



Citation for published version:

De Souza, R & Harrington, G 2023, 'Revisiting point defects in ionic solids and semiconductors', *Nature Materials*, vol. 22, no. 7, pp. 794-797. <https://doi.org/10.1038/s41563-023-01583-4>

DOI:

[10.1038/s41563-023-01583-4](https://doi.org/10.1038/s41563-023-01583-4)

Publication date:

2023

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

Unspecified

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **Revisiting point defects in ionic solids and semiconductors**

2

3

4 Roger De Souza[†] and George Harrington[§]

5 *Institute of Physical Chemistry, RWTH Aachen University, Aachen, Germany*

6

7 [†]Corresponding author. e-mail: desouza@pc.rwth-aachen.de

8 [§]Current address: Department of Chemistry, University of Bath. e-mail: gh778@bath.ac.uk

9

1 **The study of point defects in non-metallic crystals has become relevant for an increasing number of**
2 **materials applications. Progress requires a foundation of consistent definitions and terminology.**
3 **This comment clarifies the underlying definitions of point defects, encourages the correct use of**
4 **relative charge for their description, and emphasises their recognition as quasiparticles.**

5
6 To the layman and to the newcomer, the word “defect” has only unpleasant connotations. Defects are
7 bad, always undesirable. If unavoidable, their number should be minimised. To those who study defects
8 in crystals, however, the word has various connotations. In some cases, defects are detrimental, but in
9 many others, they are beneficial or even essential, imparting function to ‘functional’ materials. The
10 layman’s reductive standpoint of “bad” defects is thus replaced by the defect scientist’s detailed, nuanced
11 picture of detrimental or useful defects.

12 Interest in the “defect” picture has flourished in recent years, principally in terms of defect studies in
13 myriad new systems, but also in terms of defect studies with new experimental and computational
14 methods. And yet, and perhaps inevitably, the interest is accompanied by developments that tend to blur
15 rather than illuminate some part of the picture. This comment reviews and refines certain fundamental
16 aspects of point defects in ionic solids and semiconductors, in order to counteract three unwelcome
17 developments: a lack of awareness of how defects are defined, a lack of awareness of why defects are
18 defined, and a lack of awareness of how defects can be described in a concise and unambiguous manner.

19

20 **The increasing importance of point defects**

21 The burgeoning interest in defects can be readily understood by considering the applications for
22 materials in which defects, either by design or unintentionally, determine the performance. Not so long
23 ago, materials with large defect concentrations, often referred to as non-stoichiometric compounds, were
24 only found in a handful of niche applications, the most widespread arguably being potentiometric sensors
25 in automotive exhaust systems. In recent years, the situation has changed dramatically. Non-
26 stoichiometric compounds have fully permeated our society. Li-ion batteries with their intercalation
27 electrodes, having revolutionised portable electronics, are set to reform automotive transport and can

1 now be readily found in applications from bicycles to power tools. The future will require materials with
2 defect-determined properties to play an even larger role, if pledges in line with the Paris Agreement are
3 to be met. Hydrogen production at-scale is needed within the next 30 years (at most) in order to
4 decarbonise fertiliser, steel and cement manufacturing, and electrolyser cells based on oxide-ion
5 conducting solids are one of the foremost technologies capable of achieving this¹. Replacing thermal
6 power plants with photovoltaics is also required for decarbonisation, and hybrid-perovskite cells, already
7 offering advantages in terms of processing simplicity, are now challenging Si-based cells on efficiency²;
8 device performance and lifetimes are limited, however, by defect-dependent phenomena^{3,4}. Even in
9 computing, as the physical limits of Si-based MOSFET technology are pushed after more than 60 years
10 of miniaturisation, memristors based on the field-driven migration of ionic defects may usher in a new
11 wave of advanced programmable logic devices and physical neural network-based computing^{5,6}. In other
12 words, it is looking increasingly likely that many of the developments in transport, energy, and
13 computing over the coming decades will stem from our ability to manipulate the type, number and spatial
14 distribution of point defects in non-metallic crystals.

15

16 **Definition, Terminology and Rationale**

17 Definitions constitute a cornerstone of a discipline by providing, first and foremost, a common language.
18 In the case of defects, this cornerstone often seems to be overlooked, and consequently without the
19 shared vocabulary, it is sometimes unclear if we are all talking about the same thing.

20 Defects are generally defined by subtracting an ideal (or reference) crystal from a real crystal: defect =
21 real – ideal. For point defects this definition reveals deviations from the ideal structure of non-metallic
22 crystals that can be divided into missing ions (termed vacancies), additional ions (termed interstitials),
23 foreign ions (termed dopants or impurities depending on the context; and situated substitutionally or
24 interstitially), and native ions sitting on the wrong site (termed antisites). Point defects can also be
25 identified from thermodynamics, namely as any atomic entity that adds to the configurational entropy of
26 the ideal crystal. Importantly, this second definition shares key features with the first definition: it also
27 requires a reference state; it is also formulated in terms of a difference; and it yields the same entities
28 (vacancies, interstitials, etc).

1 This brings us on to the question of why we define defects. This simple answer is because we have to.
2 The pioneering work of Frenkel, Schottky, Wagner and Jost in the 1930s^{7–13} revealed that at finite
3 temperature a perfect crystal is thermodynamically unstable: Point defects are an equilibrium component
4 of crystals. Indeed, chemical potentials can be defined for point defects (as building units), and this
5 permits chemical reactions involving point defects to be formulated (which thus established the field of
6 defect chemistry). A second answer to the question of why we define defects is that it is extremely
7 useful^{14–17}. Or to put it in a way that is rarely appreciated: point defects are quasiparticles. At the level of
8 nuclei and electrons, or even at the level of ions, solids are complex many-body systems, and quantitative
9 treatments of their properties rapidly become intractable if the descriptions are conducted at the level of
10 the constituent particles. By defining quasiparticles, we greatly simplify quantitative treatments. The
11 most well-known quasiparticle is the electron hole in a semiconductor. Rather than considering all the
12 electrons in the material's valence band, we focus on those electrons that are missing. Point defects are
13 quasiparticles because, as with other quasiparticles, they do not exist outside of the crystal, they obey
14 Fermi–Dirac statistics, and their definition allows for quantitative treatments of materials' properties.
15 Instead of treating a huge number of strongly interacting ions, say, we focus on a drastically smaller
16 number of point defects, which may be non-interacting in the simplest case, or if they do interact, can be
17 treated through pair-wise rather than multi-body interactions. We define point defects, therefore, for
18 necessity and for simplicity.

19 Another cornerstone of a discipline is the consistent use of terminology to ensure that entities under
20 discussion adhere to the basic definitions. In this case *crystalline* or *crystallographic point defects* is the
21 term used. This distinguishes such entities from higher dimensional or extended crystalline defects, such
22 as dislocations and grain boundaries, as well as from defects in amorphous structures (as discussed
23 below). In most cases, as we do here, the term *point defect* is used as an abridged form. Unfortunately, in
24 some instances this leads to ambiguity and confusion, with claims of “defect-free” materials (not
25 thermodynamically stable at finite temperatures).

26 The word ‘defect’, as we noted above, has unpleasant connotations in general; its alternatives
27 (imperfection, flaw, mistake, error, blemish) are hardly an improvement. It is tempting, therefore, to
28 suggest that an entirely new term should be used. While this might seem radical, it would not be

1 unprecedented. The word ‘hole’ was originally used to describe a missing atom within a crystalline
2 material before it was appropriated in the 1940s (during the development of semiconductor electronics)
3 to describe an electron missing from the valence band. A new word was required to fill the hole (pun
4 intended) and ‘vacancy’ became the agreed-upon term¹⁸. Replacing the word ‘defect’, however, would in
5 our view be fundamentally misguided. Crystal defects are defects. Using a different word would only
6 provide a superficial veneer; and it would entail, arguably, a shift in focus away from the underlying
7 definition.

8 Based on both definitions given above for identifying crystalline point defects, it would appear that there
9 can be no point defects in an amorphous material because there is no uniquely defined amorphous
10 reference structure and hence no unambiguous method for identifying deviating atomic entities.

11 Alternatively, from the point of view of a reference crystal, the entire amorphous material is, essentially,
12 a crystalline defect. That said, electronically active centres are known in many amorphous
13 semiconductors (and associated with dangling bonds)^{19,20}. In addition, there are amorphous materials for
14 which a topological treatment yields atomic-scale deviations²¹. Such approaches are unlikely to apply to
15 all amorphous materials, however, in the way that the point-defect definitions are universal to all
16 crystalline solids. Furthermore, applying different approaches to an amorphous system may not
17 necessarily identify the same species (cf. the two different definitions for point defects that both yield the
18 same species). Consequently, it may be more helpful to use different terms for defects in amorphous
19 materials: vacancy-like or quasi-vacancy and interstitial-like or quasi-interstitial are reasonable
20 possibilities²², but completely different names would perhaps be more appropriate; in any case, vacancies
21 and interstitials should be reserved for crystals. Certainly, point-defect notations that indicate a
22 crystallographic site (see below) should be avoided in the case of amorphous materials because a
23 reference crystal will not have been used to identify deviating atomic entities. Lastly, we note that
24 defining point defects in crystals is not simply an academic exercise: it has unquestionable and
25 demonstrable benefits. One should ensure that this is also the case when defining defects in amorphous
26 systems.

27

28

1 Defect notation and defect charge

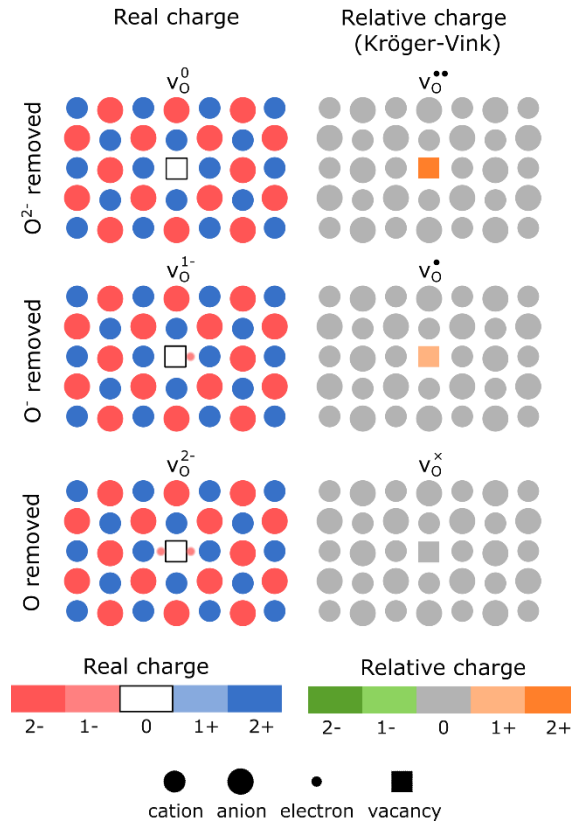
2 A notation is used to convey specific information in a concise, unambiguous form. Writing Ca^{2+} , for
3 example, conveys that a particular chemical element in a specific charge state is being considered. In the
4 case of point defects, the definitions described above (with respect to an ideal structure) demand a
5 different notation, as does the need to distinguish defects from conventional chemical species. Various
6 notations have been proposed^{23,24} over the years, but it is the notation derived by Kröger and Vink²⁵ that
7 has established itself almost universally. It consists of a main symbol, indicating the chemical species,
8 the electronic species (e is used for an electron and h for an electron hole) or a vacant site; a subscript,
9 indicating the site in the crystal or an interstitial site; and a superscript indicating the relative charge (\times
10 for neutral, \cdot for relative positive charge, and $'$ for relative negative charge), or more seldom, the real
11 charges (0 for neutral, $^+$ for real positive charge, and $^-$ for real negative charge).

12 Standard Kröger–Vink notation uses V for a vacancy and I for an interstitial site. This is clearly
13 problematic because it leads to confusion with the chemical elements vanadium and iodine. Norby²⁶
14 proposed to remove these ambiguities by using lower-case symbols, v for a vacancy and i for an
15 interstitial site. The benefits are easily recognised by considering, as Norby suggested, point defects in
16 vanadium iodide. In the standard notation, V_I is not unambiguous, for it could be one of four species; in
17 Norby's version, the four species are clearly identified: v_I (an iodine vacancy), V_i (a vanadium
18 interstitial), v_i (a vacant interstitial site) and V_I (a vanadium antisite defect on an iodide-ion site).

19 As mentioned above, the burgeoning interest in defects has led to some unwelcome developments. One
20 such development is the use of O_V to indicate v_O or of I_O to indicate O_i (focussing on anion defects as
21 examples). In addition to being in conflict with the standard (Kröger–Vink) notation, this approach
22 relegates the most immediately important information, the species being described, to the subscript. This
23 relegation creates further confusion when discussing substitutional or antisite defects. For example, Gd
24 sitting on a Zr site would have to be written as Zr_Gd rather than Gd_Zr to be consistent, which is needlessly
25 confusing to the reader.

26 The second unwelcome development is the use of real charge to indicate relative charge. This is wrong
27 because they are two different quantities: the former is defined relative to free space; the latter, relative to
28 a crystal lattice (see above). And since there is a difference, the symbols used to indicate relative charge

1 have to be different to those used to indicate real charge, otherwise there is the possibility of ambiguity,
 2 or worse, error. In order to make the difference clear, we consider in the following an oxide crystal, in
 3 which either O^{2-} , O^- or O has been removed from a regular O^{2-} site. (NB: In talking about real charges we
 4 use for good reason²⁷ the nominal charge of the ions; and similarly, for the relative charge of defects.)



5
 6 **Fig. 1 | Describing point defects with real charge or relative charge.** Three different charge states of
 7 an oxygen vacancy in a metal oxide MO are shown with their real charges or their relative charges (in
 8 Kröger-Vink notation^{25,26}). One benefit of the relative-charge description is that the relevant entities (the
 9 quasiparticles) are immediately apparent. Another benefit is that the response of defects to an applied
 10 electric field is also immediately apparent: $v_O^{\bullet\bullet}$ and v_O^\bullet move; v_O^\times and O_O^\times don't, as they are neutral. If real
 11 charges are used, the treatment needs to include O_O^{2-} , as well as v_O^0 , v_O^- and v_O^{2-} , and motion is restricted
 12 to those ions that have a neighbouring vacancy in the direction of ion drift. After careful consideration,
 13 O_O^{2-} jumping to v_O^0 is found to constitute charge transport, whereas O_O^{2-} jumping to v_O^{2-} does not. By
 14 using real charges rather than relative charges, we obtain the same information but we spurn one of the
 15 benefits of defining defects — the simplicity of the treatment.

16
 17 The removal of these three O species generates the defects, in Kröger–Vink notation, $v_O^{\bullet\bullet}$, v_O^\bullet and v_O^\times . In
 18 terms of real charge, these same three species are v_O^0 , v_O^- and v_O^{2-} . Both descriptions indicate the same

1 information: an oxygen site missing an oxygen atom and missing 2, 1 or 0 electrons (Figure 1). The use
 2 of real charge to indicate relative charge, i.e. v_O^{2+} , v_O^+ and v_O^0 to indicate $v_O^\ddot{}$, v_O^\cdot and v_O^\times , should be
 3 avoided because, besides being inconsistent with how defects are defined (defect \neq real – ideal, see
 4 above), it fails as a notation: It does not convey information correctly. By using a real charge to indicate
 5 a relative charge, it gives neither the correct real charge nor the correct relative charge. In fact, taken at
 6 face value, v_O^{2+} denotes a site at which one oxygen atom and four electrons have been removed, whereas
 7 what the writer of v_O^{2+} wants (but fails) to indicate is a site missing one oxygen atom and two missing
 8 electrons.

9 Indicating the correct relative or absolute charge of defects is also important because defects can change
 10 their charge state by trapping or de-trapping electronic species. Taking the examples of isolated anion
 11 and cation vacancies in the oxide MO, we have v_O^\times donating one electron (e') to the conduction band to
 12 become v_O^\cdot , and then a second to become $v_O^\ddot{}$ (oxygen vacancies are donor-type defects); and v_M^\times
 13 accepting one electron from the valence band (i.e. donating a hole) to become v_M' , and then a second to
 14 become v_M'' (cation vacancies are acceptor-type defects). The energies of these reactions correspond to
 15 energy levels within the electronic bandgap. These levels are correctly described as v_O^\times/v_O^\cdot and $v_O^\cdot/v_O^\ddot{}$
 16 transitions, and v_M^\times/v_M' and v_M'/v_M'' transitions.

17 Writing defect reactions with the wrong notation reveals a further problematic aspect. The anti-Frenkel
 18 reaction in an oxide, written with relative charges, is

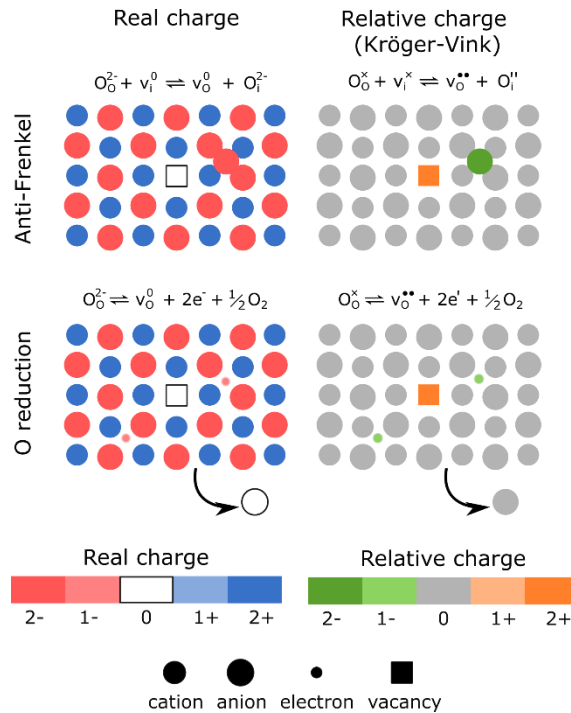


20 or with real charges,



22 Both equations communicate the same information, and in most cases, the first equation will be
 23 preferable (see Figure 2). If we start, however, with v_O^{2+} (the wrong notation for an oxygen vacancy
 24 bearing a double relative charge), the charge-balanced reaction is $O_O^0 + v_i^0 \rightleftharpoons v_O^{2+} + O_i^{2-}$. That is, we are
 25 forced to write on the reaction's left-hand side that an oxide ion sitting on its regular site has zero real
 26 charge, even as we imply through both oxygen defects on the right-hand side that the oxide ion is doubly

1 charged. This problem is not restricted to the anti-Frenkel reaction. The reduction of an oxide to form
 2 electrons and oxygen vacancies reads: $O_O^x \rightleftharpoons v_O^{\cdot\cdot} + 2e' + \frac{1}{2}O_2$, in terms of relative charge; or $O_O^{2-} \rightleftharpoons$
 3 $v_O^0 + 2e^- + \frac{1}{2}O_2$, in terms of real charge. The wrong version, $O_O^0 \rightleftharpoons v_O^{2+} + 2e^- + \frac{1}{2}O_2$, again requires
 4 an oxide ion sitting on its regular site to have zero real charge. A consistent, unambiguous defect notation
 5 avoids, therefore, the use of real charges to indicate relative charges.



6
 7 **Fig. 2 | Describing point-defect reactions with real charge or relative charge.** Equivalent descriptions
 8 of anti-Frenkel disorder and oxide reduction in terms of real charges and relative charges. Again, one
 9 benefit of the relative-charge description is that the relevant entities (the quasiparticles) are immediately
 10 apparent.

11
 12 We have seen that point defects can be described unambiguously either with relative charge or with real
 13 charge. In most cases, relative charge is preferred, but not in all. Our recommendation is to use one or the
 14 other. Mixing relative and real charge will generally result in errors or cosmetic coverings of errors, in
 15 particular when writing defect reactions but also in other cases. For example, when considering the
 16 relaxation of ions surrounding an oxygen vacancy in an ionic crystal, it is sometimes incorrectly claimed
 17 that there is a repulsive interaction between the positively charged oxygen vacancy and the positively
 18 charged cations. This is incorrect because the defect has a *relative* positive charge whereas the cations

1 have a *real* positive charge. Avoiding the combination of real and relative charge, one readily finds that
2 there is no interaction: in terms of relative charge, the cations are neutral and the defect is charged,
3 whereas in terms of real charge, the cations are charged, but the vacancy is neutral. The relaxation of the
4 cations away from the vacancy is in fact best understood in terms of real charge. From the direction of
5 the vacancy, the cations experience less Coulomb attraction (the oxide ion is missing), and this results in
6 the cations relaxing away from the oxygen vacancy. A second case for which the treatment in terms of
7 real charge is preferred is the transfer of species between two phases, e.g. between an electrolyte and an
8 electrode.

9 **Building on solid foundations**

10 The era of point defects started, arguably, close to a century ago, with Pohl's work on colour centres in
11 alkali halides²⁸. Frenkel, Wagner, Schottky and Jost⁷⁻¹³, through their contributions to point-defect
12 thermodynamics and kinetics, put the subject on a firm foundation. The period from 1949 to 1959 has
13 been called by Nowick¹⁸ a golden age in the development of concepts in crystalline defects, specifically
14 because definitions crystallised and a notation²⁵ emerged. It is on these foundations that the interest in
15 crystal defects grew continually over the intervening decades.

16 The coming years will undoubtedly witness a further increase. Even more people will be confronted with
17 materials systems and devices, in which point defects play a central role. If progress is not to be
18 hindered, a common language in terms of definitions, terminology, and notation is indispensable.
19 Attention needs to be paid to the (reliable and consistent) foundations. Only in this way will we
20 accelerate developments, perhaps even ushering in a new golden age of point defects.

21

22

23

24 The authors declare no competing interests.

25

1 References

- 2 1. Nikolaidis, P. & Poullikkas, A. A comparative overview of hydrogen production processes.
3 *Renewable and Sustainable Energy Rev.* **67**, 597–611 (2017).
- 4 2. Correa-Baena, J.-P. *et al.* Promises and challenges of perovskite solar cells. *Science* **358**, 739–744
5 (2017).
- 6 3. Senocrate, A. & Maier, J. Solid-State Ionics of Hybrid Halide Perovskites. *J. Am. Chem. Soc.* **141**,
7 8382–8396 (2019).
- 8 4. Zhou, Y., Poli, I., Meggiolaro, D., De Angelis, F. & Petrozza, A. Defect activity in metal halide
9 perovskites with wide and narrow bandgap. *Nat. Rev. Mater.* **6**, 986–1002 (2021).
- 10 5. Waser, R. & Aono, M. Nanoionics-based resistive switching memories. *Nat. Mater.* **6**, 833–840
11 (2007).
- 12 6. Yang, J. J., Strukov, D. B. & Stewart, D. R. Memristive devices for computing. *Nat. Nanotechnol.*
13 **8**, 13–24 (2013).
- 14 7. Frenkel, J. Über die Wärmebewegung in festen und flüssigen Körpern. *Z. Phys.* **35**, 652–669
15 (1926).
- 16 8. Wagner, C. & Schottky, W. Theorie der geordneten Mischphasen. *Z. Phys. Chem. B* **11**, (1930).
- 17 9. Wagner, C. Theorie der geordneten Mischphasen II. *Z. Phys. Chem. B*, 177, (1931).
- 18 10. Wagner, C. Theorie der geordneten Mischphasen. III Fehlordnungerscheinungen in polaren
19 Verbindungen als Grundlage für Ionen- und Elektronenleitung. *Z. Phys. Chem. B* **22**, 181–194
20 (1933).
- 21 11. Jost, W. Diffusion and Electrolytic Conduction in Crystals (Ionic Semiconductors). *J. Chem.*
22 *Phys.* **1**, 466–475 (1933).
- 23 12. Schottky, W. Über den Mechanismus der Ionenbewegung in festen Elektrolyten. *Z. Phys. Chem.*
24 **29B**, 335–355 (1935).
- 25 13. Jost, W. *Diffusion und chemische Reaktion in festen Stoffen.* (T. Steinkopff, 1937).
- 26 14. Rickert, H. *Electrochemistry of Solids: An Introduction.* (Springer-Verlag Berlin, 1982).
- 27 15. Schmalzried, H. *Chemical Kinetics of Solids.* (Wiley-VCH Verlag GmbH, 1995).
- 28 16. Smyth, D. M. *The Defect Chemistry of Metal Oxides.* (Oxford University Press, 2000).
- 29 17. Maier, J. *Physical Chemistry of Ionic Materials: Ions and Electrons in Solids.* (Wiley, 2004).
- 30 18. Nowick, A. S. The Golden Age of Crystal Defects. *Annu. Rev. Mater. Sci.* **26**, 1–19 (1996).
- 31 19. Mott, N. F. Defects in non-crystalline materials. *Phil. Mag. B* **51**, 177–182 (1985).
- 32 20. Stutzmann, M. The defect density in amorphous silicon. *Phil. Mag. B* **60**, 531–546 (1989).
- 33 21. Hiraoka, Y. *et al.* Hierarchical structures of amorphous solids characterized by persistent
34 homology. *Proc. Natl. Acad. Sci.* **113**, 7035–7040 (2016).
- 35 22. Popescu, M. Defect formation in amorphous structures as revealed by computer simulation. *Thin*
36 *Solid Films* **121**, 317–347 (1984).
- 37 23. Schottky, W. & Stöckmann, F. Comparative considerations on the nature of defects in
38 semiconductors and phosphors. in *Halbleiterprobleme, Advances in Solid State Physics*, 80–106
39 (1954).

- 1 24. Rees, A. L. G. *Chemistry of the Defect Solid State*. (Methuen, 1954).
- 2 25. Kröger, F. A. & Vink, H. J. Relations between the Concentrations of Imperfections in Crystalline
3 Solids. *Solid State Physics - Advances in Research and Applications* **3**, 307–435 (1956).
- 4 26. Norby, T. A Kröger-Vink Compatible Notation for Defects in Inherently Defective Sublattices. *J.*
5 *Kor. Ceram. Soc.* **47**, 19–25 (2010).
- 6 27. De Souza, R. A. & Mueller, D. N. Electrochemical methods for determining ionic charge in
7 solids. *Nat. Mater.* **20**, 443–446 (2021).
- 8 28. Pohl, R. W. Electron conductivity and photochemical processes in alkali-halide crystals. *Proc.*
9 *Phys. Soc.* **49**, 3–31 (1937).

10

11

12 **Acknowledgements**

13 R.A.D. acknowledges discussions with K. Rushchanskii. R.A.D. gratefully acknowledges support from
14 the DFG (German Research Foundation) within the framework of the collaborative research centre
15 “Nanoswitches” (SFB 917) and within the Priority Programme “FieldsMatter” (SPP1959) under project
16 DE 2854/9-2. G.F.H. gratefully acknowledges funding from the European Union’s Horizon 2020
17 research and innovation programme under the Marie Skłodowska-Curie grant agreement No.
18 [101031819 - OPTICS].

19

20