Nanoscale thermal transport. II. 2003–2012

Citation

As Published
http://dx.doi.org/10.1063/1.4832615

Publisher
American Institute of Physics (AIP)

Version
Final published version

Accessed
Sun Apr 24 14:02:58 EDT 2016

Citable Link
http://hdl.handle.net/1721.1/97398

Terms of Use
Creative Commons Attribution 3.0 Unported Licence

Detailed Terms
http://creativecommons.org/licenses/by/3.0/
Nanoscale thermal transport. II. 2003–2012

David G. Cahill,1,a) Paul V. Braun,1 Gang Chen,2 David R. Clarke,3 Shanhui Fan,4
Kenneth E. Goodson,5 Pawel Keblinski,6 William P. King,7 Gerald D. Mahan,8
Arun Majumdar,9 Humphrey J. Maris,10 Simon R. Phillpot,11 Eric Pop,12 and Li Shi13
1Department of Materials Science and Engineering and the Frederick Seitz Materials Research Laboratory,
University of Illinois, Urbana, Illinois 61801, USA
2Department of Mechanical Engineering, MIT, Cambridge, Massachusetts 02139, USA
3School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA
4Department of Electrical Engineering, Stanford University, Stanford, California 94305, USA
5Department of Mechanical Engineering, Stanford University, Stanford, California 94305, USA
6Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180, USA
7Department of Mechanical Sciences and Engineering, University of Illinois, Urbana, Illinois 61801, USA
8Department of Physics, Penn State University, University Park, Pennsylvania 16802, USA
9Department of Mechanical Engineering, University of California, Berkeley, California 94720, USA
10Department of Physics, Brown University, Providence, Rhode Island 02912, USA
11Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611, USA
12Department of Electrical and Computer Engineering, University of Illinois, Urbana, Illinois 61801, USA
13Department of Mechanical Engineering, University of Texas, Austin, Texas 78712, USA

(Received 6 April 2013; accepted 7 August 2013; published online 14 January 2014)

A diverse spectrum of technology drivers such as improved thermal barriers, higher efficiency thermoelectric energy conversion, phase-change memory, heat-assisted magnetic recording, thermal management of nanoscale electronics, and nanoparticles for thermal medical therapies are motivating studies of the applied physics of thermal transport at the nanoscale. This review emphasizes developments in experiment, theory, and computation in the past ten years and summarizes the present status of the field. Interfaces become increasingly important on small length scales. Research during the past decade has extended studies of interfaces between simple metals and inorganic crystals to interfaces with molecular materials and liquids with systematic control of interface chemistry and physics. At separations on the order of ~1 nm, the science of radiative transport through nanoscale gaps overlaps with thermal conduction by the coupling of electronic and vibrational excitations across weakly bonded or rough interfaces between materials. Major advances in the physics of phonons include first principles calculation of the phonon lifetimes of simple crystals and application of the predicted scattering rates in parameter-free calculations of the thermal conductivity. Progress in the control of thermal transport at the nanoscale is critical to continued advances in the density of information that can be stored in phase change memory devices and new generations of magnetic storage that will use highly localized heat sources to reduce the coercivity of magnetic media. Ultralow thermal conductivity—thermal conductivity below the conventionally predicted minimum thermal conductivity—has been observed in nanolaminates and disordered crystals with strong anisotropy. Advances in metrology by time-domain thermoreflectance have made measurements of the thermal conductivity of a thin layer with micron-scale spatial resolution relatively routine. Scanning thermal microscopy and thermal analysis using proximal probes has achieved spatial resolution of 10 nm, temperature precision of 50 mK, sensitivity to heat flows of 10 pW, and the capability for thermal analysis of sub-femtogram samples. © 2014 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [http://dx.doi.org/10.1063/1.4832615]

TABLE OF CONTENTS

I. INTRODUCTION ........................................ 2

II. THERMAL CONDUCTANCE OF INTERFACES .................... 3

A. Vibrational thermal energy transport at interfaces.................. 3
   1. Overview ..................................... 3
   2. Interfacial thermal transport via molecular dynamics techniques.... 3
   3. Role of interfacial bonding and structure...................... 4
   4. Interfaces in composite materials and nanoparticle suspensions.... 5

a)Electronic mail: d-cahill@illinois.edu

0021-8979/2014/1(1)/011305/45 1, 011305-1 © Author(s) 2014

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 18.51.1.88 On: Thu, 11 Jun 2015 13:40:59
I. INTRODUCTION

The relentless decrease in the size of devices and structures, the increase in their operating speeds and frequencies, and the ever more aggressive thermal conditions imposed upon them requires sophisticated understanding and control of thermal transport at the nanoscale. While thermal performance itself is a key metric for applications such as thermal barrier coatings, nuclear fuels, and materials for thermoelectric energy conversion, there are many other technologies for which the ability to manipulate heat is vital even though the primary goal of the structure is non-thermal. Some examples of technology drivers for the science of nanoscale thermal transport are phase change memory devices, heat assisted magnetic recording, thermal management of wide variety of high power and nanoscale electronic and optoelectronic devices, and the proposed use of intensely heated nanoparticles in medical therapies.

The objective of this review is to provide perspectives on new developments in the applied physics of nanoscale thermal transport that have resulted from advances in experiment, theory, and simulation over the past decade; to identify gaps in current state-of-knowledge; and discuss directions for future research. This review is an update to an earlier review, co-authored by some of the same authors, which surveyed progress in the field through 2002. The topics we have selected emphasize the importance of interfacial phenomena in nanoscale thermal transport and avoid topics that have been extensively reviewed by others in recent years, e.g., nanostructured thermoelectric materials and the transport properties of isolated graphene and carbon nanotubes. We limit our discussion of thermoelectric materials to the lattice thermal conductivity of Si nanowires. Our discussion of carbon allotropes focuses on the effects of interfaces on thermal transport. We also omit discussion of non-linear effects and transport in low-dimensional systems as these topics were recently reviewed by Li and co-workers. Experimental methods for measuring the transport properties of nanowires were recently reviewed by Weathers and Shi. Thorough development of the fundamentals of nanoscale thermal transport are available in recent textbooks.

For the purposes of this review, the term nanoscale is interpreted liberally, embracing characteristic dimensions ranging from the atomic up to sub-micron. Another way of envisaging the definition of nanoscale is the regime in which sub-continuum effects are important; that is, the system cannot be adequately described by bulk transport properties and continuum heat transport equations that do not explicitly consider interfaces. For example, while it has long been recognized that grain boundaries and interphase boundaries provide significant thermal resistance, there continues to be significant efforts to understand the transport of thermal energy across interfaces. One of the major advances in the past decade has been the extension of experiment and theory to interfaces that involve molecular materials and liquids. Systematic control of the strength of interface bonding has revealed the importance of interface chemistry and physics on heat conduction.

We include discussion of radiative transport by thermally excited electromagnetic modes in this review because we believe that there is an increasing intersection between the science of near-field radiative transport through nanoscale gaps and the thermal conductance of interfaces $G$ between materials when $G$ is small because of weak bonding, interface roughness, or a large mismatch in vibrational density of states. The coupling of electronic and vibrational modes at an interface is usually thought to be limited to atomic scales but recent theory suggests that longer-range coupling across the interface can be significant.
Heterostructures provide both an approach to reducing thermal conductivity in the cross-plane direction and a laboratory for studies of highly anisotropic thermal transport. Anisotropy in engineered superlattices has now been extended to natural superlattice materials, such as those in layered perovskite structures. Anisotropy is most extreme in two-dimensional or near two-dimensional materials, of which carbon-based materials (nanotubes, and most recently graphene) are the most prominent; however, considerable work is now being performed on layered metal chalcogenide systems, such as MoS$_2$ and WSe$_2$.

Historically, theory and simulation have been able to provide predictions with spatial and time resolutions for which corresponding experimental techniques did not exist or were not fully developed. In the last decade, there have been significant improvements in the spatial and temporal resolution of key techniques, and a much improved ability to determine temperatures and heat flow. This is now allowing theory and experiment to be confronted with experimental results at an unprecedented level. Advances in density functional theory and high performance computing have reached the point where parameter-free predictions of the thermal conductivities of simple crystals are now possible.

We begin our discussion with a review of theory and experiment on the transport of heat at an interface between two materials. In Sec. III, we focus on thermal transport in nanostructures and nanoscale electronic devices. The related topic of low thermal conductivity in nanostructured materials, including anisotropic crystals and natural superlattices, is addressed in Sec. IV. We explore issues associated with thermal measurement science and thermal processing at the nanoscale in Sec. V.

II. THERMAL CONDUCTANCE OF INTERFACES

A. Vibrational thermal energy transport at interfaces

1. Overview

A thermal boundary resistance is found for the heat flow between two materials in contact along a planar interface. The heat flow is perpendicular to the planar interface

\[ J_Q = G \Delta T, \]

where \( \Delta T \) is the small temperature difference between the two sides of the interface, and \( G \) is the boundary conductance; \( G \) has the units of power per area per degree.

A common interface in thermal transport is between two insulators. Then, the heat is carried only by phonons. The thermal boundary resistance \( G^{-1} \) from phonons was first calculated by Khalatnikov.\(^{17,18} \) He considered only longitudinal phonons. His calculation is called the acoustic mismatch theory (AMT). It is found to be a good approximation at low temperatures, where \( G \propto T^3 \). At room temperatures, it is a poor approximation. Young and Maris\(^{19} \) provided the first theory of \( G \), which included all phonons on both sides of the junction. They derived only the special case that both sides of the interface had face-centered-cubic structures, with the same lattice constant. Pettersson and Mahan\(^{20} \) extended the Young-Maris theory to include any kind of lattices on the two sides of the interface, but assumed they were in intimate contact. These theories are based on phonon scattering from the interface atoms. Persson and Ueba\(^{21} \) recently derived the phonon-phonon case using surface Green’s functions. That allowed them to consider rough surfaces.

Superlattices are periodic arrays of alternating materials. One layer in the array will have \( n \)-atomic layers of a material, while the alternating layers have \( m \) atomic layers. Thermal conductivity can be measured parallel to the layers, or perpendicular to them. For a crystalline superlattice, the thermal conductivity parallel to the layers is typically comparable to the average of the two bulk materials, as long as the interfaces between the layers do not have a large density of defects. The interesting measurement is for heat flow perpendicular to the layers, which is called the “cross-plane direction.” Hyldgaard and Mahan\(^{22} \) originally predicted that the thermal conductivity of a superlattice in the cross-plane direction could be ten times smaller than the values of the bulk materials that makeup the superlattice. This reduction is generally observed. The thermal conductivity is sometimes lower in value than a random alloy of the same material. The large reduction in heat flow is due to the numerous interfaces of the superlattice. The thermal boundary resistance is a major component of the thermal resistance. Later Simkin and Mahan\(^{23} \) suggested that a superlattice would have a minimum thermal conductivity as a function of the number of atomic layers within one superlattice layer. Several experiments have tested these ideas.\(^{24-28} \) (A summary of experimental studies of superlattices prior to 2002 is given in Ref. 1.)

Until recently all theory and experiments on superlattices were for insulating layers. In this case only phonons carry heat, and the boundary resistance is due to phonon scattering. Recently, there has been experimental work on metal-metal interfaces\(^{29} \) and multilayers,\(^{30} \) where the electrons are the important carriers of heat. Some work has been done also on metal-insulator superlattices.\(^{27,31} \)

2. Interfacial thermal transport via molecular dynamics techniques

Modeling and simulation of interfacial heat transport has been the focus of numerous computational efforts in the past decade. The most prominent approach is to use molecular dynamics (MD) simulations where atoms and molecules follow classical dynamics based on the numerical solution of the Newton’s second law of motion. The forces between atoms and molecules are derived from interatomic potentials.\(^{32} \) In the most popular scheme, equal power, \( P \), heat source, and sink are introduced on both sides of the interface, and when the steady state is established, the interfacial thermal conductance, \( G \), is evaluated from the temperature drop at the interface, \( \Delta T \), via \( G = P/(\Delta T) \), where \( A \) is the area of the interface. While the heat source-sink method allows determination of the value of the conductance, detailed understanding of the thermal coupling across the interface can be obtained from the equilibrium MD simulations.\(^{33,34} \)

Specificially, a Green-Kubo relationship is employed, which
relates the fluctuating cross interfacial heat power, $P(t)$, to the interfacial conductance via

$$G = \frac{1}{AK_BT^2} \int_0^\infty \langle P(t)P(0) \rangle \, dt,$$  \hspace{1cm} (2)

where $\langle P(t)P(0) \rangle$ is the interfacial heat power autocorrelation function. The interfacial heat power can be easily obtained by monitoring the total energy of the material on the one side of the interface, and then taking its time derivative. Also, the upper limit of the integral given by Eq. (2) is in practice finite as the autocorrelation function decays to zero. In addition to the value of $G$, the full functional form of $\langle P(t)P(0) \rangle$ provides dynamical information about coupling between molecular vibrations in the two phases forming the interface.36

Equilibrium and non-equilibrium MD thermal conductivity techniques characterize either macroscopic or interfacial behavior “integrated” over all vibrations. To explicitly provide information on underlying mechanism of heat flow in terms of individual vibrations, recently developed phonon dynamics MD based techniques have proved to be useful. The basic idea of phonon dynamics simulations is to generate wave packets of lattice vibrations of a narrow range of frequencies and well-defined polarization, using the normal modes of the perfect crystal structure, and then to allow them to propagate through the system. Following the scattering process, one can decompose the transmitted and reflected energy into phonons and thus can identify the nature of the scattering processes and to calculate the transmission coefficient for each phonon.38 For example, scattering of high frequency phonons at interfaces involves a mixture of “acoustic scattering” where the reflected phonon has the same character as the incident phonon and “diffusive scattering” where the correlation with the incident phonon wave packet is lost.38

The phonon dynamics technique relates to Green’s functions methods.39 Both methods enable calculations of the phonon transmission coefficients. However, due to computational limitations and difficulties with the correct description of anharmonic effects, Green’s functions calculations were mostly performed on simplified interfacial models, such as junctions,40 or nanowires.41 The transmission coefficients for longitudinal acoustic phonons determined by phonon dynamics simulations are similar to those obtained from GF calculations for a simple one dimensional chain model of masses connected by springs.37,42 Green’s function calculations on fully 3-dimensional structures, on the other hand, lump together all phonons at a given frequency, whereas phonon dynamics simulations can discriminate among different directions and polarizations.

3. Role of interfacial bonding and structure

The acoustic mismatch model (AMM) and diffusive mismatch model (DMM) have traditionally provided a basis for understanding interfacial thermal resistance. These models provide prescriptions for calculating the transmission coefficient of a phonon with which the interfacial conductance can be obtained by appropriate summations over all phonons.19,43 While both AMM and DMM provide values for transmission coefficients, they calculate them purely based on the bulk properties of the two solids forming the interface; thus they do not account explicitly for the strength of the interfacial bonding. Consequently, MD simulations were used to model the effect of interfacial bonding on the interfacial conductance. In particular, a study of a self-assembled monolayer (SAM) water interface (Fig. 1) exposed the relationship between interfacial bonding and interfacial conductance, suggesting a more or less linear relationship between the two and saturation on the conductance at high interfacial bonding.44 Interestingly, in this case, the MD results are in quantitative agreement with the experiment.45 A similar relationship between the interfacial conductance and the interfacial bonding strength was observed in a number of MD simulation on interfacial systems, including hard solid-polymer,42 solid-solid,46 solid-SAM-solid,47,48 carbon nanotube on silica substrate,49 and for gas-solid interfaces.50

Within the theoretical treatment, the effect of interfacial bonding on the phonon transmission coefficients and consequently on the interfacial thermal conductance can be demonstrated via an analysis of a simple model of semi-infinite one-dimensional chains of masses connected by springs.51 A calculation based on this model shows that the interfacial phonon transmission coefficient for a typical phonon representing those carrying the majority of the heat strongly increases with increasing interfacial stiffness.46 Furthermore, the transmission coefficient saturates at strong bonding at the value coinciding which by the prediction of the AMM. A similar conclusion was reached in the case of a fully 3 dimensional model, which developed an analytical extension of the AMM that includes the effects of weak interface bond spring constants.52 Analysis of the one dimensional chain junctions and of the extended AMM model shows that in the limit of very weak interfacial bonding the conductance is...
mismatch between bulk properties the two materials forming
facial thermal transport beyond the limits set by the inherent
relations to explore the potential of enhancing and tuning inter-
the interface are closely matching. This motivated MD simu-
is maximized when the vibrational spectra on both sites of
is dominant, the DMM predicts that the interfacial transport
with strong interfacial bonding. When the diffuse scattering
is not important, the AMM model predic-
tion is an upper limit of the possible conductance associated
enhanced by the insertion of an interfacial film with media-
ting vibrational properties, or by intermixing of the two solid
at the interfacial region. Another example of such a media-
ing layer is an organic SAM used in the bonding of solids,
or solids and liquids. In particular, MD simulations demon-
strated that a confined water layer between quartz solids
could lead to very good thermal transport properties. Possible effects of surface nano-patterning on thermal
transport were also explored, generally indicating benefits
for thermal conduction due to increased surface area and bet-
ter vibrational matching at the interface.

4. Interfaces in composite materials and nanoparticle suspensions

Interfaces generally present a resistance to heat flow in
addition to the resistance of bulk materials forming the inter-
face. This, due to high density of interfaces, can be particu-
larly limiting to heat flow in nanocomposite materials. This
effect manifests itself most spectacularly in the case of car-
bon nanotube (CNT) and graphene composites. Despite the
fact that the reported value of thermal conductivity of CNTs
and graphene is 2000 to 6000 W m\(^{-1}\) K\(^{-1}\) along axial and
in-plane direction while most of the polymers have ther-
mal conductivity of less than 0.5 W m\(^{-1}\) K\(^{-1}\), an addition
of nanocarbon fillers leads to rather limited thermal conduc-
vity enhancements. MD simulations combined with transient
laser heating measurement indicated that the interfacial con-
ductance is very low, of the order of 20 MW m\(^{-2}\) K\(^{-1}\). The
associated resistance is equivalent to the resistance of
of ten of nanometers the matrix materials. Analysis of the MD
results shows that such high resistance is due to weak bond-
ing between carbon nanofiller and the matrix and very strong
bonding within the nanocarbon filler.

In the context of the composite thermal transport, MD
simulation exposed differences in value of the interfacial
thermal conductance, depending on the mode of the heat
transfer across the interface. In particular, it was found that
the graphene-organic matrix cross-plane interfacial thermal
resistance determined via thermal relaxation method is about
20 MW m\(^{-2}\) K\(^{-1}\) (Ref. 62) and consistent with experimental
laser based pump-probe measurements on CNTs water. Such
consistency lies in the fact that the relaxation method
mimics pump-probe measurements. In fact, in both exper-
iment and standard MD simulations, the energy is pump pre-
dominantly to high frequency modes. As demonstrated by
simulations, before dissipating to the matrix, first the energy
flows from high to low frequency modes within graphene.
Consequently, the interfacial thermal resistance in such ther-
mal relaxation can be understood in terms of a sum of inter-
nal to the graphene resistance and the “external” resistance
associated with the heat flow across the interface.

In the heat source-sink simulations where the heat is
added and removed in the matrix away from the interface the
interfacial conductance for the heat flow perpendicular to the
graphene matrix interface was determined to be about
170 MW m\(^{-2}\) K\(^{-1}\), i.e., almost order of magnitude higher
than that obtained from the relaxation simulations. This high
conductance can be related to the “external” resistance alone,
and heat can flow from the matrix via graphene to the matrix
on the other side without employing high frequency modes
in graphene.

Yet another value of the effective interfacial thermal
conductance of about 30 MW m\(^{-2}\) K\(^{-1}\) was calculated for
the case of heat entering a high-aspect ratio carbon nanofiber
in order to propagate along the fiber. In these cases, the
heat enters the fiber predominantly via low frequency modes
involving vibrations normal to the interface. However, rapid
transport along the carbon fiber is facilitated by acoustic
modes with no or little normal the interface displacement.
Consequently the heat flow in this configuration involves
energy redistribution over modes within the fiber and is thus
affected by associated internal to the fiber thermal resistance.

All the cases described above show that interfacial ther-
mal transport can involve a complicated process in which the
phonons carrying most of the heat across the interface are
not necessarily the phonons carrying the heat in the bulk or
along the fiber, which in turn makes phonon-phonon cou-
pling within each material forming the interface an important
factor.

In the context of thermal transport in nanofluids (colloi-
dal suspensions of solid nanoparticles in heat transfer fluids),
it was postulated that at the interface an ordered liquid layer
might have higher thermal conductivity than bulk liquid. Molecular dynamics simulation on simple liquids concluded
that such an interfacial layer does not contribute significantly
to heat transfer. For strong solid-liquid interactions, typical
of those in nanofluids with metallic nanoparticles, a percolat-
ing network of amorphous-like fluid structures can emerge.
This can facilitate additional thermal conduction paths. However, a discernible increase in thermal conductivity is
possible only for exceedingly small colloidal particles, lim-
ited to a few tens of atoms. Furthermore, interpretation of the
cooling rates of Au nanoparticles suspended in water and or-
ganic solvents does not appear to require any unusual
thermophysical properties of the surrounding liquid to explain the experimental results.67

5. Heated nanoparticles

A potential for using metallic nanoparticles and electromagnetic radiation induced nanoscopic heat sources is perhaps the ultimate frontier of the nanoscale thermal transport. In particular, when large laser heating is applied, the local temperature gradients can be as high at 100 K per nanometer, creating a unique opportunity of studying highly nonequilibrium systems. In this case, molecular dynamics simulations demonstrated the fluid next to the small nanoparticle can be heated well above its boiling point without a phase change, while on flat surface, a vapor formed readily.68 These differences in heat transfer were explained by the curvature-induced pressure close to the nanoparticle, which inhibits boiling. When the nanoparticle temperature is much larger than the critical fluid temperature, a very large temperature gradient develops, resulting in close to ambient temperature just a radius away from the particle surface. The behavior reported allows us to interpret recent experiments where nanoparticles can be subjected to transient heating up to and above 1000 K, without observing boiling of the surrounding liquid.69,70 In cases of such extreme heating and associated high thermal gradients, a significant temperature drop occurs the interface. Moreover, nature of a supercritical interfacial liquid layer can contribute to the interfacial resistance and well to the stability. These issues are highly intriguing and remain largely unexplored.

6. Experiment

Interfaces between elements or simple inorganic crystals represent the easiest system to fabricate and model and have attracted considerable attention. Care must be taken however to ensure that the interface is atomically abrupt and clean; otherwise, interpretation of the data and rigorous modeling is difficult. Reactive sputter deposition of TiN on MgO substrates71 at high temperatures produced the highest thermal conductance observed to-date for individual interfaces at room temperature, $G \approx 700 \text{ MW m}^{-2}\text{K}^{-1}$. Nearly ideal interfaces for fundamental studies of thermal transport were prepared by Horn-von Hoegen and colleagues using deposition of Bi on atomically clean Si under UHV conditions.72,73 The thermal conductance was measured in-situ by heating of the Bi layer with a laser pulse and measurement of the Bi temperature by ultrafast electron diffraction (UED). $G$ for the Bi-Si interface was determined to be as low as low as 13 MW m$^{-2}$ K$^{-1}$, comparable to $G$ for Bi deposited on diamond.74 Interestingly, $G$ for Bi-Si was nearly the same for (111) and (001) oriented Si substrates, indicating that details of the crystallographic structure of the interface have a minimal effect on $G$. $G$ was also independent of thickness even for Bi layers only 2.5 nm thick.75 Going forward, UED can probe interfaces and structures, e.g., non-planar Ge-Si interfaces76 that are not accessible in more conventional approaches based on pump-probe optical methods.

Although single interfaces are often easier to fabricate, a key advantage of a multilayer systems is that although the effect of each interface may be small, in a system which consists of a “solid of interfaces,” the effect of interface conductance on thermal transport can be dramatic. In one multilayer system, atomic layer deposition and magnetron sputter deposition were used to synthesize thin-film multilayers of W/Al$_2$O$_3$. The high interface density produced a strong impedance to heat transfer, giving rise to a thermal conductivity as low as $\kappa \approx 0.6 \text{ W m}^{-1} \text{K}^{-1}$ as the interface density approached 0.5 nm$^{-1}$. Even lower conductivity, $\kappa < 0.5 \text{ W m}^{-1} \text{K}^{-1}$, was observed in Ta/TaO$_2$ multilayers.77 At an interface density of 0.14 nm$^{-1}$, Mo/Si multilayers78 of the type used for mirrors in the extreme-ultraviolet have a thermal conductivity of $\approx 1 \text{ W m}^{-1} \text{K}^{-1}$. A multilayer approach has also been applied to the study of organic/metal interfaces, namely, copper phthalocyanine interfaces with Ag and Al.79,80

Common to all four studies76–79 is the use of a metallic layer as a component in the multilayer. The thermal conductivity of ultrathin metallic layers in a metal/dielectric multilayer depends on the strength of electron-phonon coupling.77,78,81,82 The electronic contribution to thermal conductivity $\kappa_{el}$ is suppressed when the metal layer thickness is smaller than $\sqrt{\kappa_{el}/g}$, where $g$ is the electron-phonon coupling parameter43 and $\kappa_l$ is the lattice contribution to the thermal conductivity. Similarly, the interface conductance of a metal-dielectric interface has an additional series conductance81 of $\sqrt{\kappa_{el}/g}$. Au has weakest electron-phonon coupling and lowest lattice thermal conductivity among the metal elements: at room temperature, $g \approx 3 \times 10^{16} \text{ W m}^{-3} \text{K}^{-1}$ and $\kappa_l \approx 3 \text{ W m}^{-1} \text{K}^{-1}$.84 Therefore, for an interface with Au, the additional series conductance is $\sqrt{\kappa_{el}/g} \approx 300 \text{ MW m}^{-2} \text{K}^{-1}$. For all other metals, $\sqrt{\kappa_{el}/g}$ is significantly larger.

The lowest thermal conductivity observed in a fully dense solid at room temperature is $\kappa < 0.05 \text{ W m}^{-1} \text{K}^{-1}$ for the cross-plane thermal conductivity of thin films of WSe$_2$ grown from vacuum deposition of alternating W and Se layers.85 This conductivity is 30 times smaller than the $c$-axis conductivity of WSe$_2$ single crystal but more significantly is a factor of $\approx 5$ times lower than minimum thermal conductivity86 calculated from the atomic density and the speeds of sound. The origin of this “ultralow” conductivity not yet fully understood—for example, it is not clear if the transport can be described as a the series sum of the thermal conductance of closely spaced interfaces acting independently—but the strong anisotropy of the elastic constants that reduces the density of states of phonon modes that propagate in the cross-plane direction37,88 is thought to be an important factor.

The thermal conductivity of amorphous soft materials (e.g., polymers) is inherently low ($\kappa \approx 0.2 \text{ W m}^{-1} \text{K}^{-1}$), and thus the effect of interfaces is less dramatic. There are specific cases however where interface conductance is still important, e.g., surfactant stabilized nanoparticles. Most inorganic nanoparticles are made dispersible in a solvent by coating their surface with a thin layer of organic ligands, and thus are best described as a system containing two interfaces. The first is the interface between the inorganic nanoparticle and the surrounding ligand, and the second is the interface between the ligand and the solvent phase. A motivation for
understanding heat flow surrounding nanoparticles comes from the broad interest in nanoparticle photothermal therapies. In nanoparticle photothermal therapies, nanoparticles, which are designed to target various tumors, are injected into the body. Then, near IR radiation is used to heat the nanoparticles and locally kill tissue. Simple calculations show, however, that for any reasonable combination of optical power, nanoparticle loading, and background optical absorption, it is very challenging to create a nanometer-scale local temperature excursion using a cw laser source. Excitation with pulsed lasers with low duty-cycle can localize temperature excursion in the immediate vicinity of a nanoparticle. At short time-scales, interface thermal conductance plays an important role in the heat transfer from nanoparticle to surroundings.

Model systems consisting of nanoparticles dispersed in water and other solvents have been studied to investigate the effect of the organic stabilizing group on the thermal conductance $G$ of the particle/liquid interface. In a series of experiments on heat conduction from ligand-stabilized nanoparticles to solvents, $G$ ranged from $5 \text{ MW m}^{-2} \text{ K}^{-1}$ for alkanethiol stabilized AuPd alloy nanoparticles in toluene to $100–300 \text{ MW m}^{-2} \text{ K}^{-1}$ for various metal nanoparticles in water. The similar values of $G$ for all nanoparticles in water indicates that the thermal coupling between nanoparticles and water is strong regardless of the self-assembled stabilizing ligand. This was true for tiopronin (anionic), ethylene glycol (nonionic), and cetyltrimethylammonium bromide (cationic bilayer) ligands on particles ranging from 3 to 24 nm.

In principle, the measured value of $G$ of a nanoparticle-surfactant-solvent system could include at least four contributions: (i) the flow of heat from the nanoparticle into the stabilizing group; (ii) the flow of vibrational energy along the length of the stabilizing group or across the molecular layer; (iii) the transport of vibrational energy from the stabilizing group to the surrounding fluid; and (iv) the possibility of the direct transfer of heat from the nanoparticle to fluid molecules that penetrate the self-assembled molecular layer of the organic stabilizing group. The large difference in the values of $G$ for toluene and water suspensions suggests that either the third or fourth mechanism listed above is playing a critical role since we do not have any reason to believe that the first two mechanisms would have a strong dependence on the solvent. The large values for the thermal conductance $G$ of nanoparticles in aqueous solution may either indicate that the thermal coupling between hydrophilic molecular layers and water is strong, while the thermal coupling between alkanethiol monolayers and toluene is much weaker, resulting in a smaller $G$, or that water penetrates deeply within the molecular layer and interacts directly with the nanoparticle. This second explanation appears unlikely in light of the cetyltrimethylammonium bromide results since this molecule forms a bilayer on the surface of the nanoparticle with a hydrophobic core, and the water content of the interior of the bilayer is small.

With nanoparticle systems, it is not possible to explore heat transfer between hydrophobic surfaces and water, or highly polar surfaces and a non-polar solvent. These systems are not stable and aggregate rapidly. However, by moving to a planar system, such experiments are possible. Using time-domain thermoreflectance (TDTR), Ge and co-workers measured the transport of thermally excited vibrational energy across planar interfaces between water and solids that have been chemically functionalized with both hydrophobic and hydrophilic self-assembled monolayers. In these systems, it is useful to think of the interfacial conductance in the form of a Kapitza length—i.e., the thermal conductivity of water divided by $G$. The Kapitza length at hydrophobic interfaces (10–12 nm) is a factor of ~3 larger than the Kapitza length at hydrophilic interfaces (36 nm), indicating a much stronger thermal coupling between hydrophilic surfaces and water than hydrophobic surfaces and water.

Hard-soft-hard interfaces (e.g., two inorganic layers sandwiching an organic layer) are interesting both because they can be experimentally realized in systems with high degree of structural and chemical control, and for the technological reason that they may provide a route to dramatically alter thermal conductances in composite materials with attractive physical properties. A key question in hard-soft-hard systems is the relationship between interfacial bonding character and thermal conductance. MD simulations as well as analytical models predict interfacial bond stiffness can cause $G$ to vary by an order of magnitude. This has been probed by sandwiching a SAM between two surfaces. In one example, AuSAMGaAs junctions were made using alkanedithiol SAMs and assembled by transfer printing: Thermal conductance was measured using the $3\omega$ technique, and no thermal conductance dependence on alkane chain length was observed over a relatively narrow chain length (8–10 carbons). The thermal conductances ranged from 25 to 28 MW m$^{-2}$ K$^{-1}$. Because the character of the bonding does not change, this experiment provides only one data point for the chemical effects on $G$.

To explicitly study the effect of interfacial bonding character, a gold film was transfer-printed to a self-assembled monolayer with systematically varied termination chemistries. Using a combination of ultrafast pump-probe techniques (time-domain thermoreflectance and picosecond acoustics) and laser spallation experiments, changes in bonding strength and heat flow at the gold-SAM interface were independently measured and correlated. It was shown that varying the density of covalent bonds within this single bonding layer modulates both interfacial stiffness and interfacial thermal conductance. Heat transport at Au/SAM/Qz interfaces using SAMs of the same length and different $\omega$-end-groups: (1) Methyl (CH$_3$–C$_{11}$–Si $\equiv$) and (2) thiol (SH–C$_{11}$–Si $\equiv$) were studied. The thiol functionality is expected to form a strong covalent-like bond to the gold film while only a weak van der Waals attraction should attach the gold film to the methyl-terminated surface, $G = 68 \text{ MW m}^{-2} \text{ K}^{-1}$ for the Au/SH-C$_{11}$–Si$\equiv$Qz interface and $G = 36 \text{ MW m}^{-2} \text{ K}^{-1}$ for the Au/CH$_3$–C$_{11}$–Si$\equiv$Qz interface, confirming that interfacial bonding directly impacts thermal conductance. Particularly striking is the similarity between $G$ measured for Au/SH-C$_{11}$–Si$\equiv$Qz ($G_{\text{avg}} = 65 \text{ MW m}^{-2} \text{ K}^{-1}$) and the recent calculations by Keblinski for a Au/SAM/Si interface with strong bonding at both SAM end-groups ($G = 60 \text{ MW m}^{-2} \text{ K}^{-1}$).
Hsieh and co-workers continuously varied the strength of interfacial bonding using high hydrostatic pressure in a SiC anvil. An Al-graphene-SiC interface was formed on the face of the SiC anvil using transfer printing of a graphene followed by sputter deposition of Al. As the pressure increases from ambient to 10 GPa, the value of $G$ increases by nearly an order of magnitude from $G \approx 30$ to $G \approx 200$ m$^{-2}$ K$^{-1}$. At high pressures, $G$ for Al-graphene-SiC is the same as $G$ for Al deposited directly on SiC cleaned at high temperatures.

Hard-soft-hard interfaces also offer the potential for switchable nanoscale thermal conduction. Recently, it was shown that the thermal conductivity of a boron nanoribbon bundle can be modulated between high and low thermal conductivity states by wetting the interface between the nanoribbons comprising the bundle. When the ribbons are welded together cleanly by the solvent (ethanol/water mixture), phonons transmit across the interface, leading an increase in the thermal conductivity. Once this solvent is removed and replaced with isopropyl alcohol (IPA), the thermal conductivity decreases. This process is reversible. IPA leaves residues after evaporation, and it is thought these residues separate the nanoribbons, making them effectively non-interacting.

B. Near-field radiation transport across a gap

1. Overview

Photons can transfer energy between two distant bodies that are at different temperatures. The formula for the thermal conductance contains the factor $\sigma T^3$, where $\sigma$ is the Stefan-Boltzmann constant. This energy transfer is called a far-field transfer. There is also a near-field energy transfer that is electro-magnetic in origin. It operates on a distance scale less than ten micron, i.e., less than the wavelength of thermally excited infrared radiation. This near-field energy transfer has been much studied. A review of experiment and theory of near-field thermal transport is given in Secs. II B 2 and II B 3.

Polder and van Hove wrote out the basic equations in terms of local dielectric functions. The dielectric function for a homogeneous material is a function of wave vector and frequency $\epsilon(q, \omega)$. The dielectric formulations typically neglect the wave vector dependence and just use $\epsilon(0, \omega)$ to match the electromagnetic fields at the surface. For insulating solids, that is a good approximation since the wave vector dependence of the dielectric response is small, but for metallic samples, we expect deviations when the gap becomes comparable to atomic length scales.

2. Experiment

In the far field, radiation heat transfer between macroscopic bodies is limited to below that given by the Plank blackbody radiation law for the spectral-based flux or the Stefan-Boltzmann law for the total flux, both of which will be referred to as the blackbody limit. When two objects are brought close to each other, radiation can exceed the blackbody limit because of the tunneling of evanescent waves. There are two types of evanescent waves. One decays exponentially from one-side of the interface but is a propagating wave on the other side of the interface, and the other decays exponentially on both sides of the interface. The former happens when electromagnetic waves inside a material with a high and positive dielectric constant propagate across the interface with a material of a lower but positive dielectric constant at an angle of incidence larger than the critical angle as determined by the Snell law, creating an exponentially decaying electromagnetic field, i.e., an evanescent wave, on the other side. The latter is called a surface wave, which happens when the dielectric constants on both sides of an interface are opposite signs, and the absolute value of the negative dielectric constant is larger than that of the positive dielectric constant. A negative dielectric constant occurs when the frequency of photons falls in between the transverse and longitudinal optical phonon frequencies for dielectrics or below the plasma frequency for metals, when there are resonances between material waves (e.g., electrons of optical phonons) and the electromagnetic waves.

Examples of surface waves are surface plasmon-polaritons or surface phonon-polaritons, which are resonance modes of the electrons and electromagnetic waves, and the optical phonons and electromagnetic waves, respectively. Evanescent waves normally do not carry a net energy flux and decay rapidly away from the interface over a distance of the order of the wavelength, and hence do not contribute to radiation heat transfer between objects separated at macroscopic distances. However, when the two surfaces are brought close by, evanescent waves can reach the other medium and enhance heat transfer.

Measurement of near-field radiation is challenging because (i) thermal radiation heat transfer rate is small, and (ii) near-field effects start to be important only when the separation between the two surfaces are comparable to the dominant wavelength. Using Wiens displacement law $\lambda T = 2898$ $\mu$m K as a guideline, at room temperatures, blackbody radiation peaks at the wavelength of $\approx \lambda = 10$ $\mu$m. Measuring small heat transfer at small separations of two surfaces is difficult. Tien and co-workers reported nearfield radiative heat transfer at cryogenic temperatures between two copper disks with a diameter of 8.5 cm, whose cold and the hot sides were at 4.1 K and 15.1 K, respectively. Due to the low temperatures, the near-field effect can be observed at large gaps. They observed increasing radiative heat transfer as the size of the gap decreases from 1 mm to 50 $\mu$m, consistent with theoretical expectations. Their experiments were stimulated by cryogenic multilayer insulations. Due to the low emittance of metallic surfaces, the maximum heat transfer observed is $\approx 1/29$ of the Planck theory prediction. Hargreaves carried out near-field radiation heat transfer measurements between two chromium surfaces with a diameter of 4.0 cm, to gaps down to 1 $\mu$m at room temperature, demonstrating an enhancement from $\approx 1.75$ in the farfield to $\approx 2.95$ W m$^{-2}$ K$^{-1}$ at a 1 $\mu$m gap, which is $\approx 50\%$ of blackbody radiation. More recently, Krallik et al. carried out near-field measurements between tungsten surfaces, each 35 mm in diameter, with a separation distance between $\approx 1$ $\mu$m to 1 mm. The hot-side surface was maintained between 10 and 100 K and cold side maintained at 5 K.
reported a three orders of magnitude increase in the emittance between the two surfaces.

Metallic surfaces were used in these experiments because capacitive sensing can be used to control the separation and parallelism between the surfaces. Reducing the separation between surfaces less than microns, however, is challenging because of the difficulties in maintaining the surfaces parallel and the presence of dust particles and other imperfections. Xu et al. took the strategy of reducing the surface area of two metallic surfaces using a scanning tunneling microscope (STM) configuration. One side of the surface is a thin-film thermocouple, and the other side of the surface is an indium tip flattened by pressing it against a glass flat. Their experiment, however, was not conclusive due to the significantly reduced heat transfer with smaller surface area and the low sensitivity of the thermocouple heat flux sensor. Kittel et al. used an improved STM configuration to measure heat transfer of a gold tip with a gold substrate and with a GaN substrate for tip-substrate separation in the range of 1–100 nm. They observed saturation of heat transfer when the tip-substrate separation is less than 10 nm.

To study near-field heat transfer between non-metallic surfaces, especially between two surfaces supporting surface phonon-polaritons, Hu et al. used two quartz surfaces with a diameter of 1.27 cm separated by polystyrene beads with a nominal diameter of 1 μm. They reported near-field heat transfer between the quartz surfaces exceeding that of the blackbody radiation. Recently, Ottenso et al. reported measurements of near-field radiation between two large sapphire surfaces, 5 × 5 cm² in area, for distance between 2–100 μm near room temperature. They maintained the surface separation and parallelism using small capacitive sensors deposited on the surfaces. Their experiment observed increasing heat transfer with decreasing separations between two surfaces starting from 10 μm, and heat transfer coefficient exceeding that between blackbodies.

The Chen group at MIT developed a sphere-plate experimental configuration to investigate near-field heat transfer. This experimental configuration overcomes the difficulties of maintaining parallelism in the two parallel plates and unknown tip geometry in the STM experiment. Their experiment took advantage of the proven extreme thermal sensitivity of bi-layer cantilevers used for atomic force microscopes, which compensates for a smaller area involved in radiative transfer due to the curvature of the sphere compared to the parallel plate configuration. They reported experimental results between a glass sphere and different substrates including glass, silicon, and gold, for sphere and plate separation down to 30 nm. Such experiments were extended to smaller separations and between metallic surfaces. Rousseau and co-workers used a slightly modified configuration and reported near-field heat transfer between a glass sphere and a glass substrate. A subtle but significant difference between the two experiments is that in the MIT experiment, the cantilever is heated while the surrounding is maintained at ambient temperature, while in the experiment of Rousseau et al., the substrate is heated. In the former case, only near-field heat transfer is measured, since the far-field contribution is included in the initial bending of the cantilever. In the latter case, the influence of the heated substrate to the sphere, and cantilever is modeled and used as a fitting parameter. Theoretical simulation by Otey and Fan shows that the far-field component creates significant ambiguity in the interpretation of the data.

Figure 2 summarizes some of the reported near-field experimental data. Experimental data in different papers were presented in different forms: power, conductance, effective emittance, and heat transfer coefficients. These data are converted into heat transfer coefficients. For sphere-plate experiments, the conversion of conductance and heat transfer coefficient between two parallel plates relies on the proximity-force approximation, also called the Derjaguin approximation, which treats interactions between curved surfaces as locally parallel surfaces. Exact solutions of the problems have shown that this approximation is valid as long as the far-field contributions are properly separated. Figure 2 does not include the conductance data from Refs. 120 and 126, as the former has undefined geometry and the latter is complicated by the large far-field contributions.

Although the existing experiments have clearly established the phenomena of enhanced radiation heat transfer in...
the near-field, there are still many unanswered questions that require further development of experimental techniques and theoretical understanding. In the sphere-plate experiments, as the minimal separation between the sphere and the plate enters into a few to tens of nanometers regimes, surface roughness of the sphere and the plate becomes important. The impact of this roughness on experimental results and data interpretation are not clear. Experimentally, determining accurately the separation between the two surfaces, especially between nonconducting surfaces, is a nontrivial problem that has not yet been solved. In addition to thermal signal, the force interactions between the sphere and the plate can also complicate the experiments. The convolution between the force and thermal signals becomes more severe at small separations and better experiments or new designs of cantilevers are needed to separate the force and thermal signals.

At nanometer separations, the dielectric constant used to describe averaged response of a medium may lose its validity. A few theoretical studies consider the nonlocal description of dielectric constant, although the nonlocal theory on the dielectric function is not well-established. In the limit when two surfaces are in contact with each other, heat transfer is described by the heat conduction picture as discussed in Sec. II A. A complete description of the transition between radiation when the surfaces are nanometers apart to heat conduction when the surfaces are in contact has yet to be accomplished.

To close this section, we discuss the connection between the near-field radiation heat transfer and the Casimir force, a generalization of the van der Waals force between surfaces. As Casimir first pointed out, the van der Waals force between two objects can also be interpreted as the change in the zero point energy stored in the space between the two objects. As the separation between the objects shrinks, the density of states of the electromagnetic wave changes and the energy stored in the space changes, leading to a net force between the surfaces. A generalized theory for the Casimir force was developed by Lifshitz and co-workers based on the fluctuational electrodynamics theory of Rytov. As we discuss in more detail further, the fluctuational electrodynamics theory is also the foundation for near-field heat transfer. The Casimir force is generated by the zero-point fluctuations in the two interacting media and described by the Maxwell tensor, while the near-field radiation heat transfer is generated by the thermal fluctuation and described by the Poynting vector. A large amount of theoretical and experimental work has been done in the context of the Casimir force and the two areas can benefit from each other through increased interactions.

3. Theoretical formalism

The theoretical treatment of near-field electromagnetic thermal transfer aims to calculate the heat flux between two bodies separated by a vacuum gap. (Fig. 3) These bodies are maintained at temperatures $T_1$ and $T_2$, respectively. The term “near-field” refers to the regime where the characteristic width of the vacuum gap is below the thermal wavelength at these temperatures.

At a fundamental level, near-field thermal transfer can be treated using the formalism of fluctuation electrodynamics. In this formalism, the geometries are modeled by a dielectric constant distribution $\epsilon(r, \omega) = \epsilon'(r, \omega) + i\epsilon''(r, \omega)$, where $r$ represents spatial location and $\omega$ is the frequency. The imaginary part of the dielectric constants, $\epsilon''(r, \omega)$, represents material loss. By the fluctuation-dissipation theorem, such material loss is directly connected to the presence of thermally induced current sources $j_x(r, \omega)$ with an ensemble-averaged fluctuation of

$$\langle j_x(r, \omega) j_p(r', \omega') \rangle = \frac{4}{\pi} \omega \Theta(\omega, T) \delta(r-r') \delta(\omega - \omega') \epsilon''(r, \omega) \delta_{\beta \beta},$$

(3)

where

$$\Theta(\omega, T) = \frac{\hbar \omega}{\exp(\hbar \omega / k_B T) - 1}$$

(4)

is the mean thermal energy for a given mode at a frequency $\omega$, $\alpha$, and $\beta$ label different spatial directions. With the source defined, the spectral density of the heat flux is then calculated by integrating the Poynting vector as generated by these fluctuating currents on a plane in the vacuum region (labelled as ‘A’ in Fig. 3) between the two bodies

$$S(\omega) = \frac{1}{2} \Re \int_A \int d\omega' \epsilon^* (r, \omega) \times H(r, \omega').$$

(5)

Here, $\Re$ refers to the real part. The total heat flux is determined as...
As outlined above, the computation of near-field electromagnetic thermal transfer is reduced, in principle, to a sequence of calculations on the electromagnetic radiations from various sources. This process of calculating thermal flux assumes that the material properties are accurately characterized by the dielectric constant of the material. Moreover, the dielectric constant is typically assumed to be local, i.e., the dielectric constant depends on the frequency but not the wave vector of the excitation. One expects that such a local dielectric function should be applicable unless one is in the deep near-field regime where the size of the vacuum gap is reduced to a few nanometer.\textsuperscript{105,142}

The implementation of fluctuational electrodynamics treatment of near-field electromagnetic thermal transfer involves integration over large number of sources and frequencies and thus in general is very demanding computationally. While the basic formalism of fluctuational electrodynamics was developed in the 1950s,\textsuperscript{141} for several decades afterwards, exact calculations of near-field electromagnetic thermal transfer were restricted to planar geometries.\textsuperscript{101,102}

For non-planar geometries, a number of approximate methods have been developed.\textsuperscript{103,132,133,143,144} The dipole approximation was developed to treat the thermal transfer between a deep sub-wavelength nanoparticle and a planar substrate.\textsuperscript{143,144} The proximity approximation, which was also extensively applied, approximate the two bodies in terms of a series of parallel planes.\textsuperscript{132} One should note however that the proximity approximation is an uncontrolled approximation in the sense that there is no systematic process through which one can improve the approximation towards an exact result. Also, the proximity approximation does not reproduce the correct far-field results when the two bodies are far apart from each other.\textsuperscript{133}

4. Exact treatment of near-field thermal transfer in non-planar geometries

One of the most significant developments in the past decade for near-field thermal transfer is the development of techniques that implement fluctuational thermodynamics in a numerically exact fashion in non-planar geometries. Narayanaswamy and Chen solved the thermal near-field radiative transfer between two spheres.\textsuperscript{145} The sphere-plate geometry (Fig. 4), which directly connects to some of the recent experiments, has been solved independently by Krüger et al.,\textsuperscript{146} and by Otey and Fan.\textsuperscript{115} Direct comparison between numerically exact calculations and the experiment as well as several previously used approximations has been done in Ref. 115. The exact calculations did agree with the experiment of Ref. 132, if one were to modify some of the fitting parameters that were used to interpret the experiment. The exact calculations also correctly reproduced the far-field limit when the sphere-plate separation is significantly larger than the thermal wavelength (Fig. 4), as well as the results of the dipole approximation when the sphere is small. There is, however, a significant discrepancy between the exact calculations and the proximity approximation (Fig. 4).

The works of Refs. 115, 145, and 146 all involve the expansion of the fields outside the bodies in terms of either spherical or cylindrical harmonics that are solutions of the individual bodies. Therefore, these approaches are applicable only for relatively simple geometries such as spheres, cylinders, and plates. As an alternative to these approaches, finite-difference time-domain (FDTD) simulations of near-field radiative heat transfer between photonic crystal slabs has been performed in Ref. 147. In this approach, the current-current correlation of Eq. (3) is implemented directly in the simulation by introducing randomly fluctuating currents with the fluctuation of these currents simulated by a random number generator with a correlation determined from the fluctuation dissipation theory. By calculating the flux as emitted from these random sources, and by performing a time average over a long simulation time, as well as an ensemble average over different random source configurations, the thermal flux can be calculated. Boundary element approach has been implemented in Ref. 148, which may significantly reduce the computational cost compared with the FDTD approach, since in the boundary elements one only stores a current distribution on the surface of the body, thus only a two-dimensional discretization is required for a three-dimensional object. This is in contrast with the FDTD methods where storage of fields in full three dimensions are required.\textsuperscript{147}

FIG. 4. Left panel: The sphere-plate geometry. Right panel: Thermal conductance between a 20 μm radius SiO\textsubscript{2} sphere and a SiO\textsubscript{2} substrate. The sphere is at a temperature of 321 K, the plate is at 300 K. The dots are exact numerical results taken from Ref. 115. The solid line is from proximity approximation. The dashed line is the far-field limit.
5. Active modulation and control of near field thermal transfer

An emerging recent direction is the theoretical exploration of active control of near-field thermal transfer. Otey et al. have theoretically shown that significant thermal rectification can be achieved in near-field thermal transfer between two planar bodies of SiC. When the near field of a thermal body, due to electromagnetic surface excitation, the entire thermal electromagnetic field can have a spectrum that consists of a single resonant peak. Otey et al. showed that a strong rectification behavior can be achieved by exploiting the temperature dependency of the resonance frequency of such resonant peaks. Predictions about thermal rectifications using doped silicon, as well as the possibilities of negative differential thermal conductance in near-field thermal transfer, and significant modulation of near-field thermal transfer, have also been made. Further developments of theory will need to take into account the variation of the phonon-polariton line width as function of temperature, which can be significant in a number of polar materials. These works have indicated that many of the thermal circuit concepts, which have been originally explored in phonon thermal transport, may be useful in controlling photon-based thermal transport as well.

6. Image charge theory

Stoner and Maris and Lyeo and Cahill measured the conductance $G$ between various metals and insulators. Some of their values at room temperature are summarized in Table I. Mahan was able to explain these results by calculating the heat flow between a metal and an insulator when their parallel surfaces are in contact. He invoked image charge theory, which is a type of near-field effect. His theory did not account for dielectric matching but instead evaluated the effects of the electric fields on the conduction electrons in the surface layer. His theory used large wave vectors in the metals, small frequencies in both materials, and is not given by dielectric theory.

When the ions in the insulator vibrate, they create local electric fields. At the metal surface, the fields are given by classical image charge theory. These fields interact with the conduction electrons at the surface of the metal. The theory involved only parameters of the insulator such as dielectric constant, ion mass, ion Szigeti charge, and lattice constants. In the metal, the only parameter is electron density, which determines the Fermi wave vector. He calculated the boundary conductance $G$ for several alkali halide crystals, where the metal is aluminum, since all parameters are known. He found room temperature values of $G$ in the range of 40–100 MW m$^{-2}$ K$^{-1}$. These fall in the same range as the values in Table I.

Mahan’s theory does not apply to the case where the insulator is diamond. Diamond is not infrared active, and its vibrations do not create oscillating electric fields. There is, as yet, no quantitative explanation of the experimental results in Table I for diamond as the insulator although three-phonon scattering at the interface may be particularly important at interfaces between materials with highly dissimilar phonon density of states such as Pb/diamond and Au/diamond.

III. THERMAL TRANSPORT IN NANOSCALE STRUCTURES AND DEVICES

A. Physics of phonon transport

In this section, we give a brief review of the basics of phonon propagation and scattering. An understanding of these two topics is essential for the development of a theory of the heat flow in a small structure. To develop a quantitative understanding of the heat flow in a small structure, there are at the present time two rather different approaches. The first is to use some version of molecular dynamics. For a review of some of the variations of this technique, see Ref. 34. This involves, as a first step, the construction of a model for the interatomic potential; this can be based on a fit to measured quantities such as the phonon dispersion relation or it can be calculated from first principles. The interatomic potential has to include terms up to at least third order in the atomic displacements so that anharmonic phonon-phonon interactions can be correctly modeled. Various methods can be used to calculate heat flow. The most straightforward is to connect each end of the nanostructure to heat baths with different temperatures and then calculate the steady state flow of heat from one bath to the other. As an alternative, one can calculate the fluctuating thermal currents in a structure and relate these to the thermal conductivity through some form of the fluctuation-dissipation theorem. (For early work, see Ref. 33).

As a tool for calculating heat flow, molecular dynamics has two limitations. The first problem is that for some nanostructures the number of atoms that need to be included in the simulation may be rather large. As an example, suppose that it is desired to calculate how the thermal conductivity of a nanowire varies with diameter in the range of diameters where there is a crossover from ballistic propagation to diffusive propagation. If the phonon mean free path is 50 nm, it is necessary to consider wires of diameter up to at least 100 nm and lengths of these wires of perhaps 200 nm. In such a volume, they are on the order of 100 million atoms, making a simulation challenging. The second problem is that the calculation is classical. As a result, in the simulation, the average energy of each phonon mode is equal to $k_BT$, whereas the correct energy distribution is given by the Planck distribution. For modes with frequency such that $\hbar\omega < k_BT$ the error is small, but the error becomes large for any solid in which the phonon spectrum extends up to frequencies such

<table>
<thead>
<tr>
<th>Material</th>
<th>Al</th>
<th>Au</th>
<th>Pb</th>
<th>Pb (Ref. 74)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BaF$_2$</td>
<td>100</td>
<td>40</td>
<td>62</td>
<td>—</td>
</tr>
<tr>
<td>Al$_2$O$_3$</td>
<td>105</td>
<td>45</td>
<td>55</td>
<td>28</td>
</tr>
<tr>
<td>Diamond</td>
<td>46</td>
<td>40</td>
<td>31</td>
<td>19</td>
</tr>
</tbody>
</table>

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 18.51.1.88
that $\hbar \omega > k_B T$. It is hard to judge the effect of this error. The fact that the simulation overestimates the specific heat will lead to an overestimate of the conductivity. But it is also true that the error in the average energy of the high frequency modes will result in an overestimate of the rate of phonon-phonon scattering; this leads to an underestimate of the heat flow. For materials such as silicon with a Debye temperature of 600 K, the error may not be too large for simulations at room temperature, but for materials such as diamond ($\Theta_D \sim 2000$ K), carbon nanotubes and graphene, most of the phonon modes have such a high frequency that they are poorly described by classical mechanics. For a recent discussion of attempts to correct molecular dynamics for quantum effects, see Ref. 157.

Molecular dynamics has been used to successfully model the heat flow across interfaces (see Sec. II A) and heat flow in superlattices158–161. For these applications, the method has the great advantage that it is easy to include in the model different aspects of the microscopic structure of an interface, such as the roughness and mixing.165

The second approach is based on the phonon Boltzmann equation166

$$\frac{\partial n(\vec{k}j)}{\partial t} = \frac{\partial n(\vec{k}j)}{\partial t}_{\text{coll}} + \vec{v}(\vec{k}j), \nabla n(\vec{k}j),$$

where $\partial n(\vec{k}j)/\partial t|_{\text{coll}}$ is the rate at which the phonon distribution $n(\vec{k}j)$ changes due to collisions, $\vec{k}$ is the wave vector, $j$ is an index indicating the polarization, and $\vec{v}(\vec{k}j)$ is the group velocity. Since we are primarily interested in heat flow as a result of small and constant temperature gradients (gradients such that the temperature changes by a small amount over a distance of one phonon mean free path, and over a distance equal to the transverse dimensions of the sample), we can use the time-independent and linearized version of this equation

$$\frac{\hbar \omega(\vec{k}j)}{k_B T^2} \tilde{n}(\vec{k}j)(\bar{n}(\vec{k}j) + 1)\vec{v}(\vec{k}j), \nabla T = \frac{\partial n(\vec{k}j)}{\partial t} |_{\text{coll}},$$

where $\bar{n}(\vec{k}j)$ is the equilibrium number of phonons of wave vector $\vec{k}$ and polarization $j$ frequency at temperature $T$. This equation has to be supplemented by appropriate boundary conditions at the surface of the sample. These conditions correspond to diffuse scattering or specular scattering, etc.

To solve the Boltzmann equation one therefore needs to know the phonon dispersion relation, the group velocity, and the rate at which collisions occur. The phonon dispersion relation is known from neutron scattering for most materials of interest. Frequencies at special points in the Brillouin zone can be measured by light scattering. Recently, inelastic X-ray scattering has also become useful for measuring phonon frequencies.167 This has the important advantage that X-rays from a synchrotron source can be focused into a very small volume, thereby enabling small samples to be used. The phonon group velocity can be obtained from the dispersion relation by

$$\vec{v}(\vec{k}j) = \frac{\partial \omega(\vec{k}j)}{\partial \vec{k}}.$$

The variation of group velocity with wave vector is complicated; there are focusing directions in which the group velocity is stationary with respect to variations in $\vec{k}$ and so there are a large number of phonons which propagate in nearly the same direction.168–170 But to evaluate the collision term is much harder. It is important to note that to construct the collision term of the Boltzmann equation it is not sufficient to calculate just the lifetime $\tau(\vec{k}j)$ of each phonon $\vec{k}j$. A collision of a phonon $\vec{k}j$ with another phonon inevitably destroys not only the phonon $\vec{k}j$ but also another phonon. In addition, a new phonon is produced. Thus, one needs to know not just the lifetime but also the collision kernel $C$ defined such that

$$\frac{\partial n(\vec{k}j)}{\partial t} |_{\text{coll}} = -\frac{n(\vec{k}j)}{\tau(\vec{k}j)} + \sum_{\vec{k}j'} C(\vec{k}j; \vec{k}j') n(\vec{k}j') d^3 \vec{k}j'.$$

The kernel is not symmetric with respect to the interchange of $\vec{k}j$ with $\vec{k}j'$; instead,171 it can be shown that

$$C(\vec{k}j'; \vec{k}j) = C(\vec{k}j; \vec{k}j') \frac{n(\vec{k}j') [n(\vec{k}j') + 1]}{n(\vec{k}j)[n(\vec{k}j) + 1]}.$$  

To calculate the rate of anharmonic processes involves the following steps:

1. A model for the harmonic part of the interatomic potential has to be constructed. This can be done using first-principles theory or more simply by choosing a model for the interatomic forces, which gives reasonably good agreement with the measured dispersion relation.

2. One then needs a model for the anharmonic part of the potential. For the purpose of calculating three-phonon processes, this needs to contain terms that are cubic in the atomic displacements. Again, this part of the potential can be obtained from first-principles or by choosing a model that gives agreement with experimentally measured parameters. These parameters might include, for example, the third order elastic constants (known accurately for many materials), and the thermal expansion coefficient and its variation with temperature. In addition, for many materials, experimental values exist for the variation with applied elastic strain of the frequency of zone center optical phonons. (See, for example, the early work on Si and GaP by Weinstein and Piermarini.173)

3. From these two models, the matrix elements and the rate of three-phonon scattering can be calculated. These matrix elements involve not just the anharmonic potential and phonon wave vectors but also the polarization vectors of the interacting phonons. The calculation of the matrix elements and the collision kernel $C$ is straightforward but complicated.174–176 Because $C$ involves the polarization vectors (quantities which are not directly measurable), it is not easy to estimate the accuracy of the calculation.

The general formula for the phonon lifetime is given in a classic paper by Maradudin and Fein.177 A phonon $\vec{k}j_1$ can
be scattered by a phonon \( \vec{k}_{3j} \) to produce another phonon \( \vec{k}_{j1} \). In this process, energy and crystal momentum are conserved so that

\[
\hbar \omega(\vec{k}_{j1}) + \hbar \omega(\vec{k}_{3j}) = \hbar \omega(\vec{k}_{3j})
\]

\[
\vec{k}_1 + \vec{k}_2 = \vec{k}_3 + \vec{G},
\]

where \( \vec{G} \) is a reciprocal lattice vector. The phonon \( \vec{k}_{j1} \) can also decay into phonons \( \vec{k}_{3j} \) and \( \vec{k}_{j1} \) in which case

\[
\hbar \omega(\vec{k}_{j1}) = \hbar \omega(\vec{k}_{3j}) + \hbar \omega(\vec{k}_{j1})
\]

\[
\vec{k}_1 = \vec{k}_2 + \vec{k}_3 + \vec{G}.
\]

Processes in which \( \vec{G} = 0 \) are called Normal (N) processes, and those for which \( \vec{G} \neq 0 \) are Umklapp (U) processes. Umklapp processes are of special importance because these collisions directly reduce the heat flux. Early work on phonon lifetimes was summarized in review articles by Klemens,\textsuperscript{178} and Carruthers,\textsuperscript{179} and in a book by Ziman.\textsuperscript{180} Much of this work focused on determining approximate expressions for the phonon lifetime in the limit of low or high frequency, or when the temperature was much less than or greater than the Debye temperature.

Some important results are the following:

1. When the temperature is zero, the phonon lifetime is always limited by boundary scattering. It is also limited by the spontaneous anharmonic decay process governed by the conservation laws (Eqs. (12) and (13)). For longitudinal and fast transverse acoustic phonons, the rate \( \tau^{-1} \) of this process varies as \( \omega^5 \).\textsuperscript{181} this assumes that the wave number is small compared to the wave number at the Brillouin zone boundary. A rough estimate of the magnitude of this rate for a number of crystals is given in Ref. \textsuperscript{182}, and more detailed calculations have been performed by Tamura and Berke et al.\textsuperscript{184} For slow transverse phonons, the rate varies as \( \omega^5 \) when the wave vector lies in some directions, but in other directions, the decay rate is zero because there is no solution to the conditions of conservation of energy and momentum.\textsuperscript{185–188}

2. A discussion of the lifetime of optical phonons at zero temperature has been given by Klemens.\textsuperscript{189}

3. For temperatures much less than the Debye temperature \( \Theta_D \) and provided that \( \omega \) is much less than the Debye frequency, the phonons have wavelength which is much larger than the lattice parameter. The calculation of the phonon lifetime is then simplified because the anharmonic coupling can be expressed in terms of the third order elastic constants. These constants have been measured for many crystals of current technical interest. The phonon lifetime varies with frequency \( \omega \) and temperature \( T \) according to the law

\[
\tau^{-1} \propto T^5 f(\hbar \omega/k_BT).
\]

For transverse phonons with \( \omega \ll k_BT/\hbar \), \( \tau^{-1} \) is proportional to \( \omega T^2 \). Herring\textsuperscript{190} has shown that for longitudinal phonons, again with \( \omega \ll k_BT/\hbar \), \( \tau^{-1} \propto \omega^3 T^{5-a} \), where \( a \) is an integer, which is larger than 1 and dependent on the symmetry of the crystal. The lifetime is determined by the collisions of the longitudinal phonon with a transverse phonon and producing a transverse phonon of different polarization. These collisions are called Herring processes. For cubic crystals, the inverse lifetime can be calculated in closed form as\textsuperscript{191}

\[
\tau^{-1} = \frac{3 \xi(3) \hbar k^2}{4\pi^2 \rho c^2} \left( \frac{k_BT}{\hbar} \right)^3 C_{11} + C_{12} \frac{1}{C_{11} + 2C_{12} + C_{44}} \left( \frac{k_BT}{\hbar} \right)^{-1} \\
\times \left[ \{1 - 3G\} C_{11}^2 [E(R) - (1 - R)K(R)] + 4G C_{11}^2 [K(R) - E(R)] \right],
\]

where \( \xi(3) \) is the Riemann zeta function, \( C_{ij} \) are second order elastic constants, \( c \) is the sound velocity of transverse waves in the (100) direction, and \( C_1 \) and \( C_2 \) are related to the second and third order elastic constants by

\[
C_1 = C_{155} - C_{144}, \quad C_2 = C_{456} + C_{444}.
\]

\( K(R) \) and \( E(R) \) are complete elliptic integrals of the first and second kind with

\[
R = (C_{11} - C_{44})(C_{11} + 2C_{12} + C_{44})/(C_{11} + C_{12})^2,
\]

and \( G = \rho (m^2 + n^2 + l^2) \) where \( l, m, n \) are the direction cosines of \( \vec{k} \).

4. For temperatures greater than \( \Theta_D \) the inverse lifetime for all phonons is proportional to temperature. This result holds provided that three-phonon processes are the most important; at sufficiently high temperatures (for example, near the melting point), there may be a significant contribution from four-phonon processes.

5. In addition to these general results, phonon lifetimes have been calculated for the nearest-neighbor central force model of a fcc crystal with one atom per unit cell. This model has the advantage that the harmonic and anharmonic forces are each described by only one parameter. These parameters can be related to the Debye temperature and the Grüneisen constant \( \gamma \). The phonon lifetime based on this model was first considered by Maradudin et al.,\textsuperscript{192} and later investigated in more detail.\textsuperscript{193} The inverse lifetime can be expressed as

\[
\frac{1}{\tau(kj)} = \frac{k_BT^2 \rho^{5/3}}{M^{5/3}} G_{\text{total}},
\]

where \( \rho \) is the mass density, \( M \) is the mass of each atom, and \( G_{\text{total}} \) is a dimensionless function of wave vector, temperature, and phonon polarization. As an example of the results obtained, Fig. 5 shows \( G_{\text{total}} \) for longitudinal phonons propagating along the [100] direction for five different values of the wave number. In the low temperature limit such that \( \hbar \omega \gg k_BT \), the inverse lifetime is determined by spontaneous quantum decay; if the frequency is much less than the Debye frequency but still satisfying \( \hbar \omega \gg k_BT \), the rate is proportional to \( \omega^5 \) at high temperatures, the decay rate is proportional to temperature and increases monotonically with increasing wave number. The results in Fig. 5 are for the total
FIG. 5. Calculated value of the parameter $G_{\text{tot}}$ giving the lifetime of longitudinal phonons propagating in the [100] direction as a function of temperature using the results of Ref. 194. These results are for a simplified model of a fcc crystal. $k_{\text{max}}$ is the phonon wave number at the zone boundary.

decay rate, i.e., they include both Umklapp and normal processes. It turns out that the fraction $f_U$ of the scattering events, which are U-processes, varies with wave number in a surprisingly complex way (Fig. 6). At the zone boundary ($k = k_{\text{max}}$), there is no difference between a U and an N-process so $f_U = 0.5$. U-processes can occur when $k = 0.8k_{\text{max}}$ or $0.2k_{\text{max}}$, but there is a range of $k$ between these values in which U-processes cannot occur. The reason for this is discussed in Ref. 193. Because of these complexities, it is difficult to judge the validity of calculations based on highly simplified expressions for the lifetime of phonons such as Klemens’ result Eq. (32) that is discussed in Sec. IV. Certainly, that formula does not agree with the exact results for an fcc crystal shown in Fig. 5. It seems likely that within the next few years, as the number of first-principles calculations based on the Boltzmann equation increases, we will have a better understanding of the usefulness of the approximate formulas.

6. It is important to note that all of the results discussed so far assume implicitly that we are interested in the lifetime of a phonon of frequency $\omega$ such that $\omega \tau > 1$, where $\tau$ is the average lifetime of the thermal phonons. To understand this, first note that for a very low frequency phonon, the lifetime is dominated by collisions with thermal phonons (see Eq. (12)); decay processes (Eq. (13)) are negligible. The finite lifetime of phonons $\tau_{j\beta}$ and $\tau_{j\beta}$ makes their energy uncertain by amounts of $\hbar / \tau_{j\beta}$ and $\hbar / \tau_{j\beta}$, respectively. If the energy uncertainty is comparable to or larger than the energy of phonon $k_{\text{max}}$ the conservation laws break down. One can show that under these conditions, the inverse lifetime varies as $\frac{1}{\tau} \propto \omega^{2} \tau$; this is referred to as the Akhiezer regime. In this regime, one can consider that the inverse lifetime $\tau^{-1}$ is the sum of two processes, both of which can be viewed in macroscopic terms. The first is the so-called thermoelastic damping. As the low frequency phonon propagates, it expands and contracts the lattice resulting in temperature fluctuations. The flow of heat between the heated and cooled regions results in entropy production and a damping of the phonon at a rate

$$\tau^{-1}_{\text{thermo}} = \frac{e_{x}(k)}{\rho v_{x}^{2} C}$$

where $e_{x}$ are the elements of the phonon polarization vector, $\kappa_{ijkl}$ are the elements of the thermal conductivity tensor, $C_{ijkl}$ are the second order elastic constants in tensor notation, and $C$ is the specific heat per unit volume. The second mechanism is by phonon viscosity, also called the Akhiezer process. The fluctuating strain associated with the low frequency phonon raises the frequency of some of the high frequency ($\omega \sim k_{\beta}T / \hbar$) phonons and lowers the frequency of others. This creates a non-equilibrium phonon distribution which produces entropy as it relaxes. The damping rate by this process is

$$\tau^{-1}_{\text{vis}} = \frac{e_{x}(k)}{\rho v_{x}^{2} C}$$

where $v_{x}$ are the elements of the phonon viscosity tensor. Both of these processes give a damping proportional to $\omega^{2} \tau$, but generally the viscosity contribution is larger than the thermoelastic term. The damping rate in the transition regime where $\omega \tau$ goes from less than one to greater than one is complicated, and is discussed in Refs. 197 and 198. It is important to note that these are two distinct regimes in which separate formulas apply. Thus, in the $\omega \tau > 1$ regime, one should calculate the damping rate using the phonon collision approach; a contribution from the thermoelastic and Akhiezer mechanisms should not be added in.
The important scattering process in crystals which are free of defects is isotope scattering. In a cubic crystal with a single atom per unit cell, it is easy to show that the phonon lifetime is given by

\[ \frac{1}{\tau(\vec{k})} = \frac{\pi}{6} V_0 g_2 \omega^2(\vec{k}) D(\omega(\vec{k})) \],

(21)

where \( V_0 \) is the volume per atom, and \( g_2 \) is given by

\[ g_2 = \sum_i f_i \left( \frac{\delta m_i}{\langle m \rangle} \right)^2, \]

(22)

where \( i \) labels the different isotopes, \( f_i \) is the fraction of isotope \( i \), \( \delta m_i \) is the deviation of the mass of isotope \( i \) from the average mass \( \langle m \rangle \). \( D \) is the phonon density of states per unit volume. It can be shown that the same formula also holds for silicon and germanium. Isotope scattering has some similarities to the Barkla effect for X-rays. The scattering from state \( \vec{k} \) to \( \vec{k}' \) is proportional to \( |\vec{\varepsilon}(\vec{k}) \cdot \vec{\varepsilon}(\vec{k}')|^2 \). Thus, for example, for scattering at an angle of \( \pi/2 \) a longitudinal phonon \( \vec{k} \) is most likely to be scattered into a transverse phonon with polarization vector in roughly the same direction as \( \vec{k} \).

Isotope scattering is another example of a process for which it would be helpful to have a better understanding of the usefulness of simplified formulas. If we compare Eq. (21) with the Klemens’ formula Eq. (33) discussed in Sec. IV, one can see immediately that the difference between Eqs. (21) and (33) is simply that the phonon density of states has been approximated by the Debye model, i.e., has been taken to be proportional to frequency squared. In fact, of course, the density of states \( D(\omega) \) goes as \( \omega^2 \) only at low frequencies. At larger \( \omega \), \( D(\omega) \) usually contains strong peaks, which come from regions of the Brillouin zone where the phonon group velocity is small. In the frequency range of these peaks, the true value of \( D \) is much larger than the Debye estimate. Thus, one can see that Eq. (33) underestimates the scattering rate of the phonons with low group velocity and probably overestimates the scattering of phonons with less dispersion.

Once \( C \) is calculated, the final step is to solve Eq. (7) to find the phonon distribution. The heat current \( j \) can then be obtained from

\[ j = \sum_{\vec{k}j} n(\vec{k}j) \vec{v}(\vec{k}j) \hbar \omega(\vec{k}j). \]

(23)

From the discussion above, it is clear that even for a bulk crystal the calculation of the thermal conductivity is challenging. One way to simplify the problem is to ignore the collision kernel and write

\[ \frac{\partial n(\vec{k}j)}{\partial t} \bigg|_{coll} = -\frac{n(\vec{k}j)}{\tau(\vec{k}j)}. \]

(24)

This leads to the “relaxation time approximation” for the thermal conductivity tensor

\[ \kappa_{\alpha\beta} = \sum_{\vec{k}j} c(\vec{k}j) v_\alpha(\vec{k}j) v_\beta(\vec{k}j) \tau_{\vec{k}j}, \]

(25)

where \( c(\vec{k}j) \) is the contribution to the specific heat from the phonon mode \( \vec{k}j \). For a cubic crystal, the conductivity simplifies to become

\[ \kappa = \frac{1}{3} \sum_{\vec{k}j} c(\vec{k}j) \tau_{\vec{k}j}. \]

(26)

There are several well known difficulties with the use of this formula; an old but still very useful discussion of some of the problems is given in Ziman’s book. The major difficulty with this formula is that it is not clear whether every type of scattering process should be included in the calculation of \( \tau_{\vec{k}j} \). Anharmonic processes of the N type do change the direction of the scattered phonon but leave the total flux of energy in a given direction essentially unchanged. Thus, one can argue that these processes do not contribute to the thermal resistance and so should be ignored and only U processes considered. But suppose that there is a particular group of phonons which cannot be scattered in a U process. Then, if N processes are ignored, the thermal conductivity would be infinite, which is clearly impossible. To treat this situation correctly, it is necessary to go beyond the relaxation time approximation and include the effects of both N and U processes in an appropriate way.

To overcome these difficulties, quantitative calculations using the full Boltzmann equation have been made. In 1996, Omini and Sparavigna reported calculations for Ar and Kr using an accurately calculated collision kernel together with a solution of the Boltzmann equation. For these materials, the interatomic potential is relatively simple. More recently, Broido and coworkers have published a series of papers in which they report calculations of this type for silicon, germanium, and diamond. These calculations are considerably more difficult because of the more complicated form of the interatomic potential. The results are in good agreement with experiment.

Of course, in the numerical calculations of both Sparavigna and of Broido et al., the information about the lifetime and the collision kernel is contained within the program. It would be very valuable to extract this information and in this way learn more about the detailed phonon processes that take place when heat flows. Which phonons carry most of the heat? How does the phonon lifetime vary with wave vector and polarization? What is the ratio of normal processes to Umklapp processes and which phonons are involved in U-processes? Sparavigna does not provide a plot of this information, but Ward and Broido do present a plot of the lifetime of longitudinal phonons at 300 K averaged over all directions (Fig. 7). They also divide the contributions to the lifetime into U and N-processes. Lindsay and Broido have found correlations between the region of phase space available for three-phonon processes and the thermal conductivity. Study along these lines has the potential to give an indication about the validity of some of the approximations which are often made in theories of the bulk thermal conductivity. For example, it is usually assumed that although the optic modes may affect the heat flow by contributing to the scattering rate of the acoustic phonons, the actual heat carried by these modes is unimportant. It would be good to know how reliable an approximation this is.
To calculate the heat flow in a simple nanostructure, such as a nanowire composed of a single material, it is sufficient to use the same phonon lifetime and collision kernel as for a bulk material. This approach will break down only when the lateral dimensions of the structure become comparable to the phonon wavelength. However, the calculation of heat flow is much more difficult than for the bulk because now the phonon distribution function $n(\tilde{k} j)$ varies across the transverse dimensions of the sample. In a cylindrical nanowire, for example, the phonon distribution varies with distance $r$ from the axis, and the solution for the phonon distribution function has to satisfy the Boltzmann equation at each value of $r$. Calculations for a diamond nanowire have recently been reported; it was necessary to make some approximations in order to reduce the computer time that was needed. In a single-walled carbon nanotube (SWCNT), the situation is simpler, since the phonons only exist at one value of $r$. Calculations of the thermal conductivity in a SWCNT and in graphene have been performed by Lindsay et al. They do not give details of the phonon lifetimes but do discuss the relative importance of N and U processes.

Of course, if the mean free path is very large compared to all sample dimensions (the Casimir regime) the problem is tractable. It is then sufficient to simply add up the contribution from each of the phonon modes which are thermally excited. For each mode, one can calculate the distance $\Lambda_{\tilde{k} j}$ that a phonon leaving a point on the surface of the structure travels in the direction of heat flow before being scattered at another point on the surface; this distance depends on the particular point from which the phonon starts. For a uniform rod of arbitrary cross-section, one can show that the conductivity is given by the formula

$$\kappa = \frac{h^2}{16 \pi^2 k T^2 A} \int dl \sum_j \int \int \bar{v}_{\tilde{k} j} \cdot \bar{v}_l n_{\tilde{k} j} \times (n_{\tilde{k} j} + 1) \Lambda_{\tilde{k} j}^2 \hbar \omega_{\tilde{k} j} k^2 dk dc \cos \theta d\phi, \quad (27)$$

where $A$ is the cross-sectional area, the line integral runs around the boundary of the rod, and $\bar{v}_l$ is a vector normal to the surface at the position $l$. Note that the contribution from a mode depends, amongst other things, on the square of $\Lambda_{\tilde{k} j}^2$. A change in modes for which the group velocity vector lies in a direction close to the direction of heat flow clearly make a large contribution. In fact, one can show that this contribution is so large that for heat flow in the in-plane $x$-direction of a plate of thickness in the $y$-direction $D$ and width in $z$ of $nD$, the thermal conductivity diverges logarithmically as $n$ goes to infinity. Quantitative calculations of the Casimir regime heat flow in cylindrical nanowires and plates of Si and GaAs allowing for the details of the phonon dispersion relation and the variation of the phonon group velocity with wave vector and polarization have been presented in Ref. 206.

For a nanostructure with a complex geometry and composed of more than one material, new issues arise. As an example, consider a superlattice with repeat distance $d$. If the phonon mean free path in the bulk materials is larger than $d$, then it is necessary to first calculate the phonon dispersion relation in the superlattice and then calculate the phonon lifetime and collision kernel. Ward and Broido have calculated the thermal conductivity of a series of Si/Ge and GaAs/AlAs superlattices. For the GaAs/AlAs structures, they find a theoretical conductivity substantially higher than the experimentally measured thermal conductivity. They attribute this to the effect of roughness of the interfaces.

B. Nanowires and graphene: Experiment and theory

Over the past decade, the establishment of several nano-scale thermal transport measurement techniques has made it possible to use nanostructures as model systems to test the phonon transport theories reviewed in Sec. III A. Compared to bulk crystals, phonon transport in nanostructures is unique in the importance of phonon-boundary scattering as well as confinement of phonons in low-dimensional structures such as nanotubes and graphene. In the following, the assumptions behind the particle transport theories of phonon-boundary scattering are examined, followed by a discussion of a number of phonon wave interference models that have emerged over the past decade for explaining the thermal conductivity of some nanostructures. Besides the central issue regarding whether wave interference effects are observable in the thermal conductivity measurements, the size dependence and interface scattering of phonon transport in one-dimensional (1D) nanotubes and two-dimensional (2D) graphene are reviewed.

1. Particle transport description of phonon-boundary scattering

Casimir employed a radiation analogue and the Debye approximation to analyze phonon transport in a rod in the low temperature limit where internal scattering of phonons was negligible. The rod surface was treated as a black-body that absorbed all the incident phonons and re-emitted phonons diffusely according to the local equilibrium distribution. Under these assumptions, the average phonon-
boundary scattering mean free path $\Lambda_B$ is the diameter $d$ for a circular rod, and $1.12W$ for a rod with a square cross-section of width $W$.

Building upon Casimir’s result\textsuperscript{208} and several other studies,\textsuperscript{209–211} Ziman\textsuperscript{180} developed a solution of the phonon Boltzmann transport equation (BTE) in a rod under the relaxation time approximation. (Although the Debye approximation was mentioned in Ziman’s derivation,\textsuperscript{180} the assumption of an isotropic crystal is sufficient to arrive at his result without the need of assuming a frequency-independent phonon group velocity.) For diffuse boundary condition at the rod surface and absence of internal scattering, the boundary mean free path given by the BTE solution\textsuperscript{180} is the same as Casimir’s result ($\Lambda_{B,C}$), and is essentially an average over all orientations and positions of the distance traveled by phonon quasi-particles originated from the cross section and intercepted by the rod surface. For a partially specular and partially diffuse surface, the specular parameter $p$ is used to describe the probability for specular scattering at the surface. Under the isotropic assumption, the phonon BTE can be solved to obtain the boundary scattering mean free path as $\Lambda_B = \Lambda_{B,C}(1 + p)/(1 - p)$\textsuperscript{180}.

Because the phase velocity is not collinear with the group velocity in an anisotropic crystals, McCurdy et al.\textsuperscript{207} calculated that the $\Lambda_B$ of a Si rod of diffuse boundary condition ranges between $1.66\Lambda_{B,C}$ and $0.97\Lambda_{B,C}$ when the rod axis is aligned along the $\langle 100 \rangle$ and $\langle 111 \rangle$ direction, respectively, and obtained low-temperature thermal conductivity data that revealed this phonon focusing effect.

With internal scattering due to phonon-phonon processes or impurities accounted for, the BTE can be solved under the isotropic assumption to yield that the total mean free path ($\Lambda$) differs only slightly from that given by the Matthiessen’s rule, $1/\Lambda = 1/\Lambda_{B} + 1/\Lambda_{i}$, where $\Lambda_{i}$ is the mean free path due to internal scattering processes.$^{180,210}$ Wang and Mingo\textsuperscript{212} recently presented BTE solutions of $\Lambda$ for nanowires of different cross sections and 2D nanoribbons with two diffuse edges, and verified the accuracy of the Matthiessen’s rule only for the nanowire case, because diffuse boundary scattering along does not lead to a finite mean free path in 2D nanoribbons, similar to the case of the logarithmic divergence for an infinite thin film discussed at the end of Sec. III A.

These particle transport theories have been used to analyze the thermal conductivity data of nanostuctures measured by a number of methods that have been established over the past decade.$^{8,9,12}$ In addition to the axial thermal conductivity of different nanowire structures, the thermal conductivity of thin films,$^{213}$ and periodic nanoporous membrane structures,$^{214,215}$ along the in-plane direction have been obtained with the use of suspended micro-devices with built-in resistance thermometers.$^{216}$ TDTR techniques have also been employed to measure the cross-plane thermal conductivity of nanowire arrays.$^{217,218}$ The theoretical framework of Casimir and Ziman can explain well the suppressed thermal conductivity found in InAs nanowires,$^{217,219}$ Bi$_2$Te$_3$ nanowires,$^{220}$ SnO$_2$ nanobelts,$^{212}$ and Si nanowires$^{222}$ grown by a vapor liquid solid (VLS) method with a diameter larger than about 30 nm, as well as Bi nanowires,$^{223}$ InSb nanowires,$^{224}$ and SiGe nanowires$^{225}$ when additional impurity and defect scattering were considered.

Figure 8 includes a comparison between the measured thermal conductivity of some Si and SiGe nanowires and two calculated thermal conductivity results based on Callaway-type models, Matthiessen’ rule, bulk phonon dispersion, and diffuse phonon-surface scattering. Because the thermal conductivity of bulk Si is dominated by phonons with mean free paths longer than 100 nm, the size effect on the thermal conductivity can be readily observed in Si nanowires. Because the relative thermal conductivity contribution of low-frequency phonons with long mean free paths is decreased by diffuse boundary scattering, the relative contribution from the high-frequency or short-wavelength zone boundary phonons increases in a nanowire. Consequently, the calculated thermal conductivity becomes more sensitive to the choice of the cutoff frequency in the Callaway-type model for nanowires than for bulk crystals.$^{226}$ As shown in Fig. 8, the use of the same bulk cutoff frequency given by Morelli et al.$^{227}$ combined with diffuse phonon-boundary scattering yields approximately the experimental thermal conductivity values reported by VLS Si nanowires with a diameter larger than 30 nm.$^{222}$

Using a cutoff frequency adjusted to fit the data for one nanowire, Wang and Mingo$^{228}$ obtained somewhat lower thermal conductivity values for the diffuse boundary condition. Based on the results of Wang and Mingo,$^{228}$ phonons scattered by the surface of the VLS Si nanowires of diameter larger than 30 nm should contain a small specular component together with a large diffuse component. In addition, the low-temperature thermal conductivity data measured by Heron et al.$^{229}$ for patterned Si nanobeams were explained by an increased specular component with increasing phonon wavelength or decreasing temperature. By accounting for alloy...
scattering of short-wavelength phonons and diffuse boundary scattering of all phonons, the calculation by Wang and Mingo\textsuperscript{228} was used to fit the experimental thermal conductivity values reported for Si$_{1-x}$Ge$_x$ nanowires\textsuperscript{223} which suggested that alloy scattering instead of surface scattering should be the dominant scattering mechanism. However, without considering scattering by high-concentration defects, impurities or their clusters, these Callaway-type models suggest that the Casimir limit of diffuse surface scattering alone is insufficient to suppress the thermal conductivity to the experimental values reported for electroseeded rough Si nanowires\textsuperscript{230} as shown in Fig. 8.

Other types of scattering phase function at the surface can in principle lead to shorter $\lambda_0$ than the diffuse boundary condition. For example, the Monte Carlo (MC) phonon particle transport simulation result of Moore \textit{et al.}\textsuperscript{233} suggested that adding periodic sawtooth features of a sufficiently large height to width ratio to a Si nanowire of $22 \times 22$ nm$^2$ cross section can suppress the thermal conductance despite the increase in the solid volume of the nanowire. This effect has been attributed to the presence of a large backscattering component of the scattered phonons relative to the phonons incident on the sawtooth features. However, the room-temperature thermal conductivity obtained for the $22 \times 22$ nm$^2$ nanowire with 4-nm-high, 2-nm-wide sawtooth features is still four times larger than the reported experimental value shown in Fig. 8 for the ~50 nm diameter rough Si nanowire. In addition, a scattering phase function with a large component normal to the nanowire axis can also reduce the average $v_x^2$ to be smaller than the $|v_g|^2 / 3$ value for the diffuse scattering case in an isotropic crystal, where $v_x$ and $v_g$ are the group velocity and its component along the nanowire axis. However, within the particle transport description used in the relaxation time approximation or MC solution of BTE, no mechanisms have been found to yield a backscattering or normal scattering component that is sufficiently large to produce the experimental thermal conductivity values reported for rough Si nanowires.

2. Wave interference models of phonon transport in silicon nanostructures

The seeming discrepancy between the experimental data and the phonon particle transport theories has stimulated a number of investigations of different wave interference effects that are not accounted for by the particle description of phonon transport. For example, it has been suggested that phonon interference can result in phononic band gaps and reduce the phonon group velocity, which has been used to explain the experimental values of the in-plane thermal conductivity of Si nanomeshes.\textsuperscript{214} The reported room-temperature upper bound value of 2 W m$^{-1}$ K$^{-1}$ for the Si nanomeshes is about a factor of two lower than the already surprisingly small thermal conductivity of $\sim$3.5 W m$^{-1}$ K$^{-1}$ for patterned Si nanowires of a comparable critical dimension of $\approx$20 nm.\textsuperscript{214} If the measurement result of the Si nanomesh of periodic holes was analyzed with the use of effective medium theories\textsuperscript{234,235} that do not account for any size effects on the solid thermal conductivity, the obtained room-temperature solid thermal conductivity would be as low as about 1.4 W m$^{-1}$ K$^{-1}$, comparable to those reported for amorphous Si and SiO$_2$.\textsuperscript{236}

It is worth noting that the reported thermal conductivity of the nanomesh increases with temperature at near room temperature. However, if reduced phonon group velocity was the only reason for the very low thermal conductivity of the nanomesh, the room-temperature thermal conductivity would be expected to decrease with increasing $T$ because of increasing umklapp scattering.

Moreover, confinement of phonon waves along the radial direction in a small-diameter nanowire can in principle modify the phonon dispersion because of interference of the waves reflected by the nanowire surface. Some earlier theoretical works suggested that the phonon group velocity component along the nanowire axis of a 20-nm-diameter Si nanowire is reduced to be about half of the $v_g$ value in the bulk\textsuperscript{237} suppressing the thermal conductivity. However, the axial component of the phonon group velocity in a nanowire is equivalent to the bulk $v_x$ instead of the bulk $v_g$. For Si nanowires of a diffuse surface and a diameter of 20 nm and larger, Mingo\textsuperscript{226} calculated that the phonon dispersion is rather close to the bulk dispersion and the room-temperature thermal conductivity calculated based on the actual dispersion are similar to those calculated based on the bulk phonon dispersion, because the energy separation between 1D phonon sub-bands in the nanowire is much smaller than $k_BT$ for temperature $T > 40$ K in a 20 nm Si nanowire.

In addition, variation of the cross section due to the surface roughness of a nanowire can in principle result in a perturbation in the allowed wave vector states ($q_{nm}$) perpendicular to the transport direction in the nanowire, because $|q|/d(x)$ scales with $\pi/d(x)$ with $d(x)$ being the local nanowire diameter. This modulation is proportional to $\sigma/D$, where $\sigma$ is the root mean square (rms) height of the surface roughness and $D$ is the average nanowire diameter. The effect is equivalent to a perturbation of the local phonon dispersion $\omega(q)$ along the nanowire. Martin \textit{et al.}\textsuperscript{238} employed first order perturbation theory, also referred as the Born approximation, to calculate the volumetric scattering rate caused by this perturbation. Their results\textsuperscript{238} suggest that the thermal conductivity in the surface roughness scattering-dominant regime is proportional to $(D/\sigma)^2$.

Sadhu and Sinha\textsuperscript{239} employed a different Bourret approximation and the Green’s function approach to calculate the phonon transmission coefficient in 50-nm-diameter Si nanowires of different $\sigma$ and $L_r$ values. Besides the $\sigma/D$ ratio, their result suggested that the $\sigma/L_r$ ratio is another critical parameter. By setting $\sigma = 2.5$ nm and $L_r = 30$ nm to increase the $\sigma/L_r$ ratio, Sadhu and Sinha\textsuperscript{239} obtained a transmission coefficient that decreases with increasing frequency to be well below the Casimir limit, and used the result to match the experimental data for rough Si nanowires with ~50 nm diameter.

A different atomistic Green’s function calculation by Carrete \textit{et al.}\textsuperscript{240} suggested that the transmission coefficient can be suppressed to be one order of magnitude smaller than that given by Casimir’s limit only for a very large $\sigma/D$ value exceeding 20%. These authors\textsuperscript{240} emphasized that the Born approximation is invalid when the surface roughness is
on the order of wavelength and can underestimate the transmission coefficient by one order of magnitude.

It is worth noting that the surface amorphous layer on a nanowire was shown to considerably reduce the thermal conductivity of a 15 nm diameter Si nanowire in a molecular dynamics simulation and in an experiment with Si crystals. The presence of an amorphous native oxide on the surface of Si nanostructures can destroy the phase coherence of phonons transmitted into the amorphous shell and back into the crystalline nanowire core.

A recent MD simulation has suggested that interior defects such as grain boundaries are needed to obtain the abnormally low thermal conductivity in the experimental Si nanowires. While it is unclear how grain boundaries could have formed in the nanowires etched from single crystal silicon, it has been found that the electroless etching process can create nanopores in nanowires etched from highly doped Si wafers. In addition, RIE was shown to cause damage in a 30-nm-thick surface layer of Si. Although nanopores were not observed by transmission electron microscopy (TEM) in nanowires etched from low-doped wafers, TEM is effective in detecting extensive defects such as grain boundaries or pores but not local defects such as point defects, impurities and their clusters.

Hence, further experimental and theoretical studies are needed to clarify the cause of the unusually low experimental thermal conductivity of Si nanostructures. The recent combined structure-thermal transport measurement of Si nanowires has represented an important progress toward this direction. It is found that the thermal conductivity of the wires has represented an important progress toward this combined structure-thermal transport measurement of Si nanostructures. The recent combined structure-thermal transport measurement of Si nanostructures has represented an important progress toward this direction. It is found that the thermal conductivity of the wires has represented an important progress toward this direction.

3. Phonon scattering in graphene and nanotubes

Graphene and carbon nanotubes are unique in their atomically smooth surfaces and confinement of phonons to the atomic scale along one or two dimensions. These distinct features can also be found in the hexagonal boron nitride counterparts. The phonon dispersion of suspended single-layer graphene (SLG) is different from that of graphite in the absence of the cross-plane modes. Appreciable difference also appears in the basal-plane modes between graphene and highly oriented pyrolytic graphites (HOPG), also referred to as Bernal or 3D graphite. In addition to the approximately linear longitudinal acoustic (LA) and in-plane transverse acoustic (TA), the out-of-plane polarized, in-plane propagating transverse acoustic modes, the so-called flexural or ZA modes, show a quadratic dispersion in SLG. Such quadratic ZA dispersion can also be found in graphite except in the long-wavelength limit, where interlayer interaction breaks the rotational symmetry and the energy cost for rotation of the graphene sheet results in a linear term in the ZA dispersion that dominates over the quadratic term. Although the acoustic branches of graphite are similar to those in SLG, and the low-frequency longitudinal optical (LO) and transverse optical (TO) branches of HOPG are nearly degenerate with the LA and TA branches, the additional low-frequency out-of-plane optical (ZO') branch in HOPG lies appreciably above the ZA branch. Because the cross-plane modes and ZO' modes can scatter the basal-plane acoustic modes, the absence of these scattering channels should result in higher basal-plane thermal conductivity of suspended flat SLG than that of graphite.

Moreover, because of the reflection symmetry in flat suspended SLG, three-phonon scattering processes involving the ZA branch must involve an even number of ZA phonons. The selection rule restricts the phase space for ZA phonon scattering. Because of the reduced ZA scattering rate and the large density of states of the ZA phonons, it was suggested that the ZA contribution dominates the thermal conductivity of flat suspended SLG, contradicting earlier calculations that did not use this selection rule or calculate the actual phase space for phonon-phonon scattering. Together with scattering by cross-plane and ZO modes, it was suggested that breaking of the reflection symmetry in few-layer graphene (FLG) relaxes the selection rule, decreasing the basal plane thermal conductivity to approach the graphite limit when the FLG layer thickness is increased to be above five layers.

When a graphene sheet is rolled up and joined into a single-walled CNT, the curvature can also break the reflection symmetry, and increase the ZA phonon scattering rate. It has been reported that this effect results in a decreasing thermal conductivity when the SWCNT diameter is decreased to a critical value, 1.5 nm for a 3-μm-long long SWCNT. For diameters below this critical value, increasing confinement of the phonon waves along the circumference of the SWCNT increases the energy separation between 1D phonon subbands and pushes the optical modes to higher frequencies. Consequently, scattering between the optical modes and the acoustic modes is suppressed in small-diameter SWCNTs, increasing the thermal conductivity with decreasing diameter. Besides the FLG thickness and SWCNT diameter dependence, it has also been suggested that the thermal conductivity of SLG and SWCNTs increases with the lateral dimension, because first-order three-phonon processes are insufficient to prevent the divergence of the thermal conductivity contribution from low-frequency acoustic modes that play a larger role in lower dimensions. The predicted size dependence is for the case of clean, suspended defect-free graphene and nanotubes, because supporting or embedding graphene and nanotubes on a substrate or in a medium may cause interface phonon-scattering as well as breaking of the reflection symmetry. In addition, higher-order phonon-phonon processes and scattering by defects, isotope impurities, and other structure disorders can also remove the predicted length dependence.

Over the past decade, the establishment of a number of nanoscale thermal transport measurement methods has enabled the experimental investigation of phonon transport in nanotubes and graphene, as discussed in several recent reviews. Results from a micro-Raman spectroscopy measurement have been used to support the theoretical
prediction of increasing basal plane thermal conductivity with decreasing FLG layer thickness and suggest that the thermal conductivity approaches the graphite value when the FLG layer thickness increases to about eight layers,\textsuperscript{255} as shown in Fig. 9. Other Raman measurements\textsuperscript{256–259} yielded lower optical absorbance and thermal conductivity values for suspended SLG than the initial work,\textsuperscript{59} as shown in Fig. 10. In addition to large uncertainties in the optical absorbance, the limited temperature sensitivity of the Raman thermometry method has prevented experimental verification of the lateral size dependence of the thermal conductivity of suspended graphene.\textsuperscript{258} It is also necessary to verify the reported thickness dependence with other experimental techniques.

Different designs of micro-devices with built-in resistance thermometers have been employed to observe decreasing thermal conductivity of FLG supported on or embedded in SiO\textsubscript{2} with decreasing thickness, as shown in Fig. 9. Raman spectra obtained on the supported SLG\textsuperscript{260} and FLG\textsuperscript{262} samples do not contain D peaks associated with defects, with a large electron mobility found in the supported SLG sample.\textsuperscript{260} In addition, the experimentally determined acoustic phonon dispersions of a graphene monolayer weakly interacting with a substrate\textsuperscript{264,265} are similar to those of suspended SLG except for ZA phonons near the zone center. Hence, the thermal conductivity suppression found in the supported graphene has been attributed to phonon scattering by the graphene-SiO\textsubscript{2} interface. Compared to the results reported for suspended FLG, the opposite trend is found for supported and embedded FLG and indicates much stronger interface scattering than interlayer scattering within the FLG. The results from the supported FLG further suggests that the interface scattering can impact the transport of phonons in graphene sheets more than tens of layers away from the interface. A key question is whether the strong interface scattering is simply caused by the roughness of the SiO\textsubscript{2} surface, or largely due to transmission or leaking of the scattered phonons across the interface.\textsuperscript{253}

C. Applications to Information Technology and High-Power Density Electronics

Modern semiconductor devices and their constituent materials and interfaces present some of the most promising and challenging opportunities for the study of nanoscale heat conduction. Example technologies and devices include highly scaled planar or 3-dimensional (3-D) field-effect transistors,\textsuperscript{266,267} phase change memory devices based on chalcogenides,\textsuperscript{268,269} photonic crystal lasers and waveguides for optical interconnects,\textsuperscript{270,271} and composite substrates for power semiconductor devices.\textsuperscript{272–274} The temperature distributions in these structures, especially near interfaces, consume an increasing fraction of the available “thermal budget” for electronic, optoelectronic, and amplifier circuitry. While a sub-continuum analysis is not always essential for the accurate simulation of these structures, this situation is gradually changing with continued dimensional scaling, heterogeneous materials integration, and geometries targeting electron and photon band manipulation. Section III C summarizes three areas of special relevance and proposes future nanoscale experiments and simulations.

1. Data Storage

While heat generation and transport in logic\textsuperscript{266,275} and interconnects\textsuperscript{276,277} have received much attention, thermal
phenomena in data storage technologies are also very relevant due to leakage in volatile memory and to the sheer number of memory elements employed. In this context, phase change materials (PCMs) have attracted wide interest for memory and reprogrammable circuits with low voltage operation, fast switching, and high endurance.268,269 PCMs are chalcogenides like Ge2Sb2Te5 (GST) which have amorphous and crystalline phases with a large ratio (>100×) of electrical resistivity. The data in electrically programmable PCMs are stored as changes in bit resistance, which can be reversibly switched with short voltage pulses and localized Joule heating. However, a drawback of PCMs has been their high programming current (>0.1 mA), as Joule heat must be delivered to a finite bit volume.

Figure 11 shows the distribution of amorphous and polycrystalline material in a phase change random access memory (PCRAM) cell that has been subjected to varying levels of programming current. The key metrics for contemporary devices include the switching energy and time, which are governed by heat conduction and crystallization kinetics in regions of GST alloys with dimensions below 10 nm. For exploratory phase change devices, such as those developed for brain-inspired computing, nanoscale heat conduction governs the staged transition from the amorphous to the crystalline phase during pulsed electrical heating. The temporal stability of the amorphous region is another challenging research topic, which is critical for advanced cell designs including those targeting multibit operation.

Thermal transport studies in PCMs examined the impact of interface properties, phase purity, and the relative contributions of electrons and phonons to transport.284 The temperature-dependent thermal conductivities of the amorphous and polycrystalline face centered cubic (fcc) and hexagonally close packed (hcp) phases of GST films and several stoichiometric variants were measured using time-domain thermoreflectance and nanosecond transient thermoreflectance.285 Electrons contribute ∼1 Wm−1K−1 to the hcp crystalline phase and help to nearly quadruple the thermal conductivity relative to the amorphous and fcc phases. By exploring stoichiometric variations including pure GeSb, the electron contribution can be varied by nearly a factor of 5. Electrothermal measurements using suspended nanostructures and the 3ω method identified strong anisotropy in mixed phase films, which was attributed to the alignment of amorphous material along vertically oriented grain boundaries.288 Thermal property data remain urgently needed for the melted condition.

Thermal boundary resistances (TBRs) of PCMs with surrounding passivation and electrode materials (e.g., SiO2, SiNx, and TiN) are also key for the energy-efficient operation of a memory cell.289 The TBRs of Ge2Sb2Te5 are difficult to predict owing to the comparable contributions of both electrons and phonons to heat flow in the hcp phase and the large mismatch of acoustic properties with respect to common metals and dielectrics. Early work identified a strong reduction of TBR with disordered ZnS:SiO2 films due to transition to the crystalline phase.290 More recent work measured the relatively large TBR with respect to device electrode materials, and simulated their large impact on the writing current for a vertical cell geometry.291,292

Thermoelectric transport can have a large impact on temperature fields in PCM nanodevices and can strongly influence the write current.293 For the reset process to the amorphous phase, Peltier heating at the electrode interface and Thomson heating or cooling in the crystalline phase of the chalcogenide material can govern the size of the amorphous region resulting from a given current pulse. Nanoscale electrothermal simulations are complicated by size and temperature dependence of thermoelectric properties of Ge2Sb2Te5. Experimental data exhibit differing magnitudes
and temperature derivatives for the Seebeck coefficient as the material is converted from the amorphous to the crystalline phase through annealing.\textsuperscript{280,294} Furthermore, when the material is predominantly crystalline, a strong thickness dependence of the Seebeck coefficient may be linked to the variation of the phase purity with film thickness for a given annealing history.

Recent work has also explored the fundamental scaling limits of PCRAM using electrodes made from carbon nanotubes rather than metals.\textsuperscript{295–297} CNTs can have a higher electrical conductivity and can be more reliable than metals down to diameters of only \( \approx 2 \) nm; thus, they can form electrodes contacting PCM bit volumes of just tens of cubic nanometers. Figure 12 shows such a device,\textsuperscript{296} where the CNT is “cut” in the middle, forming a small nanogap (\( \approx 10 \) to 200 nm). Sputtering \( \approx 10 \) nm GST\textsuperscript{295} fills the nanogap, forming the extremely small PCM bit with two CNT electrodes. Such devices switch with programming currents of \( \approx 1 \) \( \mu \)A, two orders of magnitude lower than the previous state-of-the-art,\textsuperscript{268,296} and programming energy \(< 100 \) fJ per bit. Testing hundreds of such structures\textsuperscript{296} has uncovered trends suggesting that PCRAM could be scaled down to achieve sub-fJ per bit operation in sub-5-nm bits. Thus, the fundamental operating limits of PCMs have not yet been reached, and these limits are set by phase transformations and heating at single nanometer dimensions.

At these scales, the influence of even subtle atomistic changes (e.g., grain boundaries, and stoichiometry) on the thermophysical properties of such materials are unknown, and their contact resistances (electrical and thermal) with the surrounding materials play a very important role. Moreover, the ultimate packing density of such bits in memory arrays is limited by thermal disturb between adjacent bits,\textsuperscript{298,299} and their contact resistances (electrical and thermal) with the thermophysical properties of such materials are unknown, changing (e.g., grain boundaries, and stoichiometry) on the phase transformation and heating at single nanometer dimensions.

Recent work has also explored the fundamental scaling limits of PCRAM using electrodes made from carbon nanotubes rather than metals.\textsuperscript{295–297} CNTs can have a higher electrical conductivity and can be more reliable than metals down to diameters of only \( \approx 2 \) nm; thus, they can form electrodes contacting PCM bit volumes of just tens of cubic nanometers. Figure 12 shows such a device,\textsuperscript{296} where the CNT is “cut” in the middle, forming a small nanogap (\( \approx 10 \) to 200 nm). Sputtering \( \approx 10 \) nm GST\textsuperscript{295} fills the nanogap, forming the extremely small PCM bit with two CNT electrodes. Such devices switch with programming currents of \( \approx 1 \) \( \mu \)A, two orders of magnitude lower than the previous state-of-the-art,\textsuperscript{268,296} and programming energy \(< 100 \) fJ per bit. Testing hundreds of such structures\textsuperscript{296} has uncovered trends suggesting that PCRAM could be scaled down to achieve sub-fJ per bit operation in sub-5-nm bits. Thus, the fundamental operating limits of PCMs have not yet been reached, and these limits are set by phase transformations and heating at single nanometer dimensions.

FIG. 12. (a) CNT electrodes have been used to contact the smallest and most energy-efficient phase-change material memory devices to date, see Refs. 295–297. The CNT spans two Pd contacts on an oxidized Si wafer. A nanogap is formed electrically or mechanically with the tip of an atomic force microscope. Amorphous Ge\textsubscript{2}Sb\textsubscript{2}Te\textsubscript{5} (a-GST) is sputter-deposited to fill the nanogap, and voltage pulses are used to reversibly switch its state. (b) AFM image of experimental device (scale bar is 500 nm); zoom-in shows the top surface of the PCM bit. (c) Experimental current-voltage (I–V) curves reveal switching at \(< 1 \) \( \mu \)A, two orders of magnitude lower than previous state-of-the-art PCM devices.\textsuperscript{296} VT marks the threshold voltage of the amorphous state. (d) Simulation of nanogap bit cross-section confirms that small dimensions are essential for achieving low-power operation.\textsuperscript{297} TBR: Seebeck effects, and non-equilibrium transport of both electrons and phonons are crucial for energy-efficient operation at sub-10 nm dimensions and sub-10 ns time scales.

Another memory technology for which thermal transport plays a central role is heat-assisted magnetic recording (HAMR), which involves near-field radiation, ballistic thin-film conduction, interface resistance, and nanoscale lubricant transport.\textsuperscript{300} Near-field radiative heating, intensified and localized through plasmon excitation, must heat a \( \approx 10 \) nm spot in the recording medium above its Curie temperature (generally above \( \approx 500 \) °C) in \( \approx 1 \) ns. The recording medium consists of metal and dielectric nanolayers of thickness 10–40 nm along with the functional magnetic film. The effective thermal conductivities of this stack are highly anisotropic and governed by electron and phonon scattering and energy conversion. One challenge is to optimize the complex multilayer system, considering both photon absorption and transient, anisotropic thermal conduction, such that the lateral dimensions of the hotspot are minimized. Recent HAMR progress includes detailed electromagnetic simulations of the near-field transducer, studies of the interacting thermal and magnetic properties of the recording medium,\textsuperscript{301} approximate analyses of the combined optical and thermal transport in the medium,\textsuperscript{302} and even studies of the heated lubricant flow in the gap below the head.\textsuperscript{303} A systematic characterization of HAMR materials for optical and thermal performance remains important for future work.

2. Nanoscale CMOS and interconnects

Heat conduction in electronics begins with the energy dissipated in nanoscale transistors and interconnects.\textsuperscript{275} Complementary metal-oxide-semiconductor (CMOS) field-effect transistors (FETs) are the building blocks of modern electronics, typically based on Si technology with features down to \( \approx 28 \) nm.\textsuperscript{304} However, scaling devices to sub-10 nm dimensions will likely involve new materials from groups III–V such as GaAs, InAs, or InSb,\textsuperscript{305} or group IV such as Ge or C,\textsuperscript{306,307} which have electronic properties (e.g., mobility)
superior to Si. Further scaling will also involve new geometries such as multi-gate (FinFET) devices, nanowires, or nanotubes.

In all cases, power dissipation begins with electrons or holes that are accelerated by high electric fields, of the order $\sim 10 \text{ V}/\mu\text{m}$. These “hot” carriers scatter with the lattice, emitting acoustic and optical phonons (OPs). Modern devices operate at voltages near $\sim 1\text{ V}$, strongly emitting OPs which have energies $\sim 35 \text{ meV}$ in Ge or GaAs, $60 \text{ meV}$ in Si, 180 meV in CNTs and graphene. Typical OP lifetimes are of the order $\sim 1 \text{ ps}$, however, the thermal time constants of CMOS devices are of the order $\sim 10 \text{ ns}$ due to the larger volumetric heat capacity. It is worth noting that the cooling time of a single carbon nanotube is $\sim 0.1 \text{ ns}$, likely representing the smallest thermal time constant achievable in a CMOS device. The thermal time constant of a nanoscale device represents how quickly it cools down or heats up after an electrical transient, and such temperature excursions determine both the device leakage and long-term reliability.

Once a nanoscale device reaches thermal steady state, the pathway of heat conduction to the environment is key in determining its temperature rise and ultimately its reliability. The temperature rise of a modern device can reach $\Delta T > 100 \text{ K}$ above room temperature, being determined by its power input $P$ and the thermal resistance $R_{TH}$, as $\Delta T = PR_{TH}$. For instance, the thermal resistance of a single device on the surface of a Si wafer is approximately $R_{th} \approx 1/(2\kappa_S D)$, where $D \approx (LW)^{1/2}$, $L$ and $W$ are the length and width. However, as devices are scaled down, four additional heat flow considerations must be taken into account. First, heat conduction into contacts or the gate electrode could play a greater role. Second, novel devices (e.g., FinFET, nanowire, and CNT) contain many interfaces between the semiconductor body where dissipation takes place, and the external heat sink. Thus, the associated thermal boundary resistances must be carefully taken into account. Third, nanoscale devices are smaller than the bulk phonon mean free path, $\Lambda \approx 100 \text{ nm}$ in Si and comparable in most semiconductors. Thus, the thermal conductivity is lower than well-known bulk values, for instance, that of a $\sim 20\text{ nm}$ thin silicon-on-insulator (SOI) film is $\kappa_{SOI} \approx 25 \text{ W m}^{-1}\text{K}^{-1}$ compared to the thermal conductivity of bulk Si, $\kappa_S \approx 150 \text{ W m}^{-1}\text{K}^{-1}$. Fourth, modern devices have electrical switching transients of the order 10–100 ps, ranging between those of carrier scattering events (0.1–1 ps) and the thermal relaxation of the lattice ($\sim 10\text{ ns}$). Thus, nanoscale devices operate in a complex temporal fashion, where non-equilibrium situations can arise between the electrons, the OPs, and the lattice whose heat capacity is primarily determined by acoustic phonons.

Figure 13 illustrates the various heat conduction mechanisms in transistor-like or interconnect-like nanoscale devices. For simplicity, the figure ignores heat loss through a top gate, which is discussed later in Fig. 14. The key questions answered here concern whether heat dissipation from a modern nanoscale device takes place along the channel, through the contacts, or via the substrate. The various heat flow paths are denoted by thermal resistances $R_L$ along the device, $R_C$ through the contacts, $R_{ox}$ for the supporting oxide (of an SOI-like device), $R_{Si}$ for the silicon substrate, and $R_B$ for the channel-oxide thermal boundary resistance. The figure also clarifies the relative roles played by intrinsic thermal properties (such as thermal conductivity of the channel and
substrate) versus extrinsic properties (such as the contacts and TBR), and how these change as the device length ($L$) and width ($W$) are scaled toward sub-10 nm dimensions.

Numerical calculations in Fig. 13 are based on the models in Refs. 317–319, and correspond to a graphene device on SiO$_2$ with thickness $t_{\text{ox}} = 50$ nm, graphene-SiO$_2$ boundary resistance $R_B = 2 \times 10^{-8}$ m$^2$K$^{-1}$W$^{-1}$, and graphene thermal conductivity $\kappa_g = 150$ Wm$^{-1}$K$^{-1}$, consistent with graphene nanoribbons or encased graphene. The numerical analysis is similar for an ultra-thin SOI device with thickness $t_{\text{SOI}} = 5$ nm and thermal conductivity $\kappa_{\text{SOI}} = 10$ Wm$^{-1}$K$^{-1}$ (same $tk$ product). Thus, the trends observed in the figures are also typical for other modern and future nanoscale transistors, including ultra-thin body SOI, multi-gate (FinFET), CNT, or nanowire FETs.

Figure 13(a) shows a cross-section of the device along its length, and Fig. 13(b) shows the cross-section across its width. The heat conduction pathways can be both “vertical” into the substrate and “lateral” along the device and into the contacts, as shown by red arrows. Figure 13(c) displays the total thermal resistance $R_{\text{TH}}$, in units of temperature rise per unit input power, of a typical device as a function of channel length $L$; in this example, for a device width $W = 4L$. The thermal resistance of “longer” devices is dominated by the vertical heat flow path which includes the TBR ($R_B$), the thermal resistance of the oxide ($R_{\text{ox}}$), and, to a smaller degree, the Si substrate ($R_S$). However, the thermal resistance of “shorter” devices (particularly with $L < 100$ nm) becomes limited by heat flow along the channel and into the contacts ($R_C + R_{\text{contact}}$). The cross-over of these two regimes occurs at a device length that is approximately three times the thermal healing length ($L_{\text{TH}}$) along the channel. The results of Fig. 13(c) are consistent with experimental data summarized in Fig. 5(a) of Ref. 275.

The fractional contribution of each thermal pathway to the total thermal resistance is shown in Fig. 13(d) as a function of channel length $L$. For long-channel devices, the thermal resistance is approximately 30% due to TBR ($R_B$) and the rest dominated by the thermal resistance of the underlying oxide ($R_{\text{ox}}$) for the combination of parameters used in these estimates. (The relative importance of TBR would be larger if thinner underlying oxides are used.) The contact thermal resistance ($R_C$) becomes increasingly important in sub-100 nm devices, almost regardless of the thermal properties of the device or substrate.

Figures 13(e) and 13(f) show the fractional contributions of various thermal resistances as a function of device width $W$, for “short” ($L = 100$ nm) and “long” devices ($L = 1$ nm), respectively. In both cases, we see the increasing relative importance of the TBR between the device and the substrate as device widths are narrowed. However, for longer devices, this TBR component can almost entirely dominate as device widths approach $W \sim 1$ nm, which has been experimentally confirmed for the case of CNTs.

We also examine the role of top gates in heat sinking from transistors in Fig. 14. Finite-element simulations suggest that top gates are not significant heat sinks, in part due to the additional TBRs associated with the channel-dielectric and dielectric-gate interfaces, and in part because the main heat conduction pathway remains into the substrate and towards the back-side heat sink. This can be seen as the top gate assumes nearly the same temperature as the channel in Fig. 14. However, top gates and the dielectric capping layers covering them add thermal capacitance near the device channel, thus lengthening the device thermal time constants. For instance, the thermal time constant of the graphene device is approximately $50$ ns without a capping layer, and $>100$ ns with the capping layer shown (here taken as $200$ nm of SiO$_2$). These results illustrate that the exact heat dissipation picture is complex and must often be evaluated for the particular, 3-D device designs.

Unlike transistors, nanoscale interconnects are typically metals like Cu with liners or interfaces of TaN, TiN, or W. In these materials, both electrons and phonons can contribute to heat conduction, while the interfaces, grain boundaries (GBs), and line edge roughness (LER) play an increasingly important role as dimensions are reduced. Importantly, both the magnitude of thermal conductivity ($\kappa$) and the relative contributions of electrons and phonons ($\kappa = \kappa_e + \kappa_p$) are modified by GB and LER scattering. For instance, the electrical conductivity ($\sigma$) of Cu nanowires decreases by $\sim 5 \times$ from the bulk value at linewidths of $\sim 10$ nm.

However, data for Pt films and nanowires reveal that $\sigma$ and $\kappa$ do not always scale proportionally, as would be predicted by the Wiedemann-Franz Law (WFL) in metals. Platinum films with thickness down to 7.3 nm fabricated using atomic layer deposition (ALD) show a stronger reduction of $\sigma$ (by $4.5 \times$) than of $\kappa$ (3.4 $\times$) with respect to bulk values. The authors suggested that the excess thermal conductivity in the Pt films, above that predicted by the WFL, may be due to phonon transport. However, electrochemically deposited Pt nanowires with diameter $\sim 100$ nm reveal a weaker decrease of $\sigma$ (by $2.5 \times$) than of $\kappa$ (3.4 $\times$), which suggests that the WFL overpredicts the electron contribution.
to the thermal conductivity. This result is surprising because the WFL is generally assumed to work well for boundary scattering owing to the elastic nature of this process. These studies suggest that electron transport at nanoscales in metals is governed by subtle differences in sample and interface characteristics. Similar observations have been made by recent thermal measurements of graphene nanoribbon (GNR) interconnects, where simulations have also pointed out that LER scattering is likely to affect phonons more strongly than electron conduction in narrow GNRS with zigzag edges. Much research remains to be carried out to understand the relative contributions to heat conduction of electrons and phonons at sub-10 nm dimensions.

3. Composite substrates for power electronics

Nanoscale heat transfer can be particularly important in semiconductor devices that deliver or modulate power. Examples include semiconductor lasers, power field-effect switches, and radar amplifiers, which can have power densities exceeding 1 kW/cm². Improvements have been particularly rapid for quantum cascade lasers (QCL), which achieve efficient, tunable mid-infrared wavelength emission through quantum confinement in GaInA/AlInAs multilayers. The challenge for materials science is clear: strongly reduce defect concentrations within nanometers of a complex interface.

Figure 15 and Table II summarize the impact of the thermal resistances in GaN composite substrates on their relative heat spreading capabilities, as well as the phonon scattering mechanisms that contribute to the thermal resistance. To significantly outperform GaN-SiC, the GaN-diamond interface resistance must approach the DMM predictions under the Debye approximation, which is near 1 m² K GW⁻¹. While the Debye approach yields some error due to phonon dispersion, the challenge for materials science is clear: strongly reduce defect concentrations within nanometers of a complex interface.

The challenge for the nanoscale thermal transport community is to measure and identify the specific microstructural causes for resistances at interfaces roughly a micrometer below the surface. The effective interface resistance sums the components at the discrete material interfaces with the effects of defect scattering in the two primary substrate materials and the transition film. Diamond composite substrate fabrication procedures include direct growth of the semiconductor on diamond, in some cases with a transitional layer, and growth of the diamond on the semiconductor or a very thin adhesion film. There has been some progress on relating the concentrations of point defects and dislocation segments in the transition region to metal-organic chemical vapor deposition (MOCVD) growth conditions of AlN/GaN on SiC. Future work for diamond composites should relate concentrations of point defects and dislocation segments in the transition region to metal-organic chemical vapor deposition (MOCVD) growth conditions of AlN/GaN on SiC.

<table>
<thead>
<tr>
<th>TABLE II. Summary of the minimum thermal resistances observed in several studies of composite substrates and comparisons to the prediction of the DMM.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>GaN-diamond</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>GaN-SiC</td>
</tr>
<tr>
<td>Diamond-Si</td>
</tr>
</tbody>
</table>
specific types of defects and impurities to the thermal resistance. The optimal fabrication procedures for GaN-diamond, i.e., those that yield the lowest defect concentrations in the semiconductor and the lowest thermal resistance at its interface with diamond, remain undetermined. For the case of diamond growth on the functional semiconductor the transition film, defects in the diamond can impair heat spreading. There has been progress on measuring and modeling the impact of phonon scattering on defects near the deposition interface with silicon, and optimization of nucleation parameters to minimize the effective interface thermal resistance.\(^\text{341}\)

IV. THERMAL CONDUCTIVITY OF NANOSTRUCTURED MATERIALS

In this section, we review progress in nanostructuring materials for low thermal conductivity applications, and materials having low thermal conductivity as a result of their intrinsic atomic structure and interatomic bonding. The former includes nano-grain sized materials as well as those doped to create point defect scattering and second phases by thermal treatment. The latter includes crystal structures that constitute natural superlattices with superlattice periodicities of the order of nanometers.

In addition to their intrinsic scientific interest, materials having low thermal conductivity are important for several technologies, and for understanding larger scientific issues. For instance, there is a world-wide search for oxides with low thermal conductivity at high temperatures for use as thermal barriers for gas turbines;\(^\text{342}\) the thermal conductivity having low thermal conductivity are important for several technologies, and for understanding larger scientific issues. In identifying oxide materials with unusually low thermal conductivities at high temperatures as well as predicting the lowest attainable thermal conductivity of a compound under extreme burn-up, guidance comes from two closely related models based on classic kinetic theory. The first, proposed by Kittel in 1949 (Ref. 344), pertains to amorphous solids and argues that the length scale associated with the atomic disorder determines the mean free path for phonon scattering. The other, the Cahill and Pohl model,\(^\text{346}\) describes the minimum thermal conductivity of a crystal when the mean-free path approaches the inter-atomic spacing and is usually described in terms of the longitudinal and transverse sound velocities, and, respectively in the high temperature limit

$$\kappa_{\text{min}} = 0.40 k_B \left( \frac{N}{\Omega_0} \right)^{2/3} (v_{L,s} + 2v_{T,s}).$$  \hspace{1cm} (28)

Here, $\Omega_0$ is the unit cell volume and $N$ is the number of atoms per unit cell. At the minimum thermal conductivity, one can view heat transport in the spirit of an Einstein oscillator model as a process involving the hopping diffusion of energy from one atom to a nearest neighbor. This limiting behavior can occur at high temperatures even in crystalline materials as well as in glasses.

For crystalline materials below temperatures at which the minimum thermal conductivity, $\kappa_{\text{min}}$, is reached, the general principles underlying the effects of many microstructural effects on phonon-mediated thermal transport were largely worked out in the 1950s and 1960s by Klemens,\(^\text{345,346}\) Callaway,\(^\text{347,348}\) and others.\(^\text{180}\) These analyses led to expressions for the thermal conductivity in terms of the phonon relaxation times for different phonon scattering processes integrated over the phonon energy distribution. Assuming a Debye density of states distribution, the thermal conductivity described by Eq. (26) can be written as

$$\kappa = \frac{1}{2\pi^2 v_D} \int_0^{\omega_D} C_v(\omega)\tau(\omega)\omega^2 d\omega,$$

\hspace{1cm} (29)

where $\omega_D$ is the Debye frequency, $v_D$ is the Debye velocity of sound, and $C_v(\omega)$ is the frequency dependent heat capacity. In practice, the mean-free-path for phonon scattering cannot be smaller than approximately half the phonon wavelength and, in turn, smaller than the inter-atomic spacing, as pointed out by Roufosse and Klemens.\(^\text{349,350}\) This requires the integral to be split into two parts

$$\kappa = \frac{1}{2\pi^2 v_D} \int_0^{\omega_D} C_v(\omega)\tau(\omega)\omega^2 d\omega + \frac{1}{2\pi^2 v_D} \int_{\omega_D}^{\infty} C_v(\omega)\tau(\omega)\omega^2 d\omega,$$

\hspace{1cm} (30)

where the second term is, in essence, the same temperature-independent minimum conductivity described in greater detail by Cahill and Pohl.\(^\text{346}\) In materials in which multiple phonon scattering processes occur, the contributions from each process are added inversely through Matthiessen’s rule\(^\text{345,351,352}\)

$$\frac{1}{\tau} = \frac{1}{\tau_u} + \frac{1}{\tau_d} + \frac{1}{\tau_b} + \cdots.$$  \hspace{1cm} (31)

Klemens\(^\text{345}\) derived specific scattering times associated with phonon-phonon Umklapp scattering $\tau_u$, phonon scattering from point defects $\tau_d$, dislocation strain fields $\tau_s$, and grain boundaries $\tau_{GB}$

$$\frac{1}{\tau_u} = A\omega^2 T^3 \exp(-B/T),$$

\hspace{1cm} (32)

$$\frac{1}{\tau_d} = V_d c_i \left( \frac{m_i - m_{\text{ave}}}{m_{\text{ave}}} \right)^2 \frac{\omega^4}{4\pi v^4},$$

\hspace{1cm} (33)

$$\frac{1}{\tau_s} = \frac{2}{3} N_d b^2 \gamma^2,$$

\hspace{1cm} (34)

$$\frac{1}{\tau_{GB}} = \frac{d}{v_s}.$$  \hspace{1cm} (35)

Equation (32) is a commonly used approximation of the anharmonic scattering rates that are discussed in detail in Sec. III A, in particular, see Eq. (18) and Fig. 5. Equation (6) is an approximation of Eq. (21) that assumes a Debye density of states.

There are several important results pertinent to nanostructured materials that follow directly from the Klemens-Callaway model\(^\text{348}\) in Eq. (30) and the form of the relaxation
times. First, in the absence of any defects and at intermediate temperatures, below 2 to 5 times the Debye temperature, \( T_D \), the thermal conductivity decreases approximately as \( 1/T \) and asymptotes to a temperature independent value given by \( \kappa_{\text{min}} \). Microstructural defects, including point defects, anti-site defects, and grain boundaries, all decrease the scattering time with contributions that are independent of temperature. As the thermal conductivity due to anharmonic phonon scattering decreases with increasing temperature, all forms of microstructural defects have a greater apparent effect in lowering thermal conductivity at lower temperatures than at the higher temperatures, decreasing the observable temperature dependence. In the extreme, when the minimum conductivity is reached, they can result in a temperature-independent conductivity over a very large temperature range. We will return to this point later in discussing grain size effects and high-defect concentrations.

### A. Doping and point defects

As Eq. (33) shows, the effect of doping depends on the square of the normalized mass mismatch of the dopant and the host lattice ions. This effect has been widely investigated for low concentrations of dopants for many years and is clearly manifested in a number of materials systems. The thermal conductivity of natural diamond (containing 1.1\% of \(^{13}\text{C}\)) at room temperature is \( \approx 2000\, \text{W m}^{-1}\, \text{K}^{-1} \). This isotopic concentration, and high normalized mass mismatch, has a large impact on the thermal conductivity: diamond, isotopically enhanced such that it contains only 0.001\% \(^{13}\text{C}\) has a considerably higher thermal conductivity. Isotopic effects have little effect for higher mass systems. For example, natural uranium is an isotopic mixture of 99.3\% \(^{235}\text{U}\) and 0.7\% \(^{238}\text{U}\); atomic-level simulations of isotopic effects in uranium dioxide, \( \text{UO}_2 \), fuel show that this leads to only \( \approx 4\% \) decrease in the thermal conductivity. Turning to off-stoichiometry, \( \text{UO}_2 \) can easily accommodate oxygen interstitials compensated by \( \text{U}^{5+} \) ions. As Fig. 16 shows, such off-stoichiometry can result in a significant reduction in the thermal conductivity. Very similar quantitative behavior is observed in the \( \text{Gd}_{8-x}\text{Ca}_{x}\text{SiO}_{4} \) \( x=0,2,3+\) apatite system where ranges of solid solution with off-stoichiometry accommodated by either oxygen vacancies, oxygen interstitials or a combination can be produced by varying the \( \text{Gd}:\text{Ca} \) ratio.

At higher doping and solute concentrations, the details of phonon scattering are more complex. A case in point is the decrease in thermal conductivity of yttria-stabilized zirconia (YSZ) with doping. Without any yttria, zirconia exhibits the typical \( 1/T \) dependence of many other oxides; however, as the yttria doping concentration is increased, its conductivity decreases and the temperature dependence becomes weaker. The yttrium ions substituting on the \( \text{Zr} \) cation sub-lattice do not themselves introduce appreciable point defect scattering because their atomic mass is similar to that of \( \text{Zr} \); however, the oxygen vacancies they create do, just as do \( \text{O}_2 \) in \( \text{UO}_2 \). At 7–8 wt. \% yttria, 3YSZ, which is the prototypical material for thermal barrier applications, has a tetragonal crystal structure and its conductivity is almost temperature independent. For fully dense 3YSZ at 1373 K, a typical thermal barrier coating (TBC) operating temperature, the thermal conductivity is 2.5 W m\(^{-1}\) K\(^{-1}\). Deliberately introduced porosity further reduces this to 1.5–2.0 W m\(^{-1}\) K\(^{-1}\), depending on the deposition method used. Also, stabilization of the tetragonal phase by oxygen vacancies has the highly desirable effect of imparting high fracture toughness as a result of stress-induced ferroelasticity. Analysis of the phonon properties on both \( \text{UO}_2 \) (Ref. 354) and YSZ (Ref. 360) showed that chemical disordering through the presence of high densities of point defects means that for most vibrational modes, it is not possible to clearly define a polarization or a wave-vector, and hence consider simple Rayleigh scattering. Rather, from a heat transport perspective, the thermal-transport properties systems behave in a similar manner to structurally amorphous materials.

### B. Grain boundaries

For many materials, the simplest approach to nano-structuring is to reduce grain size. Grain boundaries act as obstacles to the movement of phonons through a material by scattering phonons (Eq. (35)), and lower the thermal conductivity by decreasing the time between phonon scattering events. A simple two-state model for the overall thermal conductivity of a polycrystal in terms of the thermal conductivity of the bulk and of a grain-boundary region leads to

\[
\kappa(T) = \frac{d\kappa_i(T)}{d + R_K \kappa_i(T)},
\]

where \( R_K \) is the interfacial (Kapitza) resistance, \( \kappa_i \) is the thermal conductivity of the bulk and \( d \) is the grain size. Figure 17 shows the effects of grain size for \( \text{SrTiO}_3 \). There is a substantial reduction in the thermal conductivity as the grain size decreases, particularly at lower temperatures, at which the anharmonic phonon-phonon scattering effects are less significant, consistent with the expectations of the Debye model discussed above. Also, as expected, the difference...
between the conductivities at high temperatures becomes less marked for larger grain sizes. Similar results have been obtained with yttria-stabilized zirconia.\textsuperscript{368} This is an important result since it means that grain size manipulation for grain sizes in excess of $\sim$1\,$\mu$m is a much less effective strategy for decreasing thermal conductivity for high temperature applications than for low-temperature applications.

While grain growth at high temperatures is familiar, decreases in grain size can also occur during the operational lifetime of a material. For example, microstructural evolution of UO$_2$ during in-reactor burn-up involves the development of very small grain sizes (averaging $\sim$130 nm, but with a significant fraction less than 100 nm)\textsuperscript{363} and high porosity in the outer regions of the fuel pellet, i.e., the “rim effect.”\textsuperscript{343,364} The combination of high GB area, segregation of fission gases, and high porosity in this region leads to a significant decrease in the thermal conductivity.\textsuperscript{365}

The interfacial thermal (Kapitza) conductance can be determined from data on the grain size and temperature dependence of thermal conductivity. Figure 18 is an example obtained from low temperature measurements from cubic zirconia\textsuperscript{356} and high temperature data for tetragonal 7YSZ zirconia.\textsuperscript{367} Up to about the Debye temperature (475 K), the Kapitza conductance increases with temperature, whereas above the Debye temperature, it is temperature independent. This temperature independence is largely due to the mean free path of the phonons approaching inter-atomic dimensions near the Debye temperature, with the result that the grain boundaries can only scatter the long wave-length phonons.

The interfacial thermal resistance is the result of the interactions of an ensemble of phonons with the GB. At the simplest level, the effect of a grain boundary can be characterized in terms of the phonon transmittance, $t$, the fraction of phonon energy that passes through the interface. One model for this transmittance based on phonon hopping is that it is related to the Kapitza conductance by\textsuperscript{368}

$$G_K = \frac{1}{R_K} = \frac{0.735k_B}{4(\Omega_0/N)}vt.$$  \hspace{1cm} \text{(37)}

Here, $v$ is the phonon velocity, $\tau$ is the transmittance, and the other symbols have the same meaning as in Eq. (28). The transmittance for YSZ grain boundaries deduced from Fig. 18 is 0.89; this is of course, averaged over all of the grain boundaries in the polycrystal and over all of the phonon branches and frequencies.

To dissect into individual phonon-grain boundary interactions, it has proved instructive to characterize the interactions of packets of phonons from a single branch and a narrow frequency range with specific interfaces. As is well-known, the standard models in which to interpret the phonon/grain boundary interactions are the DMM and the AMM.\textsuperscript{43} In the DMM, it is assumed that all of the energy in a phonon is trapped by the grain boundary, and then transmitted or reflected with probability proportional to the available density of states on the two sides of the grain boundary. In the AMM, the transmission coefficient is determined from the acoustic impedances, $Z = \rho v$, of the crystals on the two sides of the GB. Thus, for example, for a twist GB in electronic insulators and semiconductors (e.g., Si) the DMM predicts a transmission coefficient of 0.5, while the AMM predicts a transmission coefficient of 1.0. Detailed analysis of the transmission coefficient of transverse acoustic phonons through a $\Sigma 29$ twist GB in silicon showed that the AMM is valid for long wavelength phonons, whereas the DMM is valid for short wavelength phonons; the transmission coefficient shows a more complex behavior at intermediate wavelengths.\textsuperscript{38}

There thus remains considerable need for a more comprehensive understanding of phonon transmission through
interfaces and their relationship to the crystallography and orientation relationships. This is especially so for non-planar boundaries with high average curvature as they dominate in nano-crystalline materials. Furthermore, more comprehensive modeling is required to gain a more complete understanding of the thermal resistance across interfaces between dissimilar phases having different crystal structures and different mean atomic masses. This would be extremely useful in guiding the selection of two and three phase materials for the next-generation thermal barrier coatings.

C. Anisotropic crystals and natural superlattices

The engineering of artificial superlattices allows significant reduction in the thermal conductivity. It also introduces significant anisotropy between the cross- and parallel-plane thermal conductivities of such structures. It thus seems reasonable to expect that crystal structures having a strong structural anisotropy might also display anisotropic thermal transport. This has been found to be the case for several natural superlattices and is well illustrated by perovskite-based compounds. Natural superlattices are compounds whose unit cell consists of alternating crystallographic blocks having different atomic structure. For instance, bismuth titanate (BiT, Bi4Ti3O12), the prototypical Aurivillius phase, consists of (Bi2Ti3O9)2− perovskite blocks separated by (Bi2O2)2+ layers.

Similarly, the prototypical Ruddlesden-Popper (RP) phase consists of one or more SrTiO3 perovskite blocks separated by a SrO rocksalt layer. Simulations for various values of n, the number of perovskite layers between each SrO layer, show that for large values of n, the system responds as a classical interfacial system and the thermal conductivity normal to the modulation directions decreases linearly with density of the interfaces. For the lowest values of n, the system behaves as a monolithic single crystal with a larger unit cell. The thermal conductivity of rocksalt SrO is higher than that of perovskite-structured SrTiO3. Thus, for these low values of n, the thermal conductivity decreases slightly with increasing n. As a result, there is a minimum in the thermal conductivity at a critical value.

Figure 19 summarizes the conductivity of a range of recently discovered low thermal conductivity crystalline oxides. For comparison, the conductivity of fully dense 8YSZ, the current material of choice for thermal barrier coatings, is shown. In addition to having low thermal conductivity is attributed to reductions in thermal conductivity from the fluorite-derived 8YSZ due to doping. The case of (Zr0.6Y0.2Ta0.2)O2 is particularly instructive, since by doping with equal concentrations of a trivalent and pentavalent cation, the structure does not contain oxygen vacancies. (As discussed above, the low thermal conductivity of YSZ had been attributed to the presence of oxygen vacancies). LaPO4 and β–La2Mo2O9 are examples of crystal structures with internal distortions. The compound β–La2Mo2O9 is unusual because it has an exceptionally low elastic modulus, half that of fused silica and almost one-tenth of sapphire has unusual seven-fold coordination for the Mo ion and is also a fast ion oxygen conductor.

The data in Fig. 19 illustrate that materials can manifest low intrinsic thermal conductivity for a variety of reasons and that the well-established method of introducing a high concentration of point defects, by alloying or altering stoichiometry, or by refining the grain size, can also be used to decrease thermal conductivity. The strategy of selecting a refractory, high mean atomic mass compound and then alloying with cations of different atomic mass to increase phonon scattering has proven to be a viable approach to lowering thermal conductivity in many oxides. Solute alloying on the A and B cation sites of the pyrochlore oxides has been explored in considerable detail. Of these alloyed oxide systems, Sm2Zr2O7 with Yb substitution is especially interesting as there is a drop in conductivity associated with possible lattice softening at the composition where the ordered pyrochlore transitions to the disordered fluoride structure. There is also evidence in the Yb3+ alloyed La2Zr2O7 solid solution that the low thermal conductivity can be attributed to the Yb3+ behaving as a “rattler.”

Evidently, defect-free oxide crystal structures that are natural superlattices can have low thermal conductivity perpendicular to the stacking plane as shown by both experiment and simulation. The origin of the lower conductivity in terms of the relative contribution from phonon scattering and zone folding as well as from possibly weaker bonding perpendicular to the stacking plane represented by a different Gruneisen parameter remains to be established. In addition, the increased complexity and number of atoms in the unit cell of some of the oxides, such as Gd6+x+yCa2+x+y(SiO4)6O2+3x/2+y apatites and β–La2Mo2O9, for instance, are consistent
with the expectation that more of the thermal energy is carried by optical modes and contribute less to the thermal conductivity. To make further headway in understanding thermal conductivity in complex oxides such as these, more simulations and detailed inelastic neutron scattering studies of the phonon dispersion curves are required.

V. METROLOGY AND PROCESSING TOOLS

A. Time-domain thermoreflectance

1. Advances

During the past decade, TDTR has emerged as a powerful, general-purpose tool for the measurement of the thermal transport properties of materials. TDTR is applicable to materials with a wide range of properties and sample geometries. TDTR has been applied across the entire range of thermal conductivity from diamond and high thermal conductivity metals to the ultralow thermal conductivity of disordered layered crystals and fullerene derivatives. Essentially the same method can be applied to bulk materials, thin layers, and individual interfaces with both high (e.g., TiN/MgO and Al/Cu) and low (e.g., Bi/H-terminated diamond) thermal conductance. High-throughput mapping of thermal conductivity as a function of position is enabled by the micron-size focus of the laser spots and high signal-to-noise. Applications of thermal conductivity mapping include the analysis of metallurgical phase diagrams, and characterization of thermal barriers and coatings for nuclear fuels. Since TDTR is an optical, non-contact method, TDTR can be directly applied to samples in optical cryostats or high temperature microscope stages; and samples exposed to other extreme conditions such as the high pressure environment of a diamond anvil cell.

More than 25 years ago, the research groups of Maris and co-workers—the deposition of a metal film transducer on the sample of interest, heating of the transducer by a pump optical pulse, and measurement of the changes in reflectivity of the transducer for thermometry—is essentially the same as what we use today. The key advance in the past decade has been in a method for data analysis that starts from an analytical solution to the diffusion equation in cylindrical coordinates and includes the effects of so-called “pulse accumulation,” i.e., the fact that the signals are measured in a TDTR are not only a function of the pump-probe delay time but are also a function of the modulation frequency and the laser repetition rate. The use of single objective lens and integration of the pump and probe beams into an optical imaging system have made the alignment and focusing of the pump and probe beams fast and convenient.

The determination of thermal transport properties by TDTR is typically accomplished by adjusting free parameters in a thermal model to obtain the best fit between the predicted and measured thermal response of the sample. The free parameters are the unknown thermal transport properties, e.g., the thermal conductivity of a thin layer or the thermal conductance of an interface. Thus, TDTR data analysis proceeds in much the same manner as the analysis of ellipsometry data where reflectivity data as a function of wavelength and polarization is used to measure optical constants and layer thicknesses; or the manner in which x-ray reflectivity data as a function of scattering angle are analyzed to determine mass densities and layer thicknesses.

Consider a typical measurement of the thermal conductivity of a material. The sample is coated with an 80 nm film of Al. The unknowns in the thermal model are the thermal conductivity of the sample κ and the thermal conductance G of the Al/sample interface. The other parameters in the thermal model are independently measured or taken from literature values. The thickness of the Al film h is measured by picosecond acoustics; the heat capacity C of the Al is known from literature values; and the thermal conductivity of the Al film can be estimated based on the in-plane electrical conductivity and the Wiedemann-Franz law. The rms value of the pump and probe beam radius is measured by a spatial correlating the pump and probe or by measuring the transmission as a knife edge is translated through the focused pump and probe. The heat capacity of the sample is usually known from literature values or can be estimated based on data for related materials.

2. Limitations

TDTR requires that the surface of the sample is smooth enough so that any undesired modulation of diffuse scattering by thermoelastic effects does not produce a significant contribution to the desired signal coming from the change in specular reflectivity produced by the thermoreflectance of the metal film transducer. The in-plane length-scale or correlation length of the surface morphology also plays a role in these effects but an approximate rule-of-thumb is that the rms surface roughness should be <15 nm. Measurement errors created by modulation of diffuse light scattering are clearly observed in experiments on rough samples but our understanding of the mechanisms are qualitative at best and have not been systematically studied.

TDTR is mostly sensitivity to the thermal conductivity in the through-thickness direction, i.e., heat transport in the direction anti-parallel to the surface normal. For thermally anisotropic materials such as superlattices, textured polycrystalline films, and anisotropic crystals, knowledge of the in-plane thermal conductivity is also desired. In special cases, e.g., highly anisotropic materials or a relatively high thermal conductivity thin layer on a low thermal conductivity substrate, the in-plane thermal conductivity can be accurately measured using tightly focused laser beams and an analysis of how the TDTR signal varies as the pump beam is translated relative to the probe. The development of methods to accurately determine the in-plane thermal conductivity of thin layers is an active area of research.

Applicability of TDTR at low temperatures, <30 K is limited because the small heat capacity of the metal film transducer and sample requires a greatly reduced average laser power to keep the temperature excursions created by
the pump optical pulses below 10% the absolute temperature. At low temperatures, this constraint means that the maximum laser power that can be used in the experiment scales at the $\propto TC$, where $T$ is absolute temperature and $C$ is the heat capacity per unit volume of the metal film transducer. This scaling places a severe constraint on measurements at low temperatures. (High temperature measurements, on the other hand, do not present any fundamental difficulties. The technical challenges in high temperature measurements are the chemical and physical stabilities of the metal film transducer, measurement of the sample temperature, and mitigating heating of the microscope objective by air convection and thermal radiation from the hot sample environment.)

A key concept in understanding a TDTR measurement is the penetration depth $\delta$ of thermal waves in the sample at the modulation frequency $f$, $\delta = \sqrt{D/(\pi f)}$, where $D$ is the thermal diffusivity, $D = \kappa/C$, where $C$ is the heat capacity per unit volume. For the measurement of a thin film of thickness $h_f$ deposited on a substrate, we describe the layer as “thermally thin” if $h_f < \delta$ and “thermally thick” if $h_f \geq \delta$. For a thermally-thin sample, a possible limitation in the measurement is the fact that thermal conductance of the film/substrate interface cannot usually be separated from the thermal conductance of the film, $\kappa/h_f$. For a thermally thick film, the measurement is equally sensitive to both $\kappa$ and $C$ so any uncertainty in the estimation of $C$ propagates linearly into an uncertainty in $\kappa$.

3. Implementation

The most common implementation of TDTR uses a Ti:sapphire laser oscillator as the light source. The laser oscillator produces a train of sub-picosecond optical pulses at a high repetition rate, typically 80 MHz. The output of the laser is split into two optical paths, a “pump” path and a “probe” path, and the pump and probe are focused on the sample by a single microscope objective. The pump path includes an optical delay line and electro-optic modulator capable of chopping the pump beam at frequencies beyond 10 MHz. The probe path typically includes a audio-frequency mechanical chopper to facilitate the removal of coherent background signals. The term “thermoreflectance” refers to the fact that the probe measures changes in the temperature of the sample through the dependence of the optical reflectivity $R$ of a metal film transducer on temperature, $\tau = dR/dT$ is the coefficient of thermoreflectance. An Al thin film has an unusually large thermoreflectance at probe wavelengths generated by Ti:sapphire lasers, $dR/dT \approx 10^{-4}$ near room temperature at wavelengths near 800 nm. This fact, combined with the strong sensitivity of Al reflectivity to strain, makes Al ideally suited for both TDTR and picosecond acoustics.

A typical temperature excursion from the metal film is on the order of 1 K and therefore the signal (the change in reflectivity of the probe) is $\Delta R \sim 10^{-4}$. This signal is well above the noise floor of laser oscillators (the noise of Ti:sapphire laser oscillators is typically $<10^{-5}/\sqrt{\text{Hz}}$ at frequencies $>1$ MHz), but care has to be taken to suppress light from the modulated pump beam from reaching the detector.

The polarizing beam splitter that combines the pump and probe beam suppresses the reflected pump beam by a factor of $\sim100$ as long as optical elements between the beam splitter and the sample have small birefringence. (This is usually, but not always, the case; for example, in diamond anvil cells, inhomogeneous stress fields can produce significant birefringence in the diamond anvil.) Additional suppression of the reflected pump beam is provided by one of two approaches that can be combined: (i) spatially separating the pump and probe and passing only the specularly reflected probe beam with an aperture; or (ii) spectrally separating the pump and probe and blocking the reflected pump light with an optical filter. (If an aperture is used, care must be taken that the aperture does not partially block the probe; if the probe is partially blocked, thermal lensing and small deflections created by misalignment of the pump and probe beams can create artifacts in the signals and errors in the measurement.) At Illinois, we choose, in most cases, to spectrally separate the pump and probe by a small amount, $\approx 7$ nm using what we call a “two-tint” approach. Others groups have chosen to frequency double either the pump or probe. Double-modulation, i.e., adding a mechanical chopper to the probe path and audio-frequency lock-in detection of the output of the rf-lockin, eliminates any remaining background from the pump beam and also removes coherent pick-up by the rf-lockin.

Fiber laser oscillators operating at 1.55 $\mu$m (Er:fiber) or 1.05 $\mu$m (Yb:fiber) have also been used for TDTR measurements. The advantage of fiber lasers over Ti:sapphire lasers is the smaller footprint and lower cost. Also, Si is transparent at 1.55 $\mu$m, and therefore, an Er:fiber source can be used in experiments where it is desired to pass the pump and probe beam through the backside of a Si wafer. At 1.55 $\mu$m, Si has the useful property that the transient absorption near zero delay time can be used as optical thermometer with high bandwidth and high spatial resolution. At Illinois, we have not been fully satisfied these fiber lasers for general-purpose TDTR experiments because of problems in reproducing low-noise mode-lock conditions and the lack of suitable metal film transducers with the desired properties of high thermoreflectance, strong picosecond acoustics signals, high thermal conductivity, and small optical absorption depth that are provided by Al at Ti:sapphire wavelengths. While some metals have large thermoreflectance, the picosecond acoustic signals are typically weak and therefore the determination of the metal film thickness (the most important parameter in the data analysis) has to be accomplished by some other method such as x-ray reflectivity or Rutherford backscattering spectrometry. Since uncertainty in the thickness of the metal film transducer is usually the dominant source of uncertainty in the measurement, the inability to measure the transducer thickness by picosecond acoustics is a significant disadvantage.

4. Signal analysis and sensitivities

Signal analysis in TDTR is complicated by the presence of multiple time and frequency scales associated with the repetition rate of the laser and the modulation of the pump.
beams. Because the temperature excursions are a small fraction of absolute temperature, the system response is linear in both the pump and probe powers, and therefore, a frequency domain and time-domain descriptions of the signals are equivalent. In the frequency domain, the output of the laser oscillator is a comb of delta functions separated by 80 MHz. Modulation of the pump beam at $f = 10$ MHz produces side bands at 10, 70, 90, 150, 170 MHz, etc. The action of the lock-in amplifier is to evaluate the convolution of the frequency comb of the probe with the thermal response of the sample at these side bands. The time delay between pump and probe is equivalent to a phase shift in the frequency domain. The in-phase signal $V_{in}$ is thus given by a sum of the real part of the frequency response at these side bands. When the modulation frequency is $\approx 1/8$ of the repetition rate of the laser oscillator, pulse accumulation effects for $V_{in}$ are minimized and $V_{in}$ approximates the thermal response of the sample to a single pump optical pulse. The out-of-phase signal $V_{out}$ is mostly dominated by the imaginary part of the frequency response at the lowest frequency component (10 MHz) with much smaller contributions to $V_{out}$ coming from the differences between side bands at higher frequencies, e.g., the difference between the imaginary parts of frequency response at 90 and 70 MHz.

In most cases, the greatest sensitivity to the thermal properties of the material under study comes from $V_{out}$, not the time-evolution of $V_{in}$ as was conventionally analyzed prior to 2004. $V_{in}$ is mostly dependent on the heat capacity per unit area of the metal film transducer $hC_r$, which is known from independent measurements, and the thermal conductance of the transducer/sample interface $G$. The sensitivity of $V_{in}$ to the thermal conductivity of the sample $\kappa$ is usually small because only a minority fraction of the heat deposited by the pump optical pulse has moved from the metal film transducer to the sample at large delay times. $V_{out}$ is mostly dependent on $\kappa$, and to a lesser extent, $G$ and $hC_r$. The out-of-phase signal does not vary significantly with delay time. Analysis of the ratio $V_{in}/V_{out}$ is essentially a method of normalizing $V_{out}$ by the known value of $hC_r$ while at the same time providing a measurement of $G$ and eliminating artifacts created by changes in the rms spot size or small variations in the alignment of the pump and probe as a function of delay time.

It is useful to consider the relative magnitudes of an effective thermal conductance per unit area that is characteristic of the metal film transducer, the interface, and the sample. For the metal film transducer, we define $G_t = hC_{W0}$, where $\omega$ is the angular frequency of the modulation of the pump beam. For typical parameters, $G_t \approx 10$ MW m$^{-2}$ K$^{-1}$. A common interface conductance is $G = 200$ MW m$^{-2}$ K$^{-1}$. For $\kappa$ ranging from 0.1 to 1000 W m$^{-1}$ K$^{-1}$, the effective conductance of the sample $G_i = \sqrt{hC_r \kappa}$ varies between $G_i \approx 3$ and $G_i \approx 300$ MW m$^{-2}$ K$^{-1}$. Sensitivity to $\kappa$ is typically largest when $G_i < G < G_t$. (Similarly, TDTR measurements are most sensitive to the interface conductance $G$, when $G_t < G < G_i$.) Thus, TDTR measurements have reduced sensitivity when $\kappa$ is very small $G_t < G_i$ or when $\kappa$ is very high, $G_i > G$.

We can make the term “sensitivity” quantitative by defining a sensitivity coefficient that is the logarithmic derivative of the ratio signal $\phi = -V_{in}/V_{out}$ with respect to the parameter of interest. For example, the sensitivity coefficient for thermal conductivity is $S_\kappa = d\ln(\phi)/d\ln(\kappa)$. Figure 20 shows how $S_\kappa$ and $S_{G_t}$ vary with $\kappa$ for two values of $G$ that bracket common values for $G$ encountered in TDTR measurements, and two values of the delay time.

Pump-probe measurements that use two laser oscillators with slightly different repetition rates to rapidly scan the delay time between pump and probe are often referred to as asynchronous optical sampling (ASOPS). While ASOPS has clear advantages in accessing long delay times over the more traditional form of modulated pump-probe metrology where the delay time is limited by the length of a delay stage, the lack of a frequency component in the measurement less than the repetition rate reduces the sensitivity in many thermal measurements. In other words, a typical ASOPS experiment can only access the frequency response of the sample at multiples of the laser repetition rate while a conventional TDTR experiment accesses the frequency response at the modulation frequency in addition to the side bands that straddle each multiple of the repetition rate. This extra frequency

![FIG. 20. Absolute value of the sensitivity coefficients for thermal conductivity $S_\kappa$ (solid lines) and interface thermal conductance $S_{G_t}$ (dashed lines) for a prototypical TDTR measurement of a thermally thick sample of thermal conductivity $\kappa$. The Al film thickness is 80 nm; the $1/e^2$ radius of the laser spot is 10 μm; and the heat capacity of the sample is 2 J cm$^{-2}$ K$^{-1}$. In panel (a), $G = 200$ MW m$^{-2}$ K$^{-1}$; in panel (b), $G = 40$ MW m$^{-2}$ K$^{-1}$.](image)
component is critical in enhancing the sensitivity to the thermal conductivity of most samples. (Dilhaire and co-workers have recently described an approach that combines ASOPS with high frequency modulation of the pump beam.

Furthermore, ASOPS evenly weights all delay times in the data acquisition, while the TDTR data acquisition can be weighted to optimize the signal-to-noise in the regions of delay time that are most important for the measurement. For example, we typically acquire data in 1 ps intervals between 20 and 40 ps to accurately determine the relative phase of the rf lockin reference channel and measure the film thickness by picosecond acoustic signals, and then acquire data from 40 ps to 4 ns with ≈50 data points that are spaced uniformly in log(t).

Schmidt et al. analyzed the sensitivities of frequency-domain thermoreflectance (FDTR), i.e., experiments performed with a modulated cw laser as the pump and a second cw laser as the probe. The maximum sensitivity to the thermal conductivity \( \kappa \) of a bulk sample in an FDTR experiment is \( \approx 0.15 \); for comparison, the maximum sensitivity to \( \kappa \) in a TDTR experiment is \( \approx 0.5 \), see Fig. 20. FDTR has recently been extended to frequencies as high as 200 MHz using heterodyne detection to improve the signal-to-noise, an approach the authors refer to as broad-band frequency-domain thermoreflectance (BB-FDTR). Measurements of the thermal response of a sample between 200 kHz and 200 MHz are essentially equivalent to accessing time-scales for thermal transport between \( \approx 1 \mu s \) and \( \approx 1 \) ns. FDTR cannot provide measurements of the picosecond time-scale acoustics that are available in TDTR but has advantages in the simplicity of the laser sources and the ability to continuously scan frequency over a wide range. Information about the high frequency response is also embedded in the delay-time dependence of the in-phase TDTR signal; analysis of TDTR data is, however, more complex because it involves summations of the thermal response over many frequencies.

5. Forefront issues

All analysis of TDTR data to-date has depended on the assumption that (i) the diffusion equations is an adequate description of the temperature fields and heat fluxes; and (ii) boundaries between materials can be described by a radiative boundary condition that relates a temperature drop at an interface \( \Delta T \) to the heat flux crossing the interface \( J_Q \): \( J_Q = G A T \). Neither assumption can be rigorously justified unless the mean-free-paths of all of the significant heat carriers are small compared to the thermal penetration depth and couplings between various heat carriers is sufficiently strong that the distribution of heat carriers can be described by a single temperature.

When the thermal conductivity \( \kappa \) extracted experiment varies in an unexpected way, e.g., when \( \kappa \) varies with frequency or the size of the laser spot, the approach has been to interpret the result as a value that is modified by phonons that have mean-free-paths larger than a spatial scale in the experiment. This is, of course, not a completely satisfying situation because the modified value of \( \kappa \) is extracted from the data using an analysis based on the diffusion equation while simultaneously concluding that the diffusion equation cannot accurately describe the experiment. While semiconductor alloys and specific type of amorphous Si show pronounced frequency dependence in TDTR measurements, pure single crystals generally do not. It is possible that the frequency dependence is present but too weak to observe clearly. The spot-size effects at low temperature are large. Frequency domain measurements at very high frequencies may provide new insights.

The assumption that an interface in a TDTR experiment can be described as a simple radiative boundary condition on the diffusion equation has received little critical attention. This assumption ignores the true complexity of the problem—and misses a great deal of interesting science—because it assumes that the temperature is well-defined at each side of the interface; in other words, this description assumes that all excitations of each material are in equilibrium with each other at all distances from the interface. Using current methods and materials, we cannot usually resolve these effects. One exception was an experiment that combined TDTR with time-resolved x-ray diffraction where a significant fraction of the heat leaving an Al/GaAs interface was interpreted as evidence of phonons that did not equilibrate in a near-surface layer, because the mean-free-paths of these heat-carrying phonons were larger than the thickness of the near-surface layer. The effects of ballistic carriers on heat transport near a surface has continued to attract a great deal of attention in both experiment and theory, but we do not yet have a self-consistent method of incorporating these ideas into the analysis of TDTR data.

B. Scanning thermal microscopy

In scanning thermal microscopy (SThM), a nanometer-scale probe tip measures the temperature of a surface with nanometer-scale spatial resolution. The temperature measurement is spatial and temperature resolution governed by the size of the tip, size, and shape of the substrate features, and the heat transfer between tip and substrate. Over the last decade, there have been significant improvements in spatial resolution to \( \approx 10 \) nm, temperature precision to \( \approx 50 \) mK, and measurements of nanometer-scale heat flows to \( \approx 10 \) pW.

The temperature can be measured using either a temperature sensor in the cantilever tip or by measuring local thermomechanical strains in either the sample or the cantilever. The temperature sensor integrated into a scanning probe tip can be either a thermocouple or a thermistor. Noncontact SThM near-field optical probes can measure temperature with spatial resolution as small as 10 nm. These near-field optical measurements do not rely on thermal conduction between tip and sample, but instead are governed by near-field electromagnetic interactions between the tip and the sample, which are still not fully understood, see Sec. II B.

For both thermocouple- and thermistor-based SThM temperature measurements, the key technical challenge is to understand how the measured temperature signal relates to the substrate temperature. Several groups have investigated heat transport between a nanometer-scale tip and a
C. Nanometer-scale thermal analysis and thermal nano-manufacturing

Advancements in thermal scanning probe tools has enabled new capabilities in nanometer-scale materials characterization and nano-manufacturing. When the heated tip of an atomic force microscope (AFM) cantilever tip is in contact with a solid substrate, the tip-substrate interface forms a hotspot that is as small as a few nanometers in diameter. The temperature of this hotspot can be accurately controlled over a large temperature range. The ability to control the nanometer-scale hotspot at the end of a scanning probe tip has enabled fundamental measurements of nanometer-scale heat flow as well as new applications, such as nanometer-scale thermophysical property measurements and nanometer-scale thermal manufacturing.

Nano-thermal analysis (nano-TA) uses a heatable AFM cantilever to measure the phase change temperature of a material with nanometer-scale spatial resolution. In nano-TA, the AFM system holds the heated tip in contact with the surface while raising the cantilever tip temperature. As the material in contact with the tip softens, the tip penetrates into the surface and this tip motion is detected by the AFM. Nano-TA can be combined with other AFM-based measurements of substrate topography, adhesion, or stiffness. The combination of these techniques allows spatially resolved measurements of phase change temperature and temperature-dependent mechanical stiffness and loss modulus. The amount of material that can be interrogated using Nano-TA is on the order of attograms to femtograms, which is much smaller than state of the art calorimetry, thermogravimetry, or rheometry all of which require milligrams to gram scale quantities. Unlike these macroscopic techniques, Nano-TA can be performed on thin film samples and as-prepared material and can be used for failure analysis or the study of multi-component systems such as multi-layer film stacks or nanocomposites. Nano-TA measurement techniques have been aided by rigorous electro-thermal calibration and modeling as well as by improvements in cantilever design, fabrication, and instrumentation.

The nanometer-scale hot spot at the end of a heated AFM tip can be used for nanometer-scale thermal manufacturing. Thermal nano-manufacturing can form nanostructures with speed faster than 1 mm s⁻¹, which is close to that of commercial electron beam lithography tools. The high speed of thermal nano-manufacturing is due to the high temperatures that can be achieved to increase mass diffusivity or chemical conversion speed. The size of the features that can be written are governed by the size and shape of the temperature field near the tip, and in some cases, the mechanical properties of the substrate being processed.

In thermal dip-pen nanolithography (tDPN), a heated AFM tip can be used to deposit solid meltable ink from a tip onto a substrate with sub-100 nm feature size. Materials that have been written using thermal dip-pen nanolithography include many polymers, low melting temperature metals, and nanoparticle-polymer composites.

In thermochemical nanolithography (TCNL), a heated tip is brought into contact with a material that undergoes a temperature-dependent chemical reaction. For example, a polymer having a thermally removable functional group can be made reactive when its temperature is
raised. The reactive region can be smaller than 15 nm in size, owing to the very small heat affected zone near the heated AFM tip. Chemicals presented to the reactive size can form a chemical bond only in the thermally processed region. Examples of nanopatterned materials include nanoparticles and biological molecules. Another example of TCNL uses the heated tip to locally decompose a molecular glass to form three-dimensional nanopatterns.459

Finally, the heated tip can locally convert the electronic function of a material. For example, graphene oxide can be converted to graphene or an insulating organic layer can be locally converted into a semiconducting organic.456 With thermal nano-manufacturing, it is possible to control high temperature materials processing even on low temperature substrates, since the heat load on the substrate is small. For example, one study synthesized electroceramics directly on plastic substrates.458

VI. OUTLOOK

We discuss specific examples of outstanding scientific questions and recommendations for future work throughout this review but provide a brief summary of some selected examples here.

Research in the past decade has demonstrated the importance of interface physics and chemistry for thermal transport, but advances in experiment are needed to provide improved control and characterization of interface structure. Non-local contributions to the conductance—e.g., coupling of the electronic excitations of a material on one side of the interface to the vibrational excitations of a material on the other side of the interface—are not fully understood in theory and have not yet been systematically studied in experiment. Non-local contributions may be particularly important for weakly bonded interfaces, interfaces containing sub-nm gaps, and interfaces between materials with highly dissimilar lattice vibrations. Computer simulations of these channels for thermal transport will require extensions of molecular dynamics that include the interactions between quantum mechanical electronic states and vibrational states.

Extension of experiment and theory of near-field radiation to gaps on the order of 1 nm is challenging. Extremely perfect surface morphologies and more accurate methods of measuring the gap spacing are needed. Current theory is based on the long wavelength limit of the dielectric susceptibility \( \epsilon(q, \omega) \). As the gap size decreases, the spatial frequency dependence, \( \epsilon(q, \omega) \), becomes important.

Metal nanoparticles can be heated rapidly to extremely high temperatures using lasers with short pulse durations. In many cases of interest, the temperature excursion exceeds the boiling point or even the critical point of the solvent. Phase transformations under conditions of fast transient heating and steep temperature gradients are only beginning to be explored.

First principles theory can now predict the thermal conductivity of simple crystals. A challenge for the future is to extend this work to complex crystals that are of interest for thermoelectric energy conversion. Rigorous treatment of disorder and boundaries will require further advances in theory and computational methods. It will be useful to make detailed comparisons between the results of first principles calculations and the phenomenological scattering rates and integral equations that are typically used by experimentalists to model their data.

Thermal conductivities far below the value expected for boundary scattering have been observed in Si nanowires prepared by metal-assisted chemical etching of Si wafers. The mechanism of the suppression is controversial and may be explained by the physics of phonon waves interacting with rough surface morphologies. Metal-assisted etching process is, however, a complex process and the condition of the near-surface region of etched Si nanowires requires further study.

Carbon nanotubes and graphene have extremely high thermal conductivity when fully isolated, but the conductivity is strongly suppressed when supported by a substrate, embedded in a matrix or otherwise encased in surrounding materials. Advances in theoretical understanding of the mechanisms that produce this suppression and fabrication methods that mitigate the effects of surrounding materials will be important for thermal management of nanoelectronics and the synthesis of nanocomposites with high thermal conductivity.

Ultralow thermal conductivity has been observed in materials that combine disorder, high interface density, and strong anisotropy of elastic constants. The conductivity of disordered WSe\(_2\) is far below the conventional prediction of the minimum thermal conductivity of an amorphous material. Thus, the lower limit of the conductivity of a fully dense solid is not known. The materials that combine ultralow thermal conductivity with high temperature stability are needed for improved thermal barriers.

TDTR has emerged as a powerful technique for measuring the thermal conductivity of thin layers and bulk materials with high spatial resolution, but analysis of TDTR data is currently based on a solution of the diffusion equation with interfaces treated as a radiative boundary condition on heat flux and temperature. Rigorous methods for analyzing TDTR data that incorporate models of phonon scattering rates and non-equilibrium between various heat carriers at interfaces are needed to better understand the results of TDTR experiments.

ACKNOWLEDGMENTS

D.G.C., P.V.B., G.C., S.F., P.K., and W.P.K. were supported by Air Force Office of Scientific Research MURI FA9550-08-1-0407. D.R.C. acknowledges support from the Office of Naval Research (N000-14-121-0993).

K.E.G. appreciates input from Mehdi Asheghi, Elah Bozorg-Grayeli, Jungwan Cho, Jaeho Lee, and Zijian Li. K.E.G. also thanks Duygu Kuzum, Rakesh Jeyasingh, and H. S. Philip Wong of the Stanford University Electrical Engineering Department for sharing unpublished microscopy and resistance-current data for a phase change nanocell.

The work of S.R.P. was supported as a subcontractor (AC) of the U.S. Government under DOE Contract No. DE-AC07-05ID14517, under the Energy Frontier Research


