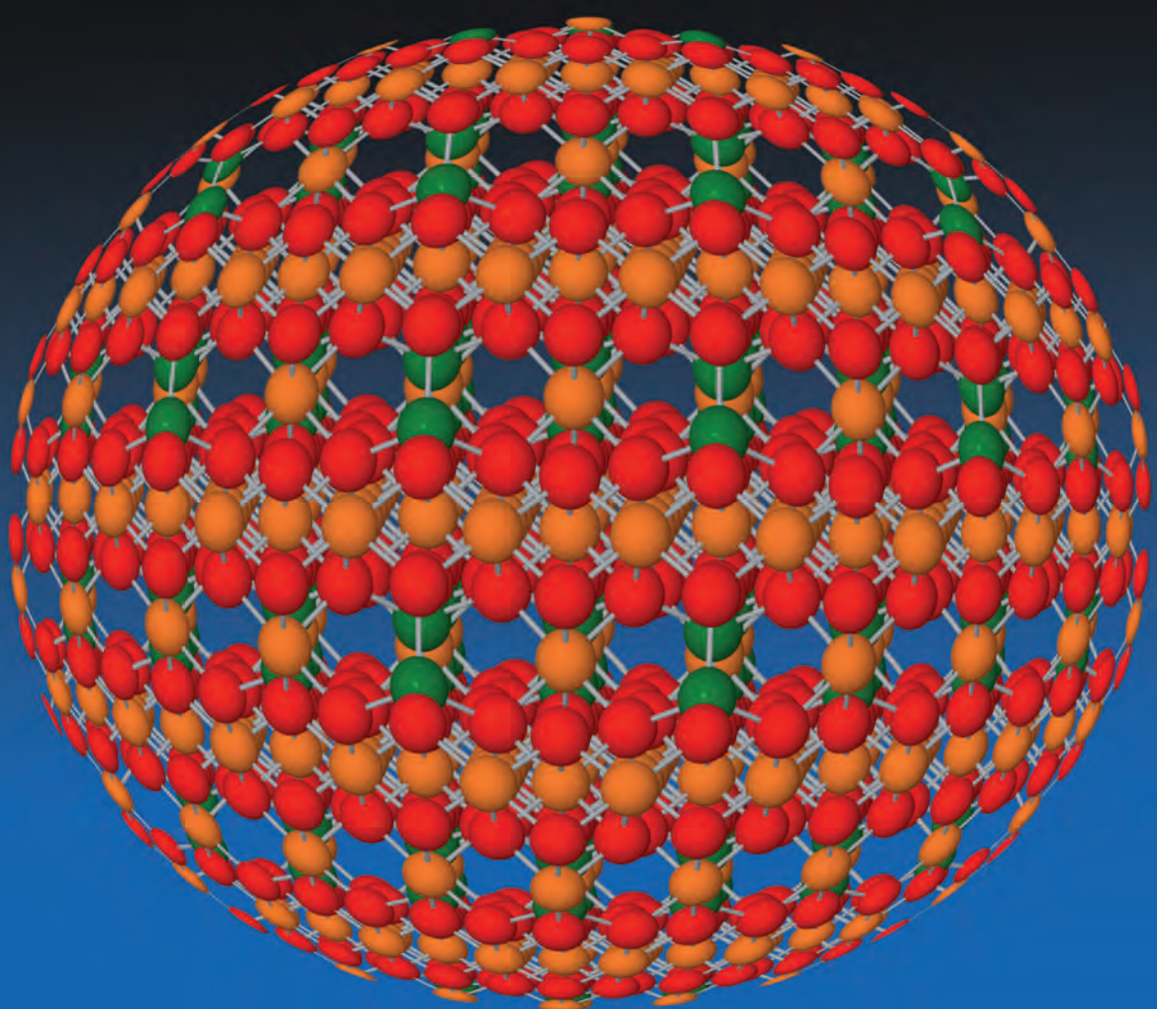


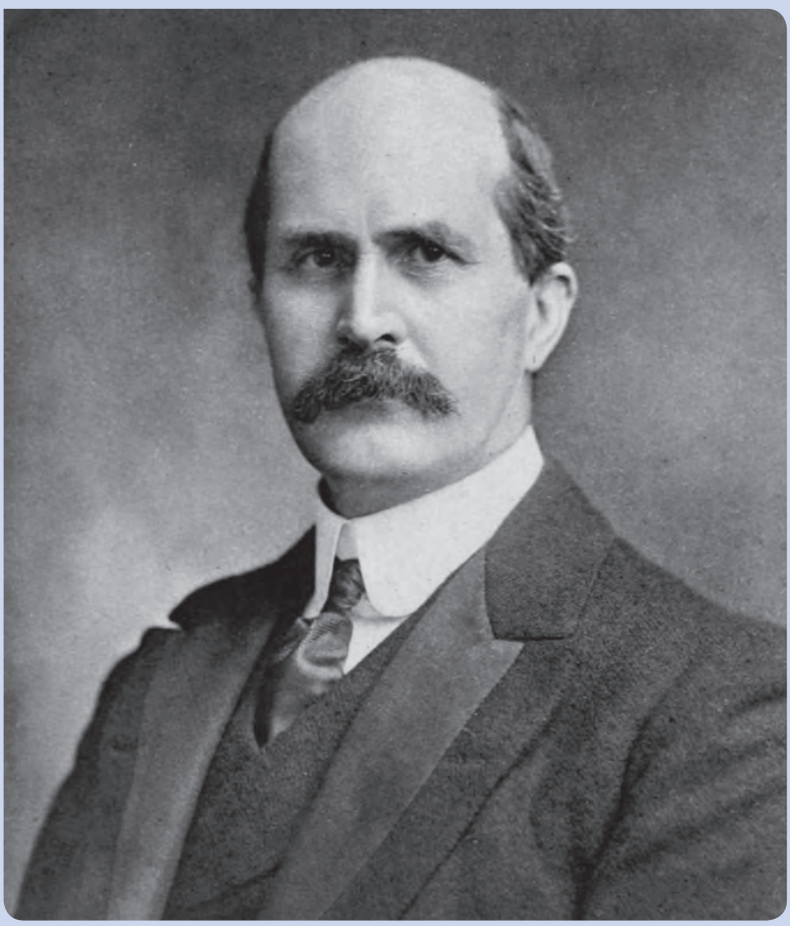
A Focus on Crystallography

Modern crystallography – creation and development of crystal structure databases using the Inorganic Crystal Structure Database (ICSD) as an example

Advancing Science

FIZ Karlsruhe

Leibniz Institute for Information Infrastructure



The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.

William Henry Bragg

Title Photo: Spinel ©FIZ Karlsruhe

Table of contents

Preface	4
Inorganic Crystal Structure Database (ICSD)	6
Extended symbols of coordination polyhedra in ICSD and their pictorial representation	11
Deposition of crystal structures at FIZ Karlsruhe	18
Analysis of spinel compounds in ICSD	20
Bond lengths statistics as a tool for crystal structure analysis – a case study with a crystalline silicophosphate	23
Using ICSD to predict new properties, compounds and modifications	27
Crystals and mathematics	37
Three centuries of crystallography through the lens of zbMATH	44
Facts and figures	51

Intas Exhibit 1021
Page 3 of 52
IPR2025-00799

Preface

Sabine Brünger-Weilandt
President & CEO of FIZ Karlsruhe

About one hundred years ago, Max von Laue discovered the diffraction of X-rays by crystals. This won him the Nobel Prize in Physics in 1914. At almost the same time William Henry Bragg and his son, William Lawrence Bragg, recognized that X-rays could be used to identify crystal structures and were awarded the Nobel Prize in Physics in 1915.

Today, 100 years later, these ground-breaking discoveries are honored and remembered on the occasion of the International Year of Crystallography. Over the years, crystallography has steadily evolved into one of the most basic and most important sciences of our times, which is reflected, among other things, in more than twenty Nobel Prizes based on a crystallographic background.

The crystal structure is essential for all solids. It determines the difference in hardness and appearance between graphite and diamond. It serves to detect the effect of drugs; important Nobel Prize-winning findings were the identification of the structure of penicillin, insulin, and, last but not least, of the DNA. Today, the Curiosity Rover measures the composition of crystalline soil samples on Mars. Crystallography impacts the development of everyday items such as innovative materials for the automotive or the aviation industry. It can also help to cope with the challenges of the German "Energiewende" (i.e. transition from fossile to alternative or renewable energy resources); Crystallography serves, for example, to develop new insulating materials or to discover suitable material structures for a new generation of cost-saving and efficient solar cells.

Material properties depend on the crystal structure. Therefore it is important, when developing new materials or products, to know how the atoms are aligned in the compound and which properties result from this. Crystallography helps answer questions on how certain compounds can be modified in order to obtain specific properties.

FIZ Karlsruhe has been operating and developing the Inorganic Crystal Structure Database (ICSD) for almost 30 years, its beginnings dating back to 1978. The oldest structures contained in ICSD are those of common salt and diamond, published by

William Henry Bragg and his son William Lawrence Bragg in 1913. The structures described and recorded today are much more complex, but the measurements are still based on the same assumption of the periodicity of crystal lattices and on the same measuring procedures. However, the mathematical methods applied are more sophisticated, and modern technology is used. In particular, the use of computers has significantly shortened the time required to identify a structure from several weeks to one or two days.

FIZ Karlsruhe

An interesting side phenomenon are the quasicrystals. They had already been known as a mathematical model in earlier years, but their existence could only be proven by modern technology. The discovery of the quasicrystal structures was awarded with the Nobel Prize in 2011. Their most distinctive property compared to all other known solids is their non-periodic crystal lattice. This makes it difficult to record such compounds in traditional structure databases.

Today, ICSD is the most comprehensive database for inorganic crystal structures worldwide. It serves not only as a reference work; the complex search algorithms also allow for comparative structure determination, for analysis, or for using the data for the development of new materials.

This brochure describes the ICSD's evolution and shows its potential applications. It also demonstrates the close relationship between crystallography and mathematics, and it answers questions, e.g., on how the bibliographic database zbMATH (formerly "Zentralblatt für Mathematik") can be used to gain an overview on the development of crystallography-related mathematical methods.

Crystallography will maintain its high importance in the future. True to its motto "Advancing Science", FIZ Karlsruhe as the producer of ICSD will live up to the new challenges of modern crystallography and will continue to provide inorganic chemists and materials scientists with a highly sophisticated tool enabling them to generate scientific value and innovation.

Ongoing popularity of factual databases even in times of large search engines

Inorganic Crystal Structure Database (ICSD)

Stephan Rühl

Introduction

In spite of manifold possibilities to gain expert information today, the use of factual databases is still essential. Because of highly specialized requests nowadays many tailor-made products are available [1].

Crystal structures in general play an important role in understanding physical properties of materials. A lot of special databases cover particular interests like zeolite structures [2] or mineral structures [3], but for general purposes overall collections of data are required. Examples to mention here are ICSD [4], CSD [5] or PDB [6]. The huge amount of valuable information stored in crystal structure databases helps researchers in many ways, for example to provide input for a Rietveld refinement [7] or data-mining parameters for structure prediction [8] or optimization procedures. But crystal structure databases contain a lot more valuable information than the obvious unit cells and atomic coordinates. For example, ICSD can be used to find similar structures by comparing certain features, like the space group or the ANX formula, that define different structure types [9, 10]. For many more applications of ICSD see article "Using ICSD to predict new properties, compounds and modifications" by Schön on page 27.

For a database to be of any use to researchers, it has to cover several important aspects. The first essential aspect is that data can be easily compared and hence the data have to be supplied in some standardized form. For crystallographic data this kind of standardization is partly inherited from the principles of crystallography itself and further enforced by the application of standardization tools to the published crystal structure. Even for the exchange of crystallographic information a generally accepted format is defined (Crystallographic Information File – CIF) [11]. Another important point is the completeness of the information provided. A statistical interpretation based only on a small subset will likely produce unreliable results. Also it is at least highly inconvenient having to check several databases and combining results. Unfortunately, this step might become necessary if the scope of a database only partly includes the field of interest. The third and most decisive factor for a database is the quality of the data. Therefore, carefully checking and evaluating new information is fundamental.

History of ICSD

The original idea of ICSD goes back to an initiative of Prof. Günter Bergerhoff in 1978 at the Institute for Inorganic Chemistry of the University of Bonn, Germany [4]. FIZ Karlsruhe started to maintain the database in collaboration with the University of Bonn in 1985. In 1989 a joint venture between the Gmelin Institute, Germany, and FIZ Karlsruhe, both non-profit institutions, took over the responsibility for ICSD. Since 1997 ICSD is produced cooperatively by FIZ Karls-ruhe, Germany, and the National Institute of Standards and Technology (NIST), U.S.A. The production, quality control and general software development is done at FIZ Karlsruhe, while the still widely used PC-Windows based graphical interface FindIt is generated under the responsibility of NIST. A first WWW interface was developed by Alan Hewat at the Institute Laue-Langevin, Grenoble, but was replaced in 2009 by a new interface developed by FIZ Karlsruhe. Since many years FIZ Karlsruhe has been working together with Prof. Rudolf Allmann concerning the quality assurance of the database entries and the development of some enhancements of ICSD, e.g., the Structure Types of the database. Prof. Allmann was awarded with the Will-Kleber medal of the German Association for Crystallography for his extraordinary achievements for ICSD. FIZ Karlsruhe is grateful for the tremendous commitment Prof. Allmann has shown towards ICSD during these years.

Prof. Rudolf Allmann

At about the time that the database was initiated the Crystal Structure Deposition at FIZ Karlsruhe began. It started as an archive of crystal structure

records because the important crystal structure information was missing in many relevant publications. Since then, many publishers provide footnote references to the crystal structure depot of FIZ Karlsruhe, where additional structure information can be obtained on request using a Crystal Structure Deposition number. Since February 1999, there has been an agreement between the Cambridge Crystallographic Data Centre (CCDC) and FIZ Karlsruhe that all organic and organometallic compounds should be deposited at CCDC and all inorganic and intermetallic compounds at FIZ Karlsruhe. The deposited data are also used as input for the respective databases.

Content of ICSD

The Inorganic Crystal Structure Database (ICSD) contains information on structures

- that have no C-C and no C-H bonds,
- whose atomic coordinates have been fully determined or were derived from a corresponding structure type.

Since 2003, crystal structures of metallic and inter-metallic compounds have been included into ICSD. The number of records in ICSD has roughly doubled in the last 12 years (Fig. 1). Apart from currently more than 6,000 new entries per year FIZ Karlsruhe is working continually on filling in the gaps in older data (Fig. 2). In 2001, ICSD changed to a relational database system with at that time 25 tables and about 200 database fields. Continual extension of the content results in more than 35 tables and more than 300 database fields. This allows for queries with much more criteria. Innovations like the inclusion of structure types and calculation of standardized data provided new search options.

ICSD records provide the full structural information published, the bibliographic references, and additional value-added content. Especially the structure types, which were introduced in 2005 [12] and were subsequently expanded during the last years, represent valuable information in comparing inorganic crystal structures.

The two defining properties for several crystal structures to be considered belonging to one structure type are that the structures are isopointal and isoconfigurational. These rather unhandy properties are broken down to some easily checkable properties like the ANX formula, Pearson symbol, Wyckoff sequence, etc. A complete description of this procedure would be beyond the scope of this paper and can be found in the article by Allmann and Hinek [12].

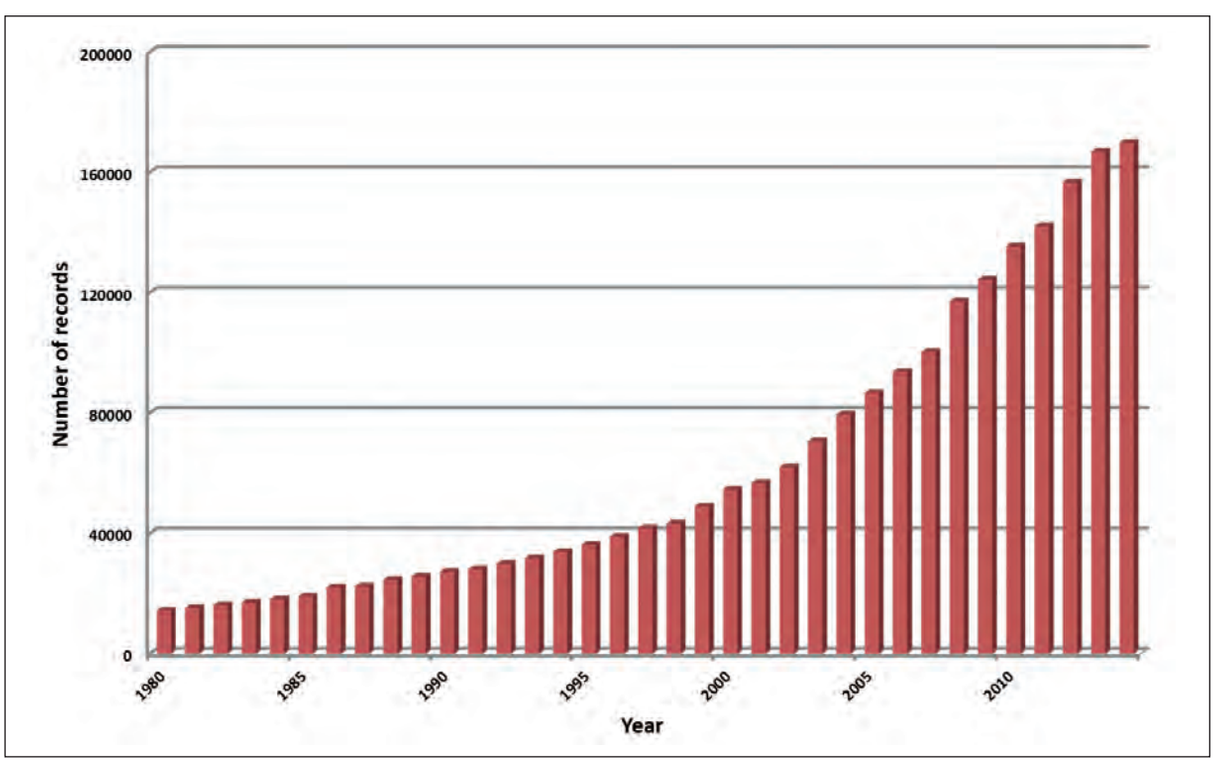


Fig. 1: ICSD growth since 1980

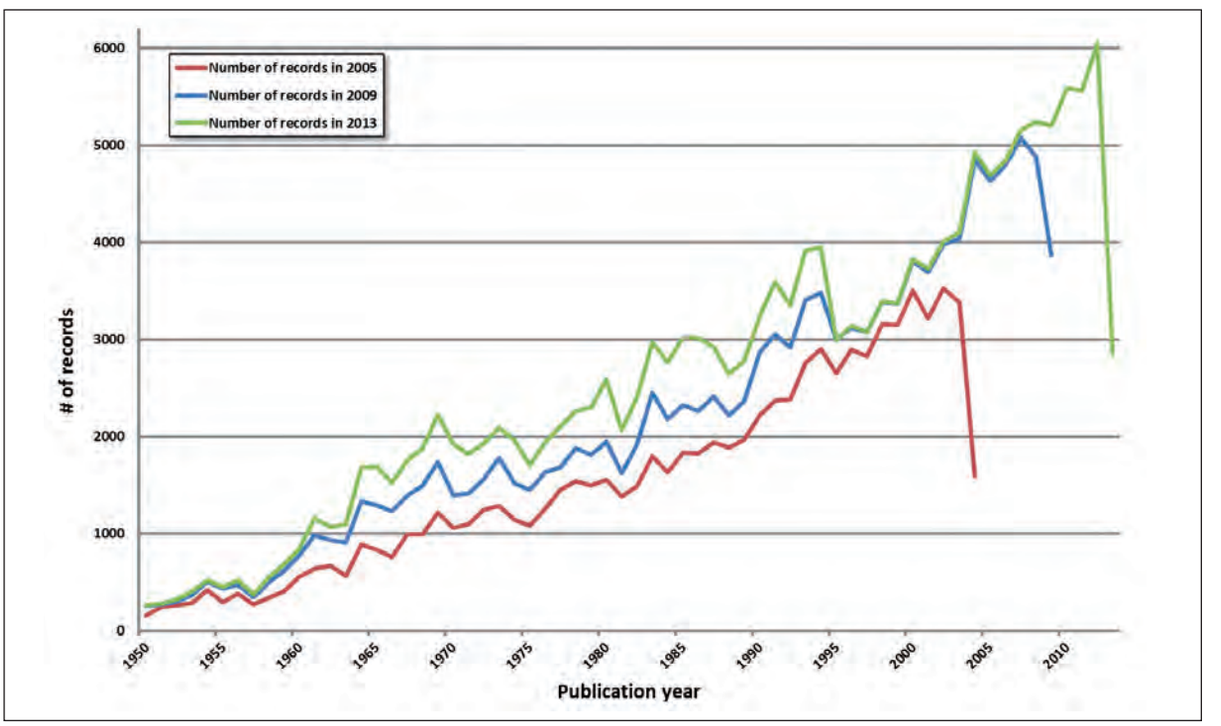


Fig. 2: Continually filling the gaps in older data – distribution of ICSD records by publication year

ICSD now contains a comprehensive and searchable list of 8,230 structure types according to which approximately 79,8% of the crystal structures have been assigned. For each structure type one representative is used as a prototype. For all prototypes the atomic environments of all atomic positions (except for Hydrogen) are given, i.e. the kind and number of neighboring atoms (coordination number CN). In questionable cases the neighbors are calculated by constructing the Dirichlet-domains for all atomic positions (more details on this can be found in the article "Extended symbols of coordination polyhedra in ICSD and their pictorial representation" by Allmann et al. on page 11). Additionally, the shape of the coordination polyhedron is given for all atoms of the prototype using the nomenclature of Lima-de-Faria [13] and an extension by Allmann and Hinek [12].

Input and data evaluation

The data acquisition for the input into the database utilizes a wide range of procedures like screening of the original papers in crystallographic journals on relevancy, searching in the Chemical Abstracts File, and including the information from the Crystal Structure Deposition depot or direct contributions from authors and crystallographic experts. Each structure determination reported in the literature leads to a separate record in the database.

Even with evaluating all these sources of information it seems clear that comprehensiveness is a difficult task so that some structures are still missing and any feedback from the scientific community is highly appreciated.

During the production process the data are excerpted by scientists and checked automatically by computer programs and manually by our experts. This evaluation procedure includes, but is not limited to, checks for correct syntax (in case of CIF files), duplicate records, missing fields and, most importantly, plausibility. In addition, some details of the new entry can easily be verified like the space group being compatible with the cell or the assigned oxidation states fulfilling electroneutrality in combination with the multiplicity and site occupations for all atoms. Also very important are checks for interatomic distances which should be within the range of similar distances in the database. This is just a short overview of the evaluation process; more details on the checking routines are described by Behrens [14].

Lastly, additional content not directly given in the article is generated either automatically or manually. Among these data items are the Wyckoff sequence, Pearson symbol, molecular formula, molecular weight, calculated density, ANX or AN formula, lists and histograms of interatomic distances, the reduced cell, the standardized cell (if necessary), mineral names or mineral groups.

Interfaces

ICSD is offered as a stand-alone version for local PC installation (e.g. FindIt, produced in cooperation with the National Institute of Standards and Technology), a local intranet version for small user groups, and as a Web version hosted by FIZ Karlsruhe.

We have developed a locally installable version based on the ICSD Web interface (Fig. 3 and 4). This version will replace the FindIt interface which was originally developed for Windows XP.

Developing a new local version based on an existing interface offers many synergies and will speed up the inclusion of new features for both interfaces. In the near future the intranet version will also be based on this framework. The user can then switch from one interface to any other interface without having to learn how to use this specific interface.

All interfaces offer search options for the relevant crystallographic information as well as the added content, visualization of the crystal structure, simulation of the powder pattern, and export as CIF (Crystallographic Information File).

Apart from FindIt, which is limited to Windows operating systems, ICSD can be used on all major platforms. Although there is no special ICSD App one can access the ICSD Web version on mobile devices via the internet. As mobile operating systems do not support Java applets so far, the visualization of crystal structures is not yet possible.

Outlook

ICSD offers a wide spectrum of useful information for chemists, physicists, crystallographers, mineralogists or geologists, and it is widely used from teaching basic inorganic structural chemistry to research on new materials in academia and industry. As new knowledge is gained, the needs of the users change and so the content of databases in general is adjusted to reflect these increasing demands.

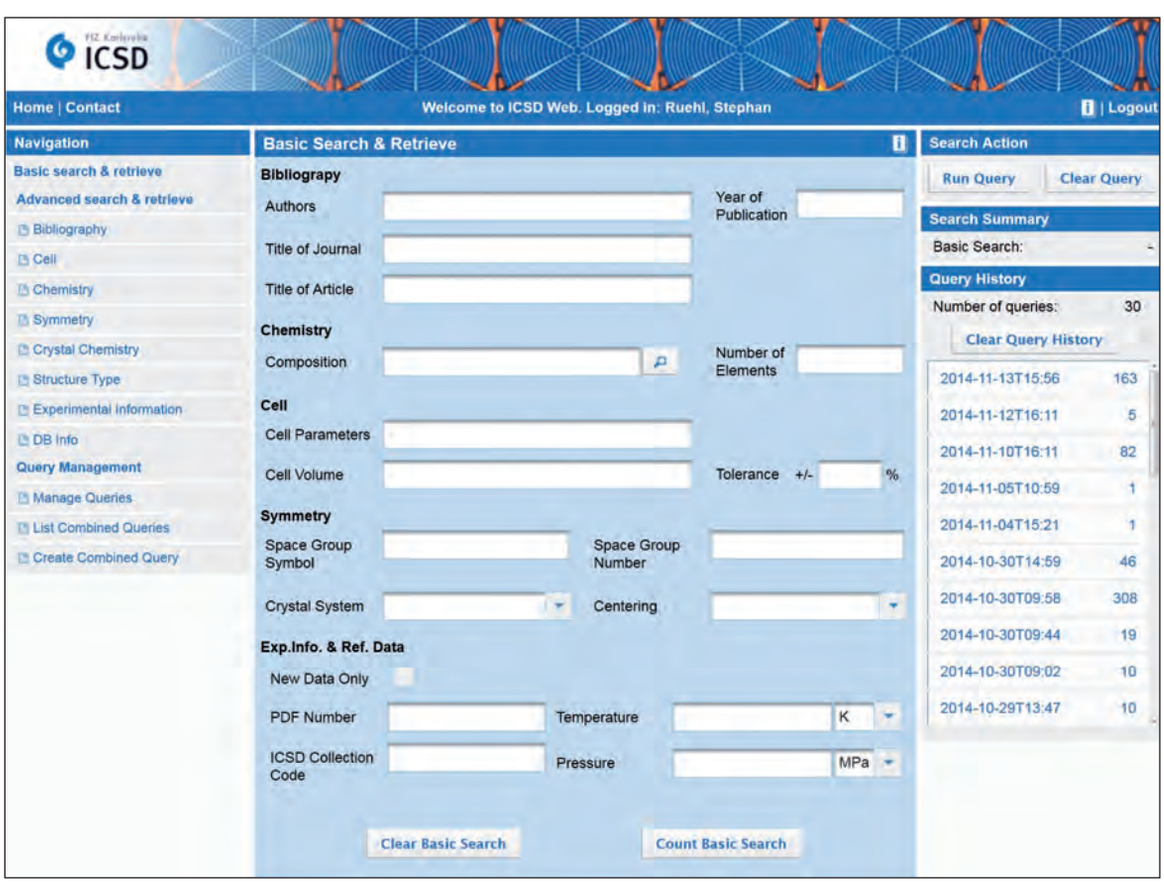


Fig. 3: Search screen of ICSD-Desktop/Web interface

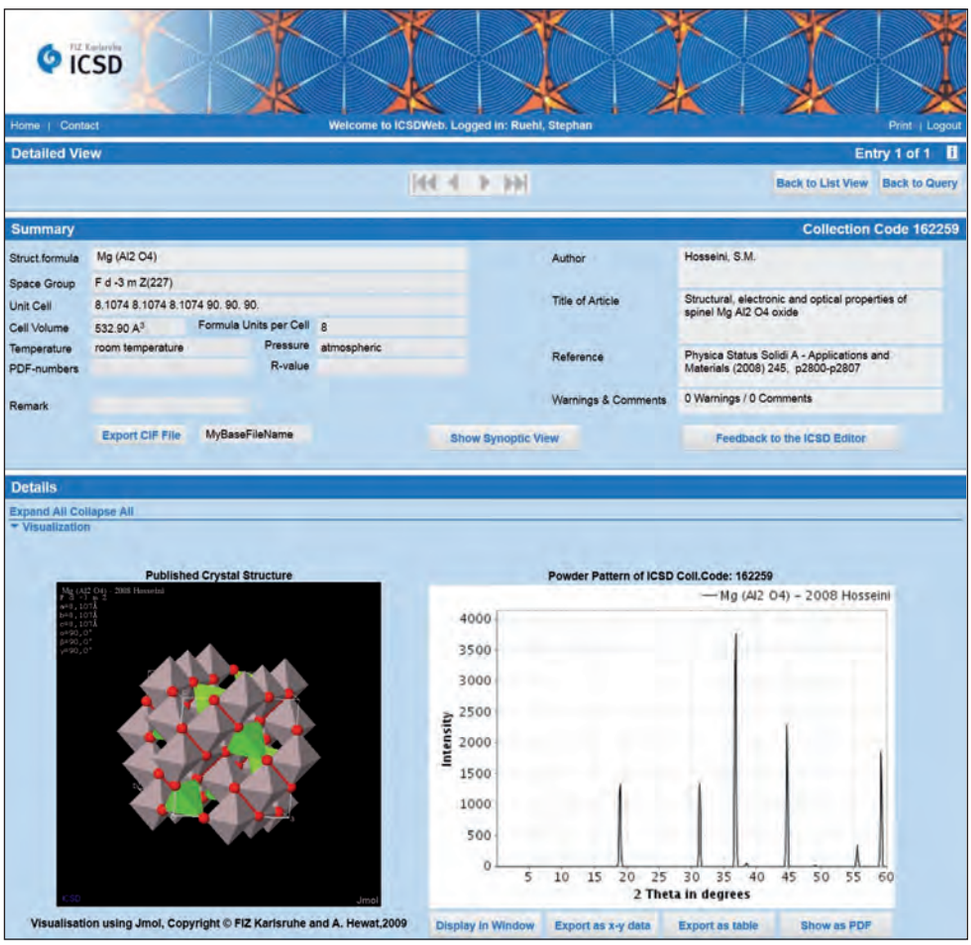


Fig. 4: Results page of ICSD Web interface

The scope of ICSD has changed several times over the last 35 years. Some smaller improvements are continuously being made, e.g., the inclusion of DOIs for easier access to the primary literature. Further extensions with additional information are planned.

Significant changes are currently made to the ICSD Web and ICSD-Desktop interfaces, for which some important new features and improvements of existing features are included. We have some more improvements – based on the wishes of our users – planned for the near future. Among these will be more export options for crystal structure data and more customization options on the display. In addition we are evaluating the option to directly include external programs in our interface to offer better analysis of the results obtained in the search. This could also include simple tools to compute information not directly contained in the database like bond valence sums.

References

[1] Böhme, U.; Tesch, S. (2014). Nachr. Chem. 62, p. 43.

[2] Baerlocher, C. ; McCusker, L.B. Database of Zeolite Structures: http://www.iza-structure.org/databases/ (November 25, 2014).

[3] Downs, R.T. ; Hall-Wallace, M. (2003). Am. Mineral. 88, p. 247.

[4] Bergerhoff, G.; Brown, I.D. (1987). Crystallographic Databases, International Union of Crystallography, Chester. pp. 77-95; Belsky, A.; Hellenbrandt, M.; Karen, V.L. ; Luksch, P. (2002). Acta Cryst. B58, p. 364.

[5] Allen, F. H. (2002). Acta Cryst. B58, p. 380.

[6] Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.; Meyer Jr., E. E.; Brice, M.D., Rodgers J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. (1977). J. Mol. Biol. 112, p.535.

[7] Rietveld, H.M.; (1967). Acta Cryst. 22, p. 151; Rietveld, H. M. (1969). J. Appl. Cryst. 2, p. 65.

[8] Hofmann, D.; Kuleshova, L. (2012). Data Mining in Crystallography, Springer Verlag, Heidelberg; Buchsbaum, C.; Höhler-Schlimm, S.; Rehme, S. (2010) Struct. Bond 134, p. 37.

[9] Kaduk, J.A. (2002). Acta Cryst. B58, p. 370.

[10] Abrahams, S.C. (1994). Acta. Cryst. A50, p. 658.

[11] Hall, S.R.; Allen, F.H.; Brown, I.D. (1991). Acta Cryst. A47, p. 655.

[12] Allmann, R.; Hinek, R. (2007). Acta Cryst. A63, p. 412.

[13] Lima-de-Faria, J.; Hellner, E.; Liebau, F.; Makovicky, E.; Parthe, E. (1990). Acta Cryst. A46, p.1.

[14] Behrens, H. (1996) J. Res. Natl. Inst. Stand. Technol. 101, p. 365 .

Classification of related structures in ICSD

Extended symbols of coordination polyhedra in ICSD and their pictorial representation

Rudolf Allmann, Roland Hinek, Stephan Rühl

In 2014 the Inorganic Crystal Structure Database (ICSD) contains 173,473 structures in total and about 79% are assigned to one of the 8,230 distinct structure types. The structure types were introduced in ICSD in 2007 as an approach to describe structural similarity between inorganic structures. ICSD structure types can best be classified as isoconfigurational, according to the definition by the IUCr [1]. In this definition isoconfigurational structures must conform to certain structural parameters (descriptors) like space group, Pearson symbol, Wyckoff sequence, c/a ratio, ANX formula. Of these descriptors the Wyckoff sequence is crucial, because it describes the atomic arrangement in the unit cell. Inorganic structures are often visualized by coordination polyhedra of single atoms or ions. Each structure type is characterized by some or all of the descriptors as has been explained in detail by Allmann and Hinek [2].

For each structure type a "prototype" is selected from all structures belonging to this structure type. The prototype and thus the name of the structure type is mostly based on the first publication of this structure. The prototypes play an important role in ICSD as the atomic environment (AE) of suitable non-H-atoms in this structure is listed as an additional comment. The AE represents the first coordination sphere and includes all first neighbors forming the coordination polyhedron. For example, in the spinel structure Al_2MgO_4 the oxygen atoms are arranged in a cubic close packed lattice while the Al and Mg occupy every second octahedral void and every eighth tetrahedral void, respectively. The comment for this entry in ICSD (collection code 31373) includes the AE:

AE: Al: 6o O6; Mg: 4t O4; O: 4t: Al3 Mg; *c.c.p. O (O: 12co O12)*

In this systematic description, the atomic environment for Al, Mg and O is specified. The aluminum is octahedrally (6o) coordinated by six oxygen ions (O6); the magnesium shows a tetrahedral (4t) coordination by four oxygen ions (O4) and the oxygen is also tetrahedrally (4t) coordinated by three aluminum ions and one magnesium ion (Al3 Mg). The text in *italics* describes the second coordination sphere around the oxygen, but usually only the first sphere is given in the prototype entries.

At a conference in Kiel (Germany) in 1964, Donnay, Hellner and Niggli introduced the denotation of short symbols for polyhedra, which at that time were formed exclusively by symmetry-equivalent atoms in each of the 230 space groups [3]. Parthé extended this nomenclature to the first neighbors (first coordination sphere) of an atom, which in addition may include non-symmetry-equivalent neighbors [4] (e.g. a tetragonal pyramid "5y" having four equivalent

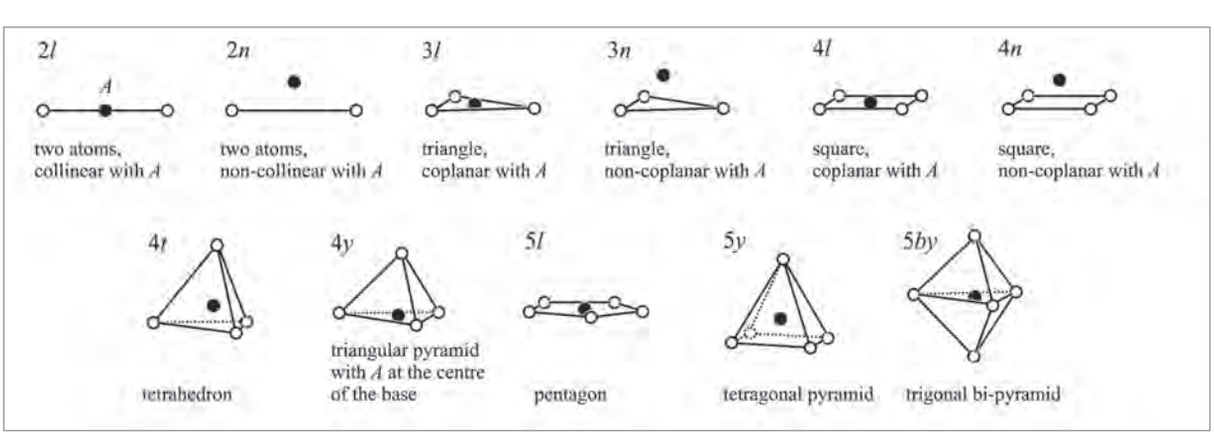


Fig. 1: Coordination polyhedra proposed by Parthé

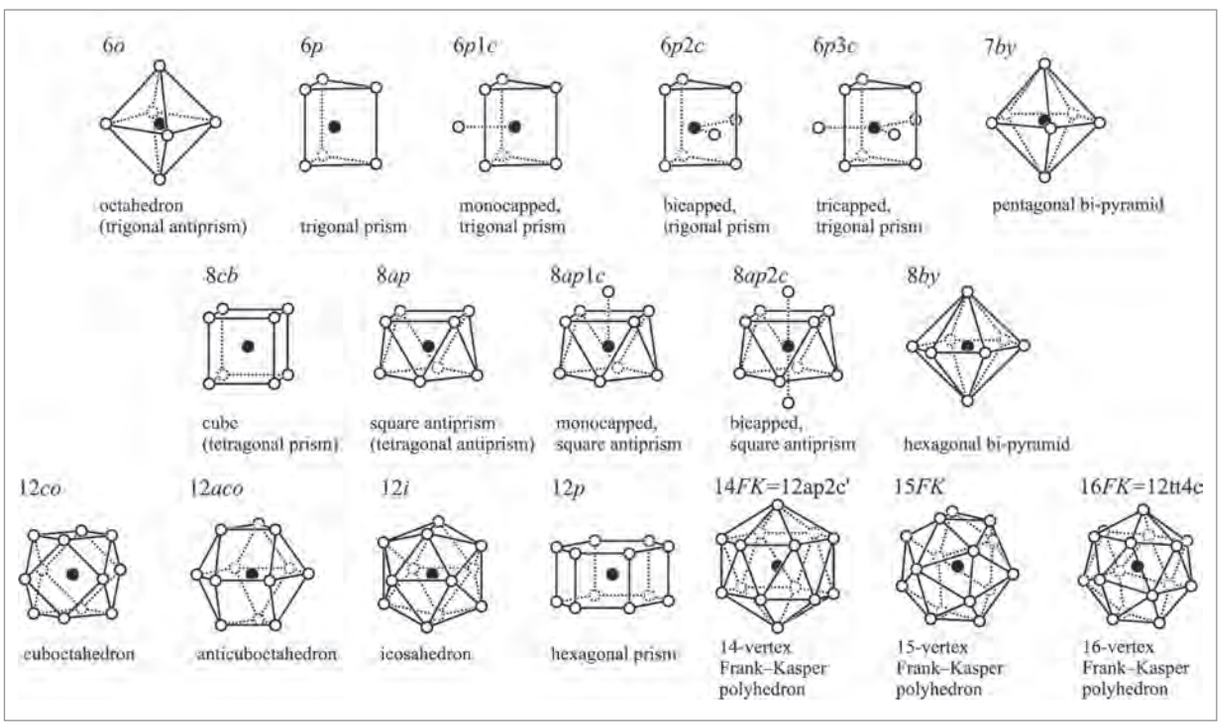


Fig. 1: Coordination polyhedra proposed by Parthé

atoms at the base and one non-equivalent atom at the top of the pyramid). In this book Parthé showed 29 frequently observed coordination polyhedra with their short symbols (Fig. 1). This denotation with short symbols was accepted by the nomenclature commission of the IUCr in 1990 and 1998 [1].

In 2010, Allmann and Hinek introduced a more precise definition of the symbol "c" describing the presence of "capping" atoms. The capping atoms in prisms and antiprisms can be differentiated by being centered on equatorially (c) or axially located faces (c') [5].

The following letters are used to distinguish between the coordination polyhedra:

l	coplanar or collinear
n	non-coplanar or non-collinear
p	prism
ap	antiprism
y	pyramid
by	bipyramid
t	tetrahedron
o	octahedron
cb	cube
i	icosahedron
tt	truncated tetrahedron
pd	pentagon-dodecahedron
co	cuboctahedron
aco	anticuboctahedron
rd	rhombic dodecahedron
bds	bisdisphenoid
tds	trisdisphenoid
to	truncated octahedron
tcb	truncated cube
snubcb	snub cube
FK	Frank-Kasper polyhedra
c	capped faces for non-prisms/non-antiprisms
	equatorially capped faces for prisms/antiprisms
c'	axially capped faces for prisms/antiprisms
(outside):	The central atom is outside of the polyhedron.

The following list shows for each coordination number the symbols observed in ICSD, and figures 2a and 2b show the 40 most observed polyhedra in an idealized geometry:

2:	2l, 2n
3:	3l, 3n
4:	4l, 4n, 4t, 4y (central atom at the center of the base)
5:	5l, 5n, 5y, 5by
6:	6l, 6n, 6o (deformed: 6by, 6ap), 6p, 6y
7:	7y, 7by, 6o1c (= 6ap1c), 6p1c; 6p1c'
8:	8cb, 8ap, 8bds, 8by, 6p2c (similar to 8ap), 6p2c', 4l4t
9:	6p3c (similar to 8ap1c'), 8p1c', 6n3n
10:	10p, 10ap, 8p2c', 8ap2c', 3n6n1c' (=7y3c')
11:	6p3c2c', 10p1c', 10ap1c'
12:	12co, 12aco, 12i, 12p, 10p2c', 10ap2c', 12tt, 12tds, 6p6c
13:	12p1c', 12ap1c', 13FK(= 10ap3c')
14:	14FK (= 12ap2c'), 12p2c', 14rd
15:	15FK(=12ap3c'), 10p5c
16:	16FK (=12tt4c)
17:	10p5c2c'
18:	12p6c
20:	12p6c2c', 20pd
24:	24to, 24tcb, 24snubcb

The coordination numbers in ionic structures usually do not exceed 12. Very frequent coordination polyhedra are 4t (tetrahedron) and 6o (octahedron). Cations are surrounded by anions and vice versa. The anion-anion contacts are neglected despite the fact that they play an important role in space filling and the unit cell dimension. In spinels, for example, each oxygen has 12 O-O contacts as second neighbors resulting in the coordination polyhedron "12co". Assuming an O-O distance of 2.8 Å this corresponds to a cubic lattice constant of about 7.92 Å (=2.8*sqrt(2)*2), which is in good agreement with the cell parameter a=8.08 Å for the ICSD entry with collection code 31373.

In structures of metals frequent coordination numbers are 12 to 16 (12i-16FK) due to the closer packing in comparison to ionic structures. A metal structure consists of cations embedded in an electron gas. The existence of energy bands of the electronic states in metal structures is the reason for characteristic properties like thermal and electrical conductivity. The variety of ion-radii of atoms in metallic structures is smaller than for ionic ones and the anions are generally greater than the cations.

Usually the first and second coordination spheres of a given atom are well separated by a significant "distance gap" (Fig. 3).

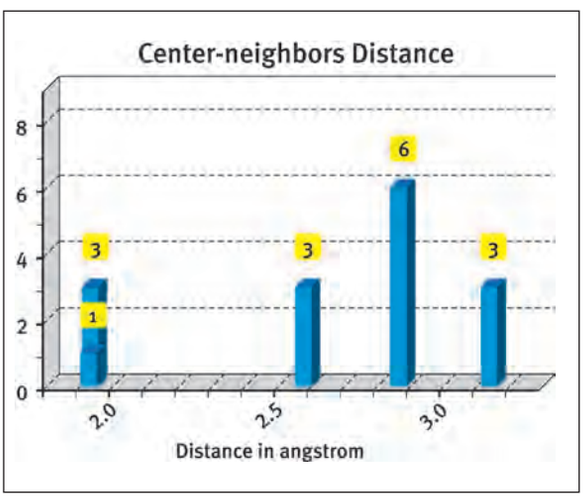


Fig. 3: ICSD entry with collection code 31373 (spinel type): distance histogram around O with a gap between first (3Al +1 Mg) and second coordination sphere (c.c.p. O (O: 12co O12))

The best way to determine the first neighbors is by means of Dirichlet domains [6]. The Dirichlet domain of an atom is an artificial construction of the part of space which is nearest to the central atom. It depends on the specific atomic environment, i.e., of its nearest neighbors. The Dirichlet domain itself is a polyhedron. It is enclosed by planes located midway between the central atom and each nearest neighbor and perpendicular to the distance between them. Figures 4a and 4b illustrate the creation of the Dirichlet domain for the cuboctahedron as the second coordination sphere around an oxygen in the spinel as discussed earlier. If only the distances from the central oxygen to the 12 nearest oxygens forming the cuboctahedron in figure 4a are drawn, the resulting image consists of only the red lines in figure 4b. Now planes are constructed that are each perpendicular to one of the red lines and include the center of this line. The central atom is then enclosed by the Dirichlet domain. This method is used in the program TOPOS [7].

Another method to generate the Dirichlet domain does not place the planes midway between the central atom and the coordinated atom but takes into account the possibly different sizes of the atoms. This concept of

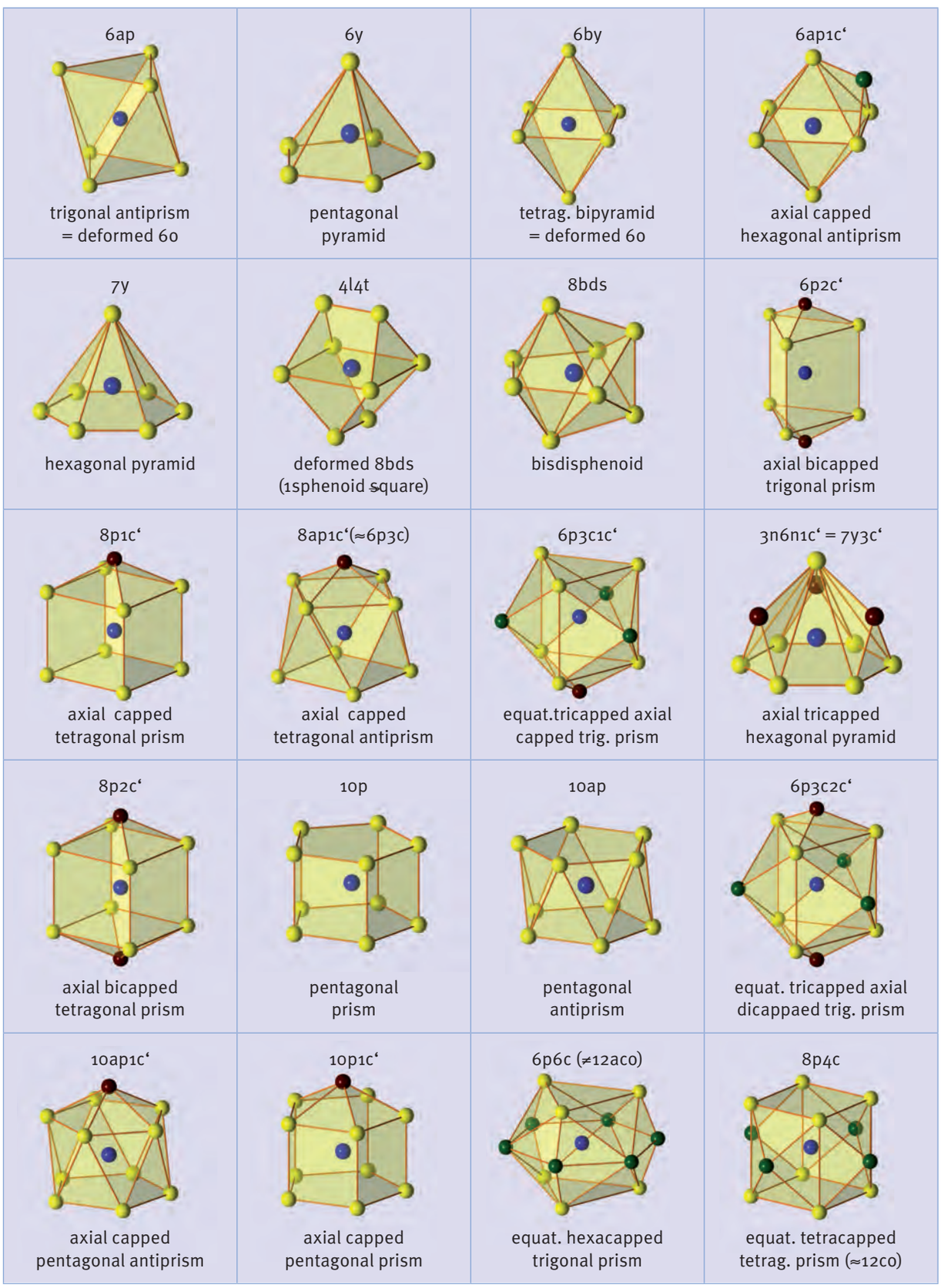


Fig. 2a: Additional polyhedra observed in ICSD

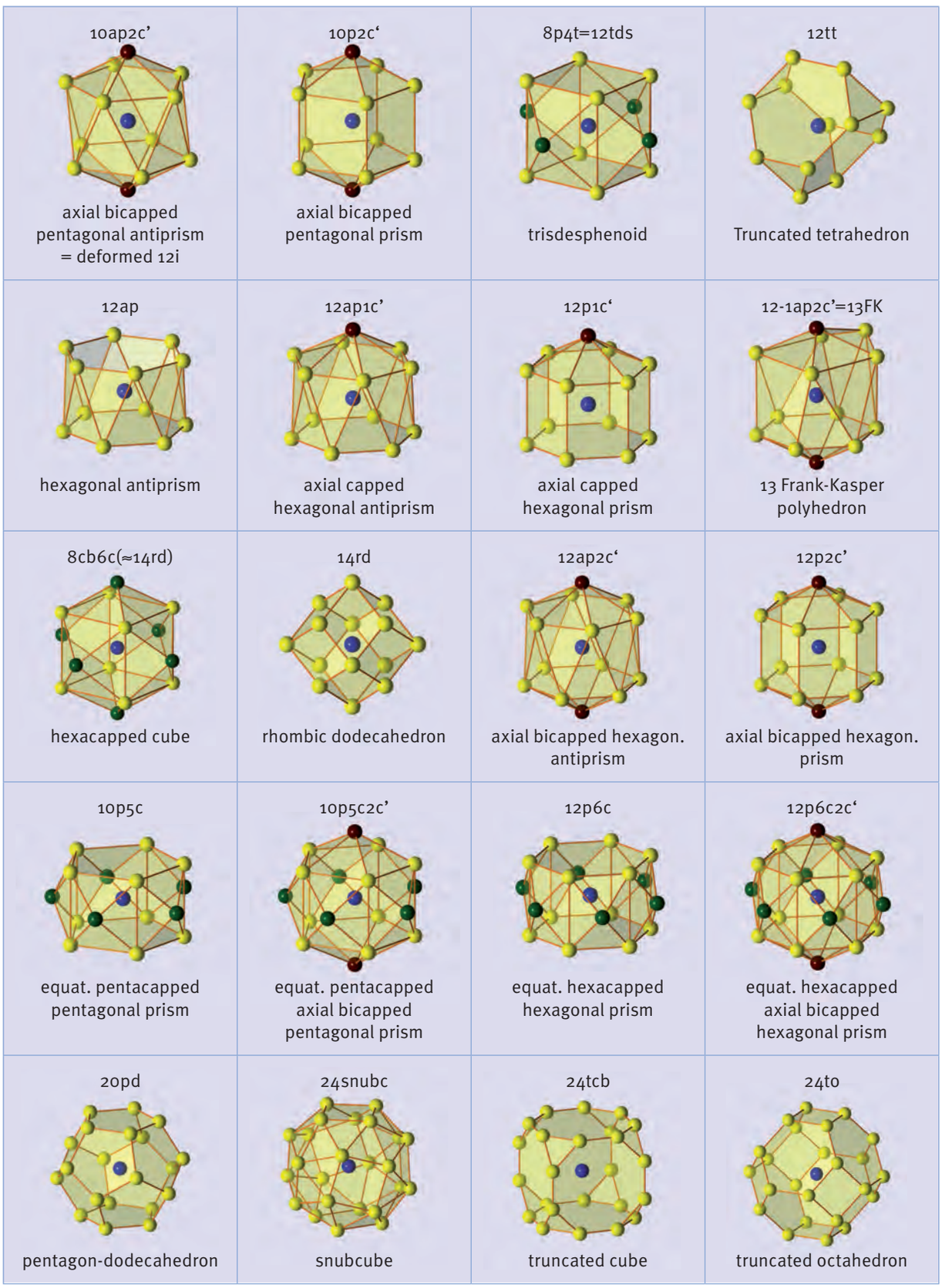


Fig. 2b: Additional polyhedra observed in ICSD

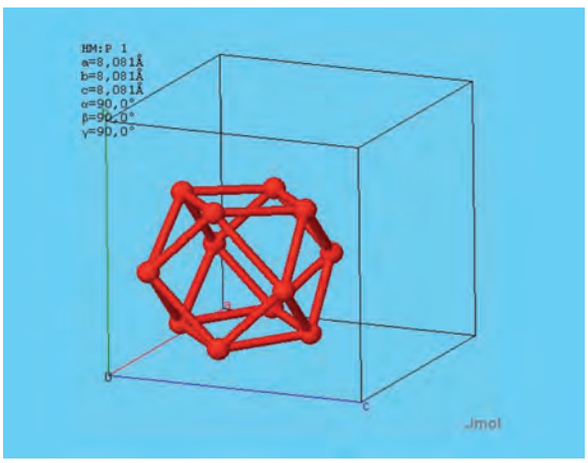


Fig. 4a: spinel-2nd AE: O O12

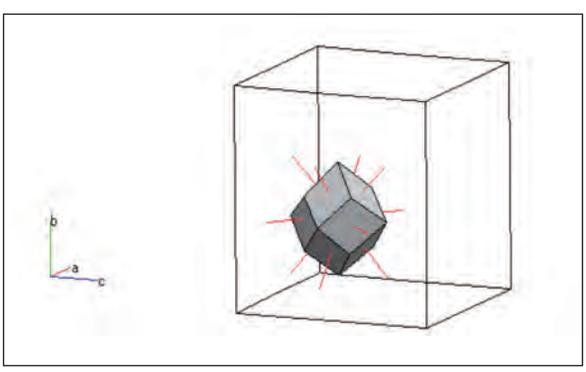


Fig. 4b: spinel-2nd Dirichlet domain: O O12 (TOPOS [7])

radical planes is used in the program DIDO [8]. The radical planes generally are located closer to the smaller atoms. This program is mostly used to calculate the atomic environments in ICSD.

Synonyms of the term Dirichlet domain are Voronoi polyhedron or Wigner-Seitz cell. A Brillouin zone is a Dirichlet domain constructed in reciprocal space instead of real space.

Dirichlet domains and coordination polyhedra (AE representation) are often "dual" to each other, i.e., they depend on each other in the way that the number of vertices and faces are interchanged. Examples of dual pairs of polyhedra are 4t-4t, 6o-8cb, 12i-20pd, 12co-14rd.

AE representation ← "DUAL" → Dirichlet domain (Dc)

Cuboctahedra "12co" ← "DUAL" → "14rd = 8cb6c"

The "strength" of a bond can be measured by the spatial angle of the corresponding face of the Dirichlet domain projected onto a sphere. Spatial angles are usually given as a percentage of the surface of the sphere. The spatial angles are large for short distances between two atoms and the corresponding face of the Dirichlet domain is then also large. The area of the face of the Dirichlet domain can be used instead of the spatial angle as an estimate for the strength of the bond.

For example, in a body-centered cubic structure each atom has 8+6 neighbors forming a rhombic dodecahedron (14rd) coordination polyhedron. The spatial angles for each of the 8 nearest neighbors is 9.84% and for the 6 farer away neighbors 3.54% corresponding to a truncated octahedral Dirichlet domain (which is not dual to the rhombic dodecahedral coordination polyhedron). The spatial angles of all 14 neighbors sum up to 100%.

Although there is a strong geometrical correspondence between the coordination polyhedron and the Dirichlet domain for an atom, there is also a remarkable difference between both constructions: The Dirichlet domains of all atoms fill the space completely like, e.g., unit cells, while the coordination polyhedra do not.

In ICSD, the coordination polyhedra are specified in the comments and are only given for the structure type prototypes. A coordination found in a prototype will also appear in all members belonging to this structure type.

Here are some examples of how to search for coordination polyhedra using the different retrieval interfaces:

Cuboctahedron	(12co)
Icosahedron	(12i)
Cube	(8cb)

In FindIt comments can be searched on the Reference tab using the comments search field available when the radio button "Free text" is selected (Fig. 5). In FindIt automatic truncation is applied to the search terms, so the number of hits tends to be higher (see examples in table 1). In ICSD Web or ICSD-Desktop this search field is accessible in the Advanced Search on the Experimental Information mask (Fig. 6). No automatic truncation takes place. This means that the given terms are searched exactly, if no wildcards have been explicitly set by the user.

Polyhedron	FindIt	ICSD Web/ ICSD-Desktop	Hits
12co	AE*12co	AE* *12co*	572
12i	AE*12i	AE* *12i*	361
8cb	AE*8cb	AE* *8cb*	489

Table 1: Examples of how to search for polyhedra in FindIt and ICSD Web/ICSD-Desktop

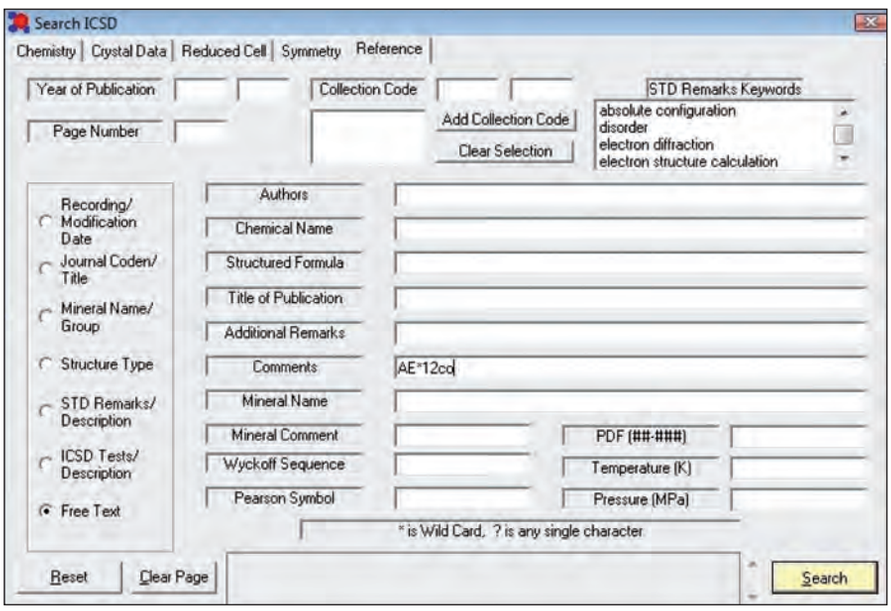


Fig. 5: Screenshot of a search for a polyhedron in FindIt

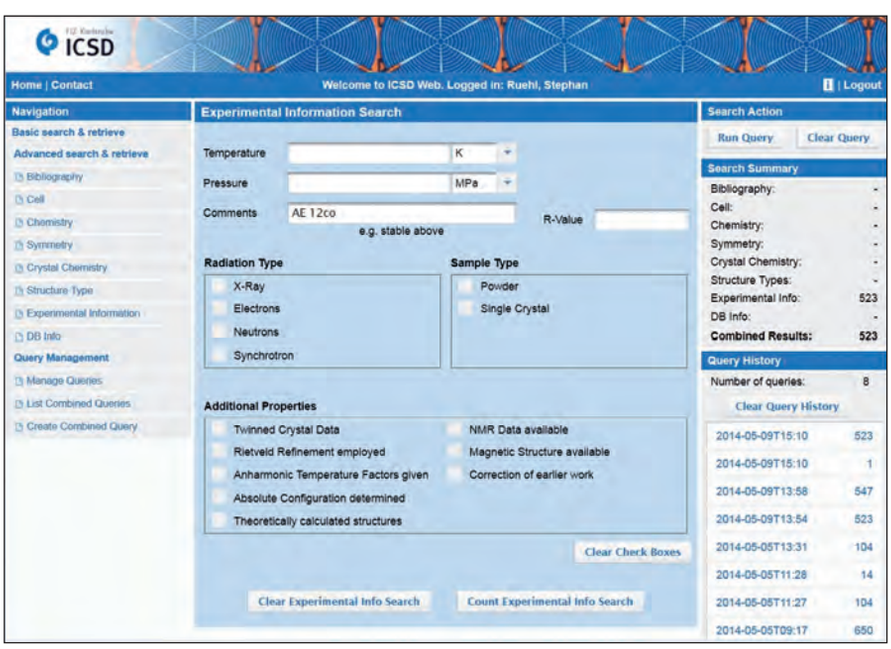


Fig. 6: Screenshot of a search for a polyhedron in ICSD Web/ICSD-Desktop

References

[1] Lima-de-Faria, J.; Hellner, E.; Liebau,F.; Makovicky, E.; Parthè, E., (1990, 1998). Report of the International Union of Crystallography Commission on Crystallographic Nomenclature Subcommittee on the Nomenclature of Inorganic Structure Types.

[2] Allmann, R.; Hinek, R., (2007). The introduction of structure types into the Inorganic Crystal Structure Database ICSD, Acta Cryst., A63, pp. 412-417.

[3] Donnay, J.D.H.; Hellner, E.; Niggli, A., (1964). Coordination polyhedra, Z. Kristallogr. 120, pp. 364-374.

[4] Parthé, E. (1964). Elements of inorganic structural chemistry, 2nd edition, K. Sutter Parthé Publisher, Petit-Lancy, Switzerland; Lima-de-Faria, J. (1994). Structural mineralogy, Kluwer Academic Publishers, Dordrecht, Netherlands.

[5] Allmann, R.; Hinek, R.,(2010). Extended symbols for coordination polyhedra after Parthé, Acta Cryst., A66, p. 174.

[6] Shannon, R.D. ; Prewitt, C.T. (1969.) Effective ionic radii in oxides and fluorides, Acta Cryst. B25, pp. 925-946.

Shannon, R.D. (1976). Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, Acta Cryst., A22, pp. 751-767.

[7] Blatov, V.A. (2004). Voronoi-Dirichlet polyhedra in crystal chemistry: theory and applications, Cryst. Rev. 10, pp. 249-318.

[8] Koch, E.; Fischer, W. (1996). DIDO95 and VOID95 - programs for the calculation of Dirichlet domains and coordination polyhedra, Z. Kristallogr. 211, pp. 251-253.

The central address for crystallographers in inorganic chemistry

Deposition of crystal structures at FIZ Karlsruhe

Helmut Müller

For now 30 years FIZ Karlsruhe has been offering its service to deposit crystal structures for free. At the beginning, and this means the period when digital networks were new and storage capacity was expensive, deposition at FIZ Karlsruhe was a valuable aid for researchers worldwide, providing access to (mostly) inorganic crystal structure data at a central point so that they could be used by all interested scientists.

Nowadays it has become more important than ever to have FIZ Karlsruhe as an independent manager of research data, so that everyone can request structure information, and regarding the increasing number of incoming structures it seems that there will be a further demand for this free service.

Moreover, the storage of research data like structural information in crystallography is nowadays a real hot topic in many branches. Today it is clear that the availability of scientific data and the reproducibility of scientific findings are essential for scientific progress. Besides, deposition of structures means securely storing them at a neutral institution with reliable access for the next decades.

When the depot started, all structures were collected, but since 1999 there has been an agreement between the Cambridge Crystallographic Data Centre (CCDC) and FIZ Karlsruhe that all organic and organometallic compounds should be deposited at CCDC, and all inorganic and intermetallic compounds belong to FIZ Karlsruhe. So, crystal structures should not contain C-C or C-H bonds if submitted to FIZ Karlsruhe.

During the first years, data were transmitted and filed in paper form including partially handwritten entries (Fig. 1), and we were always glad to get clearly legible notes.

From this period a lot of shelves with dusty folders still exist at the FIZ basement archive (Fig. 2).

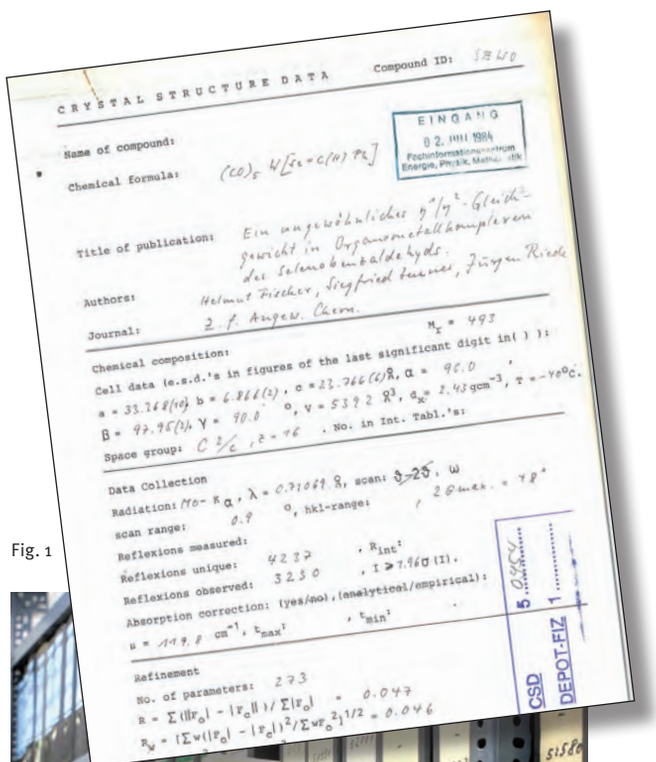


Fig. 1

Fig. 2

As a next step on the route towards electronic data, hard copies on fan-fold paper (Fig. 3) were sent to FIZ Karlsruhe. Now, in the last decade, the electronic submission of crystal structure data via CIF file is preferred, which significantly facilitates the daily work. Nevertheless there is still a demand for really old datasets. Since most of these structures are also stored in the ICSD database, FIZ Karlsruhe can in most cases send these data in electronic form, too.

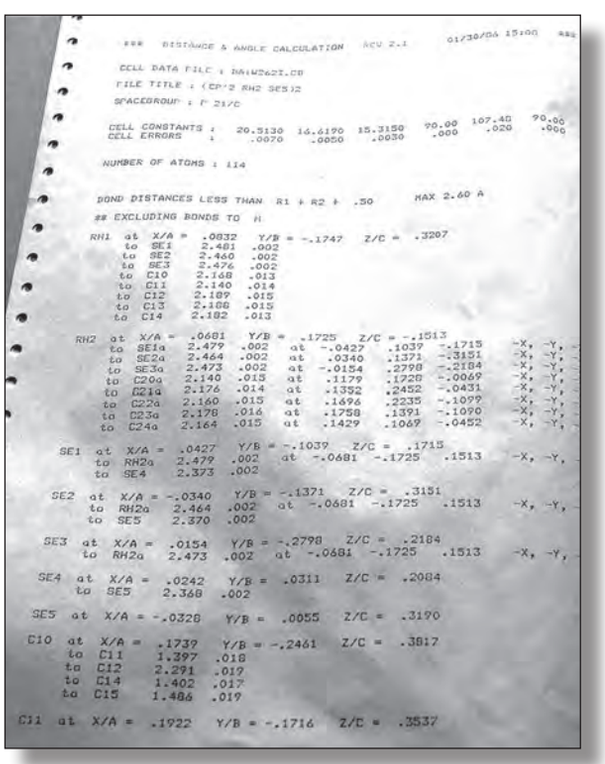


Fig. 3

The CIF format (Crystallographic Information File) was developed by the International Union of Crystallography (IUCr) and has become a worldwide standard for crystallographic information (http://www.iucr.org/iucr-top/cif/index.html).

During the last years, about 1,300 to 1,500 structures were deposited on average, and an increasing number of about 600 to 1,300 structures were called up by users.

By the way, registered users of the Inorganic Crystal Structure Database (ICSD) can also find (with a short production-dependent delay) the deposited inorganic crystal structure data sets in the database, provided that they have been published.

FIZ Karlsruhe is still interested in depositing as many structures as possible, irrespective of whether they are also collected on a publisher's server or not. Furthermore, authors may send their own data sets not intended for publication in a journal (e.g. private communications) to FIZ Karlsruhe for deposition.

In summary, regarding the use of the depot by the crystallographic community it is obvious that the deposition of structure data by FIZ Karlsruhe is still needed and highly appreciated. FIZ Karlsruhe, on its part, is pleased to offer its free service to the scientific community in the future.

Searching across all compound classes

Analysis of spinel compounds in ICSD

Andreas Barth

The Inorganic Crystallographic Database ICSD on STN International contains 169,800 records (with update May 2014). Each record consists of bibliographic information, chemical compound information, and crystallographic data. The specific combination of information supports searches which are difficult to perform in other compound databases. For example, it is possible to search for text in titles in combination with chemical compound information. This provides an easy possibility to search for compound classes which cannot be described through specific compound information such as the molecular formula. An interesting example for such a compound class is provided by the spinel compounds. The spinel compound class is named after the magnesium aluminum spinel $MgAl_2O_4$ and it describes a specific crystallographic structure. $MgAl_2O_4$ is a colorless, almost transparent crystal, but it can become a very beautiful colored gemstone resulting from traces of other chemical elements, e.g., in rubies and sapphires. The oxygen atoms (O_2-) build a cubic face-centered crystallographic system with the Mg_2+ atoms occupying 1/8 of the tetrahedral sites and the Al_3+ atoms occupying 1/2 of the octahedral sites in the lattice.

The different spinel compounds may differ considerably in their molecular composition. Hence, it is not possible to search for combinations of chemical elements. In the literature, spinels are often associated with the formula type AB2X4 and with the Pearson symbol [1] cF56 (cubic – all faces centered with 56 atoms in the unit cell). A search in ICSD will show how these concepts are related to each other. A simple search for these terms shows that there is indeed an overlap, but spinel compounds may have other formula types than AB2X4 and are not restricted to the Pearson symbol cF56 (s. Table 1). The total number of spinel compounds (search for spinel*) is 3182 and the overlap with both AB2X4 and cF56 is 1821. A search is rather complicated since the word "spinel" is an unspecific text term, the molecular formulas may differ significantly, and neither the Pearson symbol nor the formula type are unique. An illustration of the relationship between spinels and the Pearson symbol cF56 is shown in Figure 1.

Search Term	# Hits
Spinel* (*: wildcard)	3182
cF56	3608
AB2X4	6704
Spinel* AND cF56	2465
Spinel* AND AB2X4	1968
Spinel* AND cF56 AND AB2X4	1821

Table 1: Correspondence between spinel compounds, the formula type AB2X4, and the Pearson symbol (PRS)

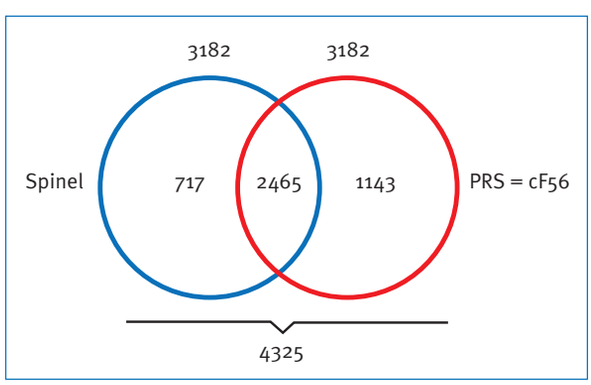


Fig. 1: Search results for "spinel*" vs. Pearson symbol PRS = cF56 (see also Table 1)

In addition to the "standard" spinel compound with molecular formula $MgAl_2O_4$ there are many variations of elements and stoichiometries; the most frequent molecular formulas (in Hill order) are shown in Table 2. $MgAl_2O_4$ is described in 201 documents, $MgFe_2O_4$ occurs 145 times, and $NiAl_2O_4$ 105 times. In positions 17, 18, and 20 there are also spinels with Cr atoms and small impurities of Mn and Zn.

It has been shown that the formula type of spinel compounds is not restricted to the type AB2X4 (s. Table 1). There are 116 different types in total, and the top 20 formula types are given in Figure 2.

The oxide and sulfide spinels with three elements were further analyzed in order to see which elements may be substituted for Mg (an analogous analysis for the substitution of Al can as well be made) and the results are shown in a heat map in Figure 3. Beside Mg, the element

TERM #	# DOC	% DOC	MF
1	201	6,32	$Al_2 Mg_1 O_4$
2	145	4,56	$Fe_2 Mg_1 O_4$
3	105	3,30	$Al_2 Ni_1 O_4$
4	96	3,02	$Fe_3 O_4$
5	58	1,82	$Mg_2 O_4 Ti_1$
6	52	1,63	$Al_2 O_4 Zn_1$
7	43	1,35	$Al_2 Fe_1 O_4$
8	40	1,26	$Al_2 Mg_{0.99} O_4$
9	35	1,10	$Fe_2 O_4 Zn_1$
10	25	0,79	$Al_2 Co_1 O_4$
11	23	0,72	$Al_1 Fe_1 Mg_1 O_4$
12	23	0,72	$Al_2 Cu_1 O_4$
13	23	0,72	$Ni_2 O_4 Si_1$
14	16	0,50	$Co_3 O_4$
15	16	0,50	$Cr_2 Mg_1 O_4$
16	16	0,50	$Cr_2 O_4 Zn_1$
17	15	0,47	$Al_{1.834} Cr_{0.162} Mg_{0.995} Mn_{0.002} O_4 Zn_{0.007}$
18	15	0,47	$Al_{1.868}$ Cro.123 $Mg_{0.999} Mn_{0.002} O_4 Zn_{0.007}$
19	15	0,47	$Li_1 Mn_2 O_4$
20	14	0,44	$Al_{1.928} Cr_{0.066} Mg_{1.003} Mn_{0.001} O_4 Zn_{0.002}$
20	14	0,44	$Al_{1.928} Cr_{0.066} Mg_{1.003} Mn_{0.001} O_4 Zn_{0.002}$

Table 2: Distribution of molecular formulas in spinel compounds

Fe shows the highest occurrence in this table. It seems that Mg is not well represented in the other spinel compounds.

Obviously, the most important spinel is the "traditional" spinel $MgAl_2O_4$; it occurs in 201 documents (s. Table 2, with mostly different crystallographic data) but there are only 26 different stoichiometries of Mg-Al-O compounds (Mg ranges from 1 to 0.35 while Al ranges from 1.998 to 2.44). A further analysis of the stoichiometries of the Al-Mg-O spinels shows a distinct relationship between the two elements Al and Mg as shown in Figure 4. The ratio Mg/Al ranges from 0.5 (1:2) to 0.14 (0.35:2.43) with two clusters around 2.05 and 2.4 Al atoms. This means that Al can partly substitute Mg in spinel compounds.

The other Al-R-O systems (with 4 oxygen atoms, R: any element except Mg) do not show this linear stoichiometric relationship: the R/Al ratio is mostly around 1:2 and sometimes 2:1 (e.g. for R=Fe both ratios are possible).

A visualization of the unit cell of two different Mg-Al-O systems shows how Mg is partly substituted by Al (see Figure 5). The first unit cell has been generated from the crystallographic data of $MgAl_2O_4$ (e.g. see [2] and the second is based on data for a synthetic defect spinel with the molecular formula $Mg_{0.388} Al_{2.408} O_4$ [3].

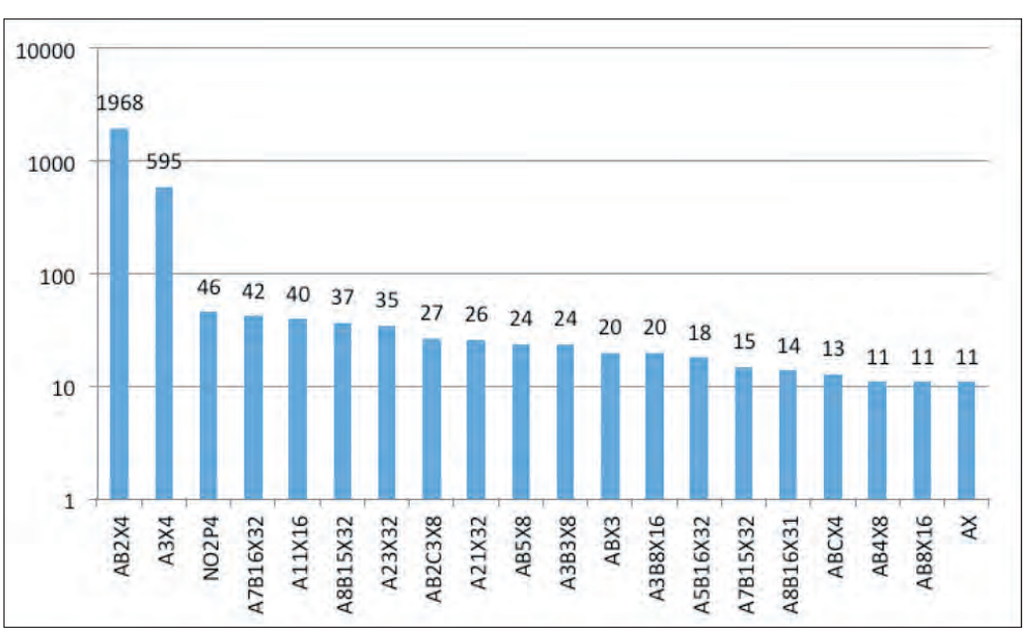


Fig. 2: Distribution of Formula Types in Spinel Compounds (logarithmic scale)

ELS	Chalc	Li	Mg	Ca	Yb	Ti	V	Cr	Mo	Mn	Fe	Co	Ni	Cu	Zn	Cd	Hg	Al	In	Sn
Al	O	1	26				2			2	8	3	4	2	2	1				
Al	S						1	3	2	1					3	1	1		1	
Ga	O		6							1	1	1	1	1	1	1				
Ga	S					3		5	4										1	
In	O		2	1						1	3					1				
In	S		2		2			2		1	3	1	1	5	1	1				5

Fig. 3: Heatmap of oxide and sulfide spinels with elements of group 13 and a total number of 3 elements. The number in the cells designates the number of compounds, e.g. there are 26 Al-Mg-O compounds with different stoichiometries

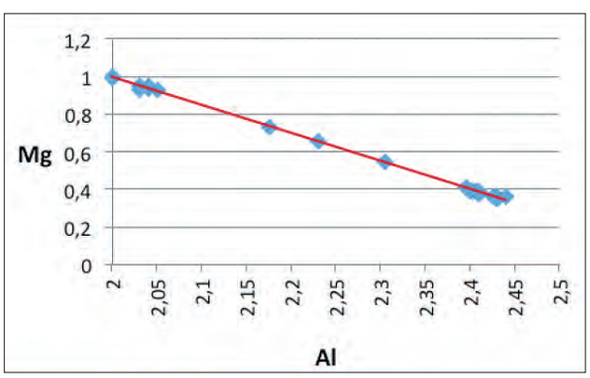


Fig. 4: Stoichiometric relationship between Al and Mg in Al-Mg-O spinels (all systems have four oxygens)

In the latter case the aluminum atoms occupy not only the octahedral sites but also some of the tetrahedral sites in the lattice. The Pearson symbol in this case is cF54.

Finally, it should be noted that molecular analysis can also be performed in the highly comprehensive databases from Chemical Abstracts Service (except for the crystallographic data and the formula type). The analysis is a little bit more complicated since the concepts have to be searched in two separate databases: CAplus^{SM} and Registry^{SM} [4]. However, a search in these databases may provide more comprehensive results.

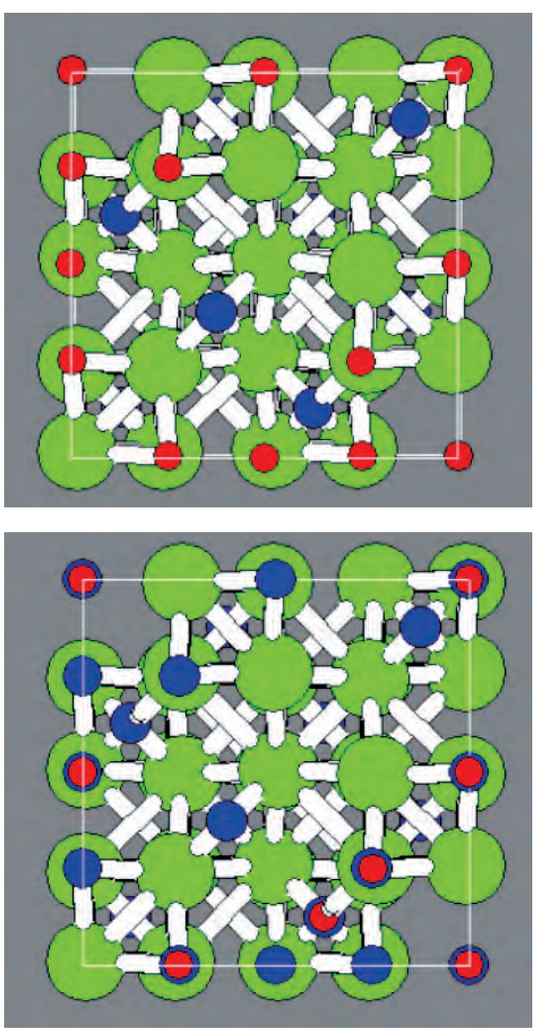


Fig. 5: Comparison of unit cells from $MgAl_2O_4$ and $Mg_{0.388}Al_{2.408}O_4$ (Mg: blue, Al: red, oxygen: green)

References

[1] Hubbard, C. R.; Calvert, L. D. (1981). The Pearson Symbol. Bulletin of Alloy Phase Diagrams Vol. 2 (2), pp. 153-157.

[2] Zorina, N. G.; Kvitka, S. S. (1968). Refinement of the structure of the spinel $Al_2Mg O_4$. Kristallografiya, 13, pp. 599-600.

[3] Basso, R.; Carbonin, S.; della Giusta, A. (1991). Cation and vacancy distribution in a synthetic defect spinel. Zeitschrift fuer Kristallographie, 194, pp. 111-119.

[4] Barth, A.; Marx, W. (2012). Stimulation of ideas through compound-based bibliometrics: Counting and mapping chemical compounds for analyzing research topics in chemistry, physics, and materials science. ChemistryOpen Vol. 1 (6), pp. 276–283. http://onlinelibrary.wiley.com/doi/10.1002/open.201200029/pdf. (November 25,2014)

Using the Cambridge Structural Database to assign atom types

Bond lengths statistics as a tool for crystal structure analysis – a case study with a crystalline silicophosphate

Uwe Böhme* and Sandra Jähnigen*

Abstract

The assignment of atom types is one of the first steps after solving a single crystal X-ray structure analysis. It is often difficult or even impossible to assign correct atom types if there are atoms with similar atomic number present in the structure solution. The Cambridge Structural Database (CSD) allows to perform statistical analyses of bond lengths and delivers an elegant solution for the problem of ambiguous atom types. Mean bond lengths of the structural units SiO_4, SiO_6, PO_4, and PO_6 were calculated from CSD data. These were compared with bond lengths from the structure solution of a hitherto unknown silicophosphate. This procedure allowed for an unambiguous assignment of atom types in the structure solution.

The Cambridge Structural Database

The Cambridge Structural Database (CSD) is a complementary database to ICSD. The CSD is compiled and distributed by the Cambridge Crystallographic Data Centre (CCDC). This is a non-profit company and has been a registered charity since 1987. The producers of the CSD describe it as the world's repository of experimentally determined small-molecule organic and metal-organic crystal structures. The database contains the results of over half a million x-ray and neutron diffraction analyses. This unique database of accurate crystal structures has become an essential resource to scientists around the world. Each crystal structure undergoes extensive validation and cross-checking by expert chemists and crystallographers to ensure that the CSD is maintained to the highest possible standards. Also, each database entry is enriched with bibliographic, chemical and physical property information, adding further value to the raw structural data. These editorial processes are vital for enabling scientists to interpret structures in a chemically meaningful way. The CSD is continually updated with new structures (>40,000 new structures each year) and with improvements to existing entries [1].

We use the in-house version of the CSD. Searches are performed with the software ConQuest. Geometric parameters obtained from a search with ConQuest are exported to the software module VISTA for statistical analysis. Further analysis of the data is possible with Mercury. This software offers a comprehensive range of tools for structure visualization, the exploration of crystal packing and further statistical analyses.

The problem with the silicophosphate structure

Silicophosphates (aka SiPOs) are promising materials for the use as lasers, optical fibers and orthopaedic implants. That is because of their special material and optical properties such as low melting points and high refractive indices. Furthermore, they contribute to the understanding of geological processes. Until now, many different silicophosphate structures were discovered. Remarkably, therein the silicon atoms are octrahedrally-coordinated to six oxygen atoms. This is in contrast to most silicates in nature where silicon occurs in tetrahedral coordination. All known SiPOs are composed of SiO_4 and PO_4 tetrahedra as well as SiO_6 octahedra. These structural building blocks are connected differently to each other and yield a multitude of various silicophosphate structures [2]. The classical synthesis of such silicophosphates is performed by melting processes, chemical transport reactions or sol-gel-techniques. Recently, we discovered another way to prepare silicophosphate materials with octahedral silicon structures. The syntheses were performed at room temperature and normal pressure in diethylether solution. The reaction of tetraethoxysilane (TEOS) and anhydrous H_3PO_4 gives amorphous Si-O-P materials with SiO_4, PO_4 and SiO_6 structural units. On the other hand, chloroethoxysilanes and anhydrous phosphoric acid yielded oligomeric silicophosphates.

* TU Bergakademie Freiberg, Institut für Anorganische Chemie

From the latter reactions one crystalline product was obtained. The X-ray structure analysis of this product leads to several problems which will be discussed in the following sections.

The structure solution delivered a large complex ion which consists of silicate and phosphate subunits. Furthermore, tetraethylammonium and solvent molecules were identified. After several refinement cycles it became obvious that the unambiguous assignment of the atom types phosphorous and silicon in the structure would be no simple task (Fig. 1). Both elements have similar atomic radii, comparable electron densities, and are able to create the same coordination geometry with oxygen. Especially the subunits SiO_4 / PO_4 and SiO_6 / PO_6 could both be present in the structure. It would be most helpful to discriminate these subunits on the basis of different bond lengths. Therefore we were in need of reliable bond lengths for these four substructures.

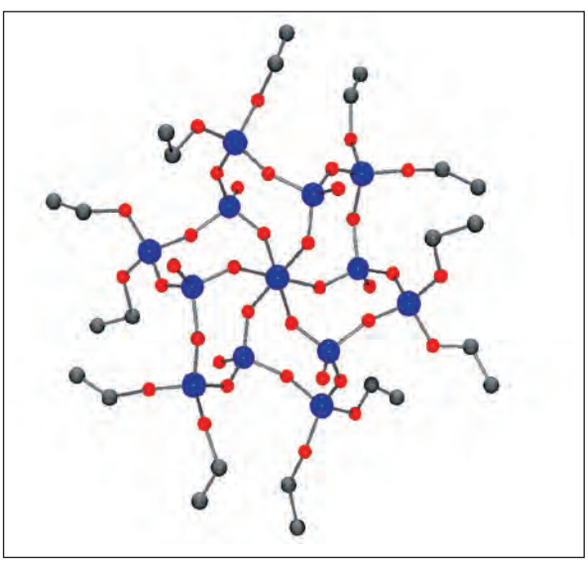


Fig. 1: Structure of the silicophosphate anion after isotropic refinement (dark grey – carbon, red – oxygen, blue – silicon or phosphorous?)

Bond lengths statistics from the CSD

A search for the substructures SiO_4 / PO_4 and SiO_6 / PO_6 was performed with the CSD in order to obtain statistically reliable bond lengths derived from experimental data. For that purpose the structural elements were drawn within the module ConQuest and the number of bonded atoms was fixed to 4 or 6. The option “ADD 3D” was used to extract the bond lengths from the structures in the database (Fig. 2). The search for the SiO_4 unit gave 587 hits. These were analysed in the software module VISTA. The mean bond lengths for this hit list is 1.618 Å with minimal and maximal values of about 1.53 and 1.71 Å, respectively (see Fig. 3). The outliers were analyzed. These outliers could be eliminated from the statistical analysis within the VISTA software. But this should be done only for good reason. The same procedure was applied to the structural units PO_6, PO_4, and SiO_6. The result of this statistical analysis is summarized in Table 1.

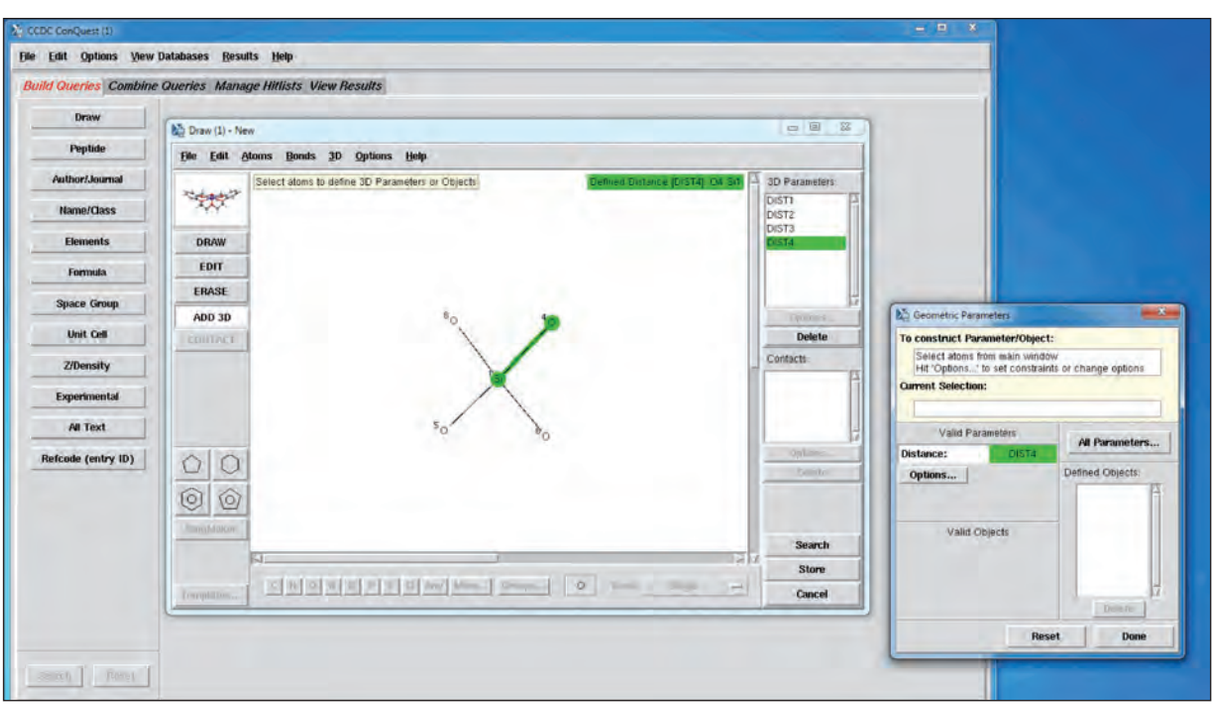
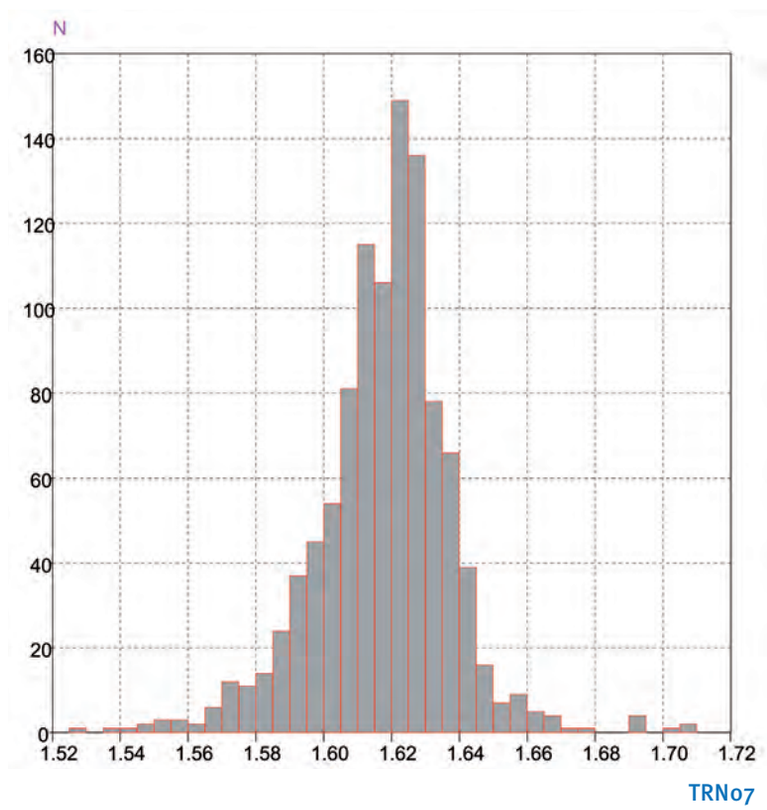


Fig. 2: Graphical user interface of the CSD

TRN07



Plot Data
File=/home/aoch/aschv
Test=1
Tot.Obs.=1036
Obs.=1036
Supp.=0

X-axis
Min.=1.527
Max.=1.709
Range=0.181
Mean=1.618
Mean SE=0.001
Sample SD=0.020

Histogram
Median=1.620
Skew=0.190
Quantile=10.000
LQ=1.593
HQ=1.639
Bin Width=0.005
Max. Bin=149.000

Fig. 3: Bond lengths statistics for the SiO_4 unit

structural unit	mean bond lengths in Å	hits	standard deviation	lower quartile	upper quartile
PO_4	1.535	3217	0.016	1.523	1.546
SiO_4	1.618	587	0.020	1.593	1.639
PO_6	1.717	58	0.020	1.705	1.719
SiO_6	1.782	53	0.039	1.766	1.788

Tab. 1: Mean bond lengths of structural units derived from the CSD

The assignment of atom types in the silicophosphate structure

The mean bond lengths derived from the CSD were used to assign the correct atom types in the X-ray structure under investigation. Three different atom types X, Y, and Z had to be assigned. These are schematically shown in Figure 4. One hexacoordinate atom X is situated at the heart of the complex. The X-O bond lengths are between 1.757 to 1.770 Å. These bonds are too long for a PO_6 unit (see Table 1). That means the central atom X corresponds to a silicon atom. The central SiO_6 octahedron is surrounded by six YO4 tetrahedra. These feature bond lengths of 1.513 to 1.574 Å and thus are too short for Si-O bonds. Furthermore, the atoms Y are supplemented by a terminal oxygen atom, which is typical for a phosphate unit. Therefore the atoms Y were assigned as phosphorous atoms. The peripheral atoms Z each link to two phosphate groups and are bound to two diethoxy groups. The bond lengths Z-O are between 1.579 and 1.633 Å. This is substantially longer than the previous set of bond lengths and corresponds to silicon-oxygen bond lengths of tetrahedral silicon atoms. In this way the atoms types in question were assigned.

The final assignment of atom types is shown in Figure 5. Obviously, a polynuclear silicophosphate complex was formed. Within the complex anion a hexacoordinated silicon atom (Si1) is in the middle of the complex. The SiO_6 octahedron is bonded to six PO4 tetrahedra (P1 to P6). These in turn are bridged by six diethoxysilicate groups (Si2 to Si7).

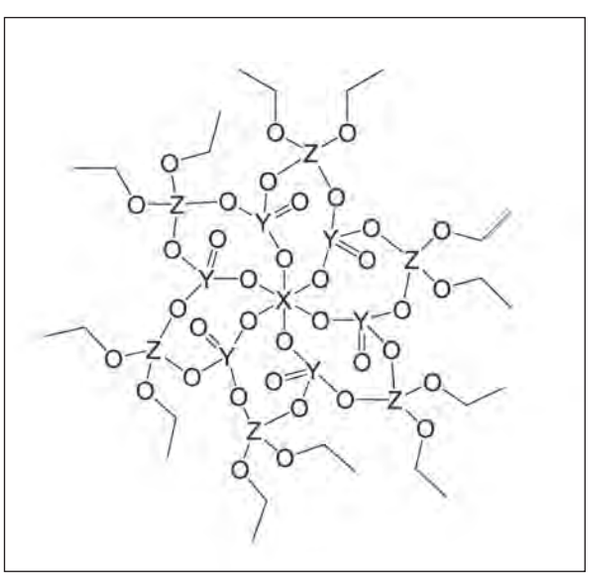


Fig. 4: Schematic representation of the structural problem

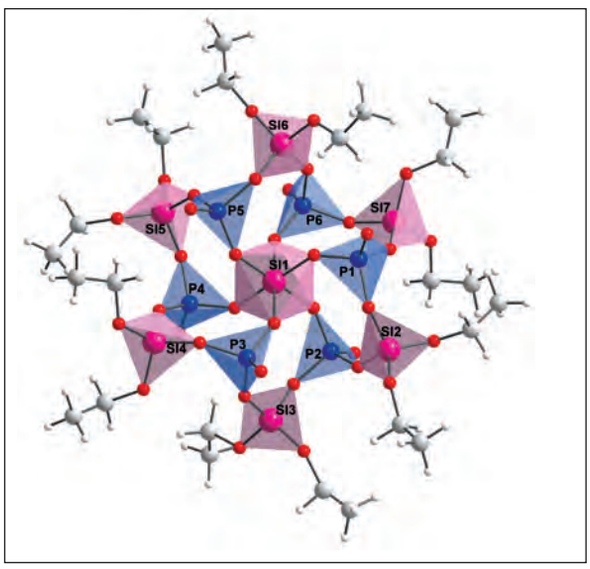


Fig. 5: $[Si^{VI}(PO_4)_6(Si^{IV}O_4Et_2)_6]$-Ion with final assignment of atom types

Summary and Outlook

It was possible to assign the correct atom types in the X-ray structure of a new silicophosphate with the help of bond length statistics drawn from the Cambridge Crystallographic Database. The structure has been published in Chemical Communications [2] and inspired our further work [3]. This case study demonstrates the ability of the CSD as a helpful tool in structure elucidation. Until now such analyses are not possible with ICSD. Here, a search can be done by adding names, specific compositions or structural formula. It would be desirable to perform bond lengths statistics in the future also with ICSD.

References

[1] Information taken from the homepage of CSD: www.ccdc.cam. ac.uk/Solutions/CSDSystem/Pages/CSD.aspx (November 25, 2014)

[2] Jähnigen, S.; Brendler, E.; Böhme, U.; Kroke, E. (2012). Synthesis of silico-phosphates containing SiO_6-octahedra under ambient conditions – reactions of anhydrous H_3PO_4 with alkoxysilanes, Chem. Commun. 48, pp. 7675–7677.

[3] Jähnigen, S.; Brendler, E.; Böhme, U.; Heide, G.; Kroke, E.(2014). Silicophosphates containing SiO_6 octahedra – anhydrous synthesis under ambient conditions, New J. Chem. 38, pp. 744-751.

Predicting new compounds by combining and analyzing large data volumes

Using ICSD to predict new properties, compounds and modifications

J. C. Schön*

An overview over possible prediction related applications of ICSD as a searchable database is given.

Introduction

ICSD constitutes a large database of crystalline structures found in solid state chemistry and materials science for a multitude of chemical systems. In particular, one finds an overview over the possible modifications in chemical systems that have been discovered experimentally (plus some theoretically predicted structures). Experience has shown that there appear to exist some correlations between the structure(s) a chemical compound takes on and the other physical properties we observe for the compound.

This had led to optimistic suggestions that one can establish perfect structure-property relationships which will produce easy ways for predicting new materials. In practice, we note, however, that there exist numerous examples where the same crystal structure is found for e.g. both insulators and metals, and thus these considerations need to be taken with a large grain of salt. Nevertheless, it is often reasonable to assume that if two compounds belong to the same general class of chemical systems, then they will share some of their properties, and thus some trends in these properties may truly be correlated with the various modifications possible in these systems. This might well allow us to predict the properties of not-yet-synthesized compounds from the structure type they crystallize in.

But this requires us to first take the step to figure out, whether a given structure type will constitute a stable modification in a particular chemical system. For practical purposes, this is equivalent to predicting, which possible crystalline modifications can exist in a given chemical system. Such modifications can be divided into two groups: those that agree with structure types that have been observed in other chemical systems (which may be either chemically similar to the system under consideration, or actually quite different and only show the same composition), and those structure types that have never been observed in any chemical system or only been realized in chemical systems that are so different chemically from the compound under consideration that "chemical intuition" will not suggest these structure types to us.

Clearly, in the case where a completely new structure type is going to appear during the synthesis of a new compound, a database can only be of very limited assistance in predicting the structure. In that case, we have to turn to fundamental physical principles that tell us that every (meta)stable chemical compound corresponds to a locally ergodic region on the energy landscape of a chemical system. At low temperatures such a region corresponds to a local minimum of the potential energy in the space of atom configurations, but at elevated temperatures such a region can be quite large encompassing several or even many local minima. Over the past twenty, twenty-five years finding such local minima of the potential energy via global optimization methods has been developed to a certain proficiency, with different research groups proposing different global optimization algorithms (see e.g. [1–4] and references therein).

However, the number of such local minima grows exponentially with the size of the system. Even if one excludes defect minima and amorphous structures by focussing on only small periodically repeated simulation cells containing up to, say, ten formula units, the number of minima can be quite overwhelming, and the danger that one misses an important modification increases with system size. Furthermore, we often need to employ simplified energy functions for the global search, refining the resulting minima afterwards with more accurate ab initio energy functions using Hartree-Fock- or Density Functional Theory (DFT)-based computer codes since the ab initio energy calculations are too time-consuming for large global searches.

* MPI for Solid State Research, Stuttgart

In this case, information drawn from databases like ICSD can be a valuable resource to heuristically guide us to find promising candidates for modifications in a given chemical system.

In this essay, we are going to outline some ongoing and possible future approaches to structure and property prediction with the help of database analyses, specifically with the analysis of ICSD. Many of the concepts and approaches discussed in this essay have already been presented in the literature in some fashion, together with applications. Thus, we are not going to discuss methods or their applications in detail, but focus on providing a general presentation of the use of databases for structure prediction and closely related questions that might inspire the reader to proceed further along the directions mentioned.

Energy landscape, chemical similarity, and structure prediction

Energy landscape

The world of all conceivable atom arrangements for N atoms is called the configuration space ($N \approx N_{Avogadro}$). A point in the configuration space can be visualized as a vector $\vec{R}$ with 3N coordinates (each atom contributes its position vector $\vec{r} = (x, y, z)$). For each such configuration, we can compute the potential energy, and the 3N-dimensional hypersurface of the energy over the configuration space is the so-called (potential) energy landscape. As we know from classical mechanics, the dynamics of the chemical system is given by the forces acting on the atoms, i.e. the gradient of the potential energy.

Clearly, if one picks such a vector $\vec{R}$ at random, the structure associated with it will be a random arrangement of atoms as one finds in the gas or liquid phase, and the crystalline structures listed in ICSD are singular points that will never be seen by randomly picking points in configuration space. However, most of these randomly selected atom arrangements are quite high in energy, and physics tells us that the chemical system will preferentially occupy those regions of configuration space that are local minima of the free energy $F = E - TS$ at a given temperature. Specifically, a metastable modification of a chemical compound corresponds to a locally ergodic region of configuration space, i.e. a region containing many (similar) atom configurations, which is locally equilibrated with a low free energy and kinetically stable enough such that the system does not leave this region, on experimental time scales.

Quite generally, the sets of configurations around a minimum or groups of minima on an energy landscape represent the locally ergodic regions at low or intermediary temperatures, respectively, and the kinetic stability of these regions grows with the height of the energetic and entropic barriers surrounding the region. Furthermore, the (crystal) structure that we associate with such a metastable modification is the average over all the configurations in the locally ergodic region; this average is, at least partly, reflected in the so-called thermal ellipsoids derived for the atoms in a crystal structure from the experimental data.

Structure prediction

From a physical point of view, structure prediction is therefore equivalent to finding all the locally ergodic regions of a chemical system. Typically, one proceeds by first finding all minima, and then identifying the barriers surrounding them to estimate their stability.

Usually, the focus is on the energetically low-lying minima, and thus the first step of the search is equivalent to a global optimization on a highly complex multi-minima energy landscape. Such a procedure is very time-consuming, the computational effort typically growing exponentially with the size of the system. This remains true even if one employs as many simplifications as possible, such as simulation cells containing few formula units with periodic boundary conditions and simple fast-to-evaluate cost functions instead of ab initio energies. As long as we are only interested in crystalline structures for systems where the energy landscape is well-approximated by empirical potentials, this simplification is reasonable, but if one wants to analyze amorphous or defect-controlled compounds much larger simulation cells are needed, of course.

Much progress has been made in developing efficient ways to perform such global optimizations for crystalline chemical systems, but as one knows from ICSD, there exist many crystal structures that contain more formula units (Fig. 1) units or cannot be well described by empirical potentials, such that even highly refined global search methods cannot guarantee success within reasonable computation times. Furthermore, as we have seen, in general it is not sufficient to find only the global minimum because in many instances it is a metastable modification, perhaps only realizable as a nano-crystal or thin film, which exhibits the desirable properties. One way to address this problem is by exploiting our knowledge about chemical systems collected in databases such as ICSD.

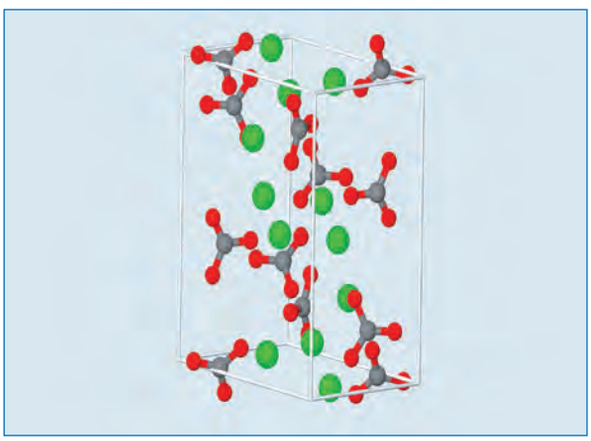


Fig. 1: Example of a structure with z=10 ($CaCO_3$ at high pressure)

While the purist might reject such heuristic procedures, it is obvious that one should not refuse any methodology that can assist us in identifying candidates for metastable compounds in a chemical system, even if it is only heuristic and clearly limited. In particular, if one's goal is not to prove the feasibility of a particular prediction methodology but to find all possible modifications for a specific given chemical system, then one should employ every method available. For more information about the many global optimization methods that are currently being employed, we refer to the literature (see references mentioned above), and for the remainder of this essay, we are only going to discuss approaches that rely on the availability of such databases.

Similarity of chemical systems

From an abstract point of view, every chemical system is different, and thus there is no a priori reason to assume that the possible modifications are going to exhibit the same or even a similar structure as any other system. And if one defines a "structure" via the exact cell parameters of a crystalline unit cell and the specific atom positions within this cell, the numerical values of these coordinates will be different from those found in any other chemical system. But the experience of the practicing crystallographer, chemist and materials scientist has shown that among the structures observed in the world of crystalline compounds there are many, which differ from each other only by slight changes in atom parameters and/or cell parameters, such that they appear to be the same, at least visually, once one disregards the specific chemical identity of the atoms involved (Fig. 2). The crystallographic literature contains many algorithms that have been suggested to quantify this similarity, e.g. based on symmetry [5], geometry [6], or topology [7] considerations, each being most useful in specific contexts. But at the moment the decisive issue is that once we subsume all these similar structures under one structure type, we realize that instead of hundreds of thousands of crystalline structures, "only" tens of thousands remain, some of which are found to occur in hundreds of different chemical systems. A nice such classification into isopointal structure types based on symmetry considerations, with a refinement to isoconfigurational structure types based on additionally using cell parameters and atom coordinates, has recently been introduced in ICSD [8]; by now most of the structures found in ICSD have been assigned to one of these types.

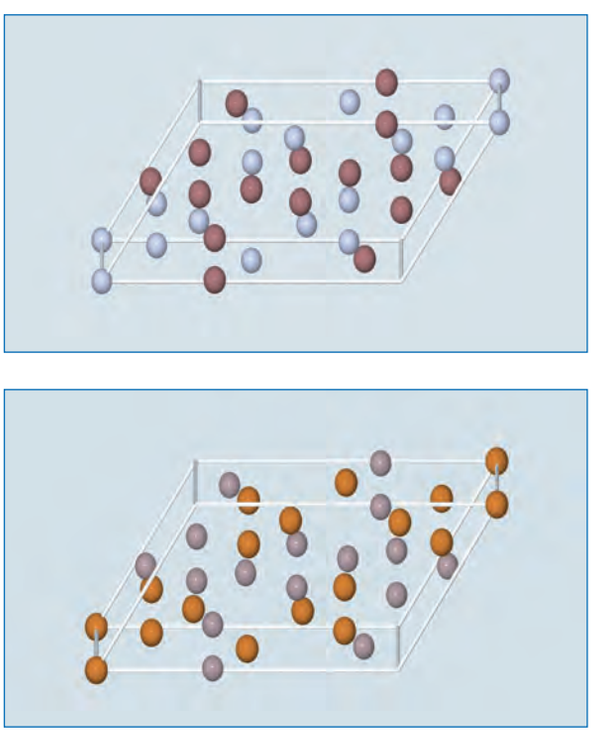


Fig. 2a + b: Structures of InPt(a) and AlCu(b) which both belong to the same structure type

From a physical point of view, the occurrence of the same structure type $\hat{T}$ in two different chemical systems means that the energy landscapes of these two systems each possesses a local minimum that belongs to this structure type $\hat{T}$. And if, as not infrequently happens, several of the modifications present in one of the systems are also observed in the second system, then it is a reasonable expectation that the two energy landscapes are going to be quite similar. Then one can hypothesize that many of the other modifications that are only found in one of the two systems are nevertheless capable of existence in the other one.

Clearly, such an educated guess is very helpful in our search for new modifications in a particular chemical

system since we can right away perform a local minimization to verify whether this structure type is present as a kinetically stable modification in the second system, which is much faster than the global search necessary otherwise. We note here that this assumption of landscape similarity with the consequence of a highly desirable transferability of the results found for one landscape to another one, has also been supported by many global landscape explorations.

One caveat is, of course, that in most cases the global search has been performed only with simplified energy functions, and thus we are comparing only models of the true energy landscapes (belonging to similar classes, e.g. hard-sphere two-body interaction models or all-electron ab initio calculations or density functional pseudo-potential models) of the two systems, which might result in misleading similarities due to the inherent features of the models and interactions. Thus, one must be aware that the true landscapes of the two systems are usually still going to be different in certain aspects: for example, not all minima found in one system will be present in the other, or the energy rankings of those minima present in both systems may be different.

From a chemical point of view, the existence of modifications with the same structure type in two different chemical systems can often be correlated with the "chemical similarity" of the two systems. Quantities such as differences in electronegativity, cation-anion-(size) ratios or total valence electron concentration are commonly employed to categorize chemical compounds, and are used to postdict, and sometimes even to predict, by analogy or "chemical intuition" (if the straightforward analogy is not an obvious one), the structure of a new chemical compound. This is a time-honored and very successful approach in experimental chemistry, as most famously demonstrated by the success of the periodic table that still is one of the most valuable organizational tools of the practicing chemist.

In this context, we note that from our experience with the energy landscapes of many chemical systems we have found that even for relatively small simulation cells with few atoms the minima that are associated with known structure types constitute not more than half of the local minima found. But among the lowest-energy minima, this proportion can increase to 90%, at least for systems with simple A_nB_m-type compositions (Fig. 3). Of course, this mainly highlights the fact that, with regard to synthesis, we have already carefully explored most of the binary systems with simple composition ratios n : m. Thus the number of unknown structure types in this class of systems is rather small (especially if one

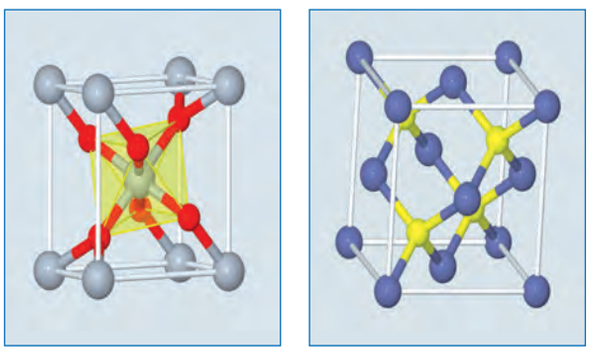


Fig. 3a + b: Example of simple structure types: (a) TiO_2 (Rutile) and (b) ZnS (Zincblende)

allows for a certain amount of distortions within a structure family). But this also supports the suggestion that the energy landscapes of chemically similar systems can show a relatively high degree of similarity. However, this high proportion is much less frequently found for compounds with more complex compositions.

ICSD-based structure prediction for given chemical systems

One obvious way to exploit the putative similarity of the energy landscapes of many chemical systems with the same composition formula consists in replacing the atoms in a known structure type by their analogues in the chemical compound of interest, and then to perform a local minimization, in order to check for the kinetic stability of this type in the new system. By repeating this procedure for all appropriate structure types found in ICSD, we can gain a certain overview over the minimum structures of the new system.

A more refined way is to use these minima as starting points [9] for a global optimization technique such as the threshold algorithm [10], which explores the configuration space below an energy lid that is accessible from a given energy minimum. By increasing this limiting lid, one can, in principle, globally explore the whole energy landscape.

We note that one important preliminary task required before such an ICSD-based search can be performed is the identification of all structure types that are present in ICSD by using one of the comparison algorithms mentioned above. To a certain extent, the classification already provided by ICSD can be employed, but since this classification is based primarily on symmetry one needs to be careful. Hence, one would first classify all ICSD structures into structure families via a geometry-based structure comparison, and then define the center of this

group of structures as the structure father or structure type representative [11]. Once this has been achieved, one can take these prototypical structures as starting structures for the further global search or local minimizations, respectively.

Here, we note that for the use in energy landscape exploration and structure prediction, a geometrical comparison criterion is usually most appropriate: the neighbors in the real configuration space of atom arrangements are defined via sums over the Euclidean distances between corresponding atoms – a geometrical quantity – while topological similarity is based on the bond network, which only rather indirectly correlates with the shape of the energy landscape, and symmetry-based similarity cannot really be mapped to the energy landscape at all. Nevertheless, following the dictum that one should not discard any systematic way to generate structure candidates, one should not hesitate to add all those structure types that have been gained on the basis of topological and symmetry arguments to the starting points that are selected based on geometrical similarity.

In this context, one should keep in mind that there exist many chemical correspondences between e.g. binary and ternary or higher phases. Thus, it might well be that the unknown structure of the new modification or compound in a binary A/B system might be identical to the structure of some ternary A'/B'/C' system, if one only identifies A' with A and B' and C' with B! (Fig. 4) To explore this, we have in the past analyzed ICSD with regard to similarity between structures and structure types in binary, ternary, quaternary, etc. systems.

This procedure will result in a family relationship tree that connects every structure not only with other structures of the same type (forming a structure family), but also with structures belonging to related families in the sense described above. This procedure can provide us with additional ICSD-based structure types, even if they do not exist, strictly speaking, so far in form of a synthesized compound.

Trying to predict structures by chemical analogy has been the oldest method in the literature [12]. If we consider the statistics mentioned above, it is no surprise that there have been quite a number of successes, especially in the field of high-pressure structures, where the knowledge that two chemically similar compounds have the same structure at standard pressure and system one has a certain high-pressure structure, makes it a relatively easy target to predict that the second system will also exhibit a high-pressure structure with the same structure type. And rules such as the pressure-coordination rule allow us a pretty well-educated guess about the similarity of the structures of chemically analogous systems, e.g. a high-pressure modification of system two exhibiting a structure identical with the one of the standard pressure modification of system one.

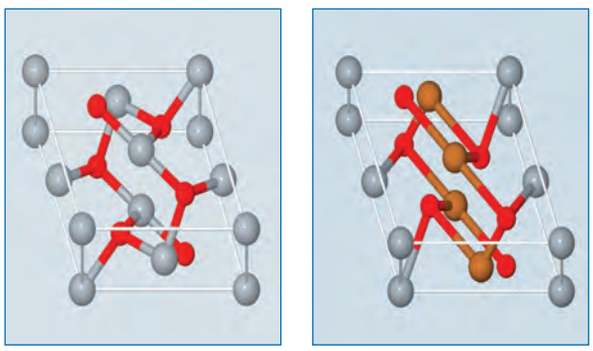


Fig. 4a + b: A binary (Ag_2O_2 (a)) and a ternary ($AgCuO_2$ (b)) structure belonging to the same structure type (AgO)

From the point of view of exploiting ICSD, this combination of chemical intuition and hard structural data in the database has been a reasonably successful one, leading many people to the conviction that database mining is all that is needed for successfully predicting the structures of new compounds. In particular, groups around Curtarolo and Ceder [13] have been developing this approach, culminating in the so-called Materials Genome Project [14]. And molecular crystal structure prediction and the prediction of secondary and ternary structures of proteins have also been guided by this kind of data-mining approach (for a review see e.g. [15]): In the case of molecular crystals, the statistics has shown that the symmetry groups for about 90 % of all structures of molecular crystals belong to only about ten different space groups, and that many of these structures can be described with only one or two molecules in the asymmetric unit. As a consequence, many search methods for molecular crystal structures exploit this information by e.g. systematically scanning all possible atom arrangements that can be generated by applying these space group symmetries to one or two molecules (where the scanning includes the systematic change of the cell parameters and the orientation of the molecules with respect to the cell axes and to each other).

Similarly, for protein structure prediction, much success has been achieved by taking the primary sequence of the bases of a protein, and then comparing pieces of these sequences, or the whole sequence, to corresponding pieces of known protein structures, deriving in this fashion the secondary structure elements like the α-helices and β-sheets for the unknown protein, and even establishing good guesses regarding the tertiary folded

structure. Again, the large number of already solved protein sequences and structures serve as a confidence-building foundation for such ”structure prediction by analogy”. One even has attempted to use statistical methods to estimate, how many ”basic protein sequences” and ”protein structures” are still missing, although there is clearly a self-reinforcing feedback going on: proteins that are similar can be solved with regard to their sequence and structure by similar methods, thus giving too much weight to the ”known” structures, while the set of the unknown, and also the unsolved, structures is more likely to contain more not-yet-known structure types than one would expect from extrapolation from the set of known and solved proteins.

Nevertheless, even with this caveat, it is clear that employing similarity analyses plus chemical intuition will quite likely continue to contribute to our ability to predict the structures of not-yet-synthesized compounds, and especially to the structure solution of synthesized but not-yet fully analyzed compounds. But one should not forget the fact that the lack of simple but relatively rigid bonding rules as one finds in covalent compounds like molecules, makes it much harder to feel fully confident that extrapolation from known cases will result in the truly new structure also in the case of bulk solid compounds. Again, the statistics is looking too good in some way, because those systems that are easy to make tend to be the ones that are similar to other chemical compounds whose structures are known. Thus the more complex compounds and the seemingly simple but not-yet-synthesized compounds are statistically more likely to exhibit new structure types never seen so far. These will often even be difficult to post-dict or ”explain” after a synthesis, no matter how much one distorts simple structure motifs such as dense sphere packings or basic coordination polyhedra.

Prediction of appearance of given structure types

Another complementary task where database information can be put to good use is the issue of predicting the appearance of a given structure type $\hat{T}$ (for an example, see e.g. [16]).

Here, the goal is to figure out, which chemical system might support a particular modification that might e.g. fit structurally to a technologically useful layered compound (Fig. 5), giving us more ways to fine-tune the properties of an electronic device. In some ways, this task is the dual to the standard structure prediction problem. Up to now, one usually has only relied on chemical intuition and trial-and-error experience, but computational capabilities have increased to the point where they can assist in this task.

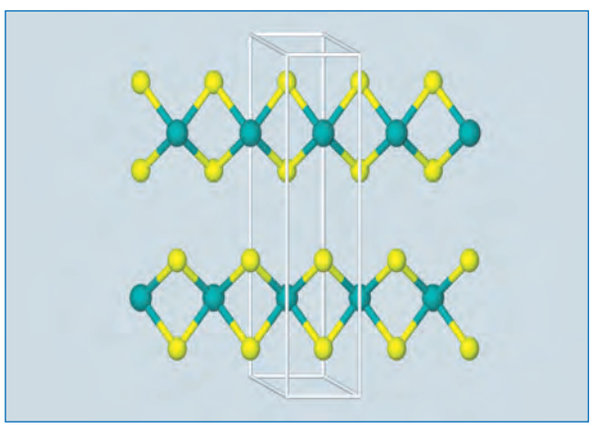


Fig. 5: Example of a layered compound (MoS_2)

Clearly, one can pursue the brute force approach: Minimize the desired structure type for all chemical systems of correct composition type that would be compatible with the $\hat{T}$ structure type. This is usually quite a computational effort, especially if one keeps in mind that one needs to a) also verify that the desired modification is kinetically sufficiently stable, and b) perform at least some global search to check how many competing modifications with lower energy exist in the chemical system. Furthermore, in this brute force approach, one would not really exploit all the pre-knowledge one possesses regarding the structure type, which might be much more extensive than just the overall composition type. Thus, one would like to perform some pre-selection in order to focus on the most likely candidates.

Such a pre-selection should take both structural information about the desired structure type and chemical information about the systems considered into account. For example, certain coordination polyhedra might be present in $\hat{T}$, and thus one would want to first look at those chemical systems which are known to crystallize in structure types that also contain these polyhedra (Fig. 6). Scanning ICSD for compounds of the right composition with modifications where these polyhedra have been observed, would thus be a fruitful data-mining approach. Of course, this would require us to automate the search for such coordination polyhedra through the whole database, but with an efficient use of scripts this can be achieved.

In principle, one might even want to include such coordination-polyhedra information for each structure directly in the database itself.

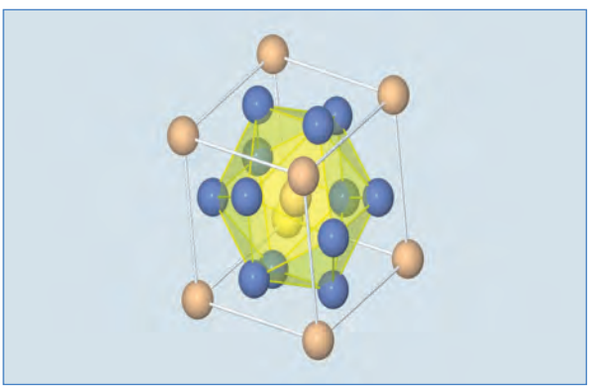


Fig. 6: Example of a structure (Cr_3Si) with a (slightly distorted) icosahedral polyhedron

Similarly, chemical intuition will be of help by providing some heuristic rules such as the radius-ratio rule for ionic compounds – again, information regarding the ionic radii can be taken either from tables compiled in the past or be generated from a systematic perusal of the structures in ICSD. Other kinds of chemical information would be the valence-electron concentration or the number of covalent bonds expected or involved for the structure type $\hat{T}$ – again a systematic pre-scanning of the chemical systems with respect to these attributes could be very helpful for restricting the search range.

Finally, a third way to attack this problem might be called the "reverse approach": In this method, we look at those systems $S^{(k)}$ where a modification with the type $\hat{T}_1$ is known to exist, and find all additional structure types $\hat{T}_{i \neq 1}$ that appear as modifications in these systems. Now, we argue from a possible similarity of the energy landscapes: If in another chemical system $S^{(j \neq k)}$ (where we have not yet observed type $\hat{T}_1$) one of the modifications of structure type $T_{i \neq 1}$ exists, then the likelihood is increased that in addition the modification with structure type $\hat{T}_1$ is also a local minimum, at least compared to a randomly selected system that only exhibits the same composition type. Thus, we select such a system $S^{(j)}$, and compute/relax all structure types $\hat{T}_i$ in these systems, to see whether $\hat{T}_1$ might be stable in such a system, and how it compares with the other competing structure types $T_{i \neq 1}$.

In most cases, we will find that the already known structures in system $S^{(j)}$ are the thermodynamically stable structures, but experience has shown that in many cases the desired structure $\hat{T}_1$ is at least kinetically stable and often quite competitive energy-wise. We note that this approach can e.g. also be used to suggest various high-pressure phases.

Prediction of multinary phases

In many applications in materials science one deals with complex multinary phases. Many of these actually constitute solid-solutions, and while their structure prediction can be addressed by computational means (for a review see e.g. [2]), here we will focus only on the case of ordered crystalline modifications where ICSD can serve as a resource. In general, we can employ ICSD for structure prediction of multinary compounds completely analogously to the case we discussed above in the general structure prediction section. However, there are not really enough such structures available in ICSD to make this approach as reliable as it has been for binary and even ternary compounds (Fig. 7). Thus, one needs to find different ways to tackle the prediction of the structures of multinary compounds.

The most promising approach consists in reversing the "family" analysis from the previous section where we found e.g. possible A/B structure candidates from a contraction of known A'/B'/C' structure types. Instead we now consider chemical systems A'/B'/C' where we expect their structure (based on chemical/structural information, such as e.g. local coordination polyhedra) to be related to some AB-structure type, and assign the atoms of types A', B' and C' to locations of atoms A and B in the known AB-structure. Next, we minimize the energy for all the structure candidates generated in this fashion and check their kinetic and thermodynamic stability. This approach for predicting structures of multinary phases from related binary or ternary phases by substitution of ions/atoms has a long tradition in experimental chemistry. Nevertheless, it is not clear to what extent one has, in the past, really systematically exploited all the data available in ICSD that can be used to predict new compounds in this fashion.

One issue one should keep in mind is that there exists an enormous number of ways to perform such substitutions – even if one focusses on small unit cells –, and thus chemical intuition, for what it is worth in this case, needs to be appealed to, in order to control the combinatorial explosion of possible substitutional atom arrangements. Sometimes, simple arguments based on e.g. minimizing the electrostatic repulsion of the ions will lead to a successful restriction in the number of possible configurations that need to be checked regarding their kinetic and thermodynamic stability. But in many cases the number of candidate structures will still be overwhelmingly large. Furthermore, as mentioned earlier, the chances of encountering a truly new structure type will increase with the complexity of the chemical system. For

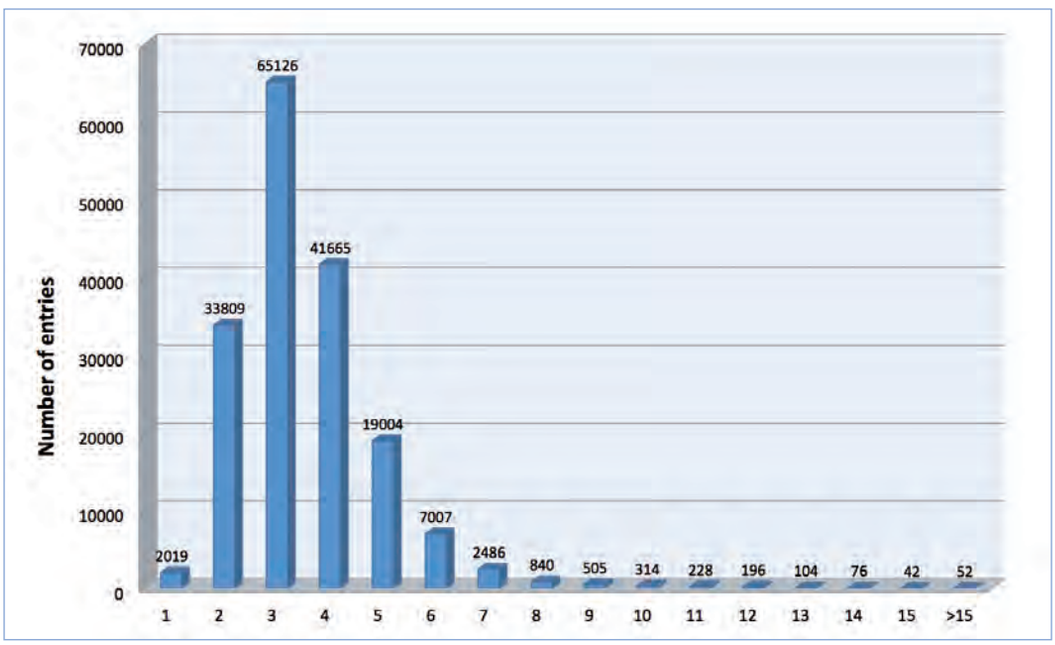


Fig.7: Decline of the number of entries in ICSD for crystal structures with many different elements present

example, due to the slightly different ionic radii of the substituting atoms, it may well be that the real structure of the multinary system will contain a different overall arrangement of coordination polyhedra than those in any of the binary or ternary systems found in ICSD, and thus a deduction of the new type from substituting into one of the known structure types cannot succeed.

Search for missing compounds

So far, we have discussed the use of ICSD to aid in the prediction of the existence and structure of specific chemical compounds, or, alternatively, the search of possible chemical systems where a particular crystalline structure is realized. But we can also use ICSD as a negative-positive screening tool, i.e. we search ICSD for holes, and try to use theory and/or experiment to fill these lacunae. What we mean by a hole are "missing" chemical compounds, i.e. we want to identify those chemical systems where one would expect a crystalline compound to exist but nothing is registered in ICSD. Every practicing chemist knows of such missing compounds, the most famous being perhaps C_3N_4 (Fig. 8), which is one of the first compounds for which a structure was predicted about twenty-five years ago [12].

But there are many other chemical systems, not only multinary ones but even among the binary and ternary compounds, which one might expect to exist, but which have not yet been synthesized. So far, such systems have been identified in a more or less haphazard fashion by chemical intuition and personal preference, but by now the computational tools have become efficient enough for a systematic exploration of their energy landscapes once such promising missing compounds have been identified via analysis of ICSD.

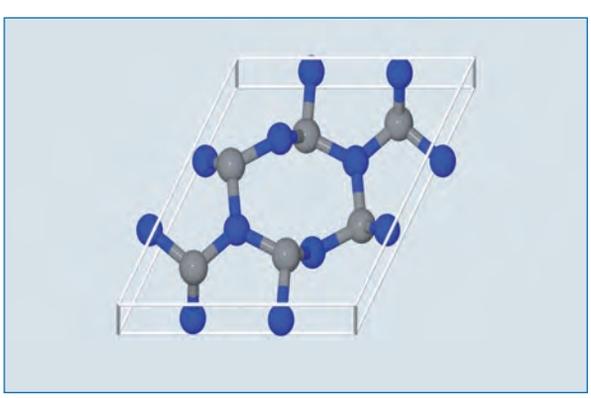


Fig.8: The predicted crystal structure of β-C_3N_4

Of course, one can again follow a brute force route and e.g. for binary systems pick every possible A_nB_m composition observed in ICSD for some chemical system, and perform local optimizations for every chemical system for all A/B-structure types found in ICSD.

In this fashion, we would be able to identify a plethora of kinetically stable binary compounds that would provide synthesis targets for the experimentalists. But while this might just be possible for binary systems, already for ternary systems, it will be extremely expensive com-

putationally, although the Materials Genome project mentioned earlier goes some way in this direction. Furthermore, most of the crystalline modifications predicted will be only marginally kinetically and thermodynamically stable, resulting in a truly momentous heap of essentially dead data.

Preferably, one would want to preselect promising candidates from among the suggested chemical systems, their chemical compositions and for each among the structural modifications, and then perform local ab initio energy minimizations for only these candidates, at least at the beginning. Such a preselection step should occur only after the analysis of the database, of course. In this step, we would employ chemical intuition about the bonding situations in the various chemical systems and their compositions (ideally quantified via a statistical analysis of ICSD and not only by "experience"), and furthermore treat each of the chemical systems again as we did for the standard structure prediction for a given chemical compound of interest: Look in ICSD for modifications that occur in more or less analogous chemical systems, and use these as the first candidates for feasible polymorphs in the system under investigation.

Of course, if we proceed like this, the usual caveats apply: if the thermodynamically stable modification in one of the chemical systems we study does not have an analogue among the already synthesized (or perhaps simulated!) crystalline structures, then we are stumped. Still, a systematic perusal of ICSD for "missing chemical compounds" should be of great help in our attempts to explore and understand the world of crystalline chemical compounds.

Network structures

So far, we have only talked about the whole structure of a chemical compound and the many similarities with other compounds noted in ICSD. But there is another aspect of structural similarity frequently employed when trying to understand crystal structures: the local coordination polyhedra around cations and/or anions formed by their counter ions, or the presence of complex anions and cations that can similarly be represented by a "rigid" polyhedron (Fig. 9). Superficially similar to the local bonding coordination of atoms in molecular chemistry, one often finds the same local building blocks in many compounds that contain the same types of atoms as anions and cations. Frequently, these polyhedra exhibit slight distortions for different compositions, or when different additional cations/anions are present in the structure but are not part of the polyhedra. Once we

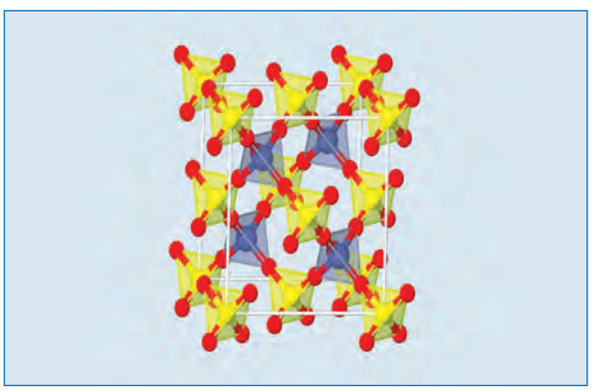


Fig.9: Example of a structure with polyhedra formed by anions around their counter ions

have investigated all structures in ICSD with respect to the presence of such "typical" building blocks, we can then, guided by the analogy to molecules, select all, or only the most common, building blocks to generate structure models for new not-yet-synthesized compounds. There exist several such "coordination graph" based approaches to structure prediction in the literature [17, 18].

While this type of method seems to avoid the global search we talked about above by yielding many apparently feasible structure candidates for a given chemical system, we quickly realize that there are very many ways one can combine these building blocks, even if one uses certain heuristic rules to restrict the exponentially large number of such polyhedra networks. Furthermore, it is not obvious that every candidate structure can be reduced to a network of known building blocks, and the addition of hypothetical coordination polyhedra might be required to globally scan the landscape, increasing the number of candidates enormously. Thus one can quickly get overwhelmed by the number of local minimizations required - after all there is no a priori information available, which of the network models are high or low in energy.

Prediction of structure-property relationships

Finally, we return shortly to the issue of using ICSD to predict materials with specific properties. With the availability of truly gigantic computers and computer farms, one can provide for each entry in the ICSD database not only structural but also many other physical properties by direct computation. By proposing to employ massive computations of this type, the materials genome type projects have set themselves the goal to provide the materials scientist with a large smorgasbord of compounds that might be useful at some point.

But going beyond just accumulating large amounts of data, the grand aim here would be to not only compute various physical properties of interest for all compounds in ICSD, but also to use this information to automatically set up structure-property relationships.

In particular, after computing the properties of many thousands of structures, the database should be large enough that we can go beyond chemical intuition when choosing the structural parameters in these structure-property plots. Instead we would use statistical analyses on a large scale to establish significant relationships among structural, chemical and physical properties of various groups of chemical compounds.

In the end, this type of information will then guide us in the choice of which chemical systems will be the most likely ones that exhibit certain physical and chemical properties. In particular, this will be true even for instances when the compound has not yet been synthesized but has only been predicted to exist as a kinetically stable modification.

Conclusion

In philosophy and pedagogy one sometimes considers phase transitions in our understanding: The abstract transition from quantity to quality, as Hegel and his successors have discussed, or the deeper understanding of a mathematical theorem gained once we have studied enough examples of its applications, or even in the study of chemistry the profound intuition acquired by the experimental chemist due to the many syntheses he or she has performed. Also in computer learning, such transitions are observed during the teaching of a neural network whose parameters are optimized by feeding it a multitude of teaching examples until the network is ready to correctly analyze and classify instances of input never presented before. Similarly, one often speaks in physics of the new paradigm of emergent properties in complex systems where there is no longer a simple direct relation between the e.g. macroscopic and/or long-time features of the system and the microscopic and/or shorttime aspects of the system that are ruled by e.g. the laws of quantum mechanics. Perhaps the most exciting such dynamical transition is suspected to lie behind the emergence of life and consciousness in biological systems that are governed on the microscopic level by physical and chemical laws.

One can propose that this kind of transition will have its counterpart in the field of database applications: Once the amount of information inside the database has grown large enough, then there exists, at least in potentia, the foundation for the emergence of higher order structures among the wealth of information contained in the database. In particular, with regard to chemical databases like ICSD, we can hope and anticipate that the combination of large amounts of data and sophisticated statistical analysis tools will result in deeper insights into the relationships among chemical systems and their structural, chemical and physical properties. The final aim would be that with the help of the database we can answer questions about chemical compounds that are not included in the database or not even known to be possible to exist.

Acknowledgements

My thanks go to W. Marx for suggesting the topic of this essay, and to H. Schier for valuable comments, and to Stephan Rühl for providing the figures for this essay. Most of the ideas presented in this essay are the outcome of discussions with colleagues and friends, too numerous to list, and I would like to thank them for their exciting conversations and contributions.

Note: This essay is based on an article recently published in *Zeitschrift für anorganische und allgemeine Chemie* (DOI: 10.1002/zaac201400374)

References

[1] Woodley, S. M.; Catlow, C. R. A. (2008). Nature Mater. 7, p. 937.
[2] Schön, J. C.; Jansen, M. (2009). Int. J. Mat. Res. 100, p. 135.
[3] Oganov, A. R., ed. (2011). Modern methods of crystal structure prediction (Wiley VCh, Weinheim).
[4] Wang, Y.; Ma, Y. (2014). J. Chem. Phys. 140, p. 040901.
[5] Burzlaff, H.; Malinovsky,Y. (1997). Acta Cryst. A53, p.217.
[6] Hundt, R.; Schön, J. C. ; Jansen, M. (2006). J. Appl. Cryst. 39, p. 6.
[7] Blatov, V. A. (2006). IUCr CompComm Newsletter 7, p. 4.
[8] Allmann, R; Hinek,R. (2007). Acta Cryst. A63, p. 412.
[9] Schön, J. C. (2004). Z. Anorg. Allg. Chem. 630, p. 2354.
[10] Schön, J. C. (1996). Ber. Bunsenges. 100, p. 1388.
[11] Sultania, M.; Schön, J. C.; Fischer, D.; Jansen, M. (2012). Struct. Chem. 23, p. 1121.
[12] Liu, A. Y.; Cohen, M. L. (1990). Phys. Rev. B 41, p. 10727.
[13] Curtarolo, S.; Morgan, D.; Persson, K.; Rodgers, J.; Ceder, G. (2003). Phys. Rev. Lett. 91, p. 135503.
[14] Kalil T.; Wadia, C. (2011). Materials genome initiative for global competitiveness, www.whitehouse.gov (November 25, 2014)
[15] Schön, J. C.; Jansen, M. (2001) Z. Krist. 216, p. 307.
[16] Zagorac, D.; Schön, J. C.; Jansen, M. (2013). Process. Appl. Ceram. 7, p. 111.
[17] Delgado-Friedrichs, O.; Dress, A. W. M.; Huson, D. H.; Klinowski, J.; Mackay, A. L. (1999). Nature 400, p. 644.
[18] Winkler, B.; Pickard, C. J.; Milman, V.; Klee, W. E.; Thimm, G. (1999). Chem. Phys. Lett. 312, p. 536.

A historical outline of the interaction between two disciplines

Crystals and mathematics

Gert-Martin Greuel

People have always been fascinated by crystals and surprised and delighted by their regular geometry, their special symmetry, and their mysterious color refraction. Many natural minerals are crystals. They are parts of stones, but due to the irregular grain boundaries their crystalline nature is often hidden. In cavities, however, minerals can freely grow into recognizable crystals. The crystal shape is also obvious in snowflakes, the growth of which is not limited. Crystals are often found on display in museums or private collections, or as polished gems. Less known is that the largest part (about 98%) of the solid ground is crystalline. This means that crystals are a stable state of the condensed matter.

The term "crystal" is derived from the Greek "krystallos" (meaning "ice") and was used for the first time in connection with rock crystals. People probably thought that crystals formed in extreme cold, which is not true. Most crystalline minerals form at high temperatures and under extreme pressure during the cooling process of magma from the interior of the earth.

Crystals are solids whose components (atoms, ions, molecules) constitute a regular crystal structure – mathematically described as a lattice with a basis. In modern crystallography, a crystal is defined and its geometrical structure described by means of diffraction patterns obtained through exposure to x-rays or other sources of radiation.

This article will briefly outline the history of crystallography and the mathematical-geometrical description of crystal structures.

1. The beginnings of research on crystals

The first known systematic treatise dedicated to minerals is that of natural scientist Theophrastos of Eresos (371-287 B.C.), a student of Plato and Aristotle. It is contained in his publication "On Stones" (see [1]).

In the encyclopedia "Naturalis Historia" by Pliny the Elder (23-79 AD), a comprehensive scientific work consisting of 37 volumes, there is also a treatise on rare stones and minerals (cf. [2]), which was the basis of mineralogical knowledge until the Middle Ages. During the Middle Ages, crystals and stones were also thought to have healing properties – a belief still shared by some New Agers today.

As neither the chemical composition, nor the optical characteristics of crystals could be determined before the beginning of the 19th century, these old studies are only of historical interest to scientists.

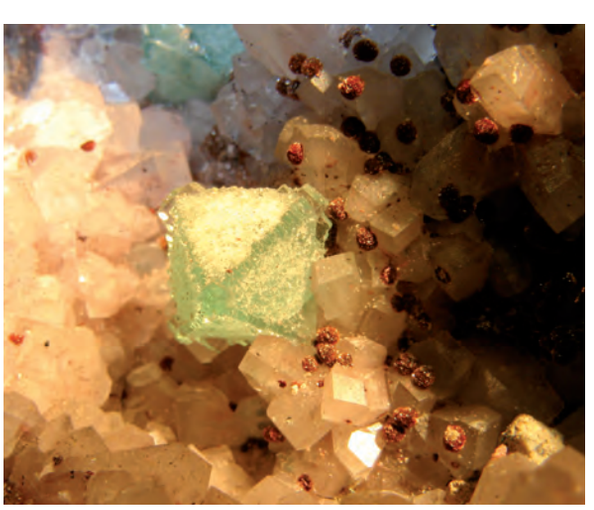


Fig. 1: Fluorite octahedron on calcite

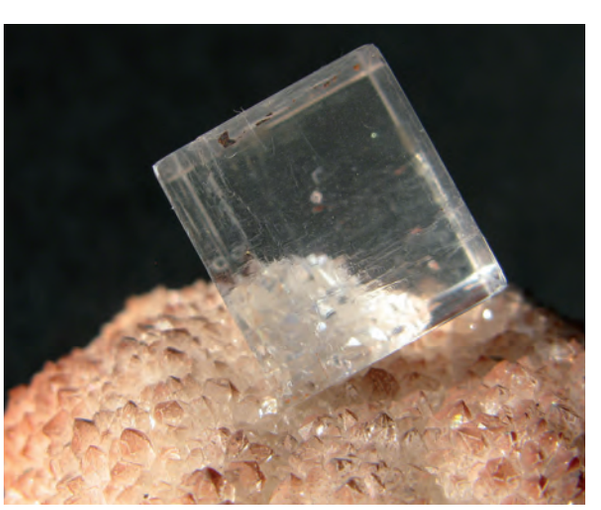


Fig. 2: Fluorite crystal – Grube Clara

2. Crystal shapes and polyhedra

The regular shapes of crystals immediately remind mathematicians of regular, convex polyhedra, especially the Platonic solids with congruent faces and the same number of faces meeting at each vertex. There are exactly five Platonic solids (tetrahedron, hexahedron or cube, octahedron, dodecahedron, and icosahedron, see Fig. 3); their faces are equilateral triangles, squares, or pentagons. The classification of regular polyhedra (claiming that there are only these five regular polyhedra) by the Greek mathematicians, in particular Plato and Euclid, may be considered the first strict approach to classification in mathematics.

In fact, there are minerals whose crystals are almost proper, regular tetrahedra, hexahedra (see Fig. 2), or octahedra (see Fig. 1). Dodecahedra and icosahedra, however, can never be crystal shapes (for reasons of translational symmetry, see chapters 8 and 10), but they can be found in living organisms, e.g., viruses.

Non-regular dodecahedra and icosahedra, on the other hand, exist as crystals. Pyrite, for example, can have the shape of a cubic pentagondodecahedron which by its appearance can be easily mistaken for a regular dodecahedron. Each of its faces is a pentagon, but with four shorter and one longer side. It has cubic symmetry.

An important role beside the Platonic solids play the *Archimedean solids*. They consist of various types of regular polygons and are symmetric in such a way that every vertex can be taken to any other vertex by rotation or reflection. Archimedean solids can be constructed as truncated Platonic solids, i.e. by cutting away a corner of the Platonic solid so that a regular face is formed (e.g., a regular triangle in place of the corner of a cube).

Even more important for crystallography are the *Catalan solids* or *Archimedean duals* that consist of only one type of face. They have, however, different types of vertices, because duality means that the face of a Catalan solid is formed at the vertex of an Archimedean solid. For example, the rhombic dodecahedron is a Catalan solid which is dual to the cuboctahedron. The rhombic dodecahedron is a typical crystal form and is present in garnets. Other frequent crystal structures are prisms, which consist of two parallel, congruent, regular polygons and the lateral surfaces that connect them.

3. Crystals and Greek mathematics

The Greek mathematicians certainly knew the crystal shapes and their symmetries. However, their studies on regular polyhedra were neither motivated by, nor related to crystals. At least, it seems that there is not any proof for this ([3], p. 342). On the contrary – platonic philosophers considered the total symmetry of the regular polyhedra an expression of perfect, everlasting beauty, independent of short-lived, earthly manifestations (cf. [3], p. 340).

In this, Greek philosophy shows some similarities to parts of modern, in particular to pure, mathematics, the research of which is directed towards mathematics-inherent structures and does not serve any particular purpose. It is only committed to its own axioms and logical deductions, without any need or regard for practicability or experience. Still we note that mathematics, also pure mathematics, is – today more than ever – indispensable for understanding scientific processes, and that it has become the driving force of industrial and economic innovation.

Fig. 3: Platonic solids

However, there is most certainly a connection between the crystal shape, which is considered beautiful by man, and the fact that Greek mathematicians were dealing with regular polyhedra. After all, they could have examined other things instead. But according to Plato, the search for everlasting perfection, for the Good, is related to beauty and regularity. *„The good is always beautiful, and the beautiful never lacks proportion“. (Platon, Timaios, 87c, 4-5)*

4. Kepler and the snowflake

It seems that the connection between examinations of the crystal shapes and the ancient studies on mathematical structures of polyhedra was only made in the Renaissance ([3], p. 342). Johannes Kepler (1571-1630) used the harmony of the Platonic solids as a means to describe the harmony of the celestial spheres. However, his attempt to prove that the distances of the planets can be defined by a model of nested, regular polyhedra encased by spheres failed. He therefore abandoned this model and developed for the first time the model of elliptical planetary orbits around the sun that is still in use today.

Kepler's interest for crystals is also reflected in his studies on the symmetry of snowflakes ([4], see Fig. 4). He discovered that the unique geometry and six-fold symmetry of the snowflake is caused by natural forces. The exact physical reason behind this, the fact that the components of matter – atoms and molecules – are always aligned in such a way that they are in a state of minimal energy, was still unknown at his times. This results in beautiful symmetrical structures in snowflakes and other crystals.

Kepler‘s attempt to describe the hexagonal structure of snowflakes as a structure of miniscule particles with minimal distance between them, caused him to study the maximum density of circle and sphere arrangements, published in the same work. He assumed that the most closely packed arrangement of spheres is the cubic close packing (as seen, for example, with oranges stacked in form of a pyramid at a market stall). Only some 400 years later, in 2003, this *Kepler conjecture* was proven by the American mathematician Thomas Hales, partly by complex computer calculations.

5. The beginnings of crystallography

As early as around 1669, Nicolaus Steno (1638-1686) discovered the law of constancy of the interfacial angle when examining quartz. This means that the angles between two identical faces of a crystal are always the same, independent of its size, shape, and the conditions under which it was formed. Steno assumed this law of

Fig. 4: Snowflake

constancy of the interfacial angle for all crystals, and his examinations marked the beginnings of crystallography, i.e. the science dealing with crystals.

The general validity of the law of constancy of the interfacial angle was empirically proven in later years, around 1783, by Jean Baptist Romé de L'Isle through systematic examination and detailed description of about 500 different crystals.

One step further goes the “Traité de minéralogie etc.“, by René-Just Haüy, published in Paris in 1801. Based on the observation of fragments of a piece of calcite shattered on the ground, Haüy described the crystals as consisting of smaller units (which he called “integrating molecules“) that always have the same structure as the crystal itself. That the external shape of crystals can be derived from a periodical alignment of elementary components is a fundamental finding for crystallography. For example, it makes it easy to explain the law of constancy of the interfacial angle. Therefore, Haüy is nowadays also referred to as the “Father of Crystallography”.

6. Modern crystallography

The proof that crystals really consist of regularly aligned units, however, was only furnished by Max von Laue (1879-1960) and his colleagues in 1912. They discovered the *diffraction of x-rays by crystals* and the crystal-dependent, regular pattern of points resulting from it. This discovery and its theoretical substantiation won von Laue the Nobel Prize in Physics in 1914.

Already in 1913, William Henry and his son William Lawrence Bragg had shown that x-rays could even be used to exactly define the position of atoms within a crystal, thus identifying its three-dimensional structure. They both were awarded the Nobel Prize in Physics for this discovery in 1915.

While up to then geometrical optics had been the most popular analytical method in crystallography, *x-ray crystallography* or other sources of radiation are almost exclusively used in modern crystallography to determine diffraction patterns. So, today some hundreds of thousands of such structures, ranging from small inorganic and organic compounds to large biomolecules, have been stored in databases, with their number steadily growing.

The methods of x-ray crystallography were continuously refined, thus making it possible to determine also the structure of important biological molecules (e.g. cholesterine, penicillin, insulin) between 1920 und 1970. This had a great impact on healthcare. The most important discovery in this connection certainly was that of the structure of the DNA by James Watson and Francis Crick by analyzing diffraction experiments, for which they were awarded the Nobel Prize in 1962 in Physiology or Medicine together with Maurice Wilkens.

In recent times, two discoveries are particularly noteworthy: Graphene as the first example of a new class of two-dimensional crystalline materials with unique electronic and mechanical properties (Nobel Prize in Physics awarded to Andre Geim and Konstantin Novoselov in 2010) and the *quasicrystals* (Nobel Prize in Chemistry awarded to Dan Shechtman in 2011). Especially the discovery of the quasicrystals in 1982 was a total surprise for crystallography. These were materials whose diffraction patterns showed a clear pattern of points just like those of ordinary periodic crystals (unlike amorphous substances, which produce blurred diffraction patterns), but with fivefold symmetry – although this should not be possible, due to the periodicity of a crystal (see sections 9 and 10).

It is worth mentioning that nowadays, besides the mathematical crystallography (i.e., the classification of crystals by symmetry), which will be discussed in detail below, there are many more important aspects of crystallography. These include the mineralogical and biological crystallography, but also crystal physics, in particular diffraction physics and crystal growth, with manifold fields of applications, e.g. medicine and materials research.

7. Crystal lattices

The discovery made by Steno, Romé und Haüy, that crystals consist of periodically repeating basic units having the same shape as the crystal itself, leads us to the concept of the crystal lattice. However, for this purpose we are looking at *ideal crystals*. These are infinite and have translational symmetry, i.e., their basic units, or, from today's point of view, the atom positions, are merged into one another by a translational movement within space, so that the whole crystal is formed by the translations of one basic unit.

Of course, ideal crystals do not exist. A real crystal is always finite and has defects, i.e. deviations from regularity. Still it suffices for many purposes, especially to classify crystals, to look at ideal crystals.

Each translation within the three-dimensional Euclidian space is defined by a translation vector that can be described as a combination of the multiples of three (once chosen) independent basic vectors. The total of all translations that transform the basic unit of a crystal into other basic units, and transform the (ideal) crystal into itself, have three linearly independent basic vectors so that each of these translations can be described as a combination of integer multiples of the basic vectors. All translation vectors (or their end points) that transform the crystal into itself form a three-dimensional lattice within the space, the *crystal lattice*.

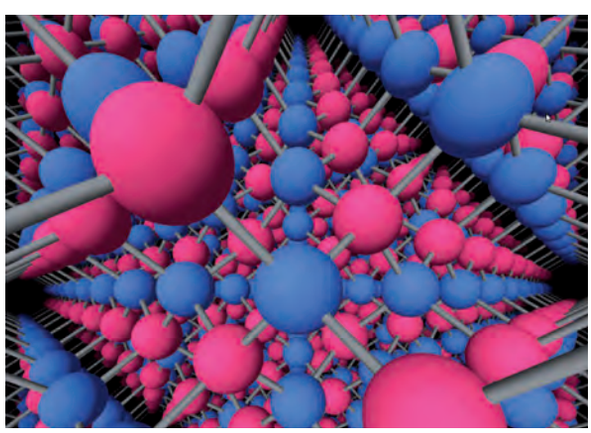


Fig. 5: Lattice of a diamond

The points of this lattice do not represent atoms but only depict the periodicity of the crystal structure. Of course, the crystal lattice depends on the outward shape of the crystal or the basic unit, but since it translates the basic units within the crystal (within the range of atom radii) it is not possible to deduce the crystal lattice from the shape of the crystal. The parallelepiped defined by the three basic vectors is called *unit cell* or *primitive cell*.

The choice of the basis of a lattice is not unambiguous. Therefore one tries to find a basis where the unit cell reflects the shape of the crystal as exactly as possible. Since this is not always possible, but, on the other hand, it is important to be able to easily view the crystal's symmetry, one waives the requirement that the vectors forming the unit cell be a lattice basis and uses three vectors forming larger (non-primitive) unit cells instead. In this case, the translations of the crystal lattice can be

described as a combination of rational multiples of these vectors. These vectors are also called a non-primitive basis of the lattice.

Around the year 1843, Auguste Bravais (1811-1863) classified the different possible crystal lattices by indicating both primitive and non-primitive unit cells. They are named *Bravais lattices* after him. In three dimensions there are exactly 14 Bravais lattices, i.e. exactly 14 translation groups of all possible ideal crystals.

8. Crystallographic groups

Crystals are classified according to their symmetric properties, i.e., we examine the isometries (distance- and angle-preserving mapping) of the three-dimensional Euclidian space that transform the crystal or crystal lattice (which is actually the same) into itself.

The inverse operation of such an isometry and subsequent random operations again result in isometries of the crystal lattice; mathematicians call this a group. The isometry group of a crystal lattice is called *crystallographic group* or *(crystallographic) space group*.

Of course, the translations of the lattice belong to the crystallographic group. They describe the "long-range order" of the crystal. However, there are also isometries of the crystal lattice with (at least) one fixed point, e.g. rotations around an axis, point reflections or reflection in a plane, or combinations thereof. This group describes the symmetry of the basic unit and thus of the crystal itself. It is called the crystal's *point group*. There are exactly 32 of such crystallographic point groups, also called *crystal classes*. Being abstract groups, these groups are of mathematical interest, but for our purposes it is important to note that they are applied to the isometry of the crystals and serve to define the different crystal classes.

Each element of a crystal's point group operates on a crystal with a finite order, i.e. after a finite number of subsequent operations the crystal is back in its initial position (because there is only a finite number of positions of a vertex, an edge or a face of a crystal). For example, the reflection in a plane has order 2, the rotation of a cube around an axis through opposite plane centers (or edge centers or corners) has order four (or two or three).

So there are 32 crystallographic point groups (crystal classes) besides the 14 translation groups (crystal lattices). There are 230 different crystallographic groups in total, i.e., isometry groups of crystal lattices (in three dimensions). They were identified in 1891 by Arthur Moritz Schoenflies and Jewgraf Stepanowitsch Fjodorow independently.

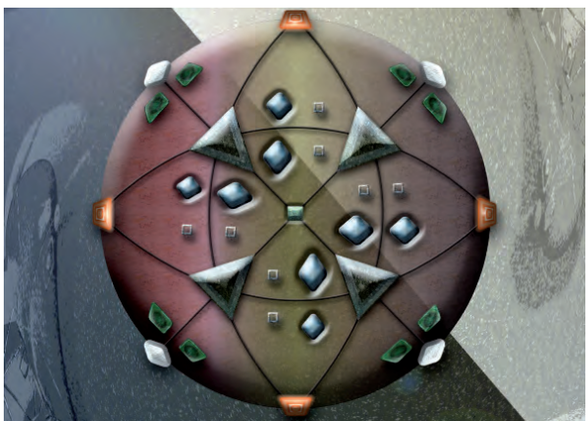


Fig. 6: Example of a point group

Different point groups are combined in one *crystal system*. Today, we differentiate between seven crystal systems: *triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal,* and *cubic*. The crystal systems were introduced by Christian Samuel Weiss (1780-1856) when he translated the textbooks by Haüys. They were based on the analysis of the alignment of particularly striking directions of the crystals, the axes "around which everything was equally distributed". Thus, the crystal systems constitute a symmetry-related classification of crystals by means of crystallographic axes of coordinates. Weiss used the axes to clearly denote for the first time the position of all crystal faces or planes within the lattice by the ratios between their intercepts on the axes, the *Weiss indices*. Nowadays we mainly use the lowest integral common multiples of the reciprocal intercepts, the *Miller indices*.

9. Mathematical groups

We can have a look at the isometries of space that transform any body into itself. These are called the isometry group or *symmetry group* of the body. Most bodies in nature are irregular or asymmetrical. In this case, there is no isometry besides the identity, which transforms the body into itself. A group that only consists of the identity is called *trivial*. Bodies with a trivial isometry group are thus asymmetrical, and bodies with a non-trivial isometry group are called symmetrical. The larger the isometry group of a body, the more symmetrical its appearance.

For example, the sphere as a symbol of a symmetrical body is transformed into itself through any isometry, and its symmetry group contains an uncountably infinite number of elements, i.e. as many as there are real numbers (e.g., every rotation around any angle and any axis and every reflection in any plane through the central point). The isometry group of a crystal is not trivial, but always countable, with the point group itself only containing a

finite number of elements. In any case, we have to note that the colloquial terms “symmetric” and “regular” are mathematically described by the concept of the group.

A (abstract) mathematical *group* consists of a set and an operation which assigns, to any two elements of the set, one element of the set. Several subsequently performed operations can be arbitrarily grouped into operations of two (put into brackets). In addition, each operation has an inverse operation, so that performing subsequent operations results in the identity, which is always one element of each group.

From a historical point of view, the abstract group concept defined in this manner (something taught today right at the beginning of mathematical studies) is still quite young and an abstraction of the concept of the symmetry group which has been in existence much longer. This is a good example of how mathematical concepts are derived from real experience (symmetrical, regular) by abstraction, and how these mathematical concepts are then exclusively defined by abstract axioms, independent of the original observation.

10. The impossibility of the icosahedron

Why can crystals not take the shape of regular icosahedra or dodecahedra? We will learn that this is due to the translational symmetry, which strongly limits the point group of the crystal.

Five regular triangles meet at each vertex of the icosahedron. If we look at a rotation around an axis through two opposite vertices, which transforms the icosahedron into itself and each triangle into the neighboring triangle, this rotation has order 5. The same applies to the dodecahedron if we are looking at the rotation axis through the centers of two opposite faces.

Now we are going to show that a crystal cannot have such a rotation of order 5. Let us look at a crystal rotating around any given axis. This axis has a finite order *n*, and transforms a unit cell of the crystal into a translated one. The case n=2, i.e. a rotation of 180°, is of course possible and we can therefore assume in the following that n is larger than or equal to 3. Now we imagine a plane located perpendicularly to the rotation axis and intersecting a vertex of the unit cell which is not located on the rotation axis. With a full rotation, the vertex describes a regular polygon with n edges within the plane. All vertices of this polygon are vertices of unit cells, i.e., they are transformed into each other by translation of the crystal grid. Each translation shifts the rotation axis so that the rotation around the shifted axis generates an adjacent n-edged polygon within the same plane. This process can be repeated for an infinite number of times until the whole plane is completely (without any gaps or overlapping) tiled with regular, n-edged polygons.

By examining the angles we will now show that this is only possible if n= 3, 4 or 6. Let us assume that r regular polygons with n edges meet at a point within the plane, and that the vicinity of this point is completely tiled with these polygons.

The angle between two adjacent edges of a polygon with n edges is 180° - 360°/n, and r(180° - 360°/n) = 360° applies accordingly. If we divide by 360°, then r/2 - r/n = 1. Multiplying by 2n and rearranging the equation gives n(r-2) = 2r. If we set r-2 = s, then n = 2 + 4/s. As n is a positive integer, s can only take the values of 1, 2 or 4, which means that n can only be 6, 4 or 3.

Therefore, the point group of a crystal can only contain rotations of orders 2, 3, 4 or 6. Thus, crystals cannot have the shape of an icosahedron or of a dodecahedron.

11. Quasicrystals

We have learned that a crystal’s point group can only contain rotations of orders 2, 3, 4 or 6. The diffraction of x-rays in crystals shows a sharp pattern of points that allows to determine the crystal’s point group. For a long time, scientist thought that purely point-shaped diffraction reflexes were only present in crystals and used this assumption to define a crystal. Accordingly, the translational symmetry or the crystal lattice were considered an equivalent for the existence of point-shaped diffraction patterns.

Therefore it came as a shock to crystallography when, in 1982, Dan Shechtman discovered clear, point-shaped diffraction patterns in materials having the same rotational symmetry as the icosahedron, i.e. of order 5. Since these new materials certainly could not be considered crystals in a traditional sense, they were named *quasicrystals*.

The discovery of quasicrystals led to a new crystallographic definition of (traditional) crystals: now, translational symmetry was required as an independent property. Quasicrystals have clear diffraction points, but no periodic translational-symmetric structure. A review article about quasicrystals published on the occasion of the awarding of the Nobel Prize to Shechtman is provided in [8].

If a quasicrystal is cut in an appropriate manner, the slice plane, e.g. in Shechtman’s aluminum-manganese alloy, shows a local five-fold rotational symmetry and a tiling of the plane that is not periodic, but “quasi-periodic”. This

quasi-periodic tiling of the plane had already been discovered in the 1970s by the British mathematician Roger Penrose and is named *Penrose tiling* after him.

The Penrose tiling has many interesting properties. For example, any patch from the tiling (independent of its size) can occur for an infinite number of times, but cannot periodically repeat itself. An easily understandable description of many mathematically interesting characteristics of quasicrystals and Penrose tilings is provided by the mathematicians Baake, Grimm, and Moody in [9].

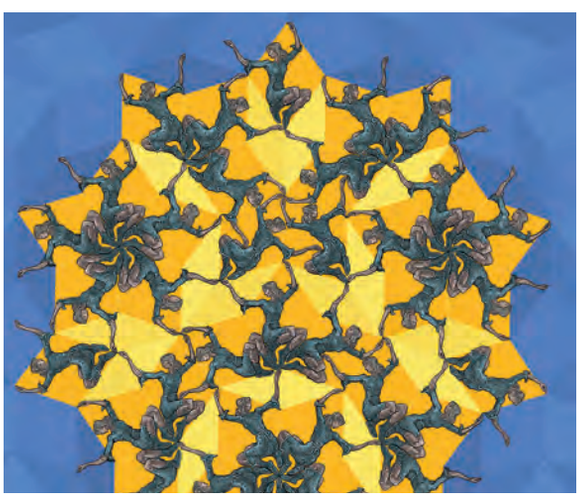


Fig. 7: Quasi-periodic rhombic penrose tiling

Quasicrystals are becoming increasingly important for technical applications because they can give special properties to materials when used as additives.

The discovery of quasicrystals has gained the Penrose tiling a lot of attention. One found out that similar quasi-periodic patterns had already been existing in Islamic ornaments of mosques and palaces from the Middle Ages. How these beautiful but complicated patterns were constructed, was only discovered a short time ago (cf. [10]).

Fig. 8: Jigsaw puzzle of a penrose tiling

Acknowledgements:

I would like to thank Prof. Stephan Klaus at the Mathematical Research Institute Oberwolfach for his useful advice and also Mr. Werner Günter, Chairman of Verein der Freunde von Mineralien und Bergbau Oberwolfach for his advice and the permission to publish his photos of crystals. Many thanks also to Regina Karl from FIZ Karlsruhe for her excellent English translation of the original German text.

The following references are only a part of the sources I used. I think they contain particularly useful further information on the topic.

References

[1] Caley, E. R.; Richards, J. F. C. (1956). Theophrastus: On Stones. Annotated edition with Greek text and English translation, Ohio State University. www.farlang.com/gemstones/theophrastus-on-stones/page_001 (November 25, 2014)

[2] Nies, A. (1884).Zur Mineralogie des Plinius. Mainz. Buchdruckerei von H. Prickarts. Full text on Archive.org. https://archive.org/details/zurmineralogied01niesgoog (November 25, 2014)

[3] Brieskorn, E. Lineare Algebra, III. 13.7, Reguläre Polyeder, Einleitung. Manuskript (soon to be found at imaginary.org). The page numbers refer to the unpublished typoscript.

[4] Kepler, J. (1611). Strena seu de nive sexangula (about the hexagonal snowflake). Keplers Gesammelte Werke, (Hrsg. Caspar, M.; Hammer, F.) Bd. IV (1941). German Translation: Goetz, D. (1987). Ostwalds Klassiker Bd. 273, Leipzig. Text in Latin and English by Hardie, C. (1966). Oxford, Clarendon P.

[5] Articles on Wikipedia: http://de.wikipedia.org/wiki/Kristall (November 25, 2014) http://de.wikipedia.org/wiki/Kristallmorphologie (November 25, 2014)

[6] Kristallographie, na klar! German translation of "Crystallography matters!" by UNESCO 2013. http://dgk-home.de/veroeffentlichungen/kristallographie-na-klar (November 25, 2014)

[7] Kristallographie in Deutschland, Deutsche Gesellschaft für Kristallographie e.V. www.kristall.ethz.ch/DGK/ und www.dgkristall.de (November 25, 2014)

[8] Trageser, G. Kristalle mit unmöglicher Symmetrie, Spektrum.de. www.spektrum.de/alias/nobelpreis-fuer-chemie/kristalle-mit-unmoeglicher-symmetrie/1129068 (November 25, 2014)

[9] Baake, M.; Grimm, U.; Moody, R. V. (2002). Die verborgene Ordnung der Quasikristalle, Spektrum der Wissenschaft 2/2002, p. 64.

[10] Schönfelder, V. Das Geheimnis der Girih, Spektrum.de www.spektrum.de/alias/angewandte-mathematik/das-geheimnis-der-girih/866284, (November 25, 2014).

Images:

Fig. 1: Fluoride octahecron on calcite, Photo: Werner Günter, Oberwolfach

Fig. 2: Fluorite crystal – Grube Clara, Photo: Werner Günter, Oberwolfach

Fig. 3: Platonic solids, Photo: Zumthie at de.wikipedia, source: http://commons.wikimedia.org/wiki/File:Platonische_Koerper_im_Bagno.jpg

Fig. 4: Snowflake, Photo: Jochen Burger, source: www.natur-portrait.de/foto-26022-frau-holles-erster-gruss.html

Fig. 5: Lattice of a diamond; screenshot from the Crystal Flight software, source: http://imaginary.org/program/crystal-flight

Fig. 6: Example of a point group; Author: Jean Constant; source: http://imaginary.org/gallery/jean-constant-crystallographic-points-groups

Fig. 7: Quasi-periodic rhombic penrose tiling; Author: Uli Gaenshirt source: http://imaginary.org/gallery/quasicrystalline-wickerwork

Fig. 8: Jigsaw puzzle of a penrose tiling; Mineralien- und Mathematikmuseum Oberwolfach (MiMa) source: http://mima.museum/mathematik-penrose.php

Practical importance and implementation of abstract mathematical concepts

Three centuries of crystallography through the lens of zbMATH

Olaf Teschke

As already outlined in the article "Crystals and mathematics", crystallography has a natural connection to mathematics: its objects are defined by inherent symmetries which can be directly translated into abstract mathematical formulations. This is quite specific, since most natural phenomena are so complex that only approximate mathematical models exist, which usually do not allow for exhaustive explicit solutions. Contrary to typical cases like problems of celestial mechanics for which closed solutions cannot be expected, chaotic systems of climate or the intricate geometries of elementary particles – the rapid advances in group theory provided already in the 19th century a complete classification of the underlying mathematical structures of crystallography. However, even these seemingly "final" results gave rise to further research which proved to be of great interest even from a purely mathematical viewpoint, but also of subsequent value in material sciences, therefore once more affirming the "Unreasonable effectiveness of mathematics" as postulated by Eugene Wigner).

In this contribution, we will outline the development of some mathematical problems arising from crystallography since the 19th century, as reflected in the database zbMATH produced by FIZ Karlsruhe, which indexes, classifies and reviews all relevant mathematical research publications since 1868. Especially the post-publication reviews will turn out to be very useful in understanding the historical developments: The special tradition of having the essential results and context of a mathematical publication reviewed by an independent expert – which has been introduced in the 19th century to handle the overwhelming complexity and density of papers and prevails in the core areas of mathematics till today – provides a highly valuable resource of information on the progress of the subject.

The notion of crystal in mathematics: direct, indirect, and apparent connections to minerals

Searching in zbMATH for "crystal*" or "kristall*" returns from the corpus of about 3.4 million documents a comparatively small, but still considerable set of ca. 14,000 publications. A first rough overview of the distinct areas where the notion arises is given by the Mathematical Subject Classification: The main areas pertaining to the notion of crystals are statistical mechanics (section 82, 3900 documents), solid mechanics (74, 3139), fluid mechanics (76, 1578) – the last two often in connection with partial differential equations (35, 1388) – and group theory (20, 1203). This coincides basically with the areas where the notion of crystal exists explicitly in the classification, namely, 74E15 (crystalline structure), 74N05 (crystals in solid mechanics), 76A15 (liquid crystals), 82D25 (crystals in statistical mechanics), 82D30 (random media and disordered materials including liquid crystals and spin glasses) and 20H15 (other geometric groups, including crystallographic groups), where the first sections basically arise from mathematical physics, i.e., mathematical descriptions of real mineral structures, and the last one from abstract structure properties of the corresponding symmetries. One should, however, be aware that this is not exhaustive: the mathematical language knows, e.g., also the notion of crystalline cohomology (with associated definitions of crystals and isocrystals, corresponding to the mathematical subject area 14F30) [1], or crystal bases of certain modules on quantum groups (mathematical subject area 17B37) [2] which notions bear only some conceptual resemblance with the classical crystal objects defined by space symmetries.

While a description of the latter mathematical concepts would be beyond the scope of this note, we will try to outline in the following some aspects of crystallography's impact on mathematics structure theory.

Crystallographic groups in the 19th century: structure theories and classification theorems

Reading 19th century scientific texts is often difficult, however, the review of Arthur Moritz Schönflies' "Krystallsysteme und Krystallstructur" [3] (which was, unusually, done by himself, due to the lack of experts), written in the Jahrbuch's spirit of making results broadly accessible, gives today's readers still a very lively impression of his breakthrough in mathematical modelling of crystal symmetries, and its historical context (see Fig. 1). It starts with

Die krystallisirte Materie unterscheidet sich bekanntlich dadurch von den übrigen festen Körpern, dass ihr physikalisches Verhalten längs verschiedener Richtungen im allgemeinen verschieden ist. Nennt man alle Richtungen, in denen sich ein Krystall in jeder Beziehung gleichartig verhält, gleichwertige Richtungen, und denkt man sich von irgend einem Punkte O des Krystalles aus eine Gerade g und alle mit ihr gleichwertigen Geraden g1 ,g2 ,... gezogen, so ist die Lage dieser N Geraden, wie die Erfahrung lehrt, durch bestimmte Symmetrieeigenschaften, wie Symmetrieaxen, Symmetrieebenen u. s. w., ausgezeichnet. Die Symmetrieeigenschaften der N Geraden g, g1, g2,... sind davon unabhängig, wie die Ausgangsrichtung g innerhalb der Krystallmasse angenommen wird; sie erhalten sich überdies während der wechselnden physikalichen Zustände, in denen sich der Krystall befinden kann. Diese Thatsache kann als das "definirende Grundgesetz der krystallisirten Materie" betrachtet werden und wird vom Verfasser zweckmässig als "Symmetriegesetz" bezeichnet.

(As is generally known, crystal matter is distinguished among other solids by its well-defined physical behaviour along distinct directions. If we denote all directions in which the crystal behaves homogeneously as equivalent, and assume that from an arbitrary point o of the crystal a line g and all its equivalent lines g1 ,g2 , ... are drawn, experience shows then that the arrangement of this N lines is characterized by certain properties like axes and planes of symmetry etc. The symmetric properties of these lines are independent on assumptions of the direction of g in the crystal, and are stable in the various physical states of the crystal. This fact can be considered the defining basic law of crystal matter, and is denoted by the author as "Law of Symmetry".)

In the following, Schönflies outlines why only a finite number of possible structures exists, and reviews the known results on the 32 classes of point symmetries and the 14 lattices reflecting the translation invariance. Both finite numbers derive from particular symmetry laws explaining different properties of the matter, and a large part of the review (as well as, naturally, the book) explains how physical structure theories are reflected in their mathematical models. The result of how both translation and point symmetries can be joined into a general symmetry law, and the explicit derivation of the 230 possible space groups, is the core of the work and defines the notion of crystallographic groups since then. Schönflies' results are also a striking proof of the conceptual power of the Erlangen program envisioned by his teacher Felix Klein, which aimed to reflect geometric properties by the classification of its transformation groups.

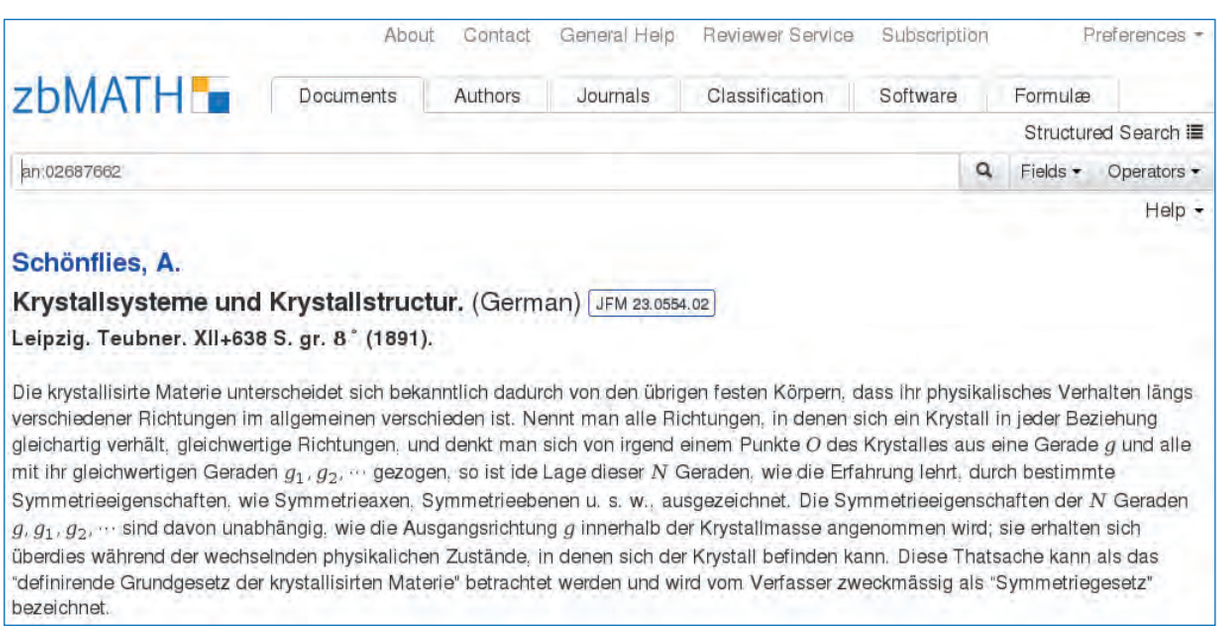


Fig. 1: The beginning of the zbMATH review of Schönflies' "Krystallsysteme und Krystallstructur"

Though from a classification viewpoint, the mathematical problems of crystallography have been "solved" by Schönflies' list, the end of his review points toward what has not been fully understood:

Die Möglichkeit, zwischen die beiden vorstehend genannten Structurtheorien noch eine Reihe anderer Structurauffassungen einzuordnen, beruht darauf, dass sich die Gesamtsymmetrie eines Krystalles im allgemeinen so in zwei Teile zerlegen lässt, dass einer von beiden der Molekel aufgeprägt wird, während sich der andere in der Anordnung, d. h. in der Art des Aufbaues darstellt. Jeder derartigen Zweiteilung der Krystallsymmetrie entspricht eine andere Structurvorstellung; je höher die Symmetrie eines Krystalles ist, um so mannigfaltiger ist daher im allgemeinen die Art, auf welche die Zweiteilung ausgeführt werden kann. Welche Structurauffassung in jedem speciellen Fall am zweckmässigsten zu Grunde gelegt wird, ist eine Frage, deren Entscheidung dem Krystallographen überlassen bleiben muss. Von mathematischer Seite könne es sich, wie der Verfasser bemerkt, nur darum handeln, die Gesamtheit aller überhaupt möglichen Structurauffassungen anzugeben und zu kennzeichnen.

(The possibility of having further intermediate structure theories between this two notions [Bravais lattices and point groups], relies on the fact that the total symmetry of the crystal can be usually split into two parts of which one is stamped onto a molecule while the second reflects its composition. Every such dichotomy of the symmetry corresponds to a different structure theory, and a larger intrinsic symmetry of a crystal allows for an increased possibility of such splits. Which structure theory is most suitable for a special crystal is a question which must be left to the mineralogist; a mathematician can only label the number of the distinct possible structure theories.)

As we will see later, the description of these possibilities has led to extended powerful mathematical theories; at this point, we will only note that the *Jahrbuch* also reflects the obstacles of scientific communication back then. Though Fedorov's "Basic Law of Crystallography" is mentioned as a title [4] it has not been reviewed due to lack of accessibility, and only in 1905 a review of Fedorov's "Lessons on Syngony" [5] appeared, ensuring the dissemination of his results within the mathematical community.

From a historical viewpoint, it is interesting to note that much information on the context of Schönflies' book can also be found through zbMATH, though in a much later review of the 1984 reprint by Johann Jakob Burckhardt [6]:

Wir bemerken, daß der Autor übersehen hatte, daß nicht Hessel (1830/31), sondern Frankenheim die 32 Klassen erstmals herleitete. Wohl zitiert der Autor zwei Arbeiten dieses Forschers und weist darauf hin, daß dieser erstmals die sog. Bravais-Gitter aufstellte. Leider übersah er aber den 1826 in der Zeitschrift ISIS von Oken erschienenen Artikel "Crystallonomische Aufsätze", worin erstmals diese Klassen hergeleitet werden. [...] Zu Recht wird daher sein Werk als einer der beiden Klassiker der mathematischen Kristallographie bezeichnet. Der andere Klassiker wäre zweifelsohne sein Zeitgenosse E. S. von Fedorov, falls dieser seine Entdeckungen in lesbarer Form veröffentlicht hätte. Allein im Ural arbeitend, fehlte ihm die Schulung im geläufigen mathematischen Ausdruck. Dies hat wohl auch Chebychev veranlaßt, die Veröffentlichung der Arbeiten abzulehnen. Glücklicherweise wissen wir aus dem Briefwechsel von Schoenflies und Fedorov, daß nur eine enge Zusammenarbeit der beiden Forscher zur endgültigen und bereinigten Anzahl der 230 Gruppen führte. Allerdings scheint der Autor nicht bemerkt zu haben, daß der Begriff der symmorphen (= arithmetischen) Klasse von Fedorov tiefer in die Struktur der Raumgruppen eingreift als derjenige der 32 Klassen, obschon er diese 73 Systeme auf Seite 598 erwähnt.

Der Nachdruck ist von hervorragender Qualität und es ist zu hoffen, daß er heute, wo das Gebiet neue Aktualität genießt, viele Leser finden wird. Sie werden reich belohnt, insbesondere auch durch die Nennung von etwa 40 früheren Autoren und deren Werken, von Hauy, Delafosse, Frankenheim, und Hessel bis zu Fedorov und Barlow, was bezeugt, daß der Autor ein ausgezeichneter Kenner der Geschichte der Kristallographie war.

(We note that the author has overlooked that not Hessel (1830/31) but Frankenheim first derived the 32 classes. Though the author quotes two of his articles and points out that he first compiled the so-called Bravais lattices, but he overlooked that these classes were derived for the first time in his article "Crystallonomische Aufsätze", published in 1826 in Oken's journal ISIS.

[...] This work is rightfully considered one of the two classics of mathematical crystallography. The other classic would have been his contemporary E. S. von Fedorov if he had been able to publish his results in a readable form. Working isolated in the Ural mountains, he lacked the knowledge of common mathematical notation, which has probably led Chebychev to the rejection of his work. Fortunately, we know from the correspondence of Schoenflies und Fedorov that only a close collaboration led both teachers to the final and simplified number of 230 groups. However, it seems that the author didn't notice that Fedorov's concept of symmorphic (=arithmetic) class catches more of the structure of space groups than the 32 classes, though he mentions these 73 systems at p. 598.

The reprint is of extraordinary quality and it remains to hope that, given the renewed interest in the subject today, it will find many readers. They will be amply rewarded also by the references to 40 earlier authors and their work, from Hauy, Delafosse, Frankenheim, and Hessel to Fedorov and Barlow, which show that the author has been an intimate expert of the history of crystallography.)

While we will explain the "renewed interest" in the next section, we would like to note that some of the historical facts pointed out by Burckhardt have been frequently ignored in earlier and later historical surveys.

Some waves of 20th century mathematical crystallography

Looking at the presence of mathematics in crystallographic journals (like the *Zeitschrift für Kristallographie*, which has been indexed in zbMATH for about 100 years) reveals some peaks of mathematical articles separated by longer gaps of stagnancy (see Fig. 2). After decades of silence following the breakthroughs of Schönflies and Fedorov, the interest of mathematicians in crystallographic questions rose again in the 1920s, followed by a very active period between 1930 und 1940. A second wave of interest took shape at the end of the 70s, and has continued ever since.

What developments in mathematical understanding have been behind this? Two main driving forces in mathematics played a major role also here: the urge to understand the underlying concepts, and the search for generalizations. A first step into both directions was taken by Pólya und Niggli in 1924 by deriving the classification of the two-dimensional analogue of crystal structures by strictly algebraic methods [7]. This has been in line with the complete reshaping of abstract algebra which took place in the 1920s in the course of Emmy Noether's ideas and concepts. Consequently, Johann Jakob Burckhardt (whom we already met as a reviewer above – in total, he contributed about 700 reviews from 1939 until 2004, at the age of 101!), who has been strongly influenced by these developments, undertook the task of adapting the new formal concepts to crystallographic questions. By introducing the notion of arithmetic crystal class (which gave a precise meaning to some concepts used by Fedorov 40 years earlier), he derived the crystallographic groups by purely algebraic means [8] – an approach which he could generalize to higher dimensions. His research on crystallographic questions cumulated in the standard book "Die Bewegungsgruppen der Kristallographie" [9] which, in the words of the reviewer György Hajós, made "for the first time the mathematical theory of crystal classes and space groups accessible also to non-mathematicians" [10].

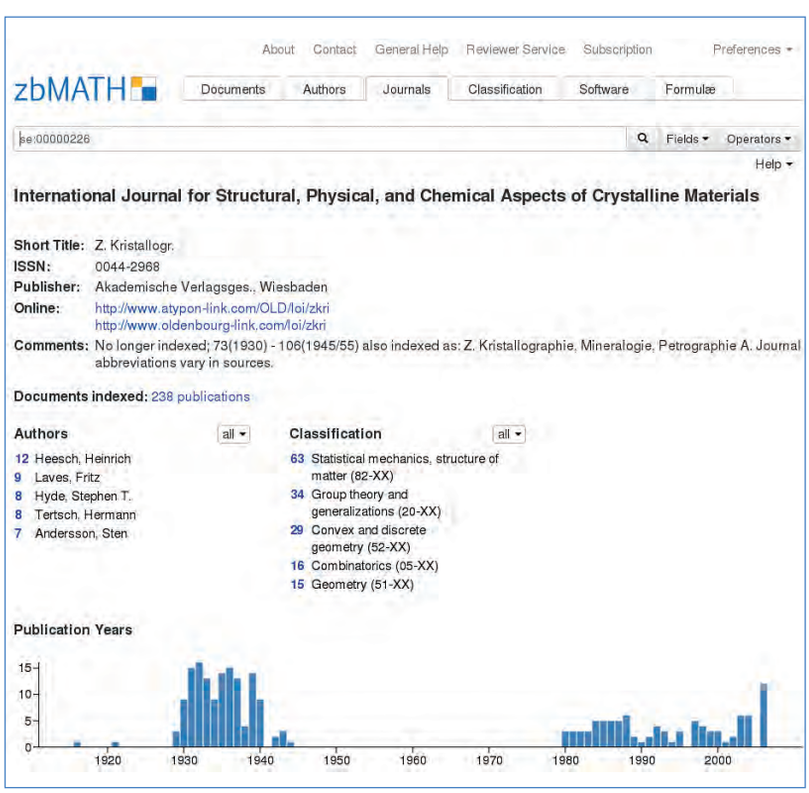


Fig. 2: Exploring mathematical crystallography through a zbMATH journal profile

While these results were again conclusive to a certain point, Burckhardt's algorithmic approach anticipated a recent driving force in the study of the subject, namely, the use of computer algebra in the study of such group structures (which will be considered in the next section).

At about the same time, a second development in the formation of abstract algebra also maintained strong connections to the origins of crystallographic groups. An essential part of its understanding was how the possible structures of point symmetries and translation symmetries fit together. From a mathematical viewpoint, this meant to consider the possible group extensions. That the underlying structure is quite involved is somewhat obvious from the fact that the 32 point groups and the 14 possible lattices make up to precisely 230 different space groups structures. Since similar questions

of how simpler structures (so-called factors) can be extended to a larger one arose also in different areas of 19th century mathematics (for instance, the question of solvability of polynomial equations in radicals led to the consider extensions of Galois groups). The successful discovery of the abstract mechanism guiding those extensions relied heavily on the concepts developed in the 1920s, leading to the notion of group cohomology which was finally discovered by Heinz Hopf and Beno Eckmann at the ETH Zürich (where also Burckhardt worked at that time), Samuel Eilenberg and Saunders Mac Lane at Columbia University, and by Hans Freudenthal at the University of Amsterdam, and which has evolved into an indispensable tool in studying such structures ever since [11].

The third parallel development does not seem to fit into the pattern of increased understanding through abstract concepts: Harold Coxeter revived the interest in the seemingly old-fashioned area of classical geometry. However, a key ingredient of his studies turned out to be the consideration of reflection groups, which are a generalization of symmetry groups arising in crystallography [12].

The Coxeter diagrams he developed for the description of these groups were not only a large conceptual step forward, they also allowed for a surprising application to the seemingly distant problem of classifying simple Lie groups of any dimension. A Lie group, contrary to the groups considered so far, is an infinite group which comes along with the geometric structure of a manifold, while simple just means to be a basic building block in the factor sense described above in the context of group extensions. Lie groups correspond to Lie algebras (a non-associative structure fulfilling a certain identity), and it turned out that they can be described by a finite "root system" of vectors encoding its geometry. The symmetries of the root system give rise to the so-called Weyl group, which is a special Coxeter group whose diagrams fulfils precisely the condition that only 2, 3, 4, and 6 are permitted as its edge labels. This corresponds surprisingly to the space-filling condition we know from crystallography (analogously to the proof of the impossibility of the icosahedron in the article "Crystals and mathematics") and has become known as the "crystallographic restriction theorem". The importance of these structures is underlined by the fact that such Dynkin diagrams (i.e., Coxeter diagrams fulfilling the crystallographic conditions) have appeared since then in many different areas of mathematics, as representation theory, singularity theory, quantum mechanics or category theory.

The appearance of computers and quasi-crystals: some recent developments

With the theoretical foundations laid in the 1930s, a complete classification of the (as proven) finitely many crystallographic structures in higher dimensions would have been possible in principle. However, it turned out that the concrete handling of the vast amount of structures was extremely difficult. Not much progress could be made even in dimension four until the availability of sufficiently powerful computers changed the game. A major milestone was the work of Harold Brown, Rolf Bülow, Joachim Neubüser, Hans Wondratschek and Hans Zassenhaus [13]. While Zassenhaus – also a descendant of the Noether school – had already found an effective algorithm in 1948 [14], it took about 30 years until sufficient computing power was available to implement it effectively. The complete list of 4895 four-dimensional crystallographic groups was one of the first major results in mathematics solely verified by computers (besides the better known case of the four-colouring of graphs in 1976), as pointed out by Burckhardt in his zbMATH review. This was a starting point of computational approaches in group theory which bore many fruits during the next decades. Since questions concerning finite groups often imply the necessity to handle extensive classification lists (many of them only for intermediate purposes) with the inherent danger of error propagation, it is an incredible help to have a formalization available which is not affected by human calculation errors. If one keeps in mind the immense effort which had been necessary for the classification of all finite simple groups (which is distributed into hundreds of research articles between 1955 and 2004, by about hundred authors, assembling to about 25,000 pages), which can hardly be surveyed by a single human anymore, the advantage of computer support becomes obvious. The arguably most powerful tool developed so far is the GAP system [15]. As can be seen from swMATH, the software facet of zbMATH, more than 800 mathematical research publications have employed this software so far [16], including a number of formal verifications of earlier human-generated lists. Included in this large project are also packages with specific crystallographic features like Cryst [17] and CARAT [18]. A typical example of ongoing developments may look like the paper "Computations with almost-crystallographic groups" of Karel Dekimpe and Bettina Eick [19] which explores the opportunities of the GAP system when dealing with almost-crystallographic groups, a natural generalization of crystallographic groups which are finite extensions of finitely generated, torsion-free nilpotent groups (instead of free abelian groups, i.e. lattices, in the case of crystallographic groups).

Fig. 3: The GAP system in swMATH

Another, probably more famous, generalization of crystalline structures are quasicrystals, which have been already introduced in other contributions of this volume. From a mathematical viewpoint, it is fascinating that the underlying abstract structures were derived before their discovery by Shechtman in the "real world" – namely, as non-periodic tilings of the plane first discovered by Robert Berger [20] and the especially aesthetic example of Roger Penrose which needed only two tiles [21] (though this discovery could also be duly attributed to the creators of the Islamic girih ornaments) [22]. Today, quasicrystals are established mathematical objects which even made their way into the Mathematical Subject Classification as a subfield of convex and discrete geometry (52C23 - Quasicrystals, aperiodic tilings). Interesting enough, one of the most powerful tools for describing quasicrystalline symmetries are again root systems and their diagrams (this time, naturally, violating the crystalline restrictions) which were constructed long before the mathematicians became aware of the existence of such objects.

It would be beyond the scope of this contribution to explore the recent ramifications of this subject which includes many different aspects of combinatorics, geometry, algebra or computational mathematics, and we recommend to explore them by browsing through the zbMATH database. We would only like to conclude with a nice example of how the aesthetics of quasicrystalline structures can be used even for the popularization of mathematics at school level: The mathematician John Horton Conway (also famous for, e.g., the "Game of Life") constructed a Penrose type tiling which uses only a pair of simple equilateral triangles as building blocks (which project to the more complicated "kite and dart"-components of Penrose) which fit together as a surface in the three-dimensional space which can be constructed by a simple and beautiful origami folding [23].

So, if you have a sheet of paper at hand (of course, apart from this brochure!) – start to explore the beauty of extended symmetry which has been long hidden both in minerals and in mathematical formalism!

Fig. 4: Quasiperiodic origami surface (source: R. J. Lang and B. Hayes, Paper Pentasia: an aperiodic surface in modular origami. Math. Intell. 35, No. 4, 6174 (2013))

References

[1] introduced by Grothendieck, A. (1968).[Crystals and the De Rham cohomology of schemes, in: Dix Exposes Cohomologie Schemas, Advanced Studies Pure Math. 3, pp. 306-358. (Zbl 0215.37102)] (https://zbmath.org/?q=an:0215.37102v) in the study of the Weyl conjectures)

[2] introduced by Kashiwara, M. (1990).[Bases Crystallines, C. R. Acad. Sci., Paris, Ser. I 311, No. 6, pp. 277-280 (Zbl 0724.17008)] (https://zbmath.org/?q=an:0739.17005)

[3] Leipzig. Teubner. XII+638 S. (1891; JFM 23.0554.02) (https://zbmath.org/?q=an:23.0554.02)

[4] St. Petersb. Miner. Ges. (2) XXXI. pp. 171-190 (1894; JFM 25.0879.02) (https://zbmath.org/?q=an:25.0879.02)

[5] Münch. Abh. 23, pp. 1-88 (1906; JFM 37.0504.01) (https://zbmath.org/?q=an:37.0504.01)

[6] Krystallsysteme und Krystallstructur. Reprint des Originals von 1891. Berlin etc.: Springer Verlag (1984; Zbl 0536.20001) (https://zbmath.org/?q=an:0536.20001)

[7] Pólya, G. Über die Analogie der Kristallsymmetrie in der Ebene, Z. Kristallogr. 60, No. 1, pp. 278-283 (1924; Zbl 06355163) (https://zbmath.org/?q=an:06355163, November 25, 2014); Die Flächensymmetrien homogener Kontinuen. Z. Kristallogr. 60, No. 1, 283-298 (1924; Zbl 06355164); (https://zbmath.org/?q=an:06355164)

[8] Comment. Math. Helv. 2, pp. 91-98 (1930; JFM 56.0867.01) (https://zbmath.org/?q=an:56.0867.01)

[9] Basel: Birkhäuser, 186 p. (1947; Zbl 0037.14903) (https://zbmath.org/?q=an:0037.14903)

[10] Many historical details can be found in Burckhardt's survey "Die Symmetrie der Kristalle. Von René-Just Haüy zur kristallographischen Schule in Zürich. Mit einem Beitrag von Erhard Scholz" [Basel etc.: Birkhäuser Verlag. 195 S. (1988; Zbl 0644.01004) (https://zbmath.org/?q=an:0644.01004).

[11] For a detailed history of the subject, see [Weibel, C. A. An introduction to homological algebra. Cambridge Studies in Advanced Mathematics. 38. Cambridge: Cambridge University Press. (1994; Zbl 0797.18001)] (https://zbmath.org/?q=an:0797.18001)

[12] Coxeter, H. S. M. (1935) J. Lond. Math. Soc. 10, pp. 21-25 (Zbl 0010.34202) (https://zbmath.org/?q=an:0010.34202)

[13] Crystallographic groups of four-dimensional space. Wiley Monographs in Crystallography. 443 p. (1978; Zbl 0381.20002) (https://zbmath.org/?q=an:0381.20002)

[14] Über einen Algorithmus zur Bestimmung der Raumgruppen. Comment. Math. Helv. 21, pp. 117-141 (1948; Zbl 0030.00902), (https://zbmath.org/?q=an:0030.00902); both works were again reviewed by Burckhardt in zbMATH

[15] Groups, Algorithms, Programming (www.gap-system.org). (November 25, 2014)

[16] www.swmath.org/software/320 (November 25, 2014)

[17] www.swmath.org/software/9514 (November 25, 2014)

[18] www.swmath.org/software/4293 (November 25, 2014)

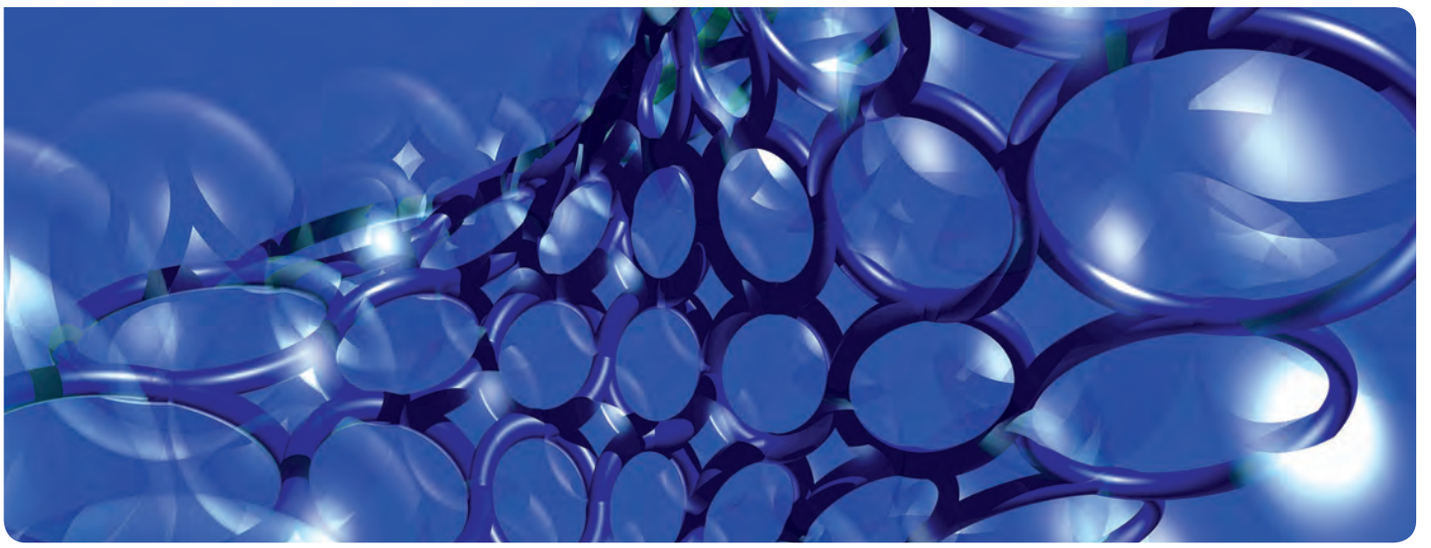
[19] Lond. Math. Soc. Lect. Note Ser. 304, pp. 110-119 (2003; Zbl 1054.20031) (https://zbmath.org/?q=an:1054.20031)

[20] The undecidability of the domino problem. Mem. Am. Math. Soc. 66, 72 p. (1966; Zbl 0199.30802) (https://zbmath.org/?q=an:0199.30802)

[21] Pentaplexity. A class of non-periodic tilings of the plane. Math. Intell. 2, pp. 32-37 (1979; Zbl 0426.52005) (https://zbmath.org/?q=an:0426.52005)

[22] Lu, P. J.; Steinhardt, P. J. (2007).Decagonal and quasi-crystalline tilings in medieval Islamic architecture. Science 315, No. 5815, pp. 1106-1110 (Zbl 1226.00028) (https://zbmath.org/?q=an: 1226.00028)

[23] Lang, R. J.; B. Hayes, B. (2013). Paper Pentasia: an aperiodic surface in modular origami. Math. Intell. 35, No. 4, pp. 61-74 (Zbl 06249373) (https://zbmath.org/?q=an:06249373)



Facts and figures

ICSD contains	Minerals
	Metals and alloys
	Inorganic compounds
Available interfaces for	Internet
	Intranet
	Local installation
Updates	twice per year
	Around 3,000 new entries per update
Content (release 2014.2)	173,473
elements	2,019
binary compounds	33,809
ternary compounds	65,126
quaternary compounds	41,665
quintenary compounds	19,004
Structure types	8,230
# of structures assigned to a structure type	138,503 (~79.8 %)
Bibliography	
Journals	1,671
Authors	65,026
Papers	79,336

FIZ Karlsruhe – Leibniz Institute for Information Infrastructure
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
Germany

Phone +49 7247 808-0
contact@fiz-karlsruhe.de
www.fiz-karlsruhe.de

Berlin
FIZ Karlsruhe's information services in mathematics and informatics are produced at the Department «Mathematics and Computer Science» in Berlin.

FIZ Karlsruhe
Franklinstr. 11
10587 Berlin
Germany

Phone +49 30 399934 0
editor@zblmath.fiz-karlsruhe.de
www.zentralblatt-math.org/zmath/en

Bonn
BINE Informationsdienst – Energy research for application is produced at the subsidiary in Bonn.

FIZ Karlsruhe
Kaiserstraße 185-197
53113 Bonn
Germany

Phone +49 228 92379 0
bine@fiz-karlsruhe.de
www.bine.info

A Focus on Crystallography
Editor: Helmut Müller
Staff of Department Mathematics and Computer Science
Staff of Department Science and Technology
Staff of Department FIZ Products and Services
Staff of Marketing Communications
Translation: Regina Karl
(Preface; "Crystals and Mathematics")

© FIZ Karlsruhe 2015

Member of the Leibniz Association

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure