Symmetry relation database and its application to ferroelectric materials discovery

Qiang Zhu, Byungkyun Kang, and Kevin Parrish, Department of Physics and Astronomy, University of Nevada, Las Vegas, NV, 89154, USA
Address all correspondence to Qiang Zhu at qiang.zhu@unlv.edu

(Received 20 April 2022; accepted 6 September 2022; published online: 30 September 2022)

Abstract
To investigate the displacive phase transition at the atomic scale, we have implemented a numerical algorithm to automate the detection of the symmetry relations between any two candidate crystal structures. Using this algorithm, we systematically screen all possible polar–nonpolar structure pairs from the Materials Project database and establish a library of ~4500 pairs that can be connected through a continuous phase transition with small atomic displacements. From this database, we identify several new ferroelectric materials. In addition, the database may also be used in other areas, such as material structure prediction and new materials discovery.

Introduction
The solid–solid phase transitions refer to the phase change between different crystalline forms of the same compound. Among them, the phase transitions driven by temperature are of note. Though these transitions often involve just minor structural modifications through displacive mechanisms, they can have a profound effect on a variety of the solid’s physical properties. According to Landau theory,[1] continuous phase transitions require that there is a group–subgroup relation between the involved crystal structures. Knowing the atomic details that occur during the phase transition plays a pivotal role in materials research. For instance, detecting possible symmetry-related structures is the basis for the identification of polar materials that are likely to have a ferroelectric to paraelectric phase transition at high temperatures.[2,3] In addition, there is a great potential to reduce the complexity of the search space in crystal structure prediction[4,5] by utilizing symmetry relations.

The group-theoretical presentation of crystal-chemical relationships have been long established. It is possible to present symmetry relations between two crystal structures in a concise manner with a Barnighausen tree.[6] To investigate a single phase transition, an experienced researcher often resorts to the text book or online server to derive the symmetry relation by hand. Despite the fact that we are entering the data-driven era in materials research, most of previous efforts in phase transition were limited to case studies. To our knowledge, there is no attempt to systematically investigate the symmetry relation from the historic materials data, which prevents more effective materials design from the knowledge of crystallographic symmetry.

In this paper, we start by describing the implementation of a numerical algorithm that enables the high-throughput detection relation from any crystallographic data. Using this tool, we search the possible polar/nonpolar structure pairs from the entire Materials Project (MP) database and establish a database containing 4514 structure pairs that possess close symmetry relations. In particular, we identify 1234 potential ferroelectric–paraelectric pairs. Lastly, we discuss the possible research directions for utilizing the symmetry relation database as a tool for future materials development.

The symmetry relation detection algorithm
Relations between crystal structures can often be expressed by group–subgroup relations. A subgroup \( H \) of a space group \( G \) is called a maximal subgroup of \( G \) if there is no proper subgroup \( M \) in between, i.e., \( G \rightarrow M \rightarrow H \). There exist two kinds of \( H \) subgroups. The translationengleiche (t) subgroups keeps all translational symmetries but belong to a lower crystal class, while the klassengleiche (k) subgroups have fewer translations but are of the same crystal class. For a given space group, its maximal t and k subgroups are well defined and can be found in the crystallography book[7] or the online Bilbao Crystallographic Server.[8] To complete the transition between a group \( G \) and its maximal subgroup \( H \), one needs to know the cell transformation matrix, as well as the Wyckoff position (WP) splitting scheme for the given atomic positions.

If the phase transition follows the immediate symmetry relation \( (G \rightarrow H) \), one needs to (i) derive the unit cell transformation; (ii) map all WPs in both \( G \) and \( H \) to ensure that they follow the required splitting relation; and (iii) compute the distortions of unit cell vectors and atomic coordinates. While this procedure for mapping two structures is straightforward for a simple case such as the transition between \( \alpha \rightarrow \beta \) phases of silica quartz [see Fig. 1(a)], many transitions may follow a chain of subgroup relations between two crystals [e.g., \( G \rightarrow M_1 \rightarrow M_2 \rightarrow H \)], see an example...
transition between $\alpha \to \beta$ phases of silica cristobalite in Fig. 1(b)]. In such a case, a more complicated search for possible transition pathways in the large search space connecting $G$ and $H$ is required. When it comes to the low symmetry cases, the mapping of the crystal lattice between transition structures is not unique and there exists multiple choices for arrangements of the cell vectors. When it comes to the mapping of the different WP sites between the low and high-symmetry structures, the atomic coordinates may be continuously shifted to better match the reference states (i.e., the $G$ and $H$ structures). In this case, finding the optimum distortion path is not trivial as well.

**Numerical implementation**

To address the aforementioned challenges, we have developed the following search algorithm to check if two input crystals follow symmetry relations [see Fig. 1(c)].

1. **Space group transition path search** In order to find the path between two arbitrary space groups ($G$ and $H$), we will adopt the depth-first search algorithm for traversing tree data structures. Starting with high symmetry $G$, we will enumerate all possible maximum $t$ and $k$ subgroups ($[H_t]$), and then the next layer subgroups $[H_{t+1}]$ will be listed from each member in the previous layer $[H_t]$. The iteration will be terminated once it reaches the maximum number of layers; the branch will stop growing when the target $H$ is found. For $k$-subgroup transitions along this path, there exist an infinite number of ways to generate the new subgroup symmetry within the same point group by replicating the unit cell along a crystallographic axis. To account for such cases that require more intermediate $k$-subgroups, we run the search twice. In the first round, we only examine the subgroup chain sequence with the shortest possible transition sequence between $G$ and $H$. If this search does not lead to any valid solution at the end of Step 6 (i.e., no small distortion path can be found between $G$ and $H$), we will explore more pathways by considering extra $k$-type subgroups for each of the navigated transitions in the first round.

2. **Compatibility check** For each given path, we will transform the high symmetry structure into a subgroup setting ($H'$) and check if the resulting WP sites in the subgroup setting are consistent with those in the low symmetry structure. The goal is to ensure that groups of similar atomic species are compatible with the Wyckoff positions of both structures.

3. **Find the matched unit cell** If the target $H$ has a tetragonal symmetry or higher, one just need to directly compare the cell vectors between $H'$ and $H$. When $H$ belongs to the orthorhombic symmetry, the permutation of crystallographic $a$, $b$, $c$ axes also need to be considered in order to achieve the best match. For the monoclinic and triclinic

---

**Algorithm**

- **User inputs ($G$ and $H$)**
- **Search space group paths from $G$ to $H$**
- **Build subgroup representation $H'$**
- **Check compatibility between $H'$ and $H$**
- **Search valid unit cell in $H'$**
- **Build equivalent Wyckoff sets $\{H'\}$**
- **Identify atom mapping in each $H'$**
- **Optimize the translation vector**

---

**Figure 1.** The schematic phase transition. (a) and (b) Illustrate the mechanism for different silica polymorphs of quartz and cristobalite, respectively. (c) Outlines the workflow for the symmetry relation detection algorithm. For quartz, it follows a one step symmetry relation between the high temperature $\beta$ (hexagonal) and low temperature $\alpha$ (trigonal) phases. On the other hand, the transition between two phases of silica cristobalite requires multiple steps of group-subgroup transitions.
cells, one also needs to check other possible unit cell settings such as \((a + b, a - b, c)\) without changing the symmetry operation. Only the structures of \(H'\) having the close unit cell within a tolerance of 1.5 Å will be considered for the next step.

4. **Enumerate all equivalent Wyckoff sets \([H']\)** In a crystal structure, the atoms may be equivalently described by different sets of WPs. For instance, a FCC Cu can be considered as either WP 4a (0, 0, 0) or 4b (1/2, 1/2, 1/2). All different Wyckoff settings can be computed by the Euclidean and the affine normalizers of a given space group.\(^6\) For each \(H'\) with a valid unit cell representation, we will then enumerate all possible equivalent discrete sets of WPs in order to identify additional compatible transition pathways.

5. **Atom mapping** For each member in \([H']\), we will navigate all likely matched pairs between WP1 in \(H'\) and WP2 in \(H\). According to the choice of \(H\), there may exist a continuous translation to obtain the equivalent Wyckoff set. If translation is possible, we will perform an overall translation between WP1 and WP2. Finally, we compute the relative atomic displacements to match all WP pairs in \(H'\) and \(H\).

6. **Optimization of translation vector (optional)** If a continuous translation is allowed, we will also perform a gradient-free optimization to minimize the atomic displacements. Within the framework of Python, this can be conveniently implemented based on the Nelder-Mead algorithm,\(^9\) in the SciPy.\(^{10}\)

This algorithm has been implemented in the open source code PyXtal.\(^{11}\) Executing the example script from Fig. 2(a) will generate a transition path (represented by 5 intermediate structures) with the smallest possible atomic distortion that satisfies the symmetry relation.

**Validation**

In a recent work,\(^3\) Smidt and coworkers proposed 413 structure pairs that may possess continuous transformation. They also provide an online dataset with detailed information characterizing the structural, energetic and polarization quantities. To validate our algorithm implementation, we systematically checked the transition path and computed the required atomic displacements for each of the pairs. Figure 2(b) shows the comparison of maximum distortion values between our results and the previous work. In general, our work identifies successful transition paths with smaller atomic distortions compared to those of the previous study. In that dataset, there exist 27 pairs without records for atomic distortion [shown as the green points in Fig. 2(b)], but our code can generate successful paths for these pairs with reasonably small atomic distortions.

**Symmetry relation database**

With the new implementation of our automated symmetry detection algorithm, we are able to screen more materials data on a large scale. In general, researchers are interested in the phase transition involving change in polarization as it is important for a number of applications, such as ferroelectricity and piezoelectricity. Therefore, we proceed to investigate symmetry relations between polar and nonpolar structure pairs with the newly developed toolkit.

**Screening of polar/nonpolar structural pairs**

From the available materials data in the MP database,\(^{12}\) we made an initial query to obtain 28,355 unique compounds belonging to polar space groups. For each polar structure, we make another query to extract the corresponding nonpolar...
structures which may possess super group symmetry relative to the reference polar structures. We find 239,101 potential non-polar-polar symmetry pairs created from 7775 low-symmetry polar and 8051 high symmetry nonpolar structures. Using the new symmetry tool, we systematically for phase transitions with a maximum atomic distortion smaller than 1.8 Å and lattice mismatch smaller than 1.5 Å in length and 10° in angle. This generates 4514 ideal structural pairs (see details in the supplementary material).

Figure 3 summarizes the overall statistics. In total, we identify 20 elemental, 934 binary, 2078 ternary, 1382 quaternary and 100 pentanary polar materials. Among the 4514 structural pairs, over 55.2% have a maximum atomic distortion smaller than 1.0 Å. In addition, most of the structure pairs (91.7%) have an energy difference ($\Delta E$) less than 200 meV/atom. Among the entire dataset, 2302 (51.0%) pairs can be considered as high quality data to study, with phase transition within the tolerances of 1.0 Å in distortion and 250 meV/atom in energy difference. Finally, Fig. 3(d) shows the frequency of point group relations. The six most frequent transitions are $\bar{1} \rightarrow 1$ (21.9%), $mmm \rightarrow mm2$ (14.6%), $2mm \rightarrow 1$ (9.9%), $2/m \rightarrow m$ (9.6%), $6/mmm \rightarrow 6mm$ (5.4%). Knowing such statistics will be useful to develop a model to predict likely continuous phase transitions for designing future new materials.

Ferroelectric materials identification

From our symmetry relation database, we obtain 1234 potential ferroelectric material pairs by applying four filters: (i) the space group relation must belong to 88 potential paraelectric to ferroelectric phase transitions; (ii) the maximum distortion is smaller than 1.0 Å; (iii) both polar and non-polar phases have nonzero band gaps; and (iv) the energy difference between the polar and non-polar phases should be smaller than 100 meV/atom. Then, we systematically compute the pair’s phase transition barrier and the polar structure’s polarization by following the automated workflow as described in previous work based on the VASP package.[14] Furthermore, we also choose some representative materials as candidates for further analysis. We calculate their free energy curves as a function of different external electric fields based on Landau theory and their electric hysteresis loops by finding optimized ionic positions using Hessian matrix and Born effective charges.[16,17]

Figure 4 shows the results for two representative materials HfO₂ and AlHO₂ from our database. The $Pca_{2}1 \rightarrow P4_{2}/nmc$ transition of HfO₂ has been extensively studied in the past.[18,19] Figure 4(a) displays HfO₂’s potential energy curve as a function of polarization for various electric fields. At a zero electric field, the energy curve shows a typical double well behavior for ferroelectric materials, and the polar phase has a polarization of 55.2 µC/cm². When electric fields are applied, the free energy of the left local minimum gradually increases and the energy barrier vanishes at the coercive field (16 MV/cm). These results are consistent with previous reports based on the computational code and DFT functional.[18] Using this workflow, we also investigate the ferro-paraelectric phase transition of AlHO₂, a system that has not been considered in any previous literature. From the MP database, we find that there is a close symmetry relation between the $Pmn21$ (mp-23807) and $Pnnm$ (mp-1078520) phases. Compared to HfO₂, the polar AlHO₂ phase has a smaller polarization (23.1 µC/cm²). However, the coercive field of AlHO₂ (15 MV/cm) is close to that of HfO₂. In addition to AlHO₂, our screening identified seven other materials that have not been studied in the past. These materials will be investigated in future.
Conclusions

In this work, we introduce the concept of a symmetry relation database. To enable the investigation of displacive phase transitions from historic materials data, we have implemented a numerical algorithm into the open-source code PyXtal that can automate the detection of symmetry relations between any two candidate crystal structures of identical chemical composition. We apply the tool to systematically screen all possible polar-nonpolar structure pairs from the entire MP database and establish a symmetry relation database of ~4500 structure pairs with possible continuous phase transition paths through small atomic displacements. From this symmetry relation database, we identify several new ferroelectric materials that have never been reported in the past.

In addition to the search of ferroelectric materials, the concept of symmetry relation can be extended to guide materials discovery in a broader sense. Most open materials databases\textsuperscript{[12,20,21]} rely on the data mining techniques\textsuperscript{[22]} augment available materials with hypothetical compounds constructed by decorating common structural prototypes with elements in the periodic table. In the past, chemists often found it useful to design or predict new structures by substituting one elements into two based on the isoelectronic concept (e.g., from cubic diamond type elemental C to compound BN,\textsuperscript{[23]} or from post-perovskite type Fe\textsubscript{2}O\textsubscript{3} to MgSiO\textsubscript{3}\textsuperscript{[24]}). But, those works were limited to only a few case studies due to the lack of a database that can systematically describe such symmetry relations. In the future, we will extend the analysis tool to support the symmetry relation detection between different structural prototypes. Thus, one can design more powerful data mining models through more sophisticated chemical substitution.

Acknowledgments

This research was sponsored by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Theoretical Condensed Matter Physics program and the DOE Established Program to Stimulate Competitive Research under Award Number DE-SC0021970. The computing resources are provided by XSEDE (TG-DMR180040). Q.Z thanks Dr. Amois Arroyo and Dr. Shunbo Hu for inspiring discussions regarding symmetry relation and ferroelectric materials.

Figure 4. The computational characterization of ferroelectric materials HfO\textsubscript{2} and AlHO\textsubscript{2}. (a) and (c) Show the free energy curves as a function of polarization for different electric fields. (b) and (d) Display the corresponding computed hysteresis curves at zero temperature.
Data availability
All data, codes and scripts generated during and/or analysed during the current study are available at https://github.com/qzhu2017/PyXtal_symmetry_relation.

Declarations
Conflict of interest
All authors declare that they have no conflict of interest.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1557/s43579-022-00268-4.

References
1. L.D. Landau, E.M. Lifshitz, *Statistical physics*, 5th edn. (Elsevier, Amsterdam, 2013)
2. P.-P. Shi, Y.-Y. Tang, P.-F. Li, W.-Q. Liao, Z.-X. Wang, Q. Ye, R.-G. Xiong, Symmetry breaking in molecular ferroelectrics. Chem. Soc. Rev. 45, 3811–3827 (2016). https://doi.org/10.1039/C5CS00308C
3. T.E. Smidt, S.A. Mack, S.E. Reyes-Lillo, An automatically curated first-principles database of ferroelectrics. Sci. Data 7, 1–22 (2020). https://doi.org/10.1038/s41597-020-0407-9
4. Q. Zhu, A.R. Oganov, A.O. Lyakhov, Evolutionary metadynamics: a novel method to predict crystal structures. CrystEngComm 14, 3596–3601 (2012). https://doi.org/10.1039/C2CE06642D
5. Q. Zhu, A.R. Oganov, A.O. Lyakhov, X. Yu, Generalized evolutionary metadynamics for sampling the energy landscapes and its applications. Phys. Rev. B 92, 024106 (2015). https://doi.org/10.1103/PhysRevB.92.024106
6. H. Wondratschek, U. Müller, Symmetry relations between space groups (International Union of Crystallography, Chester, 2006). https://doi.org/10.1107/S0003709906011714
7. M.I. Aroyo, *International tables for crystallography* (Wiley Online Library, Hoboken, 2016). https://doi.org/10.1107/97809553602060000114
8. M.I. Aroyo, J.M. Perez-Mato, C. Capillas, E. Kroumova, S. Cholia, D. Gunter, D. Skinner, G. Ceder et al., Commentary: The materials project: a materials genome approach to accelerating materials innovation. APL Mater. 1, 011002 (2013). https://doi.org/10.1063/1.4812323
9. K. Aizu, Possible species of ferromagnetic, ferroelectric, and ferroelastic crystals. Phys. Rev. B 2, 754–772 (1970). https://doi.org/10.1103/PhysRevB.2.754
10. G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. Phys. Rev. B 54, 11169 (1996). https://doi.org/10.1103/PhysRevB.54.11169
11. P. Chandra, P.B. Littlewood, A landau primer for ferroelectrics, in *Physics of ferroelectrics* (Springer, Heidelberg, 2007), pp.69–116. https://doi.org/10.1007/978-3-540-34591-6_3
12. D. Karanth, H. Fu, Large electromechanical response in ZnO and its microscopic origin. Phys. Rev. B 72, 064116 (2005). https://doi.org/10.1103/PhysRevB.72.064116
13. T. Maeda, B. Magyari-Kope, Y. Nishi, Identifying ferroelectric switching pathways in HfO2: first principles calculations under electric fields, in 2017 IEEE International Memory Workshop (IMW) (IEEE, 2017), pp.1–4. https://doi.org/10.1109/IMW.2017.7939087
14. A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder et al., Commentaries: The materials project: a materials genome approach to accelerating materials innovation. APL Mater. 1, 011002 (2013). https://doi.org/10.1063/1.4812323
15. H. Fu, L. Bellaiche, First-principles determination of electromechanical responses of solids under finite electric fields. Phys. Rev. Lett. 91, 057601 (2003). https://doi.org/10.1103/PhysRevLett.91.057601
16. D. Karanth, H. Fu, Large electromechanical response in ZnO and its microscopic origin. Phys. Rev. B 72, 064116 (2005). https://doi.org/10.1103/PhysRevB.72.064116
17. T. Maeda, B. Magyari-Kope, Y. Nishi, Identifying ferroelectric switching pathways in HfO2: first principles calculations under electric fields, in 2017 IEEE International Memory Workshop (IMW) (IEEE, 2017), pp.1–4. https://doi.org/10.1109/IMW.2017.7939087
18. P. Fan, Y. Zhang, Q. Yang, J. Jiang, L. Jiang, M. Liao, Y. Zhou, Origin of the intrinsic ferroelectricity of hfo2 from ab initio molecular dynamics. J. Phys. Chem. C 123, 21743–21750 (2019). https://doi.org/10.1021/acs.jpcc.9b04106
19. S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R.H. Taylor, L.J. Nelson, G.L. Hart, S. Sanvito, M. Buongiorno-Nardelli et al., Aflowlib.org: a distributed materials properties repository from high-throughput ab initio calculations. Comput. Mater. Sci. 58, 227–235 (2012). https://doi.org/10.1016/j.commatsci.2012.02.002
20. S. Kirklin, J.E. Saal, B. Meredig, A. Thompson, J.W. Doak, M. Aykol, S. Rühl, C. Wolverton, The open quantum materials database (oqmd): assessing the accuracy of DFT formation energies. NPJ Comput. Mater. 1, 15010 (2015). https://doi.org/10.1038npjcompumats.2015.10
21. G. Hautier, C. Fischer, V. Ehrlacher, A. Jain, G. Ceder, Data mined ionic substructures for the discovery of new compounds. Inorg. Chem. 50, 656–663 (2011). https://doi.org/10.1021/ic102031h
22. R. Wentorf Jr., Cubic form of boron nitride. J. Chem. Phys. 26, 956–956 (1957). https://doi.org/10.1063/1.1745964
23. A.R. Oganov, S. Ono, Theoretical and experimental evidence for a post-perovskite phase of MgSiO3 in earth’s D* layer. Nature 430, 445–448 (2004). https://doi.org/10.1038/nature02701

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.