



Title	第一原理計算を用いた非調和フォノン特性データベースの構築
Author(s)	大西, 正人
Citation	サイバーメディアHPCジャーナル. 2026, 16, p. 33-36
Version Type	VoR
URL	https://doi.org/10.18910/106229
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

第一原理計算を用いた非調和フォノン特性データベースの構築

大西 正人

統計数理研究所 先端データサイエンス研究系

1. はじめに

フォノンは結晶中の原子の集団振動に起因する準粒子であり、熱伝導をはじめとして電子・光・磁気・機械特性など多様な物性に重要な役割を果たす。熱マネジメントの観点から、格子熱伝導率を含む非調和フォノン特性への理解は、次世代電子デバイスのパッケージング材料や熱電変換材料の設計において不可欠である。また超電導体では電子-フォノン結合が臨界温度を決定する重要な因子であり、磁性体においてもスピン-フォノン相互作用が物性に影響する。このように、フォノン特性の理解は熱物性だけでなく、材料科学における幅広い分野で重要な課題となっている。

近年、マテリアルズ・インフォマティクス (MI) の急速な発展により、大規模データと機

械学習を組み合わせた材料探索が加速している。

Materials Project (約 20 万材料)・Alexandria (580 万構造)・ICSD などの材料データベースが整備され、電子バンド構造・形成エネルギー・誘電特性などの電子物性については大規模なデータが蓄積されてきた。一方で、フォノンに起因する非調和物性 (格子熱伝導率・フォノン寿命・散乱率・スペクトル熱伝導率など) に関するデータベースは依然として限られており、材料探索への活用が進んでいない。

格子熱伝導率の第一原理計算には大規模超格子構造 (数十~300 原子規模) に対する大量の力計算が必要であり、1 材料あたり数 100 NH を要する。このため、既存の大規模計算研究においても取り扱われる材料数は 100 材料程度にとどまっていた。近年のスーパーコンピュータ環境の発展により、大量の計算を並列・自動的に実

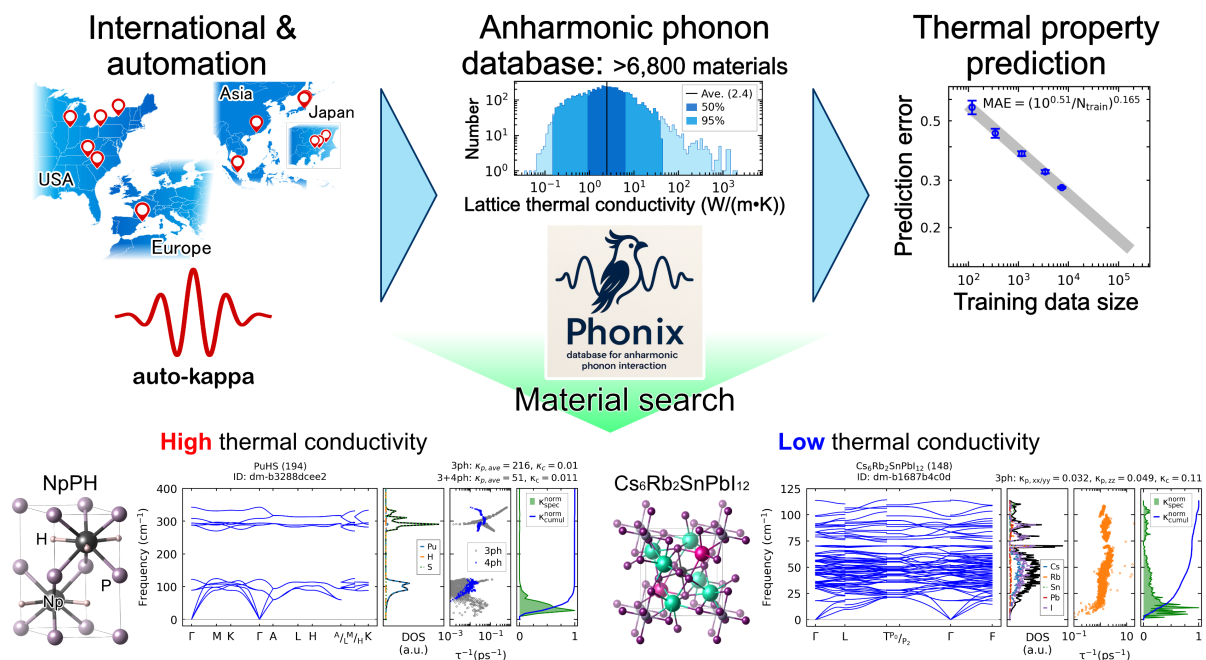


図 1. 本プロジェクトの概要：非調和フォノン特性の第一原理計算を自動化 (auto-kappa) し、国際的な協力のもと約 7,000 材料以上の熱伝導率を含むデータベース (Phonix) を構築した。構築したデータベースと GNN を用いて熱物性予測の機械学習モデルを構築し、深層学習スケーリング則を実証した。さらに予測モデルを用いて約 38 万材料をスクリーニングし、低熱・高熱伝導の候補材料を特定した。

行するハイスループット計算が現実的となりつつある。本研究では、**auto-kappa** と名付けた自動化計算ワークフローを開発し、大阪大学 D3 センターの **SQUID** をはじめとする大規模計算機を活用することで、約 7,000 材料以上の非調和フォノン特性を計算し、大規模データベース (**Phonix**) を構築することに成功した (図 1)。

また本プロジェクトでは、**auto-kappa** ソフトウェアをオープンソースとして公開するとともに、国内外の研究者との国際共同研究として大規模なデータ生成を推進した。これにより、異なる機関の計算資源を有効活用し、単独の研究グループでは困難な規模のデータセットを効率的に構築することが可能となった。本稿では、本研究の計算手法・結果・考察を報告する。

2. 計算手法・モデル等

2-1. 第一原理計算の設定

第一原理計算には平面波基底の密度汎関数 DFT コード **VASP** [1,2] を使用した。交換相関汎関数には **GGA/PBESol** 汎関数を採用し、イオンコアは射影増強波 (**PAW**) 法で扱った。カットオフエネルギーは推奨値の 1.3 倍に設定し、**k** 点数は $N_i = \max(1, \text{int}(R_k/|\mathbf{b}_i|) + 0.5)$ ($R_k = 20 \text{ \AA}$) により決定した。構造最適化は完全緩和 2 回 (**VASP** オプション: **ISIF** = 3)、次いで格子定数固定での原子位置最適化 (**ISIF** = 2) の計 3 段階で実施し、フォノン計算における虚数振動数の発生を抑制した。

2-2. フォノン特性の計算

フォノン特性の解析には **ALAMODE** [3] を用いた。慣用単位胞を基に最大 150 原子の超格子を自動生成し、有限変位法 (変位 $0.01 \text{ \AA}$) により調和力定数 (**FC2**) を評価した。調和フォノン特性では **q** 点密度を 500, 1000, 1500 **q**-points· $\text{\AA}^3/\text{atom}$ の 3 種類に設定し、Born 有効電荷から非解析補正 (**NAC**) を適用した。

非調和力定数 (**FC3**) の計算では、カットオフ距離を $\max(4.3 \text{ \AA}, \text{第 3 近接原子間距離})$ に設定し、変位量 $0.03 \text{ \AA}$ を使用した。必要な変位パタ

ーン数が 100 以上の場合は **LASSO** 回帰、それ以下では有限変位法を適用した。得られた力定数から **BTE** を解いて格子熱伝導率を算出した。虚数振動数が生じた場合は厳密な多段階構造最適化や大規模超格子 (最大 200 原子) の導入を自動的に実施した。これら一連の処理は **auto-kappa** (<https://github.com/phonix-db/auto-kappa>) により統合・自動化されている。

2-3. SQUID を用いたハイスループット計算

大規模ハイスループット計算には、主に大阪大学 D3 センターの **SQUID** の CPU ノードを約 65 万 **NH** 使用した。**auto-kappa** では、1 材料あたり 1 ノードを割り当てたジョブを多数同時に投入することでハイスループット計算を実現した。各ノード内では **VASP** および **ALAMODE** の計算を **MPI** 並列により実行し、ノード間通信のオーバーヘッドを生じさせない高効率な並列計算を実施した。**ALAMODE** によるメモリ増大にはノード内並列数の自動制御で対応した。

HPCI の枠組みでは、海外の共同研究者も計算資源を利用可能であり、米国・欧州・アジアの複数機関が参加した国際共同計算体制を構築した。この体制により、データ生成の大幅な加速と多様な材料系へのカバレッジ拡大が実現した。

2-4. 機械学習モデルの構築

得られたデータを用いてグラフニューラルネットワーク (**GNN**) に基づく機械学習モデルを構築した。スカラー量 (室温の格子熱伝導率: κ_p) の予測には **CGCNN** [4] を、スペクトル・累積熱伝導率 (κ_{cumul}) の予測には等変ニューラルネットワーク **e3nn** [5] を採用した。**CGCNN** では原子種・電気陰性度などの 9 種物性からなる 1-hot エンコーディングをノード記述子とし、**e3nn** では **SE(3)** 対称性を考慮した等変な 3 次元エッジ記述子を用いた。データは材料ベースで 80/10/10% に分割し、50 アンサンブルにより予測の不確かさも評価した。

3. 計算結果

3-1. Phonix データベースの構築

本研究では、**Phonondb** および **Materials Project** に収録された非金属・非磁性の結晶材料を中心

に約 2 万材料を対象とした。SQUID をはじめとする大規模計算機を用いたハイスループット計算により約 7,000 材料以上の非調和フォノン特性の計算を完了し、Phonix データベース (<https://phonix-db.org>) として公開した。現時点では約 7,000 材料の非調和フォノンデータが公開されており、フォノン分散・状態密度 (DOS)・格子熱伝導率・フォノン寿命・スペクトル熱伝導率などのデータを提供する。

Phonix データベースに収録された材料は幅広い結晶系・対称性を有し、ほぼ全ての空間群を網羅している。また、単位胞内の原子数は最大 160 個であり、複雑な結晶構造も含まれている。格子熱伝導率 κ_{lat} (300 K) の分布は 10^{-1} から 10^3 W/(m·K) 以上に広がり、平均値は約 2.4 W/(m·K)、全材料の 50% が 0.15~6.3 W/(m·K) の範囲に分布する。体積と熱伝導率の間には $\log_{10}(\kappa_{lat}) \propto -1.89 \log_{10}(V_{atom}) + 2.76$ の関係が確認された。また全材料の 8.4% でフォノン波動性の寄与 (κ_c) がフォノン粒子性の寄与 (κ_p) 以上になることが確認され、複雑な構造をもつ材料では波動的熱輸送が重要になることが示された。

3-2. 計算精度の検証

計算精度の検証のため、Phonondb データおよび 103 種の単結晶化合物の実験値と比較した結果、全体的に良好な一致が確認された。一部で計算値が実験値を上回る傾向があり、4フォノン散乱・スピン軌道相互作用・長距離相互作用などが現在の 3 フォノン計算パイプラインに含まれていないことが主な原因と考えられる。特に高熱伝導材料では 4 フォノン散乱の寄与が顕著であり、これを考慮した計算では実験値との一致が改善されることが確認されている。

また、超格子サイズや q 点密度の収束不足も一部の材料での誤差要因となりうる。超格子サイズを 200 原子以上、q 点密度をより高く設定した場合に精度が向上することが確認されており、今後の改善策として検討されている。機械学習モデルの活用により、こうした計算コストの高

い収束チェックを効率化することも今後の課題のひとつである。

3-3. 機械学習モデルの精度とスケーリング則

機械学習モデルの訓練結果として、学習データ数 (N_{train}) と予測誤差 (MAE) の関係において、 κ_p に対して $MAE \propto N_{train}^{-0.165}$ 、累積熱伝導率 (κ_{cumul}) に対して $MAE \propto N_{train}^{-0.141}$ というべき乗則スケーリングが確認された。これは大規模言語モデルで発見された深層学習における “neural scaling law” が熱物性予測にも成立することを初めて実証した結果であり、データ数の増加に伴い予測精度が継続的に向上することを意味する。 $N_{train} = 7,308$ のとき $MAE(test) = 0.282$ ($\log_{10}$ スケール) が達成された。

なお、高熱伝導材料の予測精度 ($MAE = 0.68$) は低熱伝導材料 ($MAE = 0.27$) と比べて低く、高熱伝導材料が単純な組成・構造を持ちモデルが学習できる構造情報が少ないことや、4フォノン散乱の無視による過大評価が原因と考えられる。

3-4. 大規模スクリーニング

構築したモデルを用いて GNoME データベース[6]の約 38 万材料をスクリーニングした結果、高熱伝導材料 ($\kappa_p^{3ph} > 200$ W/(m·K)) として 169 材料、低熱伝導材料として最低の 21 材料を特定した。高熱伝導材料の候補として六方晶系 P6₃/mmc 型の NpPH ($\kappa_{p,zz} = 277$ W/(m·K))、PaPH ($\kappa_{p,yy} = 264$ W/(m·K))、PuHS ($\kappa_p = 216$ W/(m·K)) が特定された。ただし 4 フォノン散乱を考慮すると NpPH の熱伝導率は約 78 W/(m·K) まで低下する。

低熱伝導材料の候補としては Cs₆Rb₂SnPbI₁₂ ($\kappa_{lat} \approx 0.15$ W/(m·K)) が最低値を示し、特定された 9 材料中 6 材料がセシウムを含む複雑な組成であった。これらの低熱伝導材料では異なる原子種のフォノンモードが広い周波数域に分散しており、非調和な相互作用を促進することが低熱伝導の原因と考えられる。

4. 考察

本研究により、第一原理計算に基づく大規模非調和フォノン特性データベースの構築と機械学習予測の有効性が実証された。スケーリング則の確認はデータ駆動型材料探索において重要な知見であり、大規模データ生成への継続的な取り組みが予測性能向上に直結することを示している。大阪大学 D3 センターの SQUID が提供する大規模並列計算環境は、本プロジェクトの高スループット計算を支える不可欠な基盤となった。

さらなる精度向上には、4フォノン散乱・フォノン繰り込み効果・電子-フォノン相互作用の導入が必要であり、これらの計算コストは3フォノン計算に比べて1~2桁増大する。機械学習ポテンシャル (MLP) の導入や利用ソフトウェアのGPU化による計算効率化が今後の重要な課題である。本データベースは格子熱物性に留まらず、電子-フォノン相互作用を通じた超電導転移温度・キャリア移動度や、スピン-フォノン相互作用を通じた磁気特性の理解にも活用できる可能性があり、超電導体・スピントロニクス材料など幅広い分野での応用が期待される。

Phonix データベースが幅広い空間群・原子数にわたる多様な結晶構造を含む点は、第一原理フォノンデータベースと比較して際立った特徴であり、普遍的な物性予測モデルの構築に向けた基盤データとしての価値は高い。今後は、自動実験と組み合わせた closed loop 型材料探索への応用が期待され、第一原理計算・機械学習・実験の三者を統合したデータ駆動型の材料開発プラットフォームの構築に向けた重要な基盤となることが期待される。

参考文献.

- [1] G. Kresse and J. Hafner, Phys. Rev. B 47, 558 (1993).
- [2] G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11169 (1996).
- [3] T. Tadano, Y. Gohda, and S. Tsuneyuki, J. Phys.: Condens. Matter 26, 225402 (2014).

- [4] T. Xie and J. C. Grossman, Phys. Rev. Lett. 120, 145301 (2018).
- [5] M. Geiger and T. Smidt, arXiv:2207.09453 (2022).
- [6] A. Merchant et al., Nature 624, 80 (2023).
- [7] M. Ohnishi et al., npj Computational Materials 12, 150 (2026). DOI: 10.1038/s41524-026-02033-w.