Demystifying umklapp vs normal scattering in lattice thermal conductivity
A. A. Maznev and O. B. Wright

Citation: American Journal of Physics 82, 1062 (2014); doi: 10.1119/1.4892612
View online: http://dx.doi.org/10.1119/1.4892612
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/82/11?ver=pdfcov
Published by the American Association of Physics Teachers

Articles you may be interested in
Lattice thermal conductivity of filled skutterudites: An anharmonicity perspective

Coupled theoretical interpretation and experimental investigation of the lattice thermal conductivity of Bi2Te3 single crystal

Temperature dependence of the thermal conductivity of different forms of diamond

The lattice thermal conductivity of a semiconductor nanowire

In-plane lattice thermal conductivity of a quantum-dot superlattice
Demystifying umklapp vs normal scattering in lattice thermal conductivity

A. A. Maznev
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

O. B. Wright
Division of Applied Physics, Faculty of Engineering, Hokkaido University, Sapporo 060-8628, Japan

(Received 17 October 2013; accepted 28 July 2014)

We discuss the textbook presentation of the concept of umklapp vs normal phonon-phonon scattering processes in the context of lattice thermal conductivity. A simplistic picture, in which the “momentum conservation” in a normal process leads to the conservation of the heat flux, is only valid within the single-velocity Debye model of phonon dispersion. Outside this model, the simple “momentum conservation” argument is demonstrably inaccurate and leads to conceptual confusion. Whether or not an individual scattering event changes the direction of the energy flow is determined by the phonon group velocity, which, unlike the quasimomentum, is a uniquely defined quantity independent of the choice of the primitive cell in reciprocal space. Furthermore, the statement that normal processes do not lead to a finite thermal conductivity when umklapp processes are absent is a statistical statement that applies to a phonon distribution rather than to individual scattering events. It is also important to understand that once umklapp processes are present, both normal and umklapp processes contribute to thermal resistance. A nuanced explanation of the subject would help avoid confusion of the student and establish a connection with cutting edge research.

I. INTRODUCTION

In nonmetallic solids, thermal energy is predominantly carried by lattice vibrations, or phonons. In a perfectly harmonic defect-free crystal lattice, phonons would propagate without scattering, yielding an infinite thermal conductivity. Anharmonicity of the lattice leads to interaction of lattice waves with each other, referred to as phonon-phonon scattering. Debye1 made the first attempt to explain a contribution to the thermal conductivity of solids based on anharmonicity, using an elastic continuum model. However, Debye’s paper contains a mistake, as was shown by Peierls in his fundamental work on the thermal conductivity of crystals.2 In an elastic continuum, no thermal resistance can arise from anharmonicity alone; it is the discrete nature of the crystal lattice that makes anharmonic thermal resistance possible.

An important concept from Peierls’ paper, invariably invoked in explaining how anharmonic phonon interactions in the crystal lattice lead to a finite conductivity, is that of normal vs umklapp phonon-phonon scattering.3 The majority of textbooks4–10 present a simple picture according to which normal (N) processes conserve the phonon momentum and, consequently, do not contribute to the thermal resistance. In contrast, umklapp (U) processes, in which the wave vector conservation involves a reciprocal lattice vector, change the momentum, thereby causing thermal resistance. Deceptively appealing though it seems, this description quickly runs into a number of conceptual problems once we dig a little deeper. Incidentally, Peierls himself11 emphasized that the analogy with momentum conservation “must not be taken too seriously,” and that the distinction between N and U processes depends on our convention in choosing the basic cell of the reciprocal lattice. Some textbooks12–14 follow Peierls in providing a nuanced discussion, and concede that the distinction between N and U processes is “to some extent artificial.”14 In the following, we will show that a discussion of the nuances is indeed necessary, and that the appeal of simplicity of the “momentum conservation” argument does not justify the conceptual confusion that results. In essence, this paper represents a guide to readers who may be confused by the present exposition of the subject from most sources, and also an appeal to teachers and future textbook writers to set things straight.

II. DIFFICULTIES WITH “MOMENTUM CONSERVATION”

The standard explanation of the distinction between N and U processes4–10 is illustrated in Figs. 1(a) and 1(b). Consider, for example, a three-phonon scattering process in which two phonons of angular frequencies ω1 and ω2 merge into one of angular frequency ω3. In such a process, energy conservation takes the form

$$h \omega_1 + h \omega_2 = h \omega_3 = 0.$$  (1)

In an N-process, momentum conservation expressed in terms of corresponding phonon wave vectors takes the form

$$k_1 + k_2 - k_3 = 0,$$  (2)

whereas in a U-process there is a nonzero reciprocal lattice vector \( \mathbf{G} \) on the right-hand-side:

$$k_1 + k_2 - k_3 = \mathbf{G}.$$  (3)

As can be seen in Fig. 1(b), the latter process results in redirecting the phonon momentum to a value different from \( k_1 + k_2 \), from which the German term “umklapp” meaning “flip over” was coined. (Although, strictly speaking, flipping over takes place only if \( k_1 + k_2 \) is collinear with \( \mathbf{G} \).)

It is of course recognized13,15 that the phonon “momentum” is not the true momentum. Since a lattice vibration is not
accompanied by mass transfer (unless the crystal moves as a whole), the net mechanical momentum of a phonon is in fact zero. This is particularly evident for a transverse phonon, in which the particle motion is perpendicular to the wave vector. Rather, we are talking about a quasimomentum or “crystal momentum” given by \( \pm \frac{2\pi}{C^2} k \), with \( \hbar \) the reduced Planck constant.

The phonon wave vector \( k \) is not uniquely defined in either magnitude or direction; rather it is defined modulo \( G \), a reciprocal lattice vector, which means that wave vectors \( k \) and \( k + G \) are equivalent. Consequently, the quasimomentum is only defined modulo \( \pm \frac{2\pi}{C^2} G \). Thus defined, the quasimomentum ought to be conserved in any phonon scattering process, be it an N-process or a U-process. We can always add a reciprocal lattice vector \( G \) to either side of a quasimomentum conservation equation, since wave vectors are only defined up to \( G \). What, then, is the quantity that is conserved in an N-process but not conserved in a U-process? A quick inspection shows that it is the sum of reduced wave vectors (or “reduced momenta”) defined within the primitive cell of the reciprocal lattice. However, the definition of the reduced wave vector depends on the choice of the primitive cell, and even though the first Brillouin zone (BZ), corresponding to a Wigner-Seitz cell in the reciprocal lattice, is normally taken as the primitive cell, we are not bound to this choice. A different choice of the primitive cell will turn some normal processes into umklapp processes and vice versa, for example, the U-process of Fig. 1(b) becomes an N-process in (c) and (d). But then how can there be any physical difference between N and U processes (e.g., “non-resistive” vs “resistive”) if their definition depends on the choice of primitive cell? A related problem arises from the fact that, for a given primitive cell, an infinitesimal change in a phonon wave vector will turn a non-resistive N-process into a resistive U-process.

A further difficulty is illustrated in Fig. 2, which shows a possible three-phonon scattering process in silicon where two identical longitudinal acoustic (LA) phonons merge into a longitudinal optical (LO) phonon in a normal process conforming to Eq. (2). However, note that the group velocity of the LO phonon, defined by the slope of the dispersion curve, is opposite in direction to that of the original LA phonons. Consequently, upon scattering, energy is carried in the opposite direction. Is this a non-resistive or a resistive process? Thus, under closer consideration, the seemingly simple and appealing “momentum-conservation” argument leaves one in a state of confusion, with a number of conceptual questions unanswered.

III. GROUP VELOCITY AND THE EFFECT OF A SCATTERING EVENT ON THE HEAT FLUX

We will try to resolve the conceptual difficulties on two levels: the level of an individual scattering event, and the level of the statistical properties of a phonon distribution in which many scattering events are taking place. On the first level, the question raised with respect to Fig. 2 hints that the group velocity is a key to resolving the difficulties. Indeed, the contribution of a phonon to the heat flux is equal to the product of phonon’s energy \( \hbar \omega \) and its group velocity \( v_g \). 

Fig. 1. Schematic diagrams of three-phonon scattering processes in wave vector space: (a) normal and (b) umklapp processes, defined with the first BZ chosen as the primitive cell; (c) and (d) show the same scattering event as in (b) for different primitive cell choices. Dots correspond to reciprocal lattice points.

Fig. 2. Phonon dispersion along the [100] direction in Si (after Ref. 26). Arrows show a collinear N-process whereby two identical LA phonons with frequencies \( \omega_1 \) and wave vectors \( k_1 \) merge to yield a LO phonon of frequency \( \omega_3 = 2\omega_1 \) and wave vector \( k_3 = 2k_1 \). The group velocity of the LO phonon is opposite to that of the LA phonons; therefore, the energy flow direction is “flipped.” Points \( \Gamma \) and \( X \) correspond to the BZ center and boundary, respectively.
Consequently, the total heat flux carried by phonons is given by the sum of $\hbar v_g$ over all phonons:\(^{18}\)

$$ j = \frac{1}{(2\pi)^3} \int \hbar v_g n_k d^3 k, \quad (4) $$

where $v_g$ is the phonon group velocity, $n_k$ is the phonon occupation number in wave vector space, and the summation over phonon branches is implied. Unlike the phonon wave vector the group velocity is known unambiguously, and thus Eq. (4) yields a uniquely defined measurable quantity.

In the simplest model of the phonon dispersion based on the Debye model of specific heat, phonons have a linear dispersion with a slope given by a constant velocity $v$. In this case, within the first BZ, the group velocity equals the phase velocity and the product $v n_k$ equals the reduced phonon wave vector $k$ times $v^2$. Consequently, the heat flux is indeed given, up to the factor $\hbar v^2$, by the sum of reduced wave vectors.\(^{14}\) In this case, normal processes conserve the heat flux while umklapp processes would not, because flipping the reduced wave vector would entail flipping the group velocity. However, in the single-velocity Debye model U-processes are in fact impossible because the requirements of energy and quasimomentum conservation can only be satisfied by collinear N-processes.\(^{19}\)

A modification of the Debye model that allows for the existence of separate longitudinal and transverse acoustic branches enables U-processes.\(^{19}\) Remarkably, however, as soon as we have two branches, the argument that the “momentum conservation” preserves the direction of the energy flow at the level of an individual scattering event breaks down. Consider, for example, a collinear scattering process illustrated in Fig. 3, with an LA phonon with wave vector $k_1$ decaying into LA phonon $k_1$ and TA phonon $k_2$. Energy and quasimomentum conservation for this normal process leads to relationships

$$ k_1 = -\frac{1-s}{1+s} k_3 \quad \text{and} \quad k_2 = \frac{2}{1+s} k_3, \quad (5) $$

where $s = v_T/v_L$ is the ratio of transverse and longitudinal acoustic velocities. The contribution of phonons $k_1$ and $k_2$ to the heat flux is then

$$ \sum \hbar v_g = h v_g^2 k_1 + h v_g^2 k_2 = \frac{2s^2 + s - 1}{1+s} h v_g^2 k_3. \quad (6) $$

For $s = 1/2$, the contribution of the two created phonons to the heat flux will be zero, and for $s < 1/2$, which is not uncommon, the contribution to the heat flux will flip direction with respect to $k_3$. The reason for the flipping is that although $k_1$ is much smaller than $k_3$ in magnitude, its contribution to the energy flux is larger because of the larger velocity.

Let us now address the question of whether an infinitesimal change in the wave vectors turning an N-process into a U-process produces any finite physical effect. This is indeed the case in the Debye model, which has an unphysical discontinuity of the dispersion slope at the BZ boundary that must be introduced in order to make the dispersion periodic in reciprocal space, as seen in Fig. 3. In real materials, the group velocity is continuous across the BZ boundary, hence normal and umklapp processes separated by an infinitesimal interval across the BZ boundary are indistinguishable.\(^{13}\) In the words of Peierls,\(^{11}\) “there is no important physical difference between processes in which the sum $k_1 + k_2$ just remains within the basic cell, and those in which it falls just outside.”

Thus, no clear distinction between N and U processes can be drawn, which is consistent with the fact that their definition depends on the choice of the primitive cell. Does, then, the concept of N and U processes have any utility outside the Debye model? To answer this question we need to look beyond the individual scattering event, which brings us to the main subject of Peierls’ paper,\(^{2}\) in which the concept of umklapp processes was introduced—the statistical effect of scattering processes on a phonon distribution.

### IV. LOOKING DEEPER: STATISTICAL EFFECT OF SCATTERING PROCESSES

Peierls showed that if umklapp processes are not allowed, then thermal equilibrium cannot be established, and an initial heat flux (in the absence of an external temperature gradient) will never fully dissipate.\(^{2,11–13}\) This is a statistical statement applied to the effect of many scattering processes on the phonon distribution rather than to a single scattering event. An intuitive explanation of this result is as follows.\(^{11,12}\) By restricting scattering processes to within the primitive cell we enforce the conservation of the sum of the reduced wave vectors. Since in a perturbed state, this quantity will generally be different from that in equilibrium, the return to equilibrium from an arbitrary perturbed state is thus rendered impossible.\(^{20}\) It has been shown\(^{13}\) that in the absence of U-processes a perturbed state will relax to a “displaced” Bose-Einstein distribution,

$$ m_k = \left[ \exp \left( \frac{\hbar \omega(k) - \gamma k}{k_B T} \right) - 1 \right]^{-1}, \quad (7) $$

where $k$ is the reduced wave vector and the constant vector $\gamma$ is chosen to make the sum of the reduced wave vectors correct. This distribution will, in general, support a non-zero heat flux\(^{12,13}\) that will never vanish.

Does this mean that an unambiguous distinction between N and U processes can be made after all, if only in a statistical sense? The answer is no, as the result discussed above does not depend on the choice of the primitive cell. It is not
the difference between U and N processes that matters (because any U-process can be turned into an N-process by a different choice of the primitive cell), but the fact we have effectively put a “fence” around a primitive cell (any primitive cell) that restricts phonon-phonon scattering to N-processes inside this cell; it is this restriction that prevents the perturbed system from returning to thermal equilibrium. Peierls’ result holds for any primitive cell, even though the definition of N and U processes depends on the primitive cell choice.

Even though from the formal point of view the choice of the primitive cell is arbitrary, one centered at \( \mathbf{k} = 0 \) is the natural choice (whether it is the first BZ or not). Such a choice corresponds to the only situation that involves “disallowed” umklapp scattering encountered in nature, namely, the case of low temperatures, when only low wave vector states around \( \mathbf{k} = 0 \) are populated. For a primitive cell containing a large enough neighborhood around \( \mathbf{k} = 0 \), such as those shown in Figs. 1(a) and 1(c), U-processes are “frozen out” at low temperature owing to the lack of a high wave vector phonon population. Thus, even though a rigid distinction between N and U processes cannot be made, describing low-frequency acoustic phonon interactions as N-processes is justified when analyzing thermal conductivity at low temperatures. It is in this context that the designation “N-processes” is used for practical purposes. It would still be possible to define N and U processes for the primitive cell shown in Fig. 1(d) and, if we found a way to prevent the thus-defined U-processes from occurring, N-processes would not lead to a finite thermal conductivity. However, since there is no known physical way to eliminate U-processes for this primitive cell, such a description would lack practical utility.

N-processes can only be said not to contribute to thermal resistance when U-processes are absent. However, once umklapp processes do occur, both N and U processes contribute to establishing thermal equilibrium and to a thermal resistance. Remarkably, N-processes contribute to thermal resistance even within the Debye model when they do not have any direct effect on the heat flux. An instructive case in point is the combined effect of phonon-phonon scattering and the scattering from lattice imperfections at low temperatures considered by Peierls himself. As we discussed, if the temperature is low enough, U-processes are frozen out and N-processes alone do not lead to a finite thermal conductivity. Elastic scattering from lattice defects, on the other hand, changes the direction of the reduced wave vector (and of the group velocity) and should be considered a “resistive” process. It turns out, however, that scattering by defects alone also fails to produce a finite thermal conductivity. The reason is a divergence at low frequencies owing to the defect scattering probability scaling as \( k^4 \). It is only by acting together that the two effects result in a finite conductivity.

In outline, defect scattering establishes an isotropic distribution for high-frequency phonons while N-processes establish equilibrium between low- and high-frequency phonons.

Although the formal theory of thermal conductivity due to phonon-phonon scattering has been well established through Peierls’ work, actual calculations for real materials based on this theory have until recently remained beyond reach. A significant recent development has been the advent of ab-initio calculations of lattice thermal conductivity that implement an exact iterative solution based on Peierls’ formalism and use interatomic force constants from density functional calculations. The formalism of the theory does not distinguish between N and U processes, therefore their relative role is not directly obvious from the numerical results. Ward and Broido shed light on this issue by comparing the exact calculations with the relaxation-time approximation (RTA), in which the thermal conductivity is given by

\[
\lambda = \frac{1}{3} \frac{1}{(2\pi)^3} \int C_k v_s^2 \tau_{RTA} d^3k,
\]

where \( C_k \) is the mode specific heat, \( \tau_{RTA} \) is the single-mode relaxation time, and a summation over phonon branches is implied. These researchers found that the thermal conductivity of Si above 100 K closely obeys Eq. (8) with the relaxation time given by

\[
\frac{1}{\tau_{RTA}} = \frac{1}{\tau_N} + \frac{1}{\tau_U},
\]

where \( \tau_N \) and \( \tau_U \) are obtained from normal and umklapp scattering rates, respectively. The exact solution and the RTA yielded similar results for 100–800 K. Equation (9) implies that for the purposes of calculating the thermal conductivity in this relatively high temperature case, normal and umklapp processes are indistinguishable. Moreover, it was found that about 50% of the room temperature thermal conductivity of Si is attributable to low-frequency phonons (below ~2 THz), for which normal scattering dominates and \( \tau_{RTA} \) is mainly determined by N-processes. Thus at high temperatures, when phonon states up to the BZ boundary are populated, drawing a line between N and U processes hardly serves any practical purpose. Consequently, the widely used term “umklapp resistance” is somewhat misleading; “anharmonic resistance” or “resistance due to phonon-phonon scattering” would seem to be more appropriate.

V. CONCLUSIONS

In conclusion, no rigid line between N and U processes can be drawn because their definition depends on the choice of the primitive cell of the reciprocal lattice. However, choosing a primitive cell centered around \( \mathbf{k} = 0 \) is recommended for analyzing thermal conductivity at low temperatures when U-processes become “frozen out.” N-processes alone do not create thermal resistance, but they do contribute to thermal resistance when U-processes (or scattering from defects or impurities) are present. At temperatures comparable to or greater than the Debye temperature the distinction between N and U processes becomes blurred and loses practical utility.

Conceptual difficulties arise from attempts to explain a result from statistical mechanics obtained by Peierls in a simple and visually appealing way by drawing diagrams of two individual scattering processes. Sometimes, such simplistic explanations of complicated results are useful; indeed, in the physics curriculum, nuances are often sacrificed for the sake of clarity, and rightly so. The problem with the “momentum conservation” picture given in the majority of textbooks is that the conceptual clarity is sacrificed together with the nuances. The fact that the quasimomentum is not uniquely defined in either magnitude or direction is key to understanding wave propagation in periodic media. The argument that “momentum conservation” leads to preserving the direction of the energy flow is misleading and plainly inaccurate. Equally misleading is the simplistic view that N-processes...
do not contribute to thermal resistance. It would be incorrect to think that the issues here are of a mainly theoretical nature with little practical consequence. For example, it has been argued that the phonon mean free path relevant for thermal conductivity calculations is determined by U-processes alone. This is demonstrably inaccurate: for example, equating $\tau_{RTA}$ in Eq. (8) to the umklapp relaxation time $\tau_U$ would lead to an incorrect thermal conductivity.

If a brief discussion of the subject in the classroom is desired, it should, as a minimum, invoke the group velocity. An example of such a brief treatment is given by Taylor and Heinonen. A deeper discussion should go beyond the individual scattering event, and would benefit from a consideration of the nuances pertaining to the contribution of N-processes to the thermal resistance. Discussing some of the recent work would also be instrumental in connecting the subject to state-of-the-art research.

ACKNOWLEDGMENT

The authors greatly appreciate stimulating discussions with A. G. Every and comments on the manuscript by J. K. Eliason. The contribution by A. A. M. was supported as part of the S3TEC Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award No. DE-SC0001299.


The term “umklapp” was introduced by Peierls, but he did not use the term “normal processes.” The terminology currently in use was established by the late 1950’s, see, e.g., J. Callaway, “Model for lattice thermal conductivity at low temperatures,” Phys. Rev. 113, 1046–1051 (1959).


14This important fact holds not only for phonons but for any waves in a periodic medium. Indeed, according to Bloch’s theorem, the spatial dependence of an eigenstate of any excitation in a crystal can be represented in the form $e^{i\mathbf{k}\cdot\mathbf{r}}\sum_\Gamma e^{i\mathbf{\Gamma}\cdot\mathbf{r}}$. Bloch’s expansion defines the wave vector $\mathbf{k}$, and from this definition it follows that wave vectors $\mathbf{k}$ and $\mathbf{k} + \mathbf{G}$ are equivalent. It is important to recognize a distinction between a wave in a homogeneous medium (such as an electron in vacuum or an acoustic wave in an elastic continuum), in which case the wave vector $\mathbf{k}$ defines a plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$, and a wave in a periodic medium, where the wave vector $\mathbf{k}$ generally defines not a plane wave but a superposition of plane waves given by Bloch’s expansion. Even though in both cases the term “wave vector” is used, the exact meaning of the term is different.

15We believe that “momentum conservation” is an unfortunate choice of words even if the distinction between the “crystal momentum” and the real momentum is made clear. This choice of words creates an impression of some real physical quantity conserved in N-processes only. However, conservation of a real physical quantity should not depend on the choice of the primitive cell. We would prefer to talk of the conservation of the sum of reduced wave vectors, a somewhat lengthy but precise description that underscores the artificial character of this quantity.

16This important fact holds not only for phonons but for any waves in a periodic medium. Indeed, according to Bloch’s theorem, the spatial dependence of an eigenstate of any excitation in a crystal can be represented in the form $e^{i\mathbf{k}\cdot\mathbf{r}}\sum_\Gamma e^{i\mathbf{\Gamma}\cdot\mathbf{r}}$. Bloch’s expansion defines the wave vector $\mathbf{k}$, and from this definition it follows that wave vectors $\mathbf{k}$ and $\mathbf{k} + \mathbf{G}$ are equivalent. It is important to recognize a distinction between a wave in a homogeneous medium (such as an electron in vacuum or an acoustic wave in an elastic continuum), in which case the wave vector $\mathbf{k}$ defines a plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$, and a wave in a periodic medium, where the wave vector $\mathbf{k}$ generally defines not a plane wave but a superposition of plane waves given by Bloch’s expansion. Even though in both cases the term “wave vector” is used, the exact meaning of the term is different.

17Ziman, in his 1960 book, remarked: “The Boltzmann equation is so exceedingly complex that it seems hopeless to expect to generate a solution from it directly.” It took just under half a century and modern computing capabilities to solve this “hopeless” problem! The Boltzmann transport equation for phonons, also called the Peierls-Boltzmann equation, is the form commonly used to express Peierls’ theory.


21Indeed, researchers doing actual calculations of the thermal resistance due to phonon-phonon scattering avoid using this term.

22According to Esfarjani et al., the umklapp lifetime scales as $1/k^3$ in the low-frequency limit. Plugging this dependence into Eq. (8) would result in an infinite conductivity owing to the divergence of the integral at low frequencies.