comparison to a number of correlations. After consideration of the test setup, it was concluded that these high Nusselt values were due to an inaccurate measure of the heat transfer from the top wall to the fluid. Therefore, although the base material had a low conductivity, heat transfer through the sidewalls produced the additional energy to the fluid. This second design demonstrated the assumption from the previous design that with the TLC isolated from the flow, effective measurements could be produced. However, it also demonstrated that application of accurate boundary conditions was difficult as the size was reduced.

The main difficulties with the previous two concepts were the sealing and the application of a controlled heat load. Many setbacks were observed due to the extensive in-house assembly required, which produced many failed trials. In conceiving the third design, this point was of primary concern. Readily available on the market are relatively thin-walled stainless steel circular tubes, usually produced as hypodermic needles. Tube inner diameters are available as low as 0.254 mm and with wall thicknesses as low as 0.051 mm (Small Parts). Rectangular tubing is also available and can be custom made with inner cross sections as low as 0.510 mm (Microgroup). The use of custom-made tubing is, however, expensive and requires minimum orders not necessary for small-scale experiments. For this reason, a circular tube was selected as the channel configuration for the third design. The recent use of circular steel tubes in heat transfer studies of mini and microchannels has been performed by Lelea et al. [22], and Owhaib and Palm [23] using thermocouples, as well as Hestroni et al. [10], and Hapke et al. [24] using infrared measurement.

A schematic of the circular tube test module used is shown in Fig. 4. A polycarbonate sheet was machined to produce the measurement chambers for pressure and temperature. The tube had a 1.0668 mm (0.42 in.) inner diameter and an outer diameter of 1.27 mm (0.5 in.). It was connected to these chambers using standard 0.0625 mm (1/16 in.) stainless steel compression fittings

Fig. 3 Configuration for Test Module II
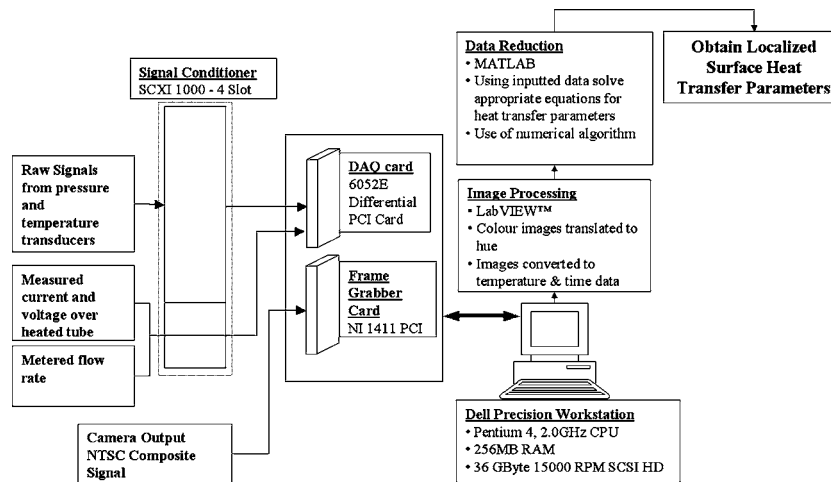
Fig. 5 Schematic of the data acquisition system

with specialty ferrules to accommodate the small diameter tubing. The ferrule material is a composite of graphite and polyimide, which has a high electrical resistance. The advantages of the compression fittings are that they are resealable and can withstand high pressures (~690 kPa rating). All fittings were insulated, and energy losses between measurement location and tube inlet were estimated to be negligible. Copper stranded wire (14AWG), was wrapped around the tube to provide the electrical lead connection. A current was applied through the leads, providing a uniform heat flux from the tube to the fluid. It should be noted that the present setup did not utilize any cover directly in contact with the TLC surface due to the small size of the tube. Attempts were made to utilize optically transparent epoxies to encapsulate the tube, however none of the ones tried were compatible with the TLC material. Rather a noncontacting cover was placed over the tube to protect the TLC coated surface from dust and the room lights. The cover was an opaque plastic half tube which is open at both ends.

2.3 Un-encapsulated Thermochromic Liquid Crystal Application. The aim of the TLC application procedure is to prepare a uniform coating, which will produce vibrant colors. A thicker coating will improve the vibrancy, while too thick of a coating may start to produce nonnegligible temperature gradients through the TLC layer. Additionally, a variably thick layer will introduce uncertainty into the observed temperature field. A relatively uniform coating has traditionally been applied through the use of an airbrush for the microencapsulated TLCs [11]. Application of the un-encapsulated TLCs as documented by Höhman and Stephan [14], Aligoodarz et al. [19] and Kenning et al. [15] has been through use of a paint brush. However, Aligoodarz et al. [19] mentioned significant difficulties in obtaining a continuous layer through this application methodology. According to Hallcrest [12], the liquid crystals (LCs) may be applied in their un-encapsulated form through dilution with an appropriate solvent, such as acetone, after which they are sprayed through an airbrush, resulting in an improved thin uniform layer. This approach, however, is usually not recommended due to the hazardous nature of most solvents. Similar methodologies of applying un-encapsulated LCs with an airbrush have been used for obtaining shear stress measurements. In these cases, the researchers diluted the LCs with petroleum ether [25] or with Freon [26] before applying them with an airbrush. No documentation, however, to the author's knowledge, is available for the application of TLCs in their un-encapsulated form via an airbrush.

Through trial runs of applying the TLCs with a paint brush, the difficulties observed by Aligoodarz et al. [19] were confirmed and therefore this methodology was considered inadequate for producing reliable results. The use of an airbrush was then selected and a variety of trials both of mixture concentration and of application were then performed to produce a sound application methodology. The solvent used was acetone and concentrations by weight of 20:1 (solvent to TLC) based on the suggestions of Hallcrest [12] were incorporated. This approach produced highly improved results in terms of coating uniformity and control of the coating. A Badger Model 100 independent action airbrush was used for application and it allowed for variation of the air to paint concentrations independently during spraying. To improve the color vibrancy of the TLC response, it is common to apply a coat of black paint before the TLC layer. This was incorporated in the current coating application using a water-based black paint and also applying through the airbrush.

2.4 Measurement Apparatus. Two 1.5-mm-diam Type-T (Omega special error limits material) thermocouples were placed in each plenum chamber to measure the bulk fluid temperature (Fig. 5). The output from these and other sensors was monitored through an automated data acquisition system using the LabVIEW™ software. The data acquisition hardware consisted of National Instrument's SCXI 1000 signal conditioning unit with the appropriate modules and the NI 6052E 16 bit, 333 kHz data acquisition card. A schematic of the data acquisition system is shown in Fig. 5. Signals from the transducers were transferred through the signal conditioning unit, then to the computer, whereas the signals from the metering devices were directly transferred through the data acquisition card to the computer. Image acquisition was performed in LabVIEW™ and the images were captured at 640×480 pixel sizes in Red-Green-Blue format. The Hue planes were simultaneously extracted and saved in tagged image file format. Post-processing of the images and acquired data was performed in MATLAB.

A detailed view of the image acquisition apparatus is shown in Fig. 6. Light originating from the illuminator box is directed through a single fiber optic cable. This has the advantage of keeping any heat generated by the light source away from the TLC coated surface. An EKE configured lamp was used with a color temperature rating of 3250 K. Light passes through a polarizer just before entering the zoom lens casing. Within the casing, it is deflected to the test surface by a beam splitter. Depending on the magnification required, the light may pass through an infinity corrected objective for high resolution measurements or else through the default auxiliary lens. Upon reflection from the surface, the light is circularly polarized. It passes into the zoom lens through the analyzer and is directed to the charge coupled device (CCD) camera. The crossed polarizing lens pair is available to obtain improved image quality. The light reflected by the TLC material is circularly polarized and thus travels through the crossed polariz-
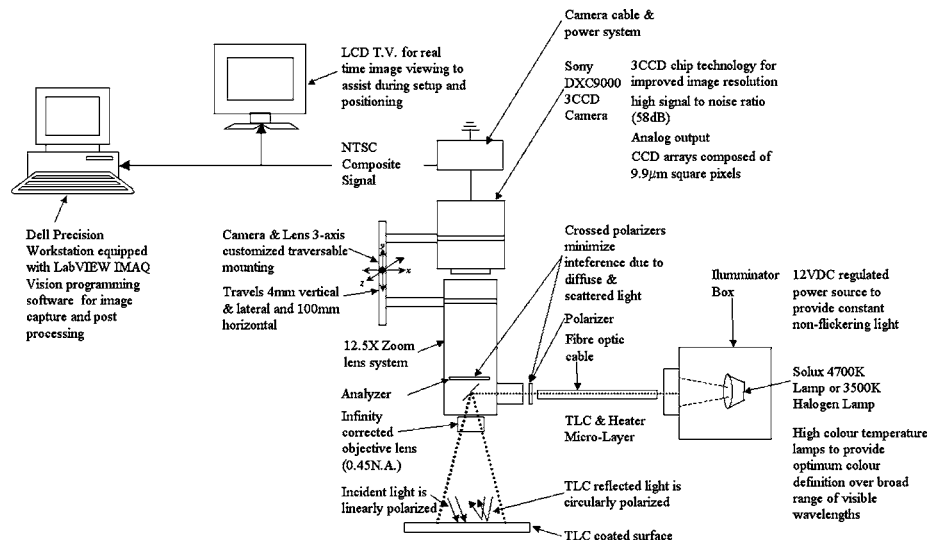
Fig. 6   Schematic of the image acquisition system

ing pair essentially unaffected. The acquired image is then transferred directly to the computer as a national television system(s) committee (NTSC) signal through a single Bayonet Neill-Concelman (BNC) cable. Although the facility has the capability of a crossed polarizing illumination setup, this was not utilized for the circular tube configuration due to the reduced intensity obtained when voltage was applied to the tube. The fiber optic illumination source was instead fixed onto the zoom lens at a fixed angle and directly illuminated the surface (Fig. 4 lighting arrangement).

A liquid crystal display (LCD) television was used for real time monitoring and for positioning. The video signal used in the television loop comes from a separate output line in the camera, and the benefit of this is to minimize the accumulated noise that can occur by transferring the desired video signal through many components. Image acquisition is obtained using a Sony 3-CCD analog camera. The camera was connected to a variable zoom microscopic video lens. The camera lens combination was mounted onto a three-axis traverse through variable length stages each with approximately a 1 μm resolution. This allowed for the length of the channel to be monitored through traversing the entire length. The lateral and vertical axis stages allow for fine tuning of the position and focusing. Heater power was obtained through two power supplies. One was a BK Precision switch mode power supply (Model 1692) with a voltage range of 2.7–15 V and a maximum current rating of 40 A. The second was a Good Will (GW) Instruments power supply (Model GPC-1850) with a voltage range of 0–5 V and a maximum current rating of 20 A. The unencapsulated liquid crystal material used was provided by *LCR*-Hallcrest. The TLC material nominally had a red start of 40 °C with a bandwidth of 10 °C.

3   Calibration

In order to obtain quantitative thermal data from the TLC response, a calibration of the material is required. The requirement of the calibration is to quantitatively relate the observed color to the material temperature. Different methods are available for quantifying the color observed and, in the current work, the hue angle is taken as the color descriptor. The definition of hue in the current work is the same as that recommended by Hay and Hollingsworth [27], which is

$$Hue(H) = \arctan\left(\frac{\sqrt{3}(G-B)}{2R-G-B}\right) \tag{1}$$

The perceived color of a TLC will depend upon the primary and background light spectral characteristics, the camera viewing angle and distance, the primary lighting angle and distance, the light's optical path, and the instance of the TLC application. For a calibration to be fully valid the above must be maintained between the calibration and the experiment. This is usually done through an in situ calibration [28]. Choices for calibration include a linear temperature gradient [29], a liquid bath or circulated liquid [27], or an isothermal block [28]. A linear temperature gradient applied to a metal block is advantageous in providing a quick calibration, however the amount of calibration points is dependent on the number of sensors employed along the gradient. An isothermal bath or circulated bath provides the best means of obtaining a truly isothermal surface, however it does require a liquid supply that is either stirred or circulated. A constant temperature block is usually difficult to work within maintaining the full viewing surface isothermal and hence not recommended.

For the present study, a calibration based on circulating the fluid through the channel was selected due to its ease of incorporation into the setup. The temperatures at the tube inlet and outlet were measured and remained within 0.5 °C of each other. The fluid was slowly heated via the preheater and images captured at incremental changes in temperature. One of the main observations from this work is that the response of the un-encapsulated TLC under increasing magnification will contain varying degrees of noise and scatter. Figure 7 shows the histogram of a color response from a constant temperature region 30×600 pixels. In order to address this in an automated fashion, an intelligent calibration process is introduced.

To start, a region of interest (ROI) needs to be selected to base the calibration curve on. Lighting, viewing angle, and coating uniformity may all influence the calibration between two different regions on the tube. To eliminate such influences a calibration curve was produced for a number of regions in the image. The minimum size for a given ROI is a single pixel, however to account for noisy pixels using statistics, a ROI size greater than one pixel was utilized. Since the camera was traversed, this meant that the ROI would be a particular tube location and that high accuracy in traversing between set locations was pertinent. The ROI should not be too large since local variability's will not be captured. For the present work, the image size was 60×640 pixels, which for the selected ROI size of 4×3 pixels, gives 3200 calibration curves. It was intended to use a fifth order polynomial to fit the calibration data, which is typical in TLC measurements. To manually verify the goodness fit of each curve would be too time con-
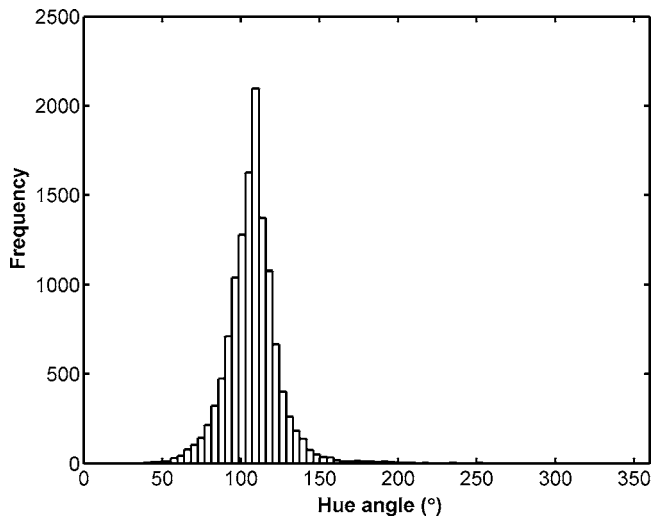
Fig. 7 Constant temperature region hue angle histogram

Fig. 9 Typical calibration curve fits at a fixed image location for a number of tube locations

suming and hence an automated approach that considers extraneous cases was incorporated. These extraneous cases occurred because of the few calibration data points captured and due to scatter.

The number of calibration points was few due to the fact that traversing occurred while the temperature was slowly rising. By the time the camera was returned to a given location, the temperature had risen significantly. If an assumption is made that within the valid calibration range the hue response will increase monotonically with increasing temperature, the main point to verify is that there are no maxima or minima occurring. If a minima or maxima is present, modification of the calibration curve needs to be addressed. It was found that using a third order polynomial will produce a reasonably accurate calibration curve without these minima or maxima locations. An automated calibration curve fitting was utilized with an initial polynomial order of 5, but if any local minima or maxima were present within a predefined region, its order was reduced to 3. Figure 8 shows a calibration curve with data points originally set to a fifth order fit then corrected to a third order fit.

The hue angle is defined on a polar system and, hence, due to scatter, hue values close to 0 deg may be present in the

300–360 deg range and vice versa for hue values close to 360 deg. Compensation similar to that of Hay and Hollingsworth [27] was incorporated whereby the negative angle was used when such scatter occurs. This can be identified, for example, when at high temperatures a low hue value is obtained, or at low temperatures, a high hue value is obtained. The criterion used was that for temperatures greater than 44°C with hue values less than 50 deg, and for temperatures less than 44°C with hue values greater 200 deg, the negative hue angle was used. This adjustment was used to prevent the fitted curve from producing large maxima and minima. Consequently, pixel values analyzed during postprocessing were only considered if they fell within this hue range. The ROI will have a particular scatter in the hue values and for the present study, the nominal hue value was based on the median of the hue values. With the above three conditions incorporated, Fig. 9 shows typical calibration curves for a given image location. Figure 10 shows a typical calibration curve for a given ROI that was considered a good fit to the data. Such good fits were not always possible due to scatter and noise as described above.

Fig. 8 Corrected calibration curve with fifth order and third order fits
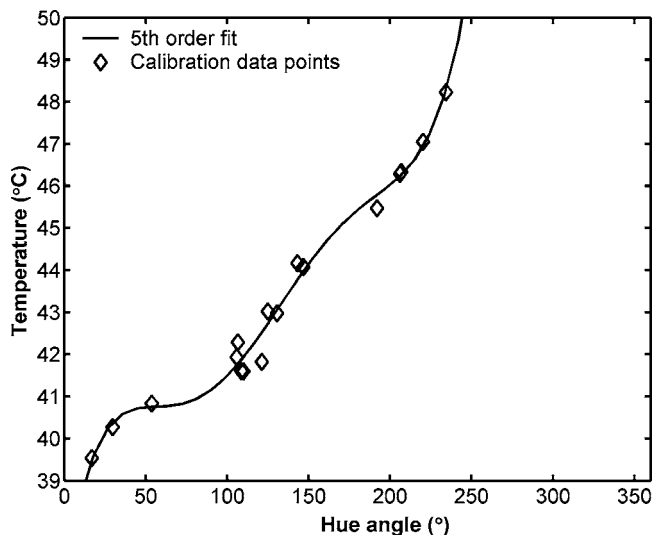
Fig. 10 Typical good-fit calibration curve
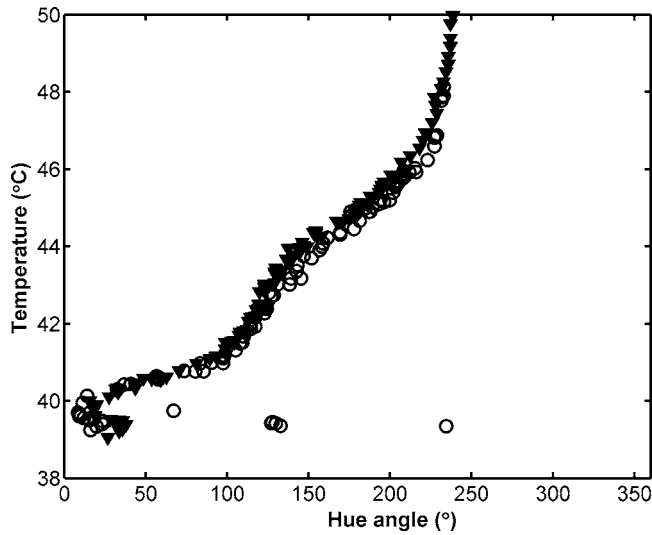
Fig. 11 Calibration curves accumulated over different time periods

4 Test Procedure and Data Reduction

4.1 Procedure. Prior to each experiment, the TLC material was calibrated as described in the previous section. To begin measurements, the flow rate was set by adjusting the bypass valve and the system was then left to run for about 10 min. Measurements were carried out over a flow range of 20–160 ml/min, with inlet temperatures ranging from 28.4 to 31.9°C and outlet temperatures ranging from 32.9 to 38.3°C. The room lights were dimmed and the tube was kept covered. Traversing began at the rear of the tube where the wall temperature would be highest. The configured magnification provided a field of view of 10.233 mm × 7.671 mm, which translated into 15.53 μm/pixel for the CCD sensor size. Although the facilities' magnification capabilities are much larger, a higher magnification would have required more traversing in order to capture the entire tube. This would require more time and provide additional information unnecessary for the purpose of the present work. After running flow through the system, measurements were ready to be taken and the cover was removed. The illumination system was turned on and the voltage was slowly adjusted until the TLC color response was in the predominantly green range (~41–46.0°C). The system was then allowed to come to equilibrium, which spanned 5–10 min, and afterwards a measurement was taken.

Three images were captured at a speed of 30 frames/s. These color images were converted to hue angle, scaled on an integer range from 0 to 255, and then averaged for each pixel location. Simultaneously, the fluid temperature at inlet and outlet, as well as the flow rate measurements, were captured. For each traversed location, 2–4 measurements were captured at a given flow rate though the heater power may be adjusted between image capture. During data reduction, the best images by visual inspection were selected for final presentation. Whenever the voltage was adjusted, the illuminating system was turned off in order to minimize any degradation to the TLC and external heat addition to the tube as the system approached equilibrium.

The TLC calibration may change slightly over time. It was important then to verify that over a measurement run, which could be as long as 3 h, the calibration curve would not shift significantly. To verify this, calibration data at a fixed location were obtained at two different times 3.5 h apart. Between this time period, the tube was left uncovered to simulate the experimental runs. Figure 11 shows the data obtained from these two calibration runs, determined from a ROI size of 4 × 3 pixels. As can be seen in Fig. 11, the change in the calibration curve is insignificant, with

a maximum temperature change of approximately 0.5°C. Hence, the calibration taken prior to the measurements runs can be used over the entire measurement period. Once an experimental run was completed, the tube remained covered.

It is noted from Fig. 11 that a significant number of calibration points were obtained compared to Fig. 10. This is due to the fact, as mentioned earlier, that this particular calibration was performed without traversing and therefore small temperature rises could be captured more easily. Performing a calibration at each location separately for the results presented was not performed since it is too time intensive.

4.2 Data Reduction. The local Nusselt number was calculated as

$$\text{Nu}_x = \frac{h_x D}{k_{\text{loc}}} \tag{2}$$

The local heat transfer coefficient (h_x) was obtained through the following convective heat transfer relation

$$h_x = \frac{q''}{(T_{w,xy} - T_{b,x})} \tag{3}$$

The heat flux was calculated based on the fluid enthalpy change given by

$$q'' = \frac{\dot{m} C p (T_{\text{out}} - T_{\text{in}})}{\pi D L_{\text{heated}}} \tag{4}$$

The local fluid bulk temperature ($T_{b,x}$) was determined from an energy balance and at each streamwise location was

$$T_{b,x} = T_{\text{in}} + \frac{\pi D q''}{\dot{m} C p} \cdot x \tag{5}$$

Properties for the above calculations, except the local Nusselt number, were based on the average fluid temperature between inlet and outlet. The properties were obtained from Kays and Crawford [30] and were interpolated between temperatures. $T_{w,xy}$ was based directly on the TLC measurement. Estimation of the internal wall temperature assuming one-dimensional heat transfer through the tube wall thickness gave a wall temperature gradient less than 0.3°C. This was well within the wall temperature uncertainty and hence was assumed negligible.

4.3 Uncertainty. The uncertainty in the TLC temperature was calculated similar to the methods discussed in Hay and Hollingsworth [27] and in Chin [17]. The wall temperature uncertainty was taken to be

$$\delta T_w = \sqrt{\left(\frac{dT}{dH}\delta H\right)^2 + \delta T_{\text{Fluid}}^2 + (2 \times \text{SEE})^2} \tag{6}$$

where dT/dH is the sensitivity in the temperature hue curve, which was either a fifth or third order polynomial as discussed earlier. The parameter δH represents uncertainty in hue for the ROI and, since the ROIs were small (12 pixels), a constant value of 1 deg was selected instead for the scatter. An additional 2 deg bias was added for the RGB to composite signal conversion uncertainty. In addition, δT_{Fluid} was the uncertainty of the thermocouple probes used to measure the fluid temperature during calibration. Finally, the standard error estimate (SEE) takes into account the error in the polynomial curve fit and is typically defined as

$$\text{SEE} = \left[\sum_{i=1}^{n} [T_i(h) - T_{\text{fit},i}(h)]^2 / (n - j - 1)\right]^{1/2} \tag{7}$$

The calibration yielded only a few points for each tube section and sometimes produced points with large scatter that may significantly alter the SEE unreasonably. A constant value was instead selected for the SEE based on repeated observations of the

Table 1 Typical parameter uncertainty for results obtained in this study

Parameter Uncertainty	
Local Nusselt number, Nu	23.7%
TLC temperature, T_w	1.1°C
Inlet/outlet temperatures, T_{fluid}	0.7°C
Reynolds number, Re	7.78%
Flow rate	7.60%
D xPr Re, x^*	9.60%
Normalized distance, x/D	16.20%

Fig. 12 Sample measured wall temperature and calculated fluid bulk temperature

calibration curves of 0.5 °C. Typical uncertainties for the parameters considered in the results are listed in Table 1.

The uncertainty in the local Nusselt number is high, but this is due to the use of a heat flux derived from the fluid enthalpy change which is a function of T_{fluid}. The use of the fluid enthalpy change rather than the applied power for determining the heat flux was selected for two reasons. First and primarily, due to the fact that the lead connection to the tubing produced a nonnegligible resistance, and so the measured power included some heat dissipation at the leads, which was unknown. The room temperature resistance of the tube measured directly was 0.22 Ω, while the room temperature resistance of the tube measured through the lead wires was approximately 0.3 Ω. This difference was not due to the lead wire lengths as their resistance was measured prior and found to be negligible. The difference was therefore due to the resistance at the lead wires connection to the tube. The heater power dissipated based on Joule heating may be obtained from $P=I^2R$, with P as the power in watts, I as the current in amps, and R as the resistance in ohms. For a given current, the ratio for two different resistances gives also the ratio of power dissipated. Furthermore, for the resistances quoted above at a given current, a power ratio of 73.3% was obtained.

The energy input based on fluid enthalpy rise compared to that based on the measured current and voltage ranged from 74 to 79%, which confirms the discrepancy observed between these two measurands. For this reason, the fluid energy rise was considered a more accurate measure of the energy transfer to the fluid. The second reason for using the fluid energy rise is that it eliminates the need to estimate the energy loss to the environment. If instead the measured power was used, an estimate of typical uncertainty in Nusselt number showed that it decreased from 23.7 to 12.7%. This value is typical of what is quoted in most works. The wall temperature uncertainty determined from the TLCs of 1.1°C is similar to that of other researchers [18].

5 Results

Figure 12 shows a sample of the surface temperature variation observed as well as the corresponding fluid bulk temperature deduced from Eq. (5). This dataset is from a condition whereby the heater flux was kept constant at each traverse. The data are from six separate images, for a range of $40-100$ x/D and is for a Reynolds number of 610. The wall temperature rises with an approximately constant slope from 41 to 44°C, as would be expected for a constant heat flux boundary condition. Such a distinct wall temperature slope was not always observed, particularly for the higher flow rates, whereby the wall temperature rise was much smaller. The wall temperature data is a streamwise average over the visible surface and has been down sampled for visual clarity. Figure 13 shows the streamwise averaged variation of the Nusselt number for the laminar flow cases, which was generated by considering approximately 90 diameters of tube length, which is a combination of nine adjacent tube sections imaged separately. For each tube section, there is a calibration curve for a region size of 4 $\times$ 3 pixels, and each averaged data point consisted of 60 spanwise pixels. The data are compared with the analytical solution for
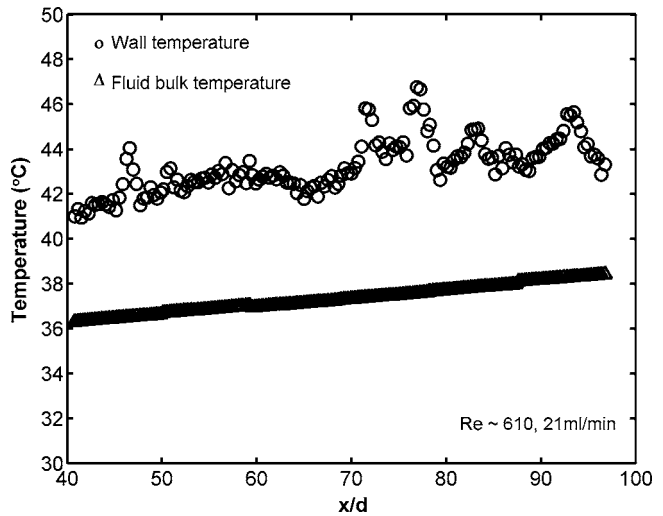
thermally developing circular tube flow with a constant heat flux boundary condition [31]. This solution has been demonstrated to have good agreement with other experimental data for tube flows down to 125 μm [22]. The results presented in Fig. 13 are in good agreement with this solution, showing a decreasing trend towards the fully developed value of 4.364. There is some scatter between the cases, which is partially due to the noisy response of the liquid crystal (Fig. 12). Other discrepancies will arise from the ability to repeatedly traverse to a given section during both calibration and experimentation.

In Fig. 14, the streamwise averaged Nusselt number is plotted for the laminar and turbulent flow rate cases. Each streamwise plot is a combination of nine tube sections imaged separately. For the laminar flow cases, the Nusselt number trend is relatively smooth, showing a slight decrease in the initial length, followed by a constant trend for the remaining length. For the turbulent cases, however, the Nusselt number is slightly oscillatory, particularly farther downstream. These oscillations are not physical phenomena since the Nusselt number should be constant. Rather they are due to noise in the images and local nonuniformities in the coating. It was observed that the application of the voltage had a minor influence on the TLCs and produced some coalescing of the
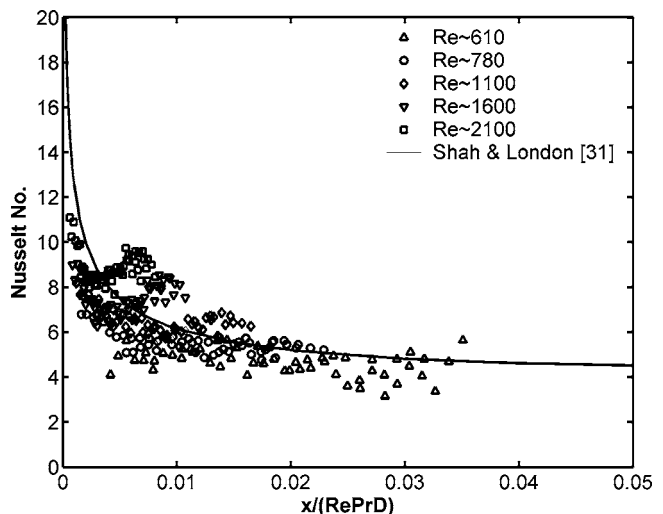
Fig. 13 Streamwise averaged Nusselt numbers for laminar flow cases
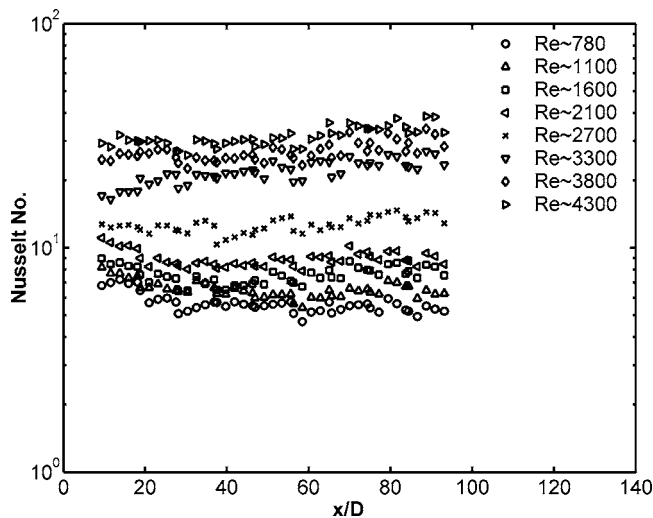
Fig. 14 Streamwise averaged Nusselt numbers for varying flow rates

coating, hence some of the local nonuniformities. Local nonuniformities also arose due to dust particles settling on the coating over time. This was a further reason for the localized calibration methodology developed. Variation between tube sections will be due to minor fluctuations during measurement in parameters such as the flow rate as well the inability to reproduce the original positions during traversing. In addition, between traversing for a fixed flow rate, the heat flux was periodically adjusted as mentioned earlier to have the wall temperature within the TLC's active and reliable range.

The averaged Nusselt number for the laminar and turbulent flow cases is shown in Fig. 15. Two sets of data are presented, from experimental runs on different days. The turbulent results are compared with two heat transfer correlations for fully developed flow in smooth circular tubes. The laminar results were compared with an analytical solution for thermally developing flow with constant heat flux boundary condition in a circular tube. For the turbulent results, one correlation is by Gnielinski [32] for macro-size tubes
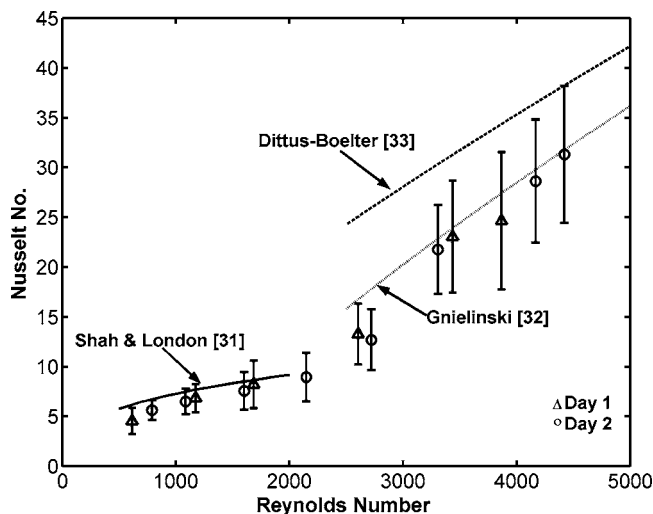
Fig. 15 Variation of the averaged Nusselt number with Reynolds number with comparison to correlations

$$Nu = \frac{(f/8)(Re - 1000)Pr}{1 + 12.7(f/8)^{1/2}(Pr^{2/3} - 1)} \qquad (8)$$

where

$$f = [1.82 \cdot \log(Re) - 1.64]^{-2} \qquad (9)$$

The second is the Dittus–Boelter [33] correlation also for macro-size tubes and given by

$$Nu = 0.024 \cdot Re^{0.8}Pr^{0.4} \qquad (10)$$

For the laminar results, the following solution, which gives the averaged Nusselt number at a streamwise location for laminar flow, was used [31]

$$Nu = 4.364 + \frac{0.046}{(x^*)^{1/3}} \qquad (11)$$

The turbulent results are in very good agreement with the correlation given by Gnielinski, whereas the Dittus–Boelter correlation appears to over predict the heat transfer. Equation (10) is in the functional form for which most experimental correlations of turbulent heat transfer are based [34]. Whereas, Eq. (8) has a similar functional form as the analytically derived relation for turbulent flow heat transfer inside tubes [30]. Both relations, however, have constants derived from empirical methods. The data suggest that the combination of the functional form and the empirically derived constants of Eq. (8), provide a better prediction for turbulent heat transfer in smooth tubes, under the present conditions. Additionally, it is noted that Eq. (8) is the recommended choice for turbulent heat transfer predictions in smooth tubes for the present conditions [30,33]. The figure also shows data from two separate days demonstrating the good repeatability. The laminar results are in very good agreement with the analytical solution, and this was expected since it is derived from the same solution used in the comparison for Fig. 13. Although limited results are presented they do suggest a laminar-turbulent transition in the range of 2000–2500. This is as expected from macrotube measurements and shows that for the presented conditions macrotube considerations are valid.

6 Conclusion and Recommendations

The development of a method utilizing un-encapsulated thermochromic liquid crystals for local heat transfer measurements in a microchannel has been described in the above. To the authors' best knowledge, this is the first use of this form of liquid crystal for heat transfer measurements in a channel this small. The un-encapsulated LC form is difficult to handle and significant scatter and noise may be present in measurements with increased magnification. A localized calibration is utilized to minimize the influence of lighting, viewing, and local variability's. A large number of calibration curves are obtained from this approach and so automation, as well as logical filtering, are necessary. The calibration data generated are demonstrated to be insignificantly altered over the course of an experimental run. The use of an airbrush is found to produce a significant improvement in coating uniformity and thickness control as compared to application with a paintbrush.

The motivation for the methods developed was the need for increased and improved measurement in mini and microchannels. Three channel configurations were discussed and, due to difficulties in the first two arrangements, attributed to the combination of the TLCs and small size, results are presented for only the third configuration, which is a circular tube. The results demonstrate that the conventional size analytical solution for laminar flow is adequate for predicting the local heat transfer in thermally developing flows inside a 1 mm nominally sized circular tube. In addition, the Gnielinski correlation for conventional size channels is adequate for predicting the average heat transfer in a 1 mm nominal size circular tube.

The circular tube design was easily configured from readily available parts and is an effective approach for channel heat transfer studies of diameters down to 250 μm. The Nusselt number uncertainty is higher than most similar work, and this is due to the definition of heat flux based on fluid enthalpy rise rather than generated energy. This definition was selected since it gave a more accurate measure of the heat flux, as opposed to the direct measure, which contained an unknown amount of heat dissipation. Some of the scatter and noise in measurements was introduced due to the need to traverse the test surface. It is noted that during application of the voltage on the circular tube geometry, the TLC material coalesced in some regions altering the surface uniformity and a slight dimming of the response intensity occurred. This was not observed in the other two geometries and was not documented in previous heat transfer experiments utilizing the un-encapsulated TLCs. This observation highlights the need for additional work related to the use of the un-encapsulated form of TLCs for engineering thermometry applications, as they are a promising potential for microthermography applications. A major drawback in the use of the un-encapsulated form of TLCs as utilized in the present work, is the considerable noise in the response. The authors have observed that significant improvement in vibrancy and continuity in spatial response can be obtained using a cover in direct contact with the TLCs. This was not possible for the circular tube, however it will be implemented in future work to be carried out on thermal microsystems Overall, the methods developed and utilized have been shown to provide the local heat transfer coefficient in a microchannel of 1 mm nominal size, and they are a viable approach for geometries of smaller sizes.

Nomenclature

Cp = specific heat, kJ/(kg K)
D = inner tube diameter, m
f = friction factor
h = heat transfer coefficient, W/(m^2 K)
i = index
j = order of curve fit
k = thermal conductivity, W/(m K)
L_h = heated length, m
$\dot{m}$ = mass flow, kg/s
n = number of data points
Nu = Nusselt number
Pr = Prandtl number
q'' = heat flux, W/m^2
Re = Reynolds number, with respect to tube inner diameter
T = Temperature, K
x^* = $x/$RePrD

Subscripts

b = fluid bulk condition
fit = curve fitted parameter
fluid = fluid condition
in = inlet condition
loc = local condition
out = outlet condition
w = wall
x = streamwise location

References

[1] Qu, W., and Mudawar, I., 2003, "Flow Boiling Heat Transfer in Two-Phase Micro-Channel Heat Sinks-I. Experimental Investigation and Assessment of Correlation Methods," Int. J. Heat Mass Transfer, **46**, pp. 2755–2771.
[2] Tuckerman, D. B., and Pease, R. F. W., 1981, "High-Performance Heat Sinking for VLSI," IEEE Electron Device Lett., **EDL-2**(5), pp. 126–129.
[3] Zohar, Y., 1998, "Microchannel Heat Sinks" in M. Gad-el-Hak., *The MEMS Handbook*, McGraw–Hill, New York, pp. 32-1–32-30.
[4] Mehendale, S. S., Jacobi, A. M., and Shah, R. K., 2000, "Fluid Flow and Heat Transfer A Micro- and Meso-Scales With Application to Heat Exchanger Design," ASME Appl. Mech. Rev., **53**(7), pp. 175–194.
[5] Bailey, D. K., Ameel, T. A., Warrington, R. O., and Savoie, T. I., 1995, "Single Phase Forced Convection Heat Transfer in Microgeometries—A Review," *Proceedings of the Intersociety Energy Conversion Engineering Conference*, **2**, pp. 301–310.
[6] Sobhan, C. B., and Garimella, S. V., 2000, "A Comparative Analysis of Studies on Heat Transfer and Fluid Flow in Microchannels," *Proceedings of the International Conference on Heat Transfer and Transport Phenomena in Microscale*, pp. 80–92.
[7] Hassan, I., Phutthavong, P., and Abdelgawad, M., 2004, "Microchannel Heat Sinks: An Overview of the State-of-the-Art," Microscale Thermophys. Eng., **8**, pp. 183–205.
[8] Thome, J. R., 2004, "Boiling in Microchannels: A Review of Experiment and Theory," Int. J. Heat Fluid Flow, **25**, pp. 128–139.
[9] Goldstein, R. J., Chen, P. H., and Chiang, H. D., 1998, "Measurement of Temperature and Heat Transfer." in W. M. Rosenhow, J. P. Hartnett, and Y. I. Cho, *Handbook of Heat Transfer*, 3rd ed., McGraw–Hill, New York, pp. 16.1–16.77.
[10] Hetsroni, G., Mosyak, A., Segal, Z., and Ziskind, G., 2002, "A Uniform Temperature Heat Sink for Cooling of Electronic Devices," Int. J. Heat Mass Transfer, **45**, pp. 3275–3286.
[11] Ireland, P. T., and Jones, T. V., 2000, "Liquid Crystal Measurements of Heat Transfer and Surface Shear Stress," Meas. Sci. Technol., **11**, pp. 969–986.
[12] Hallcrest Product Information, 2000, DS-001–DS-010.
[13] Fergason, J. L., 1968, "Liquid Crystals in Nondestructive Testing," Appl. Opt., **7**(9), pp. 1729–1737.
[14] Höhmann, C., and Stephan, P., 2002, "Microscale Temperature Measurement at an Evaporating Liquid Meniscus," Exp. Therm. Fluid Sci., **26**, pp. 157–162.
[15] Kenning, D. B. R., Konno, T., and Wienecke, M., 2001, "Investigation of Boiling Heat Transfer by Liquid Crystal Thermography," Exp. Therm. Fluid Sci., **25**, pp. 219–229.
[16] Klausner, J. F., Mei, R., and Chen, W. C., 1995, "Liquid Crystal Thermography in Boiling Heat Transfer," Proc. SPIE, **2546**, pp. 152–159.
[17] Chin, Y., 1997, "An Experimental Study on Flow Boiling in a Narrow Channel: From Convective to Nucleate Boiling," Ph.D. dissertation, University of Houston.
[18] Chin, Y., Lakshminarasimhan, M. S., Lu, Q., Hollingsworth, D. K., and Witte, L. C., 2002, "Convective Heat Transfer in Vertical Asymmetrically Heated Narrow Channels," ASME J. Heat Transfer, **124**, pp. 1019–1025.
[19] Aligoodarz, M. R., Yan, Y., and Kenning, D. B. R., 1998, "Wall Temperature and Pressure Variations During Flow Boiling in Narrow Channels," *Heat Transfer 1998, Proceedings of the 11th International Heat Transfer Conference*, **2**, pp. 225–230.
[20] Azar, K., Benson, J. R., and Manno, V. P., 1991, "Liquid Crystal Imaging for Temperature Measurement of Electronic Devices," in *Proceedings—IEEE Semiconductor Thermal and Temperature Measurement Symposium*, pp. 23–33.
[21] Piasecka, M., and Poniewski, M. E., 2004, "Influence of Selected Parameters on Boiling Heat Transfer in Minichannels," *Proceedings of the Second International Conference on Microchannels and Minichannels (ICMM2004)*, pp. 515–521.
[22] Lelea, D., Nishio, S., and Takano, K., 2004, "The Experimental Research on Microtube Heat Transfer and Fluid Flow of Distilled Water," Int. J. Heat Mass Transfer, **47**, pp. 2817–2830.
[23] Owhaib, W., and Palm, B., 2004, "Experimental Investigation of Single-Phase Convective Heat Transfer in Circular Microchannnels," Exp. Therm. Fluid Sci., **28**, pp. 105–110.
[24] Hapke, I., Boye, H., and Schmidt, J., 2000, "Onset of Nucleate Boiling in Minichannels," Int. J. Therm. Sci., **39**, pp. 505–513.
[25] Fujisawa, N., Aoyama, A., and Kosaka, S., 2003, "Measurement of Shear-Stress Distribution Over a Surface by Liquid-Crystal Coating," Meas. Sci. Technol., **14**, pp. 1655–1661.
[26] Reda, D. C., Muratore, J. J., and Heineck, J. T., 1994, "Time and Flow-Direction Responses of Shear-Stress-Liquid Crystal Coatings," AIAA J., **32**, pp. 693–699.
[27] Hay, H. L., and Hollingsworth, D. K., 1996, "A Comparison of Trichromic Systems for Use in the Calibration of Polymer-Dispersed Thermochromic Liquid Crystals," Exp. Therm. Fluid Sci., **12**(1), pp. 1–12.
[28] Anderson, M., 1999, "Thermochromic Liquid Crystal Thermography: Hysterisis, Illumination and Imaging System Effects, Image Processing and Applications," PhD Dissertation, University of California Davis, Mechanical and Aeronautical Engineering Department.
[29] Chan, T. L., Ashforth-Frost, S., and Jambunathan, K., 2001, "Calibrating for Viewing Angle Effect During Heat Transfer Measurements on a Curved Surface," Int. J. Heat Mass Transfer, **44**, pp. 2209–2223.
[30] Kays, W. M., and Crawford, M. E., 1993, *Convective Heat and Mass Transfer*, 3rd ed., McGraw–Hill, New York.
[31] Shah, R. K., and London, A. L., 1978, "Laminar Flow Forced Convection in Ducts," in *Advances in Heat Transfer*, Academic, New York.
[32] Gnielinski, V., 1976, "New Equations for Heat and Mass Transfer in Turbulent Pipe and Channel Flow," Int. Chem. Eng., **16**, pp. 359–368.
[33] Ebadian, M. A., and Dong, Z. F., 1998, "Forced Convection Internal Flow in Ducts," in W. M. Rosenhow, J. P. Hartnett, and Y. I. Cho, *Handbook of Heat Transfer*, 3rd ed., McGraw–Hill, New York.
[34] Holman, J. P., 2002, *Heat Transfer*, 9th ed., McGraw–Hill, New York.

Thermal Transport Due to Phonons in Random Nano-particulate Media in the Multiple and Dependent (Correlated) Elastic Scattering Regime

Ravi Prasher
Intel Corporation,
CH5-157, 5000 W. Chandler Blvd.,
Chandler, AZ 85226
e-mail: ravi.s.prasher@intel.com

Effects of multiple and dependent or correlated elastic scattering of phonons due to nanoparticles on thermal transport in random nano-particulate media (random phononic crystals) are investigated in this paper under various approximations. Multiple scattering means that the scattered wave from one particle is incident on another particle to be scattered again. Dependent scattering means far-field interference of the scattered waves due to phase difference, which is ignored in the independent scattering regime. Multiple and dependent scattering effects become important when the interparticle distance is comparable to the wavelength of phonons. Results show that multiple scattering primarily affects the velocity and density of states of phonons and dependent scattering primarily affects the mean free path of phonons. Effects of both multiple and dependent scattering increases with increasing volume fraction of nanoparticles. Modification of these parameters affects the equilibrium phonon intensity and the thermal conductivity of phonons. [DOI: 10.1115/1.2194036]

Keywords: nanocomposite, phononic band gap material, phononic crystals, correlated scattering, nanoscale heat transfer

1 Introduction

Materials using nano- and microparticles are going to play very important roles in future thermal technologies. Some potential applications of nanoparticles for thermal technologies have already been demonstrated [1–6]. Kim et al. [1] fabricated superlattices made with nanoscale islands of ErAs. In their study the nano-islands of ErAs were approximately 10 nm in size. Soyez et al. [2] fabricated yttria-stabilized zirconia with nanoscale grain size to reduce the thermal conductivity for enhanced insulation properties. Bao et al. [3] used nanoscale germanium/silicon quantum dot structures to enhance the thermoelectric properties of superlattices. In their study the height of the quantum dots was 4 nm. Khitun et al. [4] and Liu et al. [5] used nanoscale spherical quantum dot structures to reduce the phonon thermal conductivity to increase the efficiency of thermoelectrics. Wan et al. [6] have used nanoparticles in superconductors for which efficient thermal transport is very important.

The advent of the nanotechnology means that nanoparticles can be used in various applications at will. Depending on the size, temperature, and the properties of these nano- and microparticles, the physics of thermal transport can fall into various regimes. One such regime is where these particles serve as scatterers of heat carriers. If phonons are the heat carrier in the host medium then this can happen if the size of the particles is very small or the temperature is very low or a combination of both [7]. Increasing the number of scatterers in a medium decreases the thermal con-

ductivity of the medium in the independent scattering regime. This has been the thought behind numerous experimental works [1–5] on decreasing the thermal conductivity of materials.

The general direction of nanoscale heat transfer is to develop predictive models for thermal conductivity [8]. Scattering of phonons with particles has been studied extensively in the past [9]. Almost all the work on thermally excited phonons and their thermal conductivity in particulate media where the particulates participate as scatterers invokes three assumptions regarding elastic scattering with some exceptions [4,5,10]: (1) Scattering is isotropic, (2) scattering is independent, (3) effect of multiple scattering is negligible. Recently the author modified the equation of phonon radiative transport [10,11] to include the effects of anisotropic elastic scattering on phonon transport and showed that phonon scattering is highly anisotropic depending on the acoustic mismatch. Khitun et al. [4] and Liu et al. [5] considered the modification of density of states due to multiple scattering in the independent scattering regime assuming isotropic scattering.

If the particles or the scatterers are far from each other, it is possible to study the scattering by one particle without reference to the others, which means that scattering is *independent*. The term far from each other is with respect to the wavelength of the phonons. It is the ratio of the interparticle distance and the wavelength that decides whether scattering is independent or not. If the ratio of interparticle distance and the wavelength is much larger than 1, then independent scattering is a good assumption. In independent scattering, the phases of the scattered waves from various particles have no systematic relation between them, due to the assumption that the scatterers are distributed completely randomly. Therefore, the net effect is that for all practical purposes the intensities scattered by the various particles can be added without regard to phase [12]. If the volume fraction of the par-

ticles is high then, depending on the size of the particles, the assumption of independent scattering can break down and the effects of multiple and dependent scattering need to be considered because the wavelength of the phonons can become comparable to the interparticle distance. Therefore, dependent scattering is important in the long-wavelength regime where the ratio of the interparticle distance and wavelength is close to 1. It is to be noted that dependent scattering term is widely used in the thermal radiation literature [13–15]. Multiple scattering means that a scattered wave from one particle is incident on another particle to be scattered again. Multiple scattering affects the dispersion relation (the relationship between frequency and wave vector) in the medium (to be shown later). When the particle volume fraction increases, it cannot be assumed that the particles are distributed completely at random because they have finite size. The randomness of the particles also depends on the wavelength of the energy carrier. If the interparticle distance is comparable to the wavelength then assumption of completely random particles is not valid. Since interparticle distance is a function of the volume fraction and the particle diameter, the assumption of complete randomness will depend on the volume fraction, diameter, and the wavelength of the energy carrier. For a given volume fraction, the interparticle distance is smaller for smaller diameter of the particles. Therefore, for a given volume fraction, a system made from nanoparticles is more prone to dependent scattering effects as compared to larger particles. Similarly if the particle size is fixed, then interparticle distance is smaller for higher volume fractions. Therefore, for larger particles dependent scattering effects can be observed at higher volume fractions. Since, depending on the volume fraction and diameter of the particles, interparticle distance can become comparable to the wavelength, there is a systematic relationship between the phase of the scattered waves from various particles [16] and interference between the scattered waves from various particles becomes important. In this regime the scattering is considered to be *dependent* or correlated, i.e., scattering by each particle is not independent of other particles. Dependent scattering means far-field interference of waves scattered by the different particles due to phase difference, which is ignored in the independent scattering regime. One immediate consequence of dependent scattering is that the intensity of the scattered wave from various particles cannot be simply added, which is done in the case of independent scattering. Mostly multiple scattering and dependent scattering have been treated separately in the wave literature [17–19], however in reality both effects are present and in the most general theoretical treatment multiple and dependent scattering are interrelated [16]. Dependent scattering is important only if the distance between the scatterers is comparable to the wavelength, whereas multiple scattering is always present due to the multiple number of scatterers.

The effects of multiple and dependent scattering on thermally excited phonons and their thermal conductivity has largely been neglected in the literature primarily due to the lack of any practical need. As written in the earlier paragraph, however, due to the advent of nanotechnology, it is now easy to make composites with particle sizes ranging from nanoscale to macroscale, with volume fraction ranging from very small to very large. As the volume fraction increases, multiple and dependent scattering effects will become very important. Understanding the impact of multiple and dependent scattering on thermal transport due to phonons will become very important.

It is important to remember, however, that wave literature is full of work done on understanding the effects of multiple and dependent scattering on wave transport [16–20]. A great amount of research has been done in the area of phononic band gap materials (PhBG) [21,22]. In research related to PhBG the composites are periodic and give rise to band gaps if the wavelength of the wave is comparable to the period of the lattice. The focus of research on PhBG has been on calculating the density of states and the effective wave speed. No calculation on PhBG materials has been reported for thermally excited phonons. Similarly, no calculation has been reported on the thermal conductivity of PhBG materials in the literature. For thermal conductivity calculations both wave speed and the effective attenuation of the wave will be important. In this paper the focus is on assessing the impacts of multiple and dependent scattering on thermally excited phonons and its associated thermal conductivity in random nanoparticulate composites, i.e., random phononic crystals.

In this paper we systematically investigate the effects of multiple and dependent scattering on thermal transport due to phonons in nano-particulate media under various approximations for spherical particles. We compute the effects of multiple and dependent elastic scattering by particles on (1) the phase velocity (v) and group velocity (v_g) of phonons, (2) the density of states ($D(\omega)$) of phonons, (3) mean free path of phonons, and (4) equilibrium phonon intensity. Modification of equilibrium phonon intensity and mean free path affects the thermal conductivity of phonons. The Rayleigh scattering regime is investigated in detail, as dependent scattering effects are very important in the Rayleigh scattering regime due to larger values of phonon wavelength.

The equation of phonon transport in particulate media can be written as [10]

$$\mu\frac{\partial I_\omega}{\partial x} = K_u(I_\omega^0 - I_\omega) - \alpha I_\omega + \frac{\alpha}{4\pi}\int_{4\pi}\Phi(\Omega_i \to \Omega)I_\omega(\Omega_i)d\Omega_i \quad (1)$$

where μ is the direction cosine, I the intensity, I^0 the equilibrium intensity, ω the frequency, K_u and α are the attenuations due to inelastic scattering and particle scattering (elastic scattering), respectively, and Φ is the phase function for the scattered wave from solid angle Ω_i to Ω; ω in the subscript indicates that I is a function of ω. The equilibrium intensity in Eq. (1) is given as [10]

$$I^0(T,\omega) = \frac{1}{4\pi}v_g(\omega)\frac{1}{\exp\left(\dfrac{\hbar\omega}{k_bT}\right)-1}\hbar\,\omega D(\omega) \quad (2)$$

where T is the temperature. Equations (1) and (2) can be written for each acoustic phonon mode.

It is to be noted that Eq. (1) is valid for particulate media only if the energy transport inside the particles falls in the wave regime, i.e., the mean free path (m.f.p) of phonons is much larger than the characteristic dimension of the particle. This aspect was discussed in details in the earlier publication [10]. The scattering term due to particles (α) in Eq. (1) is derived solely based on matching the boundary conditions at the particle/medium interface assuming wave behavior. The implicit assumption inside the particle is that energy transport can still be treated like waves. If the particle diameter is much larger than the energy carrier m.f.p, then Eq. (1) is not valid for the particulate media as a whole. In that situation Eq. (1) has to be applied separately in the particle and the medium and the boundary conditions have to be matched at the particle/medium interface. Note that as long as the waves are coherent within the particle, i.e., particle size is much smaller than the phonon m.f.p, Eq. (1) is valid. The question that is still unanswered is when does one consider a particle as scatterer and use Eq. (1) to describe the whole system versus when Eq. (1) has to be applied separately in the medium and the particle. Simkin and Mahan [23] have proposed that if the characteristic dimension is less than the m.f.p of phonons, then energy transport should be treated from wave point of view, however it is still not clear as to what the demarcating criteria is. The m.f.p of phonons in silicon is approximately 250 nm at room temperature as calculated by Chen [24] and Ju and Goodson [25]. The size of the particles in various studies mentioned earlier [1–5] is approximately of the order of 10 nm or less. Therefore, it is reasonable to expect that for these particle sizes Eq. (1) is valid and the energy transport in these particles can be treated on the basis of coherent wave transport. Note that m.f.p of phonons can be as high as 100 μm at very low

temperatures in crystalline materials such as silicon, making Eq. (1) more appropriate for thermal transport in nano-particulate media at low temperatures. M.f.p of phonons in crystalline solids is primarily due to two types of scattering [11]. At very low temperatures boundary scattering dominates and at somewhat lower temperatures (typically around the thermal conductivity maxima) elastic scattering due to the presence of impurities dominates and at high temperatures three phonon Umklapp inelastic scattering dominates. The nature of transport in the particle will depend on the absolute value of the m.f.p [7,23]. Therefore, as long as m.f.p of the phonons is much larger than the particle size, irrespective of the nature of scattering in the host medium, Eq. (1) is applicable. However, since elastic scattering dominates over Umklapp scattering at low temperatures, effects of multiple and dependent elastic scattering are also expected to be more important at low temperatures. The question which this paper does not address is whether Umklapp scattering will have any dependent scattering effects.

If the medium is made of an amorphous material, then treating the energy transport inside the particles as wave transport is not valid in most cases because the m.f.p in amorphous materials is of the order of a few nanometers [26]. Therefore, dependent and multiple scattering effects will not be important for amorphous materials.

As already discussed, there are two fundamental length scales for phonon transport in particulate media: (1) m.f.p of phonons, (2) wavelength of phonons. M.f.p of phonons decides whether the energy transport in the particles in a medium falls in the wave regime or the particle regime. If the energy transport in the particles falls in the wave regime then the particles' server as the scatterer of phonons and Eq. (1) is applied over the whole system. Therefore, the ratio of m.f.p of phonons and the particle size decides whether the energy transport falls in the wave regime in the particles. If the energy transport falls in the wave regime in the particles, then the ratio of the wavelength of phonons and the interparticle distance decides whether the scattering is dependent or not. Therefore, the analysis presented in this paper is valid when the energy transport in the particles fall in the wave regime.

Under the diffusion approximation it can be shown that the thermal conductivity (Λ) can be written from Eq. (1) as [10]

$$\Lambda = \frac{4\pi}{3} \int_0^{\omega_d} \frac{1}{K_u + \alpha(1 - \langle\Phi\rangle)} \frac{dI_\omega^0}{dT} d\omega \qquad (3)$$

where $\langle\Phi\rangle$ is the asymmetry factor [10] due to anisotropic nature of scattering and ω_d is the maximum allowable frequency, which in the case of linear dispersion is the same as the Debye frequency. Computation of $\langle\Phi\rangle$ and α need the formulation of the scattering function ($F(\theta,\phi)$) where θ is the polar angle and ϕ is the azimuthal angle. $F(\theta,\phi)$ gives the angular variation of the intensity of the scattered wave.

2 Elastic Scattering From a Single Scatterer

The scalar acoustic wave equation is used in this paper instead of the vector elastic wave equation [27] to reduce the mathematical complexity and to focus on the physics of multiple and dependent scattering. The basic physics and equations for multiple and dependent scattering will remain the same for both scalar and vector wave equation. It is also to be noted that almost all the work done on phonon scattering in literature has used the scalar wave equation [5,6,9].

The scalar wave equation is

$$\nabla^2 P = \frac{1}{v^2} \frac{\partial^2 P}{\partial t^2} \qquad (4)$$

where P is the stress field and t is the time. Assuming sinusoidal wave form $P = p e^{-i\omega t}$ where ω is the frequency of the acoustic wave, Eq. (4) reduces to

$$(\nabla^2 + k^2)p = 0 \qquad (5)$$

where $k = \omega/v$ is the wave vector. The medium is denoted as 1 and the scatterer is denoted as 2. The scattered wave is $P_s = e^{-i\omega x}$. It can be shown that in the far field ($r \to \infty$) that the scattered field is [28]

$$p_s = -\frac{\sum_{n=0}^{\infty} iA_n(2n+1)P_n(\cos\theta)e^{ik_1 r}}{k_1 r} \qquad (6)$$

where $P_n(\cos\theta)$ is the Legendre function and A_n is the scattering coefficient. A_n is obtained by applying the boundary condition of stress and velocity continuity. In evaluating A_n the amplitude of the incident wave is assumed to be 1. The intensity of the wave is obtained by averaging $P_s P_s^*$ (* denotes the complex conjugate) over one period which for a since wave is proportional to the square of the amplitude [29] divided by 2. Therefore, the intensity of the scattered wave (I_s) is given by

$$I_s = \frac{p_s p_s^*}{2} = \frac{1}{k_1^2 r^2} \frac{\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A_m \bar{A}_n (2m+1)(2n+1) P_n(\cos\theta)P_m(\cos\theta)}{2} \qquad (7)$$

Since the incident wave has unit amplitude, the intensity of the incident wave (I_i) is 1/2. The scattering function is given by [12] $F(\theta) = k_1^2 r^2 (I_s/I_i)$. Therefore, for acoustic waves

$$F(\theta,\phi) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A_m \bar{A}_n (2m+1)(2n+1) P_n(\cos\theta)P_m(\cos\theta) \qquad (8)$$

and [28]

$$A_n = -\frac{j_n'(k_2 a)j_n(k_1 a) - ghj_n(k_2 a)j_n'(k_1 a)}{j_n'(k_2 a)h_n(k_1 a) - ghj_n(k_2 a)h_n'(k_1 a)} \qquad (9)$$

and k_1 is the wave vector in medium 1, k_2 the wave vector in medium 2 (scatterer), j_n the spherical Bessel function, h_n the spherical Henkel function of first kind, $g = \rho_2/\rho_1$ where ρ is the density, $h = v_2/v_1$, and a is the radius of the scatterer. The primes in Eq. (9) represent differentiation with respect to the variables in the bracket.

The scattering cross section (C) of a single scatterer is given as

$$C = \int_0^\pi \int_0^{2\pi} \frac{F(\theta,\phi)}{k_1^2} \sin\theta d\theta d\phi = \frac{4\pi}{k_1^2} \sum_{n=0}^{\infty} (2n+1)A_n\bar{A}_n \quad (10)$$

The phase function Φ is given as [10]

$$\Phi(\theta,\phi) = \frac{4\pi}{k_1^2} \frac{F(\theta,\phi)}{C} = 4\pi \frac{F(\theta,\phi)}{\int_0^\pi \int_0^{2\pi} F(\theta,\phi)\sin\theta d\theta d\phi} \qquad (11)$$

In addition, the asymmetry factor $\langle\Phi\rangle$ is given as

$$\langle\Phi\rangle = \frac{1}{4\pi} \int_0^\pi \int_0^{2\pi} \Phi(\theta,\phi)\cos\theta \sin\theta d\theta d\phi$$

$$= 2\frac{\mathrm{Re}\sum_{n=0}^{\infty} (n+1)A_n\bar{A}_{n+1}}{\sum_{n=0}^{\infty} (2n+1)A_n\bar{A}_n} \qquad (12)$$

For Rayleigh scattering (long wavelength limit) only the first two terms are important in A_n. They are given as

$$A_0 = i\frac{R_1}{3}x_1^3, \quad A_1 = i\frac{R_2}{3}x_1^3 \qquad (13)$$

where $x_1 = k_1 a = 2\pi a/\lambda_1$ (λ = wavelength) is the size parameter, $R_1 = 1/(gh^2) - 1$ and $R_2 = (g-1)/(2g+1)$. From Eq. (11) ϕ for Rayleigh scattering is

$$\Phi = \frac{(R_1 + 3R_2\cos\theta)^2}{R_1^2 + 3R_2^2} \qquad (14)$$

Scattering due to the presence of multiple, scatterers is investigated in the next four sections under various approximations and regimes: Independent, dependent, multiple, and multiple-dependent scattering.

3 Independent Scattering Approximation (ISA)

The total field, P_s^{tot}, due to scattering from N number of scatterers can be written as

$$P_s^{\text{tot}} = \sum_{j=1}^{N} p_s \cos(\omega t - \psi_j) \qquad (15)$$

where the amplitude p_s (Eq. (6)) of the scattered wave is assumed the same for each scatter (only particles of same size considered in this paper) and Ψ_j denotes the phase of the scattered waves from each scatterer. In the independent scattering, the scatterers are distributed completely at random as mentioned earlier. Therefore, variation in Ψ_j is also completely random. Total intensity (I_s^{tot}) of the scattered wave from N scatterers can be computed by averaging $P_s^{\text{tot}} P_s^{\text{tot}*}$ over one time period. Since the Ψ_j is random it can be shown by taking an average of I_s^{tot} over all possible values [29] of Ψ_j that $I_s^{\text{tot}} = I_{\text{ind}} = N p_s p_s^*/2 = NI_s$, i.e., in the independent scattering regime the intensity of the scattered wave from all the particles is added as mentioned earlier. Similarly, the effective attenuation is

$$\alpha_{\text{ind}} = z_a C = \frac{3\phi}{4\pi a^3}C \qquad (16)$$

where z_a is the number of scatterers per unit volume and ϕ is the volume fraction of the scatterers. For Rayleigh scattering case, dimensionless attenuation is (Eqs. (10), (13), and (16))

$$\frac{\alpha_{\text{ind}}}{(x_1^4/a)} = \frac{R_1^2 + 3R_2^2}{3}\phi \qquad (17)$$

In addition, it can be shown that

$$\langle\Phi\rangle = 2\frac{R_1 R_2}{R_1^2 + 3R_2^2} \qquad (18)$$

4 Dependent Scattering Approximation (DSA)

Radiation thermal transport communities [14,15] have extensively dealt with the dependent scattering phenomenon. The same learning can be applied for thermal transport due to phonons, however there are a few subtle differences between photons and phonons. The dominant wavelength of thermally excited photons by applying Wien's displacement law at room temperature is in the range of micrometers [30]. Even in the visible range the wavelength of the photons is in range of 0.4–0.7 μm [30], whereas the dominant wavelength of thermally excited phonons is in the range of a few nanometers at room temperature [31]. The dominant wavelength of phonons is in the range of 50–100 nm in crystalline solids at very low temperatures [31]. Therefore, interparticle distance also has to be of the order of a few nanometers to observe dependent scattering effects in phonons. Since the interparticle distance depends on the volume fraction and the particle size, it will be in the nanometer regime for nanoparticles for nominal volume fractions. Therefore, to observe the effects of dependent scattering in phonons, nano-sized particles are necessary, whereas
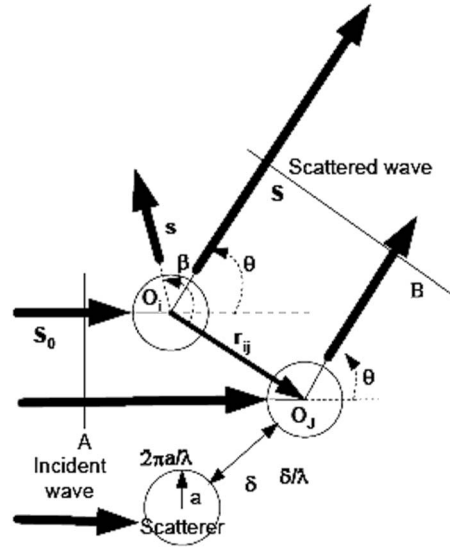
Fig. 1 Definition of vectors and angles for dependent scattering

for photons micron-sized particles can give rise to dependent scattering phenomenon as evidenced by literature on photon transport in the dependent scattering regime [32]. This aspect is discussed in details in Sec. 7.

There are two pertinent length scales involved in the scattering of plane waves as shown in Fig. 1. First length scale is the size parameter x_1 and second length scale is δ/λ where δ is the interparticle distance and λ is the wavelength of phonons. The value of δ/λ decides if the scattering is dependent or independent as discussed before.

The total field due to scattering from N number of scatterers is (Eq. (15) rewritten)

$$P_s^{\text{tot}} = \sum_{j=1}^{N} p_s \cos(\omega t - \psi_j) \qquad (19)$$

Therefore I_s^{tot} is

$$I_s^{\text{tot}} = \frac{\int_0^{2\pi/\omega} P_s^{\text{tot}} P_s^{\text{tot}*} dt}{\int_0^{2\pi/\omega} dt} = \omega/(2\pi) p_s p_s^* \int_0^{2\pi/\omega} \left(\sum_{j=1}^{N} \cos(\omega t - \psi_j)\right)^2 dt \qquad (20)$$

It can be shown that the above equation reduces to

$$I_s^{\text{tot}} = \frac{p_s p_s^*}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \cos(\psi_i - \psi_j) \qquad (21)$$

In ISA, Ψ_i and Ψ_j are assumed to be random, however in DSA no such assumption will be made and interference effects will be captured. The phase difference $(\Psi_i - \Psi_j)$ can be shown to be [33,34]

$$\psi_i - \psi_j = 2\pi \mathbf{r}_{ij} \cdot \mathbf{s} \qquad (22)$$

where $\mathbf{r}_{ij}$ is the vector connecting O_i and O_j as shown in Fig. 1. The term $\mathbf{s}$ is the difference between the unit incident wave vector $\mathbf{S_0}$ and the unit scattered wave vector $\mathbf{S}$ in the direction θ divided by the wavelength λ. Thus $\mathbf{s} = (\mathbf{S} - \mathbf{S_0})/\lambda$ and [15,33] $s = |\mathbf{s}| = 2\sin(\theta/2)/\lambda$. The phase difference arises due to different lengths traveled by the incident and scattered waves for different particles at any location. This is shown schematically in Fig. 1

between planes A and B. Equation (22) can be derived by considering the path traveled by the incident and the scattered wave for different scatterers. Therefore

$$I_s^{\text{tot}} = \frac{p_s p_s^*}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \cos[2\pi s r_{ij} \cos\beta] \quad (23)$$

β is the angle between r_{ij} and s as shown in Fig. (1). As the vector r_{ij} takes on all positions in the space, β takes all values at random, and the effect of this random orientation can be obtained by averaging Eq. (23) over all solid angles. Thus

$$\int_0^4 \pi \cos[2\pi s r_{ij} \cos\beta] d\Omega/(4\pi)$$

$$= 1/2 \int_0^\pi \cos[2\pi s r_{ij} \cos\beta] \sin\beta d\beta$$

$$= \sin(2\pi s r_{ij})/(2\pi s r_{ij}) \quad (24)$$

Therefore

$$I_s^{\text{tot}} = \frac{p_s p_s^*}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\sin(2\pi s r_{ij})}{2\pi s r_{ij}} \quad (25)$$

In performing the above summation, it is necessary to consider one scatterer and to carry out the summation over all distances and to all scatterers of the array including the origin scatterer and then to allow the origin scatterer to be every scatterer in the array in turn. Summations for any scatterer with respect to itself lead simply to unity in each case. Therefore, Eq. (25) reduces to

$$I_s^{\text{tot}} = \frac{p_s p_s^*}{2} \left(N + \sum_{i=1, i\neq j}^{N} \sum_{j=1}^{N} \frac{\sin(2\pi s r_{ij})}{2\pi s r_{ij}} \right) \quad (26)$$

where the summation now excludes the origin scatterer. Equation (26) can be written as

$$I_s^{\text{tot}} = I_s \left(N + \sum_{i=1, i\neq j}^{N} \sum_{j=1}^{N} \frac{\sin(2\pi s r_{ij})}{2\pi s r_{ij}} \right) \quad (27)$$

where I_s is the scattered intensity due to single scatterer (Eq. (7)). Now various parameters for dependent scattering can be easily defined. Equation (27) can be written as

$$I_{\text{dep}} = N I_s \times \left(1 + \frac{1}{N} \sum_{i=1, i\neq j}^{N} \sum_{j=1}^{N} \frac{\sin(2\pi s r_{ij})}{2\pi s r_{ij}} \right) = I_{\text{ind}} H(\theta) \quad (28)$$

where $H(\theta)$ is the form factor due to dependent scattering and is given as

$$H(\theta) = \left(1 + \frac{1}{N} \sum_{i=1, j\neq 1}^{N} \sum_{j=1}^{N} \frac{\sin(2\pi s r_{ij})}{2\pi s r_{ij}} \right) \quad (29)$$

Using Eq. (28) and definition of various parameters, it can be shown that the dependent scattering function is

$$F(\theta)_{\text{dep}} = H(\theta) F(\theta)_{\text{ind}} \quad (30)$$

where $F(\theta)_{\text{ind}}$ is given by Eq. (8). Once $F(\theta)$ is known, all the relevant parameters can be computed and are given as

$$\frac{\alpha_{\text{dep}}}{\alpha_{\text{ind}}} = \frac{1}{4\pi} \int_0^\pi \int_0^{2\pi} H(\theta) \Phi(\theta)_{\text{ind}} \sin(\theta) d\theta d\phi \quad (31)$$

$$\Phi_{\text{dep}} = \frac{\alpha_{\text{ind}}}{\alpha_{\text{dep}}} H(\theta) \Phi_{\text{ind}} \quad (32)$$

$$\langle\Phi\rangle_{\text{dep}} = \frac{\int_0^\pi \int_0^{2\pi} H(\theta)\Phi(\theta)\cos(\theta)\sin(\theta)d\theta d\phi}{\int_0^\pi \int_0^{2\pi} H(\theta)\Phi(\theta)\sin(\theta)d\theta d\phi} \quad (33)$$

For independent scattering $H(\theta)=1$ and Eqs. (31)–(33) reduce to the independent scattering results. Under the assumption of a continuous and infinite distribution of scatterers with number density distribution of $z(r)$, the double summation in Eq. (29) can be replaced by an integral with respect to r [15,34]. Therefore

$$H(\theta) = 1 + \int_0^\infty 4\pi r^2 z(r) \frac{\sin(2\pi sr)}{2\pi sr} dr$$

$$= 1 + \int_0^\infty 4\pi r^2 (z(r) - z_a) \frac{\sin(2\pi sr)}{2\pi sr} dr$$

$$+ \int_0^\infty 4\pi r^2 z_a \frac{\sin(2\pi sr)}{2\pi sr} dr \quad (34)$$

where z_a is the average number density (number of scatterers per unit volume). It can be shown that second integral in Eq. (34) may be taken as zero [34]. Therefore

$$H(\theta) = 1 + \int_0^\infty 4\pi r^2 (z(r) - z_a) \frac{\sin(2\pi sr)}{2\pi sr} dr \quad (35)$$

The term $z(r)$ is also called the pair correlation function; $z(r)$ means that the local number density of the scatterers can be higher or lower than the average number density. If there is no correlation between the scatterers then $z(r)=z_a$ and $H(\theta)=1$ which is the independent scattering result. This function represents the likelihood of finding the center of neighboring particle at some distance r from the central particle. Drolen and Tien [14] have discussed various pair distribution functions. Among various pair distribution functions Percus-Yevick (PY) [14,35] integral equation with appropriate particle-to-particle interactions is the best and can be used for larger volume fractions. The PY model has been very successful in modeling the attenuation of optical radiation in packed beds [14,36]. Other pair distribution functions fail at volume fractions higher than 0.1. For a packed-sphere system the hard sphere potential applies. Unfortunately for the PY model, analytical results can only be obtained for very small values of x_1, i.e., in the Rayleigh regime. However, this is fortunate in the sense that dependent scattering is more dominant in the Rayleigh regime due to large wavelengths of the phonon. In the Rayleigh regime $\sin(2\pi sr)/2\pi sr \rightarrow 1$ because $s \rightarrow 0$ due to long wavelengths. Therefore $H(\theta)$ is independent of θ and it can be shown that for PY distribution function [35]

$$H(\theta) = \frac{(1-\phi)^4}{(1+2\phi)^2} x_1 \rightarrow 0 \quad (36)$$

Substituting Eq. (36) in Eqs. (31) and (33) for Rayleigh scattering

$$\frac{\alpha_{\text{dep}}}{(x_1^4/a)} = \frac{\alpha_{\text{ind}}}{(x_1^4/a)} \frac{(1-\phi)^4}{(1+2\phi)^2} \quad (37)$$

and $\langle\Phi\rangle_{\text{dep}} = \langle\Phi\rangle_{\text{ind}}$.

5 Multiple Scattering: Effective Field Approximation (EFA)

In the previous section one critical phenomenon was ignored, i.e., the effects of multiple scattering. The effect of multiple scattering on the scattering cross section and wave velocity was ignored in thermal radiation literature [14,15]. Thermal radiation literature only dealt with the effect of dependent scattering [14,15]
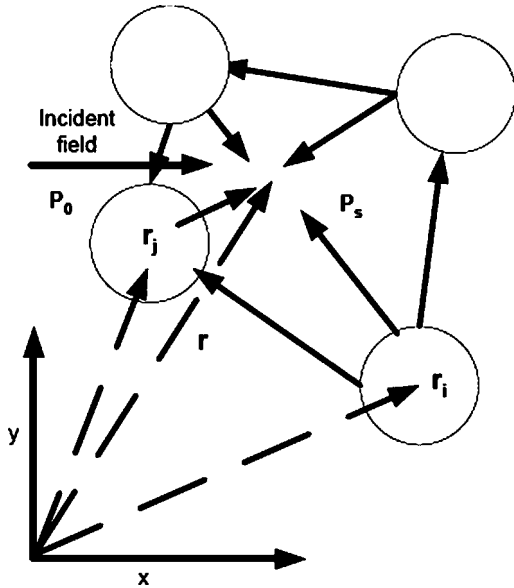
Fig. 2 Schematic to show the multiple scattering of waves

on scattering cross section. Multiple scattering effects were incorporated into thermal radiation literature in the equation of radiative transport only in terms of the in-scattering term involving the phase function, however effects of multiple scattering on other properties were ignored, as mentioned earlier. The author [37] recently considered the effect of multiple scattering on the velocity of photons and showed that it should be considered for dense particulate media. Multiple scattering means that a scattered wave from one particle is incident on another particle to be scattered again. In a dense particulate system the effective field within the medium is modified because the effective field is the sum of the incident field and the scattered field from various scatterers. It is to be remembered that dependent scattering is important only when the distance between the particles is of the order of wavelength, i.e., it depends on the closeness of the particles, whereas multiple scattering is present just because there are multiple number of particles.

Due to multiple scattering from various particles, the effective field in the medium will be different from the original field because the effective field will be the sum of the incident field and the scattered field. Multiple scattering phenomenon is schematically shown in Fig. 2. This problem was first solved by Foldy [17]. The basics of the multiple scattering approach are given briefly in this section. The basic equations (Eq. (38)–(47)) related to multiple scattering have been taken from Foldy [17] and the only reason these equations have been reproduced in the paper is to highlight the key assumptions and physics. They are also discussed because in the next section on the most general case of multiple and dependent scattering some of the following equations will be needed. One of the most important assumptions in the effective field approximation is that the scattering is independent. The total field (P) inside the medium (outside any scatterer) is given

$$P(\boldsymbol{r}) = P_0(\boldsymbol{r}) + \sum_{j=1}^{N} P_s(\boldsymbol{r},\boldsymbol{r}_j) \tag{38}$$

where P_0 is the incident field and $P_s(\boldsymbol{r},\boldsymbol{r}_j)$ is the field scattered by the jth particle to the point of observation. The exciting field ($P_e(\boldsymbol{r}_j)$) incident on scatterer j is the resultant of the incident field P_0 and the scattered field from all other scatterers so that

$$P_e(\boldsymbol{r}_j) = P_0(\boldsymbol{r}) + \sum_{i \neq j}^{N} P_s(\boldsymbol{r}_i,\boldsymbol{r}_j) \tag{39}$$

In the case of single scattering the scattered wave is related to the incident wave by Eq. (6) and has spherical wave form. Note that in deriving Eq. (6) the amplitude of the incident wave was assumed to be 1. If the amplitude of the incident wave is, for example, p^i, then p^i will appear as a multiplier on the right-hand-side of Eq. (6). Therefore Eq. (6) shows that the strength of the scattered wave is proportional to the incident wave. Similar approach is taken for the multiple scattering problems [17]. Therefore, scattered field is given by

$$P_s(\boldsymbol{r},\boldsymbol{r}_j) = \frac{A_j \exp\lfloor ik_1|\boldsymbol{r} - \boldsymbol{r}_j|\rfloor}{k_1|\boldsymbol{r} - \boldsymbol{r}_j|} = A_j E(\boldsymbol{r},\boldsymbol{r}_j) \tag{40}$$

In Eq. (40) A_j is written as $A_j = g_j(\boldsymbol{r}_j,\omega)P_e(\boldsymbol{r}_j)$ thereby making the strength of the scattered wave from a scatterer proportional to the external field acting on it. Substituting Eq. (40) in Eq. (38) and (39)

$$P(\boldsymbol{r}) = P_0(\boldsymbol{r}) + \sum_{j=1}^{N} g_j P_e(\boldsymbol{r}_j)E(\boldsymbol{r},\boldsymbol{r}_j) \tag{41}$$

$$P_e(\boldsymbol{r}_j) = P_0(\boldsymbol{r}) + \sum_{i \neq j}^{N} g_i P_e(\boldsymbol{r}_i)E(\boldsymbol{r}_i,\boldsymbol{r}_j) \tag{42}$$

The direct method of solving the multiple scattering problem would consist of solving the set of simultaneous Eqs. (42) for $P_e(\boldsymbol{r}_j)$ and substituting these in Eq. (41), thus giving the effective field $P(\boldsymbol{r})$ in the medium as function of the positions and scattering parameters of the scatterers. Taking the configurational average of this quantity would give the desired results, however this would lead to great mathematical complexity. Therefore, for multiple scattering system the method is to find equations satisfied by the average of $P(\boldsymbol{r})$. Taking the configurational average of both sides of the Eq. (41) we get

$$\langle P(\boldsymbol{r}) \rangle = P_0(\boldsymbol{r}) + z_a \int_V g_j(\boldsymbol{r}_j)\langle P_e(\boldsymbol{r}_j) \rangle E(\boldsymbol{r},\boldsymbol{r}_j)d\boldsymbol{r}_j \tag{43}$$

The quantity $\langle P_e(\boldsymbol{r}_j) \rangle$ represents the external field acting on the jth scatterer averaged over all possible configurations of all of the scatterers. The only rigorous way of evaluating $\langle P_e(\boldsymbol{r}_j) \rangle$ is to solve the simultaneous Eqs. (42) and then carry out the necessary integrations to obtain the above configurational average. This again is not feasible. Therefore, some sort of approximation is made. The approximation is to replace the external field acting on the jth scatterer averaged over all configurations of the scatterer by the average field which would exist at the position of the jth scatterer when the scatterer is not present which means replacing $(\langle P_e(\boldsymbol{r}_j) \rangle)$ in Eq. (43) by $\langle P(\boldsymbol{r}) \rangle$. Therefore, the basic equation for multiple scattering is

$$\langle P(\boldsymbol{r}) \rangle = P_0(\boldsymbol{r}) + z_a \int_V g(\boldsymbol{r}')\langle P(\boldsymbol{r}) \rangle E(\boldsymbol{r},\boldsymbol{r}')d\boldsymbol{r}' \tag{44}$$

It can be shown [16] that g is same as the forward scattering amplitude ($f(0)$) of the scattered wave from a single scatterer. Therefore Eq. (44) can be written as

$$\langle P(\boldsymbol{r}) \rangle = P_0(\boldsymbol{r}) + z_a f(0) \int_V \langle P(\boldsymbol{r}) \rangle E(\boldsymbol{r},\boldsymbol{r}')d\boldsymbol{r}' \tag{45}$$

Now applying the operator, $\nabla^2 + k_1^2$ on both sides of Eq. (45) and using the properties of Dirac-Delta function it can be shown

$$(\nabla^2 + k_1^2 + 4\pi z_a f(0))\langle P(\mathbf{r})\rangle = 0 \tag{46}$$

Equation (46) means that the effective wave vector (k) in multiple scattering particulate media is given as

$$k^2 = k_1^2 + 4\pi z_a f(0) \tag{47}$$

where k is the effective wave vector of the particulate media. Note that Eq. (47) is equally applicable for electromagnetic waves, elastic waves, acoustic waves, scalar waves or vector waves. For different types of waves $f(0)$ will be different. In this paper, since we are considering the scattering of scalar acoustic waves, $f(0)$ can be obtained from Eq. (6); $f(0)$ is from Eq. (6) and is given as

$$f(0) = -\frac{i}{k_1}\sum_{n=0}^{\infty} A_n(2n+1) \tag{48}$$

Separating A_n into real and imaginary part as $A_n = A_{Rn} + iA_{In}$ and substituting in Eq. (47) leads to

$$k^2 = k_1^2 + \frac{4\pi z_a}{k_1}\sum_{n=0}^{\infty} A_{In}(2n+1) - \frac{i4\pi z_a}{k_1}\sum_{n=0}^{\infty} A_{Rn}(2n+1) \tag{49}$$

k can be separated into real and imaginary part as $k = k_R + ik_I$; k_R gives the effective phase velocity of the medium and k_I is related to α [38] by $\alpha = 2k_I$.

Noting that [39]

$$C = -\frac{4\pi}{k_1}\sum_{n=0}^{\infty} A_{Rn}(2n+1) \tag{50}$$

Eq. (49) can be written as

$$\frac{k^2}{k_1^2} = 1 + \frac{4\pi z_a}{k_1^3}\sum_{n=0}^{\infty} A_{In}(2n+1) + \frac{iz_a C}{k_1} \tag{51}$$

It can be shown that Eq. (51) gives the independent scattering results in the dilute scatterer limit by making a binomial expansion of Eq. (51)

$$\frac{k}{k_1} = 1 + \frac{2\pi z_a}{k_1^3}\sum_{n=0}^{\infty} A_{In}(2n+1) + \frac{iz_a C}{2k_1} \tag{52}$$

Equation (52) shows that $\alpha = 2k_I = z_a C_{\text{sca}}$, which is the independent scattering result. However, Eq. (52) also shows that even in the independent scattering regime v is modified as

$$v = \frac{v_1}{1 + \frac{2\pi z_a}{k_1^3}\sum_{n=0}^{\infty} A_{In}(2n+1)} \tag{53}$$

The reason the effective phase velocity is different is that when the scattered wave is combined with the incident wave there is a progressive advancement or retardation of the phase of the disturbance and the result appears as change in the phase velocity [17]. By neglecting k_I^2 compared to other terms, for the general case Eq. (51) can be written as

$$\frac{k_R^2 + 2ik_R k_I}{k_1^2} = 1 + \frac{4\pi z_a}{k_1^3}\sum_{n=0}^{\infty} A_{In}(2n+1) + \frac{iz_a C_{\text{sca}}}{k_1} \tag{54}$$

Neglecting k_I^2 in comparison to other terms is a very good assumption and can be easily verified. Therefore, for the multiple scattering case v is given by

$$v_{\text{mult}} = \frac{v_1}{\left(1 + \frac{3\phi}{x_1^3}\sum_{n=0}^{\infty}(2n+1)A_{In}\right)^{0.5}} \tag{55}$$

In addition, v_g is related to v by
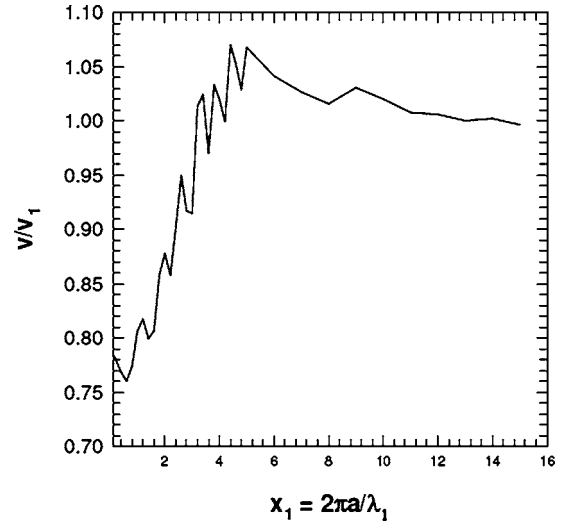
Fig. 3 Effect of multiple scattering on phase velocity of phonons for Ge scatterer in Si (ϕ=0.4)

$$v_g = \frac{v}{1 - \frac{\omega}{v}\frac{dv}{d\omega}} \tag{56}$$

α is given by

$$\alpha_{\text{mult}} = \frac{z_a C}{\left(1 + \frac{3\phi}{x_1^3}\sum_{n=0}^{\infty}(2n+1)A_{In}\right)^{0.5}} = \frac{\alpha_{\text{ind}}}{\left(1 + \frac{3\phi}{x_1^3}\sum_{n=0}^{\infty}(2n+1)A_{In}\right)^{0.5}} \tag{57}$$

Figure 3 shows the variation of v for germanium scatter in silicon for ϕ=0.4. Figure 3 shows that in the geometric scattering (large values of x_1) regime $v = v_1$. Debye speeds of Si and Ge were used for Fig. 3. Figure 3 shows that even in the linear dispersion assumption for the host medium multiple scattering can make the effective dispersion relation of particulate medium nonlinear. Figure 3 shows that in the small wavelength regime (large values of x_1), $v = v_1$, i.e., multiple scattering effects decrease with decreasing wavelength. Since

$$D(\omega) = \frac{k^2}{2\pi^2}\frac{dk}{d\omega} = \frac{k^2}{2\pi^2 v_g} \tag{58}$$

Figure 4 shows the ratio of the density of states of the particulate medium and the density of states in the pure medium. Figure 4 shows that for large value of x_1 density of states of the particulate medium approaches the density of states of the medium.

I^0 from Eq. (2) can be written as

$$I^0(T,\omega) = \frac{3\hbar\omega k_1^2\left(1 + \frac{4\pi z_a}{k_1^3}\sum_{n=1}^{\infty}(2n+1)y_n\right)}{8\pi^3}\frac{1}{\exp\left(\frac{\hbar\omega}{k_b T}\right) - 1} \tag{59}$$

It can be shown that for Rayleigh scattering v is

$$v_{\text{mult}} = \frac{v_1}{(1 + \phi[R_1 + 3R_2])^{0.5}} \tag{60}$$

v is independent of phonon frequency in the Rayleigh scattering limit. Therefore, for the Rayleigh case
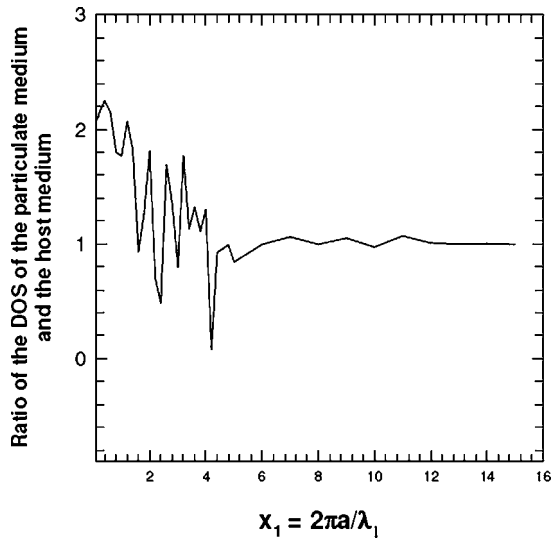
Fig. 4 Effect of multiple scattering on the density of states of phonons for Ge scatterer in Si ($\phi=0.4$)

Fig. 5 Effect of volume fraction on phase velocity of phonons in the Rayleigh regime

$$I^0(T,\omega) = \frac{3\,\hbar\,\omega^3(1+\phi[R_1+3R_2])}{8\,\pi^3 v_1^2}\frac{1}{\exp\left(\dfrac{\hbar\,\omega}{k_b T}\right)-1} \qquad (61)$$

In the Rayleigh regime I^0 is directly proportional to ϕ. For Rayleigh scattering α is given by

$$\frac{\alpha_{\mathrm{mult}}}{(x_1^4/a)} = \frac{1}{3(1+\phi[R_1+3R_2])^{0.5}}\frac{\alpha}{(x_1^4/a)}\phi \qquad (62)$$

Modification of v will also modify the Debye temperature (Θ) of the medium, which for Rayleigh regime is given

$$\frac{\Theta_{\mathrm{mult}}}{\Theta_{\mathrm{ind}}} = \frac{1}{(1+\phi[R_1+3R_2])^{0.5}} \qquad (63)$$

The $\langle\Phi\rangle_{\mathrm{mult}}$ is the same as $\langle\Phi\rangle_{\mathrm{ind}}$. Since thermal conductivity depends on I^0 and α (Eq. (3)), multiple scattering affects thermal conductivity.

6 Multiple Dependent Scattering Theory: Quasi-Crystalline Approximation (QCA)

In the most general treatment of scattering of phonons both multiple and dependent scattering effects are present. Lax [16] introduced a model called quasi-crystalline approximation (QCA) which deals with multiple and dependent scattering. Physically QCA is a combination of dependent scattering approximation and effective field approximation. This is the most general and complicated case. Under QCA Eq. (44) modifies to

$$\langle P(\mathbf{r})\rangle = P_0(\mathbf{r}) + \int_V g(\mathbf{r}')\langle P(\mathbf{r})\rangle[z(r,r')]E(\mathbf{r},\mathbf{r}')d\mathbf{r}' \qquad (64)$$

where $z(r,r')$ is the pair correlation function as defined earlier in the context of dependent scattering. Equation (64) can be written as

$$\langle P(\mathbf{r})\rangle = P_0(\mathbf{r}) + z_a\int_V g(\mathbf{r}')\langle P(\mathbf{r})\rangle E(\mathbf{r},\mathbf{r}')d\mathbf{r}' + \int_V g(\mathbf{r}')$$
$$\times\langle P(\mathbf{r})\rangle E(\mathbf{r},\mathbf{r}_j)(z(r,r')-z_a)d\mathbf{r}' \qquad (65)$$

Equation (65) shows that the first two terms on the right-hand-side are same as Eq. (44) for the multiple scattering case. The third term in Eq. (65) is the contribution due to dependent scattering. If there is no correlation between the scatterers then $z(r,r')=z_a$ and

Eq. (65) is same as the multiple scattering case (Eq. (44)). The effective wave vector is again given by Eq. (47), however the forward scattering amplitude is modified due to the dependent scattering effects. For the acoustic case under QCA, $f(0)$ (Eq. (48)) modifies to [38,40]

$$f(0) = -\frac{i}{k_1}\sum_{n=0}^{\infty} T_n A_n(2n+1) \qquad (66)$$

where T_n in Eq. (66) is due to dependent scattering and is dependent on the PY pair distribution function. This equation is substituted in Eq. (47) to derive α and v. Equation (67) shows that QCA combines DSA and EFA, and hence is called the multiple dependent scattering theory. The α and v for QCA can be derived using the same methodology as EFA. However, solution of QCA is very complicated compared to EFA for general values of x_1. For Rayleigh scattering analytical results were obtained by Tsang et al. [40] for Percus-Yevick distribution function. In the Rayleigh regime v and α are given as

$$v_{m-d} = \frac{v_1}{\sqrt{\dfrac{(1+R_1\phi)(1+R_2\phi)}{1-2R_2\phi}}} \qquad (67)$$

$$\frac{\alpha_{m-d}}{x_1^4/a} = \frac{1}{3}\left(\frac{3R_2^2(1+R_1\phi)+R_1^2(1-2\phi R_2)(1+\phi R_2)}{1-2R_2\phi}\right)$$
$$\times\frac{1}{\sqrt{\dfrac{(1+R_1\phi)(1+R_2\phi)}{1-2R_2\phi}}}\phi\frac{(1-\phi)^4}{(1+2\phi)^2} \qquad (68)$$

7 Results

Figure 5 shows the comparison of v calculated from QCA and EFA in the Rayleigh regime. Figure (5) shows that EFA and QCA are very close to each other. Figures 6 and 7 show $\alpha/(x_1^4/a)$ calculated from various approximations in the Rayleigh regime. Figures 6 and 7 show that DSA and QCA match very well, whereas EFA and ISA results are close to each other and deviate significantly in value and trend from QCA and DSA. The reason for the behaviors shown in Figs. 5–7 are that v is modified mainly due to multiple scattering, whereas α is modified mainly due to dependent scattering. Therefore v_{QCA} and v_{EFA} are close, whereas α_{DSA} and α_{QCA} are close. Effect of dependent scattering is strongest in
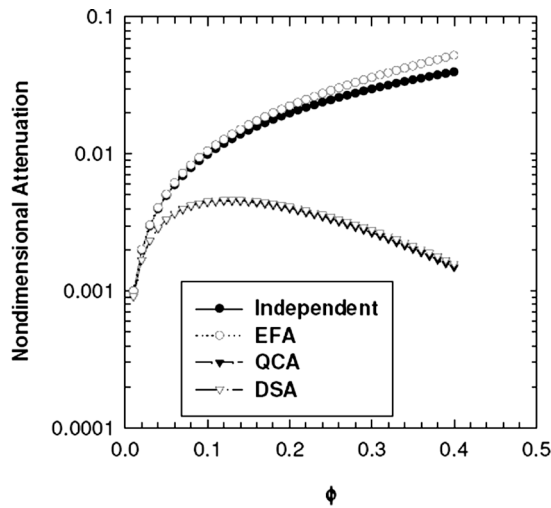
Fig. 6 Comparison between various approximations for effective attenuation for Si in Ge matrix

the Rayleigh regime. This means the m.f.p reaches a minima and then starts to increase with increasing volume fraction of scatterers. Note that this trend in scattering cross section was experimentally observed for photon transport in colloidal systems [41].

PY distribution should ideally be used only up to a volume fraction of 0.4 [38], however it is interesting to note that PY distribution function leads to correct values in the limiting cases of $\phi = 0$ and $\phi = 1$, although $\phi = 1$ is not possible for spherical particles. Putting $\phi = 1$ in Eqs. (67) and (68) lead to $v = v_2$ and $\alpha = 0$, i.e., if the medium is completely replaced by the scatterer then attenuation goes to zero and v is the same as the v_2. However, in this paper maximum value of $\phi = 0.4$ is taken as per the recommendation from Tsang and Kong [38].

Different models considered in the previous sections can be summarized as: ISA: Independent *single* scattering model, DSA: Dependent *single* scattering model EFA: Independent *multiple* scattering model, QCA: *Dependent multiple* scattering model

Multiple scattering effects are always present due to multiple number of particles, whereas dependent scattering becomes important only if the wavelength becomes comparable to the interparticle distance. QCA captures the effect of both multiple and dependent scattering but is extremely complicated except for the
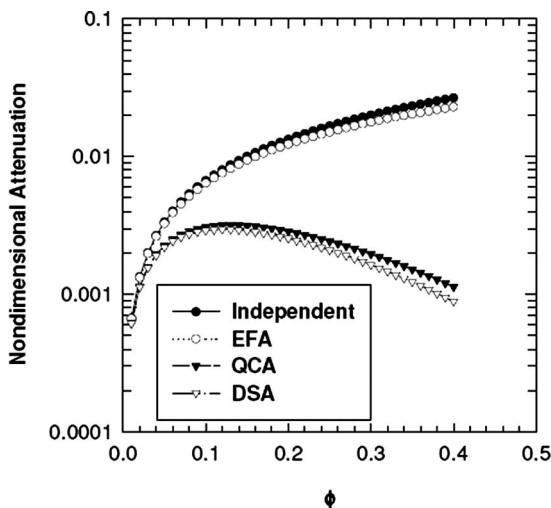
Rayleigh case. However, it is to be remembered that the Rayleigh case is the worst case for dependent scattering as the wavelength is very long. Drolen and Tien [14] have proposed that for $\delta/\lambda > 0.5$ dependent scattering effects are negligible. Since δ increases with increasing particle size, therefore if the particle size is large, then independent scattering will persist for larger volume fractions. However, the effect of multiple scattering will keep on increasing with increasing volume fractions. Since EFA is much simpler than QCA and v_p, v_g and $D(\omega)$ are mainly affected by multiple scattering, v_p, v_g and $D(\omega)$ can be obtained using EFA, whereas for α QCA needs to be used.

QCA is very complicated for the non-Rayleigh case as mentioned earlier. Since DSA matches with QCA for effective attenuation and EFA is very close to QCA for v_p, v_g and $D(\omega)$, it is proposed that effective scattering function and attenuation can be calculated using DSA and v_p, v_g and $D(\omega)$ can be calculated using EFA. It is emphasized again that for the non-Rayleigh case the difference between EFA and QCA for v_p, v_g and $D(\omega)$ will become even smaller for larger values of x_1 at the same volume fraction as the effect of dependent scattering diminishes with larger value of x_1 [14]. Therefore, for practical calculations one can use the following relations

$$k^2 = k_1^2 + 4\pi n f(0) \tag{69}$$

$$F(\theta)_{dep} = H(\theta)F(\theta)_{ind} \tag{70}$$

Equations (69) and (70) are much simpler to use than QCA. This can be very useful for calculating the scattering of phonons in the multiple and dependent scattering regime for complicated situations such as non-uniform distribution of nanoparticle sizes.

Finally a regime map for dependent versus independent scattering can be developed along the lines of the regime map for photons [14]. Inter particle distance δ can be related to the radius of the scatterer as $\delta/a \approx 1.81/\phi^{1/3}$ assuming a rhombohedral array as this pattern provides the maximum packing fraction [36]. This relation shows that δ can be in the nanometer regime if nanoparticles are used. Drolen and Tien [14], based on experimental and theoretical analysis, have proposed that for $\delta/\lambda > 0.5$ dependent scattering effects are negligible for photons. Since δ/λ is a dimensionless parameter depending only on the wavelength of the energy carrier and the interparticle distance and considering the analogy between phonons and photons, this criterion should be equally applicable for phonons. This criterion has been shown to be more appropriate for transparent particles [42]. There are few studies which have indicated this criterion may be violated for opaque particles in photonic systems [42,43] for large values of the size parameter (x_1). Opaque particles do not exist for phonons as they can penetrate through any solid. Therefore, for phonons the particles are always transparent. Therefore, $\delta/\lambda = 0.5$ can be used as the demarcation between dependent and independent scattering effects in phononic systems. The curve relating size parameter and the volume fraction that demarcates the independent scattering regime ($\delta/\lambda > 0.5$) from dependent scattering regime ($\delta/\lambda < 0.5$) can be written as [14]

$$x_1 = \pi \left(\frac{\delta}{\lambda} \right) \frac{\phi^{1/3}}{0.905 - \phi^{1/3}}, \quad \frac{\delta}{\lambda} = 0.5 \tag{71}$$

Figure 8 shows the demarcation between dependent and independent scattering. Figure 8 shows that for larger values of x_1 independent scattering results are valid for larger volume fractions as mentioned earlier. Cartigny et al. [15] also recommended that no dependent effects should occur below $\phi = 0.006$ for photons. This lower limit for dependent scattering occurs in the Rayleigh regime. To check this recommendation for phonons α is shown in Fig. 9 for various models for ϕ up to 0.0006 in the Rayleigh regime. Figure 9 shows that this recommendation is valid also for phonons as all the models are very close to each other.
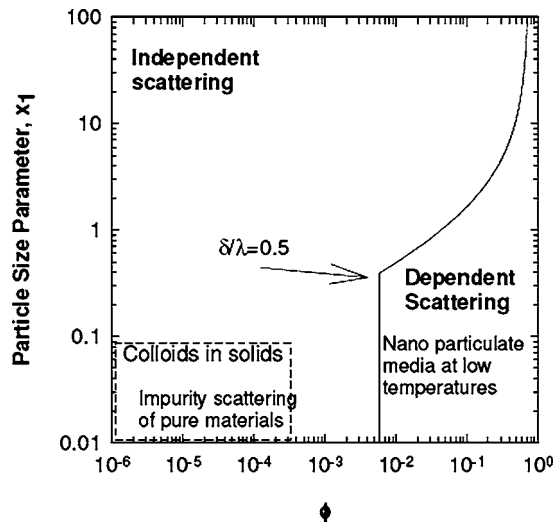

Fig. 7 Comparison between various approximations for effective attenuation for Ge in Si matrix

Fig. 8 Independent and dependent scattering regime map

The particle size parameter, x_1, depends on the diameter and the wavelength of the particles. Figure 8 shows the dependent scattering effects are more dominant for smaller value of x_1. To observe the dependent scattering of phonons at larger values of x_1, the volume fraction has to be very high. Since x_1 is inversely proportional to the wavelength, the particle diameter has to be comparable to or smaller than the phonon wavelength to observe dependent scattering effects for nominal volume fractions. The dominant phonon wavelength at room temperature in crystalline solids is of the order of 1 nm, whereas at very low temperatures it can be of the order of 50–100 nm [31]. Therefore, size parameter based on the dominant wavelength is larger at smaller temperatures as compared to higher temperatures. It is because of this reason in the regime map shown in Fig. 8, nano-particulate media such as nanocomposites at low temperature are shown in the dependent scattering regime. At higher temperatures the volume fraction has to be very high to observe this effect because x_1 based on dominant wavelength can be very large. For atomic impurities in a pure material such as Si, the volume fraction of the impurity is very small [11]. Therefore, impurity scattering falls in the independent scattering regime.
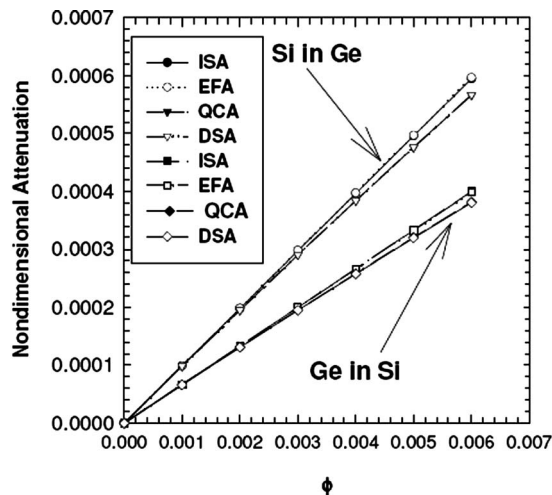
Fig. 9 Effective attenuation for $\phi<0.006$ to show that dependent scattering effects are negligible

For photons, since the dominant wavelength falls in the micrometer range, dependent scattering effects are dominant even for micron-sized particles as mentioned earlier.

Recently the author also modeled the thermal boundary resistance between a nanocomposite and a substrate using the theory of multiple and dependent scattering [44]. For thermal boundary resistance the scattering cross section and the attenuation coefficient are not important. Modification of the velocity of phonons in the nanocomposites affects the thermal boundary resistance as compared to the host medium.

8 Conclusions

Effects of multiple and dependent elastic scattering due to nano- and microparticles on acoustic phonon equilibrium intensity and attenuation were calculated using different approximations and models. The main conclusions are:

(1) Velocity, density of states and the equilibrium intensity of phonons are mainly modified due to multiple scattering;
(2) scattering cross section is mainly modified due to dependent scattering; and
(3) dependent scattering effects will be more dominant for nanocomposites at low temperatures.

Nomenclature

a = radius of the scatterer
A_n = unknown coefficient in the scattered field
C = scattering cross section
F = scattering function (Eq. (8))
g = ratio of the density in the particle and the medium
h = ratio of the sound velocity in the particle and the medium
$H(\theta)$ = dependent scattering form factor
I = phonon intensity
I^0 = equilibrium phonon intensity
I_i = intensity of the incident field
I_s = intensity of the scattered wave
k = wave vector
K_u = inelastic attenuation
P = stress field
P_s^{tot} = total scattered field
p_s = scattered amplitude from a single scatterer
R_1 = $R_1 = 1/(gh^2) - 1$
R_2 = $R_2 = (g-1)/(2g+1)$
T = temperature
v = velocity of wave
x = size parameter ($x = ka$)
z_a = average number of scatterers per unit volume
z_r = number density distribution of scatterer

Greek

Θ = debye temperature
Φ = phase function
Ψ = phase of the scattered wave
α = attenuation due to particle scattering
ϕ = volume fraction of scatterers

Subscripts

$$
\begin{array}{rcl}
1 & = & \text{medium 1} \\
2 & = & \text{medium 2 (scatterer)} \\
dep & = & \text{dependent} \\
ind & = & \text{independent} \\
m\text{-}d & = & \text{multiple dependent} \\
mult & = & \text{multiple}
\end{array}
$$

References

[1] Kim, W., Reddy, P., Zide, J., Gossard, A., Shakouri, A., and Majumdar, A., 2005, "*Beating the Alloy Limit of Thermal Conductivity in Crystalline Solids Using Embedded Nanostructures*," Materials Research Society Spring Meeting, San Francisco.

[2] Soyez, G., Eastman, J. A., Thompson, L. J., Bai, G.-R., Baldo, P. M., McCormick, A. W., DiMelfi, R. J., Elmustafa, A. A., Tambwe, M. F., and Stone, D. S., 2000, "Grain-Size-Dependent Thermal Conductivity of Nanocrystalline Yttri-Stabilized Zirconia Films Grown by Metal-Organic Chemical Vapor Deposition," Appl. Phys. Lett., **77**, pp. 1155–1157.

[3] Bao, Y., Liu, W. L., Shamsa, M., Alim, K., Balandin, A. A., and Liu, J. L., 2005, "Electrical and Thermal Conductivity of Ge/Si Quantum Dot Superlattices," J. Electrochem. Soc., **152**, pp. G432–G435.

[4] Khitun, A., Balandin, A., Liu, J. L., and Wang, K. L., 2000, "In-Plane Thermal Conductivity of a Quantum-Dot Supperlattice," J. Appl. Phys., **88**, pp. 696–699.

[5] Liu, J. L., Khitun, A., Wang, K. L., Liu, W. L., Chen, G., Xie, Q. H., and Thomas, S. G., 2003, "Cross-Plane Thermal Conductivity of Self-Assembled Ge Quantum Dot Superlattices," Phys. Rev. B, **67**, p. 165333.

[6] Wan, X., Sun, Y., Song, W., Jian, L., and Du, J., 1998, "Enhanced Flux Pinning of Bi-2223/Ag Tapes with Nano-MgO Particle Addition," Physica C, **307**, pp. 46–50.

[7] Klemens, P. G., 1991, "Thermal Conductivity of Fibre Composites at Low Temperatures," Cryogenics, **31**, pp. 238–240.

[8] Cahill, D. G., Ford, W. K., Goodson, K. E., Mahan, G. D., Majumdar, A., Maris, H. J., Merlin, R., and Phillpot, S. R., 2003, "Nanoscale Thermal Transport," J. Appl. Phys., **93**(2), pp. 793–818.

[9] Ziman, J. M., 1996, *Electrons and Phonons*, Oxford, London.

[10] Prasher, R. S., 2003, "Phonon Transport in Anisotropic Scattering Particulate Media," ASME J. Heat Transfer, **125**, pp. 1156–1162.

[11] Majumdar, A., 1993, "Microscale Heat Conduction in Dielectric Thin Films," ASME J. Heat Transfer, **115**, pp. 7–16.

[12] Van De Hulst, H. C., 1981, *Light Scattering by Small Particles*, Dover, New York.

[13] Modest, M. F., 1993, *Radiative Heat Transfer*, McGraw-Hill, New York.

[14] Drolen, B. L., and Tien, C. L., 1987, "Independent and Dependent Scattering in Packed Sphere Systems," J. Thermophys. Heat Transfer, **1**(1), pp. 63–68.

[15] Cartigny, J. D., Yamada, Y., and Tien, C. L., 1986, "Radiative Transfer With Dependent Scattering by Particles: Part 1- Theoretical Investigation," ASME J. Heat Transfer, **108**, pp. 608–613.

[16] Lax, M., 1952, "Multiple Scattering of Waves. II, The Effective Field of Dense Systems," Phys. Rev., **85**(4), pp. 621–629.

[17] Foldy, L. L., 1945, "The Multiple Scattering of Waves," Phys. Rev., **67**, pp. 107–119.

[18] Tigglen, B. A., Lagendijk, A., van Albada, M. P., and Tip, A., 1992, "Speed of Light in Random Media," Phys. Rev. B, **45**, 12233–12243.

[19] Saulnier, P. M., Zinkin, M. P., and Watson, G. H., 1990 "Scatterer Correlation Effects on Photon Transport in Dense Random Media," Phys. Rev. B, **42**, pp. 2621–2623.

[20] Meulen, F. V., Feuillard, G., Matar, O. B., Levassort, F., and Lethiecq, M., 2001, "Theoretical and Experimental Study of the Influence of the Particle Size Distribution on the Acoustic Wave Properties of Strongly Homogeneous Media," J. Acoust. Soc. Am., **110**(5), pp. 2301–2307.

[21] Krokhin, A. A., Arriaga, J., and Gumen, L. N., 2003, "Speed of Sound in Periodic Elastic Composites," Phys. Rev. Lett., **91**(26), p. 264302.

[22] Psarobas, I. E., Stefanou, N., and Modinos, A., 2000, "Scattering of Elastic Waves by Periodic Arrays of Spherical Bodies," Phys. Rev. B, **62**(1), pp. 278–291.

[23] Simkin, M. V., and Mahan, G. D., 2000, "Minimum Thermal Conductivity of Superlattices," Phys. Rev. Lett., **84**(5), pp. 927–930.

[24] Chen, G., 1998, "Thermal Conductivity and Ballistic-Phonon Transport in Cross-Plane Direction of Superlattices," Phys. Rev. B, **57**(23), pp. 14958–14973.

[25] Ju, Y. S., and Goodson, K. E., 1999, "Phonon Scattering in Silicon Films of Thickness Below 100 nm," Appl. Phys. Lett., **74**, pp. 3005–3007.

[26] Cahill, D. G., 1998, "Heat Transport In Dielectric Thin Films and at Solid-Solid Interfaces," *Microscale Energy Transport*, C-L. Tien, A. Majumdar, and F. M. Gerner, eds., Taylor and Francis, Washington, D.C., pp. 95–116.

[27] Prasher, R. S., 2004, "Mie Scattering Theory of Phonon Transport in Particulate Media," ASME J. Heat Transfer, **126**, pp. 793–804.

[28] Feuillade, C., and Clay, C. S., 1999, "Anderson (1950) revisited," J. Acoust. Soc. Am., **106**, pp. 553–564.

[29] Kerker, M., 1969, *The Scattering of Light*, Academic, New York.

[30] Siegel, R., and Howell, J. R., 1981, *Thermal Radiation Heat Transfer*, Hemisphere, Washington, D.C.

[31] Klitsner, T., and Pohl, R. O., 1987, "Phonon Scattering in Silicon Crystal Surfaces," Phys. Rev. B, **36**(12), pp. 6551–6565.

[32] Yamada, Y., Cartigny, J. D., and Tien, C. L., 1986, "Radiative Transfer With Dependent Scattering by Particles: Pat 2- Experimental Investigations," ASME J. Heat Transfer, **108**, pp. 614–618.

[33] McQuarrie, D. A., 1981, *Statistical Mechanics*, McGraw-Hill, New York.

[34] Gingrich, N. S., 1943, "The Diffraction of X-Rays by Liquid Elements," Rev. Mod. Phys., **15**, pp. 90–110.

[35] Twersky, V., 1975. "Transparency of Pair-Correlated Random Distributions of Small Scatterers, with Applications to the Cornea," J. Opt. Soc. Am., **65**, pp. 524–530.

[36] Tien, C. L., and Drolen, B. L., 1987, Annu. Rev. Numer. Fluid Mech. Heat Transfer, **1**, pp. 1–34.

[37] Prasher, R. S., 2005, "Modification of Planck Blackbody Emissive Power and Intensity in Particulate Media Due to Multiple and Dependent Scattering," ASME J. Heat Transfer, **127**, pp. 903–910.

[38] Tsang, L., and Kong, J. A., 1982, "Effective Propagation Constants for Coherent Electromagnetic Wave Propagation in Media Embedded with Dielectric Scatterers," J. Appl. Phys., **53**(11), pp. 7162–7173.

[39] Gaunard, G. C., and Wertman, W., 1989, "Comparison of Effective Medium Theories for Inhomogeneous Continua," J. Acoust. Soc. Am., **85**(2), pp. 541–554.

[40] Tsang, L., Kong, J. A., and Habashy, T., 1982, "Multiple Scattering of Acoustic Waves by Random Distribution of Discrete Spherical Scatterers With Quasicrystalline and Percus-Yevick Approximation," J. Acoust. Soc. Am., **71**, pp. 552–558.

[41] Ishimaru, A., and Kuga, Y., 1982, "Attenuation Constant of a Coherent Field in a Dense Distribution of Particles," J. Opt. Soc. Am., **72**(10), pp. 1317–1320.

[42] Singh, B. P., and Kaviany, M., 1991, "Independent Theory Versus Direct Simulation of Radiation Heat Transfer in Packed Beds," Int. J. Heat Mass Transfer, **34**(11), pp. 2869–2882.

[43] Singh, B. P., and Kaviany, M., 1992, "Modeling Radiative Heat Transfer in Packed Beds," Int. J. Heat Mass Transfer, **35**(6), pp. 1397–1405.

[44] Prasher, R. S., 2005, "Modeling of Thermal Boundary Resistance of Nanocomposites," Int. J. Heat Mass Transfer, **48**, pp. 4942–4951.

S. Sinha
Thermosciences Division,
Mechanical Engineering Department,
Stanford University,
California 94305-3030
e-mail: sanjiv@stanfordalumni.org

E. Pop

R. W. Dutton

Electrical Engineering Department,
Stanford University,
California 94305-3030

K. E. Goodson
Thermosciences Division,
Mechanical Engineering Department,
Stanford University,
California 94305-3030

Non-Equilibrium Phonon Distributions in Sub-100 nm Silicon Transistors

Intense electron-phonon scattering near the peak electric field in a semiconductor device results in nanometer-scale phonon hotspots. Past studies have argued that ballistic phonon transport near such hotspots serves to restrict heat conduction. We reexamine this assertion by developing a new phonon transport model. In a departure from previous studies, we treat isotropic dispersion in all phonon branches and include a phonon emission spectrum from independent Monte Carlo simulations of electron-phonon scattering. We cast the model in terms of a non-equilibrium phonon distribution function and compare predictions from this model with data for ballistic transport in silicon. The solution to the steady-state transport equations for bulk silicon transistors shows that energy stagnation at the hotspot results in an excess equivalent temperature rise of about 13% in a 90 nm gate-length device. Longitudinal optical phonons with non-zero group velocities dominate transport. We find that the resistance associated with ballistic transport does not overwhelm that from the package unless the peak power density approaches 50 W/μm^3. A transient calculation shows negligible phonon accumulation and retardation between successive logic states. This work highlights and reduces the knowledge gaps in the electro-thermal simulation of transistors. [DOI: 10.1115/1.2194041]

Keywords: devices, heat transfer, modeling, nanoscale, thermophysical

1 Introduction

The question of an anomalous temperature rise due to ballistic phonon transport near the heat source in a transistor remains unresolved despite much recent attention [1–4]. It is well known [5] that localized phonon emission from hot electrons near the drain of a semiconductor device results in a heat source with dimensions on the order of 10 nm. Figure 1 shows a typical heat source in a bulk silicon metal-oxide-semiconductor field-effect transistor (MOSFET). The heat generation contours were obtained from a hydrodynamic simulation of electron transport in a 90 nm gate-length device at a supply voltage, V_{dd}, of 1.2 V (with the gate and the drain terminals biased at 1.2 V, and the source terminal grounded). The peak power density is nearly 5 W/μm^3 at saturation conditions in the device. The spatial localization of this hotspot near the drain is a consequence of the sharp peak in the electric field and the short mean free path for electron-phonon scattering. A solution to the heat diffusion equation for this source would predict the temperature field to vary on a length scale comparable to the phonon mean free path in silicon, which is clearly unphysical. Several sub-continuum treatments of this problem are available in the literature as mentioned above. The common conclusion is that there is a substantial difference between the temperature rise predicted by solving the heat conduction equation and the equivalent temperature rise predicted by solving simplified versions of the phonon Boltzmann transport equation (BTE). However, past work has failed to reach a consensus on the scale of this difference and what it implies for transistors deep in the sub-100 nm regime. This is further compounded by the difficulty in experimentally verifying any of the claimed effects.

We expect the phonon distribution function to be nonlocal [6] near the source, resulting in less heat flow than predicted by the theory of heat diffusion. However, the magnitude of this reduction varies widely depending on the assumptions about the phonon dispersion relationship and the phonon frequencies dominating the heat production. The pioneering work of Lai and Majumdar [1] on concurrent electron and phonon transport employed the hydrodynamic model for electrons and a gray-body two-step energy conservation model for optical and acoustic phonons. A temperature rise of 7 K was calculated for the acoustic phonons in a 0.24 μm gate-length silicon MOSFET dissipating 1 mW per unit micrometer width of the device. Heat generation was assumed to be only through optical phonon emission, and the group velocities of optical phonon modes were assumed to be zero. To the best of our knowledge, this assumption has been retained in all subsequent models. We note that no major sub-continuum effect was observed in this study. However, Chen [7] showed that ballistic heat conduction would decrease thermal conductivity locally in the vicinity of a heated nano-particle when the size of the particle was comparable to the phonon mean free path. Expanding upon this result, Sverdrup et al. [2] considered ballistic transport in a 0.4 μm silicon-on-insulator device. Their model showed the lattice temperature rise to be 160% higher than that obtained using the heat diffusion equation. This work assumed a two-fluid dispersion for phonons. More recent efforts [3] extended the energy-moment formulation to account for phonon polarization in the acoustic modes and include frequency-dependent relaxation rates. While the above results were based on energy-moment formulations of the phonon BTE, a recent study [4] used the ballistic-diffusive equations (BDE) to also show this extra temperature rise at the source.

Interestingly, none of the silicon MOSFETs with gate lengths less than 100 nm reported in the literature (see, for example, Ref. [8]) demonstrate any anomalous thermal behavior to the best of our knowledge. While self-heating is a standard problem in the silicon-on-insulator devices, there are no indications of any anomalous heat source size effects as far as their current-voltage characteristics are concerned. This brings up several important questions: Does the presence of a sub-continuum source affect the device characteristics at all? Alternately, do device sources really
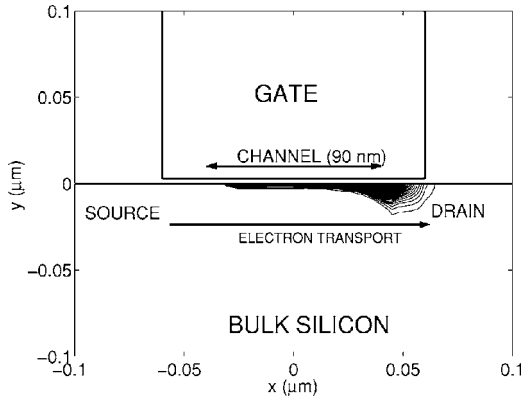

Copyright © 2006 by ASME

Fig. 1 Contours of heat generation in a 90 nm gate-length bulk silicon _n_-MOSFET are calculated using the hydrodynamic model for electron transport. The peak power density at the center is nearly 5 W/μm^3.

correspond to the sub-continuum sources that microscale heat transfer research has targeted in the past? Are the assumptions used in predicting sub-continuum phonon transport somehow responsible for the reported temperature jumps at the source? Finally, are any of these predictions realistic when it comes to practical devices?

In this paper we attempt to answer the above questions by reformulating the problem and relaxing the assumption of a graybody heat source. Going beyond this simplification requires details of electron-phonon scattering, which we obtain through Monte Carlo simulations. We solve the phonon BTE for the phonon departure from equilibrium taking into account the simulated electron-phonon scattering term as well as phonon dispersion in the acoustic and optical branches. In our formulation, we split the phonon departure from equilibrium function into two components: one that traces the evolution of the emitted phonons before they thermalize through scattering, and another that traces the diffusion of the thermalized phonons. The former is obtained by solving the ballistic BTE in a spatial region of the order of a mean free path. The latter corresponds to the solution of the BTE in the limit of diffusive transport. We compare the model predictions with existing data on hotspots in a silicon thin film. The solution is extended to predict the temperature field in a bulk silicon device. We make another departure from previous studies by using boundary conditions that take into account the thermal resistance due to the enveloping package. We show that the presence of a hotspot impedes heat conduction locally at the drain in agreement with earlier studies. However, the magnitude is much lower than earlier predictions. Based on this result, we conclude that the emission spectrum in combination with the phonon dispersion ultimately determines the magnitude of the reduced thermal conductance near the hotspot. We argue that the approach of solving either the energy moments of the phonon BTE or the BDE with no consideration of electron-phonon scattering, leads to erroneous conclusions. Finally, we solve the transient problem to investigate phonon accumulation and retardation during switching. This work aims to highlight and close key knowledge gaps in the electro-thermal modeling of sub-100 nm gate-length devices.

2 BTE Formulation for a Phonon Source

We begin the discussion on our model by first reviewing the standard form of the phonon Boltzmann transport equation. This helps us to clarify the reasoning behind our proposed model. Under the relaxation time approximation, the phonon Boltzmann transport equation at steady state is

$$\mathbf{v}_{\mathbf{k},s} \cdot \nabla N_{\mathbf{k},s} = -\frac{N_{\mathbf{k},s} - \bar{N}}{\tau_{\mathbf{k},s}} \tag{1}$$

where $(\mathbf{k},s)$ refers to the phonon mode with wave vector $\mathbf{k}$ and polarization s, $\mathbf{v}$ is the group velocity, N is the phonon occupation number, $\bar{N}$ is the equilibrium Bose-Einstein distribution function, τ is the net relaxation time from all scattering events. A common approach to solve the BTE is to write the distribution function as a departure from equilibrium function, η, added to the Bose-Einstein equilibrium distribution function

$$N = \bar{N}(T) + \eta_{\mathbf{k},s} \tag{2}$$

Substitution in the BTE yields an explicit solution for $\eta_{\mathbf{k},s}$ under the assumption that terms of second order in the temperature gradient and the term involving the gradient in the departure function are much smaller than the other terms, and may be neglected. The departure from equilibrium function thus obtained is proportional to the temperature gradient and conforms to the Fourier heat flux law

$$\eta_{\mathbf{k},s} = -\tau_{\mathbf{k},s} \mathbf{v}_{\mathbf{k},s} \cdot \nabla T \frac{\partial \bar{N}}{\partial T} \tag{3}$$

The above first order approximation works well provided the temperature gradient is small enough that the change in temperature over the relaxation length is much smaller than the absolute value of the temperature [9]. That is

$$\frac{\partial T}{\partial x} v \tau \ll T \tag{4}$$

a condition that is usually valid. However, when the second order derivative of T varies on a length scale comparable to the mean free path, the heat flux is nonlocal in the phonon distribution function. Claro and Mahan [6,10] showed that this leads to much higher temperature gradients than those predicted by the Fourier law of heat conduction.

We expect a confined high density source of phonons such as a device hot spot to also show non-local effects. In this case the nonlocality arises not due to higher order temperature derivatives but due to the gradient in the departure function itself. The departure from equilibrium function will be essentially dictated to first order by the distribution of emitted phonons in the real and reciprocal spaces of the crystal. Thus, large spatial gradients in the source function that arise when the hotspot is small compared to the phonon mean free path, translate to large gradients in the departure function as well, that is

$$\frac{\partial \eta_{\mathbf{k},s}^{*}}{\partial x^{*}} \sim \dot{n}_{\mathbf{k},s}^{*} \tau_{\mathbf{k},s} \tag{5}$$

to first order where the $\dot{n}_{\mathbf{k},s}$ is the net phonon emission from electron-phonon scattering. The asterisk in the superscript indicates a normalized quantity, with the mean free path and the equilibrium distribution function chosen as the appropriate scaling factors. The term on the right is large provided the electron-phonon relaxation rate is greater than the phonon-phonon relaxation rate. The disparity in the electron and phonon energy relaxation rates is estimated to be about two orders of magnitude in a submicrometer MOSFET operating at room temperature. The relaxation rate is about 10^{13} s^{-1} for electrons and about 10^{11} s^{-1} for optical phonons [11,12] in silicon. Thus, the gradient in the departure from equilibrium function cannot be neglected in the case of electro-thermal transport in transistors as it is in the formulation for thermal conductivity.

We modify the formulation by first including a source term in the phonon BTE that provides the net phonon emission rate due to electron scattering. The actual source function is obtained through detailed Monte Carlo simulations as described later. Other phonon scattering events are modeled through the relaxation time approxi-

mation and a reciprocal sum of the relaxation times is taken to represent the overall relaxation rate of a mode. The expressions for the relaxation times can be obtained from first order perturbation theory and empirical fits have been developed by Holland [13]. Thus, the steady-state evolution can be written as

$$\mathbf{v}_{\mathbf{k},s} \cdot \nabla N_{\mathbf{k},s} = -\frac{N_{\mathbf{k},s} - \bar{N}}{\tau_{\mathbf{k},s}} + \dot{n}_{\mathbf{k},s} \tag{6}$$

where $\dot{n}$ is the phonon emission. Proceeding as in the formulation for thermal conductivity, we write N as a small perturbation, η, over the equilibrium distribution function as given in Eq. (2). The BTE, written in terms of $\eta_{\mathbf{k},s}$, is

$$\mathbf{v}_{\mathbf{k},s} \cdot \left(\nabla \eta + \nabla T \frac{\partial \bar{N}}{\partial T} \right) = -\frac{\eta_{\mathbf{k},s}}{\tau_{\mathbf{k},s}} + \dot{n}_{\mathbf{k},s} \tag{7}$$

To obtain a solution to the non-homogeneous BTE, we first subtract the near-equilibrium departure function as given by Eq. (3) from $\eta_{\mathbf{k},s}$. The remainder is the far-from-equilibrium departure function that is due to the phonon emission source term. Thus,

$$\eta_{\mathbf{k},s} = -\tau_{\mathbf{k},s} \mathbf{v}_{\mathbf{k},s} \cdot \nabla T \frac{\partial \bar{N}}{\partial T} + n_{\mathbf{k},s} \tag{8}$$

where $n_{\mathbf{k},s}$ is the far-from-equilibrium departure function. The idea behind Eq. (8) is that the phonon flux at any point in space is due to a near-equilibrium part that obeys the Fourier law superposed on a second contribution due to the emission spectrum that does not obey the Fourier law. As we move farther away from the hotspot, we expect the second contribution to diminish strongly as the emitted phonons thermalize, and the Fourier law contribution to increase proportionately to maintain energy continuity. The BTE thus becomes

$$\mathbf{v}_{\mathbf{k},s} \cdot \left[\nabla n_{\mathbf{k},s} - \nabla \left(\tau_{\mathbf{k},s} \mathbf{v}_{\mathbf{k},s} \cdot \nabla T \frac{\partial \bar{N}}{\partial T} \right) \right] = -\frac{n_{\mathbf{k},s}}{\tau_{\mathbf{k},s}} + \dot{n}_{\mathbf{k},s} \tag{9}$$

Additionally, macroscopic energy continuity must be satisfied and this is written at steady state as

$$\nabla \cdot \mathbf{J} = \dot{Q} \tag{10}$$

where $\mathbf{J}$ is the heat flux vector and $\dot{Q}$ is the heat generation rate. These are expressed in terms of the non-equilibrium distribution functions as follows

$$\dot{Q} = \sum_s \frac{1}{8\pi^3} \int \dot{n}_{\mathbf{k},s} \hbar \omega_{\mathbf{k},s} d\mathbf{k} \tag{11}$$

$$\mathbf{J} = \sum_s \frac{1}{8\pi^3} \left[\int \mathbf{v}_{\mathbf{k},s} n_{\mathbf{k},s} \hbar \omega_{\mathbf{k},s} d\mathbf{k} - \int \mathbf{v}_{\mathbf{k},s} \tau_{\mathbf{k},s} \mathbf{v}_{\mathbf{k},s} \cdot \nabla T \frac{\partial \bar{N}}{\partial T} \hbar \omega_{\mathbf{k},s} d\mathbf{k} \right] \tag{12}$$

where the integration is taken over the first Brillouin zone. The second term in the heat flux vector can be written in terms of the thermal conductivity tensor, $\mathbf{K}$, so that Eq. (12) reduces to

$$\mathbf{J} = \sum_s \frac{1}{8\pi^3} \int \mathbf{v}_{\mathbf{k},s} n_{\mathbf{k},s} \hbar \omega_{\mathbf{k},s} d\mathbf{k} - \mathbf{K} \cdot \nabla T \tag{13}$$

With the above expression for the heat flux vector, Eq. (10) can be integrated to give

$$\sum_s \frac{1}{8\pi^3} \int \mathbf{v}_{\mathbf{k},s} n_{\mathbf{k},s} \hbar \omega_{\mathbf{k},s} d\mathbf{k} - \mathbf{K} \cdot \nabla T = \int \dot{Q} d\mathbf{r} \tag{14}$$

Equations (9) and (14) form a closed system with the unknowns being $n_{\mathbf{k},s}$ and T.

However, this system is still difficult to solve without further simplification. We note that the second term on the left in Eq. (9)
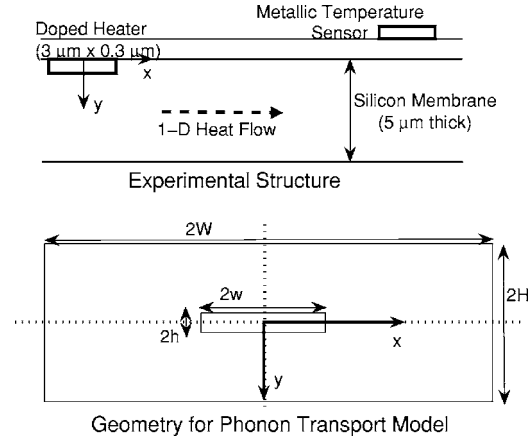
Fig. 2 A cross section of the experimental structure used to probe ballistic conduction near a doped resistor in silicon [14] is shown at the top. The resistor acted as a hotspot inside the silicon membrane. The symmetry in the problem is used to solve the BTE in the domain shown below.

is a second order term in the temperature gradient and can be neglected. This approximation causes the temperature field to be independent of the emitted phonons until they thermalize through scattering. The field does influence the far-from-equilibrium distribution indirectly through the temperature dependence of the phonon scattering rates. With the above approximation, the BTE is of the form

$$\mathbf{v}_{\mathbf{k},s} \cdot \nabla n_{\mathbf{k},s} = -\frac{n_{\mathbf{k},s}}{\tau_{\mathbf{k},s}} + \dot{n}_{\mathbf{k},s} \tag{15}$$

To solve the system we need to specify the boundary conditions. We take this up in the following sections when we develop analytical solutions for specific cases. Knowing the distribution function $n_{\mathbf{k},s}$ the continuity of energy must be solved for the temperature field.

Proceeding as above, we can also derive time dependent equations. Assuming that the temporal variation in temperature is much slower than that in the departure function, we can drop the transient temperature term in the BTE. The transient form is thus

$$\frac{\partial n_{\mathbf{k},s}}{\partial t} + \mathbf{v}_{\mathbf{k},s} \cdot \nabla n_{\mathbf{k},s} = -\frac{n_{\mathbf{k},s}}{\tau_{\mathbf{k},s}} + \dot{n}_{\mathbf{k},s}(t) \tag{16}$$

The time dependent energy conservation equation is

$$C \frac{\partial T}{\partial t} + \frac{\partial}{\partial t} \left(\sum_s \frac{1}{8\pi^3} \int n_{\mathbf{k},s} \hbar \omega_{\mathbf{k},s} d\mathbf{k} \right)$$

$$= \dot{Q} - \nabla \cdot \left[\sum_s \frac{1}{8\pi^3} \int \mathbf{v}_{\mathbf{k},s} n_{\mathbf{k},s} \hbar \omega_{\mathbf{k},s} d\mathbf{k} - \mathbf{K} \cdot \nabla T \right] \tag{17}$$

where C is the lattice heat capacity.

3 Comparison With Data for Ballistic Transport

In this section we use the above model to predict the thermal resistance associated with ballistic transport near a hotspot in silicon. Sverdrup et al. [14] used heating in a doped resistor thermometer in silicon to create a micrometer scale phonon source in a membrane structure, shown schematically in Fig. 2. In this experiment, a 3-μm-wide region in a 5-μm-thick n-type silicon membrane was p doped to a depth of about 0.3 μm. A bias applied across the terminals of the resistor forced a current through the doped region. By reverse biasing the p-n junction between the resistor and the substrate, the current was confined within the doped region. Joule heating in the doped resistor induced a temperature rise in the membrane structure. The thermal resistance of

the silicon membrane was determined from the temperature rise detected by the metallic sensors, placed parallel to the doped resistor, and by the resistor itself. The experiment was conducted in the ambient temperature range of about 100–300 K. The thermal resistance showed a large deviation from the predictions of the Fourier law at low temperatures indicating non-diffusive behavior close to the resistor.

We now calculate the thermal resistance measured in the experiment using our model. The problem is much simplified by considering the symmetry of the source and the boundary conditions. We expect scattering at the membrane boundary to be largely diffuse in the temperature range of the experiment since the phonon wavelength is comparable to the surface roughness. Thus, the BTE of Eq. (15) may be solved in the domain shown at the bottom of Fig. 2 with homogeneous boundary conditions at all four boundaries. The two-dimensional BTE in rectangular coordinates is

$$v_x \frac{\partial n_{\mathbf{k},s}}{\partial x} + v_y \frac{\partial n_{\mathbf{k},s}}{\partial y} = -\frac{n_{\mathbf{k},s}}{\tau_{\mathbf{k},s}} + \dot{n}_{\mathbf{k},s} \qquad (18)$$

We solve the above wave equation by the method of characteristics and get the solution in velocity space as follows

$$(2W - x) < -y\frac{v_x}{v_y} : G(x,y) \qquad\qquad x < y\frac{v_x}{v_y} : G(x,y)$$

$$(2W - x) > -y\frac{v_x}{v_y} : F(x,y) \qquad\qquad x > y\frac{v_x}{v_y} : F(x,y)$$

$$(2W - x) < -(2H - y)\frac{v_x}{v_y} : G(x,y) \qquad x < -(2H - y)\frac{v_x}{v_y} : G(x,y)$$

$$(2W - x) > -(2H - y)\frac{v_x}{v_y} : F(x,y) \qquad x > -(2H - y)\frac{v_x}{v_y} : F(x,y)$$

where the functions F and G are defined as

$$F(x,y) \equiv \int_0^y \frac{\dot{n}(x,y')}{v_y} \tau e^{\frac{y'-y}{\tau v_y}} dy'$$

$$G(x,y) \equiv \int_0^x \frac{\dot{n}(x',y)}{v_y} \tau e^{\frac{x'-x}{\tau v_x}} dx' \qquad (19)$$

Since the resistor was p doped, phonon emission in the experiment was from hole-phonon scattering. Unfortunately, the emission spectrum for holes at the electric fields employed in the experiment is not available in the literature. In the absence of any spectral information about the emitted phonons, we proceed by integrating the departure function over frequency to get the net power density which is known from the experiment. The net thermal resistance at the source is the sum of the diffusion and BTE resistances. Thus

$$R = R_{\text{diffusion}} + R_{\text{BTE}}$$

$$\text{with } R_{\text{diffusion}} = \frac{1}{\pi K_M}\left[\frac{\pi W}{2H} + ln\left(\frac{H}{2h + w}\right)\right] \qquad (20)$$

where w and h are the cross-sectional width and height, respectively, of the doped resistor, H is the thickness of the membrane, W is half the width of the membrane, K_M is the thermal conductivity in the membrane. Details regarding derivation of the two-dimensional thermal resistance from diffusion theory may be found in Hahne and Grigull [15].

The BTE resistance of Eq. (20) is a hypothetical resistance and is a measure of the difference between the diffusion temperature and the equivalent temperature calculated from the BTE model.

We define the equivalent temperature, T_{EQ}, similar to previous work [1,7,16] on non-equilibrium phonon transport

$$\frac{1}{8\pi^3}\sum_s \int \bar{N}(T_{EQ})\hbar\omega d\mathbf{k} = \frac{1}{8\pi^3}\sum_s \int \bar{N}(T)\hbar\omega d\mathbf{k}$$

$$+ \frac{1}{8\pi^3}\sum_s \int n\hbar\omega d\mathbf{k} \qquad (21)$$

We note that the first term on the right in Eq. (8), which is proportional to the temperature gradient, does not contribute to the energy integral. With the above definition we can represent the nonequilibrium among the modes in terms of the equivalent temperature.

In order to obtain R_{BTE} we rewrite Eq. (21) in terms of the lattice heat capacity, C, and estimate the difference in the peak lattice and equivalent temperatures. The integration over phonon frequencies is simplified by assuming phonons near the source to have a mean free path Λ and a mean relaxation time τ. Using the velocity space solution given above

$$C(T_{EQ} - T_{\text{ref}}) = C(T - T_{\text{ref}}) + \frac{1}{4\pi}\int \dot{n}\tau\hbar\omega g(\omega)d\omega$$

$$\times 4\left[\int_{\mu > \frac{W}{H}\eta} \int_{\eta > 0} \frac{d\mu d\eta}{\sqrt{1 - \mu^2 - \eta^2}}(1 - e^{-\frac{h}{\mu\Lambda}}) \right.$$

$$\left. + \int_{\eta > \frac{H}{W}\mu} \int_{\mu > 0} \frac{d\mu d\eta}{\sqrt{1 - \mu^2 - \eta^2}}(1 - e^{-\frac{w}{\eta\Lambda}}) \right]$$

$$\text{or, } (T_{EQ} - T)|_{x=0} = \frac{1}{C}\int \dot{n}\tau\hbar\omega g(\omega)d\omega \times f(w,h,\Lambda)$$

$$= \frac{\dot{Q}_{max}\tau_s}{C} f(w,h,W,H,\Lambda) \qquad (22)$$

where μ, η are the direction cosines along x and y, respectively, Λ is the phonon mean free path, and $\dot{Q}_{\max}$ is the peak heat generation rate, at x=0. The heat generated in the diode is fairly uniform across its cross section and we may use the average power instead of $\dot{Q}_{\max}$. The effective scattering time τ_s is given by the ratio of the summation of $\dot{n}\tau$ over all modes to the peak generation rate and is a fitting parameter in this calculation since the function $\dot{n}$ is unknown. The thermal resistance sustaining the peak difference between T_{EQ} and T is

$$R_{\text{BTE}} = \frac{(T_{EQ} - T)|_{x=0}}{q} = \frac{\dot{Q}_{\max}\tau_s}{qC} f(w,h,W,H,\Lambda) \qquad (23)$$

where q is the total heat current flowing to the sink. The geometry factor f is unity at room temperature, where $w, h \gg \Lambda$, and decreases slightly with temperature as the mean free path increases in relation to the source dimensions.

The dominant effect of the decreasing temperature is to increase R_{BTE}. The net increase is, in turn, due to the decrease in the heat capacity and the increase in the effective scattering time of emitted phonons. This is shown in Fig. 3 which compares the thermal resistance measured by the doped resistor at different base temperatures with predictions based on Eqs. (20) and (23). We note that in the absence of information about the emission spectrum, we have fit the model using a *constant* value for the effective scattering time for emitted phonons, τ_s, which is on the order of 10 ns. This represents a weighted average of the scattering times of all emitted phonons, the weights being the relative excitation numbers of different modes. A better fit could be obtained in principle if we include a temperature dependence of the mean scattering time. However, this is quite complicated in practice because it
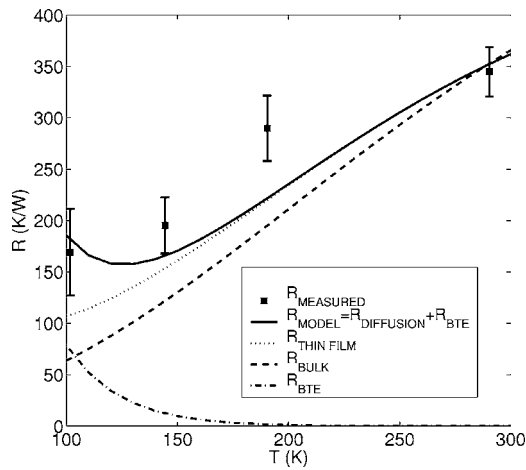
Fig. 3 A comparison of the thermal resistance measured at different temperatures. Predictions based on the proposed model are given by the continuous line. The predictions suffer from lack of information on temperature dependent scattering rates of phonons emitted by holes in silicon.

Fig. 4 The boundary conditions used in the device calculations are shown above. Heat is assumed to flow out through the metallic contacts at the top which act as fins. The heat transfer coefficient is relatively high due to this spreading effect. Most of the heat flow is toward the heat sink through the bulk silicon at the bottom.

involves the temperature dependence of not just the phonon relaxation times but also that of the weights. The latter would require detailed Monte Carlo simulations of hole-phonon scattering at each base temperature. Hence, we omit any temperature dependence of the mean scattering time in the absence of a clear intuition about the temperature dependence of the weights. The dashed line shows the thermal resistance of the membrane structure calculated from diffusion theory using the bulk thermal conductivity of silicon. Phonon boundary scattering reduces the thermal conductivity of the membrane so that the bulk predictions are significantly lower than the data. The dotted line is the thermal resistance calculated from diffusion theory (Eq. (20)) but using the thermal conductivity of the membrane measured in situ. The measurements increasingly differ from predictions of diffusion theory as the temperature decreases. We attribute this to ballistic phonon transport near the doped resistor. Phonons emitted at the heat source increasingly undergo boundary scattering as the ambient temperature decreases. This leads to a temperature slip at the heat source and causes the temperature rise to deviate from heat diffusion theory. The solid line represents the sum total of the BTE and diffusion resistances. The predictions from Eq. (20) agree well with the data at 100 and 140 K, where the departure from diffusion theory is substantial. The model is, however, unable to match the data at 190 K. Additionally, the trend curve for the data appears to have an inflection point between 140 and 190 K. The sub-continuum contribution appears to asymptote at low temperatures, whereas the model predicts ever-increasing contributions due to increasing mean free paths. We speculate that this deviation from the model may be due to the mean free path in the thin film becoming constrained by boundary scattering at low temperatures. We note that a mean free path based on the thin-film thermal conductivity and calculated from the simple kinetic theory expression of $\Lambda = 3 K / C v$ would not reproduce this trend. Thus, although the qualitative behavior of the trend in the data is understandable, we are unable to obtain a good quantitative match, primarily due to lack of information about the temperature dependence of the mean scattering time for reasons discussed above.

4 Application to a Bulk Silicon MOSFET

In this section we use our model to compute the phonon distribution function in a 90 nm gate-length bulk silicon n-type MOSFET (NMOS). We specifically choose an NMOS device because its higher power density would increase the magnitude of any sub-continuum effects. As shown in Fig. 1, the hot spot predicted
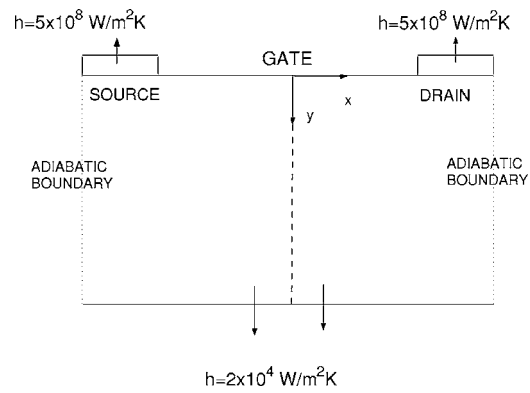
by a hydrodynamic calculation is formed in the channel under the gate toward the drain side. We note that device Monte Carlo simulations, however, predict the location to be shifted more toward the drain. The precise location is unimportant in this work since the boundaries for the thermal calculation are much larger in extent when compared to this shift in the location of the hot spot. We are more interested in the shape of the hot spot, which is predicted reasonably well by hydrodynamic calculations. The shape is approximately semicylindrical in a bulk MOSFET, with the axis aligned along the width of the device. In order to proceed toward a semi-analytical solution to our model, we approximate the location of the source to be midway between the source and the drain. As noted above, this shift in location is much smaller than the lateral extent of the device. Hence, we do not expect a major difference in our results due to this approximation. Figure 4 shows the device with the approximated source at the origin. The top boundary is assumed to be adiabatic due to the insulating gate oxide. In reality, there is a small heat flux across this interface which is ignored in our calculations to simplify the BTE solution. The side boundaries are both adiabatic due to thick isolation oxides. Some heat is lost through the metallic contacts as shown. We obtain the heat transfer coefficients by treating the metallic interconnects as fins [17]. The heat produced in the transistor mostly flows out at the bottom toward the heat sink, after passing through the bulk silicon. We assume the power density at the source to be 5 W/μm^3, scaled down by an activity factor of 0.1 (to account for average circuit activity), and the radius to be 20 nm which closely approximate realistic device hotspots. The device hot spot is further assumed to be a step function as shown in Fig. 5. The phonon spectrum in the heat production region is obtained through a Monte Carlo simulation as described below.

4.1 Electron-Phonon Scattering Using the Monte Carlo Method. The purpose behind formulating our BTE model in terms of the non-equilibrium phonon distribution function, as opposed to some fictitious temperature, is to solve it with a frequency-dependent source term without invoking questionable arguments about equivalent temperatures. Before we can solve the simplified phonon BTE, however, we need the source term, $\dot{n}_{k,s}$, appearing in Eq. (15). Ideally, the source term should couple the conservation of energy equations for electrons and phonons. In this work, we avoid that considerable complication by decoupling electron and phonon transports. In effect, we assume electron-phonon scattering to proceed between non-equilibrium electrons and equilibrium phonons. We employ Monte Carlo simulations of electron-phonon scattering to extract phonon emission rates in
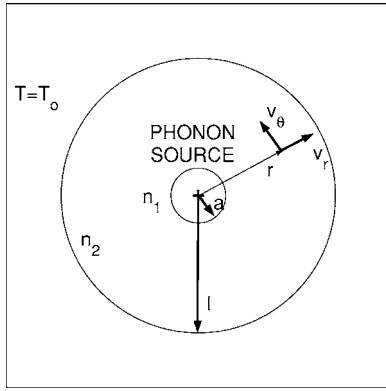
Fig. 5 The cylindrical geometry resulting from the assumption that the hotspot is located at the center of the channel serves to reduce the complexity. We further assume a step profile with the heat source confined to a radius a, as shown schematically.

silicon at a constant electric field. Details of this work are provided separately in Refs. [18,19] and are beyond the scope of this paper. We include a brief overview here for the reader's convenience.

Our Monte Carlo approach employs analytic approximations for the electron energy bands and the phonon dispersion. This band approximation for electrons is reasonable for device voltages near or below the silicon band gap (1.1 eV), such as those of future nano-technologies, and it represents an efficient implementation of the otherwise time-consuming Monte Carlo approach. The simulation treats all phonon scattering events as inelastic. Electrons exchange energy with the lattice as determined by the phonon dispersion and scattering selection rules. Scattering with intravalley longitudinal acoustic (LA) and transverse acoustic (TA) phonons, as well as with intervalley longitudinal optical (LO) and transverse optical (TO) phonons, is considered individually. The phonon dispersion relationship is used to compute the final electronic state in a manner that conserves both momentum and energy. During the simulation all phonons absorbed and emitted are tallied and net phonon emission statistics can be computed. Figure 6 shows the analytic (quadratic) phonon dispersion used in this work and the computed phonon emission spectrum at a field of 4 MV/m. To facilitate comparison the vertical axes are drawn with the same energy units. The peaks in the phonon generation
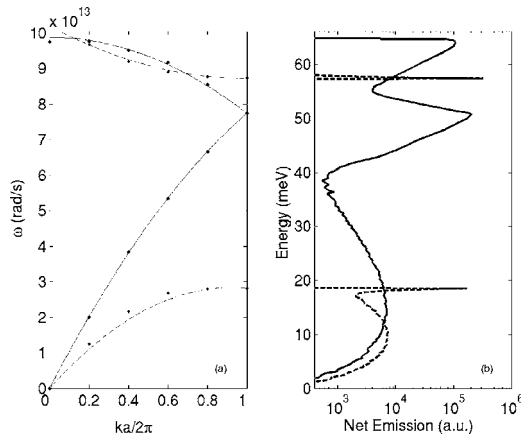

Fig. 6 The phonon dispersion in silicon along [100] is obtained from a fit to neutron scattering data from Dolling [23]. The above dispersion is assumed to hold along all directions in our calculations. At an electric field of 4 MV/m, the source term in the BTE, $\dot{n}_{k,s}$, has the frequency dependence shown on the right.

rate occur due to selection rules for the electron-phonon interaction. The relative magnitude of the peaks depends on the choice of scattering deformation potentials, which are calibrated across a wide temperature range [18]. In terms of energy, LO emissions make up about half, LA emissions about a third and TO emissions about a tenth of the heat generation rate in bulk silicon at typical electric fields. In the next section, we use the emission rate to compute steady-state phonon occupations for individual phonon frequencies.

4.2 Analytical Solution to the Phonon BTE. The two-dimensional BTE for the device problem is as given in Eq. (18). By approximating the source to be a semicylinder located at the origin, we can take advantage of the resulting radial symmetry. We convert Eq. (18) to radial coordinates as follows

$$v_r \frac{\partial n_{k,s}}{\partial r} + \frac{v_\theta^2}{r} \frac{\partial n_{k,s}}{\partial v_r} - \frac{v_\theta v_r}{r} \frac{\partial n_{k,s}}{\partial v_\theta} + \frac{n_{k,s}}{\tau_{k,s}} = \dot{n}_{k,s} \quad (24)$$

where v_r and v_θ are the radial and tangential velocities as depicted in Fig. 5. Due to axial symmetry, there is no dependence on the azimuthal angle, θ. To obtain an analytic solution we split the solution into two domains, subscripted with indices 1 and 2 as shown schematically in Fig. 5. For $r \leq a$, where a is the radius within which the source is confined, the BTE is nonhomogeneous, whereas for $r > a$ the equation is homogeneous

$$v_r \frac{\partial n_1}{\partial r} + \frac{v_\theta^2}{r} \frac{\partial n_1}{\partial v_r} - \frac{v_\theta v_r}{r} \frac{\partial n_1}{\partial v_\theta} + \frac{n_1}{\tau} = \dot{n}$$

$$v_r \frac{\partial n_2}{\partial r} + \frac{v_\theta^2}{r} \frac{\partial n_2}{\partial v_r} - \frac{v_\theta v_r}{r} \frac{\partial n_2}{\partial v_\theta} + \frac{n_2}{\tau} = 0 \quad (25)$$

The solution to Eq. (25) can be obtained by reducing the system to a form that corresponds to the BTE for electron transport in a metallic wire [20] for which the general solution is known. We find the general solution to Eq. (25) to be

$$n_1 = \dot{n}\tau \left[1 - \exp\left(-\frac{rv_r}{\tau(v_r^2 + v_\theta^2)} \right) f_1(rv_\theta, v_r^2 + v_\theta^2) \right]$$

$$n_2 = \dot{n}\tau \exp\left(-\frac{rv_r}{\tau(v_r^2 + v_\theta^2)} \right) f_2(rv_\theta, v_r^2 + v_\theta^2) \quad (26)$$

where f_1 and f_2 are arbitrary functions to be determined from boundary conditions. The boundary conditions for the problem are

$$n_2(r = l, -|v_r|, v_\theta) = 0$$

$$n_1(r = 0, |v_r|, v_\theta) = n_1(r = 0, -|v_r|, v_\theta)$$

$$n_1(r = a, \pm |v_r|, v_\theta) = n_2(r = a, \pm |v_r|, v_\theta) \quad (27)$$

Applying the boundary conditions, the final solution is

$$n_1 = \dot{n}\tau \left[1 - \exp\left(\frac{-rv_r + \sqrt{a^2(v_r^2 + v_\theta^2) - r^2 v_\theta^2}}{\tau(v_r^2 + v_\theta^2)} \right) \right]$$

$$n_2(v_r > 0) = 2\dot{n}\tau \exp\left(-\frac{rv_r}{\tau(v_r^2 + v_\theta^2)} \right) \times \sinh\left(\frac{\sqrt{a^2(v_r^2 + v_\theta^2) - r^2 v_\theta^2}}{\tau(v_r^2 + v_\theta^2)} \right)$$

$$n_2(v_r < 0) = 0 \quad (28)$$

In order to evaluate the above expressions, we need the phonon relaxation times. For the relaxation times of the acoustic modes, we use the expressions developed by Holland [13] and fit to thermal conductivity data for bulk silicon. We use a single value of 10 ps for all optical modes which is on the order of the room temperature lifetime of zone-center optical modes in silicon measured by Raman spectroscopy [12]. We now compare the model predictions with a numerical solution to the heat diffusion equa-
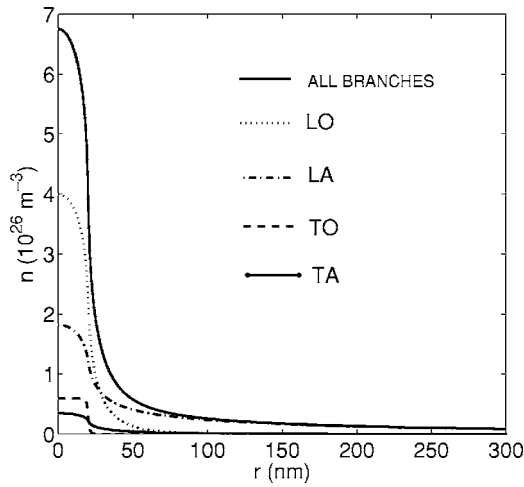
Fig. 7 The phonon number density for different branches is shown close to the hotspot, which is 20 nm in extent. The LO contribution dominates the number density and consequently, the energy density.

Fig. 9 A comparison of the temperature fields obtained from continuum heat diffusion and phonon heat conduction is shown above. Ballistic conduction augments the overall thermal resistance between the transistor and the ambient by about 13%.

tion in a 90 nm gate length device.

It is interesting to compare the non-equilibrium contribution of the different phonon branches to the number density and the non-Fourier heat flux. Due to strong emission in the LO modes, the LO contribution is more than 50% of the total non-equilibrium number density. Figure 7 shows the contribution from all four branches near the hotspot. The energy density is also proportionately higher for the LO branches since LO frequencies are higher compared to other branches. The LO and the LA branches are dominant contributors to heat conduction. This is evident from Fig. 8 where the contributions of individual branches to the heat flux are plotted. We note that the flux is zero for all branches at $r=0$ due to the imposed symmetry condition. The contributions of TA and TO are insignificant. The non-Fourier heat flux diminishes rapidly outside the hotspot, as expected from the ballistic nature of the equations. At the point where it peaks, the non-Fourier heat flux is still only 5% of the total heat flux. Therefore, heat conduction is dominated by the Fourier contribution according to this
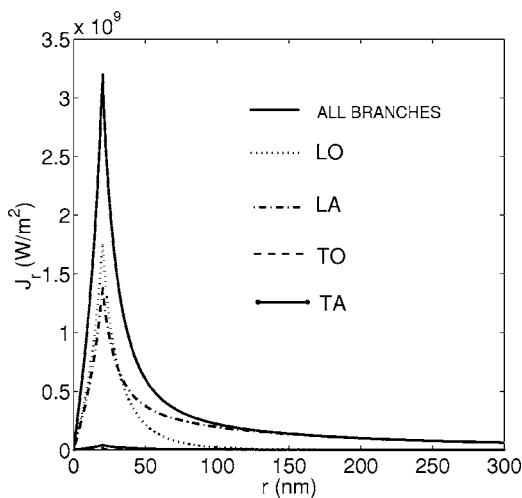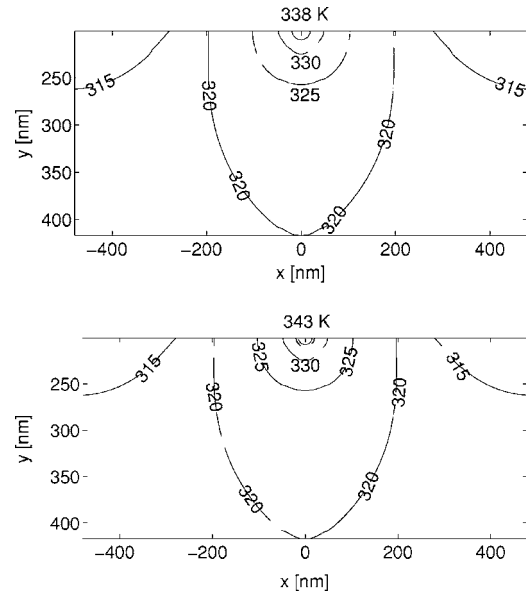

Fig. 8 The non-Fourier heat flux due to different phonon branches is shown. Ballistic heat conduction is predominantly through LO phonons. The cumulative flux is only about 5% of the total, the rest being the flux due to thermalized phonons that obey the Fourier law.

model.

The impact of the non-local departure from equilibrium is, however, quite strong near the hotspot in terms of the local energy density, which is characterized by the equivalent temperature of Eq. (21). Figure 9 compares the temperature field in the device obtained from the heat diffusion equation with that from the proposed model. The BTE model predicts an excess temperature rise of 5 K, above the peak temperature rise of about 38 K from heat diffusion. The LO branch is "hotter" than the acoustic branches, with the branch temperature being 358 K. We note that these magnitudes are very sensitive to boundary conditions. Based on these numbers, the peak equivalent lattice temperature rise is nearly 13% more than the peak temperature rise computed assuming heat diffusion for a 90 nm device. Most of the excess energy is resident in longitudinal optical phonons. The equivalent temperature rise for LO modes alone is 37% higher than that for the entire phonon ensemble, and is a measure of the nonequilibrium among phonons.

We now examine the implications of the above results in the context of the questions posed at the beginning. We emphasize that the major departure in the present treatment comes from handling of the heat source term in the phonon BTE and from solving directly for phonon distribution functions instead of hypothetical equivalent temperatures. The second aspect is critical because it altogether avoids attaching hypothetical equivalent temperatures to phonon branches as well as defining energy exchange between branches. We feel that the latter approach is arbitrary because energy mixing between phonon modes does not proceed strictly on the basis of polarization. Polarization merely decides the symmetry for selection rules in most cases. The first thing to note in the above results is the preponderance of LO phonons, not just in terms of the energy density but also in terms of the heat flux. Earlier studies have repeatedly assumed optical phonons to possess zero group velocities based on the usual flatness of their dispersion relation. We note that in reality the g-LO phonon, that is most likely to scatter with electrons, possesses a group velocity of about 1500 m/s. This is reflected in our results in terms of the dominating LO contribution to the heat flux. Further, by including boundary conditions that take into account thermal resistances all the way up to the package, we are able to gauge the importance of
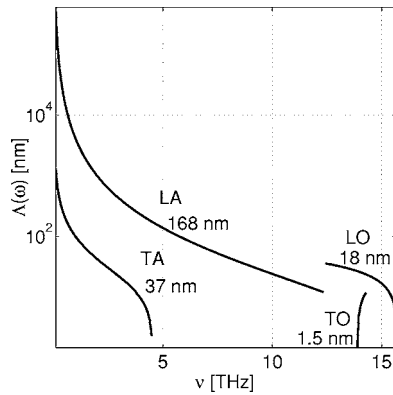
Fig. 10 The distribution of free paths in room temperature silicon as a function of phonon frequency and polarization is shown. The mean free path of phonons emitted by hot electrons in a device are also given for comparison. The use of a gray-body approximation for the heat source would lead to significant errors in predictiing the ballistic nature of transport near hotspots.

Fig. 11 A step-like phonon source symmetric about *x*=0 with a uniform power density of 5 W/μm^3 is considered for a sample transient calculation. The extent of the source, *a*, is taken to be 20 nm, consistent with device hotspots. A sink at 300 K is assumed to be present at *x*=±*l* where *l*=300 nm.

sub-continuum resistance in relation to that of the continuum junction to ambient resistance. While ballistic transport serves to create hotspots more intense than predicted otherwise by diffusion theory, the chip package remains the dominant thermal resistance. Thus, our results do not suggest the source-size effect to lead to device reliability issues in bulk devices at current electric field levels. If, however, the peak electric field and the current density increase in future devices so as to augment the peak volumetric power density by an order to magnitude (to about 50 W/μm^3), then the resultant phonon density at the hotspot would lead to reliability concerns in the drain. In this case, we estimate that sub-continuum effects would increase the thermal resistance by about 30–40%. An important aspect not considered thus far is the implication of sub-continuum effects on leakage currents. The drain to substrate junction is an important source of leakage currents. High phonon densities in the drain could promote leakage currents by increasing the thermal energy of the electrons. To the best of our knowledge, this consideration is not available in the literature. We caution that the use of equivalent temperatures obtained from BTE calculations in the usual empirical relation between leakage currents and temperature (see [21], for example) would lead to unreasonable predictions since such relations are derived on the assumption of thermodynamic equilibrium.

Finally, we examine the often-used concept of a phonon mean free path in the context of transport in silicon transistors. Generally, a phonon mean free path on the order of 100 nm is used for gray-body calculations. However, from our consideration of electron-phonon scattering and subsequent phonon transport, we find that the above figure derived from consideration of thermal conductivity of silicon is misleading, when it comes to transport near hotspots in transistors. As pointed out above, transport in the vicinity of the hotspot is far from equilibrium and has little to do with thermal conductivity of the medium. The mean free path of phonons emitted at the hot spot depends strongly on the phonon emission spectrum. Figure 10 shows the free paths of phonons in silicon as a function of frequency and polarization at 300 K based on the Holland model [13] for thermal conductivity. Also shown are the mean free paths of phonons in the emission spectrum for different branches. As evident from the figure, the phonons dominating the emission spectrum (the LO phonons) have a mean free path of about 15 nm, much smaller than the commonly used figure of 100 nm.

We explain the origin of the large deviations from heat diffusion theory reported in previous studies as follows. Essentially, none of these studies describe a pure source-size effect. The
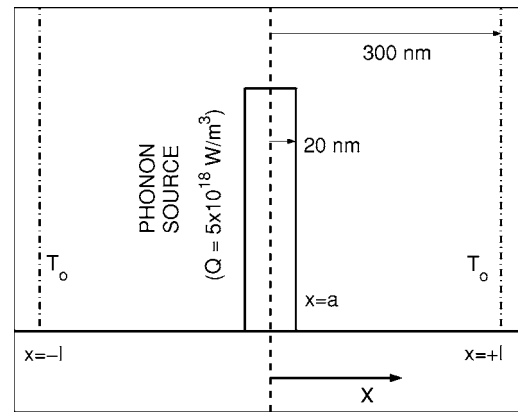
choice of a large mean free path based on a thermal phonon distribution and the selection of a thin film geometry (silicon-on-insulator device) leads to the termination of the hot phonon free paths at the film boundary. This gives rise to a large temperature slip at the heat source in the same manner as in the low temperature experiment discussed previously. *The size of the source is not the primary factor* but rather the thickness of the film in relation to the mean free path used in the study. The choice of a zero group velocity for optical phonons serves to increase the magnitude of the slip. Based on our calculations above, the phonon distribution at the heat source corresponds to a smaller mean free path. Additionally, the optical contribution corresponds to a non-zero group velocity. Thus, we do not expect the heat sources in real devices operating at room temperature to demonstrate a temperature slip unless the device is built on ultrathin film silicon with thickness down to 15 nm. This estimate of the thickness is based on the assumption that the peak electric field remains similar to that considered here. However, for a slip to occur phonons must be emitted in the direction of the thin-film boundaries. Since quantum confinement of electrons in an ultrathin film silicon channel leads to the heat source migrating away from the gate oxide interface to the middle of the channel, it is difficult to predict if the conditions required for the temperature slip would be exactly satisfied in this geometry.

5 Transient Phonon Distributions During Switching

Although the above analysis considered a steady-state phonon source, in reality, the source is time dependent if the transistor is operating in a circuit. A transistor in a digital circuit typically switches on a time scale of about 100 ps. Ignoring leakage power, a complementary metal-oxide-semiconductor device dissipates power over only a fraction of this period, referred to as the duty cycle, which is typically less than 0.3. This switching time is, however, comparable to the relaxation times of some of the phonon modes. Therefore, it is important to consider whether there is any accumulation of phonons from one switch to another. To keep the problem tractable, we ignore the device geometry completely and consider only a step-like phonon source in one dimension, as shown schematically in Fig. 11. The source is symmetric about *x*=0 where *x* is the coordinate direction. We place phonon sinks at *x*=±*l*.

The time dependent BTE to be solved is

$$\frac{\partial n_{\mathbf{k},s}}{\partial t} + v_x \frac{\partial n_{\mathbf{k},s}}{\partial x} = -\frac{n_{\mathbf{k},s}}{\tau_{\mathbf{k},s}} + \dot{n}_{\mathbf{k},s}f(t) \tag{29}$$

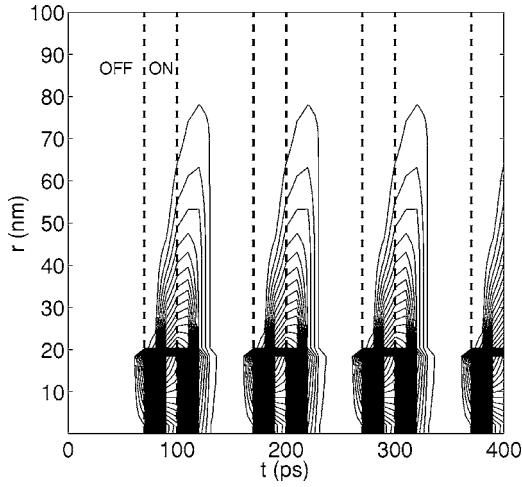where $f(t)$ is the switching function

Fig. 12 Contours of the normalized phonon number density, spaced by 0.01, are shown as a function of position and time for the longitudinal optical phonon at 14 THz. No wave retardation is evident since the emitted phonons have large enough group velocities. The accumulation of LO phonons near the source during the time period of power dissipation is clearly visible.

$$f(t) = 1 \quad t \bmod t_o < \alpha$$
$$= 0 \quad t \bmod t_o > \alpha \quad (30)$$

with t_o being the switching period and α being the duty cycle or the fraction of the time period that the device is on. Further, the boundary conditions are

$$n_{\mathbf{k},s}^+(x=0,|v_x|,t) = n_{\mathbf{k},s}^-(x=0,-|v_x|,t)$$

$$n_{\mathbf{k},s}^-(x=l,-|v_x|,t) = 0$$

$$T(x=l,t) = T_o \quad (31)$$

where $n_{\mathbf{k},s}^+$ is the departure function for phonons traveling to the right in Fig. 11 and $n_{\mathbf{k},s}^-$ is the function for phonons traveling to the left. Equation (29) is linearized by computing the relaxation time at the temperature field obtained from the heat diffusion equation. The solution to the transient problem is

$$n_{\mathbf{k},s}^+ = e^{-\frac{x}{|v_x|\tau}} \left[\int_0^x \frac{\dot{n}}{|v_x|} f\left(t + \frac{x'-x}{|v_x|}\right) e^{\frac{x'}{|v_x|\tau}} dx' \right.$$

$$\left. + \int_0^l \frac{\dot{n}}{|v_x|} f\left(t - \frac{x'+x}{|v_x|}\right) e^{-\frac{x'}{|v_x|\tau}} dx' \right]$$

$$n_{\mathbf{k},s}^- = e^{\frac{x}{|v_x|\tau}} \int_x^l \frac{\dot{n}}{|v_x|} f\left(t - \frac{x'-x}{|v_x|}\right) e^{-\frac{x'}{|v_x|\tau}} dx' \quad (32)$$

We now describe the numerical results for a hotspot with a power density of 5 W/μm^3. The numerical values for l, a, and the boundary temperature, T_o, used in these calculations are given in Fig. 11.

A phonon emitted during device switching may show a wave retardation in time if the time scale for switching is comparable to a characteristic time, obtained by dividing the length scale for emission by the group velocity of the emitted phonon. Retardation would cause the phonon distribution to be history dependent. In our computations, however, we do not find any evidence of retarded phonons. The spatial and temporal distribution for the 14 THz longitudinal optical phonon is shown in Fig. 12. We at-
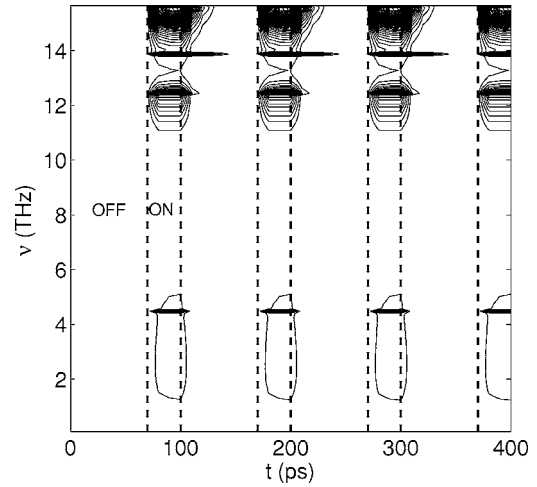
Fig. 13 Contours of the normalized phonon number density, spaced by 0.01, are shown as a function of frequency and time at $x=0$. There is no phonon accumulation for a switching period of 100 ps with a duty cycle of 30%.

tribute the absence of retardation effects to the fact that phonons dominating the emission spectrum have non-zero group velocities on the order of 1000 m/s.

Another important issue is whether the switching is fast enough to cause phonons to remain in perpetually strong nonequilibrium or to even cause accumulation over time. We do not find this to be the case for typical clock cycles. This is evident from Fig. 13 which shows the number density contours at $x=0$ as a function of phonon frequency and switching time. The emitted phonons thermalize within the off state of the device and there is no accumulation. In the absence of phonon retardation, the accumulation of phonons from one logic state to another will only occur when the time between successive states approaches the relaxation time of the dominant LO phonons. This number is about 10 ps to the best of our knowledge. Thus, unless the switch period approaches such short times, we do not anticipate any phonon accumulation. However, we note that this assertion depends on the accuracy of the relaxation times. A detailed investigation into the accuracy of scattering rates is thus necessary before making a definitive conclusion about this aspect.

6 Concluding Remarks

In this paper, we have presented a new model for determining the non-equilibrium phonon distribution function in semiconductor devices, starting from the phonon BTE. We solved a two-dimensional form of the BTE to compare our model with previous thermal resistance data on hotspots in silicon. The thermal resistance is seen to scale as the ratio of the peak power density to the lattice heat capacity. We considered a steady-state hotspot in a 90 nm gate-length bulk silicon transistor. The source distribution is taken from our prior work on Monte Carlo simulations of electron-phonon scattering. The peak equivalent temperature rise is nearly 13% more than the temperature rise from heat diffusion, with most of the excess energy resident in longitudinal optical phonons. The solution to the transient BTE shows that the emitted phonons are able to relax completely at current switching speeds but may accumulate if the switching period is reduced by half.

We find that though ballistic transport near the hotspot introduces an additional thermal resistance to that predicted by diffusion theory, the dominating contribution to the total resistance is still from the package. The sub-continuum contribution can be significant at higher peak electric fields and current densities, both being possible in future nanotransistors. Longitudinal optical phonons dominate the emission spectrum from electron-phonon

scattering and, consequently, the energy density and heat flux at the hotspot. In a clear departure from the commonly used assumption, the dominating LO modes do not have a near-zero group velocity. Also, the mean free path of the emitted phonons is significantly shorter than that for thermal phonons at 300 K. This serves to decrease the sub-continuum size effect. Additionally, the thickness of the silicon film in a silicon-on-insulator device is important in determining the onset of sub-continuum source-size effects. For a temperature slip to occur at the heat source, the mean free path of the emitted phonons should be comparable to the film thickness. The dimension of the source plays a secondary role in determining the size effect.

We are unable to gauge the impact of sub-continuum phonon conduction on leakage currents since current methods for evaluating leakage do not account for severe nonequilibrium. This is an important area for future work. Finally, we remark that, although there has been considerable progress in techniques to solve the phonon BTE, our knowledge of phonon relaxation times at large energy densities remains poor. This, however, is a prerequisite for accurately predicting non-local effects in future transistors. The relaxation rates used in this work were all derived for near-equilibrium transport and, hence, their validity at such large excitations as near a hotspot remains unknown. The authors have investigated this aspect through molecular dynamics simulations in a separate work [22].

Acknowledgment

References

[1] Lai, J., and Majumdar, A., 1996, "Concurrent Thermal and Electrical Modeling of Sub-Micrometer Silicon Devices," J. Appl. Phys., **79**, pp. 7353–7363.
[2] Sverdrup, P. G., Ju, Y. S., and Goodson, K. E., 2001, "Sub-Continuum Simulations of Heat Conduction in Silicon-on-Insulator Transistors," ASME J. Heat Transfer, **123**, pp. 130–137.
[3] Narumanchi, S. V. J., Murthy, J. Y., and Amon, C. H., 2004, "Submicron Heat Transport Model in Silicon Accounting for Phonon Dispersion and Polarization," ASME J. Heat Transfer, **126**, pp. 946–955.
[4] Yang, R., Chen, G., Laroche, M., and Taur, Y., 2005, "Simulation of Nanoscale Multidimensional Transient Heat Conduction Problems Using Ballistic-Diffusive Equations and Phonon Boltzmann Equation," ASME J. Heat Transfer, **127**, pp. 298–306.
[5] Sinha, S., and Goodson, K. E., 2002, "Phonon Heat Conduction From Nanoscale Hot Spots in Semiconductors," *12th International Heat Transfer Conference*, Grenoble, France.
[6] Mahan, G. D., and Claro, F., 1988, "Nonlocal Theory of Thermal Conductivity," Phys. Rev. B, **38**, pp. 1963–1969.
[7] Chen, G., 1996, "Nonlocal and Nonequilibrium Heat Conduction in the Vicinity of Nanoparticles," ASME J. Heat Transfer, **118**, pp. 539–545.
[8] Yeo, Y. C., Subramanian, V., Kedzierski, J., Xuan, P., King, T.-J., Bokor, J., and Hu, C., 2000, "Nanoscale Ultra-Thin-Body Silicon-on-Insulator P-MOSFET With a SiGe/Si Heterostructure Channel," IEEE Electron Device Lett., **21**, pp. 161–163.
[9] Klemens, P. G., 1951, "The Thermal Conductivity of Dielectric Solids at Low Temperatures—Theoretical," Proc. R. Soc. London, Ser. A, **208**, pp. 108–133.
[10] Claro, F., and Mahan, G. D., 1989, "Transient Heat Transport in Solids," J. Appl. Phys., **66**, pp. 4213–4217.
[11] Ferry, D. K., 2000, *Semiconductor Transport*, Taylor and Francis, New York.
[12] Menèndez, J., and Cardona, M., 1984, "Temperature Dependence of the First-Order Raman Scattering by Phonons in Si, Ge, and alpha-Sn: Anharmonic effects," Phys. Rev. B, **29**, pp. 2051–2059.
[13] Holland, M. G., 1963, "Analysis of Lattice Thermal Conductivity," Phys. Rev., **132**, pp. 2461–2471.
[14] Sverdrup, P. G., Sinha, S., Asheghi, M., Srinivasan, U., and Goodson, K. E., 2001, "Measurement of Ballistic Phonon Conduction Near Hot Spots in Silicon," Appl. Phys. Lett., **78**, pp. 3331–3333.
[15] Hahne, E., and Grigull, U., 1975, "Shape Factor and Shape Resistance for Steady Multidimensional Heat Conduction," Int. J. Heat Mass Transfer, **18**, pp. 751–767.
[16] Chen, G., 2001, "Ballistic-Diffusive Heat-Conduction Equations," Phys. Rev. Lett., **86**, pp. 2297–2300.
[17] Goodson, K. E., and Flik, M. I., 1992, "Effect of Microscale Thermal Conduction on the Packing Limit of Silicon-on-Insulator Electronic Devices," IEEE Trans. Compon., Hybrids, Manuf. Technol., **15**, pp. 715–722.
[18] Pop, E., Dutton, R. W., and Goodson, K. E., 2004, "Analytic Band Monte Carlo Model for Electron Transport in Silicon Including Acoustic and Optical Phonon Dispersion," J. Appl. Phys., **96**, pp. 4998–5005.
[19] Pop, E., Dutton, R. W., and Goodson, K. E., 2005, "Monte Carlo Simulations of Joule Heating in Bulk and Strained Silicon," Appl. Phys. Lett., **86**, p. 082101.
[20] Dingle, R. B., 1950, "The Electrical Conductivity of Thin Wires," Proc. R. Soc. London, Ser. A, **201**, pp. 545–560.
[21] Groeseneken, G., Colinge, J. P., Maes, H. E., Alderman, J. C., and Holt, S., 1990, "Temperature Dependence of Threshold Voltage in Thin-Film SOI MOSFETs," IEEE Electron Device Lett., **11**, pp. 329–331.
[22] Sinha, S., Schelling, P. K., Phillpot, S., and Goodson, K. E., 2005, "Scattering of *g*-Process LO Phonons at Hotspots in Silicon," J. Appl. Phys., **97**, pp. 023702.
[23] Dolling, G., 1963, "Lattice Vibrations in Crystals With the Diamond Structure," *Symposium on Inelastic Scattering of Neutrons in Solids and Liquids*, Chalk River, Canada, pp. 37–48.

Heat Transfer From a Translating Droplet at High Peclet Numbers: Revisiting the Classic Solution of Kronig & Brink

Douglas L. Oliver
e-mail: doliver@eng.utoledo.edu

Adham W. Souccar

MIME Department,
University of Toledo,
Toledo, Ohio 43606

More than five decades ago Kronig and Brink published a classic analysis of transport from translating droplets. Their analysis assumed that the bulk of the resistance to transfer was in the droplet phase. It considered the limiting solution as the Peclet number became very large. Their work has been cited in many subsequent studies of droplet transfer. The present work revisits their solution using numerical techniques that were not then available. It was found that only the first mode of their solution is mathematically accurate. Hence, their solution is accurate only at large times.
[DOI: 10.1115/1.2193542]

Keywords: spherical droplet, interior problem, high Peclet number, creeping flow

1 Introduction

The subject of heat and mass transfer from droplets has been investigated in many papers. There have been two excellent monographs covering the subject: *Bubbles, Drops, and Particles* [1], published in 1978 and *Transport Phenomena with Drops and Bubbles* [2], published in 1997. For brevity, the reader is referred to these two monographs for an overview of heat and mass transfer from droplets. In particular, these works present the theory of heat and mass transfer when the bulk of the resistance resides in the droplet phase (the *interior problem*).

Both of these monographs provide detailed explanations of the theory and applications associated with the classic work of Kronig and Brink [3]. Kronig and Brink's work investigated mass transfer of a solute in a droplet. In particular, they obtained a semi-analytic solution for the limiting case of the interior problem where the Peclet number is very large and the Reynolds number is very low.

Kronig and Brink assumed that as the Peclet number becomes large, the solute concentration contours (or for heat transfer—the temperature contours) become parallel with the stream function contours. The mathematical work of Kronig and Brink was directed at mass transfer. It is equally applicable to heat transfer with the appropriate dimensional conversion. The interested reader is directed to Sec. 1.3 of Sadhal et al. [2] for a discussion of the similarities in modeling heat and mass transfer. Heat transfer is modeled in the present work, however the results are equally applicable to mass transfer.

The work of Kronig and Brink has been cited in many subsequent works as a limiting bound for heat and mass transfer from droplets at low Reynolds numbers and high Peclet numbers. However, this classic work was a semi-analytic approximation using the Ritz method. The trial functions used were only quadratic functions. In addition, solutions for only two modes were sought.

The Kronig and Brink solution has been shown to be an accurate predictor of heat and mass transfer rates at large times. However, it is not clear if their solution is accurate at small times. Since this classic work has been cited often, it seems reasonable to revisit their work using more accurate numerical procedures that

were not available 50 years ago. The primary purpose of this work is to investigate what is the minimum time at which the solution of Kronig and Brink is accurate.

In this work the same assumptions are made regarding the physical parameters, (as applied to heat transfer), and governing equations as were made by Kronig and Brink. However, the solution procedure has been modified to obtain a more accurate solution at small times.

2 Problem Formulation

The steady stream function for the internal flow field for a spherical droplet at low Reynolds number is given using dimensionless coordinates by

$$\psi(r,\theta) = \frac{Ua^2}{4\left(1 + \dfrac{\mu_{\text{ext}}}{\mu_{\text{drop}}}\right)}(r^2 - r^4)\sin^2\theta \qquad (1)$$

where U and a are the droplet velocity and radius, and μ_{ext} and μ_{drop} are the dynamic viscosities of the continuous and droplet phases. The radius r is scaled by the droplet radius, a.

As stated above, the temperature contours are assumed to be parallel with the stream function contours

$$\lim_{Pe \to \infty} \Theta(r,\theta,\phi,t) = \Theta(\psi,t) \qquad (2)$$

In Eq. (2), Θ is the scaled temperature in the droplet

$$\Theta = \frac{T - T_s}{T_{\text{init}} - T_s}$$

where T_{init} is the initial droplet temperature and T_s is the surface temperature. For the *interior problem*, the bulk of the resistance to heat transfer is assumed to be in the droplet phase. As such, the temperature at the surface, T_s, is assumed to be constant.

Following the example of Kronig and Brink, Eq. (2) may be restated in terms of the dimensionless spatial variable, ξ, and the dimensionless time, τ

$$\lim_{Pe \to \infty} \Theta(r,\theta,\phi,t) = \Theta(\xi,\tau) \qquad (3)$$

where $\tau = \alpha t/a^2$ and

$$\xi = 4r^2(1 - r^2)\sin^2\theta \text{ where } 0 \leq \xi \leq 1. \qquad (4)$$

Copyright © 2006 by ASME

According to Eq. (8) of Kronig and Brink's paper (applied to heat transfer), the differential equation for heat transport may be stated as

$$\frac{\partial}{\partial \xi}\left[p(\xi)\frac{\partial \Theta}{\partial \xi}\right] = \frac{q(\xi)}{16}\frac{\partial \Theta}{\partial \tau} \qquad (5)$$

The initial and boundary conditions imposed on Eq. (5) are

$$\Theta(\xi, \tau = 0) = 1 \qquad (6)$$

$$\Theta(\xi = 0, \tau) = 0 \text{ (along the outer stream line) and} \qquad (7)$$

$$\Theta(\xi = 1, \tau) \text{ is finite at the vortex center} \qquad (8)$$

The functions $p(\xi)$ and $q(\xi)$ are given by Eqs. (14) and (15) of Kronig and Brink's paper as

$$p(\xi) = \frac{2\sqrt{1+\sqrt{\xi}}}{3}\left[(4 - 3\xi)E\left(\frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}\right)\right.$$
$$\left. - (4\sqrt{\xi} - 3\xi)K\left(\frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}\right)\right] \text{ and} \qquad (9)$$

$$q(\xi) = \frac{2}{\sqrt{1+\sqrt{\xi}}}K\left(\frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}\right) \qquad (10)$$

where $E(x)$ and $K(x)$ are the complete elliptic integrals [4].

Equation (5) is solved using separation-of-variables techniques where $\Theta(\xi, \tau)$ is assumed to be of the form

$$\Theta(\xi, \tau) = \sum_{n=1}^{\infty} A_n \Xi_n(\xi)e^{-16\lambda_n^2 \tau} \qquad (11)$$

The corresponding ordinary differential equation for Ξ_n is then

$$\frac{d}{d\xi}\left[p(\xi)\frac{d\Xi_n}{d\xi}\right] + \lambda_n^2 q(\xi)\Xi_n(\xi) = 0 \qquad (12)$$

The boundary conditions imposed on $\Xi_n(\xi)$ are

$\Xi_n(\xi = 0) = 0$ (at the outer stream function of the droplet) and

$\Xi_n(\xi = 1)$ is finite (at the droplet vortex center)

Equation (12) is solved using second-order finite difference techniques. That is, Eq. (12) is transformed into a matrix system of equations.

Both of the functions $p(\xi)$ and $q(\xi)$ are positive on the interval (0,1). In addition, $\lim_{\xi \to 0} p(\xi) = 0$. Under these circumstances, Eq. (12) is a proper Sturm-Liouville-like problem where the corresponding eigenfunctions, Ξ_n, are orthogonal on the interval (0,1) with respect to the weighting function $q(\xi)$ [5]. That is

$$\int_0^1 \Xi_m(\xi)\Xi_n(\xi)q(\xi)d\xi = \delta_{mn}\int_0^1 [\Xi_n(\xi)]^2 q(\xi)d\xi \qquad (13)$$

where δ_{mn} is the Kronecker delta function.

The eigenvalues, λ_n, may be evaluated using an energy balance analysis which requires that

$$\lambda_n^2 \int_0^1 \Xi_n(\xi)q(\xi)d\xi = \frac{8}{3}\left.\frac{d\Xi_n}{d\xi}\right|_{\xi=0} \qquad (14)$$

Once the eigenvalues and corresponding eigenfunctions are ascertained, the eigenfunctions are *normalized* so that

$$1 = \int_0^1 [\Xi_n(\xi)]^2 q(\xi)d\xi \qquad (15)$$

The coefficients, A_n, may be obtained using the orthogonality of the eigenfunctions, Ξ_n, coupled with the initial condition, Eq. (6)

$$1 = \sum_{n=1}^{\infty} A_n \Xi_n(\xi) \qquad (16)$$

The inner products of both sides of Eq. (16) are taken with $\Xi_m(\xi)$. Since the eigenfunctions Ξ_n are orthonormal with respect to the weighting function $q(\xi)$

$$A_n = \int_0^1 \Xi_n(\xi)q(\xi)d\xi. \qquad (17)$$

The bulk or average temperature, $\bar{\Theta}(\tau)$, may be calculated by taking the weighted average of the temperature, with

$$\bar{\Theta}(\tau) = \frac{\sum_{n=1}^{\infty} A_n \left[\int_0^1 \Xi_n(\xi)q(\xi)d\xi\right]e^{-\lambda_n^2 16\tau}}{\int_0^1 q(\xi)d\xi}$$

But, $A_n = \int_0^1 \Xi_n(\xi)q(\xi)d\xi$ and $\int_0^1 q(\xi)d\xi = 8/3$. Hence

$$\bar{\Theta}(\tau) = \frac{3}{8}\sum_{n=1}^{\infty} A_n^2 e^{-\lambda_n^2 16\tau} \qquad (18)$$

Solving for $\Theta(\xi, \tau)$ involves finding the eigenfunctions, $\Xi_n(\xi)$, the associated eigenvalues, λ_n, and the corresponding coefficients, A_n. Second-order finite difference techniques were used to solve Eq. (12). The values of λ_n were iteratively adjusted until the characteristic equation, Eq. (14), was satisfied. Once the eigenfunctions, $\Xi_n(\xi)$, and associated eigenvalues, λ_n, are obtained, the eigenfunctions are then normalized.

As a practical matter, Eq. (11) must be truncated with

$$\Theta(\xi, \tau) = \sum_{n=1}^{n_{max}} A_n \Xi_n(\xi)e^{-\lambda_n^2 16\tau} \qquad (19)$$

where n_{max} is the truncation limit. Finally, the coefficients A_n were obtained using Eq. (17). All integrations were performed using the trapezoidal rule. The grid for integration was equally spaced in ξ and was identical with the finite-difference grid spacing.

3 Results

When using numerical integration and finite-difference techniques, it is important to know if the numerical grid is small enough to insure convergence. To illustrate the convergence of the solution with $\Delta\xi$, the values of A_n and λ_n^2 are given in Table 1. As may be seen in Table 1, the convergence with respect to $\Delta\xi$ is rapid for the first few modes. However, the convergence for higher-order modes is slow. However, except at very small times, the higher-order terms are irrelevant. For example, with $\lambda_3^2 = 20.9646$, the third mode is insignificant (relative to the first mode) for $\tau > 0.01$. Likewise, with $\lambda_{30}^2 = 2,394.5$ the 30th term is relatively insignificant for $\tau > 0.0001$.

The first two eigenfunctions and corresponding eigenvalues may be compared with those obtained by Kronig and Brink. In the original work of Kronig and Brink, n_{max} was chosen at only two. Their approximate solution was of the form

$$A_1 = 1.32, \quad \lambda_1^2 = 1.678, \quad \Xi_1(\xi) = 0.828\xi + 0.401\xi^2 \text{ and}$$

$$A_2 = 0.73, \quad \lambda_1^2 = 9.83, \quad \Xi_2(\xi) = 4.23\xi - 5.62\xi^2.$$

For comparison purposes, plots of the product $A_i\Xi_i(\xi)$ for the first two eigenfunctions are illustrated in Fig. 1. For the first eigenfunction, the present results match that of Kronig and Brink well. However, for the second mode, the match is not good. The same is true for the corresponding eigenvalues. For Kronig and Brink's solution the eigenfunctions were limited to quadratic polynomials. In addition, numerical integration was (apparently) per-

Table 1 Convergence with respect to $\Delta\xi$

$\Delta\xi=0.01$	$\Delta\xi=0.005$	$\Delta\xi=0.0020$	$\Delta\xi=0.0010$	$\Delta\xi=0.0005$
$A_1=1.3249$	$A_1=1.3249$	$A_1=1.3249$	$A_1=1.3249$	$A_1=1.3249$
$\lambda_1^2=1.6778$	$\lambda_1^2=1.6777$	$\lambda_1^2=1.6777$	$\lambda_1^2=1.6777$	$\lambda_1^2=1.6777$
$A_2=-0.6048$	$A_2=-0.6045$	$A_2=-0.6045$	$A_2=-0.6045$	$A_2=-0.6045$
$\lambda_2^2=8.5874$	$\lambda_2^2=8.5959$	$\lambda_2^2=8.5984$	$\lambda_2^2=8.5987$	$\lambda_2^2=8.5988$
$A_3=0.3934$	$A_3=0.3936$	$A_3=0.3937$	$A_3=0.3937$	$A_3=0.3937$
$\lambda_3^2=20.9458$	$\lambda_3^2=20.9605$	$\lambda_3^2=20.9641$	$\lambda_3^2=20.9645$	$\lambda_3^2=20.9646$
	$A_{10}=-0.1156$	$A_{10}=-0.1164$	$A_{10}=-0.1166$	$A_{10}=-0.1166$
	$\lambda_{10}^2=251.7077$	$\lambda_{10}^2=258.8668$	$\lambda_{10}^2=259.8718$	$\lambda_{10}^2=260.1168$
		$A_{20}=-0.0590$	$A_{20}=-0.0586$	$A_{20}=-0.0588$
		$\lambda_{20}^2=1029.2$	$\lambda_{20}^2=1054.2$	$\lambda_{20}^2=1062.9$
			$A_{30}=-0.0397$	$A_{30}=-0.0394$
			$\lambda_{30}^2=2354.6$	$\lambda_{30}^2=2394.5$

formed using a value of $\Delta\xi=0.1$. Thus, it is not surprising the predictions of Kronig and Brink are better for the first mode than they are for the second mode.

Hence, the solution of Kronig and Brink is accurate only for the first mode. As such, this classic solution will be accurate only when the second mode is insignificant. That is, it will be accurate when

$$\exp(-16\lambda_2^2\tau)\ll 1 \text{ or}$$

$$\lambda_2^2\tau > 0.3$$

Since $\lambda_2^2=8.60$, it follows that the Kronig and Brink solution should be accurate for $\tau>0.03$.

The bulk or average temperature of the droplet may be determined using a truncated form of Eq. (18)

$$\bar{\Theta}(\tau) = \frac{3}{8}\sum_{n=1}^{n_{max}} A_n^2 e^{-\lambda_n^2 16\tau} \qquad (20)$$

The bulk temperature predicted by the Kronig and Brink model is then

$$\bar{\Theta}(\tau) = \frac{3}{8}\{(1.32)^2 e^{-26.85\tau} + (0.73)^2 e^{-157.3\tau}\} \text{ (Kronig and Brink)}$$

$$(21)$$

The bulk temperature, $\bar{\Theta}(\tau)$, predicted by the present model is compared with the bulk temperature predicted by the Kronig and Brink model in Fig. 2. Notice that for $\tau>0.003$ the bulk temperature predicted by Kronig and Brink is nearly identical to that predicted by the present model. This result may be expected, as the first mode dominates at larger times.

The good fit illustrated on Fig. 2 for $\tau>0.003$ seems to contradict the earlier statement that the Kronig and Brink solution was only accurate for $\tau>0.03$. However, the accurate bulk temperature predictions of Kronig and Brink for $\tau>0.003$ do not imply that their model is accurate with respect to temperature profiles at such low times. The temperature profiles predicted by Kronig and Brink are given by

$$\Theta(\xi,\tau) = 1.32(0.828\xi + 0.401\xi^2)e^{-16(1.678)\tau}$$
$$+ 0.73(4.23\xi - 5.62\xi^2)e^{-16(9.83)\tau}$$

or

$$\Theta(\xi,\tau) = (1.093\xi + 0.529\xi^2)e^{-26.85\tau}$$
$$+ (3.09\xi - 4.10\xi^2)e^{-157.3\tau} \text{ (Kronig and Brink)}$$

These profiles are compared with the present model in Fig. 3. In Fig. 3 the temperature profiles are plotted for $\tau=0.003$ and $\tau=0.03$. The details of the temperature profile predicted by the Kronig and Brink solution are not very accurate at $\tau=0.003$. Yet the temperature profile predicted by their model is quite accurate for $\tau\geq 0.03$ when only the first mode dominates.

Another measure of the mass transfer is the Nusselt number (or Sherwood number for mass transfer). The Nusselt number may shown to be

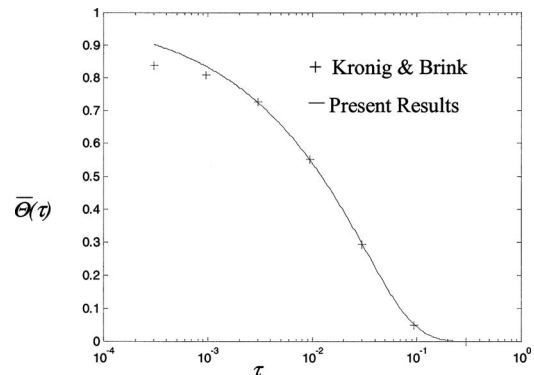**Fig. 1 First two eigenfunctions, present versus Kronig and Brink**

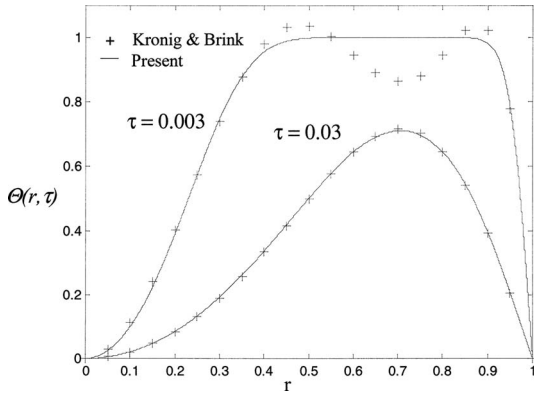**Fig. 2 Bulk temperature, $\bar{\Theta}(\tau)$**

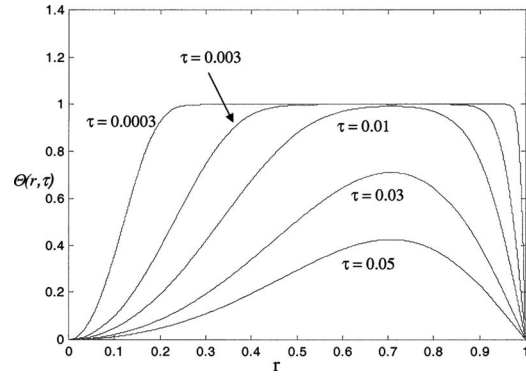**Fig. 3** $\Theta(r,\tau)$ along $\theta=\pi/2$, Kronig and Brink versus present
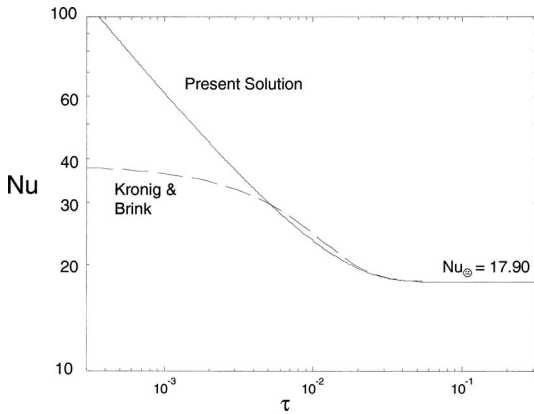

Fig. 5 $\Theta(r,\tau)$ along $\theta=\pi/2$

Fig. 4 Nusselt number, Kronig and Brink versus present

Fig. 6 Upper: locations of $\Theta=0.9$ as a function of time; lower: contours of Θ at $\tau=0.01$

$$\text{Nu} = -\frac{2\left(\dfrac{\partial \bar{\Theta}}{\partial \tau}\right)}{3\bar{\Theta}} \tag{22}$$

Substitution of Eq. (20) into Eq. (22) yields

$$\text{Nu} = \frac{32\sum_{n=1}^{n_{\max}} A_n^2 \lambda_n^2 e^{-\lambda_n^2 16\tau}}{3\sum_{n=1}^{\infty} A_n^2 e^{-\lambda_n^2 16\tau}} \tag{23}$$

Using Eq. (21) the Nusselt number based on the Kronig and Brink solution is

$$\text{Nu} = \frac{2(46.78 e^{-26.85\tau} + 83.81 e^{-157.3\tau})}{3(1.742 e^{-26.85\tau} + 0.533 e^{-157.3\tau})} \text{ (Kronig and Brink)} \tag{24}$$

The Nusselt numbers, as predicted by the present model and the Kronig and Brink model are plotted in Fig. 4 as a function of τ. The Nusselt number based on these two models differs substantially at low times. However, for $\tau > 0.03$ they converge.

For $\tau > 0.05$, *both* the Kronig and Brink as well as the present solution converge to a steady value for the Nusselt number of

$$\text{Nu}_\infty \equiv \lim_{\tau \to \infty} \text{Nu} = 17.90 \tag{25}$$

Interestingly, both of the monographs mentioned above erroneously state that the steady-state value for the Nusselt number, according to Kronig and Brink, is 17.66.[1] The steady-state value

---

[1] See Eq. (3.160) of Sadhal et al. and Eq. (3-86) of Clift et al.

for the Nusselt number of 17.90 was also reported by Oliver and DeWitt [6].

The temperature profiles predicted by the present model are plotted at various times in Fig. 5 for $0.0003 \leq \tau \leq 0.05$. Note that the center of the vortex, at $\xi=1$, is undisturbed until $\tau \approx 0.01$. As an alternative method of illustrating the development of the temperature profile, the upper part of the droplet in Fig. 6 illustrates the movement of the $\Theta=0.9$ contour front as a function of τ. Similarly, the lower part of the droplet in Fig. 6 illustrates contours of Θ at $\tau=0.01$.

Finally, for low values of τ, the present analysis has been demonstrated to be mathematically more accurate than Kronig and Brink's model. This advantage is real, yet it should not be over-emphasized. For finite Peclet numbers, initially there will be a sharp temperature gradient near the exterior of the droplet. This large gradient will result in conduction being the dominant initial mode of heat transfer. Thus for large, yet finite, Peclet numbers, the assumption that the temperature contours are parallel with the stream function contours will initially be invalid. This transition time is demonstrated by the initial oscillations in the Nusselt number predicted by previous investigators for large, but finite, Peclet numbers. (See, for example, Chap. 3 of either [1] or [2].)

4 Conclusion

The classic solution of Kronig and Brink for heat transfer from a droplet with high Peclet number has been revisited using a more accurate solution procedure. It was found that the Kronig and Brink model accurately predicts the bulk temperature of the droplet for $\tau > 0.003$. However, the Kronig and Brink solution did not accurately model the details of the temperature profile and the Nusselt number until $\tau > 0.03$. As such, the present model is more accurate for short times, ($\tau < 0.03$).

Plots of the bulk temperature, the spatially distributed temperature, and the Nusselt number were obtained. Finally, the Nusselt number for large times was shown to be 17.90, rather than the previously (and erroneously) reported value of 17.66.

Nomenclature

a = droplet radius
A_n = coefficient, see Eq. (11)
n_{max} = truncation limit
Nu = Nusselt number or Sherwood number, see Eq. (22)
$p(\xi)$ = defined by Eq. (9)
$q(\xi)$ = defined by Eq. (10)
r = radial coordinate made dimensionless with the radius a
T = temperature
U = droplet velocity

Greek Symbols

α = diffusivity
θ = tangential coordinate
Θ = dimensionless solute concentration, $\Theta = \frac{T - T_s}{T_{init} - T_s}$
$\bar{\Theta}$ = average or bulk temperature, see Eq. (18)
λ_n = eigenvalue, see Eq. (11)
μ = dynamic viscosity

ξ = dimensionless spatial coordinate, $\xi = 4r^2(1 - r^2)\sin^2\theta$
$\Xi_n(\xi)$ = eigenfunction, see Eq. (11)
τ = dimensionless time, $\tau = \frac{\alpha t}{a^2}$
ψ = stream function

Subscripts

drop = droplet or dispersed phase
ext = continuous phase
init = initial
s = droplet surface
∞ = large times

References

[1] Clift R., Grace J. R., and Weber M. E., 1978, *Bubbles, Drops, and Particles*, Academic, New York, Chaps. 1 and 3.
[2] Sadhal S. S., Ayyaswamy P. S., and Chung J. N., 1997, *Transport Phenomena With Drops and Bubbles*, Springer, New York, Chaps. 1 and 3.
[3] Kronig R., and Brink J. C., 1950, "On the Theory of Extraction From Falling Droplets," Appl. Sci. Res., Sect. A, **2**, pp. 142–154.
[4] Abramowitz M., and Stegun I., 1976, *Handbook of Mathematical Functions*, National Bureau of Standards, Eqs. (17.3.6) and (17.3.4).
[5] Hildebrand F. B., 1976, *Advanced Calculus for Applications*, 2nd ed., Prentice–Hall, Englewood Cliffs, NJ, Sec. 5.6.
[6] Oliver D. L. R., and DeWitt K. J., 1993, "High Peclet Number Heat Transfer From a Droplet Suspended in an Electric Field: Interior Problem,"Int. J. Heat Mass Transfer, **36**, pp. 3153–3155.

Effects of Multiple Reflections on Hole Formation During Short-Pulsed Laser Drilling

Michael F. Modest

Fellow ASME
Department of Mechanical and Nuclear
Engineering,
The Pennsylvania State University,
University Park, PA 16802
e-mail: mfmodest@psu.edu

Beam guiding effects during laser drilling due to multiple specular reflections inside the hole are analyzed for the case of very short laser pulses (nanosecond range). Specular reflections are valid for materials that retain a smooth surface during laser evaporation (small optical roughness compared to the laser wavelength). The problem is assumed to be two-dimensional axisymmetric (unpolarized laser), with the hole geometry defined by nodal values connected through a cubic spline. The net radiative flux onto a surface node is determined through ray tracing methods. The resulting absorbed laser flux is combined with a simple quasi-one-dimensional conduction model (to assess the minor conduction losses) and an Arrhenius evaporation rate model, to predict hole development as a function of time through iteration. To stabilize this highly nonlinear and thus unstable problem (in numerical analysis as well as in experiments) the laser beam is diffused a small amount from the specular direction (to also account for the limitation that no beam can be focused down to a point), and by periodic slight smoothing of the irradiation levels. Results show that drilling rates are increased dramatically due to beam trapping for highly reflective materials, resulting in a more pointed hole profile.
[DOI: 10.1115/1.2194035]

Keywords: laser machining, radiation, drilling, ablation

1 Introduction

Lasers have a variety of applications in modern technology because of their ability to produce high-power beams. Applications include welding, drilling, cutting, machining, medical surgery, and others. Modeling of laser drilling, cutting, and scribing has been addressed by a number of investigators. Simple one-dimensional drilling models have been given by Dabby and Paek [1] and Wagner [2]. Other approximate laser drilling models have been developed by von Allmen [3], Petring et al. [4], and others. Multiple reflections during laser drilling have also been addressed by a number of researchers. Ramanathan and Modest [5] used an approximate conduction and ablation model to simulate laser drilling with diffuse reflections, noting that the nonlinear feedback between irradiation and surface recession causes great instabilities even for purely diffuse reflections. Probably for that reason all investigations to date dealing with multiple specular reflections during laser drilling have used fake geometries to eliminate such instabilities: Milewski and Sklar [6] used ray tracing to model a V-shaped welding joint, while Lee et al. [7] used a cone to model an axisymmetric keyhole; Ho et al. [8,9] considered a paraboloid of revolution; Ki et al. [10] used a somewhat more general evolving fake shape, and, finally, Solana and Negro [11] also considered an evolving shape, but used a somewhat arbitrary surface recession model to avoid instabilities. Interestingly, specular reflections during scanning laser operation (welding, scribing) cause less instabilities, since they gradually move out of the laser interaction zone, rather than building up under the laser. Bang and Modest [12–14] studied scribing of ceramics with specular reflections (with and without polarization effects), and Ki et al. [15,16] modeled a three-dimensional (3D) keyhole for welding applications. Both Ramanathan and Modest [5] and Ki et al. [15,16]

noticed that, once the hole attains sufficient depth, the absorbed reflected irradiation density can exceed the absorbed direct irradiation by a factor of 10 or more, concentrating near the center of the beam. Apparently, this is independent of surface quality, since the first of these studies considered diffuse reflections, and the second one specular reflections. It is the purpose of the present work to study the effects of multiple, quasispecular reflections on the drilling of deep, high-aspect ratio holes with a short-pulsed laser (nanosecond range). Realizing that no real surface is a truly specular reflector, a "quasispecular" reflector reflects its energy into a small cone centered around the specular direction, with Gaussian decay away from it. This results in a somewhat smoothened distribution of reflected energy and, therefore, helps keep the predicted drilling front stable. Nevertheless, some very minor additional smoothing of the irradiation profile is required to retain stability. This is to be expected: Experimental evidence suggests that the slightest inhomogeneity (in material or beam stability) can lead to vastly different (and strangely shaped) holes. As such, the discussion here should be understood as qualitative, showing the most likely hole shape in a reasonably homogeneous material irradiated by a stable laser beam.

2 Theoretical Background

In order to obtain a realistic yet feasible mathematical description of the evaporation front that forms a high aspect ratio hole in a solid irradiated by a short-pulsed concentrated laser beam, several assumptions will be made:

1. The solid is isotropic with constant properties (evaluated at evaporation temperature). Since conduction losses are small during short-pulse irradiation (nanosecond range), this assumption should be a good one.
2. The material is opaque, i.e., the laser beam does not penetrate appreciably into the medium. This assumption is good for metals; for other materials, such as ceramics, slight penetration (of a micrometer or so) could easily be incorporated into the conduction model, but is not expected to have a noticeable effect on the results [17].

3. The change of phase from solid to vapor occurs in a single step governed by an Arrhenius reaction rate. During nanosecond pulse irradiation the thickness of the heat affected zone is of the order of 1 μm; thus, if a liquid layer exists it must be entirely too thin to modify the results [17].

4. The evaporated material escapes from the hole and does not interfere with the incoming laser beam, i.e., the optical thickness of the plume is very small at the laser wavelength and during laser on time. Assuming all material to escape is perhaps the weakest assumption: While plasma formation may be limited for nanosecond pulses, it is quite possible that some of the evaporated material cannot escape from the large aspect ratio hole, but recondenses higher up at the hole wall.

5. Heat losses by convection and reradiation are negligible (as compared to change-of-phase and conduction losses). This has been shown to be accurate even for continuous wave (CW) laser operation [18].

6. At the surface, the laser beam is reflected in a quasispecular fashion, i.e., reflection is within a small cone centered around the specular direction, with clipped-Gaussian intensity away from the specular direction. This serves two purposes: (i) no real surface can be expected to be optically smooth, and (ii) specular reflection from an axisymmetric hole wall would lead to the focussing of an irradiation ring down to a single point, which is not possible and would lead to infinite absorbed intensity at the hole center. Surface reflectance is evaluated from Fresnel's relations, based on the material's complex index of refraction m [19].

7. The laser beam is unpolarized, resulting in an axisymmetric (rather than a three-dimensional) hole. In fact, most lasers are circularly polarized, which, at the first reflection, behave exactly like an unpolarized beam. However, after several reflections (as are bound to occur in deep holes) the beam becomes partially polarized, resulting in seriously noncircular hole cross sections [20]. It was shown that trepanning (rotating work piece against laser during drilling) results in essentially perfectly round holes, which is assumed here.

The laser beam is assumed to have a spatially Gaussian power distribution with an effective radius $w(z)$, diverging away from the waist, given by Kogelnik and Li [21] as

$$w^2(z) = w_0^2 + \beta_\infty^2 (z_0 + z)^2; \quad \beta_\infty = M^2 \frac{\lambda}{\pi w_0} \quad (1)$$

where w_0 is the $1/e^2$ radius at the focal plane (the waist) through which 86.5% (or $1 - e^{-2}$) of the beam's energy passes, β_∞ is the far-field beam divergence angle, z_0 is the distance between the focal plane of the lens and the material surface ($z_0 > 0$ for focal point above surface, and $z_0 < 0$ for focal point below surface), M^2 is a beam-quality factor ($M^2 = 1$ for a Gaussian laser), and λ is the wavelength of the laser (cf. Fig. 1).

A Gaussian beam has a plane wave front at the waist (i.e., a radius of curvature of infinity) but assumes a curved wave front away from the beam waist. The radius of curvature of the wave front is

$$r_c(z) = (z_0 + z) \left[1 + \frac{(w_0/\beta_\infty)^2}{(z_0 + z)^2} \right] \quad (2)$$

If the laser beam is visualized as consisting of a bundle of rays, then the ray direction is normal to the wave front surface having curvature $r_c(z)$. The direct radiation from the laser may then be expressed as
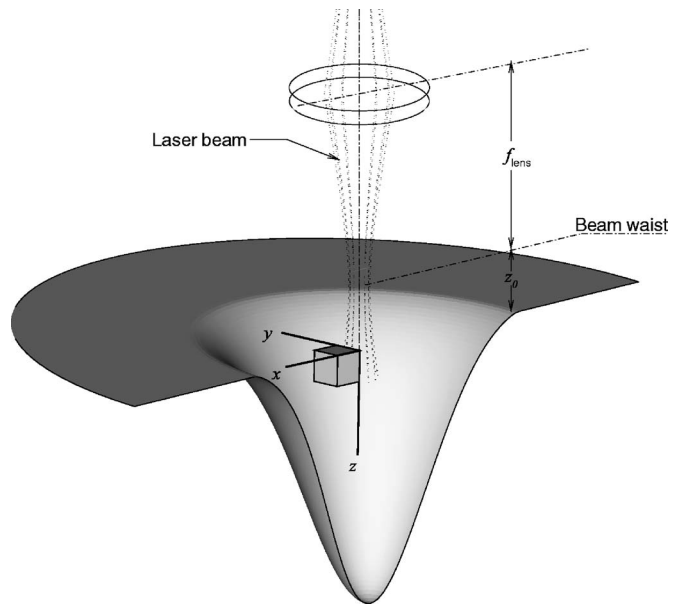
Fig. 1 Laser drilling setup and coordinate system

$$\mathbf{F}(x,y,z,t) = \left[\frac{w_0^2}{w^2(z)} \right] F_0(t) e^{-2(x^2+y^2)/w^2(z)} \frac{\hat{\mathbf{s}}(x,y)}{\hat{\mathbf{s}} \cdot \hat{\mathbf{k}}} \quad (3)$$

$$\frac{\hat{\mathbf{s}}}{\hat{\mathbf{s}} \cdot \hat{\mathbf{k}}} = \frac{x\hat{\mathbf{i}} + y\hat{\mathbf{j}}}{\sqrt{r_c^2(z) - x^2 - y^2}} + \hat{\mathbf{k}} = \tan \gamma \cos \phi \hat{\mathbf{i}} + \tan \gamma \sin \phi \hat{\mathbf{j}} + \hat{\mathbf{k}}$$

where γ is the angle between a laser ray and the z axis, ϕ is the azimuthal angle for the ray measured from the x axis in the x-y plane, and $F_0(t)$ is the flux density at the beam center at the focal plane. Since, in the present simulation, an axisymmetric hole is considered, only beams lying in the x-z plane, or $\phi=0$, need to be considered.

Two different temporal beam profiles are considered, viz., a simple top-hat profile and a clipped Gaussian profile (which fairly accurately represents the temporal power distribution of a Q-switched Nd-YAG laser)

$$\text{Top hat:} \quad F_0(t) = \begin{cases} F_{0\text{av}}, & 0 < t < t_{pw}, \\ 0 & t_{pw} < t < t_p, \end{cases} \quad (4a)$$

$$\text{Clipped Gaussian:} \quad F_0(t) = \begin{cases} F_{0\text{av}} a \dfrac{t}{t_{pw}} e^{-b(t/t_{pw})^2}, & 0 < t < 2t_{pw} \\ 0 & 2t_{pw} < t < t_p \end{cases}$$

$$(4b)$$

where t_{pw} is the nominal pulse width (on-time with power levels larger than 50% of maximum, and t_p is the total pulse duration (i.e., on and off time); $F_{0\text{av}}$ is the average flux density at the beam center over the nominal pulse width, i.e.,

$$F_{0\text{av}} = \frac{1}{t_{pw}} \int_0^{t_p} F_0(t) dt \quad (5)$$

For the clipped Gaussian profile the constants a and b are found from Eq. (5) and the condition that the pulse width at half of maximum is t_{pw}. Both profiles are compared in Fig. 2 (with the top-hat profile shifted for better comparison).

Following Modest [22,23] the transient heat conduction equation for a thick solid irradiated by a laser beam, and its auxiliary conditions, may be written in nondimensional form as (slightly

Fig. 2 Temporal laser pulse shapes: top-hat and clipped-Gaussian profiles

Fig. 3 Specular reflection cone with Gaussian decay of strength away from specular direction

modified for the present drilling scenario using a short-pulsed laser)

$$\frac{\partial \theta}{\partial t} = \text{Fo}\boldsymbol{\nabla}^2\theta \tag{6}$$

Initial condition:

$$t = 0: \quad \theta(r,z,0) = 0$$

$$s(r,0) = s_0(r) \tag{7}$$

Boundary conditions:

$$r,z \rightarrow +\infty: \quad \theta = 0; \quad z = s(r,t): \quad Q_n = N_c(V_n\text{Ste} - \text{Fo}\hat{\mathbf{n}} \cdot \nabla\theta) \tag{8}$$

Ablation condition:

$$z = s(r,t): \quad V_n = C_1 e^{C_2[1 - T_{\text{re}}/T(\theta)]} \tag{9}$$

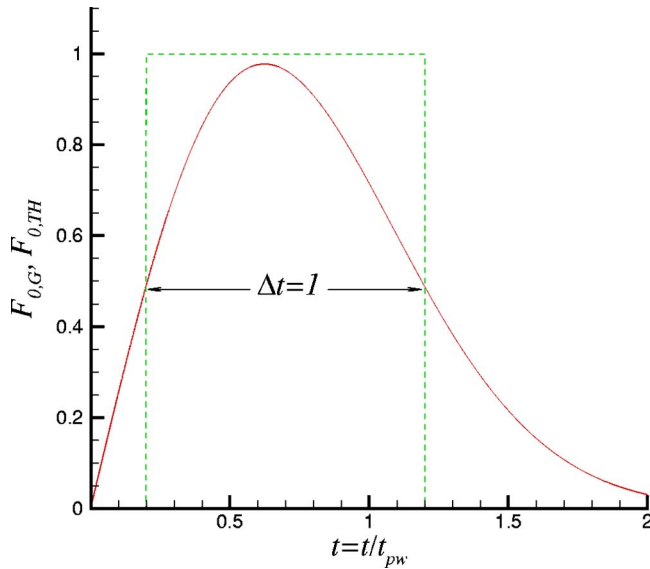with

$$r = \bar{r}/w_0; \quad z = \bar{z}/w_0; \quad s = \bar{s}/w_0; \quad t = \bar{t}/t_{pw}$$

$$\theta = \frac{T - T_\infty}{T_{\text{re}} - T_\infty}; \quad Q_n = \frac{\bar{Q}_n}{F_{0\text{av}}} = \frac{\alpha\mathbf{F}_0 \cdot \hat{\mathbf{n}} + \bar{Q}_{\text{ref}}}{F_{0\text{av}}}; \quad \text{Ste} = \frac{\Delta h_{\text{re}}}{c_p(T_{\text{re}} - T_\infty)} \tag{10}$$

$$V_n = \frac{\bar{V}_n t_{pw}}{w_0}; \quad \text{Fo} = \frac{\alpha_H t_{pw}}{w_o^2}; \quad N_c = \frac{\rho c(T_{\text{re}} - T_\infty)w_0}{F_{0\text{av}}t_{pw}}$$

Here $\bar{r}$, $\bar{z}$, and $\bar{s}$ are dimensional coordinates and groove depth (with the overbar shown only here in Eq. (10) to distinguish them from nondimensional ones), which are then nondimensionalized with the beam radius at the focal point, w_0; $\alpha_H = k/\rho c$ is the thermal diffusivity of the material, T_{re} is the equilibrium ablation (or "removal" temperature), and Δh_{re} is the energy required to remove material ("heat of removal", i.e., heat of vaporization in the simplest case). The parameter V_n is a nondimensional (transient) surface recession velocity (by ablation), Fo is the ratio between pulse duration and heat diffusion time, N_c approximates the ratio of sensible heating (due to conduction) and the incoming laser flux; and Ste is the Stefan number that compares ablation energy with sensible heat. The nondimensional absorbed radiation Q_n consists of two parts: a direct irradiation part (with α being the

directional surface absorptance as calculated from Fresnel's relations [24]), and the contribution from multiple reflections, evaluated from ray tracing.

The boundary condition at the top surface, $z = s(r,t)$, specifies that absorbed laser irradiation is used up by conduction losses and by the latent heat required to ablate material. The ablation velocity (normal to the surface) is governed by a simple reaction equation of the Arrhenius type [22].

3 Solution Approach

Equation (6) with its auxiliary conditions (7)–(9) forms a complete set of dimensionless equations in transient form to predict the forming hole shape $s(r,t)$ and temperature field $\theta(r,z,t)$. In order to find a simple, approximate solution for the conduction loss Eq. (6), we will assume that conduction takes place only in the direction of the (local) surface normal, i.e., the loss is locally one dimensional. Transforming coordinates to n, a nondimensional distance from a surface location pointing into the medium along the local surface normal (see Fig. 3), the solid will move through the origin for n with the ablation velocity V_n into the negative n direction. Thus, Eq. (6) transforms to

$$\frac{\partial \theta}{\partial t} - V_n\frac{\partial \theta}{\partial n} = \text{Fo}\frac{\partial^2 \theta}{\partial n^2} \tag{11}$$

Equation (11) is solved iteratively together with Eqs. (8) and (9) by a simple finite difference scheme (using a tridiagonal solver). In order to do so the local heat affected zone thickness must be (conservatively) estimated: Each pulse consists of a warmup period (without evaporation), during which time $n_{\max}$ is estimated as (conduction into a wall subject to constant flux)

$$n_{\max}(r,t) = 6\sqrt{\text{Fo}\,t} \tag{12}$$

and an evaporation period, during which $n_{\max}$ remains essentially constant as (quasisteady-state conduction with all energy going into ablation)

$$n_{\max}(r,t) = 6\text{Fo}(1 + \text{Ste})N_c/Q_n \tag{13}$$

For this Eq. (11) is transformed one more time, using a new independent normal variable $\nu = n/n_{max}$, leading to

$$\frac{\partial \theta}{\partial t} - \left(\frac{V_n}{n_{max}} + \frac{\partial}{\partial t} \ln n_{max}\right)\frac{\partial \theta}{\partial \nu} = \frac{Fo}{n_{max}^2}\frac{\partial^2 \theta}{\partial \nu^2} \tag{14}$$

$$z = s(r,t)[\nu = 0]: \frac{Q_n}{N_c} = V_n \text{Ste} - \frac{Fo}{n_{max}}\frac{\partial \theta}{\partial \nu} \tag{15}$$

For the finite difference solution values for θ and n_{max} must be stored for one time step.

Although the laser beam has a Gaussian intensity distribution and is best described by the wave equation [25], it may also be approximated as rays. The reflected direct irradiation is also assumed to consist of rays traveling along a straight path, i.e., the near-field phase information from wave theory is neglected. The total irradiation on the groove surface is then obtained by adding the direct irradiation calculated using Eq. (3) and the irradiation from multiple reflections. The reflected beam direction and the absorbed energy strongly depend on the surface normal and, therefore, on the method of surface representation. In the current study the hole is considered to be axisymmetric and, thus, only reflections in the x-z-plane are considered. The hole surface is described by two parametric cubic splines, each containing N_p (r,l) and (z,l) pairs, respectively, where l is arc length along the hole's surface. Such splines not only guarantee second-derivative continuity, but also allow multivalued (r,z) pairs. The latter is important for holes which have "throats," (i.e., the diameter below the surface is not monotonically decreasing, and there may be multiple z values for a given r), as may occur due to focussing effects of multiple reflections. The spline knots (l positions) are updated periodically to keep them roughly equidistant (in arc length) even in very deep holes. This is always done between pulses, when the material is cold (and updating the internal temperature is not an issue). As indicated earlier, the interaction between irradiation and surface recession is extremely strong, with the slightest variation in hole shape causing great changes in reflected intensity and vice versa. Thus, to avoid statistical scatter, ray tracing with nonrandom emission positions has been used in the current study (rather than a standard Monte Carlo method). Several schemes were tried to simulate truly specular reflections, such as focussing/defocussing of an infinitesimal beam by a curved surface, collection of reflected rays by individual cells, etc. All these attempts resulted in unstable situations, as they should: For a deep hole there will always be a ray reflected off the side wall that hits the center of the bottom. In an axisymmetric hole this implies focussing a ring onto a point, resulting in infinite intensity and violating the laws of optics. It was, therefore, decided that the surface should not be treated as perfectly specular, but only preferentially specular, with Gaussian decay away from the specular direction $\hat{\mathbf{s}}_{sp}$. Consider a beam undergoing a first reflection, and then travelling on and hitting the hole wall again after a distance l, as shown in Fig. 3. If the beam carries a total energy of Q_b, then the beam energy density in a plane normal to $\hat{\mathbf{s}}_{sp}$ at location l is assumed to be

$$q_b e^{-(\rho/d)^2} \tag{16}$$

where ρ is radial distance away from the ray center in a plane normal to $\hat{\mathbf{s}}_{sp}$, and

$$Q_b = q_b \int_A e^{-(\rho/d)^2} 2\pi\rho d\rho = \pi d^2 q_b \tag{17}$$

(where the second equality only holds if no beam shading occurs, which happens in deep grooves after several reflections and/or with too large cone angles). Thus, lines of equal intensity form circles in a plane normal to $\hat{\mathbf{s}}_{sp}$, an ellipse in a plane oblique to $\hat{\mathbf{s}}_{sp}$, and a somewhat more complicated shape in a rotationally symmetric curved body (drill hole). This two-dimensional distri-

bution of reflected energy needs to be accounted for. With some algebra and assuming, for a small cone angle θ_{sp}, that all beams in the cone hitting the hole wall travel parallel to $\hat{\mathbf{s}}_{sp} = s_x\hat{\mathbf{i}} + s_z\hat{\mathbf{k}}$, the cone radius ρ for a point on the wall $\mathbf{r} = r\cos\phi\hat{\mathbf{i}} + r\sin\phi\hat{\mathbf{j}} + z\hat{\mathbf{k}}$, with respect to a beam hitting point $\mathbf{r}_0 = r_0\hat{\mathbf{i}} + z_0\hat{\mathbf{k}}$ in the x-z-plane is determined from

$$\rho^2 = |\mathbf{r} - \mathbf{r}_0 - [(\mathbf{r} - \mathbf{r}_0)\cdot\hat{\mathbf{s}}]\hat{\mathbf{s}}|^2$$
$$= [(r\cos\phi - r_0)s_z - (z - z_0)s_x]^2 + r^2\sin^2\phi \tag{18}$$

and

$$Q_b = 2q_b \int_u \int_0^{\phi_{max}} e^{-(\rho/d)^2}(\hat{\mathbf{s}}_{sp}\cdot\hat{\mathbf{n}})d\phi\, r\, du \tag{19}$$

where $\hat{\mathbf{n}}$ is the outward surface normal at $\mathbf{r}$, and u is the arc length along the hole wall in the x-z plane. The value of the maximum azimuthal angle is $\phi_{max} = \pi$, unless the beam hits the surface from behind ($\hat{\mathbf{s}}_{sp}\cdot\hat{\mathbf{n}} < 0$), in which case ϕ_{max} is determined from

$$\hat{\mathbf{s}}_{sp}\cdot\hat{\mathbf{n}} = s_x n_x\cos\phi + s_z n_z = 0 \tag{20}$$

where $(n_x, 0, n_z)$ is the surface normal at (r,z) in the x-z-plane. The double integral in Eq. (19) is evaluated numerically for each (r,z) cell by using a simple Newton-Cotes scheme in the u direction (x-z plane), and a five-point Gaussian scheme [$\rho(\phi_{max}) < 3.5d$] or a five-point Hermitian scheme for the azimuthal integration. To conserve energy in the case of partial blockage, the q_b is adjusted so that all contributions add up to Q_b.

4 Results and Discussion

To investigate the effects of multiple specular reflections on drilling rates and hole geometry, a "typical" Nd-YAG laser, such as the Photonics Industries DS10-355 in our laboratory with a Gaussian temporal pulse profile similar to the one shown in Fig. 1 with $t_{pw} = 4$ ns was used which has an $M^2 = 1.88$ and, after focussing through a 50 mm lens onto the surface of the workpiece, has a beam radius of $w_0 = 13$ μm. For most runs a pulse energy of $E_p = 100$ μJ was employed. As a representative material
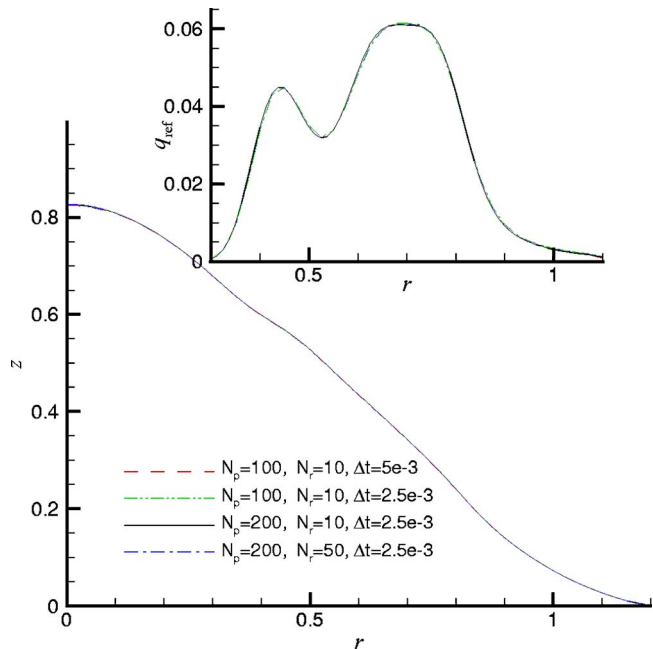

Fig. 4 Influence of time step, cell number, and ray number on accuracy of solution during nanosecond laser drilling
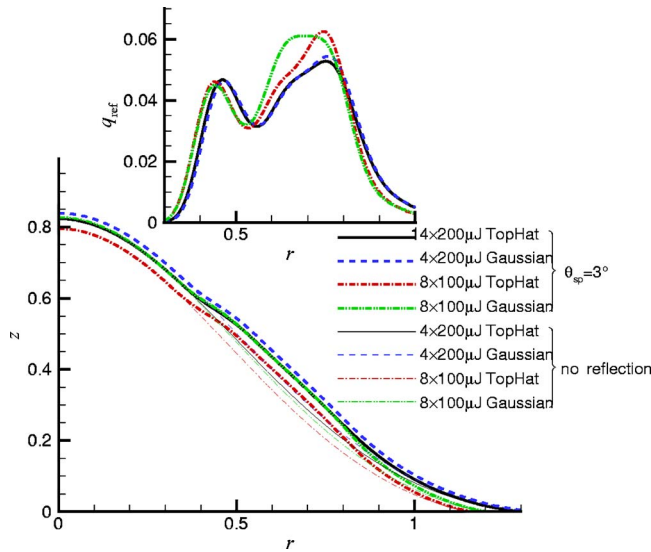
Fig. 5 Influence of pulse strength and pulse temporal shape on conduction losses during nanosecond laser drilling

Fig. 6 Influence of specularity angle on distribution of reflected energy inside laser-drilled hole

aluminum was chosen ($\alpha_H = 6.2 \times 10^{-5}$ m²/s, $\rho = 2700$ kg/m³, $c = 1200$ J/kg K, $\Delta h_{re} = 13$ MJ/kg). This results in the following nondimensional parameters:

$$\beta_\infty = 0.036; \quad \text{Ste} = 4.71, \quad \text{Fo} = 1.5 \times 10^{-3}, \quad N_c = 0.250 \tag{21}$$

For the most part a refractive index of $m = 1.4 - 5.619i$ was used (with a normal absorptance of $\alpha_n = 0.15$) but, to test the influence of absorptance, values of $m = 1.4 - 7.088i$ ($\alpha_n = 0.10$) and $m = 1.4 - 4.716i$ ($\alpha_n = 0.20$) were also employed.

Figure 4 shows the sensitivity of the solution to the number of radial nodes (N_p), number of rays traced per node (N_r) and the time step, all for eight 100 μJ pulses, i.e., at a time when multiple reflections just start appearing, carving out a small bulge near $r = 0.5$ (here shown for a specular cone angle of $\theta_{sp} = 3$ deg). Also shown is the absorbed irradiation due to multiple reflections. It is seen that $N_p = 100$, $N_r = 10$, and $\Delta t = 5 \times 10^{-3}$ is more than adequate. However, it was observed that, for very deep holes, cumulative effects from a too large time step also occurred, also causing worse instabilities. Thus, once the hole reached a certain depth, a much smaller time step of $\Delta t = 5 \times 10^{-4}$ was generally employed. In addition, $N_p = 200$ was used for deep holes to assure that nodes remained reasonably close together.

A similar sensitivity study is shown in Fig. 5, here studying the importance of conduction. It is seen that, whether four pulses at 200 μJ or eight pulses at 100 μJ are used makes relatively little difference: The four-pulse holes are slightly deeper (less conduction loss), but the difference is small, since conduction losses are small. Similarly, the differences between top-hat and Gaussian temporal profiles are small, with the Gaussian profile incurring larger conduction losses (since total pulse on time is doubled). The differences are more pronounced in the reflected radiation profiles, indicating their sensitivity to the slightest variation in hole shape.

This sensitivity can be clearly observed in Fig. 6, which shows the absorbed specular reflections for one of the previous holes (eight pulses at 100 μJ, top-hat laser profile) as a function of specular reflection angle θ_{sp}. For $\theta_{sp} = 3$ deg two mild peaks are observed, but more and sharper peaks appear for smaller θ_{sp}. Such sharp irradiation peaks burrow small dimples into the hole walls, making the problem unstable within a few time steps. In Fig. 6 for all cases the same hole geometry was used (obtained by using $\theta_{sp} = 3$ deg for drilling up to that point). If, say, $\theta_{sp} = 0.5$ deg were
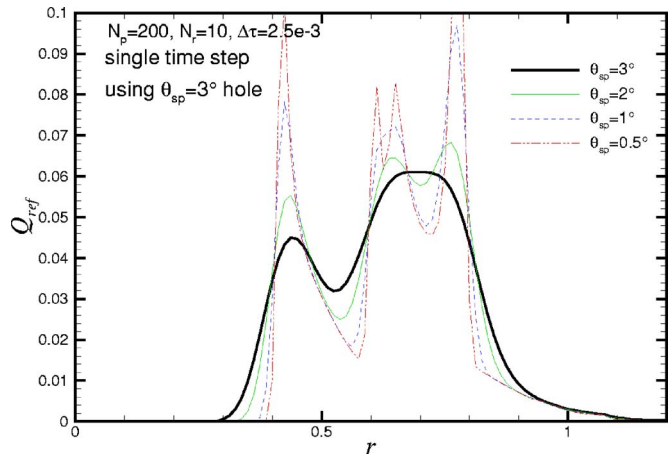
taken from the beginning, the hole shape would still look essentially identical (little ablation due to reflected energy up to that point); but the reflected energy picture already resembles random noise (not shown). Therefore, any deep drilling with a $\theta_{sp} \lesssim 5$ deg should be expected to become unstable.

During drilling, as the hole gets deeper, multiple reflections first become important when the hole depth reaches about $s(0) \approx 0.7$, with reflections hitting a ring with $r \gtrsim 0.5$ (with intensity decreasing toward $r = 1$ because of the grazing angles). This is shown in Fig. 7 for $\theta_{sp} = 5$ deg, and three typical ray paths are depicted in Fig. 7(a) for a shallow hole, showing that no beam undergoes more than one reflection. As the hole gets deeper multiple reflection intensity peaks occur, depending on specularity angle θ_{sp}; for $\theta_{sp} \gtrsim 5$ deg generally only two peaks are present as long as the hole shape is relatively smooth; still no beam undergoes more
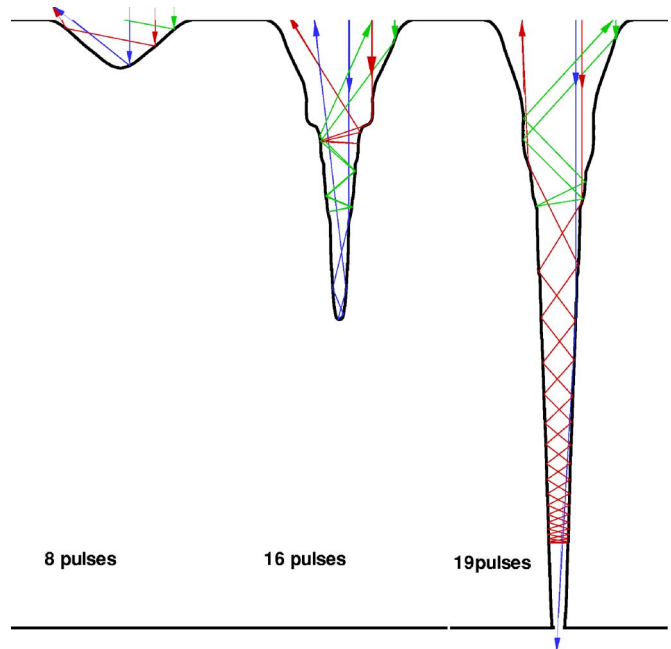
Fig. 7 Typical laser ray reflection paths for varying hole depths (Ste=4.71, N_c=0.25, α_n=0.15, θ_{sp}=5 deg); (a) hole after eight pulses (without reflection effects), (b) after 16 pulses (with bulging due to reflections), and (c) after burn-through (with strong beam trapping). Dimensions are to scale (with a material thickness of 10w_0).
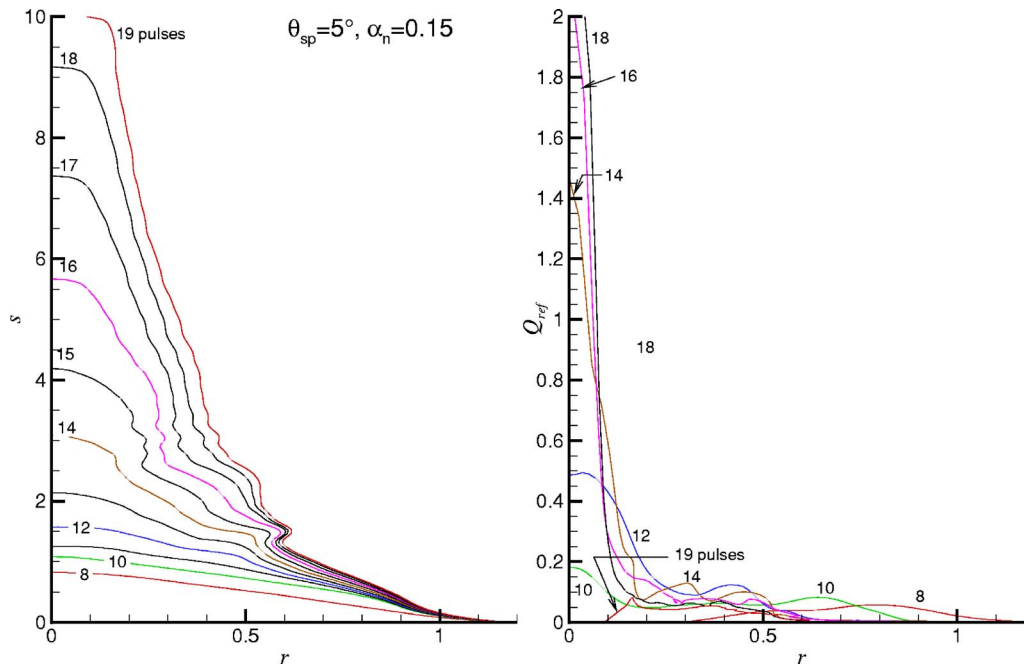
Fig. 8 Hole development as a function of pulse number (Ste=4.71, N_c=0.25, α_n=0.15, θ_{sp}=5 deg); left: hole cross section, right: absorbed, multiple-reflected laser power

than a single reflection. As the hole gets even deeper, the peak at the smaller r grows and slowly moves toward $r=0$; now a number of beams start to have two and more internal reflections. Eventually, once the hole depth exceeds $s(0) \geqslant 1.5-2$, the reflection peak at the center becomes totally dominant, due to beams channeling near the center, as seen in Fig. 7(b). Once the hole depth exceeds $s(0) \simeq 5$ the *average* number of reflections that a beam undergoes can be ten or more. However, once the material is punched through, many beams escape through the bottom after only one or two reflections (Fig. 7(c)), and the reflection peak vanishes.

As a typical example of the drilling process, Fig. 8 shows the pulse-to-pulse progress of drilling through a $10w_0=130$ μm thick aluminum film with a 100 μJ laser, using a specularity angle of $\theta_{sp}=5°$, and an index of refraction of $m=1.4-5.619i$ ($\alpha_n=0.15$). Shown are hole cross sections (left frame) and the locally absorbed amount of reflected radiation, both in nondimensional form and at the end of the (top-hat profile) pulse. As indicated earlier, multiple reflections do not occur until the seventh pulse (small single peak), and start becoming important after eight pulses (also depicted in Fig. 7(a)), showing two peaks (somewhat less pronounced than the $\theta_{sp}=3$ deg case shown in Fig. 5. As the hole deepens, a third peak appears and all peaks move toward the hole center until, after 12 pulses, the center peak dominates; however, a second, off-center peak remains for the duration of the drilling process, causing a bulge in the hole profile around $r \simeq 0.5$. Note how, after about 15 pulses beam channeling at the beam center becomes so strong to dramatically increase the drill rate per pulse. After 18 pulses a maximum Q_{ref} (= absorbed portion of multiple-reflected energy) of $2.6 \simeq 17\alpha_n$ is reached (i.e., 17 times the maximum possible direct irradiation), gradually decaying again for subsequent pulses. After burn-through, Q_{ref} drops dramatically, since a large portion of reflected beams escape through the hole, Fig. 7(c). In order to keep the calculations stable Q_{ref} was smoothened with a least-squares-error spline whenever it displayed more than three off-center maxima. However, the smoothing range was limited to small changes, i.e., the smoothing would break down in the presence of four or more *large* off-center peaks. Under this scenario slight surface roughness still can occur and, for that reason, ablation velocity V_n was also smoothened (subject to the

same conditions as Q_{ref}). This avoided smoothing of the hole shape itself (and thus avoiding possible negative ablation).

Comparison with experiment shows that the drill rate does, indeed, increase with hole depth (or material thickness) [26]. However, while Fig. 8 (as well as all later figures) shows a doubling of the drill rate between the thicknesses of four (52 μm) and ten (130 μm), experiments for aluminum saw only an increase of about 30%; for stainless steel, however, a doubling of the drill rate was observed. Experiment also showed that, for thicker specimen the drill rate decreases again (apparently due to decreased focussing and increased conduction losses) [26].

Figure 9 shows that changing the specularity angle to $\theta_{sp} = 10$ deg makes relatively little difference: The third peak after ten pulses is washed out, and the general amplification near the beam center is somewhat lower. Therefore, 20 pulses rather than 19 are required for drill through. Lowering the specularity angle to $\theta_{sp} = 3$ deg, as shown in Fig. 10, also seems to yield little qualitative change, simply accentuating peaks and bringing down the drill time to 18 pulses. However, the off-center peaks tried to dig a secondary hole maximum near $r \simeq 0.1$, which required a little smoothening of the hole shape itself. Thus, not surprisingly, smaller specularity angles cause greater instabilities that can lead to runaway if not smoothed (as frequently observed in experiments). Differences were also small due to the fact that no beam was allowed to focus to a spot smaller than $0.1w_0$, making many reflections in deep holes (with short distances between reflections) independent of small values for θ_{sp}.

Finally, Figs. 11 and 12 show the effects of raising or lowering the normal absorptance to 0.20 and 0.10, respectively. As expected, raising absorptance increases drill rates (16 pulse for drill through), while lowering α_n decreases it (26 pulses). Qualitatively, the drilling behavior is the same for all three absorptances.

Summary and Conclusions

A two-dimensional, axisymmetric model was established to study the effects of multiple, specular reflections during evaporative drilling with nanosecond laser pulses. To avoid the physically impossible focussing to a point, specularity was defined as reflec-

Fig. 9 Hole development as a function of pulse number (Ste=4.71, N_c=0.25, α_n=0.15, θ_{sp}=10 deg); left: hole cross section, right: absorbed, multiple-reflected laser power

tion into a narrow cone with Gaussian decay away from the specular direction. Results show that, independent of specularity level, reflected energy levels exhibit a dual peak profile, with a peak at the beam center becoming more and more dominant as the hole gets deeper. Thus holes are formed which (i) exhibit a bulge away from the hole center, and (ii) totally absorb the laser's energy at the beam center due to many reflections, resulting in strong increases in drilling rates.

Acknowledgment

Nomenclature

c = specific heat

Fig. 10 Hole development as a function of pulse number (Ste=4.71, N_c=0.25, α_n=0.15, θ_{sp}=3 deg); left: hole cross section, right: absorbed, multiple-reflected laser power
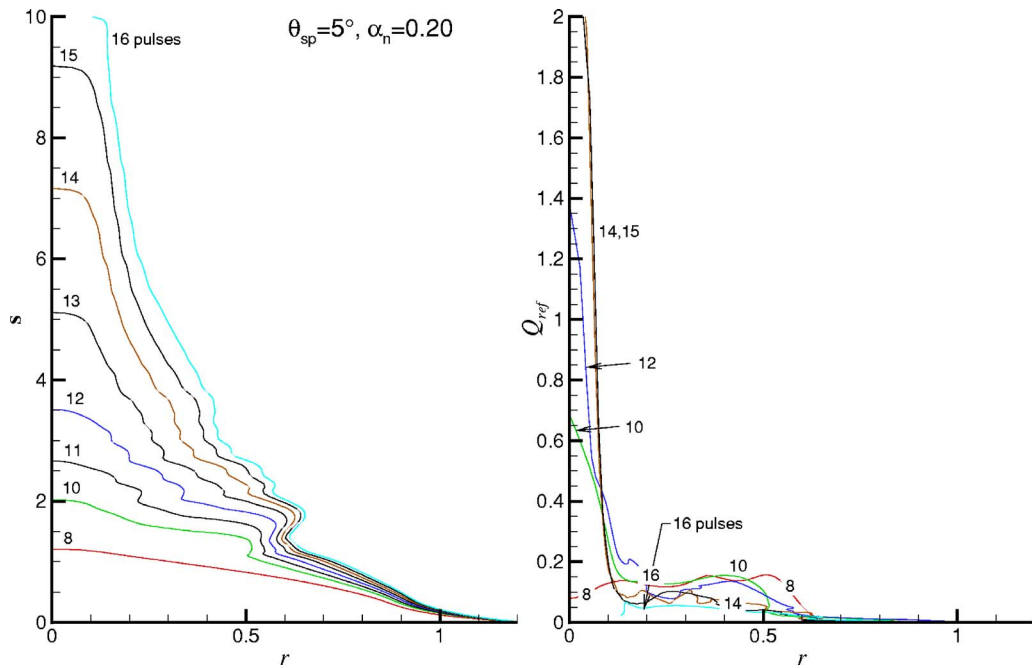
Fig. 11 Hole development as a function of pulse number ($Ste=4.71, N_c=0.25, \alpha_n=0.20, \theta_{sp}=5$ deg); left: hole cross section, right: absorbed, multiple-reflected laser power

C_1, C_2 = constants in Arrhenius relation
d = specular reflection cone diameter
E_p = total pulse energy
$\mathbf{F}$ = irradiation flux vector
F_0 = radiation flux density at center of beam at focal plane
Fo = Fourier number (pulse time-to-diffusion time ratio)
Δh_{re} = "heat of removal" (enthalpy of evaporation)
$\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ = unit vector in x, y, and z directions

k = thermal conductivity
$\hat{\mathbf{n}}$ = unit surface normal
n = distance along surface normal, measured from surface
m = complex index of refraction, $m=n-ik$
M^2 = beam quality parameters
N_c = sensible heat-to-laser power parameter
Q_n = (dimensionless) absorbed irradiation flux at surface
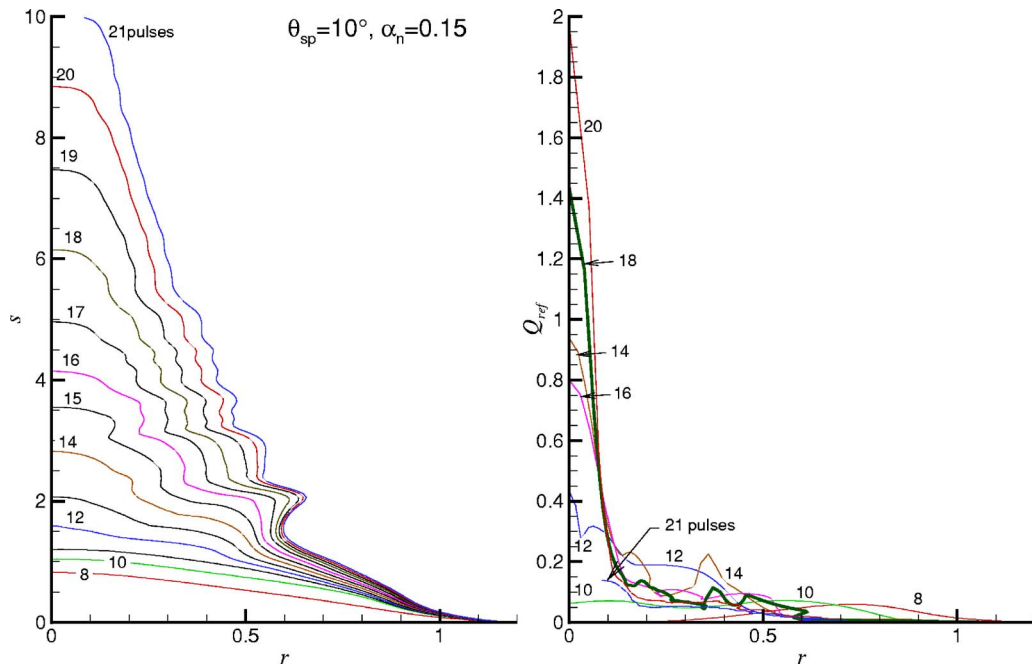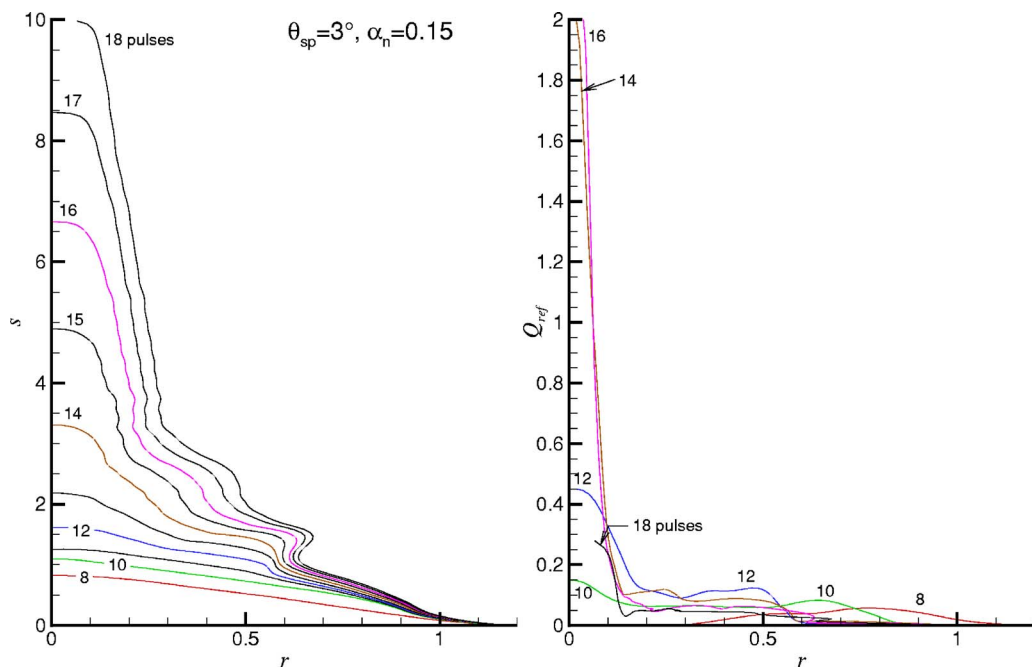r = dimensionless radial coordinate

Fig. 12 Hole development as a function of pulse number ($Ste=4.71, N_c=0.25, \alpha_n=0.10, \theta_{sp}=5$ deg); left: hole cross section, right: absorbed, multiple-reflected laser power
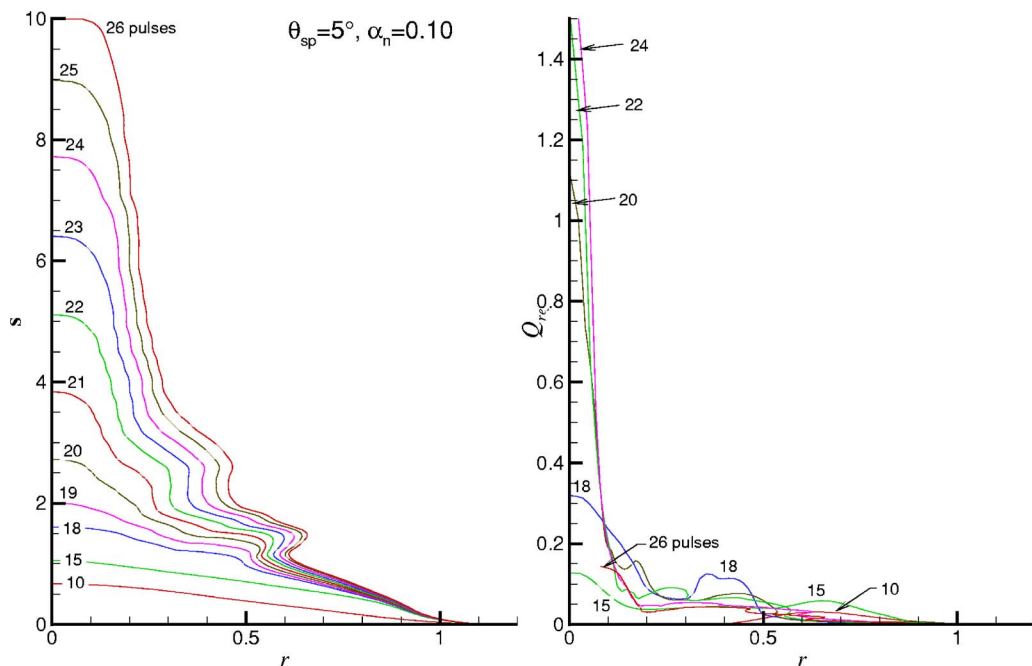
$s(r,t)$ = (dimensionless) groove depth

$\hat{\mathbf{s}}$ = unit direction vector

Ste = Stefan number (ablation energy-to-sensible heat ratio)

t = (dimensionless) time

T = temperature

T_{re} = evaporation (or decomposition) temperature

V_n = (dimensionless) ablation velocity at surface

w, w_0 = $1/e^2$ radius of laser beam (at focal plane)

x, y, z = (dimensionless) Cartesian coordinates

α_H = thermal diffusivity

α = local effective absorptance at laser wavelength and incidence direction

β_∞ = far-field beam divergence

λ = wavelength of laser radiation

ρ = density of the medium, or radial distance from specular direction

θ = dimensionless temperature

θ_{sp} = opening angle of specular reflection cone

ϕ = azimuthal angle ($x = r \cos \phi$, $y = r \sin \phi$)

Subscripts

n = normal incidence

re = evaluated at evaporation (or decomposition) temperature

0 = at focal plane

∞ = evaluated at ambient conditions, or located far away

References

[1] Dabby, F. W., and Paek, U.-C., 1972, "High-Intensity Laser-Induced Vaporization and Explosion of Solid Material," IEEE J. Quantum Electron., **QE-8**, pp. 106–111.

[2] Wagner, R. E., 1974, "Laser Drilling Mechanics," J. Appl. Phys., **45**, pp. 4631–4637.

[3] von Allmen, M., 1976, "Laser Drilling Velocity in Metals," J. Appl. Phys., **47**, pp. 5460–5463.

[4] Petring, D., Abels, P., and Beyer, E., 1988, "Absorption Distribution on Idealized Cutting Front Geometries and its Significance for Laser Beam Cutting," *High Power* CO₂ *Laser Systems and Applications*, SPIE, Bellingham, WA, pp. 123–131.

[5] Ramanathan, S., and Modest, M. F., 1992, "CW Laser Drilling of Composite Ceramics," *Proceedings of ICALEO '91, Laser Materials Processing*, San Jose, CA, Laser Institute of America, Orlando, FL, Vol. 74, pp. 305–326.

[6] Milewski, J., and Sklar, E., 1996, "Modelling and Validation of Multiple Reflections for Enhanced Laser Welding," Modell. Simul. Mater. Sci. Eng., **4**, pp. 305–322.

[7] Lee, J. Y., Ko, S. H., Farson, D. F., and Yoo, C. D., 2002, "Mechanism of Keyhole Formation and Stability in Stationary Laser Welding," J. Phys. D, **35**, pp. 1570–1576.

[8] Ho, C. Y., and Wei, P. S., 2001, "Absorption in a Paraboloid of Revolution-Shaped Welding or Drilling Cavity Irradiated by a Polarized Laser Beam," Metall. Mater. Trans. B, **32B**, pp. 603–614.

[9] Ho, C. Y., 2004, "Effects of Polarizations of a Laser on Absorption in a Paraboloid of Revolution-Shaped Welding or Drilling Cavity," J. Appl. Phys., **96**(10), pp. 5393–5401.

[10] Ki, H., Mohanty, P. S., and Mazumder, J., 2002, "Multiple Reflection and its Influence on Keyhole Evolution," J. Laser Appl., **14**(1), pp. 39–45.

[11] Solana, P., and Negro, G., 1997, "A Study of the Effect of Multiple Reflections on the Shape of the Keyhole in the Laser Professing of Materials," J. Phys. D, **30**, pp. 3216–3222.

[12] Bang, S. Y., and Modest, M. F., 1992, "Evaporative Scribing With a Moving CW Laser—Effects of Multiple Reflections and Beam Polarization," *Proceedings of ICALEO '91, Laser Materials Processing*, San Jose, CA, Laser Institute of America, Orlando, FL, Vol. 74, pp. 288–304.

[13] Bang, S. Y., and Modest, M. F., 1991, "Multiple Reflection Effects on Evaporative Cutting With a Moving CW Laser," J. Heat Transfer, **113**(3), pp. 663–669.

[14] Bang, S. Y., Roy, S., and Modest, M. F., 1993, "CW Laser Machining of Hard Ceramics—Part II: Effects of Multiple Reflections," Int. J. Heat Mass Transfer, **36**(14), pp. 3529–3540.

[15] Ki, H., Mohanty, P. S., and Mazumder, J., 2002, "Modeling of Laser Keyhole Welding: Part I. Mathematical Modeling, Numerical Methodology, Role of Recoil Pressure, Multiple Reflections, and Free Surface Evolution," Metall. Mater. Trans. A, **33A**, pp. 1817–1830.

[16] Ki, H., Mohanty, P. S., and Mazumder, J., 2002, "Modeling of Laser Keyhole Welding: Part II. Simulation of Keyhole Evolution, Velocity, Temperature Profile, and Experimental Verification," Metall. Mater. Trans. A, **33A**, pp. 1831–1842.

[17] Abakians, H., and Modest, M. F., 1988, "Evaporative Cutting of a Semi-Transparent Body With a Moving CW Laser," J. Heat Transfer, **110**, pp. 924–930.

[18] Modest, M. F., and Abakians, H., 1986, "Evaporative Cutting of a Semi-Infinite Body With a Moving CW Laser," J. Heat Transfer, **108**, pp. 602–607.

[19] Modest, M. F., 1993, *Radiative Heat Transfer*, McGraw-Hill, New York.

[20] Nolte, S., Momma, C., Kamlage, G., Ostendorf, A., Fallnich, C., Alvensleben, F., and Welling, H., 1999, "Polarization Effects in Ultrashort-Pulse Laser Drilling," Appl. Phys. A, **68**, pp. 563–567.

[21] Kogelnik, H., and Li, T., 1966, "Laser Beams and Resonators," Appl. Opt., **5**(10), pp. 1550–1567.

[22] Modest, M. F., 1996, "Three-Dimensional, Transient Model for Laser Machining of Ablating/Decomposing Materials," Int. J. Heat Mass Transfer, **39**(2), pp. 221–234.

[23] Modest, M. F., 1996, "Transient Model for CW and Pulsed Laser Machining of Ablating/Decomposing Materials—Approximate Analysis," J. Heat Transfer, **118**(3), pp. 774–780.

[24] Modest, M. F., 2003, *Radiative Heat Transfer*, 2nd ed., Academic Press, New York.

[25] Verdeyen, J. T., 1981, *Laser Electronics*, Prentice-Hall, Englewood Cliffs, NJ.

[26] Semak, V. V., Thomas, J. G., and Campbell, B. R., 2005, "Comparison of Drilling Rates and Material Removal Dynamics for Nanosecond and Femtosecond Laser Pulses," *Proceedings of SPIE, 2005 Photonics West Conference*, San Jose, CA, Vol. 5713, pp. 516–521.

Modeling and Experimental Verification of Transient/Residual Stresses and Microstructure Formation in Multi-Layer Laser Aided DMD Process

S. Ghosh

J. Choi

Department of Mechanical and Aerospace Engineering,
University of Missouri-Rolla,
1870 Miner Circle,
Rolla, MO 65409

Despite enormous progress in laser aided direct metal/material deposition (LADMD) process many issues concerning the adverse effects of process parameters on the stability of variety of properties and the integrity of microstructure have been reported. Comprehensive understanding of the transport phenomena and heat transfer analysis is essential to predict the thermally induced residual stresses and solidification microstructure in the deposited materials. Traditional solidification theories as they apply to castings or related processes, assume either no mass diffusion in the solid (Gulliver-Scheil) or complete diffusion in the solid (equilibrium lever rule) in a fixed arm space. These are inappropriate in high energy beam processes involving significantly high cooling rates. The focus of this paper is the solute transport in multi-pass LADMD process, especially the coupling of the process scale transport with the transport at the local scale of the solid-liquid interface. This requires modeling of solute redistribution at the scale of the secondary arm spacing in the dendritic mushy region. This paper is an attempt toward a methodology of finite element analysis for the prediction of solidification microstructure and macroscopic as well as microscopic thermal stresses. The computer simulation is based on the metallo-thermo-mechanical theory for uncoupled temperature, solidification, phase transformation, and stress/strain fields. The importance of considering phase transformation effects is also verified through the comparison of the magnitudes of residual stresses with and without the inclusion of phase transformation kinetics. The simulation has been carried out for H13 tool steel deposited on a mild steel substrate.
[DOI: 10.1115/1.2194037]

1 Introduction

LADMD process has demonstrated the ability to make a metal component, either homogeneous or heterogeneous, directly from a three-dimensional computer aided design (CAD) model. This process is achieved with a laser system combined with a numerical control (NC) machine tool or laser robot, CAD/CAM (computer aided manufacturing) technology, metal powder delivery system, and feedback control technology [1,2]. The three-dimensional CAD model is first sliced down to multiple two-dimensional layers. Each layer is then built using an optimized tool (laser) path pattern. In a sense, LADMD is a kind of layered manufacturing. Machine tool path for deposition can be obtained using the CAD/CAM software algorithm that provides the required G and M codes. During deposition, the deposition height is controlled by a feedback device that provides deposition height information to the controller through high speed CCD camera and photo sensor. The size of melt pool due to laser-substrate-powder interaction can be controlled by controlled feeding of metal powder through a concentric nozzle. Laser power can be controlled by the feedback system. References [1,2] have very well documented the fundamentals of LADMD process, process control, and influence of process parameters. Hence, further discussion of LADMD in this article is deemed unnecessary.

Thermal distortion and residual stresses are essentially generated by LADMD process [1–4] and it is well known that they affect the performance of the deposited structures such as brittle fracture, fatigue, buckling deformation, and stress corrosion cracking. Modeling of heat flow has become a standard practice and efforts have already been made to couple heat flow calculations to related macroscopic phenomena such as mould filling in casting, fluid flow, macrosegregation, or thermal stresses. Melting and solidification occur at very high rates in laser processing of materials. Because of the small dimension of the system and the fast rate of heat transfer, the process exhibits phenomena that are far from equilibrium. In these cases, not only microscale heat transfer, but also the microscale mass transfer and nonequilibrium phase-change kinetics become important. Macroscopic aspects are important in predicting macroscopic defects or optimizing process conditions, while microstructural features such as phase appearance, morphology, grain size, spacing, or microdefects are certainly no less important in determining the ultimate properties of the solidified product. A complete model that provides a quantitative relationship between process parameters, temperature history, phase transformation kinetics, solidification parameters, and the thermal stresses along with the deposited material microstructure is highly desirable.

In the processes incorporating phase transformation, such as laser cladding and LADMD, the change of material structure is necessary to be included in the analysis of thermal and residual stresses since the structural changes due to phase transformations are strongly coupled with the fields of temperature and stresses. This has been termed as metallo-thermo-mechanical coupling [5,6], and mode of such coupling is schematically illustrated in
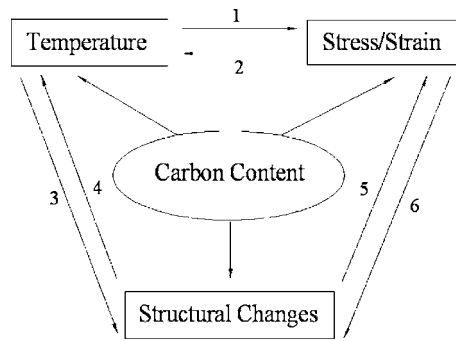
Copyright © 2006 by ASME

Fig. 1 Metallo-thermo-mechanical coupling during processes involving phase transformation

Fig. 1. As explained by the authors [5,6], when the temperature distribution in a material (and in some cases the heating and cooling rate) is unequal, thermal stress (1) is caused in the solid, and the temperature dependent phase transformation (3) leads to alternations in structure such as austenite-pearlite and austenite-martensite transformation. Local dilatation due to phase changes in the solid brings about transformation stresses (5) and also the effect of transformation plasticity interrupts the stress/strain field of the body. Contrary to these effects, arrows facing the opposite direction indicate different coupling modes. Part of the mechanical work done by existing stress in the material converts into heat generation (2), so disturbing the temperature. The acceleration of phase transformation by stress or strain, called stress-(or strain) induced transformation (6), is sometimes applied to improve the mechanical properties of metals. Arrow (4) corresponds to latent heat generation due to phase transformation, which affects the temperature distribution. In addition to such effect of coupling, content of carbon which diffuses during the carburization process is also considered to cause some influence on the fields.

Although Fig. 1 shows the complete nature of the phenomena, this paper is essentially a study of (1) a preliminary estimation of temperature history of the system and the prediction of residual stresses which also include the phase transformation kinetics of the material during its cooling and (2) the prediction of solidification microstructure based on the temperature history.

Several papers have discussed residual stresses in laser-deposited metal parts [1–4,7–9], structures with fast heating and cooling [10–16] and welded structures [17–19] in the past. Kahlen and Kar [7] have tried to model the residual stresses by analyzing the geometry of the melt pool and propagation of the solidification front and solidification angle with respect to time. However, they have considered only one component of strain rate, which depends on the solidification rate and the material dependent thermal expansion coefficient. Dai and Shaw [8] have used the commercial code ANSYS for the thermal and stress modeling of multi-material laser processing in different scanning patterns as well as different fabrication sequences. The code accounts for the strain rates due to elastic, plastic, and thermal components only. Nickel, Barnett, and Prinz [4] have tried to relate residual stresses and deformations to tool path orientation and deposition pattern. Their simulation for layered manufacturing without any material addition to the melt pool certainly requires modification. Their other approach to have an entire line of material heated at the same instant and gradually allowed to cool for a total time it takes for the laser to scan the entire line in actual experiment also does not conform to the quasi-steady state nature of heat transfer around the moving heat source. Roelens, Maltrud and Lu [17] have successfully simulated the residual stresses in submerged arc multipass welds with the inclusion of phase transformation. However, they restricted their approach to two-dimensional thermoplastic calculations in general plane strain elements.

A perspective of evolution of microstructure during solidifica-

tion is presented in this paper. Particular attention is paid to the effect of system size and resultant scale dependence of the microstructure and segregation profile. Most alloys solidify dendritically under commercially relevant conditions, and the microsegregation is defined as the compositional variation existing on the scale of the dendrite arms. Dendrites are themselves a result of a morphological instability, so it is not surprising that the result is so variable. However, the microstructure will still tend to display typical average dendrite arm spacing and an associated degree of microsegregation. Finally, it is shown how the microscopic models of microstructure formation can be coupled to macroscopic heat flow calculations in order to predict microstructural features at the scale of the whole process.

Several different theoretical models have been proposed in the literature to correlate various microstructural characteristics of solidified alloys with processing conditions [20]. Likewise, many papers [21–24] have correlated the dependence of microstructural characteristics on macro-processing conditions and attempted to draw either analytical or empirical relationships between various parameters based on the constitutive models of mass, energy, and diffusion equations. Wang and Prasad [21] emphasized the growth of velocity-dependent non-equilibrium solute partitioning coefficient. At the same time they disregarded the linear kinetics theory of local thermodynamic and chemical equilibrium at the interface which may not be valid since the interface velocity may reach or exceed the diffusion velocity. To account for such local non-equilibrium at the interface, Sobolev's hyperbolic solute transport formulation is employed. Ode et al. [23] stressed the fact that most of the theoretical models predict the relationship between secondary dendrite arm spacing, λ_2, and the local solidification time, t_f, as $\lambda_2 \propto t_f^{1/3}$ but the experiments show that the exponent is not always 1/3. They assumed a constant equilibrium partitioning coefficient in their work.

A finite element (FE) model has been developed to account for both micro solidification and residual stress effects. In the first phase of this program, the temperature and phase composition history is calculated using the heat diffusion elements of a commercial code, ABAQUS. In the second phase, stress-strain elements are used and the mechanical behavior is modeled by a subroutine. And in the third phase, microstructure calculations are made using a user subroutine.

2 Mathematical Model

2.1 Temperature Distribution. In laser aided DMD process a moving laser beam strikes on the substrate at $t=0$. The powder being added gets melted and subsequently solidified after cooling to form the clad. The transient temperature distribution $T(x,y,z,t)$ is obtained by solving the three-dimensional heat conduction equation in the substrate [25].

$$\rho c_p \frac{\partial T}{\partial t} - \rho c_p \nabla (UT) - \nabla (K \nabla T) = Q. \quad (1a)$$

Equation (1a) is a generalization of Rosenthal's derivation for a moving heat source in a quasi-steady state form. As mentioned earlier, the FE model, used to solve the conduction equation, does not directly accommodate the moving nature of heat source. However, a reasonable approximation is to divide the material addition onto the substrate into many small time steps, and apply variable flux and boundary conditions in each time step. A user subroutine is written for this approximation. The subroutine calculates the position of laser at a given time as a function of its speed (U) and the Gaussian flux distribution in terms of spatial coordinates. This approximates the quasi-steady state phenomena within the small time step and transient phenomena of the process as the time steps change. Hence, Eq. (1a) is actually used in the program in its simplified form given by Eq. (1b) that needs to be solved for every time step with new boundary conditions each time,

$$\rho c_p \frac{\partial T}{\partial t} - \nabla(K \nabla T) = Q \qquad (1b)$$

The essential boundary condition (BC) is

$$-K (\nabla T \cdot n)|_\xi = \begin{cases} \Phi_w|_\xi & \text{if } \xi \in \Psi \\ -h_c (T - T_0)|_\xi - \varepsilon_t \sigma (T^4 - T_0^4)|_\xi & \text{if } \xi \in \zeta \\ -h (T - T_0)|_\xi - \varepsilon_t \sigma (T^4 - T_0^4)|_\xi & \text{if } \xi \notin \Psi \end{cases}$$
$$(2)$$

In other words, enhanced convection due to shielding gas gushing out of the coaxial nozzle (inner nozzle for laser and the outer nozzle for shielding gas) and the radiation heat transfer due to localized melt pool must be considered for the surfaces directly below the nozzle. For simplification, the powder deposition was approximated with linear brick elements, the top surface of which has the input flux boundary condition, while the exposed faces have the enhanced convection and radiation boundary conditions. The surfaces not irradiated by the laser have been observed to have very small change in temperature. For such cases, small heat transfer coefficient of 10 W/(m² K) was assumed.

The following conditions should also be satisfied:

$$T(x,y,z,0) = T_0 \quad \text{and} \quad T(x,y,z,\infty) = T_0 \qquad (3)$$

The first of the two time dependent conditions in Eq. (3) is valid only for the start of the first pass of deposition. The second condition in this equation is required when the residual stresses at $t \to \infty$ are to be calculated.

2.2 Calculation of Microstructure Evolution. The approach in calculating microstructure evolution, such as austenite decomposition and the rate of formation of different phases during cooling process, proposed by Das, Upadhya, and Chandra [10] has been adopted. As mentioned by the authors, austenite decomposition has been modeled earlier by Kirkaldy et al. [11] based on a steady state approach using reaction kinetics for solid-state transformations. Using this approach it is possible to predict TTT (time-temperature transformation) curves for low alloy steels with a reasonable degree of accuracy. Watt et al. [12] have used this approach in combination with the austenite grain growth models developed by Ashby and Easterling [13] to predict microstructures in the HAZ of weldments. The rate of formation of different phases during the cooling process at any temperature can be estimated using Watt algorithm [12], which is given by

$$\frac{dX_i}{dt} = B(G_a, T) X_i^r (1 - X_i)^p \qquad (4)$$

where r and p are semi-empirical coefficients set to less than one to assure convergence in a form that is derived from a point nucleation and impingement growth model. Thus, Eq. (4) can be used to track the evolution of microstructure if one can determine the thermal history of the location of interest. The rate coefficient B includes the effect of grain size on the density of eligible nucleation sites. It also includes the amount of austenite super cooling and the effect of alloying elements and temperature on diffusion. The actual expressions for B, r, and p for various phases are available in Ref. [12]. As shown in Eq. (4), the effective rate coefficient B is a function of both austenite grain size G_a and the transient temperature T. Since no empirical relationship could be established to depict the coupled behavior of B with the transient nature of both G_a and T, daughter phase fractions in each time step of the simulation were obtained from the TTT curve widely available in data handbooks.

2.3 Calculation of Stresses. The total strain rate can be written as the sum of the individual components of the strain rates [6]

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}_{ij}^E + \dot{\varepsilon}_{ij}^P + \dot{\varepsilon}_{ij}^T + \dot{\varepsilon}_{ij}^{\Delta V} + \dot{\varepsilon}_{ij}^{\mathrm{TrP}} \qquad (5)$$

The various components in Eq. (5) represent strain rates due to elastic, plastic, thermal loading, volumetric change, and transformation plasticity, respectively. Here, elastic and thermal strains are normally expressed as [6]

$$\varepsilon_{ij}^E = \frac{1 + \nu}{E} \sigma_{ij} - \frac{\nu}{E} \sigma_{kk} \delta_{ij} \qquad (6)$$

and

$$\varepsilon_{ij}^T = \alpha (T - T_0) \delta_{ij} \qquad (7)$$

The plastic strain rate is reduced to the form when employing temperature dependent material parameters [6]

$$\dot{\varepsilon}_{ij}^P = \Lambda \frac{\partial F}{\partial \sigma_{ij}} = \hat{G} \left(\frac{\partial F}{\partial \sigma_{kl}} \dot{\sigma}_{kl} + \frac{\partial F}{\partial T} \dot{T} + \sum \frac{\partial F}{\partial X_i} \dot{X}_i \right) \frac{\partial F}{\partial \sigma_{ij}} \qquad (8)$$

with a temperature dependent yield function

$$F = F(\sigma_{ij}, \varepsilon^P, \kappa, T, X_i) \qquad (9)$$

$$\text{where} \quad \frac{1}{\hat{G}} = - \left(\frac{\partial F}{\partial \sigma_{mn}^P} + \frac{\partial F}{\partial \kappa} \sigma_{mn} \right) \frac{\partial F}{\partial \sigma_{mn}} \qquad (10)$$

$\hat{G}$ is termed as the hardening function.

However, to simplify the model, yield function F has been assumed to be independent of phase fraction X_i. Also, kinematic hardening was considered for calculations.

2.4 Kinetics of Phase Transformation. The incremental transformation plasticity can be written as [14]:

$$\Delta \varepsilon^{\mathrm{TrP}} = \frac{5 S_{ij} \Delta V}{4 Y V} (2 - 2X_n - \Delta X) \Delta X \qquad (11)$$

where Y is the yield stress of the weaker phase. Austenite is the weaker phase during these transformations. The strain increment due to the volumetric dilatation can be written as

$$\Delta \varepsilon^{\Delta V} = \frac{\Delta V}{3V} \Delta X \qquad (12)$$

As explained by Das et al. [10], even though in most phase transformation stress calculations volumetric strain and transformation plasticity are treated as two separate quantities, they are not unrelated effects. The term "volumetric strain" is used to represent the change in volume due to transformation per unit original volume, whereas "transformation plasticity" refers to the strains produced by the interaction between the stresses generated due to the transformation of individual grains and the macroscopic stress field that already exists in the system. Thus, there could be situations when the volumetric strain is non-zero but, the transformation plasticity is zero. These phenomena affect the stress field during the transformation from a face centered cubic to a body centered cubic structure. The present article treats both volumetric strain and transformation plasticity separately. It also assumes that both of them exist independently during solid-solid phase transformation. The last two terms in the strain rate Eq. (5) are non-zero only in the temperature range where the phase transformation occurs. The phase fractions are calculated from the phase transformation kinetics model outlined in the previous section. The fractions are then used in Eqs. (11) and (12) to update the transformation strain increments and the volumetric dilatation strain increments. The net strain increment values are then used to calculate the induced stresses. Equation (11) is only valid for a transformation process with constant temperature, involving only one transformation product [15]. In a realistic cooling process, on the contrary, several simultaneous transformation processes may be present, and the temperature is continuously changing. The time discretization is implemented according to the following approximation: *First*: The temperature decreases without any phase transformation. *Second*: Time elapses and all phase transformations take place simultaneously. When there is an amount of a third

phase present during a transformation, the most logical way of treating it is to assume that it is totally inert and does not influence the transformation in any way.

2.5 Solute-Diffusion Equation. In full diffusion equilibrium solidification that assumes no chemical potential gradients (phase compositions are uniform) and no temperature gradient, lever rule applies, i.e., a constant partitioning coefficient can be used. However, for non-equilibrium solidification wherein the dynamic effect of partitioning coefficient is to be accommodated, a simple model proposed by Aziz is used [24].

$$k_v = \frac{k + P_i}{1 + P_i} \tag{13}$$

where P_i, the interface Peclet number, is given by $P_i = \delta_i v / D_i$.

2.6 Dendrite Pattern. The solidification model presented in this section has been adopted from Kurz and Fisher [20]. They have modeled the dendritic growth assuming that the dendrite tip radius is equal to the wavelength of the critical instability of the solid-liquid interface. By minimizing the undercooling with respect to the radius of curvature of the tip, they have obtained steady-state growth velocity as given by

$$v = \frac{2D(GR^2 + 4\pi^2\Gamma)}{R^3(1-k)G - 2(1-k)C_0 m + 4\pi^2 R(1-k)} \tag{14}$$

The critical velocity of interface solidification growth is given by

$$v_{\text{critical}} = \frac{-mGD \exp\left(\frac{-E_a}{R_g T}\right)}{\Delta T_0} \tag{15}$$

where

$$\Delta T_0 = -mC_0(1-k)/k \tag{16}$$

At small velocities: $(v < v_{\text{critical}})$

$$v = \frac{2DG}{R(1-k)G - 2(1-k)C_0 m} \tag{17}$$

$$R = \frac{2D}{v(1-k)} + \frac{2mC_0}{G} \tag{18}$$

At large velocities: $(v > v_{\text{critical}})$

$$v = \frac{4\pi^2 D\Gamma}{R^2(1-k)C_0 m} \tag{19}$$

$$R = 2\pi\left[\frac{D\Gamma}{vk\Delta T_0}\right]^{1/2} \tag{20}$$

i.e., at small growth rates, the radius declines rapidly with increasing growth rate and at large growth rates, the radius falls parabolically with increasing growth rates. Assuming the overall morphology of the dendrite to be ellipsoidal, the primary dendrite arm spacing λ_1 is given by:

For small velocities: $(v < v_{\text{critical}})$

$$\lambda_1 = \left[\frac{6\Delta T'}{G(1-k)}\left(\frac{D}{v}\right) - \frac{\Delta T_0 k}{G}\right]^{1/2} \tag{21}$$

For large velocities: $(v > v_{\text{critical}})$

$$\lambda_1 = 4.3\Delta T'(1/4)\left[\frac{D\Gamma}{\Delta T_0 k}\right]^{1/4} v^{-1/4} G^{-1/2} \tag{22}$$

$$\text{where } \Delta T' = \frac{\Delta T}{(1-k)}\left(1 - \frac{GD}{v\Delta T_0}\right)$$

The secondary arm spacing λ_2 is one of the most important length scales of dendritic microstructure since it determines the
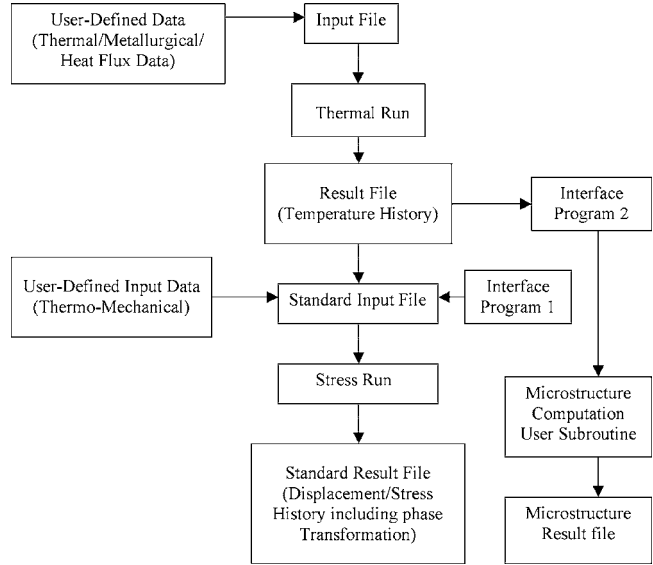
Fig. 2 Flow chart showing the microstructure computation and net residual stress calculation procedure with the inclusion of phase transformation

periodicity of solute segregation profile in the solidified material and thus influences its properties. The side branches become stable when the solute diffusional effects for instability are balanced by the stabilizing effect of the surface tension. As proposed by Kurz and Fisher [20],

$$-\lambda_2^0 R^2\left[\frac{r_0}{R} + \ln\left(1 - \frac{r_0}{R}\right)\right] = Mt_f \tag{23}$$

$$\text{where } M = \frac{-\Gamma D \ln\left(\frac{C_l^m}{C_0}\right)}{m(1-k)(C_l^m - C_0)}$$

Assuming $R/\lambda_2^0 \cong 0.7$ and $r_0/R \cong 0.8$ for a process like LADMD, the above is simplified to

$$\lambda_2 = (Mt_f)^{1/3} \tag{24}$$

3 Solution Scheme

3.1 Thermal Run. A flowchart showing various steps in the computational scheme employed in this work is given in Fig. 2. As discussed in Ref. [26], a transient thermal analysis is first performed within ABAQUS [27] to determine the temperature history in the work piece. This would need the thermal properties like, density, specific heat, conductivity, and latent heat for varying temperature range as input. This information is supplied to the user-defined interface program which computes the fraction of each daughter phase according to Eq. (4) and the strain terms related to transformation plasticity and volume change as per Eqs. (11) and (12). The subtotal of strain increments due to elastic, plastic, and thermal components are obtained at each time step from the second run, the stress run, which requires the thermo-mechanical properties—coefficient of thermal expansion, Young's modulus and Plastic kinetic of the material at different temperatures as input.

3.2 Stress/Deformation Run. In the program flow chart shown in Fig. 2, phase transformations are treated in detail including volume change and transformation plasticity. Carburizing can be treated by prescribing the carbon content as a field variable. However, for the simplicity of our model, carburizing is being neglected at present. The interface program 1 estimates the phase

kinetics at each node where the phase transformation and temperature are considered as prescribed time dependent variables. The interface program calculates the volume fractions of each phase according to the temperature-time history. In the stress analysis, the same finite element mesh as is used in the heat transfer analysis is defined and the temperature history from the heat transfer analysis is followed exactly. The size of the time steps (Δt) in the heat transfer analysis is independent of that in the stress analysis. Different time increments may be required in the two analyses in order to achieve convergence. This would require interpolation of temperatures in the stress analysis. The program automatically interpolates the temperature as necessary. Since the temperature change is very rapid during the initial heating in laser aided DMD process, a smaller time increment is necessary for convergence as compared to the time increment required for convergence during slow cooling process when the laser goes past the substrate.

All the mechanical properties – elastic modulus (E), tangent modulus in the plastic region (E_p), yield stress (S_{yield}), Poisson's ratio (ν), and coefficient of thermal expansion (α) and the thermal properties – conductivity (K), Density (ρ) and specific heat (c_p) are taken as functions of temperature from Ref. [28]. The stress-strain relation is assumed to be linearly elastic and linearly plastic described by the modulus of elasticity and the modulus of plasticity in the elastic and plastic regions respectively.

3.3 Microstructure Run. In "conventional" uncoupled thermal/mechanical situations, the stress/deformation run would use the calculated temperatures as stored on the output file directly as load specification. However, due to the number of state variables necessary in our case to describe the microstructure (phase composition), this standard temperature load input procedure cannot be used directly for the uncoupled thermal/metallurgical/mechanical problem. Instead, an interface program 2 is developed which can read the time increment organizations together with the temperature and the microstructure state variables from the thermal-run result file and produce a neutral-run database file containing all temperatures and microstructures state variables for all thermal-run nodes at all thermal-run time instants. The interface program estimates the primary dendrite arm spacing, tip radius, and secondary dendrite arm spacing at each node where the cooling rate, interface temperature gradient, local solidification velocity, and temperature are considered as prescribed time dependent variables.
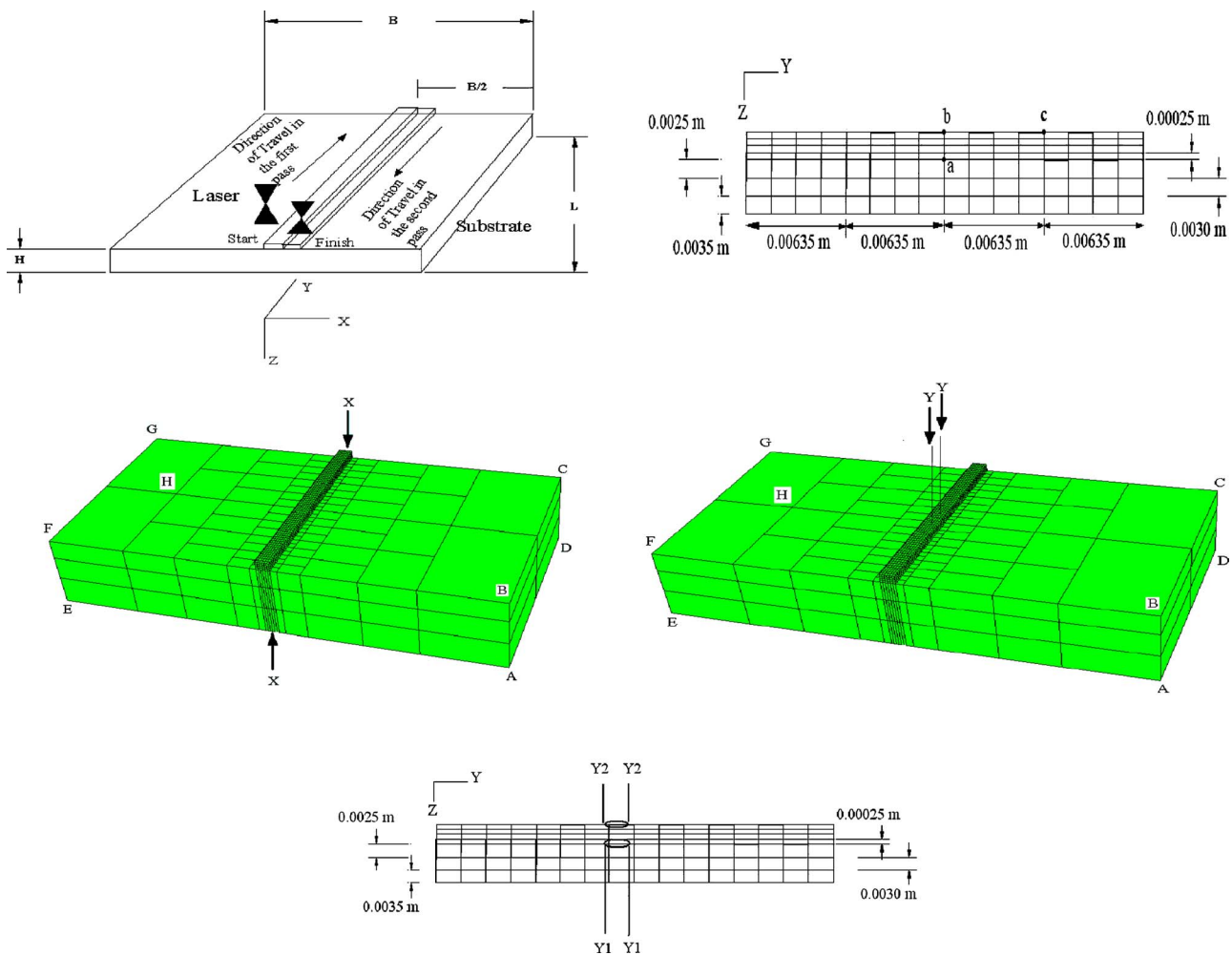
Fig. 3 (*a*) **Schematic of a double-pass laser aided DMD process,** (*b*) **section** *X-X* **chosen for stress analysis,** (*c*) **locations of the points on section** *X-X* **chosen for transient stress analysis,** (*d*) **section** *Y-Y* **chosen for solidification microstructure, and** (*e*) *Y1-Y1* **at the interface of the first layer and substrate along the direction of laser travel and** *Y2-Y2* **at the interface of the third and fourth layers along the direction of laser travel**

3.4 Finite Element Scheme and Method of Numerical Calculation. The formulated finite element equation system considering the nexus between increment of nodal temperature $\{T\}$ and displacement $\{\Delta u\}$ can be expressed in the forms:

$$[P]\{\dot{T}\} + [H]\{T\} = \{Q\} \tag{25}$$

$$[K(u)]\{\Delta u\} = \{\Delta F(T)\} \tag{26}$$

where $\{Q\}$ is the heat flux vector and $\{\Delta F(T)\}$ is the incremental thermal load vector.

As discussed earlier, the transient heat transfer analysis is first carried out. Gaussian profile for the laser beam is used and appropriate convective and radiative boundary conditions including the forced convective heat loss due to the shielding gas effect, effect of latent heat, effect of moving heat source and temperature dependent thermophysical properties are considered. The details of laser heat flux calculation, applying it to the substrate, handling the molten metal pool in the finite element model and the thermomechanical properties are available in Ref. [29].

The following sequential steps are adopted to solve the proposed model:

1. Instead of metal powder, equivalent numbers of rectangular "*elements*" are added on to the substrate. These elements represent the molten powder being added. The initial temperature of the nodes belonging to these elements is assumed to be 1703 K. The particulate nature of the cladding material has not been considered in this article. It is believed that it is altogether a difficult research topic in itself. To keep the model simple, the cladding material has been treated as a set of solid linear brick elements (finite element analysis) being added onto the substrate in any given time step. The cladding material "mass conservation" has been ensured. The cladding thickness is calculated from the powder flow rate, laser beam size, and laser traverse speed. The cladding width has been assumed to be the same as laser beam size.

2. The material addition along the laser pass is divided into small "*steps*" to simulate the quasi-steady state nature of the process.

3. In each step, input flux is applied on the top surface of the set of elements added and the necessary convective and radiative boundary conditions applied on the system to accommodate the overall transient nature of the process.

4. The temperature followed by the stress distribution in the domain is calculated by the finite element "*iterative scheme*."

5. Similar to step 4, transformation plasticity and strain due to volumetric dilatation are calculated in each time step as per interface program 1 and added to the thermal, elastic, and plastic strain rates to obtain the resultant stress.

6. Cooling rate, temperature gradient, and solidification velocity are calculated as per interface program 2 at the
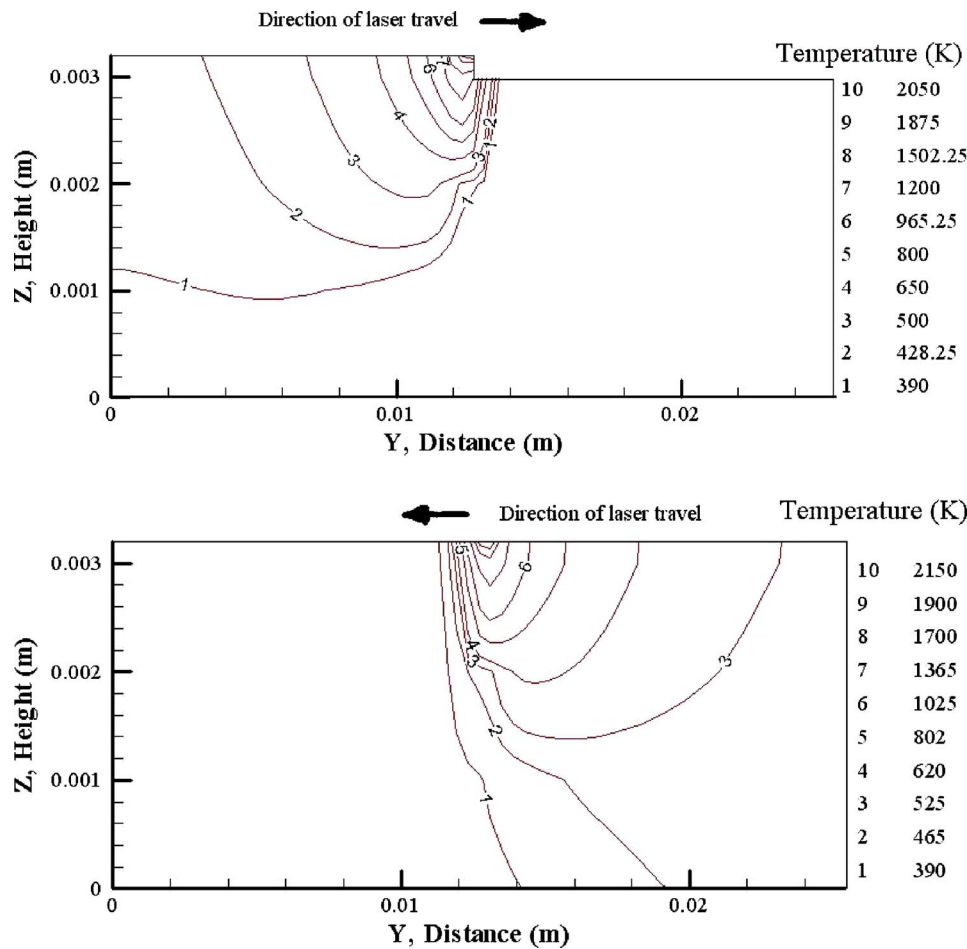

Fig. 4 Temperature contour on section *X-X* at time (*a*) *t*=1.0 s (Single-pass on the first layer) and (*b*) *t*=3.0 s (Double-pass on the first layer). All temperature values are in degrees kelvin (K).

nodal points in each time step to obtain the resulting microstructure information.

3.5 Assumptions and Considerations. A simplistic approach based on the following assumptions has been chosen to model and simulate the DMD process.

1. The melt pool dynamics and the consequent solidification phenomenon with the moving boundary between solidus and liquidus have been neglected.
2. An uncoupled FE analysis is made between thermal stresses and microstructure evolution.
3. Transient thermal stresses have been assumed to be dependent on phase fraction. However, the latent heat was assumed constant i.e., independent of the phase fractions present during the transient phenomenon.
4. The processes described as (2), (4), and (6) in the metallo-thermo-mechanical model shown in Fig. 1 have been neglected in the FE calculations.
5. Only high power laser with continuous wave (cw) and fundamental (Gaussian) mode was considered.
6. Volumetric dilatation and phase transformation plasticity are two separate quantities and exist independently during solid-solid phase transformation.
7. Both absorptivity and emissivity of deposited material were considered constant.
8. 100% austenite present at the beginning of phase transformation.

9. 10% retained austenite at the end of transformation to martensite during cooling.
10. Acceleration and deceleration during laser travel and the time to change tracks are negligible.
11. Steady state growth velocity for dendrite formation is assumed.
12. The dendrite tip has been assumed to be a paraboloid of revolution.
13. The diffusion process driven by the gradients in the liquid at the solid/liquid interface is only by thermal gradient (ΔT_t) and compositional gradient (ΔT_c), not by curvature of the tip (ΔT_r).

3.6 Example Under Consideration. The example under consideration has length=25.4 mm, breadth=50 mm, height=9 mm and laser clad width and height of 750 and 250 μm, respectively (approximated for a powder flow rate of 6 gm/min) in each pass with 2 passes in a layer and 4 layers in total, providing simulation results for 8 passes (as shown in Fig. 3(a)). The substrate and the deposited material are assumed to be made up of Mild Steel and H13 Tool Steel, respectively. A total of 4176 nodes, and 3060 eight-node linear brick type elements were considered. The laser source was assumed to travel at a speed of 12.7 mm/s with a power of 700 W, absorptivity of 0.45 and emissivity of 0.23, respectively, of the substrate-deposited material system. The entire system has been simulated for thermal and residual stresses with an initial temperature of 294 K for all nodes. The one-pass laser
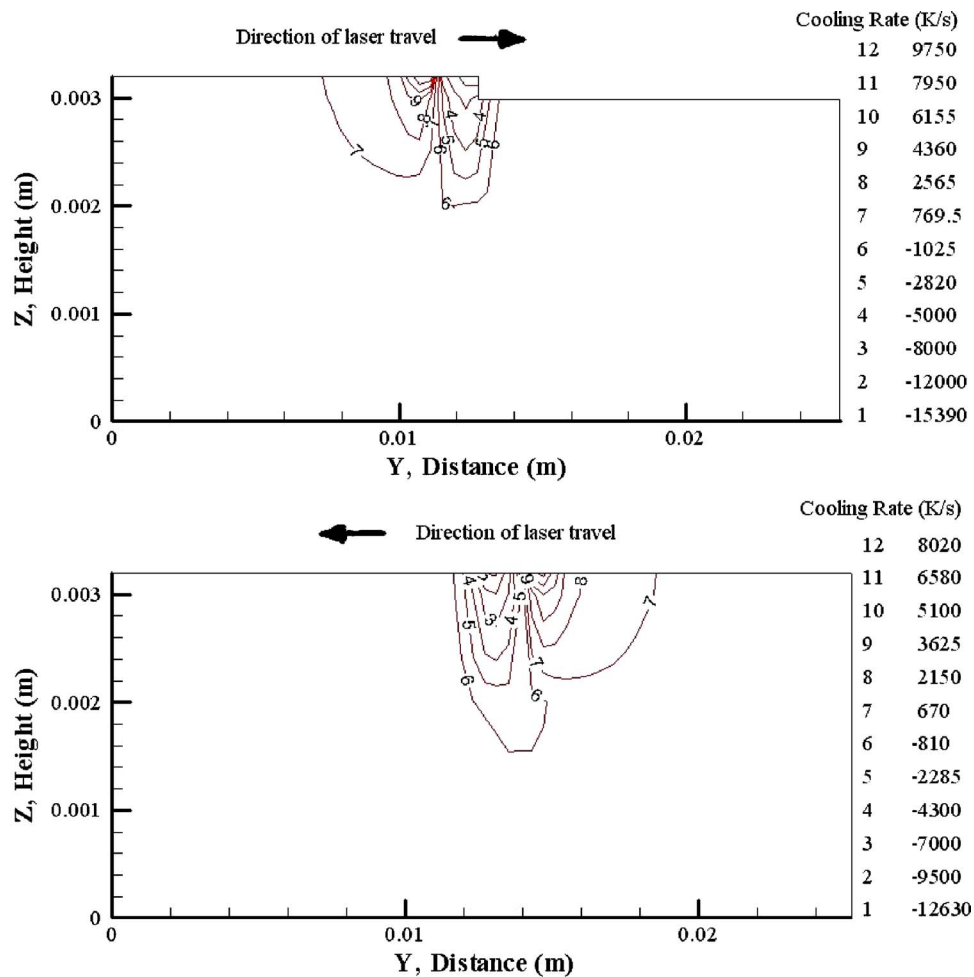

Fig. 5 Cooling rate contour on section *X-X* at time (*a*) *t*=1.0 s (Single-pass on the first layer) and (*b*) *t*=3.0 s (Double-pass on the first layer). Cooling rate values are in degrees kelvin/second (K/s).

travel of 25.4 mm has been discretized into 16 equal time steps with each time step corresponding to 0.125 s. 8 more steps corresponding to time $t=16$ s up to time $t=32$ s were added to the existing 128 steps to account for the cooling after the laser goes past the substrate. The substrate is clamped on the faces denoted by ABCD and EFGH.

Absorptivity times laser power is the power impinging on the deposited material-substrate system. The laser intensity for a Gaussian profile, at any spatial coordinate (x, y), is given by

$$I = \frac{2Q_r}{\pi R_s^2} \exp\left(-\frac{2r_s^2}{R_s^2}\right) \qquad (27)$$

where, R_s is the laser spot size (radius), usually measured between the points where the intensity has fallen to $1/e^2$ of the central value, and Q_r is the resultant energy reaching the substrate per unit time. Part of Q_r is used up to accommodate the latent heat of melting and to raise the temperature of the added powder from 294 K to above melting. The latent heat used in this case was 273.790 (kJ/kg, for H13 tool steel). The radial distance of any spatial coordinate (x, y) from the center of the beam is given by

$$r_s = \sqrt{x^2 + y^2} \qquad (28)$$

The latent heat of melting comes into picture between the temperatures of 1703 and 1733 K. This is very well accommodated in the program for temperature run.

Based on a previous model by Choi and Mazumder [18], $h = 10$ W/(m^2 K) has been used for surfaces not influenced by the shielding gas. Since the specimen is assumed to be hinged on the edges ABCD and EFGH, shown in Fig. 3(a), $h=10$ W/(m^2 K) has been used for the base surface as well. However, the following empirical relation initially suggested by Gardon and Cobonque [30] for air and later modified by Steen [31] for any gas has been used for a part of the top surface under the nozzle:

$$h = 13 * \mathrm{Re}^{1/2} \mathrm{Pr}^{1/3} k_{\mathrm{gas}} \mathrm{NPD} \qquad (29)$$

where Re=Reynolds number, Pr=Prandlt number, k_{gas} is the conductivity of the shielding gas, and NPD denotes the standoff distance. Assuming gas velocity V_{gas} to be 10 m/s, density ρ_{gas} =0.1626 kg/m^3, dynamic viscosity $\mu_{\mathrm{gas}}=19.9\times10^{-6}$ Pa s, and characteristic length i.e., effective nozzle diameter D to be 0.0025 m, Reynolds number

$$\mathrm{Re} = \frac{V_{\mathrm{gas}} D \rho_{\mathrm{gas}}}{\mu_{\mathrm{gas}}} = 204 \qquad (30)$$

For our calculations Pr=0.667 and $k_{\mathrm{gas}}=1.781\times10^{-2}$ W/(m K) were chosen. To accommodate this a user subroutine for modified heat transfer coefficient is written which relates the combined effects of the parameters mentioned earlier.

4 Simulation Results

4.1 Temperature and Cooling Rate. Thermal energy by laser raises the temperature of the substrate's surface node "a" (shown in Fig. 3(c)) beyond the melting point of 1703 K. Once the laser goes past the node it starts cooling depicting a behavior similar to the error function. The second pass of the laser causes the temperature at point "a" shoot up higher than the previous peak due to the heat accumulation from the previous pass. Figures 4 and 5 show the distribution of temperature and cooling rates on the interfacial plane. As seen clearly, the heat-affected zone is very small. However, the heat accumulation due to successive passes leads to a slightly higher melt pool depth in the ensuing passes. Very high heating rates, maximum of the order of $\sim 10^4$ K/s, have been observed in this process. The negative cooling rates, shown in Fig. 5, are actually the heating rates. The cooling rates drop drastically as time progresses. It can be contributed to the heat accumulation in the successive passes and the
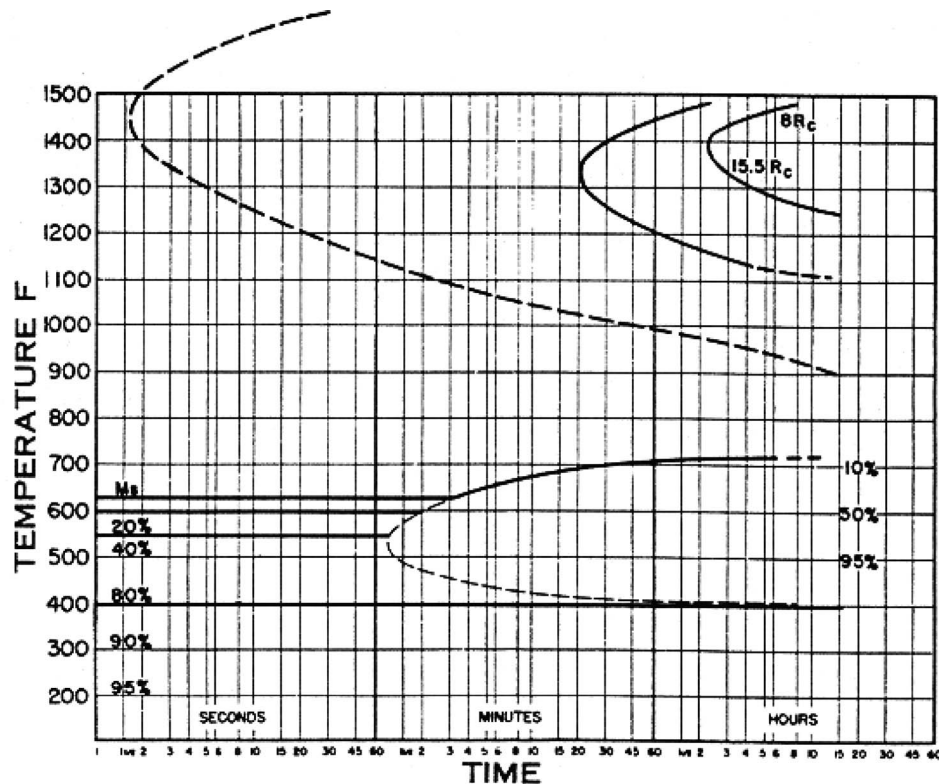

Fig. 6 Time-temperature transformation (TTT) diagram of H13 Tool Steel. (Boyer, H.E., and Gray, A.G., 1977, *Atlas of Isothermal Transformation and Cooling Transformation Diagrams*," American Society for Metals, Metals Park, OH).
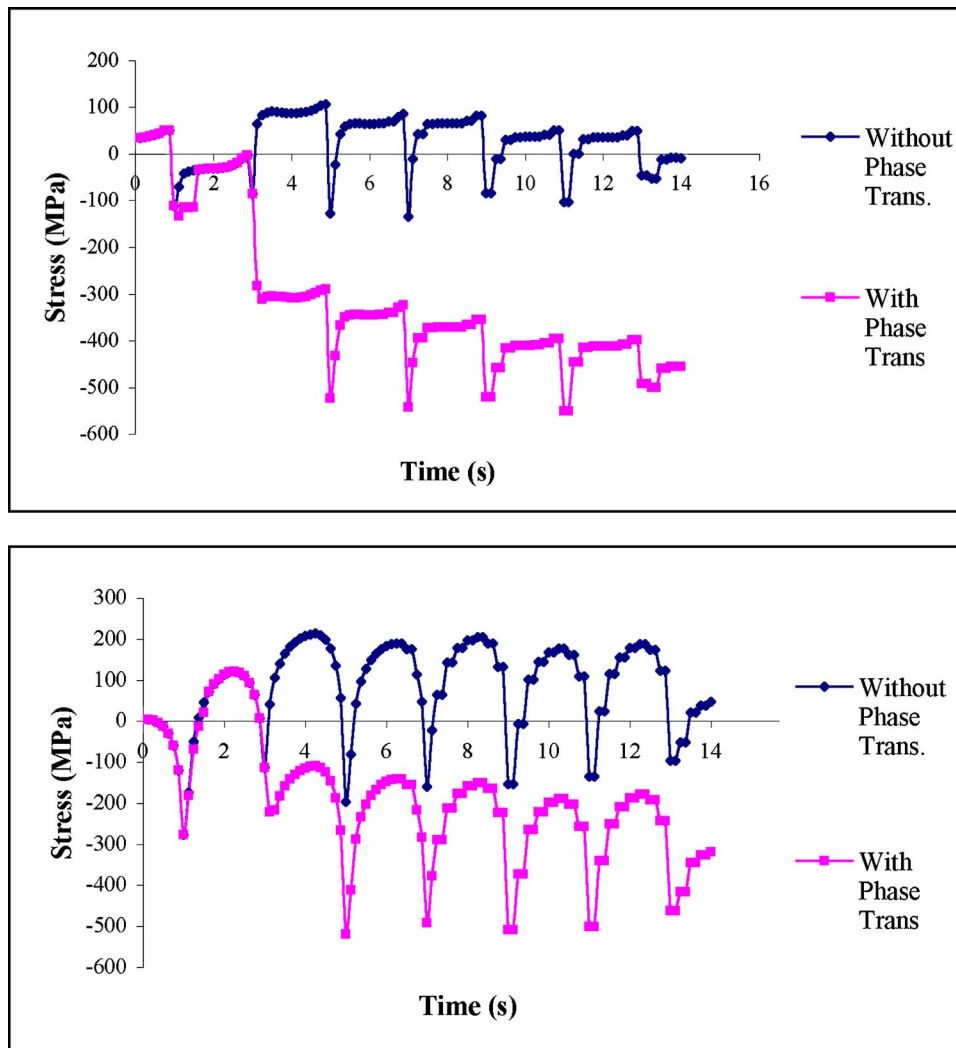
Fig. 7 Comparison of (a) transverse (S11) and (b) longitudinal (S22) transient stresses at point "a" with and without phase transformation effects

imbalance in the total amount of heat added and that dissipated. However, this phenomenon is peculiar to transient state behavior only.

4.2 Thermal Stress. From cooling rate results, it can be deduced that processes similar to laser cladding will have very high cooling rates. From the time-temperature transformation (TTT) [32] shown in Fig. 6, it is concluded that the exceptionally high cooling rates in the process will make points "a", "b", and "c" have no regions of any phases or mixed microstructures other than martensite and retained austenite phases. It has also been observed that in case the cooling rate varies a little bit or the material goes through a different phase change like δ Ferrite and γ Austenite, the necessary changes can be readily incorporated in the interface program without affecting the integrity of the main program.

In Figs. 7 and 8, the evolution of transverse (S11, perpendicular to laser travel) and longitudinal (S22, along the laser travel) transient stresses at the chosen points on section X-X are shown. Since point a is a node common both to the substrate (mild steel) and to the deposited material (tool steel), the stress at this location is the average of the stresses associated with both materials. The specimen yields as the initial thermal energy causes the surface to contract, giving rise to large tensile stresses on the surface and compressive stresses below in the core. The survey points undergo tensile stresses as the laser passes them. The small dip in the

graphs which shows compressive stress before the laser actually hits these points is due to the surface and core effect. As the specimen cools, martensitic transformation begins. The volume expansion associated with phase change tends to increase the compressive surface stress. For all the points, the effects of transformation plasticity and volumetric dilatation actually make the stress change sign from tensile to compressive. This change in the residual stress occurs since the phase transformation effects oppose the effects of thermal strains. The analysis shows that the cool elastic limit of 400 MPa is not exceeded at point b. However, point a has exceeded the limit. This is due to the heat accumulation in repeated passes.

4.3 Microstructure Formation. Figure 3(d) and 3(e) shows the sections chosen for the study of solidification microstructure. The simulated results for Y1-Y1 at time $t = 3.125$ s have been shown in Fig. 9. The simulated results for Y2-Y2 at time $t = 15.125$ s have been shown in Fig. 10. As shown in Fig. 9, for the section Y1-Y1, the local solidification time, t_f decreases near the top of the clad (approximately 2–4 ms), the dendrite arm spacings decrease and the solidification substructure becomes relatively finer. With larger values of t_f near clad-substrate boundary (approximately 20–40 ms), relatively coarser dendrite arm spacings and solidification substructures are produced. Similar pattern can
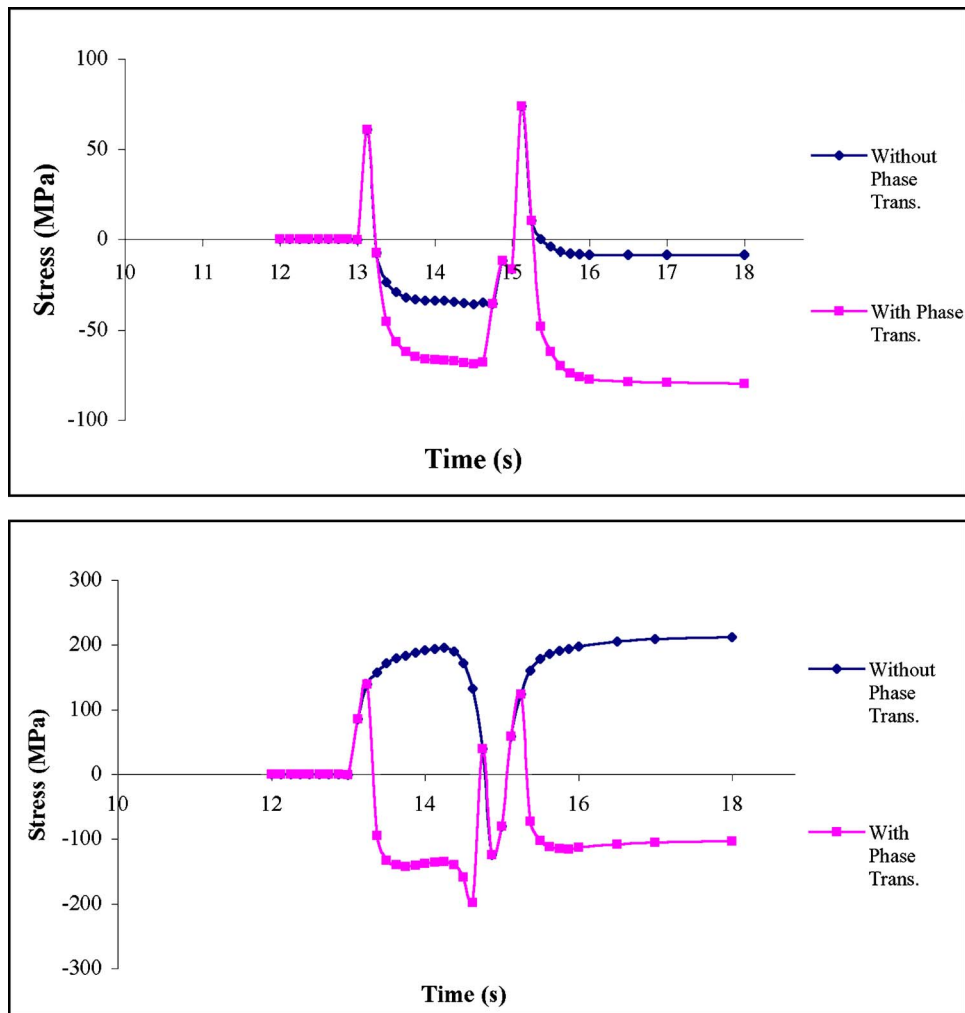
Fig. 8 Comparison of (*a*) transverse (S11) and (*b*) longitudinal (S22) transient stresses at point "*b*" with and without phase transformation effects

also be observed for section $Y2$-$Y2$. The simulation results do not show much of difference in the microstructures as the primary dendrite arm spacing (PDAS) and secondary dendrite arm spacing (SDAS) at these two locations are similar. It is convenient to view the solidification substructure as dependent on the cooling rate. Since cooling rates are tremendously high in processes like LADMD, very fine microstructures with λ_1 and λ_2 of the order of just a few micrometers can be observed. The dendrite arm spacing can also be attributed to solute segregation. Solute segregation being higher at the clad centerline reduces the rate of dendrite coarsening therein.

The growth rate of interface, also called the local solidification velocity v, is the ratio of local cooling rate to the interface temperature gradient (G). The dynamic nature of mushy zone makes the local solidification velocity change rapidly. Consequently, the local solidification time t_f, which is the ratio of the difference between liquidus and solidus temperatures to the local solidification velocity, also changes with a similar trend. For the mild steel and tool steel materials, average solidus and liquidus temperatures used were 1703 and 1733 K, respectively.

5 Comparison between Experimental and Simulation Results

5.1 Experimental Setup. The simulation program for double-pass four-layer laser aided DMD was verified with the experimental results for temperature, residual stress, and microstructure. For

temperature measurements, as shown in Fig. 11, three thermocouples were spark-welded at points *a*, *b*, and *c*. The metal deposition started at a distance of 0.008317 m from the front edge (as seen in the figure) through a total length of 0.03556 m. The substrate plate dimensions and the locations of points *a*, *b*, and *c* on it are shown in the figure. The material for the substrate was mild steel while the deposited metal was H13 tool steel. CO_2 type cw Gaussian laser beam with a traverse speed of 12.7 mm/s was chosen. The nominal laser power used was approximately 680 W, while the beam diameter and the stand-off distance were 1 and 8 mm, respectively. Varying absorptivity of 0.4–0.5 and a constant emissivity of 0.23 was assumed for the substrate-deposited material system for different simulation cases. The entire system has been simulated for thermal history with an initial temperature of 294 K for all nodes.

For residual stress and microstructure measurement, samples were made exactly like the one discussed in the simulation results earlier, with the same geometry and process variables. However, the substrate was not clamped to the work table. H13 tool steel was deposited on a mild steel substrate using a CO_2 type cw laser with Gaussian beam shape. For residual stress measurement, a Philips X'Pert MRD diffractometer was used with a Cu radiation source. The detector was a sealed proportional counter. A crossed slit collimator module with height and width both set at 1.0 mm (PW3084/62) was the primary optic module used in conjunction with the Philips high resolution goniometer (PW3050/65) and
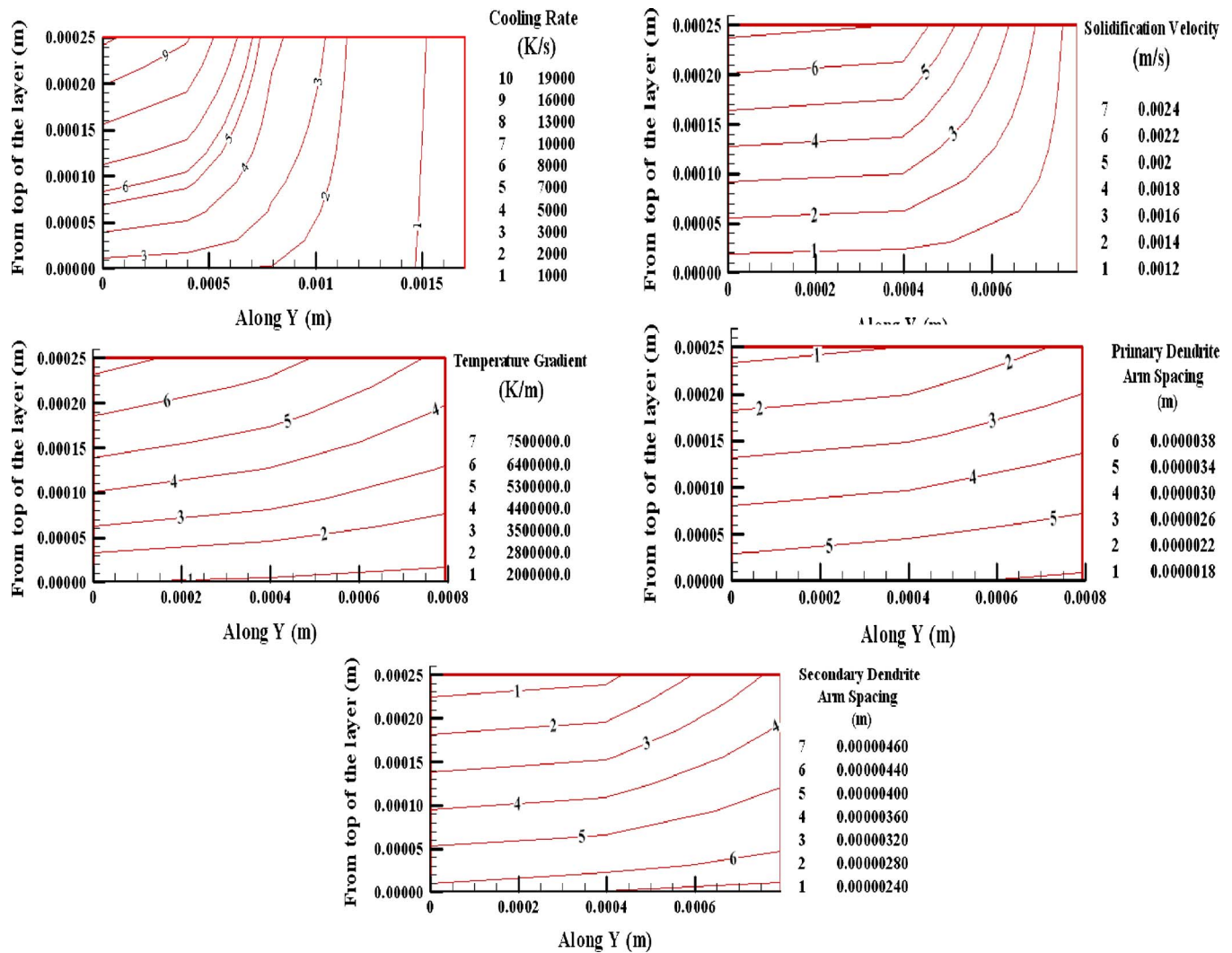
Fig. 9 (*a*) Cooling rate, (*b*) temperature gradient, (*c*) solidification velocity, (*d*) primary dendrite arm spacing, λ₁, and (*e*) secondary dendrite arm spacing, λ₂ at *Y*1-*Y*1 at time *t*=3.125 s

MRD cradle (PW3060/20). Secondary optic module used was a 0.27° parallel plate collimator (PW3098/27) equipped with a flat graphite monochromator (PW3121/00). Data were obtained over the 2-theta range of 110–125° with a step size of 0.02° and a counting time of five seconds for five tilt (ψ) angles between 0 and 70 degrees evenly spaced as a function of $\sin^2(\psi)$. These data were analyzed using Philips X'Pert stress software (version 1.0).

5.2 Temperature Results. As seen in Fig. 12, the numerical results for temperature are close to the experimental ones. Simulated results for different values of absorptivity have been compared with the experimental ones and a maximum error of approximately 13% was observed in the simulated results for absorptivity of 0.4. The uncertainty in the measurement is about 10 K. This could be slightly higher as the thermocouple beads would have been exposed to some laser energy by radiation. A closer look at the results also indicates a slight phase shift between the simulated and the experimental graphs. This could partially be contributed by (1) the acceleration/deceleration of the laser travel as well as its dwell during change of tracks were not considered in modeling and programming and (2) the laser travel was manually synchronized with the thermocouple data acquisition. It is obvious that the temperature in Fig. 12(*b*) cannot decrease in the first 2 s when the laser beam is approaching point "*b*." The discrepancy is clearly due to the mechanical delay and

the response time of the thermocouples. The approximate size of the bead was 1 mm. The thermocouples were spark welded onto the surface. Since they were not shielded or insulated with sealants (e.g., RTV (room temperature vulcanizing) silicones), it is suspected that they might have been affected by the exposure to laser radiation. It is also believed that the time steps chosen were not sufficient to show both the peaks for point *c*. However, with smaller time steps these two peaks can be distinguished clearly.

5.3 Residual Stress Results. As seen in Figs. 13 and 14, the simulation results for the longitudinal and transverse residual stresses at points "*b*" and "*c*" of Fig. 3(*c*) have been compared with that obtained by experiment. For both points, the experimental values for longitudinal stresses (S22) lied between the two simulation results, stresses with and without phase transformation. As explained earlier, this article treats volumetric strain and transformation plasticity as two separate quantities. However, there could be situations when the volumetric strain is non-zero but the transformation plasticity is zero. Analyzing these as two dependent quantities and to know when volumetric strain is non-zero while transformation plasticity is zero, requires a lot more research, and is beyond the scope of this article at present. The uncertainty in experimental stress measurement is shown as y-error bar in Figs. 13 and 14. The errors along transverse direction (S11) are higher. These are due to the limitation of the x-ray
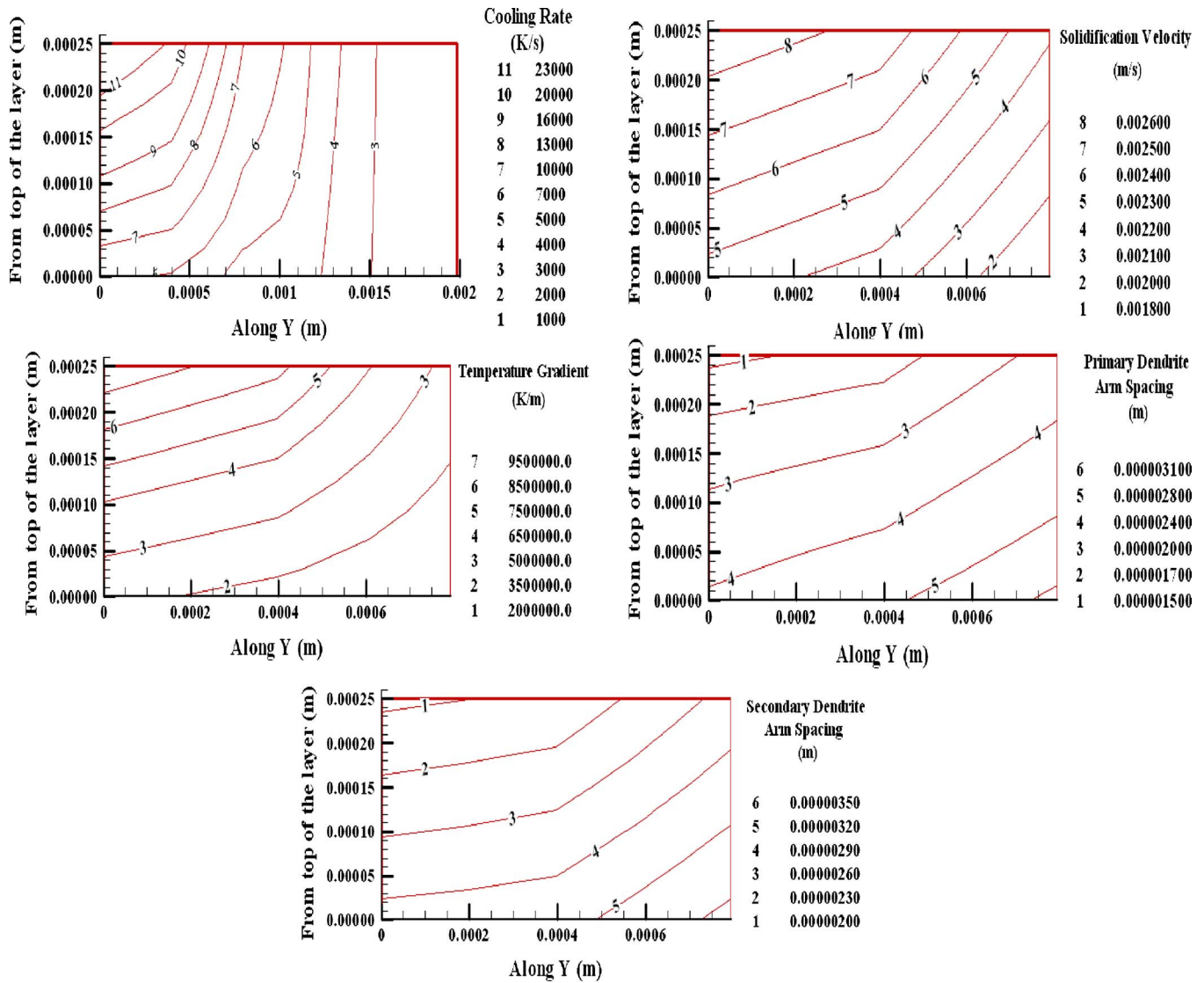
Fig. 10 (*a*) **Cooling rate,** (*b*) **Temperature gradient,** (*c*) **solidification velocity,** (*d*) **primary dendrite arm spacing,** λ_1, **and** (*e*) **secondary dendrite arm spacing,** λ_2 **at** *Y2-Y2* **at time** *t*=15.125 s

diffractometer. The experimental result was an average of many points within the x-ray beam size of 1 mm×1 mm on the top of the clad. Sample dependent errors may arise from an excessively coarse grain size, severe texture, or interference of the sample geometry with the x-ray beam. The seemingly high experimental error for S11 stresses could be due to: (1) the x-ray beam size was

higher than the spots "*b*" and "*c*" on top of the interface between the two passes of laser, (2) the experiment was done with almost 5% overlap between the two passes of the laser which forms regions of two distinct peaks and a valley. It is suspected that the x-ray beam must have missed some data points within the valley. Consequently, the average data obtained lacked some significant points within the beam size. On the other hand, since the x-ray beam measures most of the valley data points along S22 direction; reason (2) also explains why the errors along S22 direction were low.

5.4 Microstructure Results. Shown in Fig. 15 are the SEM micrographs of the H13 tool steel deposited on the mild steel substrate with the process parameters as mentioned earlier. The simulated results for solidification microstructure seem to match well with the experiment. The simulated average primary dendrite arm spacing (λ_1 or PDAS) is of the order of 2.0 to 4.0 μm as compared to 1.5 to 4 μm obtained experimentally (Fig. 15). Similarly, the simulated average secondary dendrite arm spacing (λ_2 or SDAS) is of the order of 2–4 μm as compared to 2–5.5 μm obtained experimentally (Fig. 15). The uncertainty in the experimental measurement with the image processor is 0.5 μm. Micrographs shown in Fig. 16 are for the same process parameters as

Fig. 11 Schematic of the experimental setup for double-pass single layer LADMD process

Fig. 12 Time-temperature variation at (*a*) point "*a*," (*b*) point "*b*," and (*c*) point "*c*" of Fig. 11

above, except for a higher laser traverse speed of 19 mm/s. The simulation results show a finer microstructure in this case, at both sections *Y*1-*Y*1 and *Y*2-*Y*2. Again, the simulated results for solidification microstructure seem to match well with the experiment. The simulate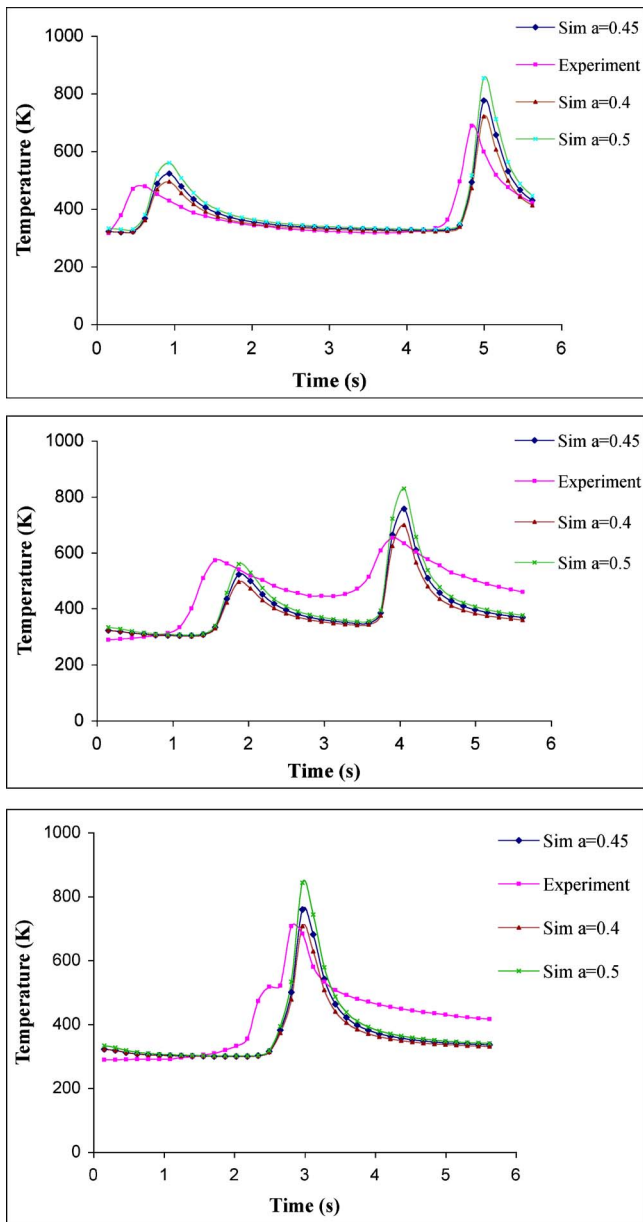d average primary dendrite arm spacing (λ_1) is of the order of 1.7–2.8 μm as compared to 1–3.5 μm obtained experimentally (Fig. 16). Similarly, the simulated average secondary dendrite arm spacing (λ_2) is of the order of 2–3.3 μm as compared to 1.5–3.5 μm obtained experimentally (Fig. 16).

Figure 17 substantiates the previous claim that exceptionally high cooling rates in the process will make points *a* and *b* have no regions of any phases or mixed microstructures other than martensite and retained austenite. Figure 17 is the plot of diffraction intensity versus 2θ for the test sample, where θ is the angle of incidence. The plot was compared with the XRD reference data for both austenite and martensite phases [33]. The experimental peaks are closer to the martensite peaks. The final microstructure

after cooling can be concluded to have 90–100% martensite and 10-0% retained austenite. This is in agreement with both our assumption and Ref. [9].

6 Discussion

Simulation results show that the inclusion of the phase transformation effects makes significant difference in the stress history and the final residual stress in the material. Points where the simple thermal analysis predicts tensile residual stress ends up with compressive residual stress when the transformation effects are included. Beyond the elastic limit, the thermal residual stress is controlled by the yield stress and the hardening parameters of the material. For all design purposes the elastic limit of stress will be considered the safe stress. Additionally, it is also presumed that any residual stress in the material exceeding the yield stress will lead to undesirable distortions and initiation of cracks. The prediction given forth by this analysis thus calls for immediate countermeasures to keep the residual stress at different points well under the elastic limit of the material. Pre-heating of the substrate leads to the reduction of residual stresses as it reduces the cooling rate and, thus, reduces the stress induced by phase transformation [34,35]. The other countermeasure could be the use of lower laser power.

The changes induced in the thermal stress state due to transformation plasticity cannot be easily foreseen. The phase transformation stress strongly depends on the macroscopic deviatoric stress state, cooling rate, phase fractions, and the volumetric strain. Since these are highly dynamic in nature, the relative order of magnitude of these stresses is also dynamic. We observed different magnitudes of these at different locations. It could be seen that the maximum contribution of phase transformation effects to the existing thermal stresses was almost up to 2.5 times.

For processes like LADMD, the thermal stress and phase transformation stresses are most important. They lead to plastic strain that causes the residual stress. Both substrate clamping load and thermal load lead to elastic as well as plastic strains (first two terms in the Eq. (5)). The phase transformation leads to volumetric strain as well as transformation plasticity (last two terms in Eq. (5)), both of which are plastic strains. Hence, these three strain factors ($\dot{\varepsilon}_{ij}^{P}$, $\dot{\varepsilon}_{ij}^{\Delta V}$ and $\dot{\varepsilon}_{ij}^{\mathrm{TrP}}$) that cause plastic strain are the most important. The elastic strain gets relaxed after substrate unclamping.

During rapid heating and cooling process, stresses are introduced by thermal gradients via volume expansion and transformation strains. If the end product is homogeneous, these stresses are only temporary and will return to zero as temperature again becomes homogeneous. But, in diffusion processes such as some phase transformations, the amount of new phase formed depends on the time spent at elevated temperature, and after quenching, the phase composition will not be homogeneous.

Simulation results for temperature, when compared with experiment, also revealed the fact that the model is extremely sensitive to material property data, process parameters, and boundary conditions. As a particular example, simulation PDAS and SDAS varied between 2 and 4 μm for laser traverse speed of 12.7 mm/s, while the same became finer and varied between 1.7–2.8 and 2–3.3 μm, respectively, for laser traverse speed of 19 mm/s. Also, a reliable convection/radiation film coefficient, absorptivity, and emissivity as functions of temperature appear to be absolutely critical to generating an accurate model, and at the same time, these are difficult to obtain. Other key data include mechanical behavior of the various phases present during the cooling process.

The temperature verification through thermocouples validates the heat transfer model. The results indicate that the melt pool dynamics have marginal effect on the three-dimensional heat flow in the system. The results also indicate that an uncoupled temperature-stress analysis is a reasonable assumption. Since the maximum error between measured and simulated temperatures was low, it is assumed that the fully-coupled analysis would only
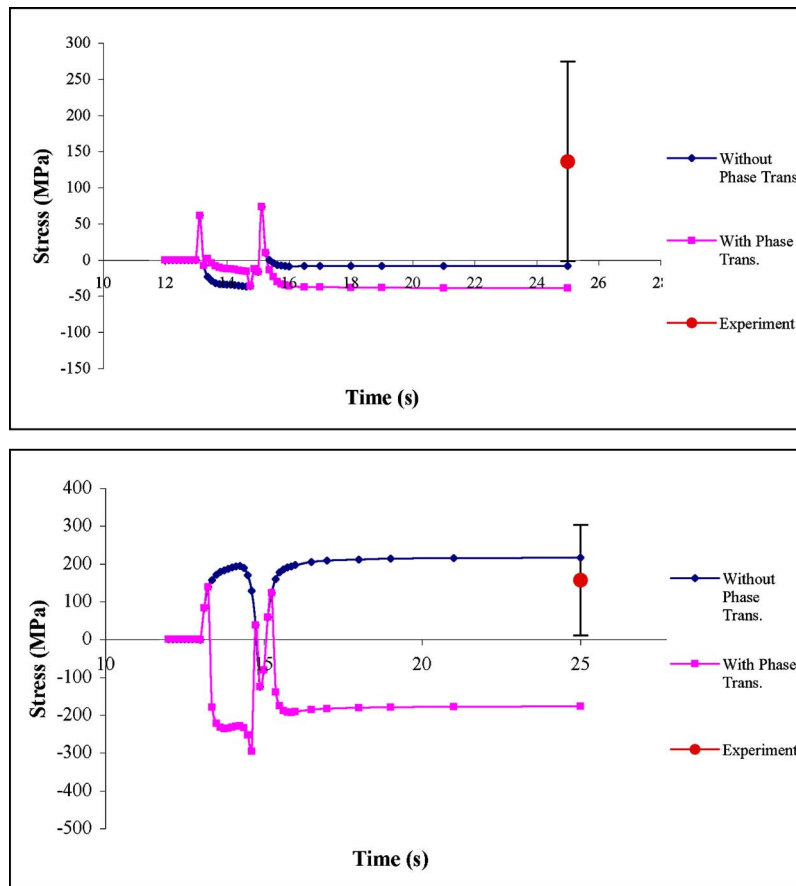
Fig. 13 (*a*) **Transverse** (**S11**) **and** (*b*) **longitudinal** (**S22**) **thermal stresses at point "*b*" of Fig. 3**

affect the results marginally. Also, constant values of latent heat, heat transfer coefficient, and emissivity lead to a small temperature error.

The microstructure verification validates our assumption that the diffusion process in the liquid at solid-liquid interface is driven mainly by thermal and compositional undercooling. The curvature undercooling was not modeled, and it did not seem to change our results much. Also, the dendrite radius tip can be better represented as a paraboloid of revolution than a hemisphere. This has been verified through simulation.

Both heat transfer and kinetics models were tested for mesh sensitivity. The mesh adopted in the present article was achieved after some trials. Since a finer mesh did not change the results much, this mesh was definitely the optimum compromise between accuracy of results and computation time.

7 Concluding Remarks

The paper presents the mathematical treatment of phase transformation effects during cooling of high strength steels in laser aided DMD process. A simple three-dimensional thermo-kinetic finite element model has been implemented to simulate the residual stresses and net distortions of the material along with the microstructure formation after solidification. Residual stresses due to transformation plasticity are usually neglected due to the complexity involved and the inability to measure the contribution of each phase constituent in the entire transient phenomenon. However, the simulation carried out for certain locations of our interest showed that the residual stresses due to the phase transformation plasticity could be alarming especially at the substrate-clad interface and may lead to crack, and finally de-bonding of the deposited material which is otherwise wrongly predicted to be very safe

if phase transformation effects were not to be taken into consideration. In processes such as laser welding, fusion cutting, or DMD in which a locally concentrated heat source is used, it is not possible to avoid fundamental problems like residual stresses and deformation. However, to cope with them successfully, it is necessary to understand them well and strategic applications of simulation are recommended as much as experiments. The results have been encouraging so far. It further stimulates us to delve into the transient and residual stress analysis for multi-pass and multilayer laser aided deposition with different combinations of tool path patterns.

The present simulation code models the thermal stresses and microstructure evolution only. Since the focus of this research was thermal residual stresses due to phase transformation, thin substrate curvature as a result of this stress was not studied. However, it is proposed to have this FE code modified to model substrate curvature as well in future.

The present model cannot predict either cracking or fracture with certainty. However, this model can predict if the residual stress at any location exceeds the elastic limit, which for all design purposes is considered the safe limit. It is reasonable to assume that a crack or fracture may initiate or undesirable deformations may occur once the residual stress goes past the elastic limit of stress.

Our stress verification showed some error between the experiment and simulation. As mentioned in the article, volumetric strain and transformation plasticity have been treated separately. However, they are not unrelated altogether. At the same time, it must be noted that it is very difficult to treat them as related terms in numerical solution. The authors are working towards a methodology of modeling them dependent on each other.
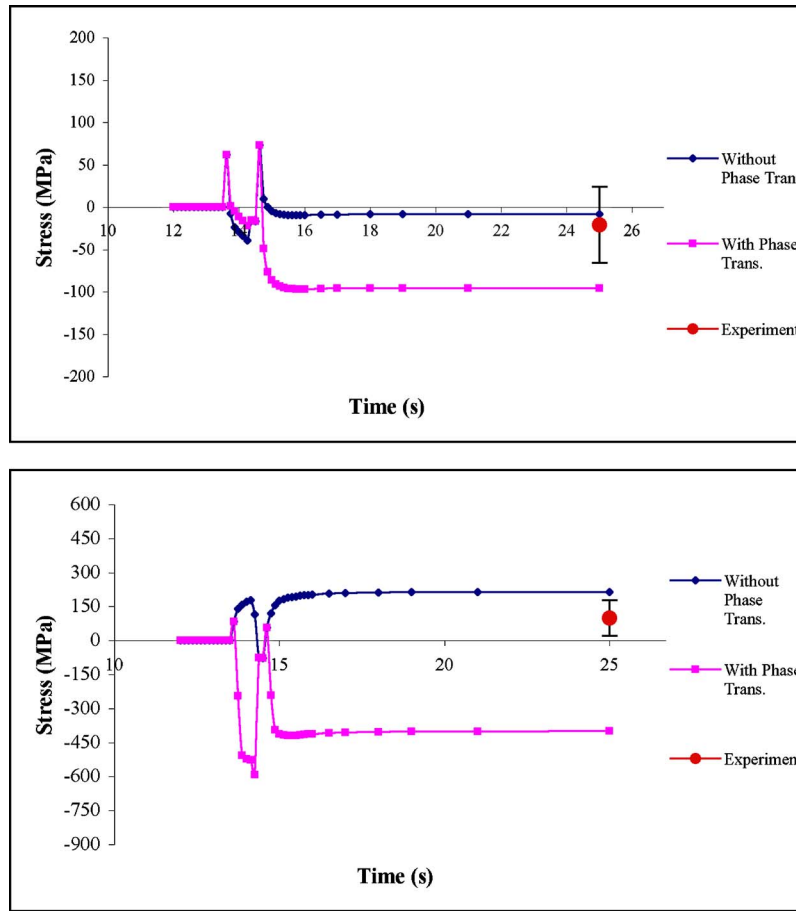
Fig. 14 (*a*) **Transverse** (**S11**) **and** (*b*) **longitudinal** (**S22**) **thermal stresses at point "***c***" of Fig. 3**

The metallo-thermo-mechanical coupling discussed by Inoue and others explains the comprehensive correlation and interdependence of stresses and microstructure formation on the thermal history the material goes through. In this paper efforts have been made to study the residual stresses and material microstructure formation due to solidification independently. However, the prediction of stresses due to changing solidification history and eventually a varying microstructure still remains a big challenge.

Nomenclature

a = absorptivity
c_p = specific heat
h = coefficient of heat transfer
h_c = coefficient of heat transfer due to enhanced cooling
k = equilibrium distribution coefficient
k_v = non-equilibrium distribution coefficient
m = liquidus slope
n = normal vector of the surface
r, p = semi-empirical coefficients known for various phases present in the steel
r_0 = critical nucleation radius
t = time
t_f = local solidification time
v = growth rate of interface
v_{critical} = critical solidification velocity based on limit of constitutional supercooling
B = effective rate coefficient
C_0 = initially alloy composition

C_l^m = ending alloy composition due to segregation
D = diffusion coefficient in liquid
D_i = interface diffusion coefficient
E = Young's modulus
E_p = tangent modulus in plastic region
E_a = activation energy for diffusion (J/mol)
F = temperature dependent yield function
G = interface temperature gradient
G_a = austenite grain size
$[H]$ = heat conduction matrix
I = laser intensity
K = thermal conductivity of the substrate
$[K(u)]$ = displacement dependent stiffness matrix
P = solute Peclet number
P_i = interface Peclet number
$[P]$ = heat capacity matrix
Q = power generation per unit volume of the substrate
$\{Q\}$ = heat flux vector
Q_r = power reaching the substrate
R = primary dendrite tip radius
R_g = gas constant (8.314 J/m K)
S_{ij} = deviatoric stress tensor
S_{yield} = yield stress
T = temperature field
T_0 = sink temperature
$\dot{T}$ = temperature rate
U = Travel speed of the source (process speed)
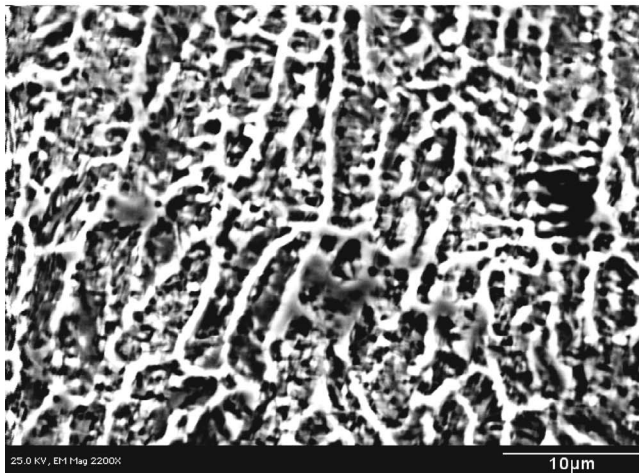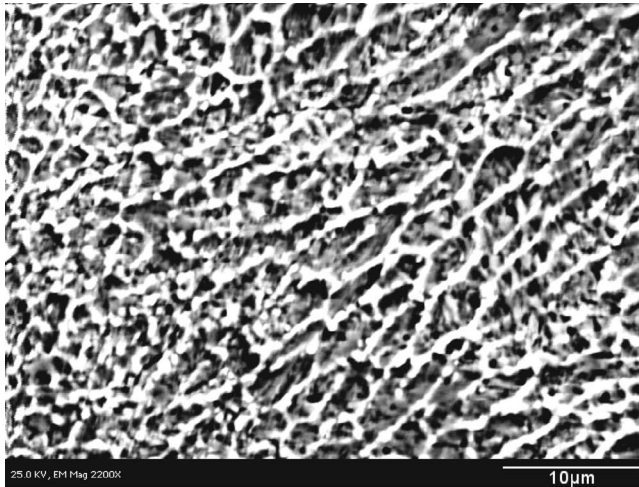
Fig. 15 SEM micrograph of deposited sample showing the microstructure: (*a*) on a surface along the laser travel and (*b*) on a surface lateral to the laser travel. Laser traverse speed is 12.7 mm/s.
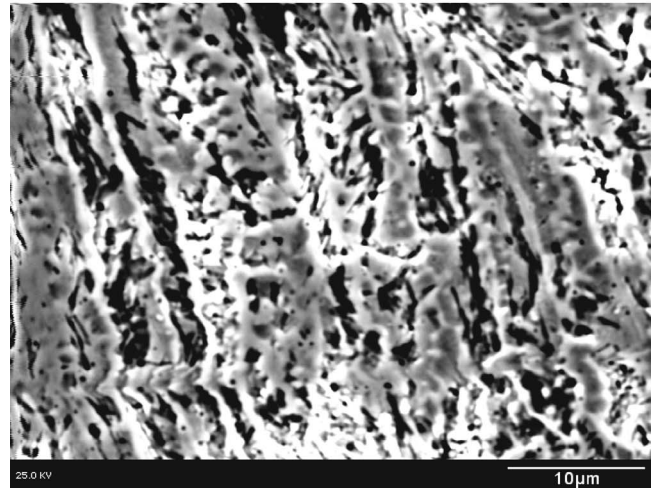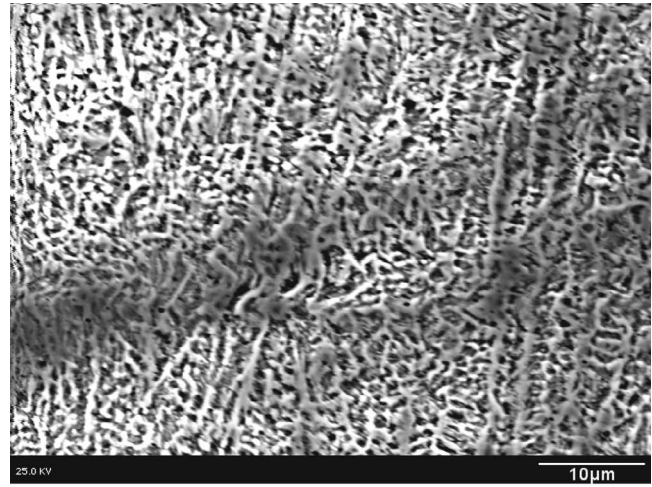
Fig. 16 SEM micrograph of deposited sample showing the microstructure: (*a*) on a surface along the laser travel and (*b*) on a surface lateral to the laser travel. Laser traverse speed is 19 mm/s.

X_i = volume fraction of the *i*th phase constituent

X_n = volume fraction of austenite already transformed

$\dot{X}_i$ = volume fraction rate of the *i*th phase constituent

Y = yield stress of the weaker phase during phase transformation

$\{\Delta u\}$ = incremental displacement

$\{\Delta F\}$ = incremental thermal load vector

$\Delta T'$ = difference between tip temperature and non-equilibrium solidus

ΔT_c = solute undercooling

ΔT_t = thermal undercooling

ΔT_r = curvature undercooling

ΔT_0 = unit undercooling

ΔT = total undercooling

$\Delta V/V$ = volumetric strain occurring during phase transformation

ΔX = amount of austenite being transformed in a particular step

Greek symbol

α = coefficient of thermal expansion

ρ = density of the material

ε_t = emissivity of the clad material

κ = hardening parameter

ν = Poisson's ratio

σ = Stefan-Boltzmann constant $[5.67 \times 10^{-8} \text{ W/m}^2 \text{ K}^4]$

ξ = surfaces of the workpiece

Ψ = surface area irradiated by the laser beam

ζ = surface area hit by the shielding gas from the co-axial nozzle

$\dot{\varepsilon}_{ij}$ = total strain rate

$\dot{\varepsilon}_{ij}^E$ = elastic strain rate

$\dot{\varepsilon}_{ij}^P$ = plastic strain rate

$\dot{\varepsilon}_{ij}^T$ = thermal (expansion) strain rate

$\dot{\varepsilon}_{ij}^{\Delta V}$ = strain rate component due to volumetric dilatation

$\dot{\varepsilon}_{ij}^{\text{TrP}}$ = strain rate due to transformation plasticity

σ_{ij} = induced elastic or plastic stress

$\dot{\sigma}_{kl}$ = induced elastic or plastic stress rate

δ_{ij} = Kronecker's delta

Γ = Gibbs-Thomson coefficient

δ_i = characteristic interface width

λ_2^0 = arm spacing before ripening

Φ_w = total flux absorbed by the substrate

λ_2 = secondary dendrite arm spacing

Fig. 17 X-ray diffraction data for point "b" (diffraction intensity versus 2θ)
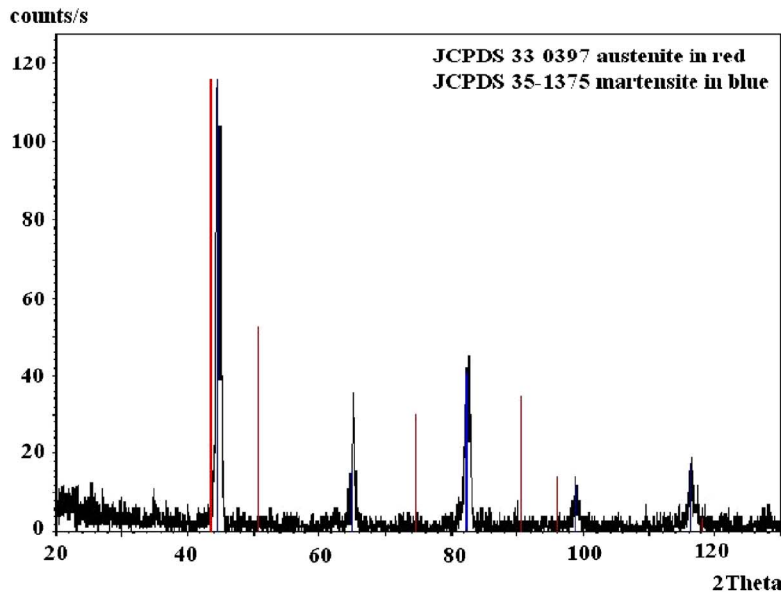
Peak	Martensite	Austenite
First	$2\theta = 44.684°$	$2\theta = 43.620°$
Second	$2\theta = 65.036°$	$2\theta = 50.836°$
Third	$2\theta = 82.382°$	$2\theta = 74.770°$
Fourth	$2\theta = 98.990°$	$2\theta = 90.791°$

Reference [33]

λ_1 = primary dendrite arm spacing

References

[1] Mazumder, J., Choi, J., Nagarathnam, K., Koch, J., and Hetzner, D., 1997, "The Direct Metal Deposition of H13 Tool Steel for 3-D Components," JOM, **49**(5), pp. 55–60.

[2] Choi, J., and Chang, Y., 2005, "Characteristics of Laser Aided Direct Metal Deposition Process for Tool Steel," Int. J. Mach. Tools Manuf., **45**(4-5), pp. 597–607.

[3] Klingbeil, N. W., Beuth, J. L., Chin, R. K., and Amon, C. H., 2002, "Residual Stress-Induced Warping in Direct Metal Solid Freeform Fabrication," Int. J. Mech. Sci., **44**, pp. 57–77.

[4] Nickel, A. H., Barnett, D. M., and Prinz, F. B., 2001, "Thermal Stresses and Deposition Patterns in Layered Manufacturing," Mater. Sci. Eng., A, **317**, pp. 59–64.

[5] Inoue, T., 2002, "Metallo-Thermo-Mechanics- Application to Quenching," *Handbook of Residual Stress and Deformation of Steel*, ASM International, Materials Park, OH, pp. 296–311.

[6] Inoue, T., Ju, D. Y., and Arimoto, K., 1992, "Metallo-Thermo-Mechanical Simulation of Quenching Process – Theory and Implementation of Computer Code – Hearts," *Proceedings of the First International Conference on Quenching & Control of Distortion*, Chicago, IL, 22–25 September, pp. 205–212.

[7] Kahlen, F.-J., and Kar, A., 2001, "Residual Stresses in Laser-Deposited Metal Parts," J. Laser Appl., **13**(12), pp. 60–69.

[8] Dai, K., and Shaw, L., 2001, "Thermal and Stress Modeling of Multi-Material Laser Processing," Acta Mater., **49**, pp. 4171–4181.

[9] Bokota, A., and Iskierka, S., 1996, "Effect of Phase Transformation on Stress States in Surface Layer Laser Hardened Carbon Steel," ISIJ Int., **36**(11), pp. 1383–1391.

[10] Das, S., Upadhya, G., and Chandra, U., 1992, "Prediction of Macro-and Micro-Residual Stress in Quenching using Phase Transformation Kinetics," *Proceedings of the First International Conference on Quenching & Control of Distortion*, ASM International, Chicago, IL, pp. 229–234.

[11] Kirkaldy, J. S., and Venugopalan, D., 1983, "Prediction of Microstructure and Hardenability in Low Alloy Steels," *Proceedings of the International Conference on Phase Transformations in Ferrous Alloys*, A publication of The Metallurgical Society of AIME, October 4–6, Philadelphia, PA, pp. 125–148.

[12] Watt, D. F., Coon, L., Bibby, M., Goldak, J., and Henwood, C., 1988, "An Algorithm for Modeling Microstructural Development in Weld Heat-Affected Zones (Part A) Reaction Kinetics," Acta Metall., **36**(11), pp. 3029–3035.

[13] Ashby, M. F., and Easterling, K. E., 1982, "A First Report on Diagrams for Grain Growth in Welds," Acta Metall., **30**, pp. 1969–1978.

[14] Oddy, A. S., Goldak, J. A., and McDill, J. M. J., 1989, "Transformation Effects in the 3D Finite Element Analysis of Welds," *Proceedings of the Second International Conference on Trends in Welding Research*, ASM International, Gatlinburg, TN, May 15–19.

[15] Jarvstrat, N., and Sjostrom, S., 1993, "Current Status of TRAST; A Material Model Subroutine System for the Calculation of Quench Stresses in Steel," *ABAQUS User's Conference Proceedings* .

[16] Denis, S., 1996, "Considering Stress-Phase Transformation Interactions in the Calculation of Heat Treatment Residual Stresses," J. Phys. IV, **6**, pp. C1 159–C1 174.

[17] Roelens, J. B., Maltrud, F., and Lu, J., 1994, "Determination of Residual Stresses in Submerged Arc Multi-Pass Welds by means of Numerical Simulation and Comparison with Experimental Results," Weld. World, **33**(3), pp. 152–159.

[18] Choi, J., and Mazumder, J., 2002, "Numerical and Experimental Analysis for Solidification and Residual Stresses in GMAW process for AISI 304 Stainless Steel," J. Mater. Sci., **37**, pp. 2143–2158.

[19] Tekriwal, P., and Mazumder, J., July 1991, "Transient and Residual Thermal Strain-Stress Analysis of GMAW," J. Eng. Mater. Technol., **113**, pp. 336–343.

[20] Kurz, W., and Fisher, J., 1990, *Fundamentals of Solidification*, Trans Tech Aedermannsdorf, Switzerland.

[21] Wang, G. X., and Prasad, V., 2000, "Microscale Heat and Mass Transfer and Non-equilibrium Phase Change in Rapid Solidification," Mater. Sci. Eng., A, **292**, pp. 142–148.

[22] Trivedi, R., and Kurz, W., 1994, "Solidification Microstructures: A Conceptual Approach," Phys. Met. Metallogr., **42**(1), pp. 15–23.

[23] Ode, M., Kim, S. G., Kim, W. T., and Suzuki, T., 2001, "Numerical Prediction of the Secondary Dendrite Arm Spacing using a Phase-Field Model," ISIJ Int., **41**(4), pp. 345–349.

[24] Aziz, M. J., 1982, "Model for Solute Redistribution during Rapid Solidification," J. Appl. Phys., **53**(2), pp. 1158–1168.

[25] Calrslaw, H. S., and Jaeger, J. C., 1959, *Conduction of Heat in Solids*, 2nd ed., Oxford University Press, Oxford.

[26] Ghosh, S., and Choi, J., 2005, "Three-Dimensional Transient Finite Element Analysis for Residual Stresses in the Laser Aided Direct Metal/Material Deposition Process," J. Laser Appl., **17**(3), pp. 144–158.

[27] ABAQUS/Standard User's Manual and Keywords Manual, version 6.4.

[28] Touloukian, Y. S., and Ho, C. Y., 1970, *Thermo Physical Properties of Matter-The TPRC Data Series by Purdue University*, IFI/Plenum Data Corp., New York.

[29] Ghosh, S., 2003, "Three-Dimensional Transient Residual Stress FE Analysis for Single and Double Pass Laser Aided Direct Metal/Material Deposition Process," University of Missouri- Rolla, M.S. thesis.

[30] Gordon, R., and Cobonque, J., 1961, "Heat Transfer between a Flat Plate and Jets of Air Impinging on it," *International Conference on Heat Transfer*, Part II, pp. 454–460.

[31] Steen, W. M., 1976, "The Printing of Laser Generated Heat Images in Cobalt Oxide on Glass Substrates," Ph.D. thesis, Imperial College, London.

[32] Boyer, H. E., and Gray, A. G., 1977, "Atlas of Isothermal Transformation and Cooling Transformation Diagrams," *American Society for Metals*, Metals Park, OH.

[33] Pfoetsch, and Ruud, Penn State Univ., Univ. Park, Pennsylvania, ICDD Grant-in-Aid, 1982 and 1984.

[34] Choi, J., and Mazumder, J., 1994, "Non-Equilibrium Synthesis of Fe-Cr-C -W alloy by Laser Cladding," J. Mater. Sci., **29**, pp. 4460–4476.

[35] Choi, J., Choudhuri, S. K., and Mazumder, J., 2000, "Role of Preheating and Specific Energy Input on the Evolution of Microstructure and Wear Properties of Laser Clad Fe-Cr-C-W Alloys," J. Mater. Sci., **35**, pp. 3213–3219.

Jun Zhou
Department of Mechanical and Aerospace
Engineering,
University of Missouri-Rolla,
229 Mechanical Engineering,
1870 Miner Circle,
Rolla, MO 65409
e-mail: jzhou@umr.edu

Hai-Lung Tsai
Department of Mechanical and Aerospace
Engineering,
University of Missouri-Rolla,
122 Mechanical Engineering Annex,
1870 Miner Circle,
Rolla, MO 65409
e-mail: tsai@umr.edu

Pei-Chung Wang
R & D Center,
General Motors Corporation,
Warren, MI 48090
e-mail: Pei-chung.wang@gm.com

Transport Phenomena and Keyhole Dynamics during Pulsed Laser Welding

Numerical and experimental studies were conducted to investigate the heat transfer, fluid flow, and keyhole dynamics during a pulsed keyhole laser welding. A comprehensive mathematical model has been developed. In the model, the continuum formulation was used to handle solid phase, liquid phase, and mushy zone during melting and solidification processes. The volume-of-fluid method was employed to handle free surfaces. The enthalpy method was used for latent heat. Laser absorptions (Inverse Bremsstrahlung and Fresnel absorption) and thermal radiation by the plasma in the keyhole were all considered in the model. The results show that the recoil pressure is the main driving force for keyhole formation. Combining with the Marangoni shear force, hydrodynamic force, and hydrostatic force, it causes very complicated melt flow in the weld pool. Laser-induced plasma plays twofold roles in the process: (1) to facilitate the keyhole formation at the initial stage and (2) to block the laser energy and prevent the keyhole from deepening when the keyhole reaches a certain depth. The calculated temperature distributions, penetration depth, weld bead size, and geometry agreed well with the corresponding experimental data. The good agreement demonstrates that the model lays a solid foundation for the future study of porosity prevention in keyhole laser welding.
[DOI: 10.1115/1.2194043]

Keywords: laser welding, heat transfer, fluid flow, modeling, free surface, keyhole, computational, VOF

1 Introduction

Welding with lasers is characterized by creating a keyhole inside the molten metal. A high-energy-density beam vaporizes the workpiece during the welding process to form a deep hole, which is called the keyhole. The keyhole increases the coupling of laser energy into the workpiece, resulting in a weld with high depth-to-width ratio and a narrow heat-affected zone. However, porosity is easy to find in laser weld, which is a major defect of laser welding. Since the porosity formation and weld quality depend on the heat transfer and the melt flow during the formation and collapse of the keyhole, it is necessary to study the detailed transport phenomena during the welding process.

So far, most of the models investigating the keyhole formation and collapse mechanisms can be divided into two categories, one based on the vaporization approach and the other based on recoil pressure research [1]. The vaporization model assumed that there is a balance between the sum of recoil pressure and vapor pressure in the keyhole and the sum of surface tension and hydrostatic pressure. It fails to describe the process of keyhole formation and predicts lower values of melt velocity as compared with experimental results [2]. The second type of vaporization keyhole model [3] assumes that the keyhole is stabilized and held open when the vaporization rate is equal to the mass flow rate of molten metal coming into the keyhole due to hydrostatic pressure. Although this type of model can be used to study the transient behavior of short-duration spot welding, the predicted behavior for a longer time process does not coincide with experimental data of keyhole dynamics [2]. The model based on recoil pressure approach assumes that the recoil pressure acts as a driving force for melt flow and keyhole formation. The recoil pressure, acting toward the metal, is a kind of reaction force, as opposed to the action force caused by

the sudden "burst" of vapor from the metal when a laser beam impinges onto it. Klemens [4] proposed a model in which the motion of the melt was a significant factor for keyhole support. Allmen and Blatter [5] proposed that a Knudsen layer was formed above the molten region and extended beyond the keyhole surface for a few mean free paths. A discontinuity in temperature, pressure, and density occurred across the evaporation front, with the flow of hot vapor propagating as a shock wave outward. Solana et al. [6] further developed a mathematical model for the ablation process. Clucas et al. [7] developed a mathematical model for the keyhole welding process by suggesting a pressure and energy balance at the keyhole walls. The pressure balance was described in terms of surface tension, recoil pressure, and hydrodynamic pressure inside the keyhole. The energy was balanced between the laser light absorption energy and the heat by a conduction and ablation process.

During the keyhole formation process, the melt flow was very complex and unstable, which was considered to be the main cause for porosity formation [8]. Chen and Bos [9] investigated the role of melt flow in determining the operating characteristics of deep keyhole welding and explained why keyhole formation was considerably easier under a reduced environment pressure. Ducharme et al. [10] developed an integrated keyhole and weld pool model for thin metal sheets welding. The predicted shape of weld pool by this model agreed well with the experiment. Sudnik et al. [11] further analyzed the three driving forces of melt flow, including the force resulting from temperature-dependent surface tension, the friction force of metal vapor escaping from the capillary, and the movement of capillary relative to the workpiece. The shape of keyhole and the melt flow velocity were both calculated numerically.

Laser-induced plasma was presented as a result of ablation vapor from the keyhole wall, which was heated up to be partially ionized under the laser radiation. Laser-induced plasma is considered to be an important factor affecting keyhole behavior. Kapadia et al. [12] thought the high vapor temperature required for plasma
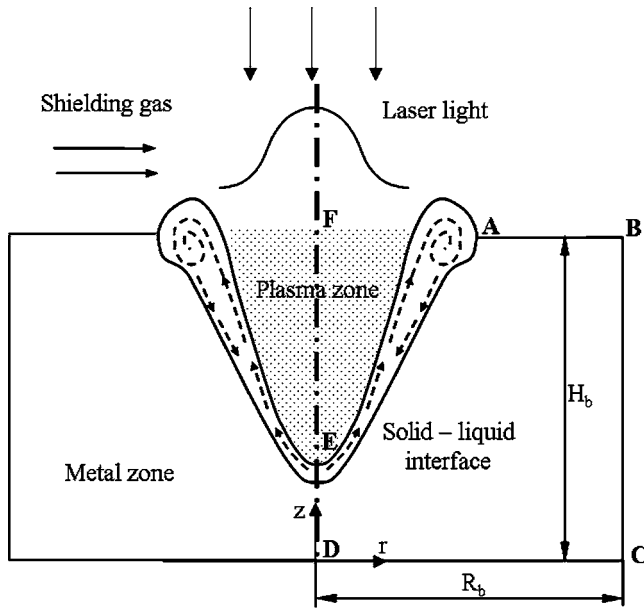
Copyright © 2006 by ASME

Fig. 1 Experimental setup and schematic sketch of static keyhole laser welding process

ignition was due to the small droplet radius and the presence of surface tension around droplet. Farson and Kim [13] simulated laser-induced evaporation and plume formation. They pointed out that a stable plume could be formed when the material surface irradiance was reduced. For the welding process at a lower flow rate of argon assistant gas, the Inverse Bremsstrahlung (IB) absorption in the plasma and the heating of the keyhole walls by the heat transfer from plasma could not be neglected [14]. Solana and Negro [15] analyzed the effect of multiple reflections on the keyhole wall, assuming the keyhole was axisymmetrical and that the keyhole wall acted as a free boundary. Metzbower [16] thought the absorption of laser energy in the keyhole was a function of laser frequency, welded material, and the temperature of the vapor.

Despite the large number of investigations referenced above, a comprehensive model of keyhole formation process that includes the simulations of temperature field, pressure balance, melt flow, free surface, laser-induced plasma, and multiple reflections has not yet been found. In this study, the transient keyhole formation and collapse process and weld pool dynamics during a pulsed laser welding has been investigated. The recoil pressure acts as a driving force for melt flow, which is considered as the primary influence factor for keyhole formation. The laser-induced plasma acts as an important role affecting keyhole behavior once it occurs.

2 Mathematical Models

Figure 1 shows the schematic sketch of a pulsed keyhole laser welding. A finite difference model employing the volume-of-fluid (VOF) technique [17] and the continuum formulation [18] is used to calculate the momentum and energy transport in the weld pool. The VOF technique can handle a transient deformed weld pool surface, while the continuum formulation can handle fusion and solidification for the liquid region, the mush zone, and the solid region. Plasma in the keyhole can be treated as the vapor of the weld material. Although the velocity and pressure change dramatically across the Knudsen layer, the generic translation vapor flow along the keyhole is neglected [14]. Meanwhile, the pressure along the keyhole can also be considered to be approximately constant [19] and will be comparable to the atmospheric pressure.

2.1 Metal Zone Simulation

2.1.1 Governing Equations. The governing differential equations used to describe heat and mass transfer and fluid flow in a cylindrical coordinate (r-z system) given by Chiang and Tsai [18] are modified and used in the current study.

Continuity

$$\frac{\partial}{\partial t}(\rho) + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{V}) = 0 \tag{1}$$

Momentum

$$\frac{\partial}{\partial t}(\rho u) + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{V} u) = \boldsymbol{\nabla} \cdot \left(\mu_l \frac{\rho}{\rho_l} \boldsymbol{\nabla} u \right) - \frac{\partial p}{\partial r} - \frac{u_l}{K} \frac{\rho}{\rho_l}(u - u_s)$$
$$- \frac{C\rho^2}{K^{0.5}\rho_l}|u - u_s|(u - u_s) - \boldsymbol{\nabla} \cdot (\rho f_s f_l \boldsymbol{V}_r u_r)$$
$$+ \boldsymbol{\nabla} \cdot \left[\mu_s u \boldsymbol{\nabla} \left(\frac{\rho}{\rho_l} \right) \right] \tag{2}$$

$$\frac{\partial}{\partial t}(\rho v) + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{V} v) = \rho g + \boldsymbol{\nabla} \cdot \left(\mu_l \frac{\rho}{\rho_l} \boldsymbol{\nabla} v \right) - \frac{\partial p}{\partial z} - \frac{u_l}{K} \frac{\rho}{\rho_l}(v - v_s)$$
$$- \frac{C\rho^2}{K^{0.5}\rho_l}|v - v_s|(v - v_s) - \boldsymbol{\nabla} \cdot (\rho f_s f_l \boldsymbol{V}_r v_r)$$
$$+ \boldsymbol{\nabla} \cdot \left[\mu_s v \boldsymbol{\nabla} \left(\frac{\rho}{\rho_l} \right) \right] + \rho g \beta_T(T - T_0) \tag{3}$$

Energy

$$\frac{\partial}{\partial t}(\rho h) + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{V} h) = \boldsymbol{\nabla} \cdot \left(\frac{k}{c_p} \boldsymbol{\nabla} h \right) - \boldsymbol{\nabla} \cdot \left(\frac{k}{c_p} \boldsymbol{\nabla} (h_s - h) \right)$$
$$- \boldsymbol{\nabla} \cdot [\rho(\boldsymbol{V} - \boldsymbol{V}_s)(h_l - h)] \tag{4}$$

The physical meaning of each term appearing in the above equations can be found in Ref. [18]. In Eqs. (1)–(4), the continuum density, specific heat, thermal conductivity, solid mass fraction, liquid mass fraction, velocity, and enthalpy are defined in Ref. [20].

2.1.2 Tracking of Free Surfaces. The algorithm of volume-of-fluid is used to track the dynamic free surface [17]. The fluid configuration is defined by a volume of fluid function $F(r,z,t)$, which tracks the location of free surface. The function F takes the value of one for the cell full of fluid and the value of zero for the empty cell. Cells with F values between zero and one are partially filled with fluid and identified as surface cells. The function F is governed by

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + (\boldsymbol{V} \cdot \boldsymbol{\nabla})F = 0 \tag{5}$$

2.1.3 Boundary Conditions. The corresponding boundary conditions of the metal zone simulation are divided into five categories according to their positions, as shown in Fig. 1.

Top surface inside the keyhole (AE in Fig. 1). For cells containing free surface, that is, cells that contain fluid but have one or more empty neighbors, in the direction normal to the free surface, the following pressures condition must be satisfied [21]

$$p = p_r + p_\sigma \tag{6}$$

where p is the pressure at the free surface in a direction normal to the local free surface, p_σ is surface tension, and p_r is recoil pressure. p_σ can be calculated as [20]

$$p_\sigma = \kappa \gamma \tag{7}$$

κ is the free surface curvature, given by [20]

$$\kappa = -\left[\mathbf{\nabla} \cdot \left(\frac{\mathbf{n}}{|\mathbf{n}|} \right) \right] = \frac{1}{|\mathbf{n}|}\left[\left(\frac{\mathbf{n}}{|\mathbf{n}|} \cdot \mathbf{\nabla} \right)|\mathbf{n}| - (\mathbf{\nabla} \cdot \mathbf{n}) \right] \qquad (8)$$

where $\mathbf{n}$ is the vector normal to the local free surface. For a pseudo-binary Fe–S system, the surface tension coefficient γ can be calculated as the function of temperature T and sulfur concentration f^α [22]

$$\gamma = 1.943 - 4.3 \times 10^{-4}(T - 1723) - RT \times 1.3$$
$$\times 10^{-8} \ln\left[1 + 0.00318 f^\alpha \exp\left(\frac{1.66 \times 10^8}{RT} \right) \right] \qquad (9)$$

The temperature-dependent Marangoni shear stress on the free surface in the direction tangential to the local surface is given by [23]

$$\tau_\mathbf{s} = \mu_l \frac{\partial (\mathbf{V} \cdot \mathbf{s})}{\partial \mathbf{n}} = \frac{\partial \gamma}{\partial T} \frac{\partial T}{\partial \mathbf{s}} \qquad (10)$$

Calculation of the evaporation-induced recoil pressure p_r is complicated by the existence of a Knudsen layer over the vaporizing surface. Based on Knight's model [24], the recoil pressure can be given by [25]

$$p_r = A P_s(T_w) = A B_0 / \sqrt{T_w} \exp(-U/T_w) \qquad (11)$$

where A is the numerical coefficient and B_0 is a vaporization constant. The coefficient A depends on the ambient pressure and its value varies from 0.55 for evaporation in the vacuum to 1 for the case of evaporation under a high ambient pressure. For practical values of the ambient pressure, the coefficient A is close to its minimal value of 0.55. B_0 is at the value of 1.78×10^{10}. T_w is the temperature of the surface liquid metal. The parameter U is defined as [25]

$$U = M_a H_v / (N_A k_B) \qquad (12)$$

where M_a is atomic mass, H_v is the latent heat of evaporation, N_A is Avogadro's number, and k_B is Boltzmann's constant.

The energy on the top free surface is balanced among the laser radiation, plasma-keyhole wall radiation, the heat dissipation through convection, and metal vaporization. In general, since the velocity of plume along the surface is zero [14], the heat loss due to convection is omitted. The energy balance can be given by the following formula

$$k \frac{\partial T}{\partial \mathbf{n}} = q_{laser} + q_{rad} - q_{evap} \qquad (13)$$

In this study, the liquid/vapor evaporation model is used due to the low intensity of laser radiation. Assuming the vaporization mechanism is surface vaporization, the heat loss due to evaporation can be given as [26]

$$q_{evap} = W H_v \qquad (14)$$

$$W = n_l \left(\frac{k_B T_l}{2\pi m_a} \right)^{0.5} \exp\left(-\frac{H_v}{k_B T_l} \right) - \theta_s n_v \left(\frac{k_B T_v}{2\pi m_a} \right)^{0.5} \qquad (15)$$

where T_l is the liquid surface temperature, m_a is the atom mass, n_l and n_v are the number of atoms per unit volume for liquid and vapor, respectively, and H_v and T_v represent latent heat of vaporization and temperature, respectively. θ_s denotes the probability that a vapor atom returning to the liquid surface from equilibrium conditions at the edge of the discontinuity layer manages to penetrate this layer to finally be absorbed on the liquid surface, which is in the range from 15% to 20%.

The laser heat flux q_{laser} comes from the Fresnel absorption of the incident intensity directly from the laser beam plus the incident intensity from the multiple reflections:

$$q_{laser} = I_0(r,z)(\mathbf{I}_0 \cdot \mathbf{n}_0)\alpha_{Fr}(\varphi_0) + \sum_{m=1}^{n} I_m(r,z)(\mathbf{I}_m \cdot \mathbf{n}_m)\alpha_{Fr}(\varphi_m) \qquad (16)$$

$$\alpha_{Fr}(\varphi) = 1 - \frac{1}{2}\left(\frac{1 + (1 - \varepsilon \cos \varphi)^2}{1 + (1 + \varepsilon \cos \varphi)^2} + \frac{\varepsilon^2 - 2\varepsilon \cos \varphi + 2 \cos^2 \varphi}{\varepsilon^2 + 2\varepsilon \cos \varphi + 2 \cos^2 \varphi} \right) \qquad (17)$$

where φ is the angle of incident light with the normal of keyhole surface, n is the total number of incident light from multiple reflections, $\mathbf{I}$ is the unit vector along the laser beam radiation direction, and $\mathbf{n}$ is the unit vector normal to the free surface. ε is a material-dependent coefficient. In CO_2 laser welding of mild steel, $\varepsilon = 0.2$ is used. $I_0(r,z)$ and $I_m(r,z)$ are the incident intensity from the laser beam and mth multiple reflection at the keyhole surface, respectively, which are given as

$$I_0(r,z) = I_c(r,z)\exp\left(-\int_0^{s_0} K_{pl}ds \right) \qquad (18)$$

$$I_m(r,z) = I_r(r,z)\exp\left(-\int_0^{s_m} K_{pl}ds \right) \qquad (19)$$

$$I_r(r,z) = I_0(r,z)(1 - \alpha_{Fr}) \qquad (20)$$

where $I_c(r,z)$ stands for the collimated incident laser beam intensity, $I_{r,m}(r,z)$ is the reflected laser beam intensity at mth reflections, $\int_0^z K_{pl}dz$ and $\int_0^{s_m} K_{pl}ds$ are the optical thickness of the laser transportation path, respectively, for the first incident and multiple reflections, K_{pl} is the plasma absorption coefficient due to the IB absorption [21]

$$K_{pl} = \frac{n_e n_i Z^2 e^6 2\pi}{6\sqrt{3} m \varepsilon_0^3 ch\omega^3 m_e^2}\left(\frac{m_e}{2\pi k_B T_e} \right)^{0.5}\left[1 - \exp\left(-\frac{\omega}{k_B T_e} \right) \right]\bar{g} \qquad (21)$$

where Z is the average ionic charge in the plasma, ω is the angular frequency of the laser radiation, ε_0 is dielectric constant, n_e and n_i are particle densities of electrons and ions, respectively, h is Planck's constant, m_e is the electron mass, T_e is the excitation temperature, c is the speed of light, and $\bar{g}$ is the quantum mechanical Gaunt factor. For weakly ionized plasma in the keyhole, the Saha equation [21] can be used to calculate the densities of plasma species

$$\frac{n_e n_i}{n_0} = \frac{g_e g_i}{g_0} \frac{(2\pi m_e k_B T_e)^{1.5}}{h^3} \exp\left(-\frac{E_i}{k_B T_e} \right) \qquad (22)$$

where g_e, g_i, and g_0 are the degeneracy factors for electrons, ions, and neutral atoms, respectively, and E_i is the ionization potential for the neutral atoms in the gas. Assuming the laser intensity distribution is ideal Gaussian-like, $I_c(r,z)$ can be calculated as [27]

$$I_c(r,z) = I_0 \left(\frac{r_f}{r_{f0}} \right)^2 \exp\left(-\frac{2r^2}{r_f^2} \right) \qquad (23)$$

where r_f is the beam radius, r_{f0} is the beam radius at the focal position, and I_0 is the peak intensity.

In laser welding, the keyhole surface temperature is much lower than that of the plume, so that the radiation and emission of surface can be omitted. Then q_{rad} can be simplified as

$$q_{rad} = \varepsilon \sigma (T_{pl}^4 - T^4) \qquad (24)$$

Here, T_{pl} is the temperature of the plasma.

Top surface outside the keyhole (AB in Fig. 1). The boundary condition on the top surface outside the keyhole is similar to that

inside the keyhole. The differences lie in the absence of plasma and multiple reflections. As shown in Fig. 1, there is a shielding gas flow above the base metal, which means that any plasma outside the keyhole will be blown away. Thus, Eq. (16) can be written as

$$q_{laser} = I_c(r,z)\alpha_{Fr}\cos\varphi \qquad (25)$$

Since there is no plasma and the temperature of shielding gas is much lower than that of the metal surface, the radiation heat flux can be given as

$$q_{rad} = -\varepsilon\sigma(T^4 - T_\infty^4) \qquad (26)$$

Here, T_∞ is the ambient temperature. Moreover, since there is a shielding gas flow over the surface, the convection term cannot be omitted. It can be given as

$$q_{conv} = h(T - T_\infty) \qquad (27)$$

Side surface (BC in Fig. 1).

$$-k\frac{\partial T}{\partial r} = q_{conv} \qquad (28)$$

$$u = 0, \quad v = 0 \qquad (29)$$

Bottom surface (CD in Fig. 1).

$$-k\frac{\partial T}{\partial z} = q_{conv} \qquad (30)$$

$$u = 0, \quad v = 0 \qquad (31)$$

Symmetrical axis (DE in Fig. 1).

$$\frac{\partial T}{\partial r} = 0 \qquad (32)$$

$$u = 0, \quad \frac{\partial v}{\partial r} = 0 \qquad (33)$$

2.2 Plasma Zone Simulation.

2.2.1 Governing Equations. In current study, the metal vapor in the keyhole is assumed to be a compressible, inviscous ideal gas. Since the heat production by viscous dissipation is rather small in laser welding, the energy equation can be simplified as [28]

$$\frac{\partial}{\partial t}(\rho_{pl}h_{pl}) = \boldsymbol{\nabla}\cdot\left(\frac{k_{pl}}{c_{pl}}\boldsymbol{\nabla}h_{pl}\right) + K_{pl}I_c\exp\left(-\int_0^s K_{pl}ds\right)$$
$$+ \sum_{m=1}^n K_{pl}I_{r,m}(r,z)\exp\left(-\int_0^{s_m} K_{pl}ds\right) \qquad (34)$$

where h_{pl} and ρ_{pl} represent the enthalpy and density of the plasma, respectively, k_{pl} and c_{pl} represent the thermal conductivity and specific heat of the plasma, respectively, s is the penetration depth of laser light in plasma, and K_{pl} denotes the plasma laser light IB absorption coefficient. When an intense laser pulse interacts with the vapor in the keyhole, a significant amount of laser radiation is absorbed by the ionized particles through IB absorption. For the laser-induced plasma inside the keyhole, the scattering effect is not significant compared with the absorbing and emitting effects. For simplicity, the plasma is assumed to be an absorbing-emitting medium and the laser intensity is exponentially attenuated inside the keyhole plasma.

2.2.2 Boundary Conditions

Bottom surface inside the keyhole (EA in Fig. 1). Close to the liquid wall inside the keyhole, there is the so-called Knudsen layer

where vaporization of material takes place. The vapor temperature across the Knudsen layer is discontinuous, which can be calculated by [24]:

$$\frac{T_K}{T_L} = \left[\sqrt{1 + \pi\left(\frac{\gamma-1}{\gamma+1}\frac{m}{2}\right)^2} - \sqrt{\pi}\frac{\gamma-1}{\gamma+1}\frac{m}{2}\right]^2 \qquad (35)$$

$$m = M_k\sqrt{\frac{2}{\gamma}} \qquad (36)$$

where T_K is the temperature outside of the Knudsen layer, T_L is the liquid surface temperature adjacent to the Knudsen layer, M_k is Mach number at the outer of the Knudsen layer, and γ is the ratio of specific heat. The value of m depends on the gas dynamics of the vapor flow away from the surface. Since it is assumed that there is no shock wave outside the Knudsen layer, the gas temperature outside the Knudsen layer is used as the boundary temperature. Thus, the boundary condition is given as [29]

$$T_{pl} = T_K \qquad (37)$$

Top surface outside the keyhole (FA in Fig. 1).

$$T_{pl} = T_\infty \qquad (38)$$

Symmetrical axis (EF in Fig. 1).

$$\frac{\partial T_{pl}}{\partial r} = 0 \qquad (39)$$

3 Numerical Method

In computation, the calculation of transport equations in metal zone and plasma zone are coupled. That is, the simulations of metal zone and plasma zone provide boundary conditions for each other. However, there are large spatial and physical differences between them. For a compromise between the result convergence and calculation time, different time resolution were used for these simulations. The governing equations (Eqs. (1)–(5) and (34)), all related supplemental equations and boundary conditions are solved through the following iterative scheme:

(1) Equations (1)–(4) are solved iteratively to obtain velocity, pressure, and temperature distributions under the associated boundary conditions for the metal zone.
(2) Equation (34) is solved iteratively to obtain the temperature field of the plasma in the keyhole under the associate boundary conditions.
(3) Iteration between steps (1) and (2).
(4) Solve VOF algorithm Eq. (5) to obtain the new domain for the metal zone and the plasma zone.
(5) Update the boundary conditions for the metal zone and the plasma zone.
(6) Advance to the next time step until the desired time is reached.

The technique for solving the partial differential equations is given by Wang and Tsai [20]. Following the Marker-And-Cell (MAC) scheme, the r and z velocity components are located at cell face centers on lines of constants r and z, respectively, while the pressure, VOF function, temperature, and absorbed laser flux are located at cell centers. Since the temperature and pressure field change more dramatically near the keyhole, a nonuniform grid system with 202×252 points is used for the total computational domain of $5.0\ mm \times 6.25\ mm$, in which smaller grids are arranged near the keyhole and larger grids for other parts. Due to the axis symmetry of the domain, only half of the grid points were used in the actual calculation. Calculations were executed on the DELL OPTIPLEX GX270 workstations with LINUX-REDHAT

9.0 OS and took 6 h of CPU time to simulate about 100 ms of real-time welding. The average time step is 10^{-4} s and the smallest time step is about 10^{-6} s.

4 Experimental Method

Bead-on-plate welding experiments were conducted by GM R&D engineers using a TRUMPF HL 4006D Nd:YAG laser system. The maximum output power of the laser at the workpiece is 4.0 kW and the wavelength is 1064 nm. The experimental setup is similar to that shown in Fig. 1. Plate specimens (40.0 mm × 3.0 mm) made of 304 standard steels were chemically cleaned and degreased. Argon was used as the protecting gas, the flow rate of which was 20 l/min. The laser beam pulse shape, modulation, and ramping are programmable or can be set externally. The diameter of the laser beam at the focus plane is 500 μm and the focus plane is on the top surface of the workpiece. The welding parameters, such as laser power, on-time duration, and beam diameter, were monitored and used in the mathematical model. Thermal profiles were recorded using K-type thermocouples welded to the symmetrical "A" (2 mm right to the laser beam center) and "B" (2 mm left to the laser beam center) locations underside of the coupon. Macroscopic analysis performed under stereoscopes and optical microscopes. An image acquisition system including a digital camera, a microscope, and a computer was used to capture and store the images. Adobe Photoshop® and Image Processing Tool Kit® software were used for the processing of the sample images. The experiments were conducted for various laser power levels and on-time durations. The weld penetration, width, and bead geometry, and temperature profiles at "A" and "B" locations for each case were measured.

5 Results and Discussion

The thermophysical properties and welding conditions used in the present study are summarized in Table 1. The following welding conditions are assumed in the model: laser power at the focus is 2.5 kW, the radius at focus is 0.25 mm, and the laser beam energy is in the Gaussian distribution. The laser power is assumed to increase from 0 to 2.5 kW in 1.0 ms and the pulse duration is 14.0 ms. The divergence of the laser beam is assumed to be negligible for a 3.0 mm metal thickness. The complete sequence of the keyhole formation and collapse and the associated weld pool fluid flow are calculated.

5.1 Keyhole Formation. Figure 2 shows the keyhole formation process. The corresponding temperature and velocity distributions are given, respectively, in Figs. 3 and 4. At the initial stage ($t<1.8$ ms), the laser energy is mainly used to heat up the base metal. As the temperature of the base metal increases, a small weld pool starts to appear under the laser beam. When the laser power reaches 2.5 kW at $t=1.0$ ms, the laser-induced recoil pressure starts to push down the molten metal under the laser beam. Since it takes time to accelerate the molten metal from a static condition, and the weld pool is thin at this time, the surface of the weld pool is nearly flat. Only part of the laser energy is absorbed by the metal via Fresnel absorption; the rest is reflected away. Heat conduction is the major heat transfer mode at this stage.

As the laser light continues to radiate, the temperature of the weld pool surface increases, as shown in Fig. 3. The large temperature difference along the free surface of the weld pool due to the Gaussian distribution of laser power leads to a high temperature-dependent Marangoni shear stress. As shown in Fig. 5, the gradient of surface tension coefficient on temperature change $\partial\gamma/\partial T$ is negative when T is above 2150 K. Since the temperature on the weld pool surface under laser radiation is more than 2150 K and the temperature gradient along the surface from the center to the outside of the weld pool $\partial T/\partial S$ is negative, the temperature-dependent Marangoni shear stress is positive along the weld pool surface, which drives the fluid to flow outwards. In

Table 1 Thermophysical properties of 304 stainless steel and process parameters

Nomenclature	Value
Specific heat of solid phase c_s (J kg^{-1} K^{-1})	700
Specific heat of liquid phase c_l (J kg^{-1} K^{-1})	780
Thermal conductivity of solid phase k_s (W m^{-1} K^{-1})	22
Thermal conductivity of liquid phase k_l (W m^{-1} K^{-1})	22
Density of solid phase ρ_s (kg m^{-3})	7200
Density of liquid phase ρ_l (kg m^{-3})	6900
Dynamic viscosity μ_l (kg m^{-1} s^{-1})	0.006
Latent heat of fusion H (J kg^{-1})	2.47×10^5
Solidus temperature T_s (K)	1670
Liquidus temperature T_l (K)	1727
Boiling temperature T_v (K)	3375
Latent heat of vaporization H_v (J kg^{-1})	6.34×10^6
Laser power P (kW)	2.5
Laser beam radius at focus r_{f0} (mm)	0.25
Laser beam radius r_f (mm)	0.25
Substrate initial temperature T_0 (K)	300
Surrounding temperature T_{gas} (K)	300
Ambient temperature T_∞ (K)	300
Thickness of substrate metal H_b (mm)	3.0
Radius of substrate metal R_b (mm)	20.0
Substrate sulfur concentration C_s (ppm)	100
Atmospheric pressure P_{atm} (Pa)	1.013×10^5
Gas constant R (J kg^{-1} mole^{-1})	8.3×10^3
Average ionic charge in the plasma Z	1
Angular frequency of the laser radiation ω (rad s^{-1})	1.78×10^{14}
Quantum mechanical Gaunt factor $\bar{g}$	1.5
Excitation temperature T_e (K)	9×10^3
Degeneracy factors for ions g_i	30
Degeneracy factors for neutral atoms g_0	25
Degeneracy factors for electrons g_e	30
Speed of light c (m s^{-1})	3×10^8
Boltzmann's constant k_B (J K^{-1})	1.38×10^{-23}
Planck's constant h (J s)	6.625×10^{-34}

the surface area where temperature is below 2150 K, surface tension coefficient gradient will change to be a positive value. Since the temperature gradient is still negative, the Marangoni shear stress forces the liquid metal to flow inwards. A crater is then formed at some point of the weld pool where the temperature is above 2150 K, as shown in Fig. 2. Moreover, higher surface temperature leads to higher recoil pressure, which pushes the liquid metal under laser radiation downwards quickly. As shown in Fig. 2, although the liquid-solid interface keeps moving downwards, the liquid region of the weld pool under laser radiation remains only as a thin layer because the strong recoil pressure squeezes the liquid metal to flow outwards. Under the hydrodynamic force and the Marangoni shear stress, this squeezed liquid metal flows upwards along the keyhole wall which facilitates the formation of a crater. As the surface level moves down and the crater level increases, a keyhole is formed in the weld pool at $t=6.0$ ms.

5.2 The Effects of Laser-Induced Plasma and Multiple Reflections. Once the keyhole comes into being, as shown in Fig. 2 at $t=6.0$ ms, some metal vapor will be trapped inside it. The incoming laser beam interacts with this metal vapor and part of laser energy is absorbed by it. Since the heat capacity of the vapor is rather low, the temperature of the vapor increases very quickly. When the temperature of the vapor exceeds around 8000 K [15], laser-induced plasma will be formed by the ionization process. The physical properties of plasma are quite different from that of the prior vapor, which absorbs more laser energy through the IB absorption process.

The laser-induced plasma plays twofold roles in keyhole formation. As shown in Fig. 6, the coefficient of the IB absorption increases with the increase of plasma temperature. Hence, once
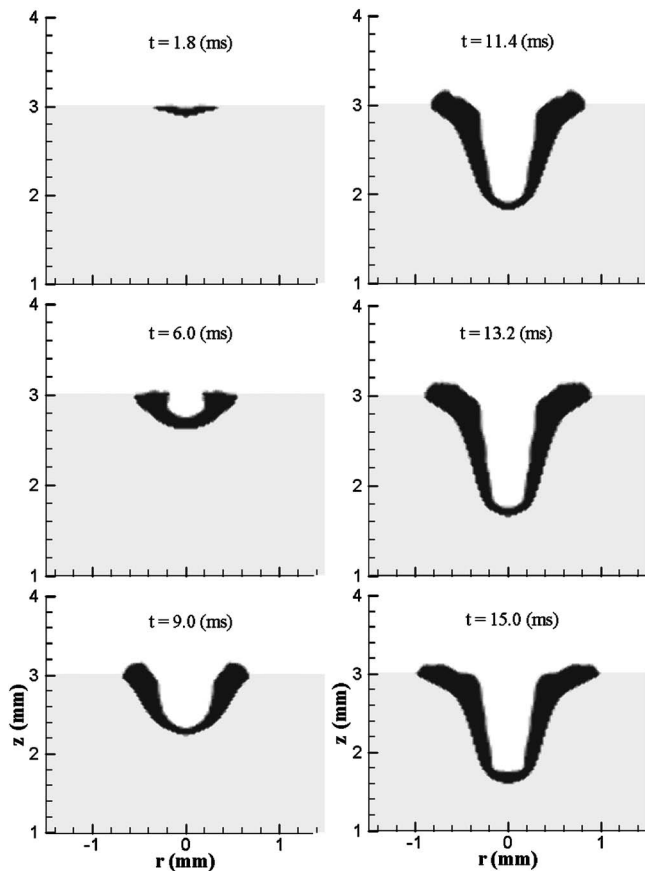
Fig. 2   A sequence of liquid metal evolution during the keyhole formation process

Fig. 3   The corresponding temperature distributions for the case shown in Fig. 2

plasma comes into being, its temperature will increase very quickly, as indicated in Fig. 3. This hot plasma separates the keyhole wall from the cold shielding gas, which reduces the heat loss from the keyhole wall to the surroundings. Moreover, the hot plasma heats up the surrounding keyhole wall via radiation. This facilitates the temperature increase on the keyhole wall when the depth of keyhole is not large. As shown in Fig. 3, the temperature on the keyhole wall under laser radiation keeps increasing. Therefore, the recoil pressure becomes stronger accordingly, which helps the depth of keyhole grow very quickly. As shown in Fig. 7, the drilling speed increases quickly once the keyhole appears, and maintains the value for the time between $t=6.0$ ms and $t=10.4$ ms.

When the plasma temperature is above 10,000 K, as shown in Fig. 6, the IB absorption coefficient decreases with the increasing plasma temperature. Although the coefficient of the IB absorption decreases with the small increase of plasma temperature, the travel length of laser light increases with the increasing keyhole depth, the overall effect of the plasma absorption becomes more and more remarkable, and it will block more percentage of the laser energy to reach the keyhole bottom wall. Meanwhile, the temperature of plasma only increases a little. The heat radiation from the plasma also increases a little correspondingly. Thus, the blocking effect on laser light becomes more and more dominant, which eventually will make the plasma become a negative factor for keyhole depth increase at certain time. As shown in Fig. 7, the keyholing speed drops quickly after $t=11.0$ ms.

The formation of the keyhole also enhances the laser light absorption through a mechanism called multiple reflections. As shown in Fig. 2, the surface of the weld pool is nearly flat at the beginning. Once laser light irradiates the flat surface, only part of the energy is absorbed by the surface via the Fresnel absorption
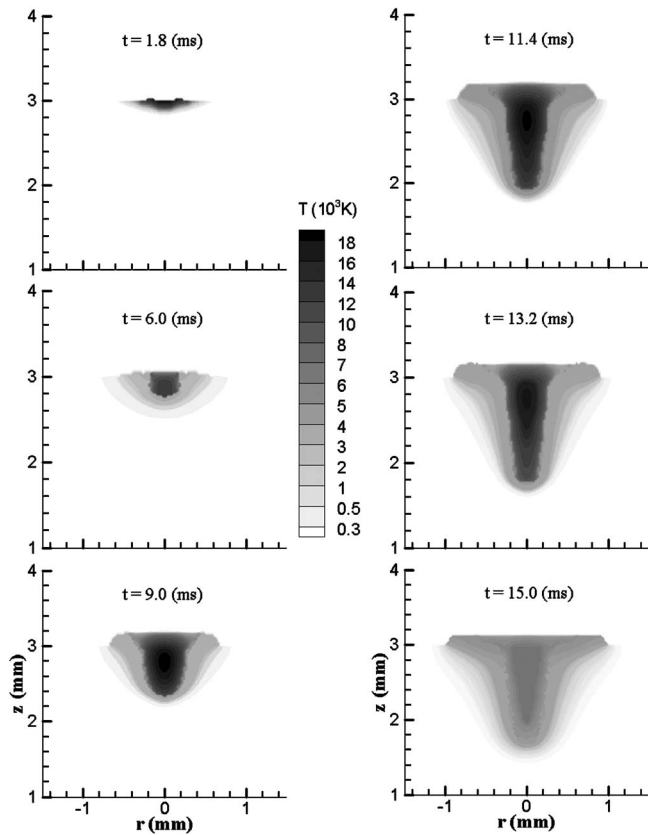
mechanism. A large amount of the laser energy is reflected back. Since the direction of the reflected light is contrary to its incident direction, this part of the laser light will never reach the weld pool surface again. However, once the keyhole is formed, the surface along the keyhole wall is not flat. Part of the reflected light will be incident on the keyhole wall again and again. During each incidence, part of the laser energy is absorbed via the Fresnel absorption mechanism. This brings more laser energy input to the keyhole wall, which leads to higher recoil pressure on the keyhole wall. Moreover, as the keyhole deepens, the possibilities and times of multiple reflections increase. When the reflected laser light travels in the plasma, part of its energy will also be absorbed by the plasma to increase the plasma temperature.

5.3   Fluid Flow and Weld Pool Dynamics. As shown in Fig. 4 at $t=1.8$ ms, since the weld pool is small and it takes time to accelerate the molten metal, the fluid flow is not significant. Once the keyhole is formed, with the strong action of the laser-induced recoil pressure, some strong velocities appear in the weld pool at $t=6.0$ ms. At $t=9.0$ ms, on the bottom of the keyhole wall, higher recoil pressure leaves a thinner liquid layer compared with that at $t=6.0$ ms. Although the hydrostatic pressure and surface tension at the bottom of the keyhole becomes larger and larger as keyhole deepens, which makes it difficult for the drilling process of the keyhole, the velocity of liquid along the bottom surface does not decrease, as shown in Fig. 4. Meanwhile, the hydrostatic pressure at the bottom of the keyhole increases as the keyhole deepens, which makes it more and more difficult for the recoil pressure to pump the liquid upwards. Moreover, the temperature difference along the sidewall of the keyhole decreases as the keyhole depth increases. This decreases the Marangoni shear stress and makes it more difficult for the squeezed liquid metal to flow upwards. As shown in Fig. 4, both the downward velocities at the bottom keyhole and the upward velocities along the sidewall of the keyhole
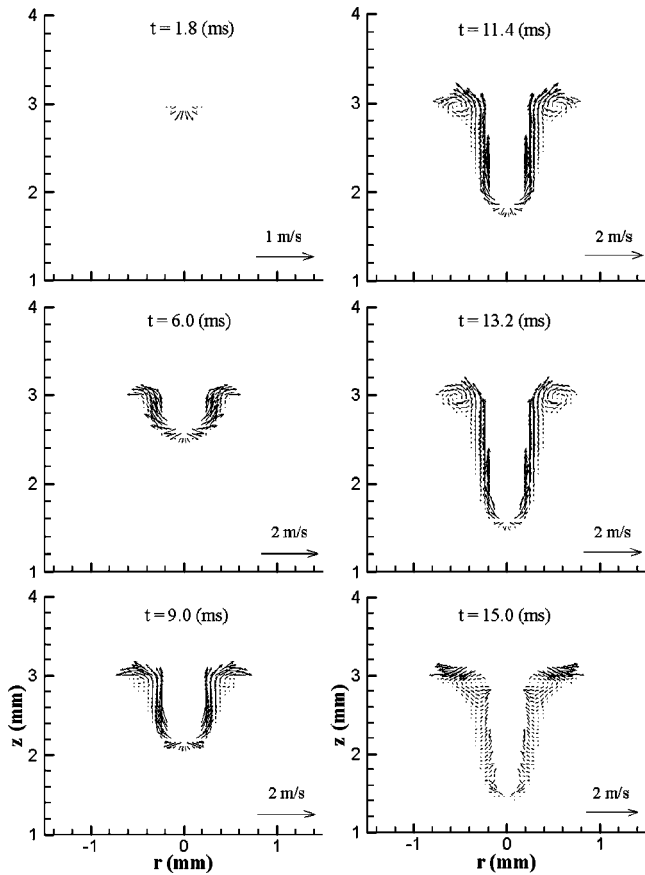
Fig. 4 The corresponding velocity distributions for the case shown in Fig. 2

Fig. 6 Coefficient of Inverse-Bremsstrahlung absorption as a function of plasma temperature from Ref. [16]

hole causes the liquid-solid interface in the keyhole to move outwards, which enlarges the heat affect zone in the weld substrate. Hence, as shown in Fig. 2, the liquid region of keyhole is enlarged at $t = 15.0$ ms compared with that at $t = 13.2$ ms and the keyhole depth does not increase remarkably at this period. Also, as noticed in the figure, the liquid metal on the bottom of the keyhole is thicker at $t = 15.0$ ms than that at $t = 13.2$ ms. This is because, with more and more laser energy being absorbed by the plasma, the recoil pressure on the bottom of the keyhole is decreasing at $t = 15.0$ ms, hence more liquid metal is accumulated there.

As shown in Fig. 4, there is a vortex in the upper part of the weld pool at $t = 13.2$ ms, which is the result of combined action of four driving forces: temperature-dependent Marangoni shear stress force, surface tension, hydrodynamic force, and hydrostatic force. On the upper part of the vortex, the liquid is driven from the hot edge to the cold edge by the Marangoni shear stress force when the temperature of liquid is above 2150 K. As the liquid metal moves to the cold edge, the temperature gradient becomes smaller and smaller which makes the Marangoni shear stress force decreases. At the point where the liquid temperature is under 2150 K, the Marangoni shear stress will change its direction. This change prevents the liquid from moving outwards. At the same time, the surface tension on the cold edge also blocks the liquid to flow outwards. Therefore, the velocity vector along the r direction

are smaller at $t = 13.2$ ms compared with those at $t = 11.4$ ms.

As shown in Fig. 7, the keyholing speed drops quickly after $t > 10.5$ ms. This means more laser energy is absorbed by the keyhole plasma, especially by that in the upper part of the keyhole. Thus, it makes the plasma at that portion reach a very high temperature, as shown in Fig. 3. This hot plasma also increases the heat input to the keyhole wall. Meanwhile, the multiple reflections also bring more and more energy to the sidewall of the keyhole as keyhole deepens. The increased heat input to the sidewall of key-
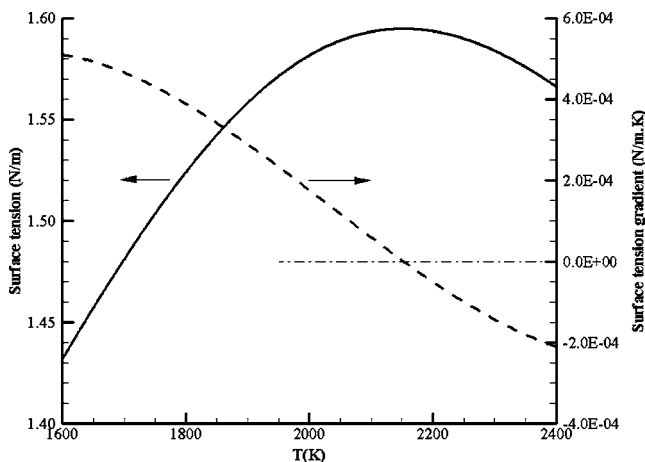

Fig. 5 Surface tension and its gradients as a function of temperature for the pseudo-binary Fe-S system with 300 ppm sulfur
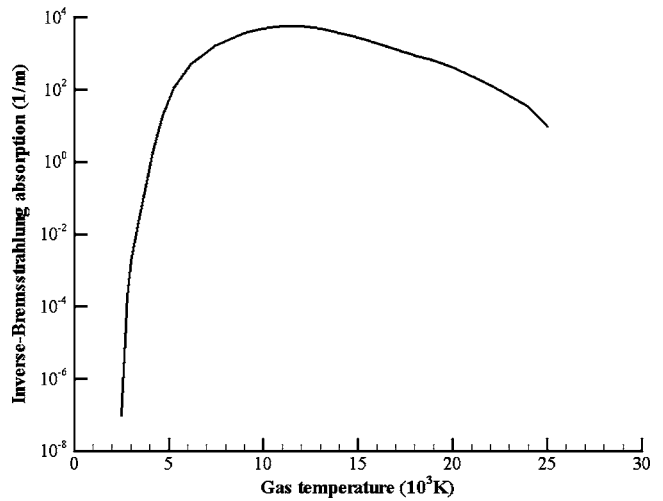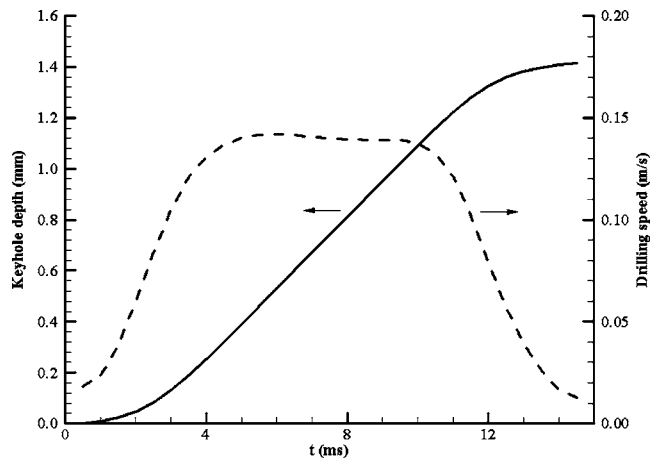

Fig. 7 Keyhole depth and keyholing speed as a function of time
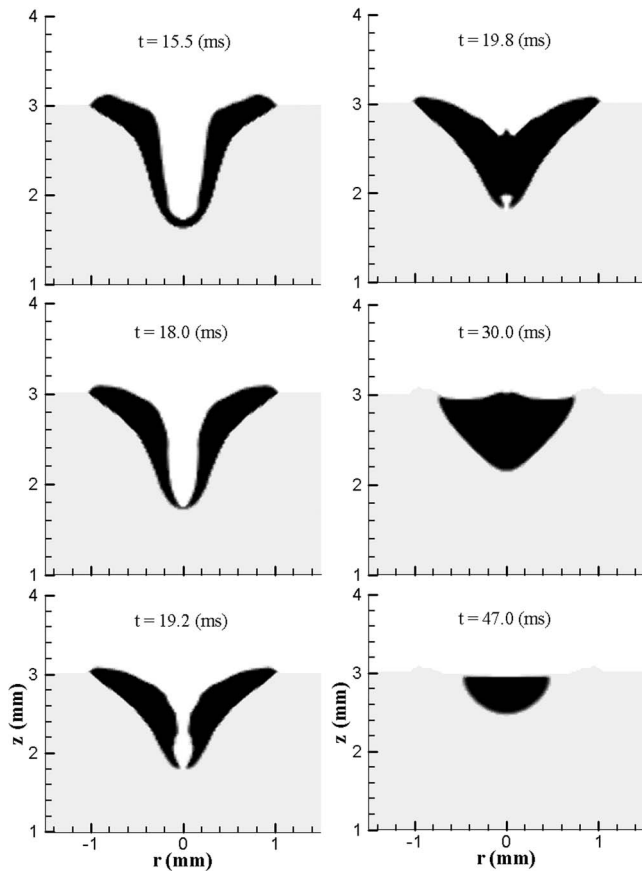
Fig. 8   A sequence of liquid metal evolution during the keyhole collapse and solidification processes

Fig. 9   The corresponding temperature distributions for the case shown in Fig. 8

becomes smaller and smaller. Meanwhile, the hydrostatic force and surface tension cause the liquid to flow downwards, which makes the velocity along the z direction become larger and larger. Finally, the liquid metal changes its flow direction to flow downwards at the upper of the cold edge. When this downward-flowing liquid metal meets the liquid-solid interface, it is blocked by this interface and flows along the interface. Meanwhile, the hydrodynamic force of the squeezed liquid from the lower part of keyhole makes the liquid flow upwards. Thus, the liquid changes its direction to flow upwards at the lower part of hot edge. As the liquid flows upwards, the hydrodynamic force is counteracted by the hydrostatic force and surface tension, which retards the upward tendency. Meanwhile, the Marangoni shear stress force pushes the liquid to flow toward the cold edge. Thus, the liquid flow rotates clockwise on the shoulder of keyhole and a vortex is formed as shown in the figure. The flow velocity is on the order of 10 cm/s, which agrees fairly well with the experimental results [11]. This vortex enhances the heat transfer along the liquid-solid interface and enlarges the liquid region on shoulder of the keyhole, as shown in Fig. 2.

5.4 Keyhole Collapse and Porosity Formation. When the laser pulse is shut off at $t=14.0$ ms, the recoil pressure disappears quickly and the keyhole collapses under hydrostatic pressure and surface tension. Figure 8 shows a sequence of the liquid metal evolution during the keyhole collapse and solidification processes. The corresponding temperature and velocity distributions are given in Figs. 9 and 10. Once the laser is shut off, the hot plasma radiation is the only heat input source for keyhole wall. However, since there is no heat input to the plasma and the heat capacity of plasma is very small, the temperature of the plasma drops very quickly. Meanwhile, the heat conduction from the keyhole wall to
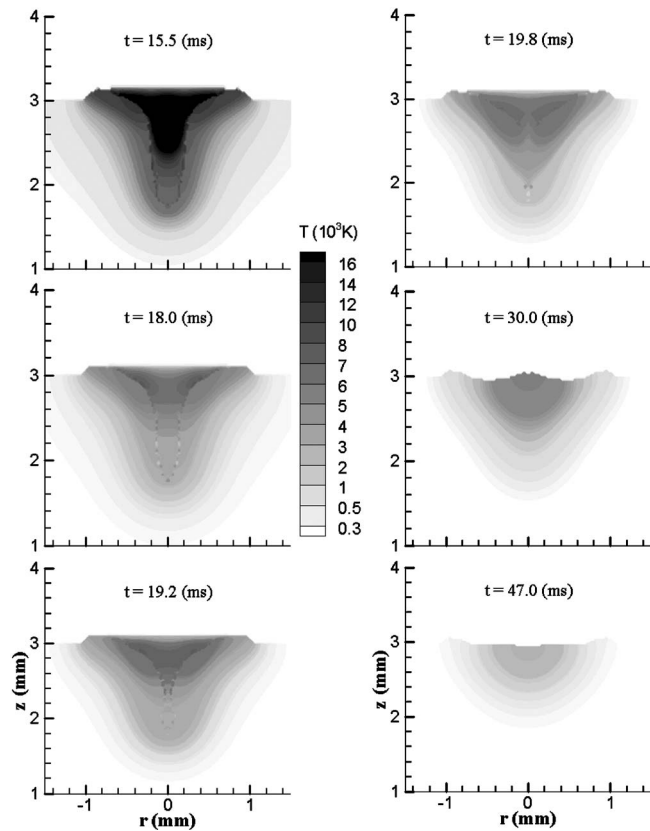
the surrounding metal is very strong due to the high aspect ratio and high temperature gradient. Hence, as shown in Fig. 9, the temperature of the keyhole wall drops very quickly, especially in the lower part of the keyhole since there is only a thin layer of the metal liquid. Due to this quick temperature drop, the thin layer of liquid metal on the bottom of keyhole completely solidifies at $t=19.2$ ms in Fig. 9.

At the initial stage of the collapse, the temperature gradient along the sidewall surface of the keyhole is negative from the bottom to the top and the temperature of the liquid metal is above 2150 K. Thus, the Marangoni shear stress is still positive, which helps the liquid to flow upwards. Since the temperature of the metal on the bottom surface drops much faster than that on the upper surface, the Marangoni shear stress there decreases quickly. Meanwhile, the hydrodynamic pressure of the squeezed liquid metal also decreases very quickly due to the removal of recoil pressure. The surface tension and hydrostatic pressure then make the liquid metal have a tendency to fill back the keyhole. However, the velocity of liquid metal needs a time to change its direction, especially on the lower part of the keyhole wall due to their original high velocities inertia. On the keyhole shoulder, the Marangoni shear stress also decreases quickly with the decreasing temperature gradient. Hence, with the action of hydrostatic pressure and surface tension, the liquid metal on the upper part of the keyhole starts to flow inwards and downwards at $t=15.5$ ms in Fig. 10. Since there is certain amount of liquid metal there and the flow friction along the liquid-solid interface is larger for the thin liquid layer, the back-filling velocities of the liquid metal on the upper part will be accelerated more easily than those on the lower part of keyhole. Thus, the keyhole is closed on the top first as shown in Fig. 8. Meanwhile, the liquid metal continues to be accelerated by the hydrostatic force to flow downwards along the keyhole wall and finally the keyhole is refilled at $t=30.0$ ms. As
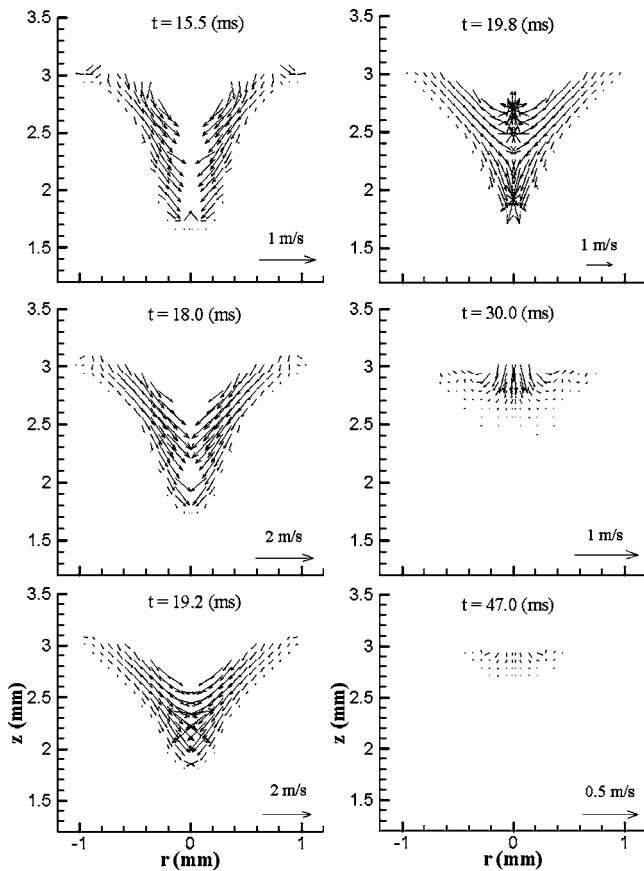
Fig. 10 The corresponding velocity distributions for the case shown in Fig. 8

shown in Fig. 8 from $t=19.8$ ms to $t=30.0$ ms, the liquid region also shrinks as the liquid refills the keyhole, especially at the bottom because of fast conduction heat loss and low heat capacity. Finally, the liquid metal on the bottom completely solidifies. During the keyhole collapse process, the hot liquid metal cannot reach the far end of the top surface due to the disappearance of recoil pressure. Moreover, the velocity of the liquid metal flowing toward the edge decreases during the back-filling process. Thus, the liquid metal at the far end of the top surface cannot fill back and solidifies there, which results in a rough top surface on the final weld, as shown in Fig. 8 at $t=30.0$ ms.

5.5 Comparison Between Model Predictions and Experiments. To validate the numerical model, experiments were performed using the parameters listed before. In this study, the size and geometry of the weld bead are compared with simulated results in Fig. 11 and the temperature distribution comparisons are
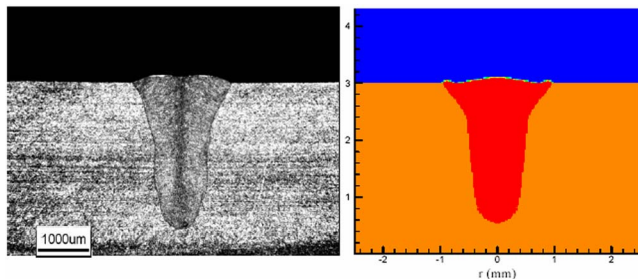

Fig. 11 The comparison of the weld bead geometry between experimental result and model prediction
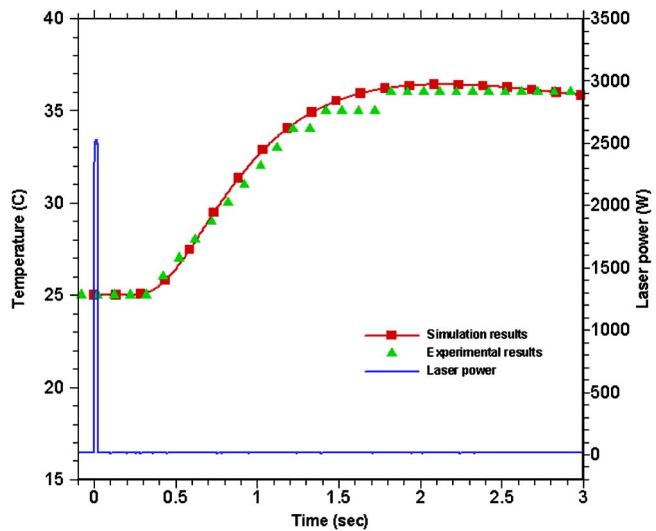
Fig. 12 The comparison of the temperature history at location "A" between experimental result and model prediction

shown in Fig. 12. As shown in Fig. 11, when the laser power is 2.5 kW and the laser pulse duration is 20.0 ms, the penetration depth predicted by the model is 2.52 mm and it is 2.56 mm as shown in the experiment. The weld bead width is 1.96 mm from the model prediction and is 2.0 mm in the experiment. The top surface shape predicted by the model is slightly different from that shown in the experiment, and the size of the welded zone shown in the upper part of the welding coupon in the experiment is slightly larger than that predicted by the model. The difference between the model prediction and experimental results may be due to the difference of the input values of parameters (laser power, pulse duration, etc.) in the model and those in the experiments and the shielding gas effect. Overall, the weld bead geometry predicted by the model agrees well with that from the experiment. Also as shown in Fig. 12, the temperature history at location "A" predicted by the model agrees very well with that measured by the experiments. More comparison results are listed in Table 2 and, as shown, good agreements between experimental and computational results are obtained.

6 Conclusions

A mathematical model has been developed to simulate the transient keyhole formation, heat transfer, and fluid flow, and keyhole collapse process and the corresponding experiments were conducted to validate its efficiency. The model is divided into two submodels. One is to calculate the mass, momentum, and energy transport in the weld pool and the other is to calculate the energy transport in plasma. These two submodels provide boundary conditions for each other. The VOF technique is used to handle the free surface, and complicated velocity and temperature distributions are calculated.

Recoil pressure plays a key role in the keyhole formation, which pushes down the liquid in the weld pool and acts as the main driving force for the keyhole formation. This laser-induced recoil force combined with the Marangoni shear force, hydrodynamic force, and hydrostatic force causes very complicated fluid flow in the weld pool. A vortex is found at the upper part of keyhole, which is helpful in increasing the heat and mass transfer during the welding process.

Laser-induced plasma plays two-fold roles during keyhole formation. The positive effect is that the IB absorption increases the energy coupling, which facilitates the keyhole formation at the initial stage. The negative effect is that, when the keyhole reaches

Table 2   Comparison of weld bead size between model predictions and experiment results

Laser Power (On-time) (kW)	Penetration (mm)		Width (mm)	
	Experiment	Calculated	Experiment	Calculated
1.8 (10.0 ms)	1.22	1.18	1.12	1.20
1.8 (20.0 ms)	2.01	2.04	1.74	1.82
1.8 (60.0 ms)	1.98	2.08	1.94	1.82
2.5 (20.0 ms)	2.56	2.52	2.00	1.96

certain depth, more laser energy is absorbed by the plasma and less laser energy could reach the keyhole wall, which prevents the keyhole from deepening further. Multiple reflections can bring more laser energy inside the keyhole to help the keyhole formation process.

During the keyhole collapse process, the keyhole closes on the top first and the liquid metal from the top then flows downwards to refill the keyhole. Porosity could be introduced during the keyhole collapse. Its formation is dependent on two factors: (1) solidification rate after the shut-off of the laser power and (2) backfilling speed of the liquid metal. The detailed discussion can be found in the forthcoming papers. The liquid metal on the far end of top surface cannot flow back in the solidification process, which results in a rough surface on the final weld. The temperature history and geometry of the weld bead predicted by the proposed model agree very well with the experimental results, which indicate the proposed mathematical model can be used as a solid base for future studies.

Acknowledgment

Nomenclature

A = constant in Eq. (11)
A_v = constant in Eq. (23)
B_0 = vaporization constant in Eq. (11)
c_p = specific heat
C = coefficient defined in Eq. (3)
d = dendrite arm spacing
E_i = ionization potential for neutral atom
f = mass fraction
f^α = sulfur concentration
F = volume of fluid function
F_{vis} = viscous energy dissipation
g = gravitational acceleration
g_e = degeneracy factor of electron particle
g_i = degeneracy factor of ion particle
g_0 = degeneracy factor of neutral atom
h = enthalpy for metal zone simulation
h_{pl} = enthalpy for plasma
h_c = convective heat-transfer coefficient
H = latent heat of fusion
H_v = latent heat for liquid-vapor
H_b = thickness of base metal
I = laser intensity
I_r = laser intensity leaving keyhole
k = thermal conductivity
K = permeability function in Eq. (2)
K_{pl} = plasma laser absorption coefficient
k_i = thermal conductivity for component i
k_B = Boltzmann's constant
M_a = atomic mass
m = times of reflection
m_e = electron mass

n = vector normal to local free surface
I = vector along the radiation direction
n_e = electron particle density in plasma
n_i = ion particle density in plasma
n_0 = neutral atom density in plasma
N_A = Avogadro's number
p = pressure for metal zone simulation
P_{atm} = atmospheric pressure
p_r = recoil pressure
p_σ = surface tension
q_{conv} = heat loss by convection
q_{evap} = heat loss by evaporation
q_{laser} = heat input flux by laser radiation
q_{radi} = heat loss by radiation
r-z = cylindrical coordinate system
r_f = laser beam radius
R = gas constant
R_b = radius of base metal
s = penetration depth of laser in plasma
s = vector tangential to local free surface
t = time
Δt = time step for metal zone simulation
Δt_{pl} = time step for plasma zone simulation
T = temperature
T_v = vaporization temperature
T_w = temperature on keyhole wall
u = velocity in the r direction
v = velocity in the z direction
V = velocity vector
V_I = ionization voltage
W = melt evaporation rate

Greek symbols

α_{pl} = fraction of laser absorbed by plasma
α_{IB} = IB absorption coefficient
α_{Fr} = Fresnel absorption coefficient
α_{mr} = plasma absorption coefficient at mth reflection
α_I = degree of ionization
β_T = thermal expansion coefficient
γ = surface tension coefficient
$\partial\gamma/\partial T$ = surface tension temperature gradient
ε = surface radiation emissivity
κ = free surface curvature
μ_l = dynamic viscosity
σ = Stefan-Boltzmann constant
ρ = density

Subscripts

0 = initial value
c = original incident laser light
IB = inverse bremsstrahlung absorption
l = liquid phase
r = relative to solid velocity
(r,m) = mth reflected laser beam
pl = plasma
s = solid phase

References

[1] Semak, V. V., Hopkins, J. A., McCay, M. H., and McCay, T. D., 1994, "A Concept for a Hydrodynamic Model of Keyhole Formation and Support During Laser Welding," *Proc. ICALEO*, pp. 641–650, Oct. 17–20, Orlando, FL.

[2] Semak, V. V., Hopkins, J. A., McCay, M. H., and McCay, T. D., 1994, "Dynamics of Penetration Depth During Laser Welding," *Proc. ICALEO*, pp. 17–20, Oct. 17–20, Orlando, FL.

[3] Hopkins, J. A., McCay, T. D., McCay, M. H., and Eraslan, A., 1993, "Transient Predictions of CO_2 Laser Spot Welds in Inconel 718," *Proc. ICALEO*, pp. 24–28, Oct. 24–28, Orlando, FL.

[4] Klemens, P. G., 1976, Heat Balance and Flow Conditions for Electron Beam and Laser Welding, J. Appl. Phys., **47**, pp. 2165–2174.

[5] Allmen, M., and Blatter, A., 1995, *Laser-Beam Interaction with Material*, 2nd ed., Springer-Verlag, Berlin.

[6] Solana, P., Kapadia, P., and Dowden, J., 1998, "Surface Depression and Ablation for a Weld Pool in Material Processing: a Mathematical Model," *Proc. ICALEO*, Sec. F, pp. 142–147, Nov. 16–19, Orlando, FL.

[7] Clucas, A., Ducharme, R., Kapadia, P., Dowden, J., and Steen, W., 1998, "A Mathematical Model of Laser Keyhole Welding Using a Pressure and Energy Balance at the Keyhole Walls," *Proc. ICALEO*, Sec. F, pp. 123–131, Nov. 16–19, Orlando, FL.

[8] Matsunawa, A., Kim, J., Seto, N., Mizutani, M., and Katayama, S., 1998, "Dynamic of Keyhole and Molten Pool in Laser Welding," J. Laser Appl., **10**, pp. 247–254.

[9] Chen, M. M., and Bos, J. A., 1998, "Melt Flow in Deep Penetration Welding," *Proc. ICALEO*, Sec. F, pp. 187–196, Nov. 16–19, Orlando, FL.

[10] Ducharme, R., Williams, K., Kapadia, P., Dowden, J., Steen, B., and Glowacki, M., 1994, "The Laser Welding of Thin Metal Sheets: an Integrated Keyhole and Weld Pool Model With Supporting Experiments," J. Phys. D, **27**, pp. 1619–1627.

[11] Sudnik, R., Rada, D., Breitschwerdt, S., and Erofeew, W., 2000, "Numerical Simulation of Weld Pool Geometry in Laser Beam Welding," J. Phys. D, **33**, pp. 662–671.

[12] Kapadia, P., Dowden, J., and Ducharme, R., 1996, "A Mathematical Model of Ablation in the Keyhole and Droplet Formation in the Plume in Deep Penetration Laser Welding," *Proc. ICALEO*, Sec. B, pp. 106–114, Oct. 14–17, Detroit, MI.

[13] Farson, D. F., and Kim, K. R., 1998, "Simulation of Laser Evaporation and Plume," *Proc. ICALEO*, Sec. F, pp. 197–206, Nov. 16–19, Orlando, FL.

[14] Miyamoto, I., Ohmura, E., and Maede, T., 1997, "Dynamic Behavior of Plume and Keyhole in CO_2 Laser Welding," *Proc. ICALEO*, Sec. G, pp. 210–218, Nov. 17–20, San Diego, CA.

[15] Solana, P., and Negro, G., 1997, "A Study of the Effect of Multiple Reflections on the Shape of the Keyhole in the Laser Processing of Material," J. Phys. D, **30**, pp. 3216–3222.

[16] Metzbower, E. A., 1997, "Absorption in the Keyhole," *Proc. ICALEO*, Sec. G, pp. 16–25, Nov. 17–20, San Diego, CA.

[17] Kothe, D. B., Mjolsness, R. C., and Torrey, M. D., 1991, "Ripple: A Computer Program for Incompressible Flows With Free Surfaces," LA-12007-MS, Los Alamos National Laboratory.

[18] Chiang, K. C., and Tsai, H. L., 1992, "Shrinkage-Induced Fluid Flow and Domain Change in Two-Dimensional Alloy Solidification," Int. J. Heat Mass Transfer, **35**, pp. 1763–1769.

[19] Dowden, J., Postacioglu, N., Davis, M., and Kapadia, P., 1987, J. Phys. D, **20**, pp. 36–44.

[20] Wang, Y., and Tsai, H. L., 2001, "Impingement of Filler Droplets and Weld Pool Dynamics During Gas Metal Arc Welding Process," Int. J. Heat Mass Transfer, **44**, pp. 2067–2080.

[21] Duley, W., 1999, *Laser Welding*, John Wiley & Sons Inc., New York.

[22] Sahoo, P., DeBroy, T., and Mcnallan, M. J., 1988, "Surface Tension of Binary Metal-Surface Active Solute Systems Under Conditions Relevant to Welding Metallurgy," Metall. Trans. B, **19**, pp. 483–491.

[23] Choo, R. T. C., Szekely, J., and David, S. A., 1992, "On the Calculation of the Free Surface Temperature of Gas-Tungsten-Arc Weld Pools From First Principles: Part II. Modeling the Weld Pool and Comparison With Experiments," Metall. Trans. B, **23**, pp. 371–384.

[24] Knight, C. J., 1979, "Theoretical Modeling of Rapid Surface Vaporization With Back Pressure," AIAA J., **17**, pp. 519–523.

[25] Semak, V., and Matsunawa, A., 1997, "The Role of Recoil Pressure in Energy Balance During Laser Materials Processing," J. Phys. D, **30**, pp. 2541–2552.

[26] Landau, L. D., and Lifshitz, E. M., 1980, *Statistical Physics*, 3rd ed., Pergamon, New York.

[27] Kaplan, A., 1994, "A Model of Deep Penetration Laser Welding Based on Calculation of the Keyhole Profile," J. Phys. D, **27**, pp. 1805–1814.

[28] Siegel, R., and Howell, J. R., 1992, *Thermal Radiation Heat Transfer*, 3rd ed., Hemisphere Publishing Corp., Washington, DC, Chap. 13.

[29] Ho, R., Grigoropoulos, C. P., and Humphrey, J. A. C., 1996, "Gas Dynamics and Radiation Heat Transfer in the Vapor Plume Produced by Pulsed Laser Irradiation of Aluminum," J. Appl. Phys., **79**, pp. 7205–7215.

Heat Transfer Performance During Condensation Inside Horizontal Smooth, Micro-Fin and Herringbone Tubes

Adriaan Lambrechts
Master's Degree Student
University of Johannesburg,
South Africa

Leon Liebenberg[1]
Associate Professor
Department of Mechanical and Aeronautical
Engineering,
University of Pretoria,
Pretoria 0002, South Africa
e-mail: lieb@up.ac.za

Arthur E. Bergles
Glenn L. Martin Institute Professor
University of Maryland,
College Park, MD 20742-3035
e-mail: abergles@aol.com

Josua P. Meyer
Professor and Head
Department of Mechanical and Aeronautical
Engineering,
University of Pretoria,
Pretoria 0002, South Africa
e-mail: jmeyer@up.ac.za

An experimental investigation was conducted into the heat transfer characteristics during in-tube condensation of horizontal smooth, micro-fin, and herringbone tubes. The study focused on the heat transfer coefficients of refrigerants R-22, R-134a, and R-407C inside a series of typical horizontal smooth, micro-fin, and herringbone tubes at a representative average saturation temperature of $40\,^\circ$C. Mass fluxes ranged from 300 to 800 kg/m^2 s, and vapor qualities ranged from 0.85 to 0.95 at condenser inlet, to 0.05 to 0.15 at condenser outlet. The herringbone tube results were compared with the smooth and micro-fin tube results. The average increase in the heat transfer coefficient of the herringbone tube, when compared with the smooth tube at comparable conditions, was found to be 322%, with maximum values reaching 336%. When compared with the micro-fin tube, the average increase in heat transfer coefficient was found to be 196%, with maximum values reaching 215%. Moreover, a new correlation was developed to predict the heat transfer coefficients in a herringbone and micro-fin tube. Semi-local heat transfer coefficients were calculated from the modified Wilson plot technique, using measurements of condenser subsection inlets and outlets, and from knowledge of the temperature gradient on the annulus side. The correlation predicted the semi-local heat transfer coefficients accurately, with 96% and 89% of the data points falling in the ±20% region for the herringbone tube and the micro-fin tube, respectively. The average heat transfer coefficients were accurately predicted, too, with all the data points for the herringbone tube and 83% of the data points for the micro-fin tube falling in the ±20% region. The derived heat transfer correlations can be used for design, especially for reversible heat pumps. This research proves that predicting the flow pattern during intermittent and annular flow is not a prerequisite for predicting the heat transfer accurately to within 20% of the measurements. [DOI: 10.1115/1.2194038]

Keywords: herringbone tube, micro-fin tube, condensation, heat transfer coefficient correlation, heat pumps

Introduction

In-tube condensation heat exchangers are of great importance in refrigeration and air-conditioning applications. In-tube condensation is used in air-cooled and evaporative condensers, and also with water-cooled condensers in reversible heat pumps and tube-in-tube condensers. In-tube condensation is especially important in large equipment operating at high pressures, typical of shell-and-tube condensers. The demand for more compact systems, higher energy efficiency, lower material costs, and other economic incentives has led to the rapid development of better and more effective heat exchangers [1]. These heat exchangers have enhanced surfaces, displaced enhancement devices, swirl-flow devices, or surface tension devices to enhance the heat transfer coefficient [2]. This implies that intensive research is currently being conducted on enhanced heat exchangers. Furthermore, the amendments to the Montreal Protocol [3] that focuses on the phase out of all chlorinated compounds, have ensured that substitutions for traditional refrigerants such as R-22 have to be found. The use of zeotropic replacement refrigerants such as R-407C further leads to heat transfer degradation during condensation, which must be accounted for. Widely used semi-empirical correlations to predict

heat transfer during condensation have been shown to be inaccurate in certain applications, and this has resulted in renewed efforts to characterize flow conditions and associated predictive procedures for heat transfer and pressure drop of condensing vapors.

It has become common practice to enhance condensation heat transfer inside horizontal tubular heat exchangers for air conditioners by employing micro-fin tubes (cf. Fig. 1(a)). Micro-fin tubes are usually made of copper and have outside diameters that range from 4 to 15 mm. They feature internal grooves (or fins), a set of 50–70 fins that are spiraled at helix angles of 6–30 deg (but usually 18 deg), with fin heights between 0.1 and 0.25 mm. These fins are usually trapezoidal in shape with apex angles of around 30 deg. Micro-fin tubes yield a heat transfer enhancement of 60–200% and a pressure drop increase of approximately 50% when compared to conventional smooth tubes [1,4]. Extensive research work has been conducted on micro-fin tubes since their inception in the 1970s, but very little experimental data are available on a recent heat transfer enhancement development, namely herringbone tubes.

Herringbone tubes (cf. Fig. 1(b)) yield an increase in heat transfer coefficient in excess of 200% under certain conditions, and a pressure drop increase of between 50% and 60% when compared to micro-fin tubes during condensation [5,6]. According to Goto et al. [7], herringbone tubes increase the heat transfer coefficient twice as much as when compared to micro-fin tubes. Miyara et al. [8] found that the heat transfer coefficient of a herringbone tube
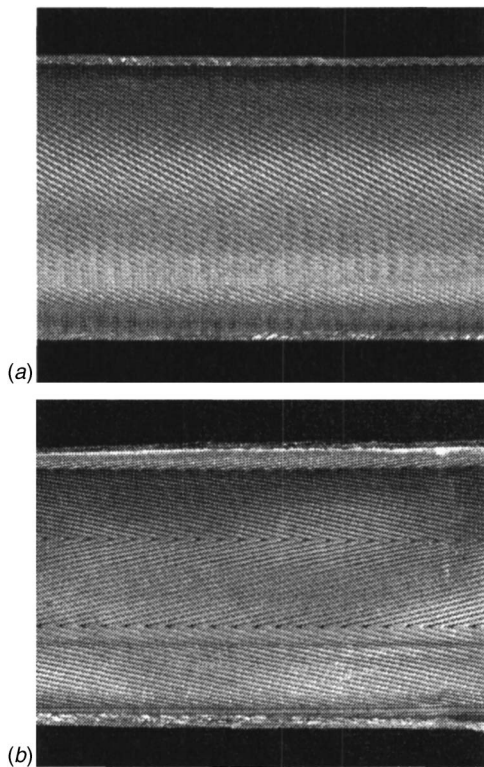
Fig. 1 Experimental (*a*) micro-fin tube and (*b*) herringbone tube

was higher than that of a micro-fin tube in the high mass flux region by as much as a factor of 2, while it had slightly lower values in the lower mass flux regions. Also, Miyara et al. proposed that the double-chevron arranged microfins of the herringbone tubes remove liquid at the fin-diverging parts and collect liquid at the fin-converging parts. Herringbone tubes have similar fin heights than their helical micro-fin counterparts, although herringbone fins are typically spiraled at 8–10 deg.

Miyara et al. [8] furthermore present the only known correlation for predicting the pressure drop and heat transfer coefficients for refrigerant condensation in herringbone tube heat exchangers. However, the correlations make use of inside wall temperatures and are sometimes difficult to implement in practice. A further difficulty with previous experimental work is that flow regime maps along with visual observation of the flow regimes were not always incorporated to validate and explain the experimental work. The refrigerants used in experimental work on herringbone tube heat exchangers were R-407C, R-410A, and R-22.

The aim of this study was to investigate the heat transfer coefficients of horizontal herringbone tube and micro-fin tube heat exchangers during condensation with refrigerants R-22, R-134a, and R-407C by using flow pattern maps and visual observations of flow regimes, and to develop a correlation for the prediction of heat transfer coefficients.

Experimental Setup

A test facility was constructed to measure in-tube condensation heat transfer and pressure drop of pure refrigerants and zeotropic mixtures [4,9], and is shown in Fig. 2. It was a vapor compression heat pump water heater comprising three flow loops: a refrigerant loop, a heating water loop (for the evaporator), and a cooling water loop (for the condenser).

The refrigerant loop consisted of four main components: the compressor, the condenser, the expansion valve, and the evaporator. The compressor was a hermetically sealed, reciprocating type

with a nominal cooling capacity of 10 kW. An oil separator was connected to the compressor in parallel with a bypass line. By manually controlling the flow through the bypass line and the oil separator simultaneously, the oil mass fraction in the refrigerant could be controlled. The water-cooled-type tube-in-tube condenser consisted of two parts: the test condenser and the after condenser. The test condenser consisted of eight separate coaxial tube test sections connected in series, which were labeled A, B, C, to H as shown in Fig. 2. This ensured that different flow regimes could be monitored in each subsection. The inner and outer tubes were separated by means of thin copper spacers to prevent sagging, which assisted in promoting turbulence in the annulus. The outer tube of each test section was a hard-drawn copper tube with an inside diameter of 17.27 mm and an outside diameter of 19.05 mm. Three types of test sections with inner tubes were manufactured, namely smooth, micro-fin, and herringbone types, all made of hard-drawn copper. The tube dimensions are given in Table 1. It will be noted that tubes with similar inside diameters were chosen to enable an effective comparison of data. High-pressure sight glasses were installed between the test sections in order to visually monitor the refrigerant flow pattern. The inner diameters of the sight glasses were identical to the inner root diameters of the test sections. All test sections were well insulated to minimize heat loss.

The cooling water cycle consisted of a centrifugal pump circulating the water from a 1000 L insulated cold-water storage tank through the system at the condensing side. A 15 kW chiller was connected to the storage tank in order to control the cooling water temperature (between 20 and 25°C). The heating water cycle was composed of a centrifugal pump that circulated the water from a 1000 L insulated hot water storage tank through the system at the evaporating side. A 20 kW electrical resistance heater was connected to the storage tank in order to constantly control the heating water temperature (between 30 and 40°C).

A bypass line connected in parallel with the test condenser was used to control the refrigerant mass flow rate through the test condenser. The after condenser was a coiled water-cooled, tube-in-tube condenser, used specifically to ensure that only liquid refrigerant entered the Coriolis mass flow meter.

A second Coriolis mass flow meter was placed at the inlet of the test condenser on the water side to measure the cooling water flow rate. A suction accumulator was placed upstream of the compressor to retain the liquid refrigerant and ensure that only vaporous refrigerant would flow to the compressor.

Temperatures were measured with Pt-100-type resistance temperature detectors (RTDs) mounted longitudinally at each condenser sub-section inlet on the outside wall of the tubes (refrigerant and cooling water loops). At each measurement location, two Pt-100s were used, one measuring the temperature at the top and another at the bottom of the outside tube walls. The average temperature between top and bottom values was used for data analysis. For the herringbone tube, one RTD was placed on the top and one on the side of the tube in order to measure the average temperatures over the tube cross section. This was done since Cavallini et al. [5] showed that the heat transfer coefficients at the sides of the herringbone tube were typically 10% to 15% higher than at the top and bottom of the tube.

High-pressure sight glasses and six mini-video cameras were also mounted in close proximity to the temperature and pressure sensors in order to observe and capture flow patterns. The sight glasses were mounted just before the tube U bends, so that there was at least a 1 m spacing between sight glasses, to negate the flow effects that U bends induce due to generation of non-negligible centrifugal forces. (Tube bends typically affect the local flow conditions, but flow typically returns to its original (prebend) characteristic approximately ten tube diameters downstream.)

Signals from the temperature sensors, pressure sensors, Coriolis flow meters, and the six mini-video cameras were recorded by a
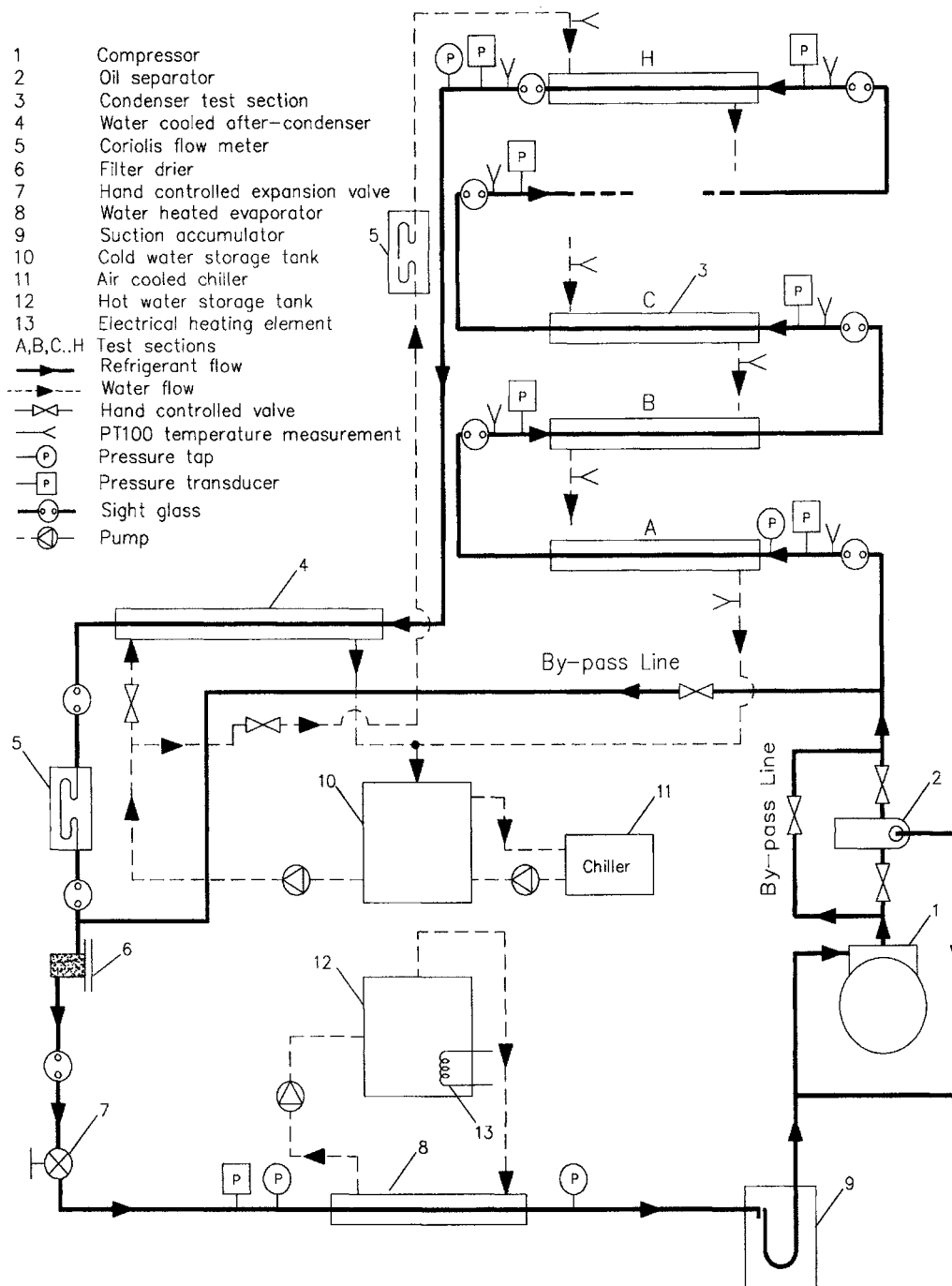
Fig. 2   Experimental setup

computerized data acquisition system, that was programmed to evaluate the data and to calculate energy balances, average saturation temperature, heat transfer coefficients, and pressure readings, all of which were displayed in real time.

Measurements for R-22, R-134a, and R-407C were conducted at a nominal saturation temperature of 40°C. The saturation temperature for R-407C was taken as the dew-point temperature. Refrigerant properties were obtained from a refrigerant database [10]. The experimental uncertainties were calculated by a propagation-of-error method and are reported in Table 2.

Experimental Procedure

Data measurements were taken upon obtaining an energy balance of better than 1%, for a minimum period of 10 min.

Table 1   Geometric parameters of experimental inner tubes

Tube no.	1	2	3
Type	Smooth	Micro-fin	Herringbone
Helix angle, β (°)	...	18	16
Apex angle, γ (deg)	...	55	25
Number of fins, n (−)	...	60	70
Outside diameter, d_o (mm)	9.52	9.55	9.51
Inside diameter, d_i (mm)	8.11	8.936	8.52
Tube wall thickness, t (mm)	1.3	0.302–0.312	0.3
Fin height, e (mm)	...	0.198–0.2197	0.2
Sub-sectional length, l (m)	1.5	0.9	0.563
Actual flow area, A_{fa} (mm²)	60.9	60.64	61.87

Measurements	Low mass flux (400 kg/m² s)	High mass flux (800 kg/m² s)
Refrigerant temperature	0.14 K	0.04 K
Water temperature	0.11 K	0.04 K
Saturation temperature	0.12 K	0.02 K
Pressure	0.28%	0.18%
Refrigerant mass flow rate	0.31%	0.15%
Water mass flow rate	0.15%	0.40%
Average quality	4.58%	1.46%
Heat transfer coefficient	14.70%	7.88%
Viscosity	0.10%	0.10%
Density	0.03%	0.04%
Reynolds number	1.09%	0.95%
Differential pressure	7.69%	5.91%

Annulus-Side Heat Transfer Coefficient (for Single-Phase Flow). A modified Wilson Plot technique [4] was used to determine the unknown Reynolds number exponent and the annulus-side coefficient for the heat transfer coefficient expressed by a Dittus and Boelter-type equation

$$\alpha_o = 0.0936 \frac{\lambda_l}{d_i} \operatorname{Re}_{lo}^{0.0742} \operatorname{Pr}_l^{1/3} (\mu_l/\mu_{l,w})^{0.14} \qquad (1)$$

The cold water-side (annulus) flow rate was varied as follows: 0.34, 0.27, 0.22, 0.16, 0.11, and 0.059 kg/s. For each of these settings, the hot water-side (inner tube) flow rate was varied as follows: 0.11, 0.08, 0.06, and 0.036 kg/s, thus delivering 24 overall settings, of which an average of ten readings per point were taken. This delivered 240 data points for each test condenser setup. The flow rates corresponded to the maximum and minimum flow rates possible, prescribed by the physical limitations of the tube length and diameters used. For the *tube-side Wilson plot*, the annulus-side Reynolds number was kept constant, and the tube-side water flow rate varied, with $(10,000 < \operatorname{Re}_i < 40,000)$.

Deduction of Average Heat Transfer Coefficient. With the heat transfer coefficient in the annulus known, the average refrigerant-side heat transfer coefficient could be calculated from [4]

$$\alpha_i = \frac{1}{A_i} \left[\frac{1}{UA} - \frac{1}{\alpha_o A_o} - \frac{\ln(d_o/d_i)}{2\pi\lambda_{Cu} l} \right]^{-1} \qquad (2)$$

where α_o is the water-side heat transfer coefficient determined from the modified Wilson plot, and the overall heat transfer coefficient U (based on the outside area of the inner tube) was determined by

$$U = \frac{Q_{ave}}{A_o \text{LMTD}} \qquad (3)$$

where LMTD is the logarithmic mean temperature difference, determined by

$$\text{LMTD} = \frac{(t_{sat} - T_{out}) - (t_{sat} - T_{in})}{\ln\left(\dfrac{t_{sat} - T_{out}}{t_{sat} - T_{in}}\right)} \qquad (4)$$

The inner diameters (d_i) for the enhanced tubes are taken as the fin tip-to-fin tip diameters.

Although the thermal resistance term (last term in Eq. (2)) was taken into consideration with calculations, it was found to be negligibly small.

For condensation inside the micro-fin tube, the heat transfer coefficient in the inner tube (Eq. (2)) was based on the inside area of a smooth tube that has an inner diameter equal to the maximum diameter of the micro-fin tube (nominal heat transfer area).

For pure refrigerants, the saturation temperature was obtained from the average of the refrigerant bulk temperature at the inlet and exit of the condensation test section. For the zeotropic mixture (R-407C), the saturation temperature was deduced by using the following correlation of Cavallini et al. [5]

$$t_{sat} = t_{dew} - \Delta t_{glide}(1 - x) \qquad (5)$$

where Δt_{glide} is the temperature glide, while x is an empirical function of the vapor quality that varies from 0 to 1.

Deduction of Semi-Local Heat Transfer Coefficient. The semi-local heat transfer coefficient for the refrigerant side was determined from [4]

$$\alpha_i = \frac{1}{\dfrac{(t_r - T_w)\pi d_i}{\dot{m}c_{p,w}\dfrac{dT_w}{dz}} - \dfrac{d_i}{\alpha_o A_o} - \dfrac{d_i \ln(d_o/d_i)}{2\lambda_{Cu}}} \qquad (6)$$

The first term in the denominator is the overall thermal conductance, the second term is the annulus thermal conductance, and the third term is the tube wall conductance. The water temperature in the annulus was fitted to a two-degree polynomial as a function of axial position. The local variation in heat flux along the length of the tube could then be related to the derivative of the temperature profile by using an energy balance.

The water temperature distribution in the annulus was expressed as a function of the axial distance z along the test section and fit to a second-order polynomial. The axial temperature variation dT_w/dz was then determined from the derivative of the second-order polynomial and evaluated at the water temperature at the respective measurement stations.

Deduction of Vapor Quality. The average sectional vapor quality was calculated from

$$x_i = \frac{h_i - h_l}{h_g - h_l} \quad \text{with } h_l \text{ and } h_g \text{ measured at } t_i$$

$$x_o = \frac{h_o - h_l}{h_g - h_l} \quad \text{with } h_l \text{ and } h_g \text{ measured at } t_o \qquad (7)$$

The average vapor quality of each test subsection was then calculated as $x = (x_i + x_o)/2$. Although there were eight condenser subsections, two-phase flow could only be obtained in four or five of these, with the other subsections comprising of either sub-cooled or desuperheated flow. Although a small change in vapor quality in the test subsections would have been preferred, the authors also attempted to simulate realistic condenser operation by using realistically long tube lengths, which introduced a 16–24% change in average vapor quality in the two-phase subsections.

Correction Factor of Zeotropes. The operation of zeotropic refrigerant mixtures in refrigeration systems produces heat transfer degradation during condensation, due to mass transfer resistance between the liquid and vapor phases. The slip velocity between the two phases, brought on by an increase in the liquid phase saturation temperature, causes nonequilibrium between the liquid and vapor phases. The correction factor of Bell and Ghaly [11] was used to compensate for the zeotropic characteristics of R-407C.

Experimental Results

Flow Regime Observations. Liebenberg et al. [4] and Olivier et al. [12] adapted a flow regime map generated by Thome [13] for condensation in the helical micro-fin and herringbone tubes (cf. Fig. 3). The flow regimes that were observed in the herringbone tube condenser validated the flow regime map predictions. The observed flow patterns were annular, slug, and plug (i.e.,
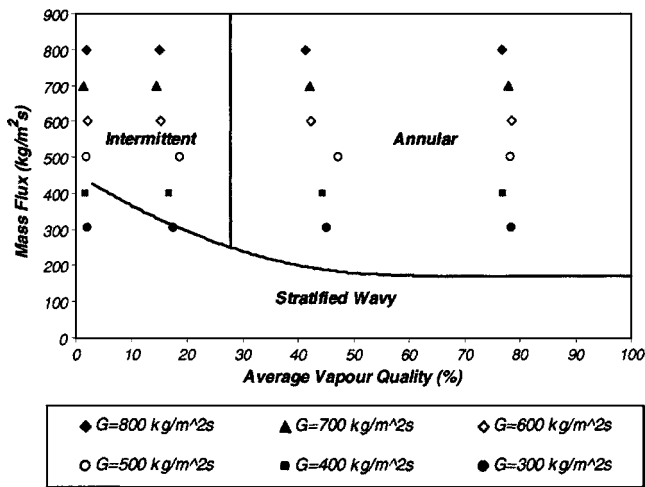
Fig. 3 Experimental data and new flow regime map for R-407C condensing in a herringbone tube

intermittent), and stratified flow (the latter only for R-407C and R-134a at 300 kg/m² s). In the micro-fin tube condenser, annular flow, slug and plug flow, and stratified flow (the latter for R-407C at 300 kg/m² s) were observed, while for the herringbone tube annular flow, slug flow, plug flow and stratified wavy flow were observed (cf. Fig. 4). During slug flow there are large waves that bridge the full cross section of the channel, and as condensation continues, the slugs coalesce into a predominantly liquid flow with large bubbles. Slug and plug flow occur towards the end of the condensation process and are termed intermittent flow for the purpose of this paper.

Since the flow regimes were similar in all three types of condensers, only a representative flow regime map is presented in Fig. 3 (i.e., for R-407C condensing in a herringbone tube), and the corresponding visual observations are shown in Fig. 4. (Similar trends for condensing R-134a and R-22 were observed and are

therefore not presented here. They are, however, available in Liebenberg et al. [4].) The adapted Thome flow regime map (Fig. 3) evidently predicted the observed flow regimes accurately. At the highest vapor qualities (i.e., $x \simeq 0.9$), the flow in the herringbone tube is annular with a stable and very thin annular condensate film, with the gas phase flowing in the central core. As expected, the flow has a twisting (or swirling) nature, induced by the helix angle of the microfins. Heat transfer would therefore be controlled by vapor shear and turbulence in this regime. It is further expected that the herringbone microfins will be very effective at mixing the liquid-vapor interface due to their proximity to this interface. Also, surface tension drainage is expected to have a large effect at these high vapor qualities.

As the condensation process progresses, liquid accumulates at the bottom of the horizontal condenser tube. At vapor qualities around 40%, all the liquid does not accumulate in a pool as it does with a smooth tube [4]. Rather, the herringbone microfins ensure that some liquid is carried to the top due to the capillary action of the fins and the swirling energy of the flow. The double-chevron shape of the herringbone microfins, however, ensures that the liquid is again carried back to the bottom of the tube, so that a thin liquid film is evident on the tube sides. The liquid pool at the bottom of the tube is thicker than that at the top, due to the action of gravity. The effective flow pattern therefore changes from an intermittent flow regime to a semi-annular flow regime, thereby further enhancing the heat transfer. Elongated slugs form at the lower vapor qualities (cf. $x = 2-10\%$). Occasionally, waves are observed at around 5% vapor quality but only at the lowest mass fluxes. It appears that the slug frequency, slug length, wave amplitude, and the wave frequency are stochastic in nature. Eventually the refrigerant becomes a subcooled liquid and no occasional bubbles appear, as was the case with the smooth tube [4].

It is apparent that the microfins delay the transition from annular flow to intermittent flow (by approximately $x = 15-20\%$) due to the combined actions of (i) swirling, (ii) surface tension drainage, and (iii) heat transfer surface area enhancement, which enhances the heat transfer relative to the smooth-tube counterpart. In addition to the aforementioned three enhancement mechanisms,
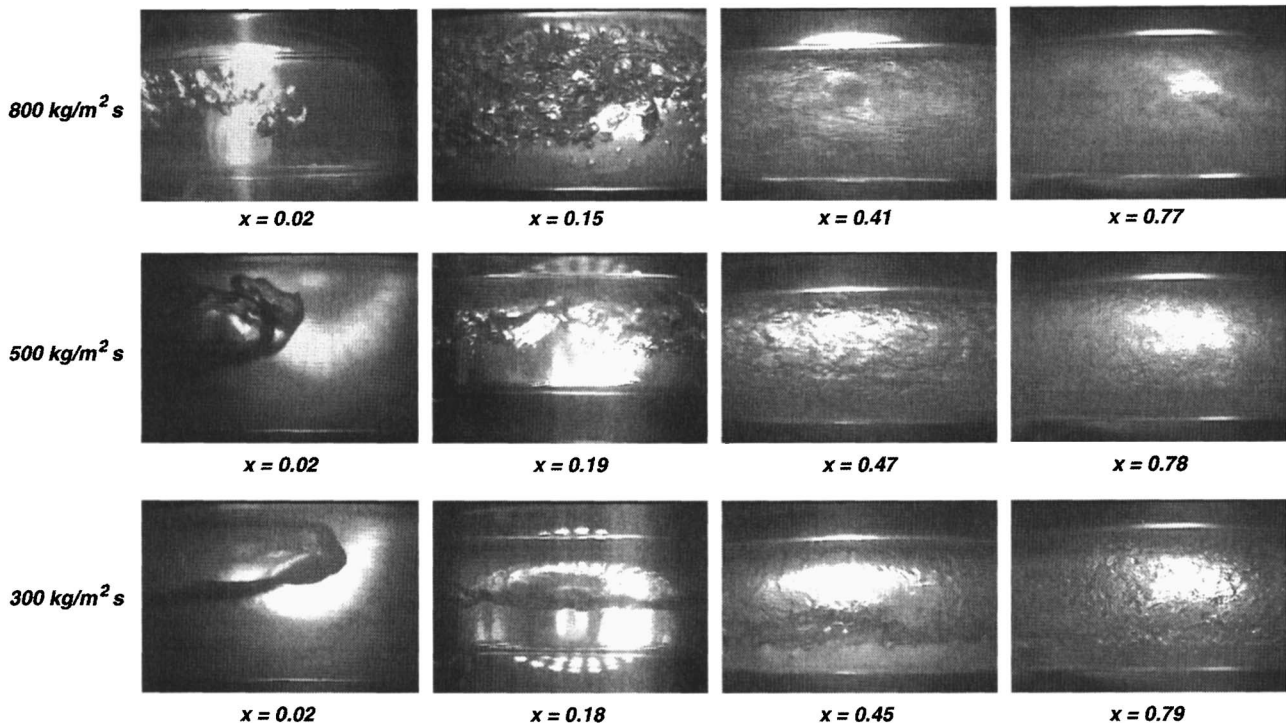

Fig. 4 Flow regime observations for R-407C condensing in a herringbone tube at 300 kg/m² s, 500 kg/m² s, and 800 kg/m² s
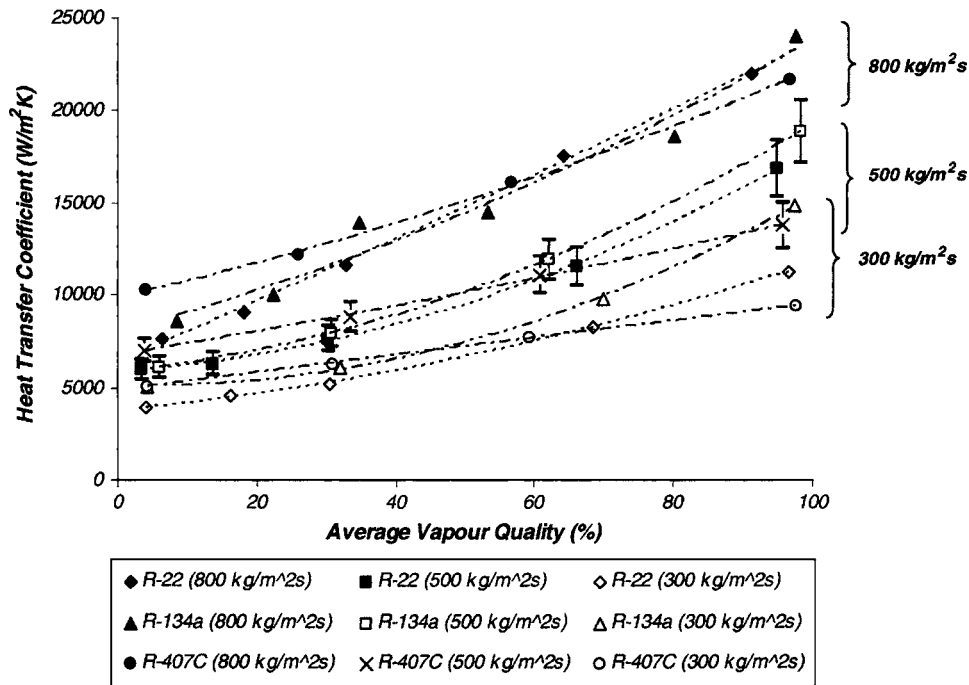
Fig. 5 Semi-local heat transfer coefficients for herringbone tube

herringbone tubes induce additional swirl (Fig. 1(*b*)), resulting in greater heat transfer enhancement than the comparative micro-fin tube. Thus, when compared with helical micro-fin tubes (Fig. 1(*a*)), an additional 10% to 20% delay in transition from annular to intermittent flow is ensured and implies the greatest heat transfer enhancement of the three tubes. Visual observations, furthermore, show that the intermittent flow regime is neither uniform nor steady.

Semi-Local Heat Transfer Coefficients. The semi-local heat transfer coefficients followed the same trend for all three condensers, with an increase in heat transfer coefficient as the vapor quality and mass flux increased. Figure 5 shows the semi-local heat transfer coefficients for the herringbone tube condenser for condensing R-22, R-134a, and R-407C. From Figs. 3–5 it follows that the semi-local heat transfer coefficients decrease as the flow regimes move from annular flow at high vapor qualities of $\pm 80\%$ to intermittent flow at lower vapor qualities of $\pm 20\%$, and to nearly total liquid phase at very low vapor qualities of $\pm 2\%$. Visual observations of the flow pattern (cf. Fig. 4) reveal the dominance of intermittent flow in the low-quality regime (i.e., for $x < 20\%$), with the flow becoming stratified wavy at very low vapor qualities (i.e., for $x < 2\%$). The main mechanism of heat transfer in this latter flow regime is conduction through the liquid film at the top of the tube. Since this liquid film is very thin, the presence of microfins reduces its thickness considerably, which in turn considerably reduces the film conduction resistance. Therefore, a large enhancement factor is expected under low mass flux conditions. Since wavy (stratified) flow prevails over the entire very-low-vapor-quality range, the enhancement factor remains relatively constant with varying vapor quality.

It can also be concluded from Fig. 5 that, in general, the heat transfer enhancement factors of R-134a are the highest, followed by those of R-407C and R-22, but that the differences are small. The better condensation performance of R-134a is ascribed mainly to its reduced vapor pressure and low vapor density (hence greater vapor-specific volume), compared to the higher-pressure refrigerants, R-22 and R-407C.

Average Heat Transfer Coefficients. The average heat transfer coefficients for the smooth tube were compared to the correlations of Cavallini and Zecchin [14], Shah [15], Dobson and Chato [16], and Tang [17]. The correlations of Shah [15] and Tang [17] fitted the smooth tube data best, with Shah [15] producing an average over prediction of 23% and a maximum over prediction of 26% for R-22. For R-134a, the average overprediction by Shah [15] was 25% and the maximum over prediction 29%. Tang [17] over predicted the average heat transfer coefficient of R-407C by an average of 21% and a maximum of 32%.

The micro-fin tube data were compared to the correlation of Cavallini et al. [5], who used micro-fin geometries similar to those in the present study. The average deviation for the R-22 data was found to be a 10% over prediction, with a maximum deviation of 18% over prediction. The average deviation for the R-134a data was a 9% under prediction (with a maximum of 17%), while for R-407C the average deviation was a 14% over prediction (maximum deviation of 38%). The Cavallini et al. correlation accurately predicts the R-22 and R-134a data, but was adapted so as to fit the R-407C data more accurately [18].

Figure 6 shows the average heat transfer coefficients for all three condenser tubes for R-22, R-134a, and R-407C. From this figure it follows that the average heat transfer coefficients increase with an increase in mass flux for all three tubes. The increase in the average heat transfer coefficients for the herringbone tube, compared to the smooth and micro-fin tube, is substantial.

When compared to a smooth tube, the herringbone tube increased the average heat transfer coefficients by an average of 284% for R-22, with a maximum of 296% at 300 kg/m^2 s. For R-134a, the average increase was 322% with a maximum of 336% at 300 kg/m^2 s and for R-407C, the increase was found to be 280% with a maximum of 297% at 800 kg/m^2 s.

When compared to a micro-fin tube, the maximum increase in heat transfer coefficient for R-22 was 198% at 800 kg/m^2 s, with an average increase of 172%. For R-134a, the average increase was 191% with a maximum increase of 209% at 800 kg/m^2 s and for R-407C, the average increase was 196% with a maximum of 215% at 800 kg/m^2 s. Overall, R-134a was the best performing refrigerant in the micro-fin and herringbone tubes, while R-407C performed best in the smooth tube.
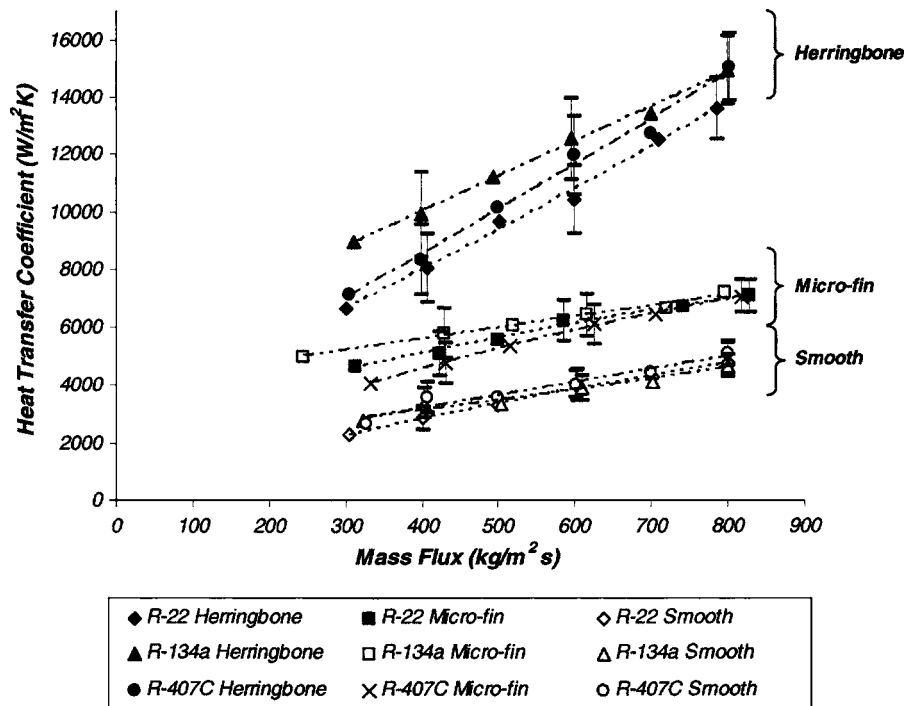
Fig. 6 Average heat transfer coefficients for smooth, micro-fin, and herringbone tube

Heat Transfer Enhancement Factors (EF) and Pressure Drop Penalty (PF) Factors. The area enhancement factor is defined here as the ratio of the surface area of the enhanced tube to that of an equivalent smooth tube. The *area* enhancement factors for the micro-fin and herringbone tube, compared to the smooth tube, were found to be 1.52 and 1.80, respectively. The heat transfer enhancement factor (EF) is defined here as the ratio of the heat transfer coefficient of the enhanced tube to the heat transfer coefficient of the smooth tube, with the same temperature difference between the wall and the refrigerant. When the herringbone tube's performance was compared to that of the smooth tube, the maximum heat transfer enhancement factors for R-22, R-134a, and R-407C were found to be 2.0, 2.17, and 1.89, respectively, with a maximum of 1.96 at 800 kg/m² s, 2.27 at 300 kg/m² s, and 2.0 at 800 kg/m² s, respectively. When compared to the micro-fin tube, the average heat transfer enhancement factors for R-22, R-134a, and R-407C were 1.27, 1.41, and 1.45, respectively, with the maximum values of 1.46, 1.55, and 1.59 at 800 kg/m² s for R-22, R-134a, and R-407C, respectively.

Olivier [12] calculated the pressure drop penalty factors (PF) for the current data as the ratio of the pressure drop in the herringbone tube to the pressure drop in the smooth tube, and the ratio of the pressure drop in the herringbone tube to the pressure drop in the helical micro-fin tube, with both tubes being of equal length. The overall performance factors were calculated by determining the ratio of the heat transfer enhancement factor to the pressure drop penalty factor (=EF/PF). The average overall performance factors for the herringbone tube compared to the smooth tube for R-22, R-134a, and R-407C were calculated as 1.10, 1.30, and 0.75, respectively, while the maximum performance factors were calculated as 1.30, 1.76, and 1.04, respectively. For the herringbone tube compared to the micro-fin tube, the average performance factors for R-407C were found to be 1.50, 1.74, and 1.18, respectively, with maximum values of 2.09, 3.26, and 1.97, respectively.

New Heat Transfer Correlation

The new correlation for predicting the heat transfer coefficients in a helical micro-fin or herringbone tube during in-tube conden-sation was developed by adapting the correlation of Cavallini et al. [19] for micro-fin tubes and by comparing the experimental results of the herringbone tube to the correlation. The correlation applies to both the annular and intermittent (slug and plug) flow regimes, due to the similar vapor shear-controlled heat transfer mechanisms. The correlation of Cavallini et al. [5] was adapted so that the average heat transfer coefficients at high mass fluxes would be equal or very close to those of the experimental data. Based on the visual observations, the main reason for the increase in heat transfer coefficients was the increase in turbulence. Thus, the correlation of Cavallini et al. [19] was adapted by initially changing the power of the equivalent-Reynolds number from 0.8 (i.e., $Re_{eq}^{0.8}$) to 0.828 (i.e., $Re_{eq}^{0.828}$).

However, the first adaptation caused the correlation to over pre-dict the average heat transfer coefficient at low mass fluxes by as much as 52% for R-22. Based on the variation of heat transfer coefficient with mass flux, and the dependence of Froude number on mass flux, the Froude number was varied. It was eventually found that the natural logarithmic function of the Froude number (namely ln (FR)) best correlated with the experimental data of the herringbone tube.

The correlation, however, over predicted the heat transfer coef-ficient over the entire mass flux range and the power of the equivalent-Reynolds number was consequently adapted from 0.828 ($Re_{eq}^{0.828}$) to 0.715 ($Re_{eq}^{0.715}$). The same process was carried out with the micro-fin tube data with the Reynolds number being adapted to 0.7 ($Re_{eq}^{0.7}$). The new correlation for predicting the heat transfer coefficients in a micro-fin or herringbone tube is defined as follows

$$\alpha = \frac{\lambda_l}{d_i} 0.05 \, Re_{eq}^s Pr_l^{1/3} TF \cdot (Rx)^2 [Bo \cdot \ln(Fr)]^{-0.26} \quad (8)$$

where Re_{eq} is the equivalent Reynolds number, defined by

$$Re_{eq} = 4\dot{m}[(1-x) + x(\rho_l/\rho_g)^{0.5}]/(\pi d_l \mu_l) \quad (9)$$

Pr_l is the liquid Prandtl number and is defined by $Pr_l = (\mu_l c_{p,l})/\lambda_l$, while $\dot{m}$ is the refrigerant mass flow rate. Equation (9) uses the following Reynolds number exponents

Table 3 Recommended ranges for new correlation

Parameter	Unit	Recommended range
Re_{eq}	...	$>15{,}000$
Pr_l	...	$3<Pr_L<6.5$
β	Degrees	$7<\beta<30$
G	kg/m^2 s	$300<G<800$

$$s = 0.715 \text{ for herringbone tubes}$$

$$s = 0.7 \text{ for micro-fin tubes} \tag{10}$$

TF is the trend factor and is given by

$$TF = (1-x) + x^{1.3} \tag{11}$$

while Rx accounts for the increase in heat transfer surface area compared to a smooth tube, and is given by [5]

$$Rx = \frac{\{[2en[1-\sin(\gamma/2)]]/[\pi d_i\cos(\gamma/2)]+1\}}{\cos\beta} \tag{12}$$

The Bond number Bo accounts for surface drainage effects, and is given by

$$\text{Bo} = \frac{g\rho_l e\pi d_i}{8\sigma n} \tag{13}$$

and the Froude number Fr is given by

$$\text{Fr} = \frac{u_{go}^2}{g d_i} \tag{14}$$

In the above, u_{go} is the velocity of the vapor phase only and is calculated by dividing the total mass flux by the vapor density. The parameter Rx is intended to account for the effect of the heat transfer area increase. The same parameter can also be used for herringbone tubes by considering the geometrical characteristics of the first set of grooves, as the additional set of grooves does not improve the heat transfer area significantly [19]. The Bond number accounts for surface tension effects, while the product of the Bond and Froude numbers directly represents the relative comparison between vapor shear and surface tension effects. The recommended ranges for the correlation are summarized in Table 3.

Comparison of New Correlation With Experimental Data.
For the herringbone tube correlation, 96% of all the semi-local heat transfer coefficient data points fall in the ±20% region. The new correlation predicts the semi-local heat transfer coefficients accurately, with an average error of 6% over prediction and a standard deviation of 10% for R-22. For R-134a the correlation under predicts the data with an average error of 4% and a standard deviation of 8%. The correlation predicts R-407C data with an average of error of 0% and a standard deviation of 6%.

For the micro-fin tube correlation, 89% of all the data points fall in the ±20% region. The average error for R-22 is a 7% over prediction with a standard deviation of 8%. The correlation under predicts the R-134a data with an average error of 1% and a standard deviation of 11%. For R-407C, the average error is 8% and the standard deviation 7%.

Regarding average heat transfer coefficients for the herringbone tube, all the data points fall in the ±20% region (cf. Fig. 7). The new correlation over predicts the R-22 data by an average error of 3% and a maximum error of 8%, while the standard deviation is 5%. For the R-134a data, the correlation under predicts the data with an average error of 3% and a maximum error of 9%, and with a standard deviation of 6%. The R-407C data is under predicted with an average error of 1% and a maximum error of 9%, while the standard deviation is 6%.

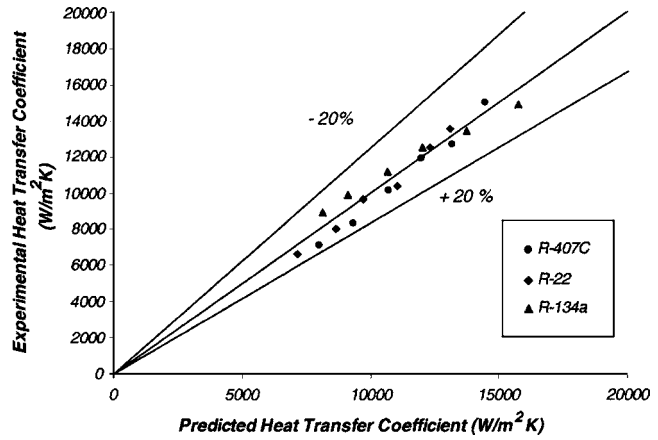The new correlation predicts all the average heat transfer coef-

Fig. 7 Experimental average heat transfer coefficients versus predicted average heat transfer coefficients for herringbone tube

ficients accurately for all three refrigerants with all the data points in a ±20% region. Figure 7 shows the experimental average heat transfer coefficients versus the predicted average heat transfer coefficients for the herringbone tube. Figures 8(a)–8(c), in turn, show the experimental and predicted average heat transfer coefficients.

For the micro-fin tube correlation, 83% of the data points fall in the ±20% region. The R-22 data are predicted accurately with an average over-prediction error of 12% and a maximum over-prediction error of 17%. The standard deviation is 6%. For R-134a, the correlation under predicts the data with an average error of 4% and a maximum error of 15%, while the standard deviation is 12%. The correlation over predicts the R-407C data with an average error of 16% and a maximum error of 19%, while the standard deviation is 3%.

The new micro-fin tube correlation thus tends to over predict the average heat transfer coefficients and is not as accurate as the herringbone tube correlation. However, the correlation still represents a slight improvement on the Cavallini et al. correlation [5] when comparing the micro-fin tube data, in the sense that the general trends are predicted more accurately for all three refrigerants. Detailed analyses can be found in Lambrechts [18].

Conclusions

An experimental investigation was conducted into the heat transfer characteristics of horizontal smooth, helical micro-fin, and herringbone tubes during in-tube condensation. The study focused on the heat transfer coefficients of refrigerants R-22, R-134a, and R-407C inside the three tubes at a nominal saturation temperature of 40 °C with mass fluxes ranging from 300 to 800 kg/m^2 s, and with vapor qualities ranging from 0.85 to 0.95 at condenser inlet to 0.05 to 0.15 at condenser outlet. The herringbone tube results were compared to the smooth and micro-fin tube results. The average increase in the heat transfer coefficient of the herringbone tube when compared to the smooth tube was found to be 322%, with maximum values reaching 336%. When compared to the micro-fin tube, the average increase in heat transfer coefficient was found to be 196% with maximum values reaching 215%.

In general, heat transfer coefficients increased with an increase in mass flux. The heat transfer enhancement, however, decreased as the mass flux was increased for both micro-fin tubes and herringbone tubes, producing the lowest heat transfer augmentation at the highest mass flux. For both herringbone and micro-fin tubes, at lower mass fluxes (ranging between 300 and 500 kg/m^2 s), a substantial heat transfer augmentation was observed for the vapor quality ranging between 10% and 30%, which corresponds to the
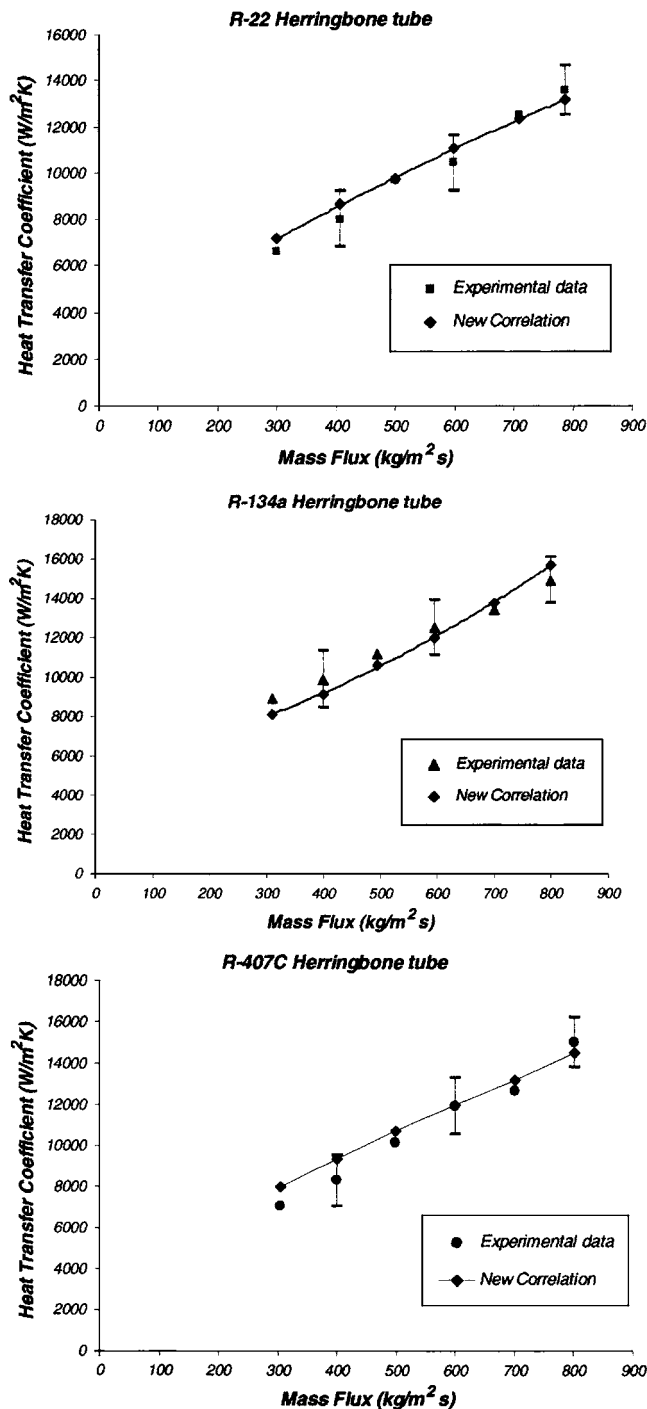
Fig. 8 Average heat transfer coefficients for condensing (*a*) R-22 in herringbone tube, (*b*) R-134a in herringbone tube, and (*c*) R-407C in herringbone tube

intermittent flow region. However, a moderate heat transfer augmentation was observed for vapor quality ranging between 40% and 90%, which corresponds to the transitional and annular flow region. For the herringbone and micro-fin tubes, at higher mass fluxes (ranging between 600 and 800 kg/m² s), a different trend was observed where moderate heat transfer augmentation occurred over the entire range of vapor quality.

A semi-empirical heat transfer coefficient correlation for annular and intermittent flow was generated for refrigerant condensation inside herringbone and micro-fin tubes, based on a widely used correlation. The new correlation predicted the semi-local

heat transfer coefficients accurately with 96% and 89% of the data points falling in the ±20% region for the herringbone and micro-fin tube, respectively. The average heat transfer coefficients were also predicted accurately, with all the data points for the herringbone tube and 83% of the data points for the micro-fin tube falling in the ±20% region. The derived heat transfer correlations can be used for design, especially for reversible heat pumps. The work shows that it is not necessary to predict the flow pattern during intermittent and annular flow in order to predict the heat transfer to within 20% of the measurements.

Acknowledgment

Nomenclature

A = area, m²
c_p = specific heat, J/kg K
d = diameter, m
e = fin height, m
g = gravitational acceleration, m/s²
G = mass flux, kg/m² s
h = specific enthalpy, J/kg
k = thermal conductivity, W/m K
l = tube length, m
LMTD = log-mean temperature difference, °C (K)
$\dot{m}$ = refrigerant mass flow rate, kg/s
n = number of fins
Pr = prandtl number,
Q = heat transfer, W
s = Reynolds number exponent
T = water temperature, °C (K)
t = refrigerant temperature/tube wall thickness, °C (K), m
TF = trend factor
U = overall heat transfer coefficient, W/m² K
u = velocity, m/s
x = vapor quality
z = axial tube direction

Greek symbols

α = heat transfer coefficient, W/m² K
β = helix angle, deg
γ = apex angle, deg
λ = thermal conductivity, W/m K
μ = dynamic viscosity, Ns/m²
ρ = density, kg/m³
σ = surface tension, N/m

Dimensionless Groups

Bo = bond number
Fr = Froude number
Pr = Prandtl number
Re = Reynolds number
Rx = dimensionless parameter
TF = trend factor

Subscripts

ave = average
Cu = copper
dew = dewpoint
eq = equivalent
fa = flow area

g = vapor
glide = temperature glide (difference between dew point and bubble point)
go = vapor only
h = hydraulic
i = inside
in = inlet
lo = liquid only
l = liquid
o = outside
out = outlet
sat = saturation
r = refrigerant
w = wall, water
w,o = outer wall

References

[1] Liebenberg, L., Bergles, A. E., and Meyer, J. P., 2000, "A Review of Refrigerant Condensation in Horizontal Micro-fin Tubes," *International Mechanical Engineering Congress*, Orlando, FL, 40, pp. 155–168.

[2] Bergles, A. E., 1988, "Some Perspectives on Enhanced Heat Transfer-Second-Generation Heat Transfer Technology," Trans. ASME, Ser. C: J. Heat Transfer, **110**, pp. 1082–1096.

[3] UNEP, 2000, "The Montreal Protocol on Substances that Deplete the Ozone Layer," United Nations Environmental Program, Nairobi, Kenya.

[4] Liebenberg, L., Thome, J. R., and Meyer, J. P., 2005, "Flow Visualization and Flow Pattern Identification With Power Spectral Density Distributions of Pressure Traces During Refrigerant Condensation in Smooth and Micro-fin Tubes," ASME J. Heat Transfer, **126**(3), pp. 209–220.

[5] Cavallini, A., Del Col, D., Doretti, L., Longo, G. A., and Rossetto, L., 2000, "Heat Transfer and Pressure Drop During Condensation of Refrigerants Inside Horizontal Enhanced Tubes," Int. J. Refrig., **23**, pp. 4–25.

[6] Ebisu, E., and Torikoshi, K., 1998, "Experimental Study on Evaporation and Condensation Heat Transfer Enhancement for R-407C Using Herringbone Heat Transfer Tube," ASHRAE Trans., **104**(2), pp. 1044–1052.

[7] Goto, M., Inoue, N., and Ishiwatari, N., 2000, "Condensation and Evaporation Heat Transfer of R410A Inside Internally Grooved Horizontal Tubes," Int. J. Refrig., **24**, pp. 628–638.

[8] Miyara, A., Nonaka, K., and Taniguchi, M., 2000, "Condensation Heat Transfer and Flow Pattern Inside a Herringbone-Type Micro-Fin Tube," Int. J. Refrig., **23**, pp. 141–152.

[9] Liebenberg, L., 2002, A Unified Prediction Method for Smooth and Micro-fin Tube Condensation Performance, Vol. 1, Ph.D thesis, October, Rand Afrikaans University (now University of Johannesburg), South Africa.

[10] REFPROP, 1998, NIST Standard Reference Database 23: NIST Thermodynamic Properties of Refrigerants and Refrigerant Mixtures (REFPROP). Version 6.01, National Institute of Standards and Technology (NIST), Gaithersburg, MD.

[11] Bell, K. J., and Ghaly, M. A., 1973, "An Approximate Generalized Design Method for Multicomponent/Partical Condenser," Am. Inst. Chem. Eng. Symp. Ser., **69**, pp. 72–79.

[12] Olivier, J. A., Liebenberg, L., Kedzierski, M., and Meyer, J. P., 2004, "Pressure Drop During Refrigerant Condensation Inside Horizontal Smooth, Helical Microfin, and Herringbone Microfin Tubes," ASME J. Heat Transfer, **126**, pp. 687–696.

[13] Thome, J. R., 2003, "Update on the Kattan-Thome-Favrat Flow Boiling Model and Flow Pattern Map," *Proceedings Fifth International Conference on Boiling Heat Transfer*, Jamaica, May 4–8.

[14] Cavallini, A., and Zecchin, R., 1974, "A Dimensional Correlation for Heat Transfer in Forced Convection Condensation," *Paper presented at the Fifth International Heat Transfer Conference*, Tokyo, 3, pp. 309–313.

[15] Shah, M. M., 1979, "A General Correlation for Heat Transfer During Film Condensation Inside Pipes," Int. J. Heat Mass Transfer, **22**, pp. 547–446.

[16] Dobson, M. K., and Chato, J. C., 1998, "Condensation in Smooth Horizontal Tubes," ASME J. Heat Transfer, **120**, pp. 193–213.

[17] Tang, L., 1999, "Empirical Study of New Refrigerant Flow Condensation Inside Horizontal Smooth and Micro-Fin Tubes," University of Maryland at College Park, Ph.D thesis.

[18] Lambrechts, A., 2003, "Empirical Study of Heat Transfer in Herringbone and Micro-Fin Tubes During Refrigerant Condensation," M.Eng. dissertation, Faculty of Engineering, Rand Afrikaans University (now University of Johannesburg).

[19] Cavallini, A., Censi, G., Del Col, D., Doretti, L., Longo, G. A., Rosetto, L., and Zilio, C., 2003, "Condensation Inside and Outside Smooth and Enhanced Tubes—A Review of Recent Research," Int. J. Refrig., **26**, pp. 373–392.

Convective Heat Transfer in Turbulent Flow Near a Gap

D. Chang
Doctoral Student

S. Tavoularis
Professor
e-mail: tav@eng.uottawa.ca

Department of Mechanical Engineering,
University of Ottawa,
Ottawa, Ontario, K1N 6N5, Canada

Convective heat transfer in a rectangular duct containing a heated rod forming a narrow gap with a plane wall has been simulated by solving the unsteady Reynolds-averaged Navier-Stokes equations with a Reynolds stress model. Of particular interest is the role of quasi-periodic coherent structures in transporting fluid and heat across the gap region. It is shown that the local instantaneous velocity and temperature vary widely because of large-scale transport by coherent vortical structures forming in pairs on either side of the rod. [DOI: 10.1115/1.2194039]

1 Introduction

Forced convective heat transfer in complex channels with narrow gap regions between solid walls flanked by wider subchannels occurs commonly in thermal equipment, particularly in nuclear reactor rod bundles. The fluid in gap regions is more likely to overheat than fluid in the adjacent subchannels, especially if the gap size diminishes, due to vibrations or thermally induced strains. Such conditions may reduce the thermal efficiency and have the potential of causing accidents, if local overheating leads to dryout in a device designed for liquid-phase operation. It is well known that axial flow near narrow gaps, and the associated heat and mass transfer across the gap, are dominated by strong, large-scale, quasi-periodic, flow pulsations, which have been characterized as coherent structures [1–8]. Cross-gap motions enhance inter-subchannel mixing and heat transfer in the gap region. This is evidenced by the fact that the local heat transfer coefficient in the gap is insensitive to the gap size, except for extremely narrow gaps [9,10]. The most detailed experimental studies of coherent structures in a gap region are by Guellouz and Tavoularis [7,8], who studied isothermal flow in a rectangular channel containing a cylindrical rod. Using conditional sampling and phase averaging, they concluded that the coherent structures formed a street of pairs of counterrotating vortices with axes alternating on either side of the gap. Due to the strong time dependence of the large-scale structures in such channels, numerical approaches using the Reynolds averaged Navier-Stokes equations (RANS) and conventional turbulence models are unsuitable, and can be only brought to some agreement with experimental results by using empirical adjustments, as, for example, the use of experimentally determined, non-isotropic eddy viscosities [11]. In a recent numerical study simulating the Guellouz and Tavoularis experiments, Chang and Tavoularis [12], hereafter referred to as CT, reproduced most experimental observations by using the unsteady Reynolds averaged Navier-Stokes equations (URANS) approach and a Reynolds stress model. The present work is an extension of this study, in the same geometry but with the rod heated passively to introduce a thermal field. The main objective of the present work is to provide a thorough understanding of the influence of coherent structures on heat transfer characteristics in a simplified rod-bundle-like configuration, which may lead to improved modeling of such processes. It is hoped that the present results would

be of benefit to safety analyses of nuclear reactors and to the design of improved rod bundles and other heat exchangers.

2 Computational Conditions and Procedures

The present numerical studies utilize the commercial software package FLUENT (version 6.1.22), based on the finite volume method. The computational domain and the velocity computation procedures were identical as those in the CT simulations [12]. Briefly, the domain consisted of a rectangular channel containing a rod with a diameter D, as shown in Fig. 1. The height, width, and length of the computational section were, respectively, $2D$, $3D$, and $L=20D$. The gap between the rod and the adjacent plane walls was set at $\delta=0.10D$ and the equidistant plane was defined by $y=1/2\delta$. The domain was sufficiently long to permit complete spatial decorrelation of turbulence statistics available from the experiments [7]. The Reynolds number, based on the bulk velocity U_b and the hydraulic diameter $D_h=1.59D$, was 108,000, nearly matching the experimental value. The unsteady velocity field was computed by solving the unsteady Reynolds-averaged continuity and Navier-Stokes equations for incompressible flow, supplemented by modeled equations for the turbulent kinetic energy and its dissipation rate and the low Reynolds number version of the Reynolds stress model (RSM) [13]. To enhance convergence and avoid instability of the solution, the computations were initiated using the more stable Renormalization Group $k-\varepsilon$ model [14]. The near-wall flow was resolved by an increasingly refined mesh, rather than modeled using wall functions. The two-layer approach of Chen and Patel [15] was employed, in which the eddy viscosity in the viscous region was computed using the kinetic energy obtained from its transport equation and an algebraic mixing length. For high accuracy and stability, a second-order upwind scheme [16] was selected for deriving the face value of different variables for space discretization. The semi-implicit pressure linked equations - consistent algorithm [17] was applied for the treatment of pressure-velocity coupling. Time discretization was achieved by a second-order implicit method (second-order backward Euler scheme) [18]. Coherent structures in the flow were identified and captured using the Q criterion, which involves second invariants of the velocity gradient tensor. Details about the Q criterion have been presented by CT.

The instantaneous local flow temperature T was computed by solving the following Reynolds-averaged energy equation for incompressible, single-phase flow
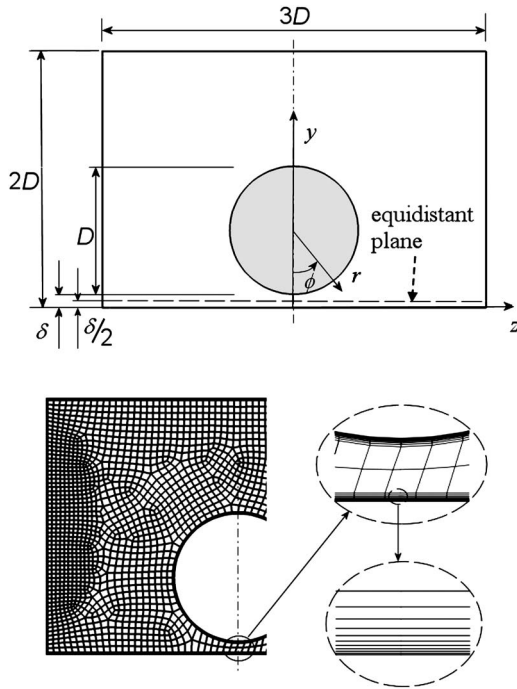
Fig. 1 Sketches of the computational geometry and mesh

$$\frac{\partial T}{\partial t} + \frac{\partial(UT)}{\partial x} + \frac{\partial(VT)}{\partial y} + \frac{\partial(WT)}{\partial z}$$

$$= \gamma\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) - \left(\frac{\partial(u\vartheta)_{nc}}{\partial x} + \frac{\partial(v\vartheta)_{nc}}{\partial y} + \frac{\partial(w\vartheta)_{nc}}{\partial z}\right) \tag{1}$$

In this equation, the Reynolds-averaged turbulent heat fluxes were modeled as [19]

$$-(u\vartheta)_{nc} = \frac{\mu_T}{\sigma_T}\frac{\partial T}{\partial x}, \quad -(v\vartheta)_{nc} = \frac{\mu_T}{\sigma_T}\frac{\partial T}{\partial y} \text{ and } -(w\vartheta)_{nc} = \frac{\mu_T}{\sigma_T}\frac{\partial T}{\partial z} \tag{2}$$

where the turbulent Prandtl number was taken as $\sigma_T = 0.85$. The resolved eddy viscosity μ_T was computed in terms of the resolved, non-coherent local turbulent kinetic energy k_{nc} and its dissipation rate ε_{nc} as

$$\mu_T = C_\mu \rho \frac{k_{nc}^2}{\epsilon_{nc}} \tag{3}$$

where $C_\mu = 0.09$.

The buoyancy force was negligible, as evidenced by the fact that the Richardson number, which represents the ratio of buoyancy and inertia forces, was less than 0.0003 [20]. The physical and thermodynamic properties of the fluid, taken to be air, were assumed to be constant and equal to those at 300 K, which was the temperature in the reference isothermal case.

Two heating modes were considered: one in which the rod surface temperature was kept uniform and equal to 305 K (mode A) and a second one in which the heat flux $\dot{q}$ from the rod surface was kept uniform and equal to 100 W/m^2 (mode B). Mode A is relevant to conventional heat exchangers, while mode B applies to nuclear reactors. In both cases, the heat flux on all plane walls was set to zero, corresponding to a thermally insulated duct.

To allow full development of the flow properties in a relatively short domain, the periodic boundary condition was applied to the flow. Heat was introduced at sufficiently small amounts for the temperature to be essentially a passive scalar and not to affect in

any measurable way the fluid properties or the velocity and pressure fields. This was a requirement for using periodic boundary conditions for the temperature [19]. In this approach, the velocity distribution on the exit plane was reintroduced on the inlet plane and the uniform pressure gradient was determined iteratively such as to correspond to a mass flow rate $\dot{m}$ that matched the specified value.

The time dependent resolved temperature $T(x,y,z,t)$ is presented in dimensionless form as

$$\Theta(x,y,z,t) = \frac{T(x,y,z,t) - \bar{T}_b(x)}{\bar{T}_{\text{rod},m}(x) - \bar{T}_b(x)} \tag{4}$$

where $\bar{T}_b(x)$ is the time average of the cross-section-averaged (bulk) temperature $T_b(x,t)$, defined as (A is the cross-sectional flow area)

$$T_b(x,t) = \frac{\displaystyle\int_A \rho UT dA}{\rho U_b A} \tag{5}$$

and $\bar{T}_{\text{rod},m}(x)$ is the time average of the circumferentially averaged rod surface temperature $T_{\text{rod},m}(x,t)$, defined as (ϕ is the azimuthal direction with respect to the rod axis)

$$T_{\text{rod},m}(x,t) = \frac{1}{2\pi}\int_0^{2\pi} T_{\text{rod}}(x,\phi,t)d\phi \tag{6}$$

The periodic boundary condition for the temperature was enforced by matching the values of the dimensionless temperature Θ on the inlet plane with corresponding values on the exit plane at all time instances. For mode A, this simply meant to match the actual temperature values at corresponding locations on the inlet and exit planes. For mode B, however, the periodic boundary condition was applied such that the temperatures at corresponding locations on the inlet and outlet planes were connected by the relationship

$$T(0,y,z,t) = T(L,y,z,t) - \frac{\dot{q}}{\dot{m}c_p}L \tag{7}$$

where c_p is the specific heat under constant pressure. Application of this periodic boundary condition for mode B ensures full development of the temperature fluctuation field while keeping the temperature rise within bounds even as heat enters continuously into the system. For the uniform rod temperature case $\bar{T}_{\text{rod},m}(x)$ and $\bar{T}_b(x)$ are independent of x, whereas for the uniform rod heat flux case $\bar{T}_{\text{rod},m}(x)$ and $\bar{T}_b(x)$ increase linearly with x at the rate $\dot{q}/(\dot{m}c_p)$. The time-averaged local temperature also increases at the same rate. Thus, in both cases, time-averaged properties of the dimensionless temperature difference Θ would be independent of x.

An additional requirement of the solution is the specification of the bulk temperature $\bar{T}_b(0)$ at the inlet, time averaged and mass weighted over the entire cross section. In the present analysis, we specified $\bar{T}_b(0) = 300$ K, which was equal to the temperature used in the isothermal simulations. The values of $\bar{T}_{\text{rod},m}$ and $\bar{T}_b$ are immaterial, as long as they are sufficiently close to each other for the heating to have a negligible effect on the velocity field.

For the uniform temperature case, the time-averaged local heat flux $\dot{q}$ from the rod surface was computed as

$$\dot{q} = \kappa \frac{\partial \overline{T}}{\partial r} \qquad (8)$$

where κ is the thermal conductivity of the fluid and r is the radial direction. Then, the local heat transfer coefficient on the surface of the rod was calculated as

$$\overline{h} = \frac{\dot{q}}{\overline{T}_{\text{rod},m} - \overline{T}_b} \qquad (9)$$

The time-averaged local heat transfer coefficient is presented in dimensionless form as the time-averaged local Nusselt number

$$\overline{\text{Nu}} = \frac{\overline{h} D_h}{\kappa} \qquad (10)$$

In our previous isothermal study [12], we have presented tests of possible dependence of the solution to mesh size and time step. It was demonstrated that logarithmic velocity profiles near the wall computed using meshes with 1.3×10^6 and 10.2×10^6 elements were in good agreement with each other, while the logarithmic velocity profile computed using 0.3×10^6 elements did not follow the same overall trend. A comparison of simulations with time steps equal to 1.9×10^{-2} and 1.9×10^{-3} t_c, where $t_c = L / U_b$ is the average flow turnover time (namely the average time that it takes for the fluid to pass through the computational domain) showed that both resolved the coherent structures during an initial computation period, but the appearance of these structures deteriorated in time when the longer time step was used, whereas it did not for the shorter time step. Additional grid independence studies for the velocity and temperature fields in the non-isothermal cases were not deemed to be necessary, in view of the facts that the heating was passive and the fluid Prandlt number was of order unity. Similar to the isothermal computations, the flow domain away from walls was represented by an unstructured mesh with roughly 1.3×10^6 hexahedral elements (see Fig. 1); in addition, a subgrid with roughly 0.65×10^6 hexahedral elements was used in the wall regions. All elements had a streamwise length equal to about $0.10D$. The adopted grid is the finest possible that would produce an accurate solution of the problem in a reasonably timely manner. The effect of abrupt changes in mesh size near the wall is not expected to have a significant effect on the solution [21].

As in the isothermal studies, the time step was set equal to $1.9 \times 10^{-3} t_c$. The solution at each time step was considered to converge when all normalized residuals became less than 10^{-4}, except for the temperature, for which the value 10^{-6} was used instead. Time averaging in the present simulations was over $3.8\ t_c$, which is equal to 2000 time steps, following full development of the flow structure inside the channel. The full development of the flow structure was ensured by examining the spanwise velocity and temperature fluctuation patterns on the equidistant plane, on which the effects of coherent structures are most apparent [12]. The time averaging extended over approximately 13 average periods of coherent structure passage, which, according to Iaccarino and Durbin [22], is sufficient for phase averages to converge. Considering that the total computing time for these simulations was already quite long (about 1400 h), extending the simulations over longer times was impractical, while also introducing the risk of solution deterioration due to the accumulation of numerical errors.

The present URANS simulations distinguish between resolved and modeled velocity and temperature fluctuations and it is important first to justify the use of the adopted terminology and second to understand the significance of the computed results. Rough estimates of relevant scales for the velocity fluctuations, determined experimentally for the present configuration [23], were as follows:

- spacing between two consecutive pairs of coherent structures: 4 to $5D$ (see also CT)

- streamwise integral length scale of the streamwise velocity (away from walls and gaps): $0.19D$
- streamwise Taylor microscale (away from walls): $0.05D$
- Kolmogorov microscale (away from walls): $0.0017D$.

A direct numerical simulation of this flow, capable of resolving motions with scales of the order of the Kolmogorov microscale, would require a mesh with roughly 2×10^{10} elements, whereas a large eddy simulation would typically be required to resolve 2% of the integral length scale, which corresponds roughly to 2×10^9 elements. Neither approach would be possible to implement with the available resources. The presently used main grid (i.e., not including the boundary layers) had elements with dimensions roughly equal to $0.03D$ on the transverse plane and $0.10D$ in the streamwise direction. Thus, this mesh was adequate in resolving the coherent motions in the gap region, but not most of the energy contained in conventional turbulent fluctuations and essentially none of the dissipative turbulent activity. Concerning the physical significance of the reported results, CT identified the velocity fluctuations resolved by the present URANS method as coherent and the ones modeled by the RSM as non-coherent, although, strictly speaking, the resolved velocity field could contain a small amount of non-coherent activity and the non-resolved field would contain some small-scale coherent activity, not related to the dominant structures of interest in the present study. This approach is based on the assumption that the additional cross-correlation terms arising due to coherent and non-coherent fluctuations are uncorrelated [24–26], which can be justified by the large difference in the time scales of the coherent and non-coherent fluctuations [24,26,27]. Small-scale coherent structures in the boundary layers, which cannot be resolved by the present scheme, would be accounted for, together with the non-coherent turbulence, by the RSM. In support of this distinction, the computed azimuthal velocity spectrum in the gap region contained a sharp peak (see discussion in the following section), associated with the cross-gap flow pulsations. The average period of the resolved coherent structures was about $0.29\ t_c$, which is two orders of magnitude larger than the temporal resolution (i.e., the time step) of the present scheme. Thus, the coherent vortices which generate cross-gap flow pulsations were resolved well both in space and in time by the present URANS approach.

Definitions of the coherent and non-coherent turbulent stresses have been discussed by CT. In a similar manner, the local variance of the temperature fluctuations is determined as the sum of coherent and non-coherent components, as

$$\overline{\vartheta^2} = \overline{\vartheta_{nc}^2} + \overline{\vartheta_c^2} \qquad (11)$$

The coherent component $\overline{\vartheta_c^2}$ is the variance of the resolved temperature T, given by

$$\overline{\vartheta_c^2} = \overline{(T - \overline{T})^2} \qquad (12)$$

where $\overline{T}$ is the time-averaged, resolved local temperature. The non-coherent temperature fluctuations cannot be determined precisely, because the present solution procedure does not solve turbulence equations for the temperature field. In order to compare approximately the magnitudes of the coherent and non-coherent temperature contributions, we have estimated the variance of the non-coherent temperature fluctuations using the following expression [28]

$$\overline{\vartheta_{nc}^2} = -\frac{k_{nc}}{C_{\vartheta 1} \epsilon_{nc}} \left((u\vartheta)_{nc} \frac{\partial T}{\partial x} + (v\vartheta)_{nc} \frac{\partial T}{\partial y} + (w\vartheta)_{nc} \frac{\partial T}{\partial z} \right) \qquad (13)$$

This expression is an algebraic model of the balance equation for the non-coherent temperature fluctuation variance, in which convection and diffusion terms were ignored, an assumption that seems appropriate for the present flow. The local, time-dependent resolved heat fluxes $(u\vartheta)_{nc}, (v\vartheta)_{nc}, (w\vartheta)_{nc}$ were determined from

the resolved temperature using Eq. (2), and the value of the coefficient $C_{\vartheta 1}$ was taken as 0.62.

3 Results and Discussion

The cross-sectional variations of the time-averaged dimensionless local temperature difference $\bar{\Theta}$ for the two heating modes examined are shown in Fig. 2(a). It can be seen that relatively high temperature regions were restricted around the rod and across the gap and that the wall facing the gap was far warmer than any of the other walls. The isocontours of $\bar{\Theta}$ were slightly closer to the rod for mode B than for mode A, but had comparable shapes for both modes. The apparent inclination of some isotherms to the wall is an artifact of the near-wall mesh refinement; unlike the appearance in this figure, in reality, all computed isotherms approached the wall perpendicularly. Figures 2(b) and 2(c) show, respectively, isocontours of the normalized standard deviations of the coherent and non-coherent temperature fluctuations. Each pair of isocontour sets had similar shapes and comparable corresponding values for the two modes. The coherent temperature fluctuations were mostly confined in the gap region, while the non-coherent fluctuations, modeled by Eq. (13), were distributed nearly uniformly around the rod with strong values near the rod. Coherent temperature fluctuations accounted for up to 50% of the estimated total temperature variance for both modes A and B (Fig. 2(d)), dominating the temperature field in the gap region, much like coherent fluctuations dominated the velocity fluctuations in the same region [12].

The spanwise variations of the time-averaged dimensionless temperature differences $\bar{\Theta}$ and two representative profiles of instantaneous resolved dimensionless temperature differences Θ across the bottom wall are plotted for the two heating modes in Fig. 3. It may be seen that, although the maximum time-averaged temperature was on the symmetry plane, the locations of instantaneous temperature maxima fluctuated within the approximate range $-0.5 < z/D < 0.5$ for both heating modes.

The streamwise variation of the time-averaged dimensionless temperature difference $\bar{\Theta}$ on the bottom wall is depicted in Fig. 4, in which two representative plots of instantaneous resolved temperature differences have also been shown for comparison. It is observed that, although the local time-averaged temperature difference on the wall was essentially independent of streamwise distance for both heating modes, its instantaneous value fluctuated by as much as 57% for mode A and by up to 69% for mode B. These instantaneous temperature fluctuations are associated with cross-gap motions caused by the passage of coherent structures. Figure 5 shows the variation of the Nusselt number around the rod, whereas Fig. 6 shows the corresponding variation along the rod, on the symmetry plane, and facing the gap. Both figures show time-averaged values and two representative instantaneous variations. Averaged in time and around the rod, the Nusselt number was about 190 for both heating modes. Local instantaneous values of the Nusselt number were significantly lower than the corresponding averages around the rod (as much as 38% lower for mode A and 20% lower for mode B). The minimum time-averaged Nusselt number was located on the symmetry plane towards the gap, while instantaneous minima spanned an angular range as wide as possibly ±60 deg around the symmetry plane. The maximum Nusselt number was not on the symmetry plane facing the top wall, but on either side of it, at about ±30 deg from this plane. The time-averaged Nusselt number was essentially constant along the bottom of the rod, approximately equal to 165 for both heating modes. In contrast, the instantaneous local Nusselt number along the same line reached values lower than the time average by as much as about 30% for both modes.

From the above discussion, it is evident that the resolved velocity fields and temperature fields are strongly time dependent. This time dependence is intimately associated with the convection of large-scale, quasi-periodic, vortical coherent structures, which
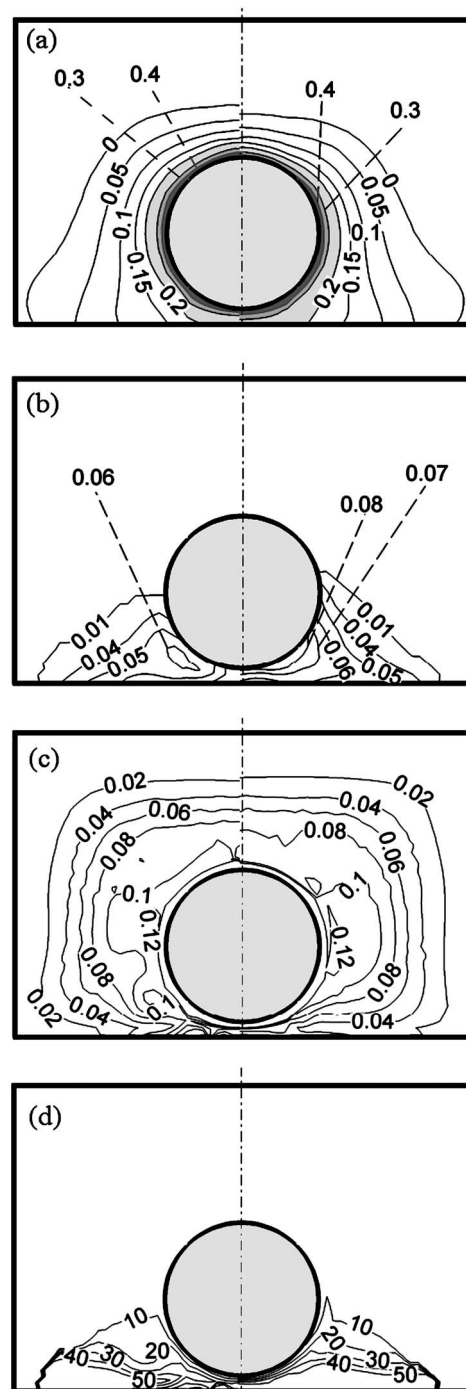
Fig. 2 Predicted isocontours of (a) the dimensionless time-averaged temperature difference $\bar{\Theta} = (\bar{T} - \bar{T}_b)/(\bar{T}_{\text{rod},m} - \bar{T}_b)$; (b) the standard deviation of the normalized coherent temperature fluctuations $\sqrt{\bar{\vartheta}_c^2}/(\bar{T}_{\text{rod},m} - \bar{T}_b)$; (c) the standard deviation of the normalized non-coherent temperature fluctuations $\sqrt{\bar{\vartheta}_{nc}^2}/(\bar{T}_{\text{rod},m} - \bar{T}_b)$; and (d) the percent ratio of the variances of coherent and total temperature fluctuations $\overline{\vartheta_c^2}/\overline{\vartheta^2} \times 100$; the left-side plots correspond to the uniform rod temperature case and the right-side plots to the uniform rod heat flux case

appear in pairs on either side of the gap between the rod and the channel plane wall. To illustrate the influence of the coherent structures on the instantaneous temperature fields, Fig. 7 shows isocontours of the instantaneous resolved temperature difference on the equidistant plane, together with the locations of coherent
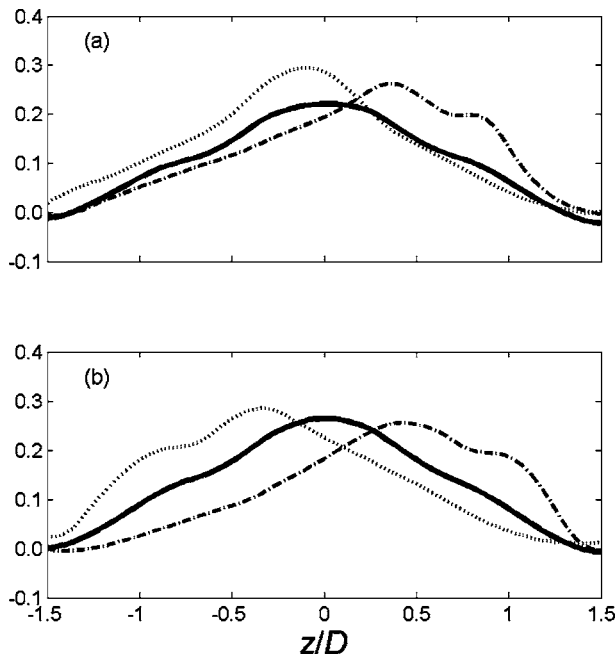
Fig. 3 Spanwise variations of the time-averaged (——) and representative instantaneous (- - -, –•–•–) dimensionless temperature differences on the bottom wall in (*a*) the uniform rod temperature case and (*b*) the uniform rod heat flux case

structures, captured using the Q criterion. The characteristics of these structures, based on phase-averaged measurements, have been presented by Guellouz and Tavoularis [8], while their numerical identification and detailed features were discussed by CT. Briefly, these structures consist of pairs of counterrotating vortices with axes alternating on either side of the gap and transporting
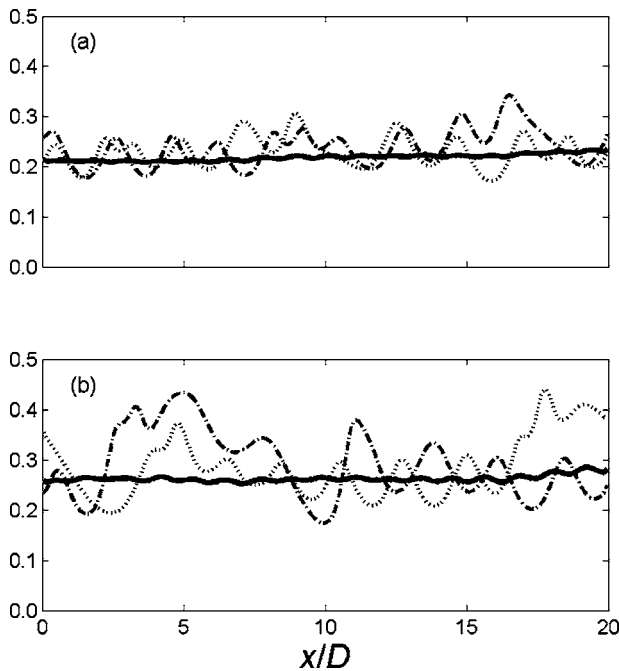

Fig. 4 Streamwise variations of the time-averaged (——) and representative instantaneous (- - -, –•–•–) dimensionless temperature differences along the bottom wall at *z/D*=0 in (*a*) the uniform rod temperature case and (*b*) the uniform rod heat flux case
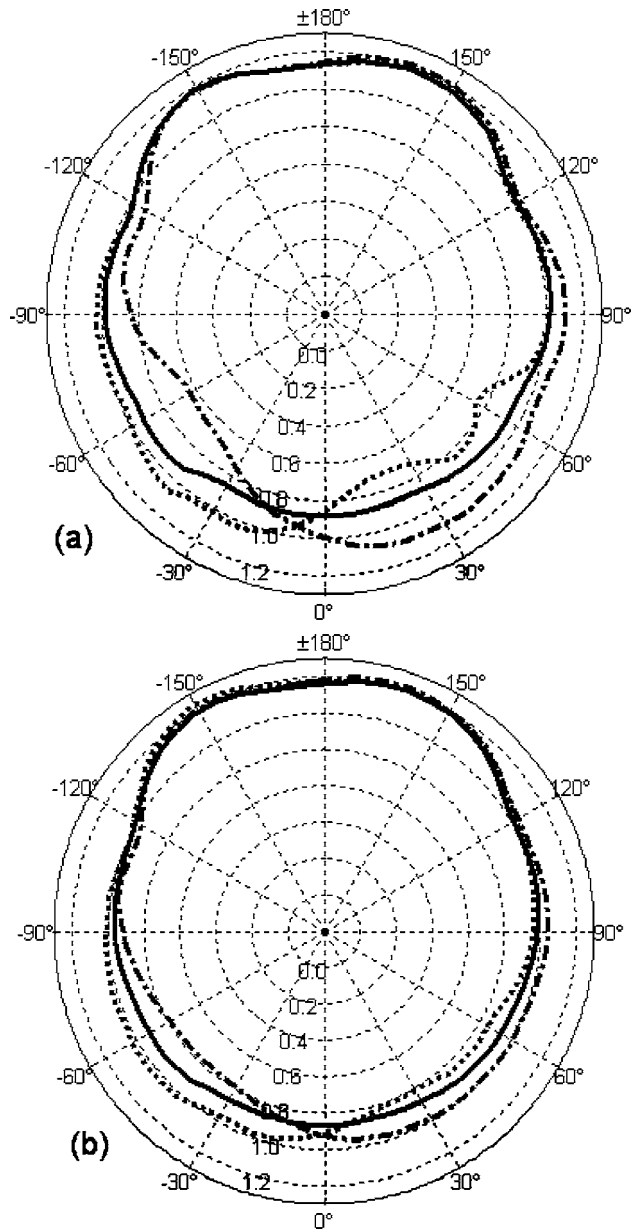

Fig. 5 Circumferential variations of the time-averaged (——) and representative instantaneous (- - -, –•–•–) Nusselt numbers around the rod in (*a*) the uniform rod temperature case and (*b*) the uniform rod heat flux case; the values have been normalized by the corresponding circumferential averages

fluid far across the gap. This transport has the beneficial effect of moderating the time-averaged temperature rise in the gap region, which otherwise would have been significantly higher. The highest local instantaneous temperatures occur in the gap region between structures, occurring at streamwise locations at which there is little cross flow across the gap.

Finally, Fig. 8 shows power spectra of the spanwise velocity and the temperature in the center of the gap. The frequency f was presented in dimensionless form as a Strouhal number

$$\text{St} = \frac{fD}{U_b} \qquad (14)$$

Note that CT used the hydraulic diameter D_h, rather than the rod diameter D, in their Strouhal number definition, but the present choice is deemed to be more appropriate for comparisons with other rod bundle geometries. The Strouhal number at the peak
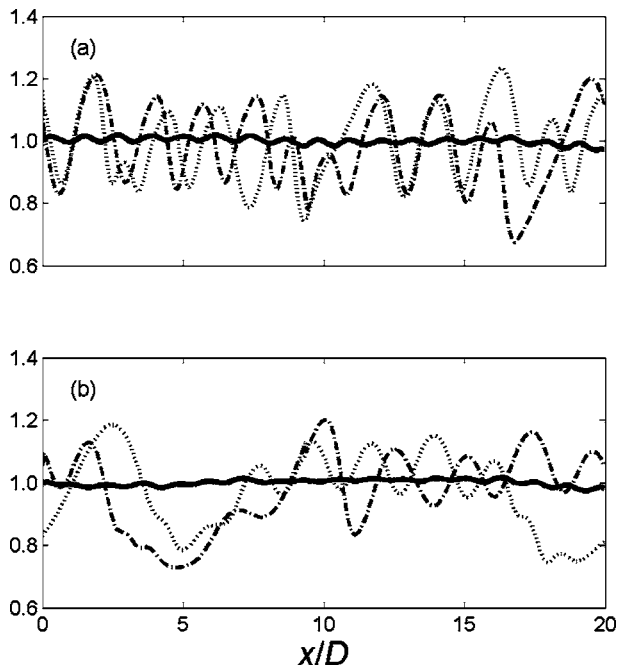
Fig. 6 Streamwise variations of the time-averaged (——) and representative instantaneous (- - -, –•–•–) Nusselt numbers along the rod on the symmetry plane and facing the bottom wall in (*a*) the uniform rod temperature case and (*b*) the uniform rod heat flux case; the values have been normalized by the corresponding streamwise averages

frequency of the velocity spectrum was about 0.17, whereas the Strouhal number at the peak frequency of the temperature spectrum for both heating modes was approximately twice as large. This is perfectly consistent with the symmetry of the channel and the balanced strengths of the two vortices in each pair: while the spanwise velocity alternates *once* across the gap during each cycle, cool fluid from the two subchannels flanking the gap region is swept across the gap *twice* in each cycle. Figure 8(*a*) also shows
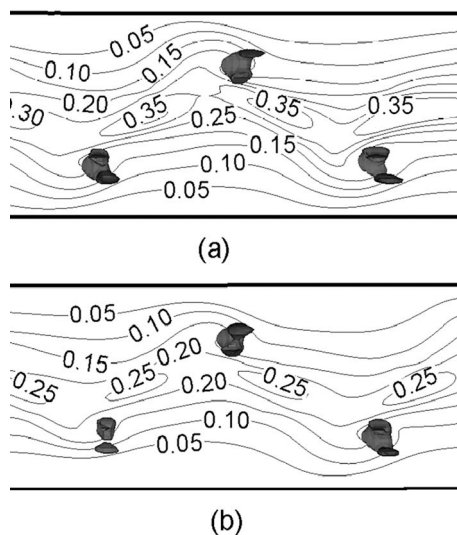

Fig. 7 Predicted isocontours of the dimensionless instantaneous temperature difference Θ on the equidistant plane in (*a*) the uniform rod temperature case and (*b*) the uniform rod heat flux case; coherent structures identified by the *Q* criterion are also shown in both plots, which represent a flow area that is 7.5*D* long
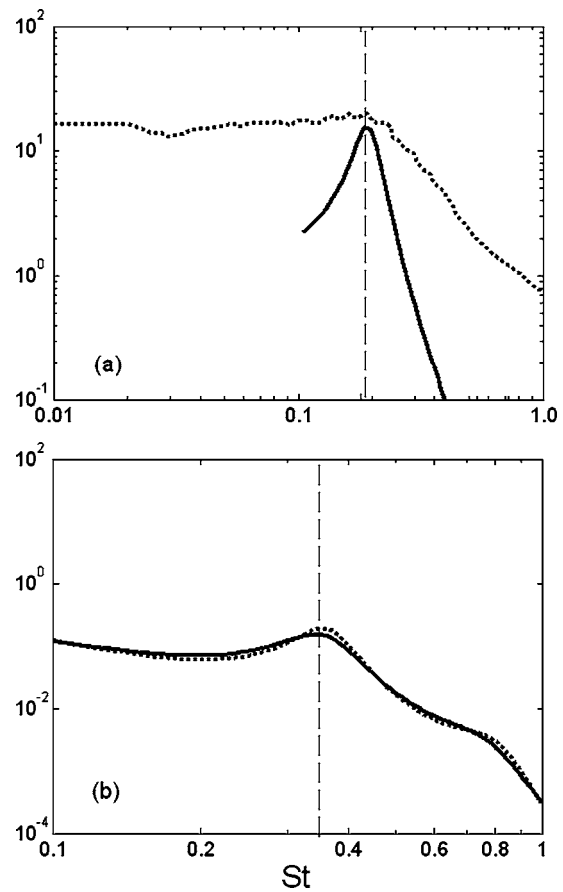
Fig. 8 (*a*) Computed (—) and measured [7] (- - - -) smoothened power spectra of the spanwise velocity in the center of the gap; (*b*) computed smoothened spectra of the resolved temperature in the constant rod temperature case (—) and the constant rod heat flux case (- - - -)

a comparison of the computed and experimental (total) spectra of the azimuthal velocity in the gap center, which demonstrates that the URANS approach acts like a bandpass filter, separating the coherent fluctuations from the non-coherent ones. The present approach produces comparable spectral effects to those generated by adaptive filters and the proper orthogonal decomposition, as described by Bonnet and Delville [26]. A note of explanation seems useful when comparing the experimental and numerical frequency spectra shown in Fig. 8(*a*). The peak in the experimental spectrum, although detectable, is not nearly as prominent as the peak in the numerical spectrum. This is due to the inability of Fourier analysis to account for phase shifts and missing cycles in the signals. Such phenomena, which can clearly be seen in the corresponding velocity signals (see Fig. 9 of Ref. [7] and Fig. 11 of Ref. [12]), reflect lateral and streamwise jitter of the coherent structures, occasional merging of them, and other distortions and broaden the spectral peaks. The measured spectral peaks become more distinct when the gap is reduced (see Fig. 10 of Ref. [7]), which indicates not only the strengthening, but also the tighter organization of the coherent structures at narrower gaps. Sharp peaks in measured spectra of the azimuthal velocity component in rod-bundle gaps have also been presented by several other researchers (e.g., [3,5,29,30]). In the present simulations, the spectral energy is concentrated at a peak because the fine structure and most of the non-coherent turbulence are filtered by the algorithm; moreover, the periodic boundary condition and the limited streamwise length of the domain effectively filter fluctuations with scales much larger than the dominant structure scales and plausibly tend to organize the dominant structures spatially to reduce spatial scat-

ter and phase shifts. As discussed by CT, the strength of the predicted coherent fluctuations in the gap center was essentially the same as the corresponding experimental value, computed by phase averaging of the measured signals [8], irrespective of the differences in the appearances of the experimental and numerical spectra.

An important question of concern to designers of thermal equipment using liquids as coolants is whether the local fluctuations in rod and channel wall temperatures would be of sufficiently long durations to permit overheating and possible phase change, which may lead to dryout. The time scale t_f for the duration of such fluctuations can be estimated as the inverse of the peak frequency of the temperature spectrum. For the present configuration, the corresponding Strouhal number was about 0.35, which gives the "period" of temperature fluctuations as

$$t_f \approx 3 \frac{D}{U_b} \tag{15}$$

There is ample experimental evidence [3] that the Strouhal number of coherent structures in rod bundles is insensitive to the Reynolds number. Consequently, expression (15) is expected to be approximately applicable to a wide range of flow conditions. On the other hand, because the temperature spectral peak is wide band, one should consider that parts of the wall would be exposed to local high temperatures over periods which may be several times longer than t_f.

4 Conclusion

The present simulations have clearly documented the significance of coherent vortical structures in heat transfer in a rod-wall gap region. These structures transport fluid across the gap, moderating the time-averaged temperature rise in the gap region but also creating substantial local variations of the instantaneous temperature and heat transfer coefficient. These fluctuations are at relatively low frequencies and may have to be taken into consideration in safety analyses.

Acknowledgment

Nomenclature

A = cross-sectional area
C_μ = constant in the eddy viscosity equation
$C_{\vartheta 1}$ = constant in the temperature fluctuation equation
c_p = specific heat at constant pressure
D = rod diameter
D_h = hydraulic diameter
f = frequency
h = local heat transfer coefficient
k = local turbulent kinetic energy per unit mass
L = streamwise length of the duct
$\dot{m}$ = mass flow rate
Nu = Nusselt number, hD_h/k
P = static pressure
$\dot{q}$ = local heat flux
Re = Reynolds number, $\rho U_b D_h/\mu$
r = radial coordinate with respect to the rod center
St = Strouhal number, fD/U_b
T = temperature
t = time
t_c = average flow turnover time
t_f = characteristic time of temperature fluctuations
U, V, W = velocity components
U_b = bulk velocity
u, v, w = velocity fluctuations

x, y, z = streamwise, transverse, and spanwise coordinates, respectively

Greek Symbols

γ = thermal diffusivity coefficient
δ = gap between the rod and duct walls
ϵ = turbulent kinetic energy dissipation rate
Θ = dimensionless temperature difference
ϑ = temperature fluctuation
κ = thermal conductivity of the fluid
μ = dynamic viscosity
μ_T = turbulence viscosity
ρ = density
σ_T = turbulent Prandtl number
ϕ = angular coordinate with respect to the rod center

Subscripts

b = bulk quantity
c = coherent quantity
nc = non-coherent quantity
rod, m = circumferentially averaged quantity over the rod surface

Other Notation

$(\bar{\ })$ = time-averaged value

References

[1] Rowe, D. S., Johnson, B. M., and Knudsen, J. G., 1974, "Implications Concerning Rod Bundle Crossflow Mixing Based on Measurements of Turbulent Flow Structure," Int. J. Heat Mass Transfer, **17**, pp. 407–419.
[2] Hooper, J. D., and Rehme, K., 1984, "Large-Scale Structural Effects in Developed Turbulent Flow through Closely-Spaced Rod Arrays," J. Fluid Mech., **145**, pp. 305–337.
[3] Möller, S. V., 1992, "Single-Phase Turbulent Mixing in Rod Bundles," Exp. Therm. Fluid Sci., **5**, pp. 26–33.
[4] Wu, X., 1995, "On the Transport Mechanisms in Simulated Heterogeneous Rod Bundle Subchannels," Nucl. Eng. Des., **158**, pp. 125–134.
[5] Krauss, T., and Meyer, L., 1996, "Characteristics of Turbulent Velocity and Temperature in a Wall Channel of a Heated Rod Bundle," Exp. Therm. Fluid Sci., **12**, pp. 75–86.
[6] Krauss, T., and Meyer, L., 1998, "Experimental Investigation of Turbulent Transport of Momentum and Energy in a Heated Rod Bundle," Nucl. Eng. Des., **180**, pp. 185–206.
[7] Guellouz, M. S., and Tavoularis, S., 2000, "The Structure of Turbulent Flow in a Rectangular Channel Containing a Cylindrical Rod – Part 1: Reynolds-Averaged Measurements," Exp. Therm. Fluid Sci., **23**, pp. 59–73.
[8] Guellouz, M. S., and Tavoularis, S., 2000, "The Structure of Turbulent Flow in a Rectangular Channel Containing a Cylindrical Rod – Part 2: Phase-Averaged Measurements," Exp. Therm. Fluid Sci., **23**, pp. 75–91.
[9] Guellouz, M. S., and Tavoularis, S., 1992, "Heat Transfer in Rod Bundle Subchannels With Varying Rod-Wall Proximity," Nucl. Eng. Des., **132**, pp. 351–366.
[10] Groeneveld, D. C., 1973, "Forced Convective Heat Transfer to Superheated Steam in Rod Bundles," Report No. AECL-4450, Atomic Energy of Canada Ltd., Chalk River, Ontario, Canada.
[11] Wu, X., 1994, "Numerical Study on the Turbulence Structures in Closely Spaced Rod Bundle Subchannels," Numer. Heat Transfer, Part A, **25**, pp. 649–670.
[12] Chang, D., and Tavoularis, S., 2005, "Unsteady Numerical Simulations of Turbulence and Coherent Structures in Axial Flow Near a Narrow Gap," ASME J. Fluids Eng., **127**, pp. 458–466.
[13] Launder, B. E., and Shima, N., 1989, "Second-Moment Closure for the Near-Wall Sublayer: Development and Application," AIAA J., **27**(10), pp. 1319–1325.
[14] Yakhot, V., and Orszag, S. A., 1986, "Renormalization Group Analysis of Turbulence: I. Basic Theory," J. Sci. Comput., **1**, pp. 3–51.
[15] Chen, H. C., and Patel, V. C., 1988, "Near-Wall Turbulence Models for Complex Flows Including Separation," AIAA J., **26**(6), pp. 641–648.
[16] Barth, T. J., and Jespersen, D. C., 1989, "The Design and Application of Upwind Schemes on Unstructured Meshes," AIAA Paper No. 89-0366, *AIAA 27th Aerospace Sciences Meeting*, Reno, Nevada.
[17] Van Doormal, J. P., and Raithby, G. D., 1984, "Enhancements of the SIMPLE Method for Predicting Incompressible Fluid Flows," Numer. Heat Transfer, **7**, pp. 147–163.
[18] Gresho, P. M., Lee, R. L., and Sani, R. L., 1980, "On the Time-Dependent Solution of the Incompressible Navier-Stokes Equations in Two and Three Dimensions," *Recent Advances in Numerical Methods in Fluids*, C. Taylor and K. Morgan, eds., Pineridge, Swansea, UK, Chap. 2.
[19] Murthy, J. Y., and Mathur, S., 1997, "Periodic Flow and Heat Transfer Using

Unstructured Meshes," Int. J. Numer. Methods Fluids, **25**, pp. 659–677.

[20] Burmeister, L. C., 1983, *Convective Heat Transfer*, Wiley, New York, Chap. 12.

[21] Schröder, K., and Gelbe, H., 1999, "Two- and Three-Dimensional CFD-Simulation of Flow-Induced Vibration Excitation in Tube Bundles," Chem. Eng. Process., **38**, pp. 621–629.

[22] Iaccarino, G., and Durbin, P., 2000, "Unsteady 3D RANS Simulations Using the v^2-f Model," *Annual Research Briefs 2000*, Center for Turbulence Research, Stanford University, Stanford, CA, pp. 263–269.

[23] Guellouz, M. S., 1998, "Turbulent Flow and Heat Transfer in Rod Bundles," Ph.D. thesis, Department of Mechanical Engineering, University of Ottawa, Ottawa, Ontario, Canada.

[24] Hussain, A. K. M. F., 1983, "Coherent Structures – Reality and Myth," Phys. Fluids, **26**, pp. 2816–2850.

[25] Wilcox, D. C., 2000, *Turbulence Modeling for CFD*, DCW Industries, La Canada, CA.

[26] Bonnet, J. P., and Delville, J., 2001, "Review of Coherent Structures in Turbulent Free Shear Flows and Their Possible Influence on Computational Methods," Flow, Turbul. Combust., **66**, pp. 333–353.

[27] Spalart, P. R., 2000, "Strategies for Turbulence Modeling and Simulation," Int. J. Heat Fluid Flow, **21**, pp. 252–263.

[28] Chen, C. J., and Jaw, S. Y., 1998, *Fundamentals of Turbulence Modeling*, Taylor and Francis, Washington, DC, Chap. 6.

[29] Wu, X., and Trupp, A. C., 1994, "Spectral Measurements and Mixing Correlation in Simulated Rod Bundle Subchannels," Int. J. Heat Mass Transfer, **37**, pp. 1277–1281.

[30] Baratto, F., Bailey, S. C. C., and Tavoularis, S., 2006, "Measurements of Frequencies and Spatial Correlations of Coherent Structures in Rod Bundle Flows," Nucl. Eng. Des. (in press).

Steady-Periodic Green's Functions and Thermal-Measurement Applications in Rectangular Coordinates

Kevin D. Cole

Mechanical Engineering Department,
University of Nebraska–Lincoln,
Lincoln, NE 68588-0656
e-mail: kcole1@unl.edu

Methods of thermal property measurements based on steady-periodic heating are indirect techniques, in which the thermal properties are deduced from a systematic comparison between experimental data and heat-transfer theory. In this paper heat-transfer theory is presented for a variety of two-dimensional geometries applicable to steady-periodic thermal-property techniques. The method of Green's functions is used to systematically treat rectangles, slabs (two dimensional), and semi-infinite bodies. Several boundary conditions are treated, including convection and boundaries containing a thin, high-conductivity film. The family of solutions presented here provides an opportunity for verification of numerical results by the use of distinct, but similar, geometries. A second opportunity for verification arises from alternate forms of the Green's function, from which alternate series expressions may be constructed for the same unique temperature solution. Numerical examples are given to demonstrate both verification techniques for the steady-periodic response to a heated strip. [DOI: 10.1115/1.2194040]

1 Introduction

Steady-periodic heat conduction, of interest for thermal-property measurements, may be treated analytically in two ways. In the time-domain approach, the solution can be stated with the Duhamel integral as a time convolution between the heating history and the step response [1]. A particular simplification is possible if the time history of the heating has a simple wave shape (on-off, sawtooth, etc.). In these cases the convolution integral can be evaluated in closed form so that the time dependence reduces to a series involving decaying time exponentials. This approach is used for determination of thermal properties by the thermal hot-strip method [2–5].

The frequency-domain approach is appropriate if the heating history is sinusoidal [6], or, if the heating is simply periodic and a phase-locked amplifier is used to select the response at the periodic frequency. This is the approach used for thermal-property measurements by photothermal techniques [7,8]. In a unique study, Aviles-Ramos [9] applied the frequency-domain approach to virtual steady-periodic data which was constructed, via Fourier series, from transient non-periodic data.

The frequency-domain approach has also been applied to the 3-omega method for determination of thermal properties [10]. In this method a metallic strip is plated on the sample, and suitable electronics are used to periodically energize the strip and simultaneously measure its temperature response. The temperature response is systematically compared with theory to deduce the thermal properties of the sample. The 3-omega method has recently been extended to two-dimensional layered composite materials [11,12] with a matrix technique originally described by Carlaw and Jaeger [1] for one-dimensional materials. In this method a matrix equation, constructed from matching conditions between adjacent layers, is solved to satisfy all the inter-layer boundary conditions simultaneously. In the traditional matrix technique, steady-periodic heating is applied to one surface of the composite and no internal heating of the material is considered. In contrast, the Green's function (GF) method allows for internal heating, such as from optical absorption within partially transparent layers. In previous work by the author, the Green's function method was combined with the matrix technique for one-dimensional heat transfer in a layered material [13] and for a layered material heated by an axisymmetric laser beam [14].

The purpose of this paper is to revisit steady-periodic heat conduction with the method of Green's functions (GFs). The GF approach provides a comprehensive set of solutions and specific strategies for improving the numerical evaluation of these solutions. The geometries under discussion, two-dimensional bodies with planar boundaries in the Cartesian coordinate system, have practical application to thermal property measurements.

There are several recent books on Green's functions applied to heat conduction. Beck et al. [15] give a large number of GFs for transient heat conduction. Both GF and temperature solutions are organized according to a numbering system for the domain shape and for the kind of boundary conditions present. Five kinds of boundary conditions are discussed. Although several two-dimensional geometries are discussed, no steady-periodic cases are given. Duffy [16] gives a large number of GFs for several differential equations. There is a section devoted to steady heat conduction (Poisson equation) in two-dimensional rectangular geometries, and boundaries of kinds 1, 2, and 3 are given. However, no two-dimensional transient solutions are treated. Sheremet [17] discusses GFs for the Lamé and Poisson equations in rectangular coordinates. Several GFs for steady two-dimensional heat conduction are given, however only boundaries of kinds 1 and 2 are treated. Mandelis [18] is devoted exclusively to steady-periodic heat conduction with the method of GF, and solutions are given for a variety of geometries and boundary conditions of types 1, 2 or 3. Although many cases are given in Cartesian coordinates, the emphasis is on one- and three-dimensional cases; only one GF is given that is two-dimensional in Cartesian coordinates.

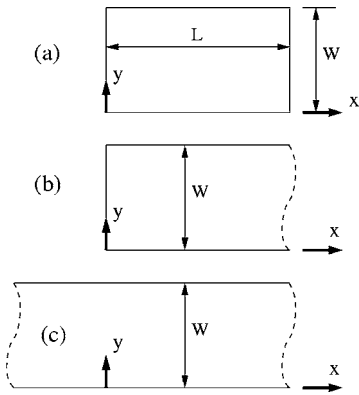The contributions of this paper are threefold. First, a great many

Fig. 1 Geometries under discussion include (*a*) rectangles, (*b*) semislabs, and (*c*) slabs

steady-periodic solutions are presented systematically with the method of GFs for rectangles, slabs, semislabs, and semi-infinite regions. All solutions presented are two dimensional. Second, five kinds of boundary conditions are treated in a unified fashion: specified temperature; specified heat flux; specified convection; a high-conductivity surface film; and a high-conductivity surface film with convection. To the author's knowledge many of the surface-film solutions for steady-periodic conduction have not been published before. Third, alternate forms of the GFs are given for several geometries, which provide for efficient numerical computation and allow for independent verification that the numerical results are correct.

The paper is divided into sections on temperature, Green's functions, measurement applications, and numerical examples. In the next section the temperature solution is given for a wide variety of body shapes in rectangular coordinates.

2 Temperature in Rectangular Coordinates

Consider the temperature in a two-dimensional domain in rectangular coordinates. The temperature distribution $\widetilde{T}(x,y,t)$ satisfies

$$\frac{\partial^2 \widetilde{T}}{\partial x^2} + \frac{\partial^2 \widetilde{T}}{\partial y^2} - \frac{1}{\alpha}\frac{\partial \widetilde{T}}{\partial t} = -\frac{1}{k}\widetilde{g}(x,y,t) \quad \text{in domain } \Omega \quad (1)$$

$$k_i \frac{\partial \widetilde{T}}{\partial n_i} + h_i \widetilde{T} + (\rho cb)_i \frac{\partial \widetilde{T}}{\partial t} = \widetilde{f}_i(x_i,y_i,t); \quad \text{at boundary } i \quad (2)$$

Domain Ω can include the rectangle, semislab and slab as shown in Fig. 1. Additional geometries included in Ω but not shown in Fig. 1 are the semi-infinite body $(0 < y < \infty)$ and the infinite body.

Index *i* represents the physical boundaries, up to a maximum of four boundaries, for body shapes considered here. The boundary condition may be one of five kinds at each boundary depending on the values of coefficients k_i, h_i, and b_i (refer to Table 1): boundary type 1 is specified temperature; boundary type 2 is specified heat flux; boundary type 3 is specified convection; boundary type 4 is specified heat flux and a surface film of thickness b_i; and boundary type 5 is specified convection and a surface film. Boundary type 0 is also used to represent a boundary at infinity.

Since in this paper the applications are limited to steady-periodic heating, we take the heating terms $\widetilde{g}$ and $\widetilde{f}$ and the resulting temperature $\widetilde{T}(x,y,t)$ to be steady periodic at a single frequency. That is, let

$$\widetilde{g}(x,y,t) = \text{Real}[g(x,y,\omega)e^{j\omega t}]$$

$$\widetilde{f}_i(x_i,y_i,t) = \text{Real}[f_i(x_i,y_i,\omega)e^{j\omega t}] \quad (3)$$

Table 1 Boundary conditions for temperature

Number	Description	Boundary condition
0	Boundary at ∞	($\partial \widetilde{T}$ bounded)
1	Specified temperature	$\widetilde{T} = \widetilde{f}_i$
2	Specified heat flux	$k \partial \widetilde{T}/\partial n_i = \widetilde{f}_i$
3	Convection	$k \partial \widetilde{T}/\partial n_i + h_i \widetilde{T} = \widetilde{f}_i$
4	Heat flux and surface film	$k\frac{\partial \widetilde{T}}{\partial n_i} + (\rho cb)_i \frac{\partial \widetilde{T}}{\partial t} = \widetilde{f}_i$
5	Convection and surface film	(see Eq. (2))

$$\widetilde{T}(x,y,t) = \text{Real}[T(x,y,\omega)e^{j\omega t}]$$

Now in Eqs. (1) and (2) replace $\widetilde{g}$, $\widetilde{f}_i$, and $\widetilde{T}$ with $ge^{j\omega t}$, $f_i e^{j\omega t}$, and $Te^{j\omega t}$, respectively, to find the steady-periodic heat conduction equation

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} - \frac{j\omega}{\alpha}T = -\frac{1}{k}g(x,y,\omega); \quad \text{in domain } \Omega \quad (4)$$

$$k_i \frac{\partial T}{\partial n_i} + [h_i + j\omega(\rho cb)_i]T = f_i(x_i,y_i,\omega); \quad \text{at boundary } i \quad (5)$$

In this paper complex-valued $T(x,y,\omega)$ is interpreted as the steady-periodic temperature (Kelvin) at a single frequency ω. For further discussion of this point see [18]. Later in the paper, results will be discussed in the form of amplitude and phase of this temperature.

The steady-periodic temperature will be found with the Green's function (GF) method. Assume for the moment that the appropriate GF in frequency space, G, is known. Then the steady-periodic temperature is given by the following integral equation (see [15])

$$T(x,y,\omega)$$

$$= \frac{\alpha}{k} \int \int g(x',y',\omega)G(x,x',y,y',\omega)dx'dy'$$

(for volume heating)

$$+ \alpha \int_{s_i} f_i \times \begin{bmatrix} -\partial G/\partial n_i' & \text{(type 1 only)} \\ \frac{1}{k}G & \text{(type 2 - 5)} \end{bmatrix} ds_i'$$

(at boundaries) $\quad (6)$

Note that the same GF appears in each integral term but it is evaluated at locations appropriate for each integral.

3 Green's Function

The GF represents the response at (x,y) to a steady-periodic point source of heat located at (x',y'). The GF associated with Eqs. (4)–(6) is defined by

$$\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} - \frac{j\omega}{\alpha}G = -\frac{1}{\alpha}\delta(x - x')\delta(y - y') \quad \text{in domain } \Omega$$

$$(7)$$

$$k_i \frac{\partial G}{\partial n_i} + [h_i + j\omega(\rho cb)_i]G = 0; \quad \text{at boundary } i \quad (8)$$

Here $\delta(\cdot)$ is the Dirac delta function. It is important to note that the GF satisfies homogeneous boundary conditions of the same kind as the temperature equation.

Table 2 Eigenfunctions for rectangles, slabs, and semislabs

Case	$Y_n(y)$
Y11, Y12, Y13, Y14, and Y15	$\sin(\gamma_n y)$
Y21, Y22, Y23, Y24, and Y25	$\cos(\gamma_n y)$
YK1, YK2, YK3, YK4, and YK5[a]	$\gamma_n W \cos(\gamma_n y)$ $+ (\lambda_1 W/k)\sin(\gamma_n y)$

[a]K=3, 4, or 5

Table 3 Inverse norm and eigenvalues or eigencondition. Note $B_i = \lambda_i W/k$.

Case	$N(\gamma_n)^{-1}$	γ_n or eigencondition
Y11	$2/W$	$n\pi/W$
Y12	$2/W$	$(2n-1)\pi/(2W)$
Y13, Y14, Y15[a]	$2\phi_{2n}/W$	$\gamma_n W \cot(\gamma_n W) = -B_2$
Y21	$2/W$	$(2n-1)\pi/(2W)$
Y22	$\frac{2}{W};\ \gamma_n \neq 0$ $\frac{1}{W};\ \gamma_n = 0$	$n\pi/W;\ n=0,1,2,\dots$
Y23, Y24, Y25[a]	$2\phi_{2n}/W$	$\gamma_n W \tan(\gamma_n W) = B_2$
Y31, Y41, Y51[a]	$2\phi_{1n}/W$	$\gamma_n W \cot(\gamma_n W) = -B_1$
Y32, Y42, Y52[a]	$2\phi_{1n}/W$	$\gamma_n W \tan(\gamma_n W) = B_1$
Y33, Y34, Y35[b] Y43, Y44, Y45[b] Y53, Y54, Y55[b]	$2\Phi_n/W$	$\tan(\gamma_n W) = \frac{\gamma_n W(B_1+B_2)}{(\gamma_n W)^2 - B_1 B_2}$

[a]$\phi_{in} = [(\gamma_n W)^2 + B_i^2] \div [(\gamma_n W)^2 + B_i^2 + B_i]$
[b]$\Phi_n = \phi_{2n} \div [(\gamma_n W)^2 + B_1^2 + B_1 \phi_{2n}]$

4 GF Number

The GF needed for a given temperature solution is determined by the body shape and by the kind of boundary conditions present. To distinguish among all the different GFs that are included in this paper, we use a "number" of the form $XijYkl$ in which X and Y represent the coordinate axes, and the letters following each axis name take on values 1, 2, 3, 4, or 5 to represent the kind of boundary conditions present at the body faces normal to that axis. Number 0 is also used to represent a boundary at infinity.

For example, number $X12$ represents boundary conditions of type 1 at $x=0$ and type 2 at $x=L$, and $X30$ represents boundary conditions of type 3 at $x=0$ on a body that extends to $x=\infty$. As another example, number $X11Y13$ describes a GF for a rectangle with three faces having type 1 boundaries ($G=0$) and the face at $y=W$ has a type 3 boundary (convection). As a final example, number $X00Y12$ represents a slab, infinite in the x direction, with boundary of type 1 at $y=0$ and a boundary of type 2 at $y=W$.

The number system described here is useful for classifying Green's functions. Extensions of this number system for classifying temperature solutions, including designations for energy generation, initial conditions, and time and space distributions in the boundary conditions, are given in the book by Beck et al. [15] and in the Green's Function Library, an internet site devoted to Green's functions [19].

5 GF for the Rectangle, Semislab, and Slab

The GF may be found as a series expansion involving eigenfunctions for bodies of finite size in at least one direction. This applies to rectangles, semislabs and slabs in rectangular coordinates in two dimensions. The GFs for these body shapes that satisfy Eqs. (7) and (8) are given below in a single-sum form

$$G(x,y,\omega|x',y') = \sum_{n=0}^{\infty} \frac{Y_n(y)Y_n^*(y')}{N_y(\gamma_n)} P(x,x',\sigma) \qquad (9)$$

The series for the GF contains eigenfunction Y_n, norm N_y, and kernel function P which will be discussed below. The $n=0$ term is needed only when zero is an eigenvalue (when $Y22$ is part of the GF number).

5.1 Eigenfunctions. The y-direction eigenfunction satisfies the following ordinary differential equation

$$Y_n''(y) + \gamma_n^2 Y_n(y) = 0, \quad (0 < y < W) \qquad (10)$$

where γ_n is the associated eigenvalue. There are three different eigenfunctions associated with the 25 possible combinations of boundary condition YKL ($K,L=1, 2, 3, 4, or 5$). Eigenfunctions $Y_n(y)$ are composed of sines and cosines, and are listed in Table 2. Table 3 contains the associated inverse norms and eigenconditions (or eigenvalues for simple cases). For case Y22 the eigenvalue may also take on the value zero.

Boundary conditions of type 4 or 5, which include a thin surface film, require special care because the eigenvalues are complex numbers and the eigenfunctions contain complex-valued sine and/or cosine. Complex-valued eigenvalues have been previously shown to occur for heat conduction in multi-layer, multi-dimensional bodies [20].

5.2 Kernel Functions. With the above choices for $Y_n(x)$ and $N(\lambda_n)$ the kernel function $P(x,x')$ must satisfy

$$\frac{d^2 P}{dx^2} - \sigma^2 P = -\frac{1}{\alpha}\delta(x-x') \qquad (11)$$

Here function P has units $(s\ m^{-1})$ and parameter $\sigma^2 = \gamma_n^2 + j\omega/\alpha$ has units, m^{-1}. The solution for P may be found using two solutions of the homogeneous equation that satisfy the boundary conditions and are joined appropriately at $x=x'$ (see, for example, [21]). The kernel functions are given by

$$P(x,x',\sigma) = \frac{S_2^-(S_1^- e^{-\sigma(2L-|x-x'|)} + S_1^+ e^{-\sigma(2L-x-x')})}{2\alpha\sigma(S_1^+ S_2^+ - S_1^- S_2^- e^{-2\sigma L})}$$
$$+ \frac{S_2^+(S_1^+ e^{-\sigma(|x-x'|)} + S_1^- e^{-\sigma(x+x')})}{2\alpha\sigma(S_1^+ S_2^+ - S_1^- S_2^- e^{-2\sigma L})} \qquad (12)$$

where the subscripts 1 and 2 represent the two boundaries at the smallest and largest x values, respectively. Coefficients S_M^+ and S_M^- depend on the boundary conditions on side M and are given by

$$S_M^+ = \begin{cases} 1 & \text{if side } M \text{ is type 0, type 1, or type 2} \\ k\sigma + \lambda_M & \text{if side } M \text{ is type 3, 4, or 5} \end{cases}$$

$$S_M^- = \begin{cases} 0 & \text{if side } M \text{ is type 0} \\ -1 & \text{if side } M \text{ is type 1} \\ 1 & \text{if side } M \text{ is type 2} \\ k\sigma - \lambda_M & \text{if side } M \text{ is type 3, 4, or 5} \end{cases}$$

The derivation of the kernel function given in Eq. (12) parallels that for steady-state GF given elsewhere [22]; however, in the present work σ is complex. The special case $\omega=0$, steady state has been treated previously for rectangles [23] and slabs [24].

5.3 Alternate Forms of GF. Although the GF is the unique solution to Eq. (7), for many geometries there exist alternate forms for the GF. These alternate forms have a very important role in numerical evaluation of the GF and the temperatures constructed from them. Specifically, the alternate GF can be used for verification that computed numerical values are correct. This usage of the word "verification" is somewhat similar to Roach [25] who discusses quantifying the error in a finite-element code by comparison with analytical solutions. In the present usage, verification is a comparison between two computer codes for numerical evaluation of alternate forms of the same analytical solution. Here the primary goal is improving one's confidence in the results,

rather than quantifying the error.

5.3.1 Alternate GF for Rectangles. In the rectangle an alternate series for the GF may be found by placing the kernel functions in the y direction and the eigenfunctions in the x direction. The alternate GF is important because at a point in the rectangle where one series converges slowly, the other series converges rapidly, and vice versa. In previous work with steady temperature, we have shown that there are locations in the domain at which the slowly converging series requires thousands of times more terms than the rapidly converging series [22,23]. A double-sum form of the GF may also be found from Fourier expansions along both x and y, however it generally converges very slowly and should not be used when a single-sum form is available.

5.3.2 Alternate GF for Slab Bodies. An alternate GF for slab bodies may be found with a spatial Fourier transform. Consider slab bodies described by cases *X00YIJ* for which I, J=1, 2, 3, 4, or 5. The solution will be found with a spatial Fourier transform defined by the following transform pair

$$\overline{G}(\beta) = \int_{-\infty}^{\infty} G(x)e^{-j\beta x}dx \qquad (13)$$

$$G(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{G}(\beta)e^{j\beta x}d\beta \qquad (14)$$

Note that variable x' has been suppressed by a change of variable, replacing $(x-x')$ by x, which is allowed under Eq. (7) which defines G. Apply the above transform to Eqs. (7) and (8) to obtain

$$\frac{\partial \overline{G}^2}{\partial y^2} - \nu^2 \overline{G} = -\frac{1}{\alpha}\delta(y - y') \qquad (15)$$

$$k_i\frac{\partial \overline{G}}{\partial n_i} + \lambda_i\overline{G} = 0 \quad \text{at boundary } i \qquad (16)$$

$$\text{where} \quad \nu^2 = \beta^2 + j\omega/\alpha \qquad (17)$$

$$\lambda = h_i + j\omega(\rho cb)_i \qquad (18)$$

Equation (15) is similar to Eq. (11) which defines the kernel function, so the solution for $\overline{G}$ is given by the kernel function from Eq. (12) with parameter σ replaced by ν and x replaced by y. That is, $\overline{G}(\beta, y, y', \omega) = P(y, y', \nu)$. Finally, the GF may be formally stated in x space by use of the inverse transform

$$G(x, y, \omega|x', y') = \frac{1}{2\pi}\int_{-\infty}^{\infty} P(y, y', \nu)e^{j\beta(x-x')}d\beta \qquad (19)$$

Here variable x' has been recovered by reversing the earlier change of variable and replacing x by $(x-x')$. In general, the inverse-transform integral must be evaluated numerically, which is possible because the integrand approaches zero as $\beta \to \pm\infty$.

6 GF for Infinite and Semi-infinite Bodies

The GFs for infinite and semi-infinite bodies are found in the same manner as the alternate GF discussed above for the slab body. For cases *X00YI0* for I=0, 1, 2, 3, 4 or 5, the kernel function P given in Eq. (12) may be simplified by taking $S_2^-=0$ and $S_2^+=1$. Then the GF for infinite and semi-infinite bodies may be written

$$G(x, y, \omega|x', y') = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{j\beta(x-x')}}{2\alpha\nu}[e^{-\nu|y-y'|} + De^{-\nu(y+y')}]d\beta$$
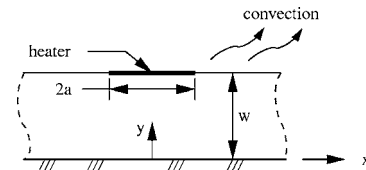

Fig. 2 Slab heated over a small area and cooled by convection

$$\text{where} \quad D = \begin{cases} 0 & \text{(infinite body)} \\ -1 & \text{(type 1 at } y=0) \\ +1 & \text{(type 2 at } y=0) \\ (k\nu - \lambda)/(k\nu + \lambda) & \text{(type 3, 4, or 5)} \end{cases} \qquad (20)$$

For some cases the β integral is known in closed form. For cases *X00Y00*, *X00Y10*, and *X00Y20*, the GF may be written [18]

$$G(x, y, \omega|x', y') = \frac{1}{2\pi\alpha}\Big[K_0\big(\nu\sqrt{(x-x')^2 + (y-y')^2}\big)$$
$$+ DK_0\big(\nu\sqrt{(x+x')^2 + (y+y')^2}\big)\Big] \qquad (21)$$

where K_0 is the modified Bessel function of order zero (with complex argument) and D=0, -1, or 1 for cases *X00Y00*, *X00Y10*, and *X00Y20*, respectively.

7 Measurement Applications

In this section the temperature in bodies heated over a small region will be studied as simple models of devices used for measurement of thermal properties. For example, in the 3-omega method, a thin metal strip is plated on a solid surface and suitable electronics are used to introduce heat in a steady-periodic fashion. The thermal response is measured on the metal strip or at other locations on the solid surface [10]. As another example, in photothermal methods, a sample is heated by a periodically modulated laser beam and the thermal response is measured by optical or acoustic methods [7,8,14].

Generally thermal properties are measured indirectly, through a type of inverse problem, in which the thermal properties are deduced by a systematic comparison between the experimental data (such as temperature) and a thermal model. The thermal model is the subject of this discussion. The inverse problem, although an important part of the measurement process, is beyond the scope of this paper. For a discussion of inverse methods associated with thermal properties, see [6,26]. Next, numerical examples are given of thermal models appropriate for measurement applications constructed with two-dimensional steady-periodic GFs.

7.1 Slab Heated Over a Small Region. In this example a slab body is heated over a small region and cooled by convection on one side. The other surface of the body is insulated. The GF number for this case is *X00Y23* and the geometry is shown in Fig. 2. The temperature is formally given by Eq. (6) with volume heating

$$T(x, y, \omega) = \frac{\alpha}{k}\int\int g(x', y')G_{X00Y23}(x, x', y, y', \omega)dx'dy' \qquad (22)$$

The heated region is of infinitesimal thickness along y and is piecewise constant along x, described by $g(x', y')=q(x')\delta(y'-W)$ where

$$q(x') = \begin{cases} q_0; & |x'| < a \\ 0; & |x'| > a \end{cases}$$

This heating function could represent an electrically heated metal film with negligible thermal mass. Substitute the above heating function into the temperature expression, Eq. (22), to obtain
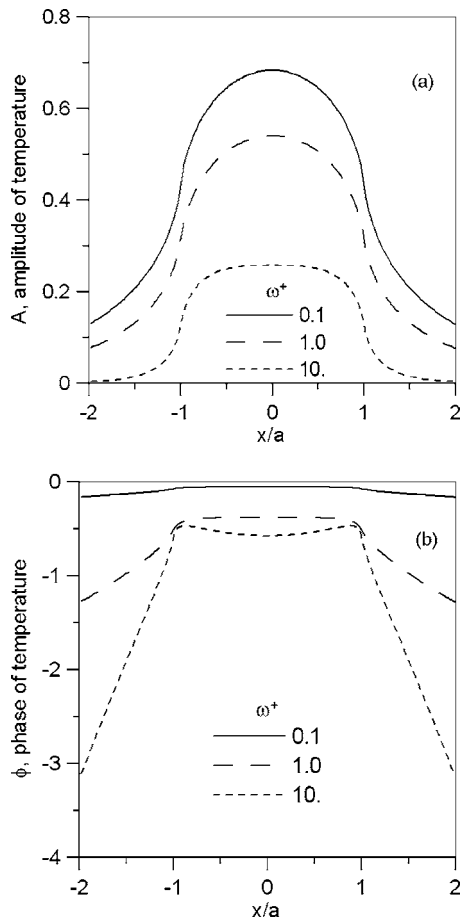
Fig. 3 Amplitude (*a*) and phase (*b*) of the temperature on the heated surface of a slab with $B_a=1$ and $W/a=1$ for three heating frequencies

Fig. 4 Amplitude (*a*) and phase (*b*) of the temperature on the heated surface of a slab with heating frequency $\omega^+=1$ and $W/a=1$ for three Biot values

$$T(x,y,\omega) = \frac{\alpha}{k} \int_{-a}^{a} q_0 G_{X00Y23}(x,x',y,y'=W,\omega)dx' \quad (23)$$

Note that the integral on y' has been stripped away by the Dirac delta function. The series form of the GF, Eq. (9), will be used to find the temperature. Substitute the appropriate eigenfunction from Table 2, norm from Table 3, and kernel function from Eq. (12) to obtain

$$T(x,y,\omega) = \frac{\alpha q_0}{k} \sum_{n=0}^{\infty} \frac{\cos(\gamma_n W)\cos(\gamma_n y)}{N_y} \int_{-a}^{a} \frac{e^{-\sigma|x-x'|}}{2\sigma\alpha}dx' \quad (24)$$

$$\text{where} \quad N_y^{-1} = \frac{2}{W} \frac{\gamma_n^2 W^2 + (B_2 W/a)^2}{\gamma_n^2 W^2 + (B_2 W/a)^2 + B_2 W/a} \quad (25)$$

Note that the integral on x' may be carried out in closed form.

Next, results are presented for the amplitude and phase of the dimensionless temperature on the $y=W$ surface of the slab. The amplitude of the temperature A, and phase of temperature ϕ, are computed as follows

$$A = [T \cdot T^*]^{1/2} \quad (26)$$

$$\phi = \tan^{-1}[\text{Imag}(T)/\text{Real}(T)] \quad (27)$$

where "Imag" and "Real" are the imaginary and real parts of the (complex) temperature, respectively. The following dimensionless variables are used for reporting results: for temperature, $T^+ = Tk/(q_0 a)$; for geometry, W/a; for convection, Biot number $B_a = ha/k$; and for frequency $\omega^+ = \omega a^2/\alpha$. Figure 3 shows results for
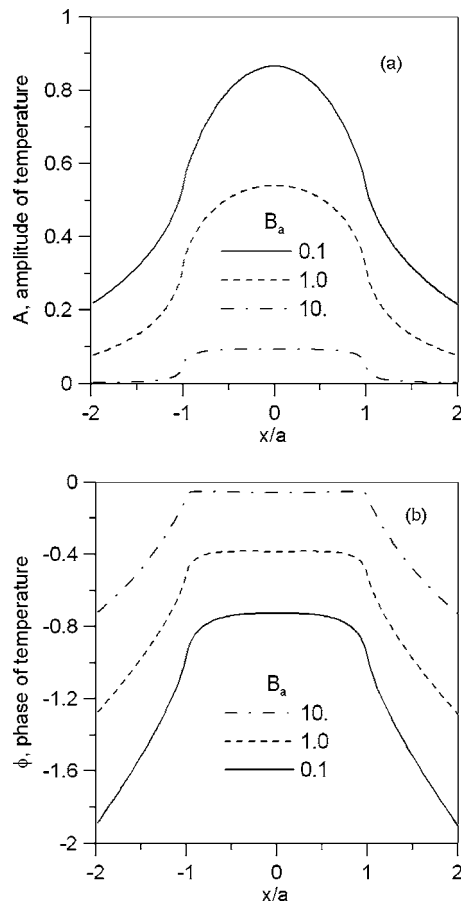
the surface temperature under conditions $B_a=1$ and $W/a=1$ for various values of the frequency ω^+. As the frequency increases the temperature amplitude decreases and the phase becomes more negative. Also for higher frequency a smaller region of the surface is participating in the periodic temperature variations. Figure 4 shows results for $\omega^+=1$, $W/a=1$ for several values of the Biot number, B_a, which controls convection cooling. At higher Biot number the amplitude decreases (convection cools the body) but the phase moves closer to zero. This is because with vigorous convection there is a smaller volume of the body participating in the heat transfer, and the phase lag will be smaller when the thermal mass is smaller. Figure 5 shows results for $B_a=1$ and $\omega^+=1$ for several values of the slab thickness W/a. As the slab becomes thinner the temperature amplitude grows, but it is more closely confined to the heated region. This is because there is less area for heat conduction along the x coordinate. As the slab becomes thinner the phase in the heated region moves toward zero, again because a smaller thermal mass participates in the heat transfer.

The alternate GF was also used to find the temperature in this case. The alternate GF is given by Eqs. (19) and (12) with $S_1^+ = S_1^- = 1$, $S_2^+ = k\nu + \lambda_2$, and $S_2^- = k\nu - \lambda_2$ as appropriate for case $X00Y23$. Then the GF, evaluated at $y'=y=W$, is given by

$$G_{X00Y23}(x,0|x',0,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-j\beta(x-x')}W(1-e^{-2\nu W})d\beta}{\alpha[\nu W + B_2 - (\nu W - B_2)e^{-2\nu W}]}$$

$$(28)$$

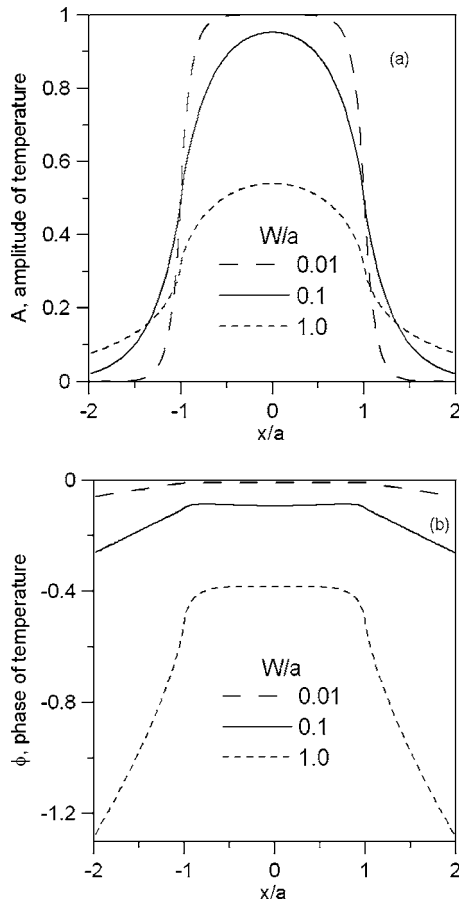The temperature is found by replacing the above GF into Eq. (23), to give

Fig. 5 Amplitude (a) and phase (b) of the temperature on the heated surface of a slab with $B_a=1$ and $\omega^+=1$ for three slab thicknesses

$$T(x,y,\omega) = \frac{q_0 W}{k} \int_{-\infty}^{\infty} \frac{e^{-j\beta(x-a)} - e^{-j\beta(x+a)}}{j\beta}$$
$$\times \frac{(1 - e^{-2\nu W})d\beta}{[\nu W + B_2 - (\nu W - B_2)e^{-2\nu W}]} \quad (29)$$

Note that the integral over x' has been carried out in closed form, but the remaining integral on β must be carried out numerically. A Romberg integration scheme was used here. In a comparison with the series form of the temperature, Eq. (24), it was found that the above integral form required much more computer time for convergence than for the series form. Because of this, the integral form, Eq. (29), was primarily used for checking that the two forms produced plots in reasonable agreement. Agreement to three decimal places was used for this purpose. Higher-accuracy agreement could have been sought by tightening the convergence criteria, however this calculation was not attempted.

7.2 Thin Film on a Thick Substrate. Consider a thick substrate with properties k, ρc which is coated with a thin surface film with high conductivity ($k_1 \gg k$) and with volume thermal capacity $(\rho c)_1$. The surface film is in perfect thermal contact with the substrate. There is a strip heater and convection cooling similar to the previous example. Refer to Fig. 6. The GF needed for this case has number $X00Y50$.

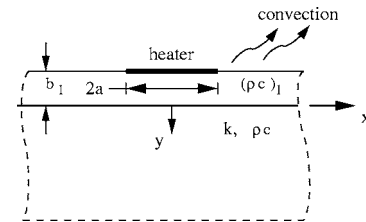The temperatures in the substrate may be formally stated with the GF solution equation given in Eq. (6)

Fig. 6 Thin film on a thick substrate, heated over a small area and cooled by convection

$$T(x,y,\omega) = \frac{\alpha}{k} \int_{-\infty}^{\infty} q(x') G_{X00Y50}(x,y,\omega | x', y'=0) dx' \quad (30)$$

where $q(x')$ is the heat distribution introduced at $y=0$. The GF for this case, given in Eq. (20), contains a β integral that in general must be carried out with numerical quadrature. One approach to evaluating the above temperature would be to find G numerically and then evaluate the integral on x' numerically. A better approach is to examine the temperature in β space where the x' integral can be evaluated in closed form. Specifically, perform a spatial Fourier transform on Eq. (30) according to Eq. (13), and apply the convolution rule

$$\bar{T}(\beta,y,\omega) = \frac{\alpha}{k}\bar{q}(\beta)\bar{G}_{X00Y50}(\beta,y,y'=0,\omega) \quad (31)$$

In this approach the x' integral is removed by the Fourier transform. The β-space GF, $\bar{G}$, has been discussed earlier. Function $\bar{q}$ must be found from the spatial distribution of heating and the Fourier transform. For the strip heater with piecewise-constant heating over ($-a < x < a$), function $\bar{q}$ is given by

$$\bar{q}(\beta) = q_0 \int_{-a}^{a} e^{-j\beta x} dx = q_0 \frac{e^{j\beta a} - e^{-j\beta a}}{j\beta} \quad (32)$$

Function $\bar{q}$ can also be found in closed form for a variety of other heating distributions (Gaussian, point source, etc.). Then, the temperature in real space is given by Eq. (31) combined with the transform-inversion integral

$$T(x,y,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{T}(\beta,y,\omega) e^{j\beta x} d\beta \quad (33)$$

7.2.1 Average Temperature on the Heater. Rather than show the spatial distribution of temperature, which has many of the same trends as the previous example, numerical results are presented below for the average temperature on the heater.

The average temperature on the heater is of great importance for thermal-property measurements for which a metal film is used simultaneously as a heater and as a temperature sensor. To find the spatial average, evaluate the above temperature expression on the body surface ($y=0$), integrate over the heated region, and divide by the length of the heated region

$$T_{av}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{T}(\beta,y=0,\omega)\left[\frac{1}{2a}\int_{-a}^{a} e^{j\beta x} dx\right]d\beta \quad (34)$$

Note that the spatial integral, shown in brackets in the above expression, may be evaluated in closed form. Now use $\bar{T}$ from Eq. (31), $\bar{q}$ from Eq. (32), and $\bar{G}$ from Eq. (20) to find

$$T_{av}(\omega) = \frac{1}{4\pi}\frac{q_o a}{k}\int_{-\infty}^{\infty}\frac{(e^{j\beta a} - e^{-j\beta a})^2}{(j\beta a)^2}\frac{ad\beta}{(\nu a + \lambda a/k)} \quad (35)$$

Although numerical quadrature for the transform-inversion integral cannot be avoided, with this approach quadrature is only
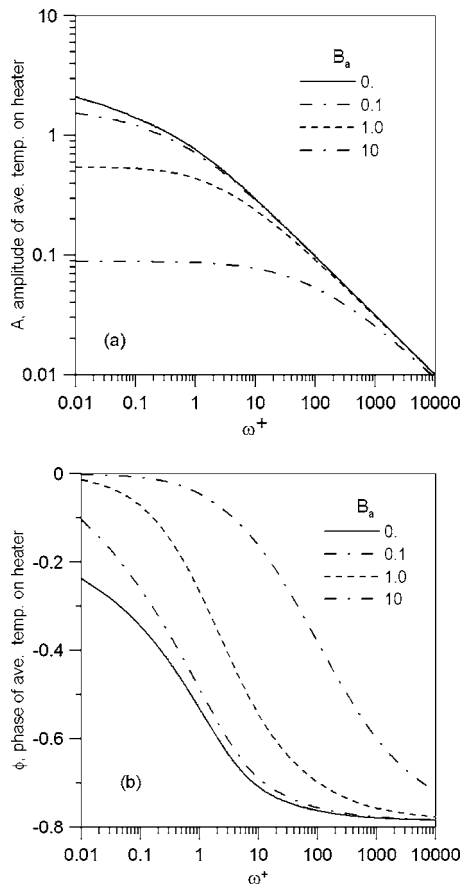
Fig. 7 Amplitude (*a*) and phase (*b*) of the spatial average temperature on the heater for a thick substrate with no surface film as a function of frequency for three convection values

Fig. 8 Amplitude (*a*) and phase (*b*) of the spatial average temperature on the heater as a function of frequency at B_a=1 for various thicknesses of films on a large substrate. For the films, $(\rho c)_1/(\rho c)$=1.

needed once.

Next numerical results will be discussed, beginning with the semi-infinite substrate with no surface film (b_1=0). Figure 7 shows the dimensionless spatial-average temperature (amplitude and phase) on the heater over a wide range of frequency values for values of the surface convection. The highest amplitude occurs for Biot = 0 representing no convection cooling (insulated surface). As the convection value increases the amplitude falls and the phase moves closer to zero. These trends agree with those discussed in the previous example.

Figure 8 shows the effect of adding a high-conductivity surface film, with $(\rho c)_1/(\rho c)$=1.0, to the substrate. All results shown in Fig. 8 are at Biot=1.0 and the heating conditions are the same as those shown in Fig. 7. Figure 8 shows that the surface film has an increasing effect as the film thickness increases and as the frequency increases. A comment is needed on the frequency behavior of the results with the surface film. There is a maximum frequency beyond which the thin-film assumption breaks down, determined approximately when $\omega b_1^2/\alpha \approx 1$ when the thermal penetration into the substrate becomes small enough to approach the thickness of the surface film. This effect may explain the shape of the phase curves in Fig. 8 at higher frequencies. The effect is more pronounced as b_1 becomes larger.

Results for several other Biot values were also explored, however, as the trends are identical to those shown in Fig. 8, results for Biot ≠ 1 are not shown to save space.

7.2.2 Computer Issues. About 2 min of computer time was required to compute the 300 temperature values plotted for each curve shown in Figs. 7 and 8. Computations were carried out on a
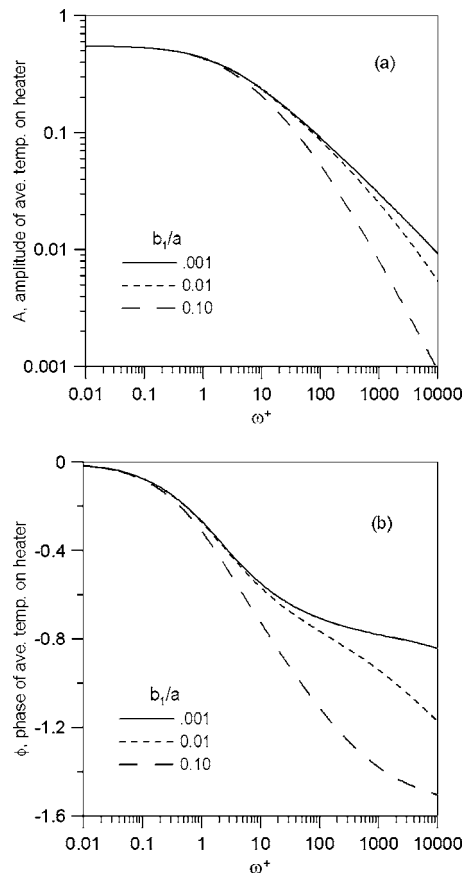
Sun Blade 2000 with dual 900 MHz processors running the Solaris operating system. The calculations were coded in Fortran 77 with variables of type double-precision complex.

Numerical integration for the inverse transform from β space to x space was carried out efficiently with the recognition that the integrand contains an infinite number of zero crossings along the β axis at regular intervals of 2π, and that the amplitude decays monotonically as $\beta \rightarrow \pm\infty$. Then, the integration over domain ($-\infty < \beta < \infty$) was carried out by a series of integrations over subdomains of size 2π beginning at β=0. Within each subdomain, a Romberg integration scheme converged rapidly. By examining the relative contribution of successive subdomains, the infinite domain was truncated while retaining control over the numerical precision of the result. The numerical-integration routine for the inverse transform was checked by computing case *X*00*Y*20 based on Eq. (20), and then comparing with the closed-form expression given in Eq. (21). Six-digit agreement was obtained.

8 Conclusions

Verification of numerical results is an important part of any numerical calculation. The Green's function method presented here provides two distinct opportunities for verification. First, a variety of Green's functions are presented for closely related geometries, which allow for the construction of useful limiting cases. For example, case *X*00*Y*30 was used to check case *X*00*Y*32. Second, the alternate Green's functions, given here for rectangles and slabs, may be used to construct alternate expressions for the same temperature. These alternate temperature expressions, although they represent the same unique solution, have different

computational behaviors. They may be checked, one against the other, to ensure that numerical values may be found not only with high precision, but with high confidence that the results are correct.

In this paper a variety of Green's function solutions are given for steady-periodic heat conduction in two-dimensional bodies in rectangular coordinates. Five types of boundary conditions are treated. Numerical results are presented for two geometries involving bodies heated over a small region related to thermal-property measurements. Because the method is computationally efficient, results were reported over 6 decades of frequency under several different conditions.

Nomenclature

a = half-width of heated region (m)
A = temperature amplitude, Eq. (26)
B_a = Biot number, ha/k
B_i = Biot number, $h_i W/k$, see Table 3.
b_i = film thickness on boundary i (m)
c = specific heat (J kg^{-1} K^{-1})
D = Coefficient in Eq. (20)
f = known effect at boundary i
G = steady-periodic Green's function (s m^{-2})
g = internal heating at frequency ω (W m^{-3})
h = heat transfer coefficient, (W m^{-2} K^{-1})
j = imaginary number, $\sqrt{-1}$
K_0 = modified Bessel function, order zero
k = thermal conductivity (W m^{-1} K^{-1})
L = length of domain in x direction (m)
n_i = outward-facing unit normal vector on boundary i
N_y = norm, Eq. (9) (m)
P = kernel function, Eq. (12) (s m^1)
q = steady-periodic heat flux (W m^{-2})
S_M = coefficient for kernel function in Eq. (12)
ds_i = dx_i or dy_i depending on boundary i (m)
$\underline{T}$ = steady-periodic temperature (K)
$\overline{T}$ = β-space temperature (K m)
T^+ = $Tk/(q_0 a)$
t = time (s)
W = width of domain in y direction (m)
Y_n = eigenfunction, Eq. (10)
Y_n^* = complex conjugate of eigenfunction

Greek

α = thermal diffusivity (m^2 s^{-1})
β = wave number, Eq. (14) (m^{-1})
γ = eigenvalue, Table 3 (m^{-1})
δ = Dirac delta function
λ = boundary parameter, Eq. (18)
ν = $(\beta^2 + j\omega/\alpha)^{1/2}$ (m^{-1})
ρ = density (kg m^{-3})
σ = $(\gamma_n^2 + j\omega/\alpha)^{1/2}$ (m^{-1})

ϕ = phase, Eq. (27)
ω = frequency (rad s^{-1})
Ω = domain

References

[1] Carslaw, H. S., and Jaeger, J. C., 1959, *Conduction of Heat in Solids*, Oxford University Press, Oxford, UK, pp. 108,111.
[2] Gustafsson, S. E., 1979, "Thermal Properties of Thin Insulating Layers Using Pulse Transient Hot Strip Measurements," J. Phys. D, **12**, pp. 1411–1421.
[3] English, A. T., Miller, G. L., Robinson, D. A. H., Dodd, L. V., and Chynoweth, T., 1978, "Pulse Nonlinearity Measurements on Thin Conducting Films," J. Appl. Phys., **49**, pp. 717–722.
[4] Gustafsson, S. E., and Karawacki, E., 1983, "Transient Hot-Strip Probe for Measuring Thermal Properties of Insulating Solids and Liquids," Rev. Sci. Instrum., **54**, pp. 744–747.
[5] Zhong, Q. Y., Favro, L. D., and Thomas, R. L., 2000, "Thermal Wave Reflections of a Pulsed Stripe Heat Source From a Plane Boundary," J. Appl. Phys., **87**, pp. 3999–4004.
[6] Haji-Sheikh, A., Hong, Y. S., You, S. M., and Beck, J. V., 1998, "Sensitivity Analysis for Thermophysical Property Measurements Using the Periodic Method," ASME J. Heat Transfer, **120**, pp. 568–576.
[7] McGahan, W. A., and Cole, K. D., 1992, "Solution of the Heat Conduction Equation in Multilayers for Photothermal Deflection Experiments," J. Appl. Phys., **72**, pp. 1362–1373.
[8] Hu, H., Wang, X., and Xu, X., 1999, "Generalized Theory of the Photoacoustic Effect in a Layered Material," J. Appl. Phys., **86**, pp. 3953–3958.
[9] Aviles-Ramos, C., Haji-Sheikh, A., Beck, J. V., and Dowding, K. J., 2001, "Estimation of Thermophysical Properties by the Spectral Method–Development and Evaluation," ASME J. Heat Transfer, **123**, pp. 24–30.
[10] Cahill, D. G., 1990, "Thermal Conductivity Measurement From 30 to 750 K: the 3ω Method," Rev. Sci. Instrum., **61**, pp. 802–808.
[11] Kim, J. H., Feldman, A., and Novotny, D., 1999, "Application of the Three Omega Thermal Conductivity Measurement Method to a Film on a Substrate of Finite Thickness," J. Appl. Phys., **86**, pp. 3959–3963.
[12] Borca-Tasciuc, T., Kumar, A. R., and Chen, G., 2001, "Data Reduction in 3ω Method for Thin-Film Thermal Conductivity Determination," Rev. Sci. Instrum., **72**, pp. 2139–2147.
[13] Cole, K. D., 2004, "Analysis of Photothermal Characterization of Layered Materials–Design of Optimal Experiments," Int. J. Thermophys., **25**, pp. 1567–1584.
[14] Cole, K. D., and McGahan, W. A., 1993, "Theory of Multilayers Heated by Laser Absorption," ASME J. Heat Transfer, **115**, pp. 767–771.
[15] Beck, J. V., Cole, K. D., Haji-Sheikh, A., and Litkouhi, B., 1992, *Heat Conduction Using Green's Functions*, Hemisphere, New York, Chap. 2, pp. 40–43.
[16] Duffy, D. G., 2001, *Green's Functions With Applications*, Chapman and Hall/CRC, Boca Raton, FL.
[17] Sheremet, V. D., 2002, *Handbook of Green's Functions and Matrices*, WIT, Southhampton, UK.
[18] Mandelis, A., 2001, *Diffusion-Wave Fields, Mathematical Methods and Green's Functions*, Springer, New York, pp. 2, 3, and 231.
[19] Cole, K. D., Green's Function Library, www.greensfunction.unl.edu
[20] Haji-Shiekh, A., and Beck, J. V., 2002, "Temperature Solution in Multi-Dimensional Multi-Layer Bodies," Int. J. Heat Mass Transfer, **45**, pp. 1865–1877.
[21] Stakgold, I., 1967, *Boundary Value Problems of Mathematical Physics*, Mac-Millian, New York, Chap. 1.
[22] Crittenden, P. E., and Cole, K. D., 2002, "Fast-Converging Steady-State Heat Conduction in a Rectangular Parallelepiped," Int. J. Heat Mass Transfer, **45**, pp. 3585–3596.
[23] Cole, K. D., and Yen, D. H. Y., 2001, "Green's Functions, Temperature, and Heat Flux in the Rectangle," Int. J. Heat Mass Transfer, **44**, pp. 3883–3894.
[24] Cole, K. D., and Yen, D. H. Y., 2001, "Influence Functions for the Infinite and Semi-Infinite Strip," J. Thermophys. Heat Transfer, **15**, pp. 431–438.
[25] Roach, P. J., 1998, "Verification and Validation in Computational Science and Engineering," *Hermosa*, Albuquerque, NM, p. 23.
[26] Beck, J. V., and Arnold, K. J., 1977, *Parameter Estimation in Engineering and Science*, Wiley, New York.

Natural Convection in a Cavity With a Wavy Wall Heated From Below and Uniformly Cooled From the Top and Both Sides

Amaresh Dalal

Research Scholar
Department of Mechanical Engineering,
Indian Institute of Technology Kanpur,
Kanpur-208 016, India
e-mail: amaresh@iitk.ac.in

Manab Kumar Das[1]

Associate Professor
Department of Mechanical Engineering,
Indian Institute of Technology Guwahati,
Guwahati-781 039, India
e-mail: manab@iitg.ernet.in

In this paper, natural convection inside a two-dimensional cavity with a wavy right vertical wall has been carried out. The bottom wall is heated by a spatially varying temperature and other three walls are kept at constant lower temperature. The integral forms of the governing equations are solved numerically using finite-volume method in the non-orthogonal body-fitted coordinate system. The semi-implicit method for pressure linked equation algorithm with higher-order upwinding scheme are used. The streamlines and isothermal lines are presented for three different undulations (1, 2 and 3) with different Rayleigh number and a fluid having Prandtl number 0.71. Results are presented in the form of local and average Nusselt number distribution for a selected range of Rayleigh number $(10^0 – 10^6)$. [DOI: 10.1115/1.2194044]

Keywords: natural convection, numerical simulation, laminar flow, complex geometry

1 Introduction

In case of natural convection heat transfer, wavy enclosure is one of the several methods employed to enhance heat and mass transfer efficiency. Heat transfer inside a cavity is relevant to large scale natural phenomena in the fields of astrophysics, geophysics, atmospheric sciences, and a wide range of engineering applications such as cooling of electronic equipment, solidification processes, growing crystals, and solar collectors. Due to surface waviness, there is a change in the flow field. This is observed in a complex corrugated-duct geometry and is absent in conventional ducts such as rectangular, circular, or annular ducts. The parameters affecting their design are amplitudes, wavelength, phase angle, number of waves, etc. The hydrodynamics and thermal behaviors of fluid inside the cavity are affected by these parameters. The effect of these configurations in the flow phenomena is the motivation for carrying out numerical study of this topic.

Yao [1] has studied theoretically the natural convection along a vertical wavy surface. He found that the local heat transfer rate is smaller than that of the flat plate case and decreases with increase of the wave amplitude. The average Nusselt number also shows the same trend. Adjlout et al. [2] studied the effect of wavy hot wall in an enclosure where other three walls were straight. One of their findings was the decrease of heat transfer with the surface waviness when compared with flat wall cavity. Mahmud et al. [3] studied flow and heat transfer characteristics inside an isothermal vertical wavy-walled enclosure bounded by two adiabatic straight walls at different Grashof number and orientations for some selected waviness of the surface. Das and Mahmud [4] investigated buoyancy induced flow and heat transfer inside a wavy enclosure. They reported that the amplitude-wavelength ratio affected local heat transfer rate, but it had no significant influence on average
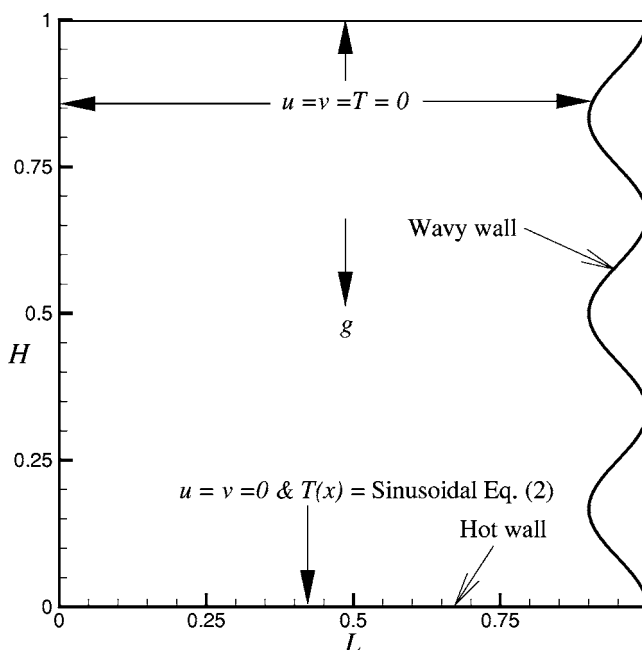

Fig. 1   Flow configuration and boundary condition for three undulations case

---

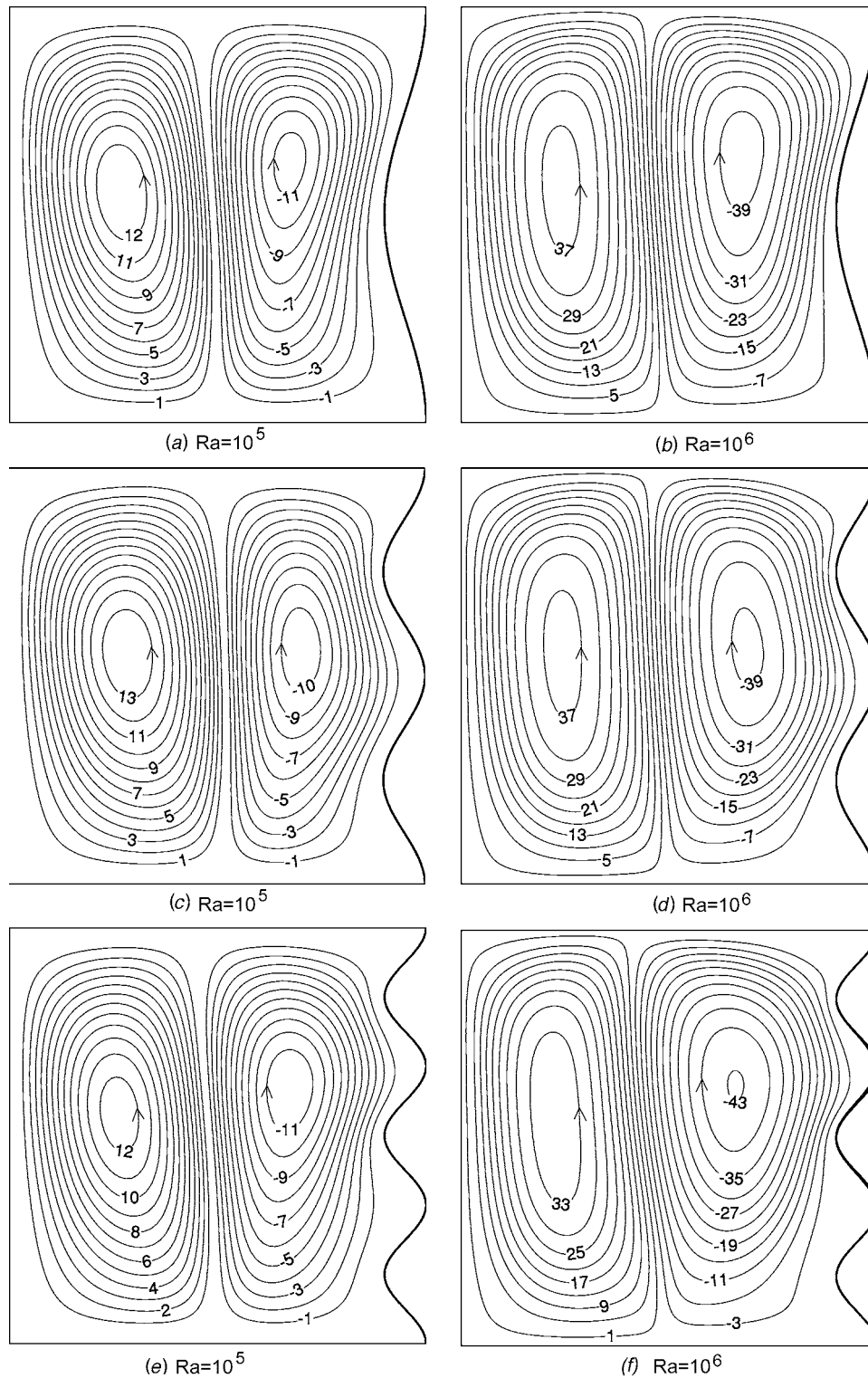**Fig. 2 Streamline contour plot for λ=0.05**

heat transfer rate. Dalal and Das [5] have considered a case of heating from the top surface with a sinusoidally varying temperature and cooling from the other three surfaces. The right vertical surface was undulated having one and three numbers. The effect of the number and the amplitude of undulation was studied. In another study, Dalal and Das [6] have made a detailed study by considering the same geometry as of [5]. The study was conducted at different inclination of the enclosure from 0 to 360 deg in steps of 30 deg. They concluded that the maximum and minimum av-

erage Nu occurs at certain orientation angles. Rathish Kumar et al. [7] have reported the effect of sinusoidal surface imperfections on the free convection in a porous enclosure heated from the side. The observations reveal that the heat transfer decreases as the amplitude of the wave increases. Also, the total heat transfer rate is less when compared with the heat transfer in an enclosure with plane walls. Rathish Kumar and Gupta [8] have analyzed the combined influence of mass and thermal stratification on non-Darcian double-diffusive natural convection from a wavy vertical wall to a
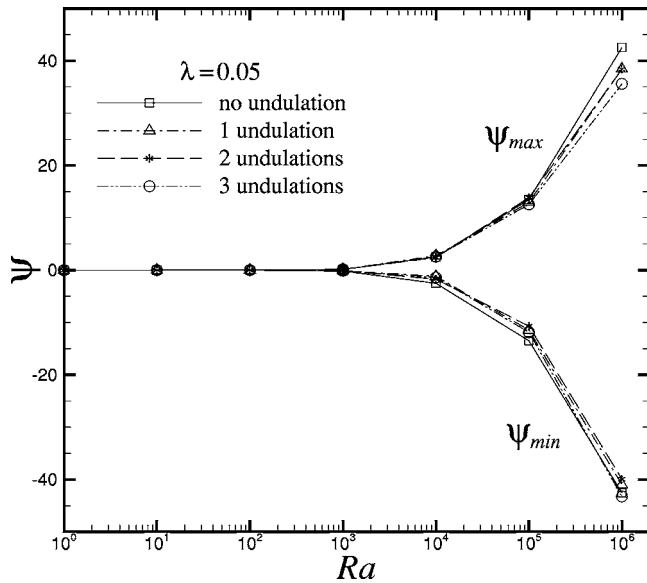
Fig. 3 Maximum and minimum stream function distribution for $\lambda = 0.05$

porous media. Extensive numerical simulations are carried out to analyze the influence of various parameters. It is observed that the presence of surface waviness brings in a wavy pattern in the local heat fluxes.

In the present investigation, a numerical analysis of natural convection in a two-dimensional cavity heated from below and uniformly cooled from the top and both sides is conducted. The cavity is having three flat walls and the right vertical wall consisting of one, two, and three undulations. The amplitude of undulations is varied from 0.00 to 0.10. The two vertical and top walls are cooled with a fixed temperature, whereas the bottom wall is heated with a sinusoidal temperature distribution in space coordinate. Air has been taken as the working fluid with Pr=0.71. The flow structure type and heat transfer rate are analyzed and discussed for a wide range of Rayleigh number (10^0 to 10^6) in this study.

2 Problem Specification

Figure 1 shows the geometry of the two-dimensional square cavity with wavy right vertical wall filled with viscous fluid. The bottom wall temperature is considered to be spatially varying with sinusoidal temperature distribution, $T_h^*(x^*)$. The other three walls are considered to be of constant temperature, T_c^*. The temperature distribution on the bottom wall is as follows

$$T_h^*(x^*) = T_c^* + \frac{\Delta T^*}{2}\left[1 - \cos\left(\frac{2\pi x^*}{L}\right)\right] \qquad (1)$$

where T_c^* is the minimum value of the imposed temperature distribution, ΔT^* is the temperature difference between the maximum and the minimum temperatures of the bottom wall, and L is the length of the enclosure. The above equation can be written in the dimensionless form as follows

$$T_h(x) = \frac{1}{2}[1 - \cos(2\pi x)] \qquad (2)$$

The right vertical wall is taken as sinusoidally varying. The expression of the wavy wall is given by

$$f(y) = 1 - \lambda + \lambda \times \cos(2\pi n y) \qquad (3)$$

where n is the number of undulations [2]. Three different cases with one, two and three undulations are studied. The wave amplitude (λ) for three cases is varied from 0.00 to 0.10. The fluid

considered in this study is air (Pr=0.71). The Rayleigh number is varied from 10^0 to 10^6.

3 Governing Equations and Boundary Conditions

Natural convection is governed by the differential equations expressing the conservation of mass, momentum, and energy. The present flow is considered steady, laminar, incompressible and two dimensional. The viscous dissipation term in the energy equation is neglected. The variation of fluid properties with temperature has been neglected, with the only exception of the buoyancy term, for which the Boussinesq approximation has been adopted. The governing equations and the boundary conditions are cast in dimensionless form using the following dimensionless variables

$$x = \frac{x^*}{L} \quad y = \frac{y^*}{L} \quad u = \frac{u^*L}{\alpha} \quad v = \frac{v^*L}{\alpha} \quad p = \frac{p^*L^2}{\rho\alpha^2} \quad T = \frac{T^* - T_c^*}{\Delta T^*}$$

$$(4)$$

The resulting equations are
Continuity equation:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \qquad (5)$$

u-momentum equation:

$$\frac{\partial(u^2)}{\partial x} + \frac{\partial(uv)}{\partial y} = -\frac{\partial p}{\partial x} + \text{Pr}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \qquad (6)$$

v-momentum equation:

$$\frac{\partial(uv)}{\partial x} + \frac{\partial(v^2)}{\partial y} = -\frac{\partial p}{\partial y} + \text{Pr}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) + \text{Ra Pr } T \qquad (7)$$

Energy equation:

$$\frac{\partial(uT)}{\partial x} + \frac{\partial(vT)}{\partial y} = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \qquad (8)$$

In addition, the velocity and temperature boundary conditions take the following form

$$\begin{array}{ll} u = v = T = 0 & \text{for } x = 0, 1 \text{ and } 0 \leqslant y \leqslant 1 \\ u = v = T = 0 & \text{for } y = 1 \text{ and } 0 \leqslant x \leqslant 1 \\ u = v = 0 \text{ and } T = \frac{1}{2}[1 - \cos(2\pi x)] & \text{for } y = 0 \text{ and } 0 \leqslant x \leqslant 1 \end{array}$$

$$(9)$$

4 Transformation of the Governing Equations

The governing equations transformed from the Cartesian system (x, y) to the boundary-fitted coordinate system (ξ, η) [9,10] are given by
Continuity equation:

$$U_\xi + V_\eta = 0 \qquad (10)$$

Generalized momentum and energy equations

$$(U\phi)_\xi + (V\phi)_\eta = S(\xi, \eta) + \left\{\frac{\Gamma}{J}(q_1\phi_\xi - q_2\phi_\eta)\right\}_\xi + \left\{\frac{\Gamma}{J}(-q_2\phi_\xi + q_3\phi_\eta)\right\}_\eta \qquad (11)$$

where $\Gamma = \text{Pr}$ for the momentum equation and $\Gamma = 1$ for the energy equation. The source term $S(\xi, \eta)$ is given by
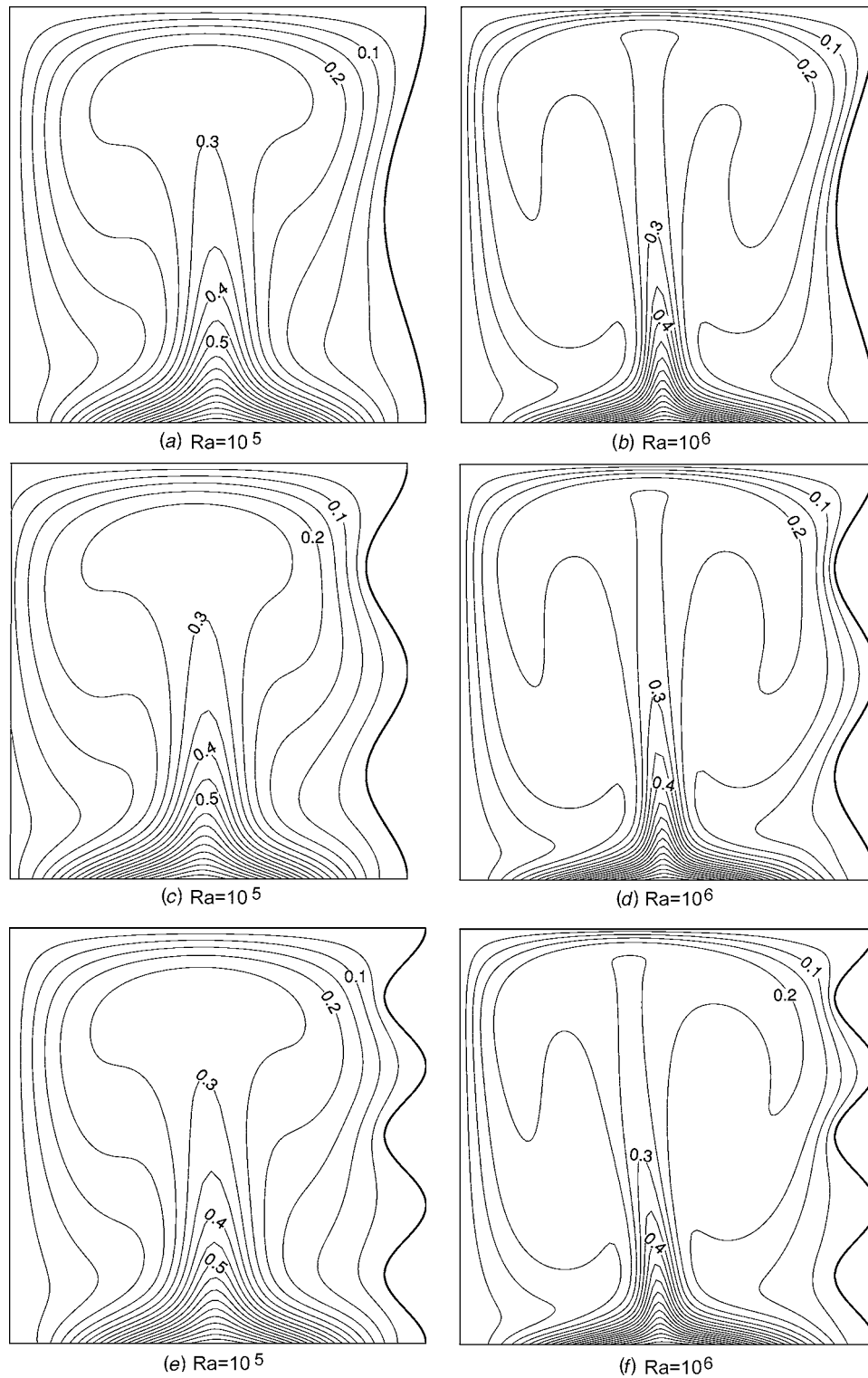
Fig. 4 Isotherm contour plot for $\lambda = 0.05$

$$S(\xi, \eta) = -y_\eta p_\xi + y_\xi p_\eta \qquad \text{for } \phi = u$$
$$S(\xi, \eta) = x_\eta p_\xi - x_\xi p_\eta + J \, \text{Ra} \, \text{Pr} \, T \quad \text{for } \phi = v \qquad (12)$$
$$S(\xi, \eta) = 0 \qquad \text{for } \phi = T$$

The relationships between the Cartesian and contravariant velocity components are

$$U = y_\eta u - x_\eta v \quad V = x_\xi v - y_\xi u \qquad (13)$$

The boundary condition given in Eq. (9) are Dirichlet type. The boundary condition for computational plane can be written as follows
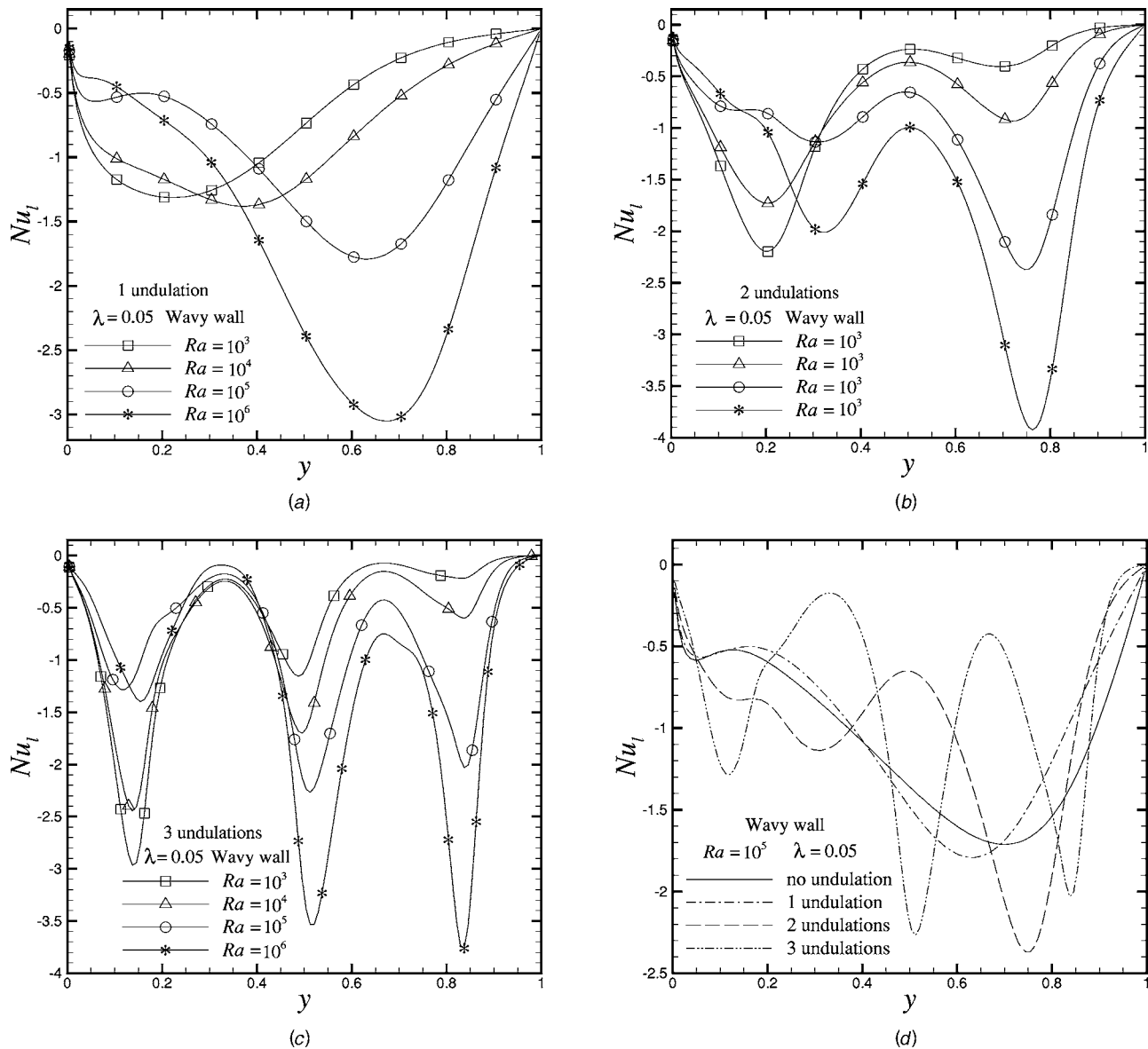
Fig. 5 Local Nusselt number distribution on the wavy wall

$$u = v = T = 0 \qquad \text{for } \xi = 0, 1 \text{ and } 0 \leq \eta \leq 1$$

$$u = v = T = 0 \qquad \text{for } \eta = 1 \text{ and } 0 \leq \xi \leq 1$$

$$u = v = 0 \text{ and } T = \frac{1}{2}(1 - \cos(2\pi x)) \quad \text{for } \eta = 0 \text{ and } 0 \leq \xi \leq 1$$

$$\text{(14)}$$

The heat transfer rate by convection in an enclosure is obtained from the Nusselt number calculation. The local Nusselt numbers in different walls are expressed as

Top wall $\quad \mathrm{Nu}_l = \dfrac{1}{J\sqrt{q_3}}(q_3 T_\eta - q_2 T_\xi)$

Right wall $\quad \mathrm{Nu}_l = \dfrac{1}{J\sqrt{q_1}}(q_1 T_\xi - q_2 T_\eta)$

$$\text{(15)}$$

Bottom wall $\quad \mathrm{Nu}_l = \dfrac{-1}{J\sqrt{q_3}}(q_3 T_\eta - q_2 T_\xi)$

Table 1 Comparison of Nu$_{\max}$ on wavy wall for $\lambda = 0.05$

	Ra=10^3	Ra=10^4	Ra=10^5	Ra=10^6
No undulation	−1.232699	−0.878886	−1.712428	−3.010422
1 Undulation	−1.313549(6.56%)	−1.381887(57.23%)	−1.792783(4.69%)	−3.051431(1.36%)
2 Undulations	−2.196742(78.21%)	−1.730059(96.85%)	−2.370553(38.43%)	−3.924854(30.38%)
3 Undulations	−2.965302(140.55%)	−2.445720(178.27%)	−2.266665(32.37%)	−3.761287(24.94%)

λ	1 undulation	2 undulations	3 undulations
0.00	−1.712428	−1.712428	−1.712428
0.02	−1.748054(2.08%)	−2.027770(18.41%)	−1.944703(13.56%)
0.04	−1.767352(3.21%)	−2.276543(32.94%)	−2.156845(25.95%)
0.06	−1.781482(4.03%)	−2.389627(39.55%)	−2.353556(37.44%)
0.08	−1.787428(4.38%)	−2.357953(37.70%)	−2.715110(58.55%)
0.10	−1.967826(14.91%)	−2.255406(31.71%)	−2.961491(72.94%)

Table 2   Comparison of Nu_{max} on wavy wall for Ra=10^5

λ	1 undulation	2 undulations	3 undulations
0.00	−1.051330	−1.051330	−1.051330
0.02	−1.032238(−1.82%)	−1.069038(1.68%)	−1.028577(−2.16%)
0.04	−0.991104(−5.73%)	−1.028949(−2.13%)	−0.896421(−14.74%)
0.06	−0.982906(−6.51%)	−0.964809(−8.23%)	−0.772838(−26.49%)
0.08	−0.923823(−12.13%)	−0.883091(−16.00%)	−0.694831(−33.91%)
0.10	−0.927771(−11.75%)	−0.851067(−19.05%)	−0.654445(−37.75%)

Table 3   Comparison of Nu_{av} on wavy wall for Ra=10^5

$$\text{Left wall} \quad \mathrm{Nu}_l = \frac{-1}{J\sqrt{q_1}}(q_1 T_\xi - q_2 T_\eta)$$

The average Nusselt number is calculated by the following expression

$$\mathrm{Nu}_{av} = \frac{1}{L}\int_0^L \mathrm{Nu}_l\, dl \qquad (16)$$

5   Numerical Procedure

Equations (10) and (11) are discretized using staggered, non-uniform control volumes and these equations are solved numerically by finite volume method. The semi-implicit method for pressure linked equation [11] is used to couple momentum and continuity equations. In order to minimize the numerical diffusion errors, the deferred quadratic upstream interpolation for convective kinetics (QUICK) scheme of Hayase et al. [12] is employed in approximating the convective terms for both the momentum equations and energy equation. The central difference scheme is employed near the boundary points for the convective terms. The tri-diagonal matrix algorithm [11] is applied for the line-by-line solution of the momentum, energy, and pressure correction equations. The pseudo-transient approach is followed for the numerical solution as it is useful for a situation in which governing equations give rise to stability problems, e.g., buoyant flows [13]. Under-relaxation factor for pressure with values of 0.01 is used and the pseudo time step is taken as 10^{-2} for Ra=10^0–10^2; 10^{-3} for Ra =10^3, 10^4; 10^{-4} for Ra=10^5; 10^{-5} for Ra=10^6.

The iterative procedure is initiated by the solution of energy equation followed by momentum equations and is continued until convergence is achieved. Euclidean norm of the residual is taken as convergence criteria for each dependent variable in the entire flow field [14]. The expression for this is

$$\|r_p\|^n \le \gamma_p\|r_p\|^0 \qquad (17)$$

where $\|r_p\|^0$ - initial Euclidean norm of residuals, $\|r_p\|^n$ - Euclidean norm after n iterations, γ_p is the residual reduction factor. The expression for mass residual R_{mass} is given by

$$R_{\mathrm{mass}} = \sum_{\text{all cells}} |(u_e - u_w)\Delta y + (u_n - u_s)\Delta x| \qquad (18)$$

The mass residual for global convergence was taken as 10^{-6}.

6   Code Validation

The present code is validated with the numerical results of de Vahl Davis [15], Markatos and Perikleous [16] and Hadjiso-phocleous et al. [17] for the buoyancy driven laminar heat transfer in a square cavity with differentially heated sidewalls. The left wall is maintained hot while the right wall is cooled. The top and bottom walls are insulated. In the present work, numerical predictions using the developed algorithm have been obtained for Rayleigh numbers between 10^3 and 10^6 on elliptic mesh with 61 ×61 grid points. Comparison of the average Nusselt number on hot wall and comparison of maximum and minimum Nusselt number on hot wall have been done. The results are in very good agreement with the benchmark solution.

7   Grid Independence Study of the Problem Concerned

The grid independence test is performed using successively sized grids, 21×21, 41×41, 61×61, 81×81, 101×101, 121×121, and 131×131 for Ra=10^4, 10^6, and λ=0.05 for one, two, and three undulations cases. It is observed that there is less change in the average Nusselt number on the wavy wall from 121×121 to 131×131 than 101×101 and 121×121. So a grid number of 121×121 is chosen for further computation.

8   Results and Discussion

A parametric study has been carried out to determine the influence of Rayleigh number on the flow field and effect of number of undulation on heat transfer. The results are for Rayleigh number of 10^0 to 10^6, Prandtl number of 0.71, and undulation amplitude of 0.00 to 0.10. The discussion of the following results concerns the streamlines, isotherms, local Nusselt number, and average Nusselt number distributions on the walls.

8.1   Flow and Thermal Field. Figure 2 shows the streamlines obtained for Rayleigh numbers 10^5 and 10^6 with amplitude λ =0.05. There is not much variation in flow patterns for Ra=10^0 –10^3 because conduction is the mode of heat transfer. A pair of counterrotating cells is formed in left and right halves of the enclosure. The fluid getting heated from the bottom wall moves up near the vertical midplane of the enclosure. Then the fluid impinges near the middle of top wall and it moves horizontally toward corners losing heat to the top wall. Finally, it descends along the cold sidewalls. The right cell is squeezed and small compared to left cell due to the presence of undulation; it has much less effect on the size of the left cell. The intensity of the recirculation pattern increases with increase of Rayleigh number because convection becomes the dominant mode of heat transfer. The maximum and minimum stream function distribution for λ=0.05 is shown in Fig. 3. It is noticed that up to Ra=10^3, the stream function value is almost zero indicating that the convection effect is much less. As Ra is increased, $\psi_{\max}$ and $\psi_{\min}$ plot shows that the

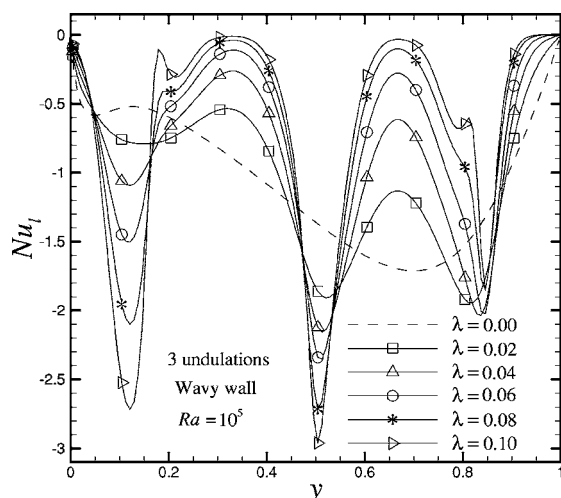Table 4   Comparison of Nu_{av} on wavy wall for λ=0.05

	Ra=10^3	Ra=10^4	Ra=10^5	Ra=10^6
No undulation	−0.584669	−0.673252	−1.051330	−1.776514
1 Undulation	−0.669581(14.52%)	−0.815491(21.13%)	−0.989409(−5.90%)	−1.594566(−10.24%)
2 Undulations	−0.661525(13.14%)	−0.736676(9.42%)	−0.999318(−4.95%)	−1.510676(−14.96%)
3 Undulations	−0.589651(0.85%)	−0.676268(0.45%)	−0.832816(−20.78%)	−1.132782(−36.24%)
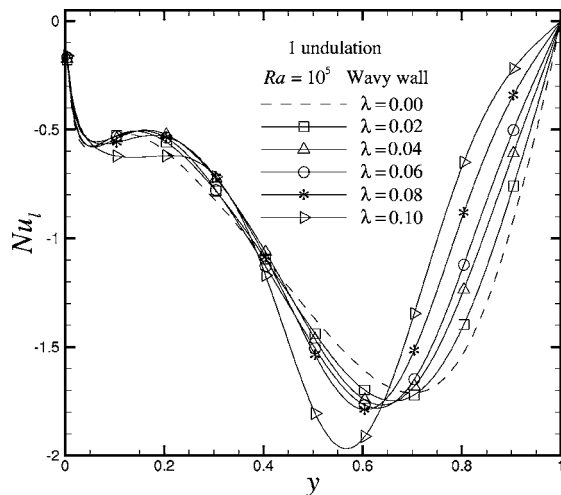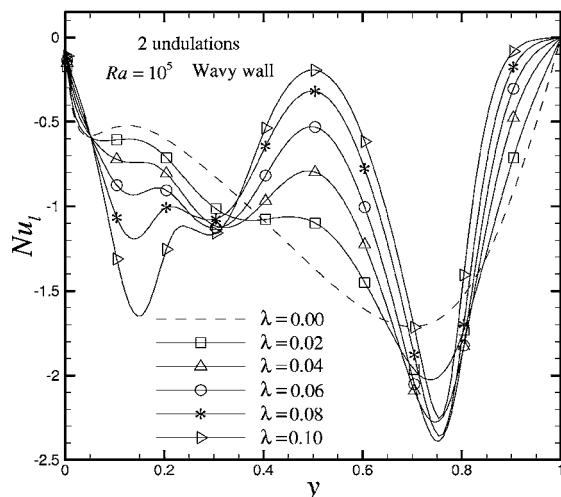
Fig. 6 Local Nusselt number distribution on the wavy wall for different amplitudes

Figure 4 shows the isotherms for Rayleigh numbers 10^5 and 10^6. For lower Rayleigh number (10^3 and 10^4) the isothermal lines are uniformly distributed inside the cavity, as conduction is the main heat transfer mechanism in this case. The isothermal lines swirl at $Ra=10^5$ and $Ra=10^6$ due to the influence of increased convection current. The growth of thermal boundary layer on the bottom heated surface is to be noticed. Also there is an increased penetration of the hot fluid towards the upper cold surface of the enclosure.

8.2 Heat Transfer Distribution. The effect of different parameters (i.e., Rayleigh number, undulation and wave amplitude) on heat transfer is described from Nusselt number distribution.

8.2.1 Local Heat Transfer. Figure 5 shows the variation of local Nusselt number along wavy wall for different Rayleigh number and $\lambda=0.05$. In the parentheses of Tables 1–4, percentage increase in Nusselt number (Nu_{max} or Nu_{av}) with respect to second row (no undulation or $\lambda=0.00$) is shown for comparison. For low Ra, the maximum Nu_l is located near the bottom heated surface. As Ra is increased, the location gradually rises up because of the increasing convection strength. An inspection of Table 1 reveals that Nu_{max} decreases as Ra is increased from 10^3 to 10^4 for no undulation case. The same is repeated for two-undulations case. The reduction in Nu_{max} continues up to $Ra=10^5$ for three undulations case. This nature is due to the combination of undulations and convection strength. The variation of local Nusselt number along the wavy wall for different wave amplitude and $Ra=10^5$ is shown in Fig. 6. From no-undulation case to $\lambda=0.08$ for one-undulation case, it is observed that there is not much variation in the Nu_{max} value. However, for $\lambda=0.1$ case, there is an appreciable change equal to 15% (Table 2). For two-undulations case, however, the Nu_{max} has an appreciable change for no-undulation case for the range of λ considered (Table 2). For three-undulations case, it increases even further (73%, Table 2).

8.2.2 Overall Heat Transfer. Figure 7 shows the average Nusselt number distribution on wavy wall for different wave amplitude and Rayleigh number. It is observed that in the conduction region (i.e., $Ra \leqslant 10^3$), Nu_{av} increases with increase in λ (Figs. 7(a)–7(c), Table 3). However, as the convection mode of heat transfer becomes dominant with increase in Ra, there is a decrement of Nu_{av} compared to that of no-undulation case. Figure 7(d) shows that for $Ra=10^6$, no-undulation case has maximum Nu_{av} and it decreases with number of undulation. From Table 4, it is seen that Nu_{av} is less by 36% for a three-undulation case compared to no-undulation case.

9 Concluding Remarks

Buoyancy-induced flow and heat transfer inside a cavity with sinusoidal temperature boundary condition on the bottom wall and constant cold temperature boundary condition on other three walls are investigated numerically. For small Ra, the heat transfer is dominated by conduction across the fluid layers. With increase in Ra, the process begins to be dominated by convection. The presence of undulation in the right wall affects both local heat transfer rate and flow field as well as thermal field. The heat rejection from the fluid to wavy wall increases up to $Ra=10^4$ for one-undulated cavity compared to square cavity without undulation and then it decreases. But the reverse scenario is observed for the left wall. With the increase of amplitude, the average Nusselt number on the wavy wall is reduced. Up to $Ra=10^4$, the heat transfer increases by undulating the wall. However, it decreases as the number of undulations is increasing. As the Ra is increased, undulations on the wall reduce the heat transfer.

convection strength is increased. However, the presence of undulation does not have any significant effect on the stream function values.
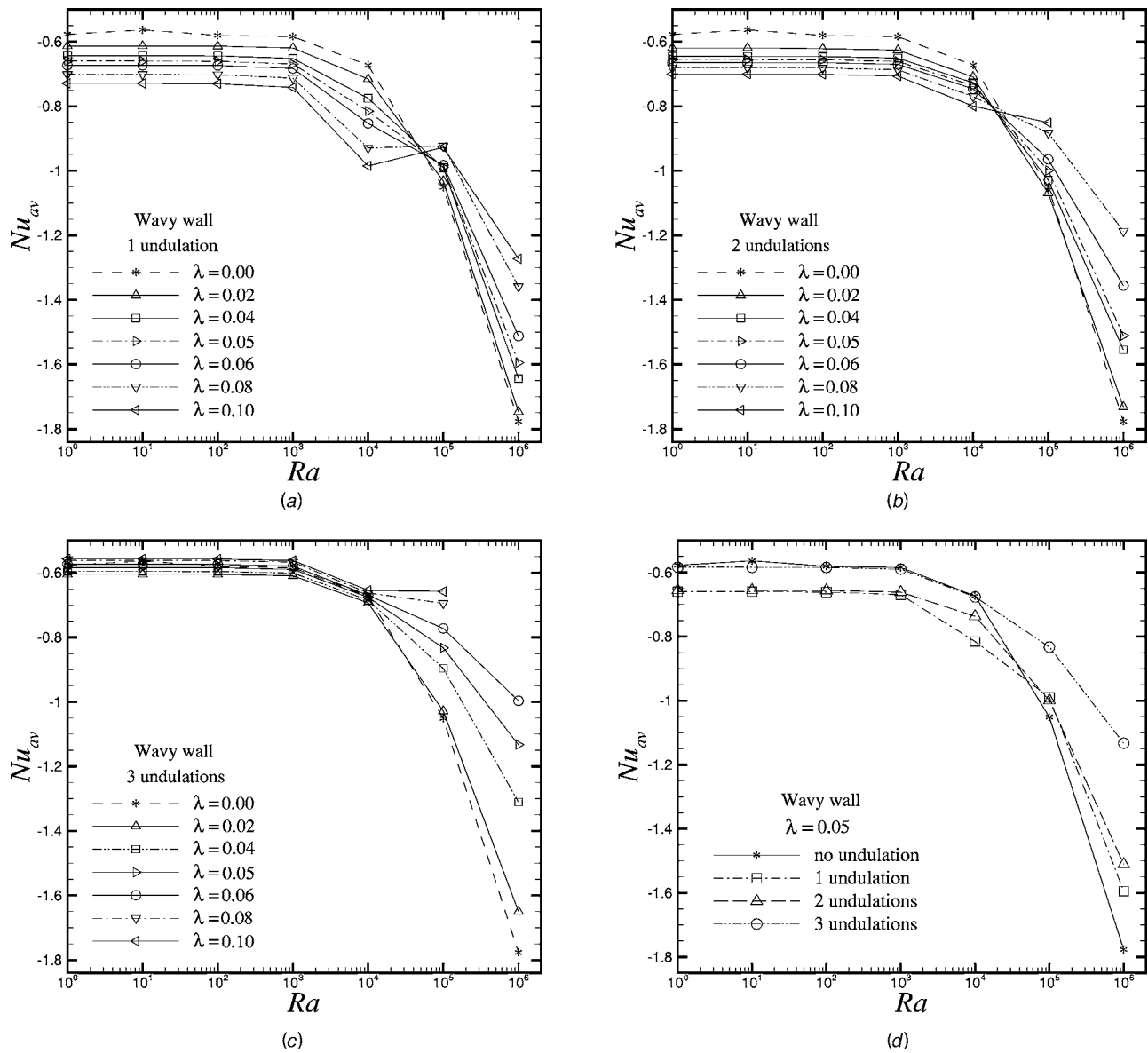
Fig. 7 Average Nusselt number distribution on the wavy wall for various Ra and amplitude

the computation facilities extended by the CFD Laboratory of the Mechanical Engineering Department of IIT Kanpur.

Nomenclature

g = gravitational acceleration
H = height of the enclosure
J = Jacobian
L = length of the enclosure
n = number of undulation
Nu = Nusselt number
p = dimensionless pressure
Pr = Prandtl number $(=\nu/\alpha)$
q_1, q_2, q_3 = geometric relations between coordinate systems
Ra = Rayleigh number $(=g\beta\Delta T L^3/\alpha\nu)$
S = source term
T = dimensionless temperature
ΔT = differential temperature, dimensionless
U, V = dimensionless contravariant velocity components in ξ and η direction
u, v

= dimensionless velocity components in x and y direction
x, y = dimensionless Cartesian coordinates

Greek Symbols

α = thermal diffusivity
ξ, η = dimensionless curvilinear coordinates
λ = wave amplitude
ν = momentum diffusivity
ϕ = general variable representing u, v and T
ψ = stream function

Subscripts

av = average
c = cold wall
h = hot wall
l = local
max = maximum
x, y, ξ, η = derivative relative to x, y, ξ, η, respectively

Superscript

$*$ = dimensional form

References

[1] Yao, L. S., 1983, "Natural Convection Along a Vertical Wavy Surface," ASME J. Heat Transfer, **105**, pp. 465–468.

[2] Adjlout, L., Imine, O., Azzi, A., and Belkadi, M., 2002, "Laminar Natural Convection in an Inclined Cavity With a Wavy Wall," Int. J. Heat Mass Transfer, **45**, pp. 2141–2152.

[3] Mahmud, S., Das, P. K., Hyder, N., and Islam, A. K. M. S., 2002, "Free Convection in an Enclosure With Vertical Wavy Walls," Int. J. Therm. Sci., **41**, pp. 440–446.

[4] Das, P. K., and Mahmud, S., 2003, "Numerical Investigation of Natural Convection Inside a Wavy Enclosure," Int. J. Therm. Sci., **42**, pp. 397–406.

[5] Dalal, A., and Das, M. K., 2003, "Laminar Natural Convection in a Complicated Cavity With Spatially Variable Upper Wall Temperature, *Proceedings of 2003 ASME Summer Heat Transfer Conferences*, Las Vegas, July 21–23.

[6] Dalal, A., and Das, M. K., 2005, "Laminar Natural Convection in an Inclined Complicated Cavity With Spatially Variable Wall Temperature," Int. J. Heat Mass Transfer, **48**, pp. 2986–3007.

[7] Rathish Kumar, B. V., Singh, P., and Murthy, P. V. S. N., 1997, "Effect of Surface Undulations on Natural Convection in a Porous Square Cavity," ASME J. Heat Transfer, **119**, pp. 848–851.

[8] Rathish Kumar, B. V., and Gupta, S., 2005, "Combined Influence of Mass and Thermal Stratification on Double-Diffusion Non-Darcian Natural Convection From a Wavy Vertical Wall to Porous Media," ASME J. Heat Transfer, **127**, pp. 637–647.

[9] Maliska, C. R., and Raithby, G. D., 1984, "A Method for Computing Three Dimensional Flows Using Non-Orthogonal Boundary-Fitted Co-Ordinates," Int. J. Numer. Methods Fluids, **4**, pp. 519–537.

[10] Özisik, M. N., 1994, *Finite Difference Methods in Heat Transfer*, CRC, London.

[11] Patankar, S. V., 1980, *Numerical Heat Transfer and Fluid Flow*, Hemisphere, New York.

[12] Hayase, T., Humphrey, J. C., and Greif, R., 1992, "A Consistently Formulated QUICK Scheme for Fast and Stable Convergence Using Finite-Volume Iterative Calculation Procedures," J. Comput. Phys., **98**, pp. 108–118.

[13] Versteeg, H. K., and Malalasekera, W., 1995, *An Introduction to Computational Fluid Dynamics, The Finite Volume Method*, Longman, Malaysia.

[14] Van Doormall, J. P., and Raithby, G. D., 1984, "Enhancements of the SIMPLE Method for Predicting Incompressible Fluid Flows," Numer. Heat Transfer, **7**, pp. 147–163.

[15] de Vahl Davis, G., 1983, "Natural Convection of Air in a Square Cavity: A Benchmark Numerical Solution," Int. J. Numer. Methods Fluids, **3**, pp. 249–264.

[16] Markatos, N. C., and Perikleous, K. A., 1984, "Laminar and Turbulent Natural Convection in an Enclosed Cavity," Int. J. Heat Mass Transfer, **27**(5), pp. 755–772.

[17] Hadjisophocleous, G. V., Sousa, A. C. M., and Venart, J. E. S., 1988, "Prediction of Transient Natural Convection in Enclosures of Arbitrary Geometry Using a Nonorthogonal Numerical Model," Numer. Heat Transfer, **13**, pp. 373–392.

Critical Heat Flux of Steady Boiling for Water Jet Impingement in Flat Stagnation Zone on Superhydrophilic Surface

Zhenhua Liu

e-mail: liuzhenh@sjtu.edu.cn

Yuhao Qiu

School of Mechanical and Power Engineering,
Shanghai Jiaotong University,
Shanghai, People's Republic of China 200030

The nucleate boiling heat transfer characteristics of a round water jet impingement in a flat stagnation zone on the superhydrophilic surface were experimentally investigated. The superhydrophilic heat transfer surface was formed by a TiO₂ *coating process. The experimental results were compared with those on the common metal surface. In particular, the quantificational effects of the flow conditions, heating conditions, and the coating methods on the critical heat flux (CHF) were systemically investigated. The experimental data showed that the nucleate boiling heat transfer characteristics on the superhydrophilic surface are significantly different from those on the common metal surface. The CHF of boiling on the superhydrophilic surface is greatly increased by decreasing of the solid-liquid contact angle.*
[DOI: 10.1115/1.2194045]

*Keywords: jet, boiling, CHF, superhydrophilic surface*

Introduction

Titanium dioxide (TiO₂), one of the photocatalysts, has recently come into the spotlight because of its very attractive nature. One of its marvelous natures is the superhydrophilicity [1–3]. When the surface coated with TiO₂ is irradiated by ultraviolet light, the contact angle for water decreases with time and finally reaches almost zero. This nature has various practical applications. By making use of the superhydrophilic heat transfer surface, we expect that heat transfer characteristics of liquid-vapor phase change phenomena like boiling can be enhanced.

The mechanism of superhydrophilicity is divided into three processes [1]. In the first process, when the surface of TiO₂ is irradiated by ultraviolet light, a pair of electron and hole is created by excitation

$$h\upsilon \rightarrow e^- + h^+$$
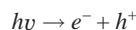
In case of usual photocatalytic reaction, this electron-hole pair reacts with oxygen or water that are adhering to the surface and then produce superoxide radical anions ($\cdot O_2^-$) and hydroxyl radical ($\cdot OH$). However, in case of photoinduced superhydrophilicity, TiO₂ crystal surface itself is reduced and oxygen vacancy is created

$$e^- + Ti^{4+} \rightarrow Ti^{3+}$$

$$4h^+ + 2O^{2-} \rightarrow O_2$$
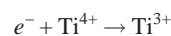
Ti³⁺ is immediately oxidized by oxygen in the air and, on the other hand, oxygen vacancy bonds water molecule in the air. Finally, hydroxyl group (OH) is created on the surface and this acts as a chemisorbed water layer.

The water jet cooling has been widely used in the iron or steel industry, nuclear power processes, the making of many microelectronic devices, and thermal management processes. In this study, the steady boiling experiments for water jet impingement in the flat stagnation zone were performed to reveal the effect of the wettability of the heat transfer surface on the boiling and the CHF.

For the boiling heat transfer of water jet impingement in the stagnation zone (the diameter of the heat transfer surface is the same or smaller than that of the jet nozzle), many steady boiling experiments using the common metal surfaces were carried out and the CHF data were included in these experimental results [4–13]. However, no reports concerning the effect of the solid-liquid contact angle between the heat transfer surface and liquid on the jet boiling were proposed.

Recently, Liu and Zhu [11] performed a theoretical study to predict the CHF of the saturated liquids jet impingement in the flat stagnation zone of the common metal surface and proposed a semi-theoretical and semi-empirical correlation as below

$$\frac{q_{c,0}}{Gh_{fg}} = 0.132\left(1 + \frac{\rho_v}{\rho_l}\right)^{1/3}\left(\frac{\sigma\rho_l}{G^2d}\right)^{1/3}\left(\frac{\rho_v}{\rho_l}\right)^{1.4/3} \quad (1)$$

For water at atmospheric pressure, since all physical properties of water are constants, Eq. (1) can be simplified as

$$q_{c,0} = 3.6 \times 10^5 \left(\frac{v}{d}\right)^{1/3} \quad (2)$$

where the units of $q_{c,0}$, v, d are W/m², m/s, and m, respectively.

For the CHF of the subcooled water jet impingement in the flat stagnation zone of the common metal surface, Liu and co-workers have proposed an empirical correlation as follows [12]

$$\frac{q_c}{q_{c,0}} = 1 + 15.18\frac{c_{p,l}\Delta T_{sub}}{h_{fg}} \quad (3)$$

Taking the previous studies into consideration, the present study objective was focused on the CHF of the saturated and sub-cooled water jet boiling in the stagnation zone on the superhydrophilic surface. The nucleate boiling heat transfer characteristics of a round water jet impingement in a flat stagnation zone on the superhydrophilic surface were experimentally investigated (see Table 1). Three influencing parameters, subcooling, impact velocity, and jet nozzle diameter, were changed and their effects on the CHF were systematically studied. The empirical correlations were obtained for predicting the boiling heat transfer and the CHF on the superhydrophilic surface.

Coating Process and Experimental Procedures

The superhydrophilic surface was made by the dipping method in the coating process. In the dipping method, TiO₂ colloid was used. The specimen was a copper block. Prior to dipping, the specimen was cleaned and dried. The heat transfer surface was polished to mirror finish and washed by hydrochloric acid and then by the acetone and purified water. When the surface became dry, the specimen was dipped into TiO₂ colloid then extracted slowly. After this dipping process, the specimen was heated by infrared lamps and maintained at about 150°C for 1 h. The thickness of the coated layer was measured by an interferometer. The thickness was about 1 μm. The specimen was ordinarily stored in a dark place to resist long term superhydrophilic characteristic. The TiO₂ coated surface would produce superhydrophilic characteristic after being exposed to ultraviolet light little by little. After 2 h, the superhydrophilic characteristic could remain stable.

The contact angle was measured by means of the sessile drop method in room temperature. The specimen was irradiated with ultraviolet lights of the peak wavelengths between 275 and

Table 1 Experimental conditions

Test fluid	Water
Jet Reynolds number	$2.5 \times 10^4 - 4 \times 10^5$
Heating condition	Steady state
Ultraviolet light	275–315 nm
Test pressure	Atmospheric pressure
Impact velocity (m/s)	0.5–8.0
Nozzle diameter (mm)	4, 8
Diameter of the heated surface (mm)	4, 8
Subcooling (K)	0–80

315 nm. For the superhydrophilic surface, the water drop would entirely expend on the superhydrophilic surface and covered the whole surface. Therefore, the accurate value of the contact angle cannot be obtained. However, it can be confirmed that the contact angle on the superhydrophilic surface is less than 5 deg. When there is no ultraviolet light, the contact angle on the TiO_2 coated surface is between 20 deg and 40 deg, while the contact angle on the copper surface is between 40 deg and 70 deg. Afterwards, the TiO_2 coated surface under the ultraviolet light is named as the superhydrophilic surface and the TiO_2 coated surface not under the ultraviolet light is named as the TiO_2 coated surface.

We have used different ultraviolet luminous intensity for irradiating the TiO_2 coated surface. It is found that the luminous intensity has no effect on the contact angle. On the other hand, the effects of the thickness and nonuniformity of the TiO_2 coated layer on the wettability also have not been found in the prior test.

In the present study, steady heat transfer experiments were carried out for the nucleate boiling regime to obtain the boiling curves and determine the CHF in the steady state. The experimental apparatus and the procedures are the same as that in the author's previous study [12], and they are no longer introduced.

Results and Discussion

Figure 1 shows the experimental results of the nucleate boiling regime for the saturated and the subcooled water jet impingement on the superhydrophilic surface for the nozzle diameter of 4 mm and the impact velocity of 1 m/s. The experimental data are plotted in the form of wall heat flux against wall superheat for the different subcoolings.

Equation (4) is an empirical correlation proposed by Wolf, Incropera and Viskanta for predicting the jet boiling heat transfer of water on a stainless steel surface [14]. The jet boiling experiment
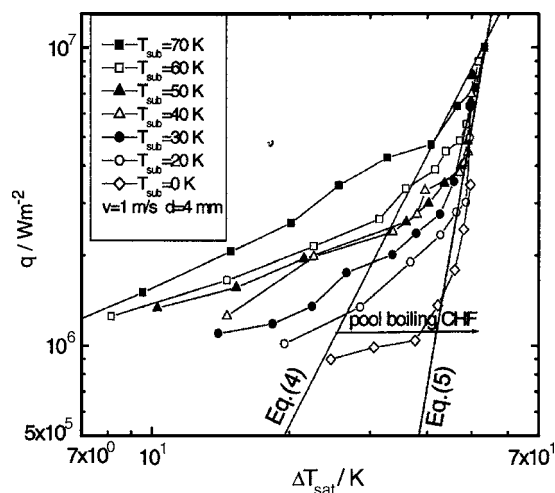

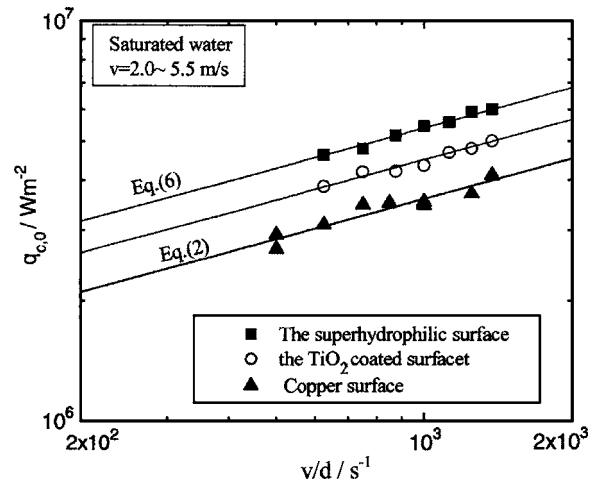**Fig. 1 Jet boiling curves on the superhydrophilic surface for Φ4 mm nozzle**

Fig. 2 Comparison of CHF between the superhydrophilic surface and copper surface for saturated water

of water on the copper surface was also carried out in the present study. The experimental data agreed reasonably well with Eq. (4). Therefore, in Fig. 1, the calculated values from Eq. (4) shown in the form of a solid line may denote the jet boiling heat transfer on the copper surface

$$q = 63.7 \Delta T_{sat}^{2.9} \; [\text{W/m}^2 \, \text{K}] \qquad (4)$$

It is found that the boiling incipience on the superhydrophilic surface is greatly delayed. The experimental range mainly belongs to the forced convective regime. In the fully developed nucleate boiling regime, the boiling curves are quite steep. To each subcooling, the superheat corresponding to the CHF is basically between 30 and 40 K. The boiling curves resulting from the different subcoolings essentially converge into a steep line. The boiling curves for the superhydrophilic surface are remarkably shifted to the rightward. The superhydrophilic surface plays an important role in the jet boiling heat transfer characteristics.

By using a least-square fit, the following empirical correlation can well predict the nucleate boiling heat transfer of water jet on a superhydrophilic surface. The vast majority of the data agree well with the predictions of Eq. (5) within the relative errors of ±6%

$$q = 2.8 \times 10^{-9} \times \Delta T_{sat}^9 \qquad (5)$$

where the units of q and ΔT_{sat} are W/m² and K, respectively.

According to Eq. (2), we have known that there exists a function relation between (v/d) and $q_{c,0}$ for the saturated water jet on common metal surface. Figure 2 shows the compared results of the CHF data of the saturated water jet boiling among the superhydrophilic surface, the TiO_2 coated surface, and the copper surface. No matter for the superhydrophilic surface, or for the TiO_2 coated surface, it is found that there still exists a good linear relation between the CHF and $(v/d)^{1/3}$. Therefore, the form of Eq. (2) is available for various surfaces.

For the saturated water jet on the superhydrophilic surface, the correlation factor in Eq. (2) was determined from the present CHF data by using a least square fit within the relative errors of ±5%. The equation is presented as follows

$$q_{c,0} = 5.4 \times 10^5 \left(\frac{v}{d} \right)^{1/3} \qquad (6)$$

here the units of $q_{c,0}$, v, and d are W/m², m/s, and m, respectively.

From the above equation, it is found that the CHF on the superhydrophilic surface is about 50% higher than that on common copper surface.
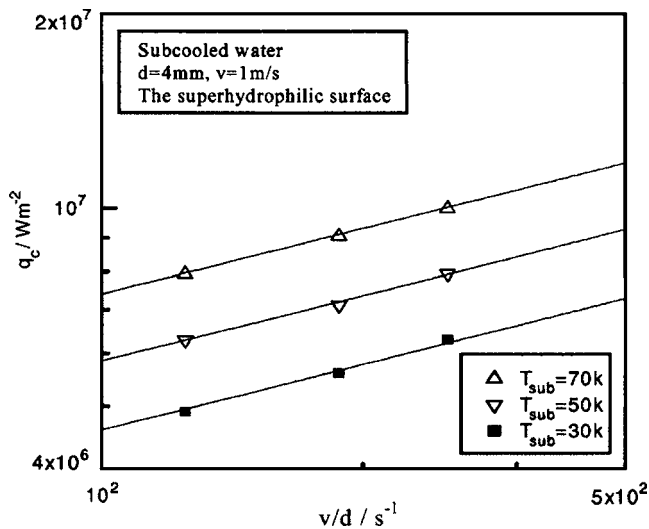
Fig. 3 Comparison of CHF between the superhydrophilic surface and copper surface for subcooled water

As shown in Eq. (6), the CHF of the saturated water is of $q_{c,0} \propto (v/d)^{1/3}$. If (v/d) could still be considered as a parameter that affects the CHF for the subcooled water, the correlation for predicting the CHF of the subcooled water would have a simple form. Figure 3 shows the relation between the CHF of the subcooled water and (v/d). It is found that all of the solid lines consisting of the data have a gradient of one third for every fixed subcooling. Therefore, there is the same relationship of $q_c \propto (v/d)^{1/3}$, for the subcooled water as for the saturated water. Meantime, it is found that the CHF of the subcooled water on the superhydrophilic surface is about 50% higher than that on common copper surface. Therefore, the effects of the superhydrophilic surface on the CHF are the same either for saturated water or for the subcooled water.

Figure 4 shows the comparison of the CHF data of the subcooled water jet on the superhydrophilic surface with Eq. (3). Here, the q_c and $q_{c,0}$, respectively, denote the CHFs of the subcooled water and the saturated water on the superhydrophilic surface. Equation (3) correlates the CHF data with the maximum
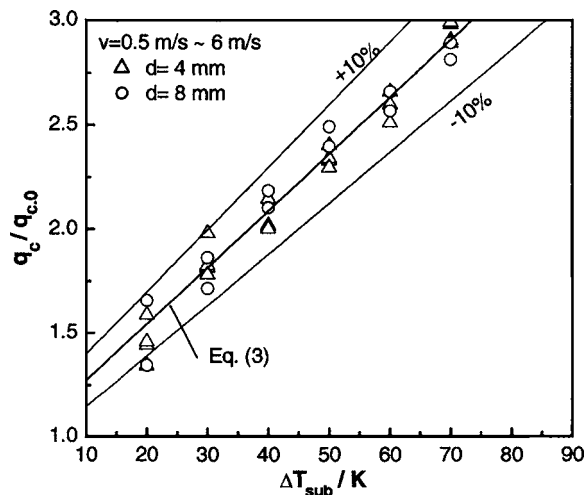

Fig. 4 Comparison of CHF data of the superhydrophilic surface with Eq. (3)

relative errors less than 10%, and can still be used for predicting the CHF of the subcooled water jet boiling on the superhydrophilic surface.

Conclusions

1. A superhydrophilic heat transfer surface was used for increasing the CHF of water jet boiling in a flat stagnation zone. The superhydrophilic surface was made by irradiating ultraviolet light to the TiO_2 coated surface. The ultraviolet light's power has no effect on the contact angle.
2. The boiling incipience on the superhydrophilic surface is greatly delayed. The boiling curves are remarkably shifted to the rightward. In the fully developed nucleate boiling regime, the boiling curves are quite steep.
3. The effects of the impact velocity, the nozzle diameter, and the subcooling of water on the CHF for the superhydrophilic surface are the same as that for the common metal surface. However, the CHF on the superhydrophilic surface is about 50% higher than that on common copper surface.
4. The empirical correlations are obtained for predicting the boiling heat transfer and the CHF of steady boiling for the saturated and subcooled water jet impingement in the flat stagnation zone on the superhydrophilic surface.

Nomenclature

$c_{p,l}$ = specific heat (J/kg K)
d = diameter of jet nozzle (m)
G = mass flux of liquid jet (kg/m^2 s)
h_{fg} = latent heat of evaporation (J/kg)
q = wall hat flux (J/m^2 s)
$q_{c,0}$ = CHF of saturated water (J/m^2 s)
q_c = CHF of subcooled water (J/m^2 s)
ΔT_{sat} = wall superheat (K)
ΔT_{sub} = subcooling of subcooled water (K)
v = impact velocity of jet flow at nozzle exit (m/s)
V = kinematics viscosity (m^2/s)
σ = surface tension (N/m)
ρ_1 = liquid density (kg/m^3)
ρ_v = vapor density (kg/m^3)

References

[1] Wang, R., Hashimoto, K., Fujishima, A., Chikuni, M., Kojima, E., Kitamura, A., Shimohigoshi, M., and Watanabe, A., 1997, "Light-Induced Amphiphilic Surfaces," Nature (London), **388**, pp. 431–432.
[2] Takata, Y., Hidaka, S., Cao, J. M., Tanaka, K., Masuda, M., Ito, T., Watanabe, T., and Shimohigoshi, M., 2000, "Boiling and Evaporation From a Superhydrophilic Surface," Therm. Sci. Eng., **8**, pp. 33–41.
[3] Takata, Y., Hidaka, S., Cao, J. M., Nakamura, T., Yamamoto, H., Masuda, M., and Ito, T., 2005, "Effect of Surface Wettability on Boiling and Evaporation," Energy, **30**, pp. 209–220.
[4] Ishigani, S. K., Nakanishi, J. K., and Hikoshichi, T. A., 1978, "Boiling Heat Transfer for a Plane Water Jet Impinging on a Hot Surface," *Proceedings of the Sixth International Heat Transfer Conference*, Paris, **1**, pp. 445–450.
[5] Nakanishi, J. K., Ishigani, S. K., Ochi, T., and Morita, I., 1980, "Two-Dimensional Water Jet Impinging for Cooling of High Temperature Plate," Trans. JSME, Series B, **46**, pp. 714–724.
[6] Kumagai, S., Sano, T., Kamata, T., Suzuki, S., and Kubo, R., 1994, "Boiling Heat Transfer to an Impinging Jet in Cooling a Hot Metal Slab," Trans. JSME, (B), **60**, pp. 259–263.
[7] Kumagai, S., Sano, T., Kamata, T., Suzuki, S., and Kubo, R., 1995, "Transient Cooling of Hot Plate with Impinging Water Jet," Trans. JSME, (B), **61**, pp. 351–356.
[8] Ma, C. F., and Bergles, A. E., 1986, "Jet Impinging Nucleate Boiling," Int. J. Heat Mass Transfer, **29**, pp. 1095–1101.
[9] Liu, Z. H., and Wang, J., 2001, "Study on Film Boiling Heat Transfer for

Water Jet Impinging on a High Temperature Flat Plate," Int. J. Heat Mass Transfer, **44**, pp. 2475–2481.

[10] Robidou, H., Auracher, H., Gardin, P., and Lebouche, M., 2002, "Controlled Cooling of a Hot Plate With a Water Jet," Exp. Therm. Fluid Sci., **26**, pp. 123–129.

[11] Liu, Z. H., and Zhu, Q. Z., 2002, "Predication of Critical Heat Flux for Convective Boiling of Saturated Water Jet Impingement on the Stagnation Zone," ASME J. Heat Transfer, **124**, pp. 1125–1130.

[12] Liu, Z. H., Tong, T. F., and Qiu, Y. H., 2004, "Critical Heat Flux of Steady Boiling for Subcooling Water Jet Impingement on the Flat Stagnation Zone," ASME J. Heat Transfer, **126**, pp. 179–183.

[13] Qiu, Y. H., and Liu, Z. H., 2005, "Critical Heat Flux of Steady Boiling for Saturated Liquids Jet Impinging on the Stagnation Zone," Int. J. Heat Mass Transfer, **48**, pp. 4590–4597.

[14] Wolf, D. H., Incropera, F. P., and Viskanta, R., 1996, "Local Jet Impingement Boiling Heat Transfer," Int. J. Heat Mass Transfer, **39**, pp. 1395–1406.