R. C. Chu
U. P. Hwang
R. E. Simons

Conduction Cooling for an LSI Package: A One-Dimensional Approach

The introduction of LSI packaging has significantly increased the number of circuits per silicon chip, and at the same time has greatly increased their heat flux density. In comparison to earlier MST (monolithic systems technology) products, the heat flux which must be removed from the new multi-chip substrates (100 or more chips) has increased by an order of magnitude or more. This paper discusses an innovative conduction-cooling approach using He gas encapsulation which has been developed in response to the new LSI technology requirements. Background is provided on the liquid-encapsulated-module technology which preceded the new approach, and the basic challenges encountered in building a thermal bridge from individual chips to the module and cold plate are described. The underlying theory of operation is presented using one-dimensional mathematical and discrete analog models. The effects of various factors such as geometry, chip tilt, He concentration, air leakage, and materials are illustrated using these models. A thermal sensitivity analysis is performed to determine variations in junction temperatures and the contributions of the major parameters. The companion paper by Oktay and Kammerer which follows this one treats the more general "multi-dimensional" approach using numerical analysis techniques.

Introduction

With the advent of integrated circuits it has become increasingly difficult to provide a proper thermal environment for circuit packages, especially those for high-performance applications. As the trend for further integration and microminiaturization continues, thermal design problems are becoming so critical that heat removal is recognized as one of the factors limiting the achievement of higher-performance packages. Three basic thermal problems are associated with achieving the required temperature control for high-performance packaging. These are heat transfer by conduction from the internal heat sources of a package to its external surface, heat removal from the external surface of a package by a cooling medium, and the maintenance of coolant temperatures for a given package or system in a chosen application environment.

This paper discusses an innovative conduction cooling approach using He gas encapsulation that has been developed as an enhanced thermal conduction path. A step-by-step one-dimensional thermal analysis of this cooling approach is summarized. A discussion of the multi-dimensional analysis is given in the companion paper by Oktay and Kammerer. An explanation of the basic concept is provided together with discussions of all factors and parameters affecting the overall performance of this approach.

Previous IBM systems such as the System/370 Model 168 and the IBM 3033 typically had peak heat flux densities of 1.5–2.5 W/cm^2 at the chip level and 0.3–0.6 W/cm^2 at the module level. At these flux levels, adequate cooling to maintain junction temperatures below a functionality limit of 85°C can be provided by a conventional air-liquid hybrid scheme [1], wherein heat is removed directly from the chip-carrying modules by means of conduction and by convection resulting from the forced flow of air. However, coolability with air is limited by both the heat transfer coefficient attainable (0.003–0.012 W/cm^2-°C) and by the inability of the air stream to absorb heat without a relatively large temperature rise.

The overall cooling requirements for the new chip and module technology planned for the IBM 3081 Processor

Unit, including a maximum junction temperature of 85°C, were complicated by an increase in the number of circuits per chip and by the development of a multi-chip ceramic substrate to carry as many as 100 or 118 chips. As a result, the peak heat flux densities increased to 20 W/cm^2 at the chip and 4 W/cm^2 at the module levels. It was clearly evident from these projections that air cooling could not be expected to satisfy future cooling requirements; therefore, development of direct liquid-immersion cooling schemes was begun in the late 1960s.

Although the early efforts concentrated on the circulating-liquid scheme [2] with a fluorocarbon coolant being pumped through modules, by 1971 the module-cooling design had evolved into a self-contained liquid-encapsulated module (LEM) [3]. The substrate carrying the integrated circuit chips was mounted within a module-cooling assembly containing a dielectric liquid coolant. Boiling at the chip surfaces resulted in very high heat transfer coefficients (0.17–0.57 W/cm^2-°C) with which to meet the chip cooling requirements. The heat was transported from the dielectric coolant to internal fins and was then transferred to water flowing through an externally attached cold plate. Although this technique appeared to be capable of meeting the requirement to cool a 4-W chip within a 300-W module, additional cooling-related concerns had arisen by the mid-1970s. It was absolutely essential that the liquid be extremely pure and that all residues from chip- and module-joining processes be removed since these contaminants could be dissolved in the coolant and redeposited at the chip-to-module interconnecting pads as part of the boiling process. The end result could be corrosion or failure of the interconnecting pads. In addition, single-chip boiling experiments have shown significant variability in initiation of chip boiling. In some instances, significant superheating of the chips occurred before boiling commenced and the desired chip temperature levels were attained. In other instances, no boiling occurred and the chip was cooled by natural liquid convection, but at unacceptably high chip temperatures. Finally, with the state of the art for boiling that existed at the time, 4 W on a 0.457 × 0.457-cm chip was considered to be the upper limit on the power density. This would allow no extendability of chip cooling capacity. If the chip power was later increased film boiling would result, with potentially catastrophic runaway of chip temperatures.

As a result of these concerns, an effort was commenced in 1975 to develop a viable cooling alternative. Conceptually, it was desired to bring the water-carrying cold-plate surface as close "thermally" to the chip heat sources as possible. At the same time, it was necessary to allow for variations in chip heights and locations resulting from the manufacturing process. In addition, allowances had to be made for nonuniform thermal expansion or contraction across whatever path was provided.

The concept was thus conceived of the spring-loaded mechanical piston touching the chip to provide a thermal path from chip to case, with point contact and minute air gaps between the chip and piston and between the piston and module housing (hat) [see Fig. 1(a) and later discussion]. Based on preliminary calculations, it was determined that temperature drops due to thermal conduction across the air gaps would be too high to satisfy cooling requirements; therefore, an interface medium with a significantly higher thermal conductivity than air would be required. This medium would have to be easy to apply and chemically compatible with the various materials used within the module. Helium gas was identified as meeting both of these requirements and was thus selected as the interface medium.

Thus the total module-cooling assembly, patented as the *gas-encapsulated module* [4] and later known as the *thermal conduction module* (TCM), provided for a multiplicity of chips with an individual piston contacting each chip and providing a thermal path to the module housing. This module is discussed in detail in the companion paper by Oktay and Kammerer [5].

Here, we give a detailed account of the thermal considerations in development of the TCM and the corresponding one-dimensional mathematical treatment used to quantify these considerations. The section on analysis of conduction cooling provides a brief description of the individual thermal resistance terms used in this analysis and presents equations for calculating their magnitude. The combination of these resistances is then discussed with respect to the resulting cooling limits. Finally, the effects of internal thermal parameters (surface roughness, chip tilt, piston tip radius, piston diameter, piston length, piston material, and He concentration) are discussed, and their effects on the associated thermal resistance are examined.

Analysis of conduction cooling

An individual module contains a multiplicity of integrated circuit chips and thermal paths to the cold plate. For purposes of analysis, it is convenient to consider a single-chip cell, as shown in Fig. 1(b). In this instance, the junction temperature T_j of the chip may be expressed as

$$T_j = \Delta T_{j\text{-}c} + P_c(R_c + R_{c\text{-}p} + R_t + R_{p\text{-}h} + R_h) + P_m R_{ext} + T_w , \quad (1)$$

where $\Delta T_{j\text{-}c}$ is the temperature drop (in °C) from the junction to the chip, T_w is the water temperature, P_c and P_m are the chip and module powers (in W), and the R are thermal resistances (in °C/W) with the following subscript definitions: c = chip, c-p = chip-to-piston, t = piston tip, p-h = piston-to-hat, h = hat, ext = external resistance from

hat to water. The following sections discuss the individual thermal resistances and their associated parameters.

• *Temperature drop from the junction to the chip* $T_{j\text{-}c}$

The rise of the junction temperature over the chip temperature results from power dissipation of the collector-base depletion region carrying the bulk of the injected current [6]. This region of power dissipation may be approximated by a rectangular parallelepiped imbedded within the chip. For a typical chip with dimensions of $0.457 \times 0.457 \times 0.038$ cm, the chip may be considered to be an infinite body in comparison with the power dissipation region, which has typical dimensions of $2.5 \times 15 \times 0.35$ μm.

Kutateladze [7] reported that the thermal resistance of a rectangular heat source in an infinite solid can be expressed as $R_{con} = [\ln (4a/b)]/(2ak)$, where a and b are the dimensions of the rectangular source, k is the thermal conductivity (W/cm−°C) of the infinite solid, and R_{con} is the thermal resistance. Since, for the device junctions under consideration, the chip may be treated as an infinite solid, the thermal resistance from the junctions to the chip R_{j-c} can be estimated by the Kutateladze equation. For $k = 1.465$ W/cm−°C (pure Si) [8], $R_{j-c} = 230$°C/W. Since $\Delta T_{j-c} = P_j R_{j-c}$, for a junction power dissipation P_j of 0.0135 W, $\Delta T_{j-c} = 3.1$°C.

• *Chip internal thermal resistance* R_c

Since the heat-dissipating devices are located near the side of the chip containing the pads and most of the heat flow is to the piston contacting the opposite side of the chip, there will be an internal temperature drop across the chip. For purposes of analysis, these heat sources may be considered to be uniformly distributed over the side of the chip with the pads, resulting in a uniform heat flux at the surface. At the other side of the chip, point contact will exist between the chip surface and the crowned surface of the piston. Most of the heat is transferred to the piston across the region surrounding the contact point, *i.e.*, the center portion of the chip (shown later). This results in a chip thermal-constriction resistance R_c as the heat flows into this confined area; R_c can be estimated using the work of Kennedy [10]. By varying the ratio of the effective heat-sink radius r to the total effective chip radius r_0 (0.2285 cm), the corresponding variation in R_c can be obtained; see Fig. 2(a). Sixty percent of the heat was found (shown later) to be conducted through the area in which $r/r_0 = 0.6$, corresponding to $R_c \approx 0.43$°C/W.

• *Thermal resistance from the chip to the piston* $R_{c\text{-}p}$

The thermal resistance in the interface between the chip and the piston, $R_{c\text{-}p}$, is a complex function of many geometric, physical, and thermal characteristics of the contacting solids and the interfacial fluids. The resistance $R_{c\text{-}p}$ can be considered as composed of three parallel thermal resistances: a conduction resistance through many small metallic contact areas, thermal radiation, and thermal conduction through the interfacial fluid.

Achieving a good, reproducible, and reliable thermal path was considered critical to the success of the module-cooling scheme. It was not considered practical within the various

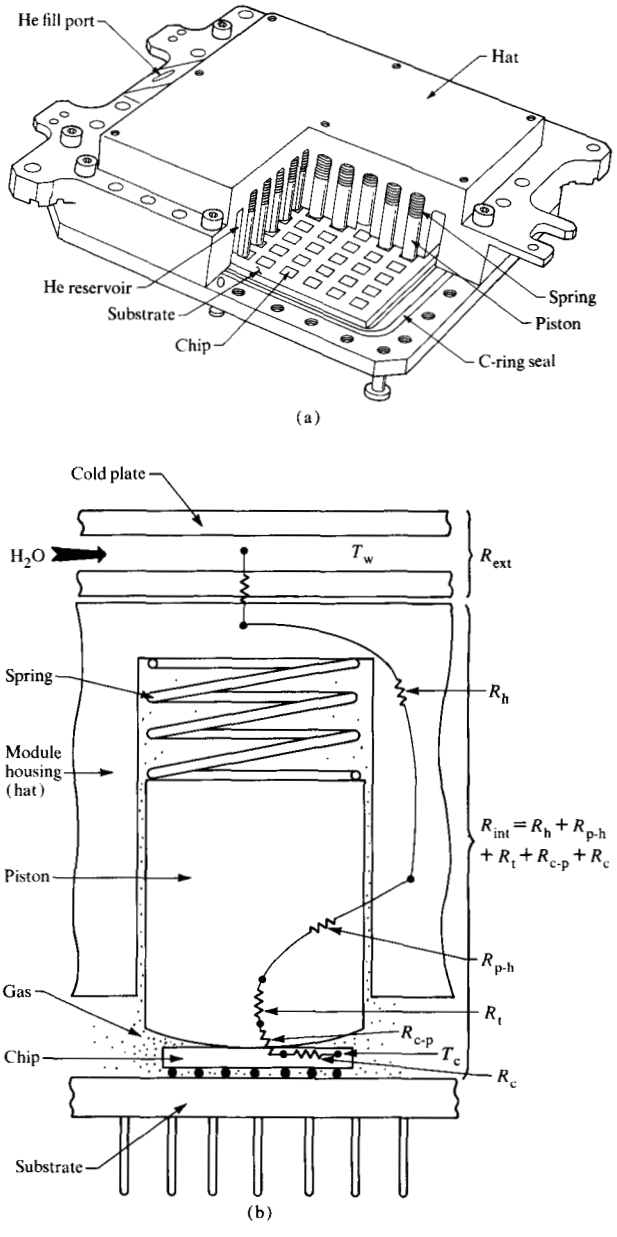

Figure 1 (a) Thermal conduction module and (b) piston conduction cooling paths. All notation is explained in the text.

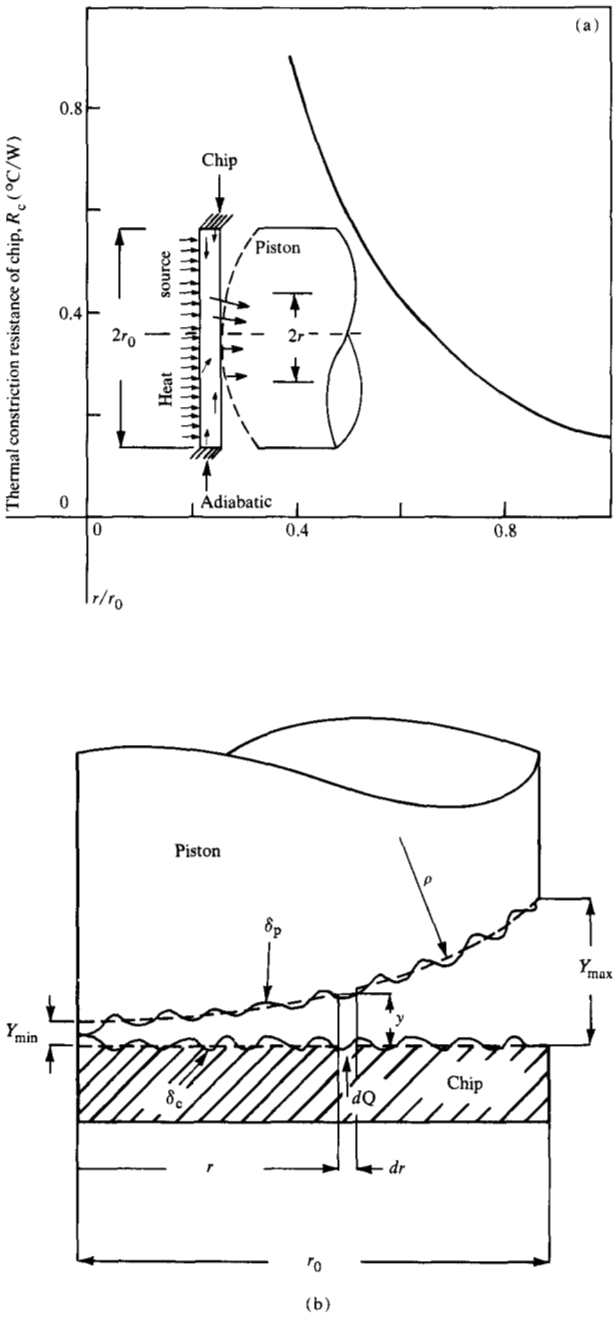

Figure 2 (a) Thermal constriction resistance of the chip R_c as a function of r/r_0. (b) Chip-to-piston interface contact geometry.

packaging and processing constraints to rely on achieving the necessary path by virtue of metallic contact areas. Similarly, it was hoped that the temperature difference across the gap could be minimized; thus, thermal radiation would not be expected to be a dominant mode. Therefore, by design, the dominant thermal path was chosen to be thermal conduction across the interfacial fluid. Consequently, the development of $R_{c\text{-}p}$ discussed in this section is based on thermal conduction across an interfacial fluid medium. As shown in Fig. 2(b), a piston with spherical crown radius ρ and piston surface roughness δ_p may be considered to make point contact with the chip of surface roughness δ_c. The minimum distance between the mean planes of the contacting rough surfaces, Y_{min}, may be expressed [9] as $\eta[(\delta_p)^2 + (\delta_c)^2]^{1/2}$, where η is a geometric parameter with typical values of 3.2–3.7 for light loads. For this analysis the following values were assumed: $\eta = 3.2$, $\delta_p = 0.4$ μm, and $\delta_c = 0.2$ μm (when $\eta = 3.7$ is used, very little difference is observed). Thus, $Y_{min} = 1.43$ μm. From geometric considerations, the maximum interface gap Y_{max} between a piston with $\rho = 14$ cm and the chip can be calculated as $\rho - [(\rho)^2 - (r_0)^2]^{1/2} = 18.6$ μm. Thermal conduction from the chip through the interfacial fluid medium to the piston, as shown in Fig. 2(b), can be written in differential form as

$$dQ = k_g(2\pi r dr)(T_c - T_p)/Y. \tag{2}$$

Here, dQ is the heat flow across the elemental area $2\pi r dr$, k_g is the thermal conductivity of the interface medium (He gas, air, or mixtures thereof), and $Y = \rho - (\rho^2 - r^2)^{1/2}$. The thermal conductivity of the interfacial gas medium is a function of the Knudsen number K_N [11], the ratio of the gas mean free path λ and the spacing between the two contact surfaces. The continuum gas thermal conductivity k_g at a reference condition, k_g^0, is related to the Knudsen number by the relation

$$k_g = k_g^0/(1 + \alpha B K_N), \tag{3}$$

where α is an accommodation parameter [11], $B = 2\gamma/\beta \times (\gamma + 1)$, $\gamma = C_p/C_v$ (the ratio of the specific heats at constant pressure C_p and constant volume C_v), and $\beta = \mu C_p/k_g^0$. If $K_N << 1$, the temperature profile from the solid surface to the gas will be continuous and $k_g = k_g^0$. The λ for pure He gas and air are 0.186 and 0.063 μm, respectively, at 15°C and 101 kPa (760 mm); therefore, pure He gas will result in the highest K_N for this application ($= \lambda/Y_{min} = 0.13$). Since $K_N << 1$, it is appropriate to consider the interface gas as behaving like a thermally conductive continuum. Equation (2) may now be integrated over the entire heat transfer region ($r = 0 \rightarrow r_0$), considering T_c and T_p as uniform chip and piston surface temperatures. The overall $R_{c\text{-}p}$ is then given as

$$R_{c\text{-}p} = (T_c - T_p)/Q$$

$$= \frac{1}{2\pi k_g \rho \ln (Y_{max}/Y_{min})} = 2.9\text{C/W}. \tag{4}$$

Equation (4) may be used to determine the relative heat transfer distribution across the interfacial surface. From the center of the chip out to radius r the fraction of the total heat transfer which occurs within that radius may be given as

$$\frac{Q_r}{Q_{total}} = \frac{R_{c\text{-}p}|_r}{R_{c\text{-}p}|_{r_0}} = 5.8\pi k_g \rho \ln (Y_{max}/Y_{min}) \tag{5}$$

for $0 < r < r_0$. The results of Eq. (5) are shown in Fig. 3. Thus, it is seen that 60% of the heat is transferred within $r/r_0 = 0.6$, or approximately 36% of the available heat-transfer area.

• *Thermal resistance of the piston* R_t

To provide the necessary mechanical allowance (travel) during module assembly, which would in turn accommodate tolerances in the substrate, chip-to-substrate pads, and chip itself as well as subsequent thermal expansion or contraction, the tip of the piston must extend beyond the surrounding hat walls. As a consequence, the heat entering the surface of the piston tip must be conducted through a distance L_t before it can begin to spread through the piston/hat gap into the cooling hat or module case. As previously discussed, most of the heat flow is concentrated in the central portion of the piston tip area. Thus, the heat must spread out to the larger area of the piston radius. The resulting spreading resistance (or thermal constriction resistance) R_t may be approximated by the equation for conduction across a truncated cone, $R_t \approx L_t/k_p\pi r_1 r_2$, where L_t is the distance over which heat conduction occurs ($\approx$0.2 cm), $k_p \approx$ 1.67 W/cm-°C for the aluminum alloy used, r_1 is the effective radius at the interface ($0.6r_0$), and r_2 is the piston radius (0.272 cm). Therefore, $R_t = 1.02$°C/W.

• *Thermal resistance from the piston to the hat* $R_{p\text{-}h}$

The heat transferred from the piston across the He gap to the module hat may be treated mathematically as heat exchange between two conductively coupled extended surfaces or fins. An energy balance equation for the element dx [Fig. 4(a)] gives rise to two differential equations:

$$\frac{d^2T_p}{dx^2} = \frac{k_g\pi D_p}{k_p S A_p}(T_p - T_h)$$

and

$$\frac{d^2T_h}{dx^2} = -\frac{k_g\pi D_p}{k_h S A_h}(T_p - T_h), \tag{6}$$

where D_p is the piston diameter (in cm), A_p is the piston cross-sectional area (in cm^2), S is the piston annular gap (cm), A_h is the cross-sectional area of the piston hat, k_h is the thermal conductivity of the hat, and T_p and T_h are the piston and hat temperatures. The following boundary conditions were also used:

$$\left.\frac{dT_p}{dx}\right|_{x=0} = \frac{-Q}{k_p A_p}, \qquad \left.\frac{dT_p}{dx}\right|_{x=L_t} = 0;$$

$$\left.\frac{dT_h}{dx}\right|_{x=0} = 0, \qquad \left.\frac{dT_h}{dx}\right|_{x=L_t} = \frac{-Q}{k_h A_h}. \tag{7}$$

Applying the Laplace transform technique to Eqs. (6) and (7) yields

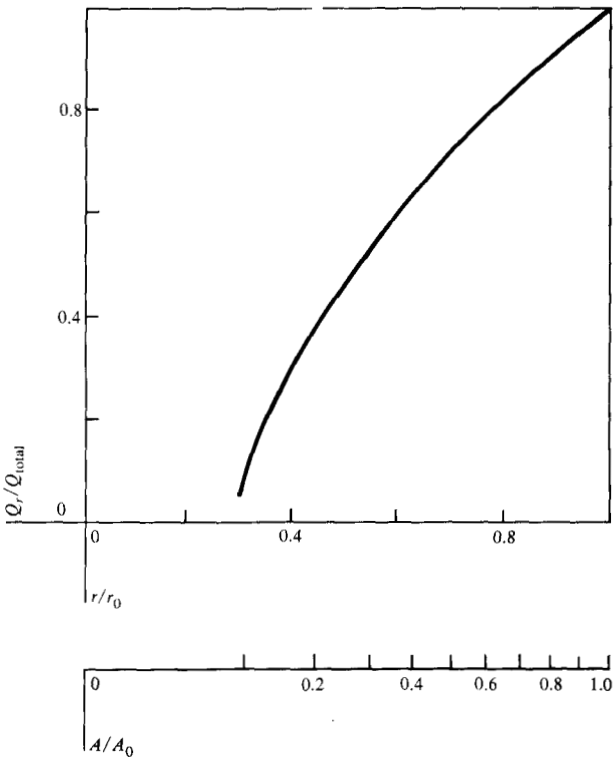

Figure 3 The effective chip-to-piston heat-transfer area; Q_{total} is the total heat transferred from the entire chip area A_0 of radius r_0; Q_r is that amount of heat transferred over the contact area A of radius r.

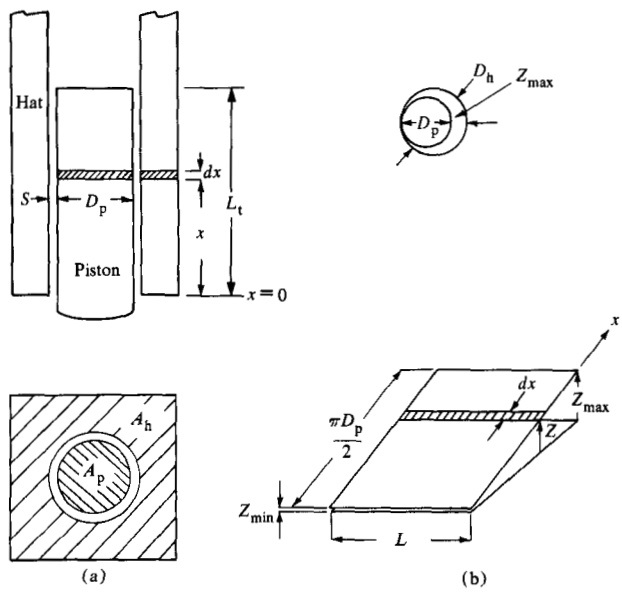

Figure 4 (a) Conductively coupled extended surface geometry and (b) the contact plane approximation.

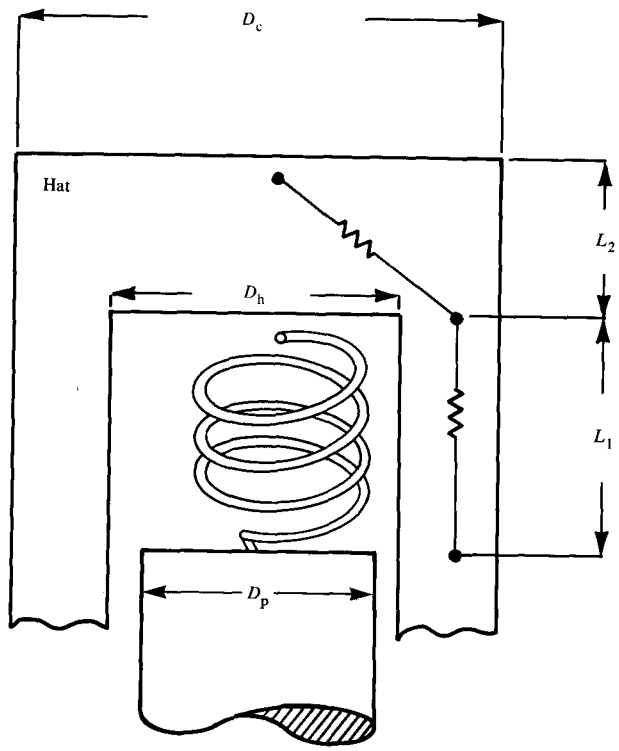

Figure 5 Thermal conduction paths within the module hat.

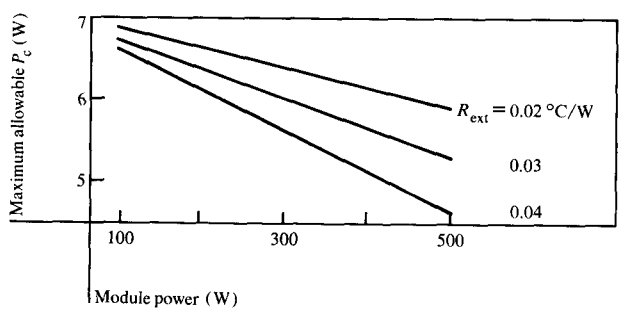

Figure 6 Chip and module cooling limits for various values of R_{ext}; $T_j = 85°C$, $T_w = 24°C$.

$$R_{p-h} = \frac{L_t}{k_p A_p} - \frac{m^2(\cosh aL_t - 1)}{a^3 \sinh aL_t}\left[\frac{1}{k_h A_h} + \frac{\cosh aL_t}{k_p A_p}\right] + \frac{m^2}{k_p A_p}\left(\frac{1}{a^3}\sinh aL_t - \frac{L_t}{a^2}\right) + \frac{\cosh aL_t}{a \sinh aL_t}\left[\frac{1}{k_h A_h} + \frac{1}{k_p A_p}\cosh aL_t\right] - \frac{1}{k_p A_p}\left(\frac{\sinh aL_t}{a}\right), \quad (8)$$

where

$$m^2 = \frac{k_g \pi D}{k_p A_p S},$$

$$n^2 = \frac{k_g \pi D}{k_h A_h S},$$

$$a^2 = m^2 + n^2.$$

In order to calculate $R_{p\text{-}h}$ with Eq. (8), the effective gap between the piston and module hat, S, must be determined or estimated. As shown in Fig. 4(b), the contact between the piston outer surface and the surrounding module hat hole surface may be approximated by two planes. The conduction heat transfer from one plane, through the interfacial fluid medium, to the second plane may be expressed as

$$dQ = \frac{k_g}{Z}(Ldx)\Delta T, \quad (9)$$

where dQ is the heat flow across the elemental area Ldx. Assuming each plane to have a uniform temperature, Eq. (9) can be integrated to obtain

$$R_g = \Delta T/Q = \frac{Z_{max}}{k_g \pi D_p L \ln (Z_{max}/Z_{min})}, \quad (10)$$

where R_g is now the thermal resistance across the piston-to-hat gap.

The equivalent gap S for use in Eq. (8) is

$$S = R_g k_g \pi L = \frac{Z_{max}}{\ln Z_{max}/Z_{min}}, \quad (11)$$

where Z_{max} is the difference between the hole diameter and the piston diameter, and Z_{min} is the distance between the rough contact planar surfaces, which may be estimated by $\eta[(\delta_p)^2 + (\delta_c)^2]^{1/2}$. For a typical piston 1.55 cm long, 0.545 cm in diameter, and with a root-mean square (rms) surface roughness $\delta_p = 0.8$ μm, in contact with a module hat with a hole diameter of 0.55 cm and an rms $\delta_h = 0.8$ μm, the equivalent gap may be calculated using the additional relations

$$Z_{max} = D_h - D_p$$

and

$$S = \frac{Z_{max}}{\ln (Z_{max}/Z_{min})} = 19\ \mu m.$$

Using this result with the previous piston and hat hole dimensions, $R_{p\text{-}h}$ for pure He gas is 2.15°C/W.

- *Thermal resistance of the hat* R_h

Within an individual chip-piston cell, the thermal conduction resistance through the module hat, R_h, may be treated as two serial resistances, as illustrated in Fig. 5. The first

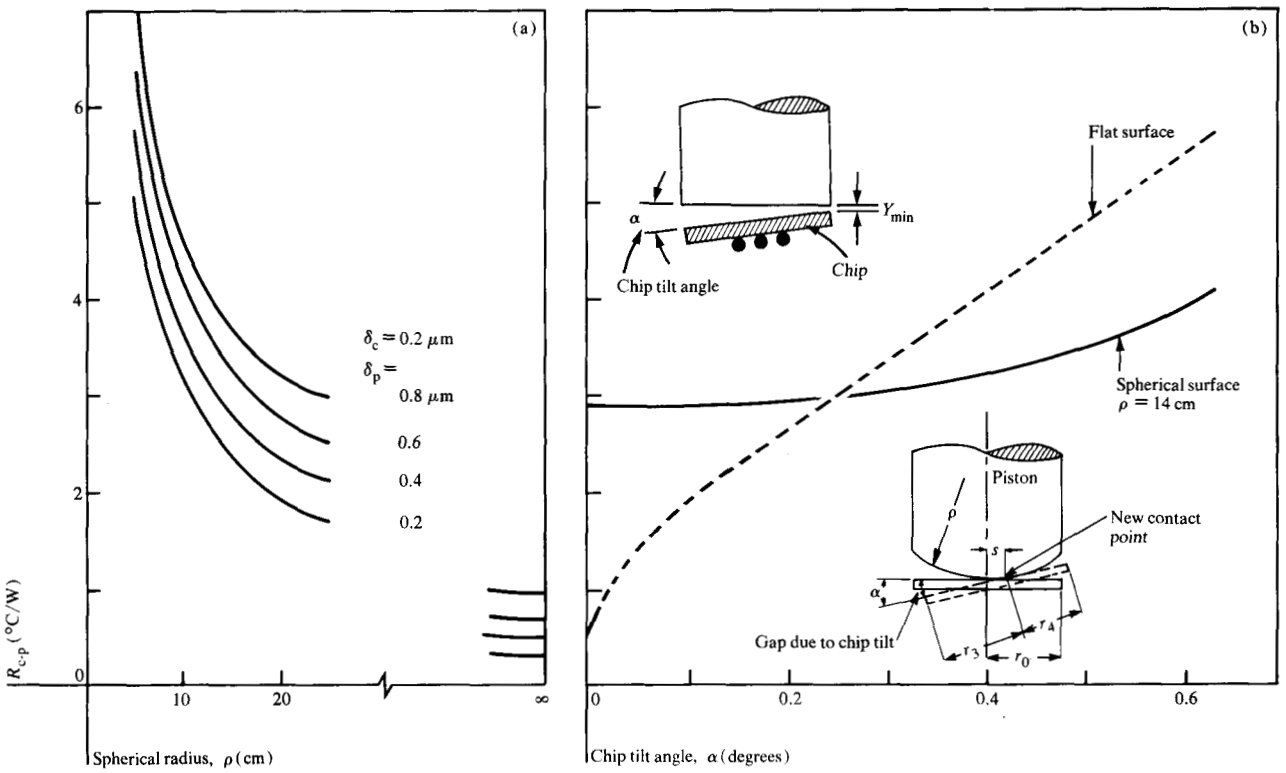

Figure 7 (a) Thermal effects of chip and piston surface roughnesses δ_c and δ_p, and the piston-to-tip spherical radius ρ. (b) Chip tilt geometry for a spherical piston (inset) and the thermal effects of chip tilt on $R_{c\text{-}p}$ for flat (- - - - -) and spherical (———) piston surfaces.

resistance is due to conduction along length L_1 (0.6 cm) of the piston well adjacent to the piston spring, with a cross-sectional area $A_h = [D_c^2 - (\pi/4)\, D_h^2]$. The second resistance is due to heat flow across L_2 (0.85 cm), consisting of the remainder of the module hat up to the cold-plate interface. Along this path, the cross-sectional conduction area expands from A_h to D_c^2. The combined hat thermal resistance can then be calculated by

$$R_h = \frac{L_1}{k_h\left(D_c^2 - \frac{\pi}{4} D_h^2\right)} + \frac{L_2}{\frac{k_h}{2}\left[\left(D_c^2 - \frac{\pi}{4} D_h^2\right) + D_c^2\right]} = 1.58°\text{C/W}. \qquad (12)$$

Cooling limits

The several thermal resistance components may now be added together to form the total module internal thermal resistance R_{int}:

$$R_{int} = R_c + R_{c\text{-}p} + R_t + R_{p\text{-}h} + R_h = 8.08°\text{C/W}.$$

Using this calculated internal resistance and the system water supply temperature of 24°C, we may calculate the maximum allowable chip power as a function of module power and module external thermal resistance from the hat to the water, R_{ext}. Thus,

$$T_j = \Delta T_{j\text{-}c} + P_c R_{int} + P_m R_{ext} + T_w,$$

giving $P_c \leq 7.16 - P_m R_{ext}/8.08$ for a maximum junction temperature of 85°C. The results of this relation are shown in Fig. 6 for a range of R_{ext} (0.02 – 0.04°C/W).

Conversely, we may use Eq. (1) with the highest chip power (2.9 W) and the highest module power (212 W) in the system and the *measured* nominal R_{ext} of 0.02°C/W to determine the highest nominal junction temperature to be expected under nominal conditions: 54.7°C, which is well within the required operating temperature range of 45–85°C.

Effects of internal thermal parameters

The analytical models of thermal resistance developed may now be used to examine the effects of various factors such as surface roughness δ, geometry, chip tilt, materials, and He concentration.

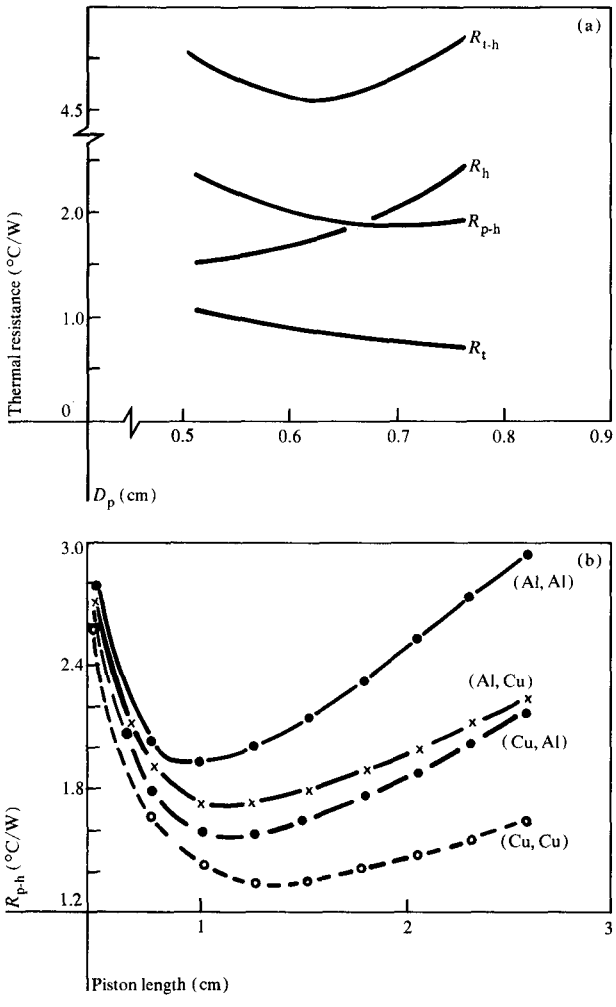

Figure 8 (a) Thermal effects of the piston diameter D_p for various resistances; $R_{t\text{-}h}$ (piston tip to hat), R_h (hat), $R_{p\text{-}h}$ (piston to hat), R_t (piston tip). (b) Effect on $R_{p\text{-}h}$ of the piston length and the materials from which the piston and hat are constructed: key—(piston material, hat material).

The surface roughness of the piston's spherical contacting surface, δ_p, can be controlled by the degree of polishing. The previous calculations assumed $\delta_p = 0.4$ μm, which resulted in a minimum interface gap of 1.43 μm. The equation for Y_{min} may be used to determine the minimum interfacial gap obtainable by decreasing the rms δ_p to the minimum value of 0.2 μm: $Y_{min} = 0.9$ μm. Using this value in Eq. (4) gives a reduced $R_{c\text{-}p}$ value of 2.5°C/W. The net effect of reducing this resistance is a 5% reduction in the overall internal thermal resistance of the module R_{int}. Similarly, increasing δ_p would result in comparable increases in R_{int}. The variations in $R_{c\text{-}p}$ as a function of δ_p are plotted in Fig. 7(a).

The magnitude of the spherical surface radius of the piston tip ρ also affects $R_{c\text{-}p}$. As ρ is increased, the piston surface tends to become flatter, resulting in a decreased $R_{c\text{-}p}$, as shown in Fig. 7(a). For a perfectly flat piston in contact with a chip with zero tilt, $R_{c\text{-}p}$ could be reduced to 0.56°C/W. However, such a design would also be more sensitive to chip tilt which results from substrate camber or nonuniform collapse of the chip pads during chip bonding.

Chip tilt will change the chip-to-piston surface contact geometry. As shown in the inset to Fig. 7(b), for a chip tilt angle α, a new contact point is established a distance away from the original position at the center of the chip. This new contact point results in two heat-transfer regions, with radii r_3 and r_4. In order to estimate the two parallel thermal resistances, Eq. (4) may be used. For example, if we assume $\alpha = 0.3°$ (0.0052 rad), the distance the contact point moves will be given by the arc length $r = \rho\alpha = 0.0728$ cm. Therefore, the two radii can be calculated from

$$r_3 = r_0 + r = 0.301 \text{ cm};$$

$$r_4 = r_0 - r = 0.156 \text{ cm}.$$

The use of these values in the relations for Y_{max} and Y_{min} yield parallel thermal resistances of 4.80 and 8.29°C/W, for an effective overall value of 3.04°C/W. Calculations were performed for tilt angles up to 0.6° and the results are plotted in Fig. 7(b) for both flat and spherically tipped pistons. This figure verifies that a spherically tipped piston is less sensitive to chip tilt than its flat counterpart.

An increase in the piston diameter would increase both the cross-sectional area available for thermal conduction within the piston and the circumferential surface area for conduction across the gas gap from the piston to the hat. This will result in an increase in the thermal resistance within the hat. These individual effects are illustrated in Fig. 8(a), which gives a plot of the thermal resistance *versus* the piston diameter D_d. Also shown is the net effect on the thermal resistance from the tip of the piston to the cold-plate attachment surface. It can be seen that the thermal resistance first decreases with increased diameter and then increases, with the minimum value occurring between D_d = 0.6–0.7 cm.

According to Eq. (8), the piston length L will have an effect on $R_{p\text{-}h}$. However, the direction this effect will take is not apparent because of the hyperbolic nature of the functions involved; numerical calculations were thus performed to determine the variation of $R_{p\text{-}h}$ with L. From Fig. 8(b), it can be seen that an optimum piston length exists for which $R_{p\text{-}h}$ has been minimized. Optimization for simultaneous consideration of both D_d and L can be made using the equations for R_h, $R_{p\text{-}h}$, and R_t. As one might expect, the thermal conductivity of both the piston and the hat will affect $R_{p\text{-}h}$. These effects are shown in Fig. 8(b) for various

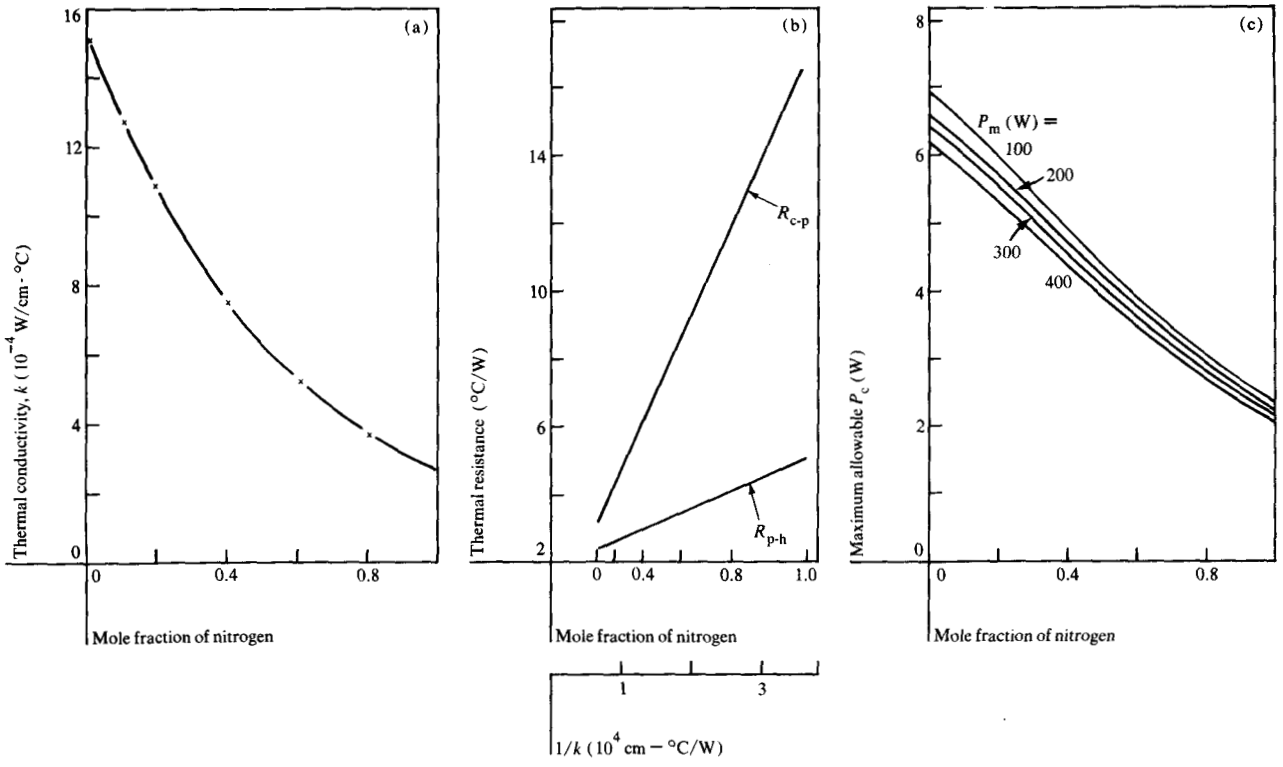

Figure 9 (a) The thermal conductivity of various He/N gas mixtures; k_{He} = 0.00152 W/cm-°C and k_N = 0.000266 W/cm-°C at 37.8°C. (b) The effect of He concentration (in terms of the mole fraction of nitrogen in the mixture and the thermal conductivity of that mixture) on the thermal resistance. (c) The effect of He concentration (in terms of the mole fraction of N in the gas mixture) on the coolability of the chips and module; T_j = 85°C and T_w = 24°C, while R_{ext} = 0.02°C/W.

combinations of materials for the piston (listed first in parentheses) and hat (listed last)—(Al, Cu), (Al, Al), and (Cu, Cu).

For anything but a perfect seal, air is expected to leak (diffuse) into the module during its lifetime, resulting in a dilution of the He concentration and a decrease in the thermal conductivity k of the interface medium, since air is 78% nitrogen and k_N = 0.000266 W/cm-°C ($k_{air} \approx k_N$; k_{He} = 0.000152 W/cm-°C). The calculated k_g for various He/N mixtures [12], including the very significant effect that this dilution has on $R_{c\text{-}p}$ and $R_{p\text{-}h}$, and the corresponding effect on the maximum allowable chip power at several module power levels, is shown in Fig. 9. It should be noted that the seal is designed to limit air ingress to the module to a maximum of 13% over the life of the module. This will result in no more than a 0.7°C/W increase in R_{int} [5]. Thermal protection, in the form of a temperature-sensitive thermistor mounted to the substrate within the TCM, has been provided to give a warning should a significant leak or any other thermal malfunction occur.

Thermal sensitivity analysis

As discussed previously, the internal thermal parameters (surface roughness, chip tilt, piston tip radius, piston dimension, and gas composition) can vary as a result of variations in the manufacturing and assembly processes. The external thermal resistance of the cold plate can also vary due to variations in cold-plate attachment and the flow rate of water through the cold plate. In addition, both chip and module powers can vary due to chip manufacturing tolerances, chip and circuit usage, and power supply tolerances. Therefore, junction temperatures throughout the system can be expected to vary, and the variance can be estimated using statistical analysis techniques [13].

The square of the standard deviation of the junction temperature (also known as the *variance*) is given as

$$\begin{aligned}\delta_{T_j}^2 &= \left(\frac{\partial T_j}{\partial P_c}\delta_{P_c}\right)^2 + \left(\frac{\partial T_j}{\partial P_m}\delta_{P_m}\right)^2 + \left(\frac{\partial T_j}{\partial R_{int}}\delta_{R_{int}}\right)^2 \\ &\quad + \left(\frac{\partial T_j}{\partial R_{ext}}\delta_{T_{ext}}\right)^2 + \left(\frac{\partial T_j}{\partial T_w}\delta_{T_w}\right)^2 \\ &= (\bar{R}_{int}\delta_{P_c})^2 + (\bar{R}_{ext}\delta_{P_m})^2 + (\bar{P}_c\delta_{R_{int}})^2 \\ &\quad + (\bar{P}_m\delta_{R_{ext}})^2 + \delta_{T_w}^2, \end{aligned} \tag{13}$$

where δ is the standard deviation of each parameter, and the overscore denotes the mean value for each parameter. Using typical values of P_c = 2.9 W ± 30%, P_m = 212 W ± 10%,

$R_{int} = 8.08°C/W \pm 20\%$, $R_{ext} = 0.02°C/W \pm 30\%$, and $T_w = 24°C \pm 10\%$, the junction temperature variance can be calculated, using (13), to be

$$3.45 + 0.02 + 2.44 + 0.18 + 0.64 = 6.73.$$

Combining this result with the junction temperature previously calculated using (1), the range of junction temperatures will be $54°C \pm 7.8°C$, where the $\pm$ tolerance represents the 3σ extremes. It can also be seen from this example that the major contributors to the variation in T_j are the variations in the chip power and the internal thermal resistances.

Summary and conclusions

This paper has reviewed the cooling requirements and the basic cooling concepts underlying the development of the He gas-filled thermal conduction module. The origin of each element of thermal resistance has been physically explained and analytically derived to provide quantitative results. From its conception, the TCM was viewed as being more predictable and amenable to analytical treatment from a thermal standpoint than the previous cooling technologies cited, and this view has been generally verified in practice (see the companion paper by Oktay and Kammerer). It is also significant to note that where deviations from predicted results have arisen, they have generally been in the direction of better cooling capability. Both the analytical results cited here and those in the companion paper have demonstrated that the TCM cooling concept is thermally superior to its *liquid-encapsulated module* (LEM) predecessor, offering the cooling capability required for current and projected circuit and chip technologies.

References and note

1. V. W. Antonetti and A. L. Pascuzzo, "Cooling Large Scale Computer Systems," *ASHRAE J.* **13,** 25–30 (1971).
2. R. E. Simons and K. P. Moran, "Immersion Cooling Systems for High Density Electronic Packages," *National Electronic Packaging and Production Conference (NEPCON) Proceedings,* Industrial and Scientific Conference Management, Inc., Chicago, IL, 1977, pp. 396–409.
3. N. G. Aakalu, R. C. Chu, and R. E. Simons, "Liquid Encapsulated Air Cooled Module," U.S. Patent 3,741,292, June 26, 1973.
4. R. C. Chu, O. R. Gupta, U. P. Hwang, and R. E. Simons, "Gas Encapsulated Cooling Module," U.S. Patent 3,993,123, November 23, 1976.
5. S. Oktay and H. C. Kammerer, "A Conduction-Cooled Module for High-Performance LSI Devices," *IBM J. Res. Develop.* **26,** 55–66 (1982, this issue).
6. R. C. Joy and E. S. Schlig, "Thermal Properties of Very Fast Transistors," *IEEE Trans. Electron Devices* **ED-17,** 586–594 (1970).
7. S. S. Kutateladze, *Fundamentals of Heat Transfer,* Academic Press, Inc., New York, 1965, p. 97.
8. The thermal conductivity k is a function of the impurity; see G. A. Slack, "Thermal Conductivity of Pure and Impure Silicon, Silicon Carbide, and Diamond," *J. Appl. Phys.* **35,** 3460–3466 (1964). The value for pure Si was used because the region in which most of the R_{con} occurs has impurity concentrations $<10^{18}$ atoms/cm^3 (*i.e.,* fairly "pure" silicon).
9. M. M. Yovanovich, "Thermal Contact Resistance in Microelectronics," *NEPCON Proceedings,* 1978, pp. 177–188.
10. D. P. Kennedy, "Spreading Resistance in Cylindrical Semiconductor Devices," *J. Appl. Phys.* **31,** 1490–1497 (1960).
11. Y. Ogniewicz and M. M. Yovanovich, "Effective Conductivity of Regularly Packed Spheres: Basic Cell Model with Constriction," *Heat Transfer and Thermal Control Systems,* L. S. Fletcher, Ed., Vol. 60 of *Progress in Astronautics and Aeronautics,* American Institute of Astronautics and Aeronautics, New York, 1978, pp. 209–228.
12. R. Bird, N. Stewart, and E. Lightfoot, *Transport Phenomena,* John Wiley & Sons, Inc., New York, 1960.
13. U. P. Hwang, V. W. Antonetti, and C. G. Keller, "Calculating IC Junction Temperatures in a Packaged System," *Electron. Packaging Prod.* **21,** 259–265 (1981).

Received July 17, 1980; revised August 28, 1981

The authors are located at the IBM Data Systems Division laboratory, Poughkeepsie, New York 12602.