UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Programa de Pós-graduação em Física

Fernando Luiz Alves Jardim

EXPLORING THE STRUCTURE OF COORDINATION COMPOUNDS AND MOFS USING HIRSHFELD AND ANOMALOUS DISPERSION REFINEMENT

Belo Horizonte
2025

Fernando Luiz Alves Jardim

EXPLORING THE STRUCTURE OF COORDINATION COMPOUNDS AND MOFS USING HIRSHFELD AND ANOMALOUS DISPERSION REFINEMENT

Dissertação apresentada ao Programa de Pós-Graduação em Física do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Mestre em Ciências.

Orientador: Carlos Basílio Pinheiro

Belo Horizonte

2025

Dados Internacionais de Catalogação na Publicação (CIP)

J37e Jardim, Fernando Luiz Alves.
Exploring the structure of coordination compounds and MOFs using hirshfeld and anomalous dispersion refinement / Fernando Luiz Alves Jardim. – 2025.
86 f. : il.

Orientador: Carlos Basílio Pinheiro.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Departamento de Física.
Bibliografia: f. 68-77.

1. Raios X, Difração. 2. Compostos de coordenação. I. Título. II. Pinheiro, Carlos Basílio. III. Universidade Federal de Minas Gerais, Departamento de Física.

CDU – 548.7 (043)

Ficha catalográfica elaborada por Romário Martins Ribeiro – CRB6 3595
Biblioteca Professor Manoel Lopes de Siqueira – Departamento de Física - UFMG

UNIVERSIDADE FEDERAL DE MINAS GERAIS

INSTITUTO DE CIÊNCIAS EXATAS

PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

FOLHA DE APROVAÇÃO

A presente dissertação, intitulada **"Exploring the Structure of Coordination Compounds and MOFs Using Hirshfeld and Anomalous Dispersion Refinement"** de autoria de **FERNANDO LUIZ ALVES JARDIM** submetida à Comissão Examinadora, abaixo-assinada, foi aprovada para obtenção do grau de **MESTRE EM FÍSICA, área de concentração Física,** em 30 de julho de 2025.

Belo Horizonte, 30 de julho de 2025.

Prof. Carlos Basílio Pinheiro
Orientador do estudante
Departamento de Física/UFMG

Prof. Ângelo Malachias de Souza
Departamento de Física/UFMG

Prof. Pedro Henrique Morais Andrade
Departamento de Física/UFMG

Prof. Willian Xerxes Coelho Oliveira
Departamento de Química/UFMG

Prof. Marcos Antônio Ribeiro
Departamento de Química/UFES

Documento assinado eletronicamente por **Marcos Antônio Ribeiro**, **Usuário Externo**, em 31/07/2025, às 14:13, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Carlos Basilio Pinheiro**, **Membro de comissão**, em 31/07/2025, às 15:42, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Pedro Henrique Morais Andrade**, **Professor(a)**, em 31/07/2025, às 16:40, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Angelo Malachias de Souza**, **Membro de comissão**, em 31/07/2025, às 19:06, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site

Acknowledgements

I am grateful to my parents, my father Moyzés and my mother Izabel Cristina, for their unwavering support despite my unconventional career path. This dissertation could never be produced without their support and love. My girlfriend, Cecilia, transformed my perspective on science, life and everything else. I cannot describe in words how important she was during the whole process.

I am also grateful to every one of my friends who took part in this journey with me. Among them, I must highlight Davi Ferreira, who was with me in every single class. We stayed together from the first semester until the end of my master's degree. If anyone understands what my journey was like, it is him.

I am grateful to my supervisor, Carlos Basílio, for the opportunities he gave me. Nearly six years ago, this journey began, and no one could have imagined how it would turn out. It was a fun journey, and it makes me happy that my very specific set of skills proved extremely useful for the laboratory.

I am also grateful to every member of LabCri, current or former, for the incredible time we shared there. Lucas was my "partner in crime" - he was there every day with me and the one who worked most closely with me. Juliana was the first to teach me, both in the lab and in life, being the wonderful person she always was. Ludmila was a great mentor when I joined the lab, and it was truly amazing to work with her, even if only for a brief period. João Neto was both a great friend and a pleasure to have in the lab; everyone knows what excellent company he is. Victor was my best bar companion from the lab, I'm grateful for every drink we shared, every scientific discussion, and every conversation we had. The Laboratory of Crystallography was like a family, and one I was proud to belong to.

I want to acknowledge the founding agencies that made this work possible, CAPES, CNPq and FAPEMIG. I would also like to acknowledge the Federal University of Minas Gerais (UFMG), the department of physics and the Laboratory of Crystallographic (LabCri) for providing infrastructure for the realization of the work.

*"**When do you think people die?** When they are shot through the heart by the bullet of a pistol? No. When they are ravaged by an incurable disease? No. When they drink a soup made from a poisonous mushroom!? No! It's when... they are forgotten.*

— Doctor Hiluluk, One Piece, Eiichiro Oda

Resumo

A caracterização precisa da estrutura cristalina e propriedades químicas, incluindo ocupação metálica e estados de oxidação, é fundamental para o planejamento de compostos de coordenação com propriedades funcionais. Este trabalho avalia a aplicação de técnicas cristalográficas, entre elas Superfície de Hirshfeld, *Hirshfeld Atom Refinement* (HAR), e refinamento de dispersão anomala (ADR), inovadoras a compostos metal-orgânicos. Dois estudos de caso serão apresentados: análise do **estado de oxidação** do centro metálico em compostos de cobalto e juglona e análise da **ocupação metálica** em de redes metalogarnicas (MOF) bimetálicas.

Difração de raios-X de monocristais foi usado para identificar o estado de oxidação do centro metálico em compostos de coordenação de Cobalto e Juglona (5-hydroxy-1,4-naphtalenodiona - jug). A análise da distância entre os átomos de Cobalto e sua primeira esfera de coordenação confirmou o estado de oxidação Co^{II} em ambas estruturas. A análise da superfície de Hirshfeld revelou importantes interações intermoleculares de grande relevância para o empacotamento dos sistemas de cobalto-juglona. O estudo também mostrou que compostos de cobalto e juglona podem ser estáveis com ou sem solventes na rede.

Também foi estudado o uso de refinamento da dispersão anomala (ADR), no contexto de difração de monocristais com fontes de raios-X convencionais, para determinar a ocupação metálica em MOFs bimetálicas. A ocupação dos pares Co/Mg, Zn/Mg e Co/Zn foi determinada através de ADR usando radiação CuKα, obtendo resultados consistentes com espectroscopia de absorção atômica. ADR apresentou desempenho superior ao modelo tradicional de desordem, especialmente para metais com número atômico similar (Co/Zn), devido ao maior contraste em f''.

O sucesso dos métodos utilizando fontes de raios-X convencionais sugere ampla aplicabilidade na caracterização de materiais metal-orgânicos sem a necessidade de uso de luz síncrotron. Esse trabalho atende à crescente necessidade de identificação precisa de posições atômicas na ciência de materiais, especialmente para materiais porosos ao ajudar a estabelecer o ADR como mais uma alternativa acessível para a exploração de substituições metálicas.

Palavras-chave: Difração de raios-X, XRD, Monocristais, ADR, MOF.

Abstract

Accurate determination of crystalline structure and chemical properties, such as metal occupancy and oxidation states, is critical for the design of functional metal-organic compounds. This work evaluates the application of novel crystallographic techniques, such as Hirshfeld surfaces, *Hirshfeld atom refinement* (HAR), and anomalous dispersion refinement (ADR), to metal-organic crystals. We demonstrate these techniques through two case studies: the analysis of metal **oxidation state** in cobalt-juglone compounds and the analysis of metal **occupancy** in bimetallic metal-organic frameworks (MOFs).

Single-crystal X-ray diffraction was employed to identify the oxidation state of the metal center in cobalt-juglone metal-organic compounds. Analysis of the cobalt atom's distance to its first coordination sphere confirmed a Co^{II} oxidation state in both structures. Hirshfeld surface analysis revealed significant intermolecular interactions contributing to structural packing in cobalt-juglone (5-hydroxy-1,4-naphtalenodiona - jug) systems. The study further demonstrates that cobalt-juglone compounds can form stable crystalline structures both with and without solvent incorporation.

Anomalous dispersion refinement (ADR) was explored in the context of single-crystal X-ray diffraction for determining metal occupancy in bimetallic MOF-74 compounds. Using CuKα radiation, we resolved Co/Mg, Zn/Mg, and Co/Zn occupations in single crystals, achieving results consistent with atomic absorption spectroscopy. ADR outperformed traditional disorder modeling, particularly for metals with similar atomic numbers (e.g., Co/Zn), due to enhanced f'' contrast.

Both methods' success with lab-scale X-ray sources suggests broad applicability for characterizing metal-organic frameworks without requiring synchrotron access. These findings address a critical need for precise atomic position analysis in materials science, particularly for porous materials, by supporting the establishment of ADR as a lab-accessible alternative for metal substitution identification.

Keywords: X-ray Diffraction, XRD, Single-Crystal, ADR, MOF.

List of Figures

List of Tables

Contents

1 Crystallography

1.1 Introduction

Crystallography, one of the oldest disciplines in modern physical sciences, originated within the field of mineralogy and became pivotal in the structural determination of crystalline materials during the 19th century. Early foundational contributions include the systematic classification of crystal forms by Samuel Weiss [3] and the introduction of Miller indices by William Miller, which provided a mathematical framework for describing crystal planes and orientations [4].

The advent of modern crystallography commenced in 1912, when Max von Laue hypothesized, during a discussion with Peter Ewald, that X-rays might exhibit wave-like diffraction upon interaction with crystalline lattices. To test this, Laue, Friedrich, and Knipping irradiated a copper sulfate ($CuSO_4$) crystal with X-rays, observing a discrete diffraction pattern on photographic plates. This landmark experiment demonstrated the wave nature of X-rays and established the critical link between diffraction phenomena and atomic periodicity, laying the groundwork for theoretical crystallography [5, 6].

Shortly after, William Lawrence Bragg and William Henry Bragg developed a complementary approach, interpreting diffraction as constructive interference from successive crystallographic planes, behaving as a reflection. Their model, illustrated in Figure 1, posited that for diffraction peaks the path difference between reflected waves must satisfy the condition:

$$2d\sin(\theta) = n\lambda \tag{1.1}$$

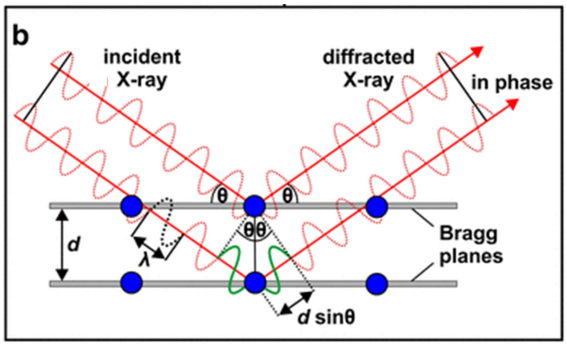


Figure 1 – Representation of Bragg's interpretation of the diffraction effect as resembling reflection. Source : Bijelic & Rompel(2018) [1]

where n is an integer, λ the radiation wavelength, d the interplanar distance, and θ the incident angle. This relation, formalized as Bragg's law (Equation 1.1), became a cornerstone for structure determination and remains fundamental in X-ray crystallography [7].

The 20th century witnessed transformative advancements in X-ray diffraction, revolutionizing the field of crystallography. Key developments included:

- Macromolecular structure determination, enabling the elucidation of complex biological structures [8, 9]
- Automation of data collection, significantly accelerating structural analysis [10].
- Synchrotron radiation sources, providing high-intensity, tunable photon energy X-rays for precision experiments [11].
- Innovative structure solution methods, such as dual-space algorithms and charge-flipping approaches [12, 13].

These breakthroughs propelled crystallography far beyond the foundational visions of Weiss and Miller, extending its reach into materials science, biology, and nanotechnology.

The intellectual trajectory of crystallography is most clearly demonstrated by the transition from morphological to diffraction-based definitions of crystals. Historically, crystals were described by their external morphology, with early theories by Huygens, Bergmann, and Haüy proposing that crystals comprised repeating unit cells, at that moment, a conceptual framework rooted in macroscopic repetition [14]. The advent of X-ray diffraction shifted the paradigm, defining crystals instead by their internal atomic periodicity, characterized by a unit cell, illustrated in Figure 2, containing atoms and the translation vectors that generate the lattice. This definition restricted permissible symmetry operations to rotations of 360°, 180°, 120°, 90°, and 60°, further discussed in Section 1.2.

The definition of crystal up until this point was based on periodicity, meaning it was inherently dependant on translation and even though it sufficed for most natural crystals and applications (e.g., pharmaceuticals, mineralogy), exceptions emerged. For instance, Penrose tilings, a 2D aperiodic structure that completely fills the space, exhibit sharp diffraction patterns despite lacking translational symmetry. Such discoveries prompted the International Union of Crystallography (IUCr) to refine the definition in 1991 [15]:

"*A solid is a crystal if it has essentially a sharp diffraction pattern. The word essentially means that most of the intensity of the diffraction is concentrated in relatively sharp Bragg peaks, besides the always present diffuse scattering.*"

This modern definition accommodates quasi-crystals and other aperiodic ordered solids, reflecting the field's theoretical and experimental progress. It is worthy noting that the unit cell is usually representd following rules determined by the International Tables [16], and it can occur that this standard representation does not cointain the entire molecule, such as in Figure 2.

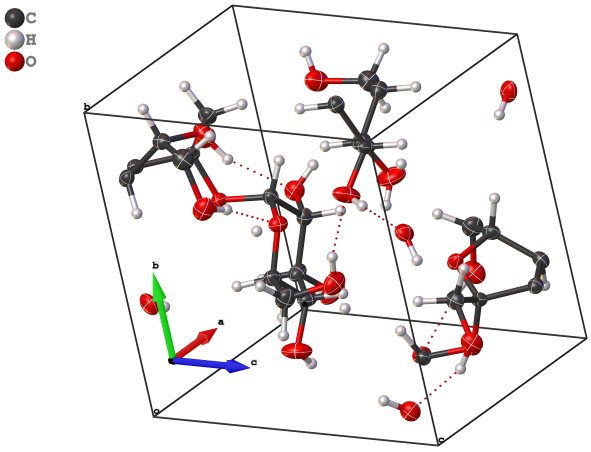


Figure 2 – Graphical representation of the unit cell of a Sucrose crystal, generated in *Olex*2 [2]. Carbon atoms in black, Oxygen atoms in red and Hydrogen atoms in white. Atoms are represented as ellipsoids with 50% probability of finding the atom. Dashed lines shows hydrogen interactions

1.2 Symmetry

A symmetry operation is defined as any transformation that leaves the crystal lattice invariant. The associated geometric entities (points, axes, or planes) about which these operations act are termed symmetry elements. The complete set of symmetry operations for a given crystal structure forms its space group [16].

To satisfy our crystallographic definition, the internal symmetries must be describable through congruent mathematical operations. These transformations establish an isometric relationship between objects, ensuring that: every point in the original object is bijectively mapped to its counterpart; the Euclidean distance between any two points remains invariant under the transformation.

For any two points $\vec{x_1}$ and $\vec{x_2}$ in object A and their images $\vec{x_1}'$ and $\vec{x_2}'$ in object B it must hold true that:

$$|\vec{x_1} - \vec{x_2}| = |\vec{x_1}' - \vec{x_2}'| \tag{1.2}$$

Such congruence operations are classified as either **direct** (or proper), when they preserve chirality or **opposite**, when they do not preserve chirality. Direct operations

are equivalent to a physical rigid-body motion. Opposite operations are, in essence, the composition of an inversion operation with a direct operation. The direct opperations fundamental to crystallography comprise three types:

(1) Translation

A translation is a rigid-body transformation that displaces every point of an object by a fixed vector $\vec{T}$, preserving all interatomic distances and angles. For any point $\vec{r}$ in the crystal lattice, the translation operation yields:

$$\rho(\vec{r}) = \rho(\vec{r} + \vec{T}) \tag{1.3}$$

In periodic crystalline systems, the periodicity condition requires the existence of three linearly independent translation vectors that respect the symmetry condition in equation 1.3.

(2) Rotation

A rotation is a symmetry operation defined by the transformation of spatial coordinates about a fixed axis $\hat{n}$ through an angle θ. A rotation can be described as:

$$\vec{\rho}' = R_n(\theta)\vec{\rho} \tag{1.4}$$

where $R_n(\theta)$ is the rotation operator. Notably, the rotation axis do not need to coincide with any crystallographic direction. As established in Section 1.1, the requirement of translational periodicity in crystalline materials imposes fundamental restrictions on permissible rotational symmetries.

To derive the crystallographic restriction theorem, we begin by considering that successive applications of a rotation must return the system to its original configuration after a finite number of operations. This periodic condition requires the rotation angle θ to satisfy $\theta = 2\pi/n$. Here n is a natural number, and we will call a rotation with such an angle a rotation of order n.

Consider a crystal lattice possessing an $n-$fold rotation axis aligned with the $z-$axis, without loss of generality. Applying rotations $R_z(\theta)$ and $R_z(-\theta)$ to the translation vector T and its opposite $-T$, respectively, yields vectors $\vec{T}'$ and $\vec{T}''$, as shown in Figure 3. The initial points T and $-T$ were translationally equivalent points. If $R_z(\theta)$ is a symmetry operation of the crystal, points T' and T'' must also be translationally equivalent points. This means the vector difference $\vec{T}' - \vec{T}''$ must be an integer multiple m of the lattice vector $\vec{T}$, which can be written as $\vec{T}' - \vec{T}'' = mT$. Since all vectors share the same magnitude, we can express this condition in scalar form:

$$\cos\left(\frac{2\pi}{n}\right) = \frac{m}{2} \tag{1.5}$$

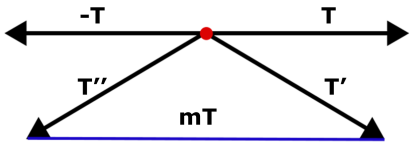


Figure 3 – Scheme described for determining possible rotations.

Equation 1.5 admits only five positive integer solutions solutions for n: $n = 1, 2, 3, 4$ and 6. These values constitute the complete set of crystallographically allowed rotation axes, as required by the compatibility between discrete translational symmetry and point group operations in three-dimensional Euclidean space.

In the Hermann-Mauguin notation [16], these rotational symmetries are denoted by the corresponding integers, where:

- 1: Identity operation (360°)
- 2: Twofold rotation (180°)
- 3: Threefold rotation (120°)
- 4: Fourfold rotation (90°)
- 6: Sixfold rotation (60°)

(3) Rototranslation

A rototranslation (or screw operation) is a crystallographic symmetry operation that combines a rotation of order n about an axis with a simultaneous translation t along the same axis. Mathematically, this operation can be expressed as:

$$\vec{\rho}' = R_n(\theta)\vec{\rho} + t\hat{n} \tag{1.6}$$

Let us examine the constraints imposed on screw operations by the translational symmetry of crystals. For a screw axis of order n, $n-$fold application of the operation must restore the system to its original configuration. If t represents the fractional translation associated with each individual screw operation, the cumulative translation after n applications must satisfy:

$$n\vec{t} = k\vec{T}$$

where k is an integer number. This constraint necessitates that the elementary screw translation must be a rational fraction of the lattice period:

$$\vec{t} = \frac{k}{n}\vec{T} \tag{1.7}$$

Consequently, any screw operation is uniquely characterized by two integers, the rotational order n and the translation numerator k. The Hermann-Mauguin representation of a rototranslation is n_k. Also notable is that, by our definition, $k < n$

When two objects are related through opposite isometries (also called improper congruent operations), they form an enantiomorphic pair. These symmetry operations, which involve a change of handedness, can be classified into three fundamental types:

(1) Inversion

Inversion is a fundamental symmetry operation that transforms any point to its antipodal position relative to a fixed center of symmetry (the inversion center). This operation generates a centrosymmetric relationship between equivalent points in the crystal structure. In the Hermann-Mauguin notation, inversion symmetry is denoted by $\overline{1}$ (read as one-bar). Inversion operation can be mathematically described by:

$$\vec{\rho}' = J\vec{\rho} \tag{1.8}$$

where J is the inversion operator in relation to a point. If the point is located at the origin of the reference system $J = -I$, where I is the identity operator.

(2) Rotoinversion

A rotoinversion of order n is a composite symmetry operation combining a proper rotation of order n about an axis with subsequent inversion through a point on that axis. Rotoinversion can be mathematically described by:

$$\vec{\rho}' = J \circ R_n(\theta)\vec{\rho} \tag{1.9}$$

where J is the inversion operator and $R_n(\theta)$ is the rotation operator. In Hermann-Mauguin notation, rotoinversions are distinguished from proper rotations by an overbar ($\overline{n}$).

The composite symmetry operation $\overline{2}$ is mathematically equivalent to a mirror reflection (m) through a plane perpendicular to that axis. Similarly, the six-fold rotoinversion ($\overline{6}$) operation decomposes into the product of a three-fold rotation (3) and a mirror reflection (m) normal to the rotation axis. This operation is represented in Herman-Maguin notation as $\frac{3}{m}$.

These relationships demonstrate how certain rotoinversion operations can be equivalently represented using other fundamental symmetry elements in crystallography [16].

(3) Glide Plane

A glide plane is a crystallographic symmetry operation that combines a reflection with respect to a mirror plane and a translation t parallel to this plane. Mathematically, this operation can be described as:

$$\vec{\rho} = M\vec{\rho} + \vec{t} \tag{1.10}$$

where M is the reflection operator. The periodicity condition requires that an even number of consecutive glide operations must restore the original configuration, which imposes the constraint $2m\vec{t} = \vec{T}$. A typical glide operator has $m = 1$, meaning that $\vec{t} = \frac{1}{2}\vec{T}$. Glide operators with $m = 2$ are also possible and are called "diamond glides". Glide operations are represented in Hermann-Mauguin notation by a lowercase letter (a,b, or c for glide planes in the respective crystallographic direction; n for a diagonal glide plane and d for a diamond glide plane).

1.3 Unit Cell

The existence of translational symmetry allows us to define a fundamental region of space that completely describes the crystal when replicated by integer multiples of the appropriate translation vectors. This region is called the unit cell, which is of paramount importance in X-ray diffraction (XRD) studies since a complete description of its electron density is sufficient to describing the entire crystal's electron density..

The unit cell is the volume enclosed by the volume generated by three non coplanar translation vectors, denoted $\vec{a}, \vec{b},$ and $\vec{c}$ (or alternatively $\vec{a}_1, \vec{a}_2,$ and $\vec{a}_3$). Those vectors are called "basis vectors" and the volume of the unit cell is the mixed product of those vectors. The position, $\vec{Q}$, of any point in the unit cell can be written as:

$$\vec{Q} = x\vec{a} + y\vec{b} + z\vec{c} = \sum_{n=1}^{3} x_n a_n \tag{1.11}$$

Any atom outside the unit cell is equivalent to some atom within the unit cell via translation, which means that $x_n \in [0, 1]$. x_n is also called the fractional position of the atom. While the unit cell origin is arbitrary in principle, Appendix A.2 demonstrates how strategic origin selection can facilitate structure solution via direct methods. The unit cell, by our definition, is intrinsically linked to translation. This means that inside the unit cell there can be atoms related by symmetry operations, depending on the space group of the crystal. Because of that, one can define the region that encloses all non-symmetry-related atoms; this region is called the **assymetric unit**.

Although the basis vectors do not form an orthonormal set, they generate a vector space with an associated dual space satisfying:

$$\vec{a}_i \cdot \vec{a}_j{}^* = \delta_{ij} \tag{1.12}$$

where δ_{ij} is the Kronecker delta function. The dual space defined by equation 1.12 is the so-called Reciprocal space. Simple arithmetic leads us to find that:

$$\vec{a}_n{}^* \equiv \frac{\vec{a}_i \times \vec{a}_j}{V} \tag{1.13}$$

where $a_n{}^*$ is positively defined and $i \neq j$. Reciprocal space is fundamental to XRD analysis because diffraction patterns represent Fourier transforms of the electron density, as discussed in Section 2.1. Notably, given that they can be written as cross product between direct space vectors, one can interpret vectors in reciprocal space as vectors normal to planes in direct space. The miller indices are integer multiples of the basis vector of the reciprocal space.

There is more than one type of unit cell. Unit cells that contain only one point, located at (0,0,0), are called primitive cells and in Hermann-Mauguin notation are designed with the letter P. The other possibilities are:

- I for body centered, when an additional point is present in the middle of the cell
- F for Face centered, when there is an additional point is present in the middle of every face
- A,B or C for centering only in the respective face

These configurations represent the Bravais lattices (illustrated in Figure 4), which are crucial for space group classification (Section 1.4). Table 1 shows the compatibility between crystal systems and centering types. Notably:

1. Triclinic systems only permit primitive lattices
2. Monoclinic systems can only be C-centered (base-centered)
3. Orthorhombic systems allow all centering types, including different centerings for A,B and C (totaling six possibilities)

Rhombohedral lattices, historically treated as hexagonal centering, now constitute a separate category but maintain dual representation in *The International Tables* [16]. A general representation for hexagonal lattices with rhombohedral axis, although possible, is not maintained in *The International Tables* [16].

Table 1 – *Compatibility between crystal systems and Bravais lattice types. A/B/C denote single-face centering variants*

	P	I	F	A/B/C
Triclinic	✓	-	-	-
Monoclinic	✓	-	-	✓
Orthorhombic	✓	✓	✓	✓
Tetragonal	✓	✓	-	-
Trigonal	✓	-	-	-
Hexagonal	✓	-	-	-
Cubic	✓	✓	✓	-

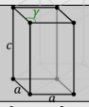
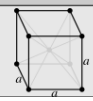
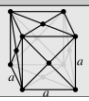
Crystal Family	Lattice System	Point Group	Lattice Type			
			Primitive (P)	Base Centered(ABC)	Body-Centered(I)	Face-Centered(F)
Triclinic		-1				
Monoclinic		2/m				
Orthorhombic		mmm				
Tetragonal		4/mmm				
Hexagonal	Rhombohedral	-3m				
	Hexagonal	6/mmm				
Cubic		m-3m				

Figure 4 – Graphical representation of each Bravais Lattice. Source: Mellish, B. Bravais Lattice, 2007. Available on: Bravais Lattice at Wikipedia Acessed on: 18/06/2025

1.4 Point and Space Groups

Point and space groups are, as the name suggests, mathematical groups. This means they are sets equipped with a binary operation (here denoted "$-\cdot-$") that always generates members of the set. The binary operation for point and space groups is composition, while the symmetry operations discussed in Section 1.2 constitute the group elements. Mathematical groups follow three axioms:

(1) Identity

There must exist an element 1 such that for every element a of the group, has the property $1 \cdot a = a$. This element is called identity.

(2) Associativity

For every a, b, c in the group $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(3) Inverse Element

For every element a in the group there exists an element b such that $a \cdot b = 1$ and $b \cdot a = 1$. This b is called the inverse of a and is commonly denoted a^{-1}.

In real crystals, multiple symmetry operations may coexist, implying the simultaneous existence of multiple symmetry elements. Initially, we will consider only symmetry operations that do not involve translation, these constitute the point groups. This nomenclature arises because these groups leave at least one point in space invariant under all their symmetry operations. The collection of crystals sharing the same point group is called a crystal class, and the crystal class symbol is identical to its corresponding point group symbol.

The simplest possible combination involves a single symmetry operation, in addition to the identity operation (1) in all other directions. We consider five direct rotational operations (1, 2, 3, 4, 6) and their corresponding opposite operations formed by composition with inversion ($\overline{1}$, $\overline{2}$, $\overline{3}$, $\overline{4}$, $\overline{6}$). Furthermore, axes may simultaneously exhibit both direct and opposite operations, yielding the following combined symmetry elements ($\frac{1}{\overline{1}} = \overline{1}, \frac{2}{\overline{2}} = \frac{2}{m}, \frac{3}{\overline{3}} = \overline{3}, \frac{4}{\overline{4}} = \frac{4}{m}, \frac{6}{\overline{6}} = \frac{6}{m}$).

Since the $\overline{1}$ and $\overline{3}$ groups were previously described, this formulation initially yields 13 distinct point groups. We may now proceed to examine combinations of axes intersecting at a common point.

We begin this analysis by examining combinations of proper axes (i.e. axes associated with direct operations). While this problem was originally solved by Leonhard Euler, we follow the approach presented by Carmelo Giacovazzo [17].

Consider the spherical triangle shown in Figure 5, where: through A and B passes symmetry axes of order m and n, respectively; angles α and β satisfy $2\alpha = \frac{2\pi}{m}$ and $2\beta = \frac{2\pi}{n}$.

According to Sylvester's theorem, a rotation of 2α about axis A followed by a rotation of 2β about axis B is equivalent to an anticlockwise rotation of 2γ about axis C, where $\gamma = \frac{\pi}{p}$ for some natural number p.

Since the sum of internal angles in any spherical triangle exceeds π, we obtain the fundamental inequality:

$$\frac{1}{m} + \frac{1}{n} + \frac{1}{p} > 1 \tag{1.14}$$

A trivial solution to this problem exists using only integer values (l, 2, 2) where $l > 1$. While this solution holds for all integer values of l, only four cases correspond to crystallographic point groups: 222, 322, 422, and 622. The remaining solutions may describe either molecular symmetries or approximate local site symmetries in crystalline

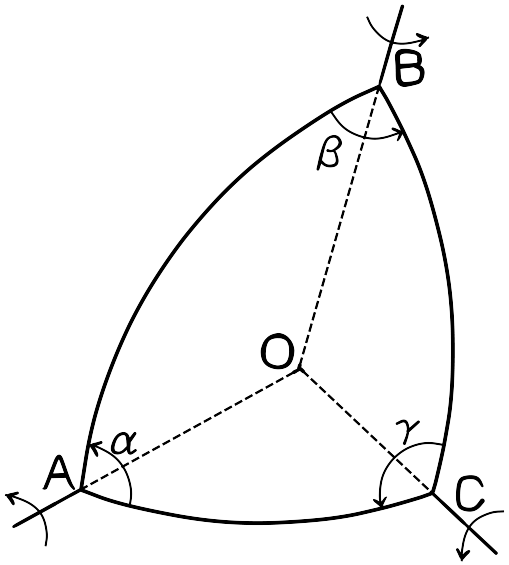


Figure 5 – A triangular square built to respect A and B symmetry orders. By: Ingrid Mitraud

environments.

For non-trivial solutions, we first consider the case of 233, with corresponding angles $\alpha = \frac{\pi}{2}$, $\beta = \frac{\pi}{3}$, and $\gamma = \frac{\pi}{3}$. This point group describes the symmetry of a tetrahedron. The solution 432, characterized by angles $\alpha = \frac{\pi}{2}$, $\beta = \frac{\pi}{3}$, and $\gamma = \frac{\pi}{4}$, represents the symmetry of both a cube and an octahedron. Finally, while 532 constitutes a valid solution, it does not correspond to any crystallographic point group and will therefore not be considered further. In total, this analysis yields six distinct point groups.

Let us now examine point groups containing both proper and improper axes. When combining proper and improper axes, we observe that the transformation results of a proper axis can only be related to those of an improper axis through another improper axis. This constraint implies that an improper axis cannot exist in isolation and that three improper axes cannot coexist without proper axes.

We therefore consider combinations of the form PII and IPI (where P represents a proper axis and I an improper axis). Note that IIP groups are equivalent to PII groups, even for the 432 solution.

For PII combinations, starting from our six original groups, we obtain: 2mm, 3mm, 4mm, 6mm, $2\overline{3}\overline{3}$, $4\overline{3}\overline{2}$. For IPI combinations, we find: $\overline{3}$2m, $\overline{4}$2m, $\overline{6}$2m and $\overline{4}$3m. This analysis adds ten additional point groups to our collection.

Regarding symmetry combinations, the remaining case involves the simultaneous presence of proper and improper axes. Beginning with our six PPP solutions, we obtain the following groups: $\frac{2}{m}\frac{2}{m}\frac{2}{m}$; $\frac{6}{m}\frac{2}{m}\frac{2}{m}$; $\overline{3}\frac{2}{m}\frac{2}{m}$; $\frac{2}{m}\overline{3}\overline{3}$; $\frac{4}{m}\frac{2}{m}\frac{2}{m}$ and $\frac{4}{m}\overline{3}\frac{2}{m}$. Among these, only three represent unique groups: $\frac{2}{m}\frac{2}{m}\frac{2}{m}$; $\frac{6}{m}\frac{2}{m}\frac{2}{m}$; $\frac{4}{m}\frac{2}{m}\frac{2}{m}$.

When combined with the 29 groups previously identified, these additions complete the set of 32 crystallographic point groups. The relationships between groups containing improper axes and those comprising only proper axes are presented in Table 2.

Table 2 – Table showing the relations of groups with improper (I) and mixed(M) axes to those with only proper(P) axis

PPP	PII	IPI	MMM
222	2mm	-	$\frac{2}{m}\frac{2}{m}\frac{2}{m}$
322	3mm	$\overline{3}$2m	-
422	4mm	$\overline{4}$2m	$\frac{4}{m}\frac{2}{m}\frac{2}{m}$
622	6mm	$\overline{6}$2m	$\frac{6}{m}\frac{2}{m}\frac{2}{m}$
233	$2\overline{3}\overline{3}$	-	-
432	$4\overline{3}m$	$\overline{4}$3m	-
6	6	4	3

Point groups alone are insufficient for describing crystalline symmetries; space groups must be introduced to incorporate translational symmetry. By systematically applying the Bravais lattices (as detailed in Table 1) to the point groups described previously, one obtains 73 symmorphic space groups.

Additional translation-specific symmetry operations can then be introduced: glide planes (replacing mirrors along certain axes); Rototranslations (replacing pure rotations along specific axes). These modifications yield 54 hemisymmorphic space groups.

The remaining 103 groups constitute the asymmorphic space groups, for which no general derivation algorithm is currently known. Complete descriptions of all 230 space groups are provided in *The International Tables for Crystallography*, Volume A [16].

2 X-ray Diffraction

Scattering phenomena can be classified into two distinct types:

- **Elastic scattering**, characterized by **identical frequencies**—and thus energy—between the incident and scattered waves
- **Inelastic scattering**, characterized by **different frequencies**—and thus energy—between the incident and scattered waves

Diffraction in the context of this work is, primarily, an elastic scattering phenomenon.

2.1 Elastic Scattering

In elastic (Thomson) scattering, the scattered wave maintains the same frequency as the incident wave. For X-ray diffraction analysis, where our primary interest lies in identifying diffraction pattern maxima rather than wave propagation, we may disregard the temporal component of the wave equation.

Considering that parallel waves with identical phase are diffracted by O and O', as illustrated in Figure 6, the phase difference ($\Delta\phi$) between waves scattered by points O and O' is given by:

$$\Delta\phi = \frac{2\pi}{\lambda}(\vec{s} - \vec{s}_0) \cdot \vec{r} = 2\pi\vec{r}^{\,*} \cdot \vec{r} \tag{2.1}$$

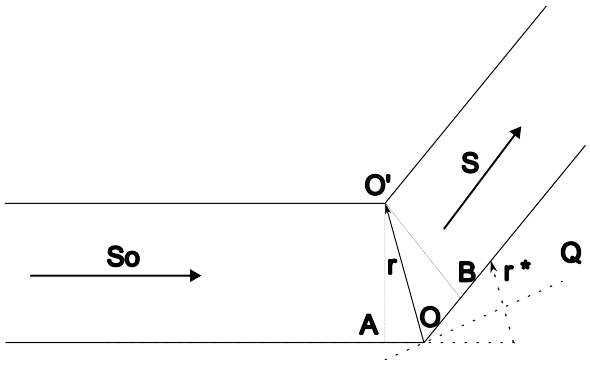


Figure 6 – Point scatterers at positions O and O', with $\vec{s}$ and $\vec{s}_0$ representing unit vectors.

Here, $\vec{s}_0$ denotes the incident wave vector, $\vec{s}$ represents the scattered wave vector, and $\vec{r}$ connects the two scatterers. If we define the scattered wave amplitude from O as A_O and from O' as $A_{O'}$, while setting the phase of the wave scattered by point O to zero, the scattered wave from O' becomes:

$$F_{O'}(\vec{r}^{\,*}) = A_{O'} \exp(2\pi i \vec{r}^{\,*} \cdot \vec{r}_j) \tag{2.2}$$

The simplified Thomson formula ($A_j/A_n = f_j^2/f_n^2$) relates the atomic scattering factor f_j of the jth atom to the diffracted wave intensity.

The **atomic scattering factor** f_j can be estimated by considering X-ray scattering from an electronic distribution $\rho_e = |\psi|^2$, which satisfies Schrödinger's equation. Under the first-order approximation that the scattered wave amplitude is proportional to the electron density, we obtain:

$$F_{O'}(\vec{r}^{\,*}) = \int \rho(\vec{r}) \exp(2\pi i \vec{r}^{\,*} \cdot \vec{r}) d\vec{r} \tag{2.3}$$

It is important to note that this integral is non-trivial to solve. In a first approximation, $\rho(\vec{r})$ can be treated as spherically symmetric, this constitutes the fundamental model for X-ray diffraction and remains the most common definition of electronic density $\rho(\vec{r})$. Known as the independent atom model (IAM), this approach has several limitations, most notably its disregard for chemical environment in the analysis of electronic distribution. More advanced models addressing this specific limitation will be discussed in Section 2.7.3.

Equation 2.3 defines the structure factor for electronic density scattering. In crystalline materials, we must account for an infinite array of scatterers related by translational symmetry. Recalling our unit cell definition, we note that describing a single unit cell suffices to characterize the entire crystal's electronic density. The crystalline electronic density $\rho(\vec{r})$ can be expressed as:

$$\rho(\vec{r}) = \rho_{cell}(\vec{r}) * \sum_{u,v,w} \delta(\vec{r} - \vec{r}_{u,v,w}) \tag{2.4}$$

where $\vec{r}_{u,v,w}$ represents generic lattice vectors, corresponding to the crystal's translation vectors, with integer coefficients and -∗- is the convolution operator. The unit cell's electronic density arises from the superposition of individual atomic densities:

$$\rho_{cell} = \sum_j \rho_j(\vec{r} - \vec{r}_j) \tag{2.5}$$

$$F_{\mathbf{h}} = \sum_j f_j \exp(2\pi i \vec{h} \cdot \vec{r}_j) \tag{2.6}$$

here, $\vec{h} = \sum_{n=1}^{3} h_n \vec{a_n}^{\,*}$ with integer coefficients h_n, and f_j denotes the atomic scattering factor for the j-th atom. The structure factor $F_{\mathbf{h}}$ emerges as the Fourier transform of the crystalline electronic density. A comparison of Equations 2.3 and 2.6 reveals important relationships between $F_{\mathbf{h}}$, f_j, and $\rho(\vec{r})$: The atomic scattering factor f_j can be determined through spatial integration of the electronic density $\rho(\vec{r})$; The structure factor $F_{\mathbf{h}}$ emerges as the Fourier transform of the electronic density.

2.2 Resonant Scattering

In conventional XRD experiments, researchers typically select a wavelength that minimizes inelastic effects, primarily to avoid absorption artifacts that could compromise the analysis. However, various techniques have been developed to exploit effects outside of Thomson scattering, such as X-ray absorption spectroscopy (XAS) [18].

This text will employ a semi-classical approach† to model X-ray inelastic scattering phenomena.

In the semi-classical model, resonant scattering will be described as the scattering of light by electrons bounded to atoms by a potential, acting as a forced oscillator. Assuming an incident X-ray beam linearly polarized in the $\hat{x}$ direction $E_{\text{in}} = \hat{x}E_0 \exp(-i\omega t)$ the movement equation for the electron is given by:

$$\ddot{x} + \Gamma\dot{x} + \omega_s^2 x = -\left(\frac{eE}{m}\right)\exp(-i\omega t) \tag{2.7}$$

Using a test solution of the form $x(t) = x_0 \exp(-i\omega t)$ in Equation 2.7, one finds the amplitude x_0 to be:

$$x_0 = -\left(\frac{eE}{m}\right)\frac{1}{\omega_s^2 - \omega^2 - i\omega\Gamma} \tag{2.8}$$

The radiated field is proportional to the acceleration in retarded time $\ddot{x}(t - R/c)$, giving:

$$E_{\text{rad}}(R,t) = \left(\frac{e}{4\pi\epsilon_0 Rc^2}\right)\ddot{x}(t - R/c)$$

Using $\ddot{x}(t - R/c) = -\omega^2 x_0 \exp(-i\omega t)\exp(iR\omega/c)$, we obtain:

$$\begin{aligned} E_{\text{rad}}(R,t) &= \frac{\omega^2}{(\omega_s^2 - \omega^2 - i\omega\Gamma)}\left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right)E_0 \exp(-i\omega t)\left(\frac{\exp(ikR)}{R}\right) \\ \frac{E_{\text{rad}}(R,t)}{E_{\text{in}}(R,t)} &= -r_0\frac{\omega^2}{\omega^2 - \omega_s^2 + i\omega\Gamma}\left(\frac{\exp(ikR)}{R}\right) \end{aligned} \tag{2.9}$$

The atomic scattering factor f_s given in Equation 2.9 is:

$$f_s = \frac{\omega^2}{\omega^2 - \omega_s^2 + i\omega\Gamma} \tag{2.10}$$

As a complex number, f_s can be expressed as $f_s = f_s' + if_s''$. Applying this decomposition to Equation 2.10 yields:

$$f_s' = \frac{\omega_s^2(\omega^2 - \omega_s^2)}{(\omega^2 - \omega_s^2)^2 + (\Gamma\omega)^2} \tag{2.11}$$

$$f_s'' = -\frac{\omega_s^2\omega\Gamma}{(\omega^2 - \omega_s^2)^2 + (\Gamma\omega)^2} \tag{2.12}$$

† While the quantum mechanical treatment provides greater mathematical rigor and consistency, it offers less physical intuition for the phenomena under study. For readers interested in the quantum mechanical approach, the author recommend Section 8.7 of *Elements of Modern X-ray Physics* by Als-Nielsen and McMorrow [19].

The function described by Equation 2.10 is analytic in the upper half of the complex plane. This implies that f'_s and f''_s are related through Kramers-Kronig relations, as discussed in Appendix A.1. Consequently, we may write:

$$f'_s = \frac{2}{\pi} \fint_0^{+\infty} \frac{\omega' f''_s}{\omega'^2 - \omega^2} d\omega' \tag{2.13}$$

$$f''_s(\omega) = -\frac{2\omega}{\pi} \fint_0^{+\infty} \frac{f'_s(\omega')}{\omega'^2 - \omega^2} d\omega' \tag{2.14}$$

2.3 Atomic Displacement Parameter

Up to this point, we have not considered the influence of temperature and defects in the mean position of atoms. But we know that the atomic position should vary due to temperature and crystal defects. To account for that a tensor T_j will be introduced to the structure factor, and this tensor will account for the variation in the mean atomic position. When accounting for the variation of the average position of the atoms in the crystal, the structure factor can be written as:

$$F_{\mathbf{h}} = \sum_j f_j(\vec{h}) T_j(\vec{h}) \exp(2\pi i \vec{h} \cdot \vec{r}_j) \tag{2.15}$$

$$T_j = \exp\left(-2\pi^2 B \left[\frac{\sin(\theta)}{\lambda} \right]^2 \right)$$

When T_j is isotropic, the factor B is given by:

$$B = 4U_{jj}$$

Otherwise, it is given by:

$$B = \sum_{i,j=1}^{3} U_{ij} {x_i}^* {x_j}^*$$

Accounting for temperature is essential for an accurate description of electronic density and can significantly influence structural interpretation.

In structural modeling, the Atomic Displacement Parameter (ADP) manifests as a deviation from the idealized spherical atomic form. Isotropic displacements are represented by spheres whose radii correspond to the probability of finding the atom within the depicted region; typically, the 50% probability isosurface. For anisotropic displacements, such as those observed in the Pb and Br atoms in Figure 7, the probability distribution is represented by ellipsoids, where the enclosed volume likewise corresponds to a 50% probability of locating the atom.

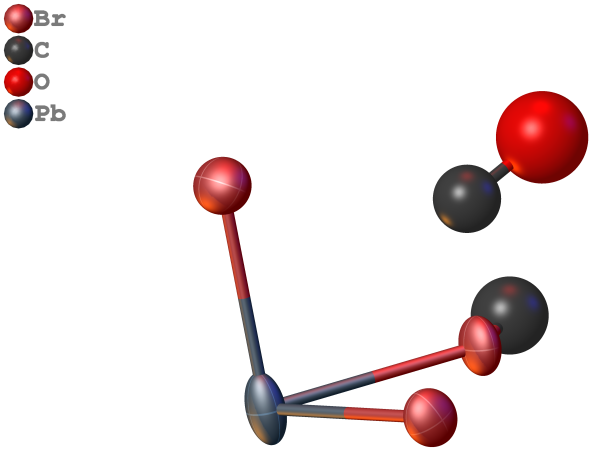


Figure 7 – Representation of isotropic (spherical) and anisotropic (ellipsoidal) Atomic Displacement Parameters (ADPs) for selected atoms. The iso-surfaces enclose regions with a 50% probability of atomic occupancy.

2.4 The phase problem

In XRD experiments, the incident radiation possesses an unknown initial phase, and detectors (such as photonic plates and CCDs) record only the intensity of diffracted rays. While the diffracted wave has both phase and amplitude components, only the latter is measured experimentally. This constitutes the fundamental **phase problem** in crystallography.

This limitation necessitates the development of models capable of accurately describing the measured intensities to assign phases and reconstruct the atomic arrangement. In the elastic scattering regime, all results for Thomson scattering remain valid, yielding some key relationships of the proportionality of the intensity of the diffracted ray:

$$I_{\vec{h}} \propto |F_{\vec{h}}|^2 \tag{2.16}$$

$$I_{\vec{h}} \propto \frac{1+\cos^2(2\theta)}{2} \tag{2.17}$$

$$I_{\vec{h}} \propto I_0$$

Here, $I_{\vec{h}}$ represents the intensity of a Bragg peak and I_0 the incident beam intensity. Equation 2.16 directly relates the measured intensity to the structure factor, while Equation 2.17 describes the polarization factor arising from the experimental geometry.

An additional geometrical consideration is the Lorentz factor, which accounts for variations in Bragg peak exposure times during XRD experiments. The exact form depends on the experimental geometry, but its simplest formulation assumes constant angular velocity of the crystal (most applicable to powder XRD):

$$L = (\sin\theta)^{-1}$$

This factor is typically combined with the polarization correction to form the Lorentz-Polarization (LP) factor. In its simplest form, the LP factor is given by:

$$LP = \frac{1 + \cos^2(2\theta)}{\sin^2(\theta)\cos(\theta)} \tag{2.18}$$

Absorption must also be considered in XRD analysis. The intensity attenuation follows an exponential decay characterized by the absorption coefficient μ, which is proportional to f'' and the path length through the crystal. This relationship can be expressed as:

$$I = I_0 \exp(-\mu x) = I_0 A(\mu, x) \tag{2.19}$$

The path-length dependence presents a challenge, as an analytical solution for the ideal absorption correction factor is generally unavailable[†].

An additional phenomenon, known as secondary extinction (E), arises from internal diffraction within the crystal. For small crystals, this effect exhibits behavior similar to absorption. Incorporating all these effects yields the complete intensity expression:

$$I = I_0 LPAE|F_{\vec{h}}|^2 \tag{2.20}$$

Following the application of Equation 2.20, the modulus of the structure factor can be determined. The electronic density $\rho(\vec{r})$ is obtained through the inverse Fourier transform of the structure factor multiplied by the unit cell volume V. This relationship can be expressed mathematically as $\frac{1}{V}\mathscr{F}[F_{\vec{h}}] = \rho(\vec{r})$ where $\mathscr{F}$ denotes the Fourier transform operator.

Consequently, the primary challenge in XRD analysis lies in solving the phase problem to determine atomic positions within the unit cell. This is achieved by modeling the electronic density as a superposition of atomic electron densities. Traditional XRD data analysis comprises two main stages:

- **Structure solution**: Identification of a preliminary structure that explains the measured intensities while estimating initial phases
- **Structure refinement**: Optimization of atomic positions, incorporating effects such as thermal displacement and, more recently, chemical deformation

During refinement, previously undetected atoms may be identified. This is particularly relevant for hydrogen atoms, which are typically not identified during the solution phase, but may also apply to other light elements in the structure.

[†] For further discussion, see Chapter 4 of *Fundamentals of Crystallography* by Giacovazzo et al. (2011) [17]

2.5 Structure Solution

Structure solution constitutes the initial and often most challenging step in XRD analysis. Historically, the first structures were determined through trial-and-error approaches, which proved both time-consuming and laborious. In this section, we explore two of the three primary structure solution methodologies: Patterson methods and direct methods.

We begin by revisiting Equation 2.6 from Section 2.1, which relates directly to measured intensities. Given a starting model, we may calculate the structure factor F_h^C as:

$$F_h^C = \sum_j f_j \exp(2\pi i \vec{h} \cdot \vec{r_j}) \tag{2.21}$$

To assess model quality, we employ the R-factor defined by:

$$R = \frac{\sum_h |\sqrt{I_h} - K|F_h^C||}{\sum_h \sqrt{I_h}} \tag{2.22}$$

where K represents a normalization constant.

The foundation of reliable structure solution lies in correct space-group assignment. Initial assessment involves analyzing intensity distributions in the diffraction pattern, as centric and acentric structures exhibit distinct probabilities for strong, intermediate, and weak reflections. Their respective probability distribution functions are:

$$\begin{aligned} P_{\bar{1}}(|F|) &= \frac{\sqrt{2}}{\sqrt{\pi\delta}} \exp\left(-\frac{|F|^2}{2\delta}\right) \\ P_1(|F|) &= \frac{2|F|}{\delta} \exp\left(-\frac{|F|^2}{\delta}\right) \\ \delta &= \sum_{j=1}^{N} f_j^2 = \frac{1}{\epsilon}\langle |F|^2 \rangle \end{aligned} \tag{2.23}$$

The set of equations in 2.23 reveals an inherent limitation: the mean value of $|F|$ varies with the Bragg peak angle θ due to geometric factors. Specifically, δ decreases as $\sin(\theta)/\lambda$ increases. To address this, we define two normalized quantities: the unitary structure factor U_h (Equation 2.24) and the normalized structure factor E_h (Equation 2.25):

$$|U_h| = \frac{|F_h|}{\sum_j f_j} \tag{2.24}$$

$$|E_h| = \frac{|F_h|}{\sqrt{\langle |F|^2 \rangle}} \tag{2.25}$$

These normalized factors yield modified probability distributions for centric and acentric structures (Equations 2.26 and 2.27, respectively):

$$P_{\overline{1}}(|F|) = \sqrt{\frac{2}{\pi}} \exp\left(- \frac{|E|^2}{2} \right) \tag{2.26}$$

$$P_1(|F|) = 2|E| \exp\left(- |E|^2 \right) \tag{2.27}$$

The distributions demonstrate that centric structures exhibit more extreme intensity variations (both strong and weak peaks), while acentric structures predominantly show intermediate intensities. This characteristic serves as a valuable tool for space group determination when centric and acentric options appear equally probable after Bragg peak indexing. However, the method remains non-deterministic as some diffraction patterns can be adequately described by either lattice type due to their similar distributions.

- **Atomicity**: As established in Equation 2.5 and discussed in Section 2.1, this principle states that the electronic density can be described by the superposition of the electronic density of atoms.
- **Non-negativity of** $\rho(\vec{r})$: While non-essential for Patterson methods, this assumption becomes crucial for direct methods.

2.5.1 Patterson Methods

Patterson methods constitute a class of structure solution techniques based on the Patterson function, denoted as $P(\vec{u})$. This function represents the self-convolution of the electronic density and is mathematically expressed as:

$$P(\vec{u}) = \rho(\vec{r}) * \rho(-\vec{r}) = \int_V \rho(\vec{r})\rho(\vec{r} + \vec{u})dr \tag{2.28}$$

By incorporating Equations 2.4 and 2.5, we can reformulate Equation 2.28 as:

$$P(\vec{u}) = \sum_j \rho_j(\vec{r}_j)\rho_j(\vec{r}_j + \vec{u}) \tag{2.29}$$

Furthermore, the Patterson function can be shown to be the inverse Fourier transform of $|F(\vec{h})|^2$[†], yielding:

$$P(\vec{u}) = \int_{v^*} |F(\vec{h})|^2 \exp(-2\pi i \vec{r}^{\,*} \cdot \vec{u}) d\vec{r}^{\,*} \tag{2.30}$$

Given the relation $|F_h| = |F_{-h}|$ and the discretization imposed by equation 2.4, this simplifies to:

$$P(\vec{u}) = \frac{1}{V} \sum_h |F_h|^2 \cos(2\pi i \vec{h} \cdot \vec{u}) \tag{2.31}$$

[†] For a detailed derivation, see Chapter 5 of *Fundamentals of Crystallography* by Giacovazzo [17]

Equation 2.31 enables the reconstruction of a complete Patterson map from experimentally measured intensities. The interpretation of this map stems from Equation 2.29. Due to the atomicity condition $P(u) \neq 0 \iff \rho_j(\vec{r}_j) \neq 0$ and $\rho_j(\vec{r}_j + \vec{u}) \neq 0$. Since $\vec{r}_j$ represents an atomic position within the unit cell, this condition requires that $\vec{r}_j + \vec{u}$ must also correspond to an atomic position. Consequently, for $\vec{u} \neq 0$, the vector $\vec{u}$ necessarily represents an interatomic vector.

A Patterson map represents all interatomic distances within the crystal, with each peak's intensity being directly proportional to the electron density of the contributing atoms. This principle forms the basis of the **heavy atom method**, where peaks associated with heavier atoms exhibit greater intensity, thereby facilitating their identification. Subsequent refinement cycles (discussed in detail in Section 2.6) enable the determination of lighter atom positions.

The method's effectiveness depends on the ratio of the heavy atom's squared electron density to that of the remaining unit cell. When this ratio approaches $\frac{1}{2}$, the probability of correctly determining the heavy atom positions increases significantly.

Patterson methods find particular utility in solving structures containing a small number of significantly heavier atom, as commonly encountered in coordination compounds and metalloproteins. These methods are classified as indirect because they recover phase information not directly from measured intensities, but rather through analysis of the Patterson map.

2.5.2 Direct Methods

Direct methods constitute a class of structure solution techniques characterized by the direct determination of phase information from measured intensities. While phase is typically unrelated to intensity in wave phenomena, X-ray diffraction presents a special case: the diffraction pattern arises from a discrete, non-negative electron density distribution. These combined assumptions enable the derivation of phase information from intensity measurements. The discrete nature of the electron density further motivates the use of normalized structure factors (Equation 2.25).

As previously discussed, the inverse Fourier transform of F_h yields $V\rho(\vec{r})$. We may similarly define a structure factor G_h associated with the squared electron density $\rho^2(\vec{r})$, whose inverse Fourier transform gives $V\rho^2(\vec{r})$. Applying the convolution theorem,

we obtain $\mathscr{F}[\rho^2(\vec{r})] = (1/V)F_h \cdot (1/V){F_h}^*$. Given the discrete nature of F_h, this leads to:

$$\begin{aligned} G_h &= \frac{1}{V}\sum_k F_k F_{h-k} \\ F_h &= f_h \sum_j \exp(2\pi i \vec{h}\cdot\vec{r_j}) \\ G_h &= g_h \sum_j \exp(2\pi i \vec{h}\cdot\vec{r_j}) \\ F_h &= \frac{f_h}{g_h} G_h \\ F_h &= \frac{\theta_h}{V}\sum_k F_k F_{h-k} \end{aligned} \tag{2.32}$$

where $\theta_h = f_h/g_h$. Equation 2.32, known as Sayre's equation, represents a foundational development in direct methods. Multiplying 2.32 by F_{-h} yields:

$$|F_h|^2 = \frac{\theta_h}{V}\sum_k |F_h F_k F_{h-k}| \exp[i(\varphi_{-h} + \varphi_k + \varphi_{h-k})] \tag{2.33}$$

Since $|F_h|^2$ is real and positive, for large $|F_h|$ values the real components on the equation's right-hand side become dominant. This assertion requires verification through probability theory, which has provided the most significant advances in direct methods development. Equation 2.33 represents a triplet invariant, whose detailed treatment appears in Appendix A.2.

The probability formula for triplet invariants in non-centrosymmetric structures is given by Equation 2.34. Here, G_{hk} represents a factor dependent on the atomic types involved in phase determination, while L serves as a normalization factor dependent on G_{hk} and $\sigma_n = \sum_j Z_j^n$, where Z_j denotes the atomic number of the jth atom. The derivation of the probability function for equal atoms is provided in Appendix A.3:

$$\begin{aligned} P(\Phi_{hk}) &= (1/L)\exp(G_{hk}\cos(\Phi_{hk})) \\ G_{hk} &= \frac{2}{\sqrt{N}}|E_h E_k E_{h-k}|, \qquad \text{For equal atomic types} \\ G_{hk} &= 2\sigma_3\sigma_2^{-3/2}|E_h E_k E_{h-k}|, \qquad \text{For different atomic types} \end{aligned} \tag{2.34}$$

Equation 2.34 follows a von Mises distribution (named after Richard von Mises), where G_{hk} acts as the concentration parameter. Under the assumption of intensity independence, knowledge of φ_k and φ_{h-k} pairs enables calculation of the phase probability distribution for φ_h

$$P(\varphi_h) = \prod_j P_j(\varphi_h) = A\exp\left(\sum_j G_{hk_j}\cos(\varphi_h - \varphi_{k_j} - \varphi_{h-k_j})\right) \tag{2.35}$$

The exponent in Equation 2.35 can be rewritten as a single cosine function $\alpha_h \cos(\varphi_h - \beta_h)$. While this eliminates the summation, it introduces two new unknowns. To relate these quantities to experimental data, we employ the trigonometric identity $\cos(a - b) = \cos a \cos b + \sin a \sin b$. Defining $\omega_j = \varphi_{k_j} - \varphi_{h-k_j}$, we obtain:

$$\begin{aligned} \alpha_h \cos(\beta_h) &= \sum_j G_{hk_j} \cos(\omega_j) \\ \alpha_h \sin(\beta_h) &= \sum_j G_{hk_j} \sin(\omega_j) \\ \tan(\beta_h) &= \frac{\sum_j G_{hk_j} \sin(\omega_j)}{\sum_j G_{hk_j} \cos(\omega_j)} \end{aligned} \tag{2.36}$$

Equation 2.36 represents the renowned **tangent formula**, which revolutionized XRD structure solution. To demonstrate its utility, we determine the most probable value for φ_h by differentiating Equation 2.35:

$$\frac{dP}{d\varphi} = A \exp(\alpha_h \cos(\varphi_h - \beta_h))[\alpha_h \sin(\varphi_h - \beta_h)]$$

This derivation shows that the maximum probability occurs when $\varphi_h = \beta_h$, confirming that Equation 2.36 yields the optimal phase estimate.

This mathematical framework enables structure solution via direct methods. The initial widely-adopted algorithm proceeded through three stages:

1. Random assignment of a small set of initial phases
2. Phase permutation to expand the starting set
3. Application of the above calculations to determine remaining phases

Subsequent advancements employed large sets of random initial phases, which typically converged to solutions more efficiently than the original approach.

As discussed by George Sheldrick [20], this presentation of direct methods exhibits some limitations:

- The approach is highly nondeterministic due to its reliance on random initial phases.
- Heavy atoms remain significantly easier to locate, often being determined at the expense of lighter atoms.
- The method frequently converged to trivial solutions in non-translational space groups (e.g., $P\bar{1}$, $C2/m$), known as "uranium atom" solutions.
- Only a small fraction of the data (approximately 15%), corresponding to the strongest reflections, was typically utilized.

Initial improvements emerged in the 1970s with the introduction of quartet invariants alongside triplets. These additional relations enabled the incorporation of more data, particularly weaker reflections. The resulting algorithms proved highly effective and remained in use for decades, with George Sheldrick's *SHELXS* [21] standing out for its computational efficiency and longevity.

Despite these advances, this generation of direct methods still lacked chemical constraints, maintaining a tendency to produce "uranium atom" solutions, especially for structures containing numerous atoms. Nevertheless, these methods successfully solved many small-molecule structures, and the *SHELXS* algorithm remained actively used until 2015. Notably, during the development of direct methods, researchers recognized that initial solution in P1 followed by subsequent space-group determination proved more efficient than attempting direct solution in the original space group. Consequently, most direct methods algorithms adopted this approach [22].

2.5.3 The path to Intrinsic phasing

The subsequent significant advancement in this field was the development of the "Shake-and-Bake" (SnB) algorithm. SnB introduced a residual function for diffracted wave phases, thereby establishing a variational principle in crystallographic structure solution. As a member of the "dual-space" solution family, SnB operates in two major steps:

1. Minimization of the residual function in reciprocal space
2. Real-space filtering, where atoms are identified in the electron density map and constrained to satisfy specific chemical requirements

This approach effectively addressed several limitations of traditional direct methods, most notably through automated atomic identification and incorporation of chemical constraints into the structural model. For many years, dual-space structure solution algorithms were divided between SnB and traditional direct methods. This paradigm persisted until the introduction of *SHELXT* [12]. Recognizing that valuable innovations were dispersed across multiple programs, George Sheldrick consolidated these advances into a unified algorithm. This integrated approach, termed intrinsic phasing, was implemented in the *SHELXT* software.

SHELXT incorporates several distinctive features. First, it implements a modified Patterson map approach by calculating minima from the superposition of two Patterson maps displaced by a strong diffraction peak. This generates a P1-expanded representation of the original structure, and according to Sheldrick [23], the initial phases obtained through this method consistently outperform random starting phases.

The subsequent step, termed dual-space recycling, employs specific algorithms for phase determination. The *SHELXT* algorithm utilizes the classical tangent formula to determine phases from an initial set. Following this, it implements a procedure adapted from protein crystallography known as "Random OMIT Maps". In this process, the algorithm randomly suppress a small portion of the Bragg peak's intensity (typically <10%), performs refinement, and generates a new difference electron density map. This approach enhances the robustness of the crystallographic model by systematically verifying its stability.

The final component involves the "free lunch" algorithm [24], which facilitates structure solution by generating synthetic data points for unmeasured Bragg peaks. These synthetic measurements are constrained to be consistent with the experimentally observed data, thereby expanding the dataset available for applications such as the tangent formula. This technique has gained widespread acceptance in XRD analysis since it maintains the integrity of the original experimental data without artificial modification. The complete algorithmic workflow is presented schematically in Figure 8.

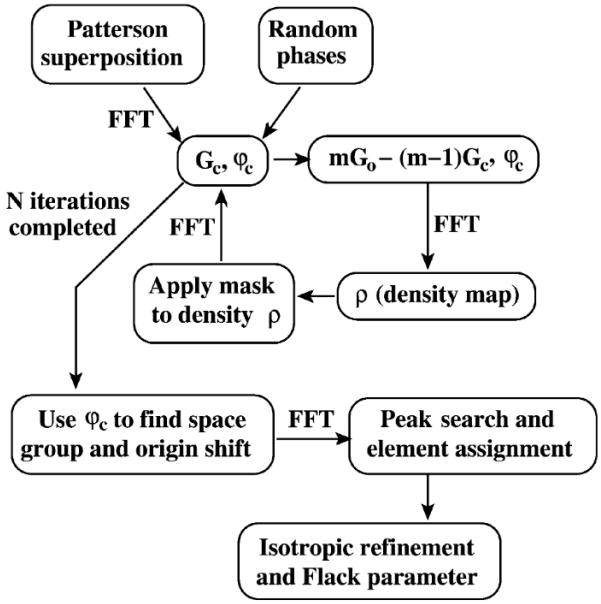


Figure 8 – Detailed scheme of the ShelXT algorithm. Source: Sheldrick (2015)

SHELXT is one of the most important contemporary structure solution algorithms and lay the basis for any future improvement in direct methods.

2.6 Refinement

Model refinement constitutes the second stage in XRD analysis. While one could theoretically initiate refinement cycles by randomly assigning atomic positions within the unit cell, this approach would likely fail to yield the correct atomic distribution, as XRD refinement is a non-deterministic optimization process.

Numerous software packages exist for single-crystal XRD refinement, with two particularly noteworthy examples: *SHELXL* [25], George Sheldrick's highly influential algorithm, and *olex2.refine* [26], a refinement engine developed for the *OLEX2* GUI [2] that incorporates most modern refinement techniques. The implementation strategy of *olex2.refine* represents a significant advancement by making sophisticated refinement methods accessible to non-specialists. The following discussion of refinement principles follows Bourhis (2015) [26].

In this framework, XRD refinement corresponds to the minimization of the function L (Equation 2.37):

$$L = \sum_{h} \omega_h (I_{0,h} - K I_{c,h})^2 + \sum_{\text{restraints}} \omega_i (T_{0,i} - T_{c,i})^2 \tag{2.37}$$

where $I_{0,h}$ represents the measured intensities; $I_{c,h}$ denotes the calculated intensities from the model; The summation over h includes all non-equivalent reflections; ω_h is a weighting function, typically $\frac{1}{\sigma^2(I_0)}$ (with σ as the standard deviation function) and K represents the scale factor

This minimization constitutes a least-squares optimization problem, where L is minimized with respect to the parameter space $[K, x]$. Here, x can encompasses all physical quantities required to compute the structure factor in single-crystal refinement [26].

Refinement using *olex2.refine* follows a two-step optimization procedure. First, the function L is minimized with respect to only the scale factor K while holding all other variables constant. Subsequently, L is minimized against the parameter space x while keeping the scale factor fixed. Notably, the parameter space x is typically smaller than the complete set of parameters required for structure factor calculation. For instance, the atomic form factor f_j is not refinable in *SHELXL*.

From a computational perspective, the refinement algorithm employs matrix formalism (for details, see Bourhis (2015)) [26]. Various minimization strategies are applicable to least-squares problems, with *olex2.refine* implementing two principal methods: Gauss-Newton (G-N) algorithm, and Levenberg-Marquardt (L-M) algorithm (both discussed further in Appendix A.4)

The L-M algorithm represents an extension of the G-N approach, incorporating gradient descent methodology to enhance reliability and expand the range of viable initial parameters compared to the basic G-N method.

The term T in Equation 2.37 relates to structural restraints that can be imposed during refinement. Such restraints may include fixed bond lengths (e.g., specific carbon-oxygen distances), constrained geometries (e.g., maintaining ring bond distances while allowing ring center movement), etc.

2.7 Hirshfeld Partitioning, Surfaces and HAR

Partitioning represents a fundamental concept in XRD and crystallography, serving to define both the unit cell in space and the contribution of individual molecules within the unit cell. Conventional single-crystal X-ray diffraction (SCXRD) typically assumes spherically symmetric atomic electron densities that remain unaffected by neighboring atoms, an approximation known as the Independent Atom Model (IAM). While IAM remains widely employed in most XRD refinements due to its computational efficiency, it exhibits several well-documented limitations. First, the model fails to account for chemical bonding effects, neglecting how atomic electron densities are perturbed by their molecular and crystalline environments [27]. Second, IAM demonstrates particular inaccuracy in hydrogen atom positioning, typically displacing them by approximately 0.2 Å [28].

To address these theoretical shortcomings, Hirshfeld introduced a stockholder partitioning approach [29] that quantitatively accounts for interatomic electron density modifications. This methodology, when properly implemented, enables more accurate determination of atomic electron densities within crystalline materials.

2.7.1 Hirshfeld Partitioning

Let us first examine Hirshfeld's original concept. His work focused on quantifying how each atom's electron density is modified throughout space. The fundamental idea involved weighting each atom's contribution to the electron density at a given spatial point based on its free-atom electron density. This concept can be mathematically expressed through a weight function $\omega_k(r)$:

$$\omega_k(r) = \frac{\sum_k \rho_k(r)}{\sum_j \rho_j(r)}$$

Here, the index k enumerates only the atoms of interest (whether a single atom or molecular fragment), while j indexes all atoms in the molecule, including those in subset k. For his analysis, Hirshfeld defined the resulting electron density $Q_j(r)$ for each atom as:

$$Q_j(r) = \int \omega_j(r)\rho_j(r)dr \tag{2.38}$$

2.7.2 Hirshfeld Surfaces

Initially, Hirshfeld partitioning was computationally impractical to implement. However, with substantial increases in computational power by the late 1990s, these calculations became feasible in the context of general molecular analysis. Building upon

this development and extending the original concept, Spackman adapted the Hirshfeld stockholder partitioning to redefine atomic electron densities within molecules. While employing the same formulation as Equation 2.38, Spackman additionally established that the electron density of atom j corresponds to the spatial region where $\omega_j(r) \geq 0.5$. This definition implies that an atom's (or fragment's) electron density occupies the volume where it contributes more than half of the local electron density.

This innovative approach was subsequently advanced when Spackman and Jayatilaka introduced the concept of the "Hirshfeld surface" in their 2009 publication [30]. The Hirshfeld surface represents the isosurface where $\omega_f(r) = \text{cte}$ for a selected molecular fragment. As demonstrated in their work, the surface generated by $\omega_f = 0.5$ proves particularly valuable for supramolecular analysis, with numerous applications detailed in the original publication [30].

This analysis focuses on the normalized distance parameter (D_{norm}), a key property for visualizing Hirshfeld surfaces. D_{norm} incorporates both internal ($D_{|i|}$) and external ($D_{|e|}$) normalized distances to the nearest atoms for each surface point, calculated as:

$$D_{|i| \text{ or } |e|} = \frac{D_{i \text{ or } e} - r_{vdw}}{r_{vdw}} \tag{2.39}$$

$$D_{\text{norm}} = D_{|i|} + D_{|e|} \tag{2.40}$$

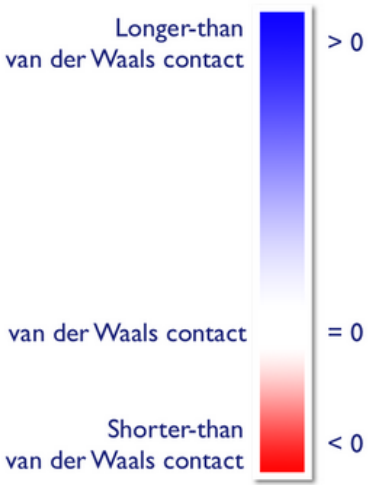
where r_{vdw} represents the van der Waals radius of the closest atom, thereby normalizing the distance against this radius. For experimentally measured compounds, Hirshfeld surfaces are typically generated by first calculating the electron density using density functional theory (DFT), followed by application of the Hirshfeld partitioning scheme discussed in Section 2.7.1. Figure 9 demonstrates a representative Hirshfeld surface colored by D_{norm}.

Beyond its value for identifying supramolecular interactions, another significant advantage of Hirshfeld surfaces lies in their ability to condense complex three-dimensional information (such as interaction distances) into two-dimensional fingerprint plots. These plots depict $D_{|i|}$ versus $D_{|e|}$, enabling rapid identification of surface features. Software packages like *CrystalExplorer* [31] facilitate this analysis by allowing selective highlighting of surface points associated with specific atomic pairs, as shown in Figure 10.

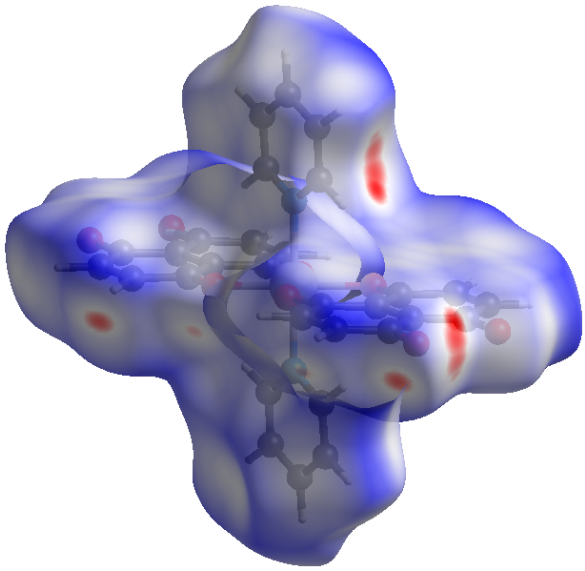
All Hirshfeld surface analyses in this work were performed using *CrystalExplorer*, the established software package for such calculations.

2.7.3 Hirshfeld atom refinement (HAR)

Hirshfeld atom refinement (HAR) extends this approach by applying the electron densities obtained through Hirshfeld stockholder partitioning schemes to XRD refinement. This method modifies the atomic scattering factor f_j in Equation 2.6 to account



(a) Scale used in the D_{norm} surface



(b) Hirshfeld Surface of the $[Co(Jug)_2Py_2]$ compound colored by D_{norm}

Figure 9 – Example of Hirshfeld surface

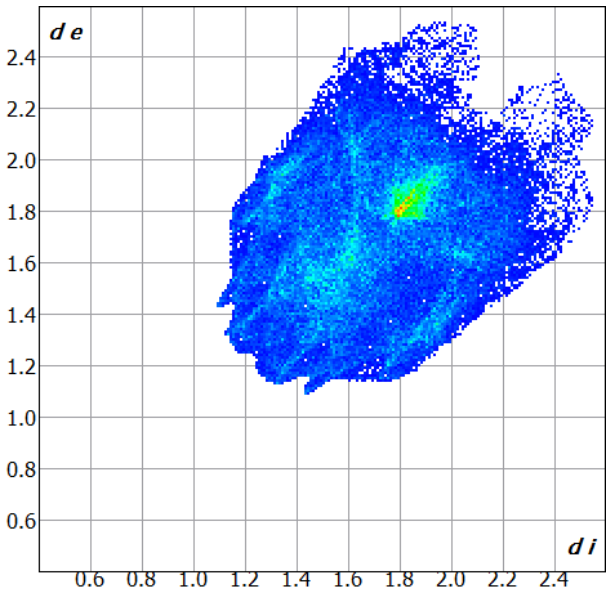


Figure 10 – Fingerprint of the surface in Figure 9

for chemical environment effects. In the Independent Atom Model (IAM), f_j remains spherically symmetric and depends solely on atomic type, as evidenced by its mathematical representation where f carries only the atomic-type subscript j and lacks dependence on the Miller index $\mathbf{h}$.

The HAR model introduces two significant modifications to this conventional approach. First, the resulting scattering factor becomes a complex-valued function requiring indexing through Miller indices [27], rather than maintaining real values throughout all space and being symmetrical in respect to $\mathbf{h}$. Second, and crucially for HAR's methodology,

atomic information cannot be pre-tabulated since such tabulation would eliminate the essential chemical context information. Consequently, wavefunction calculations must be performed for each molecular system analyzed using HAR, representing a notable departure from conventional XRD refinement pipelines.

HAR typically commences following conventional refinement, once a reliable starting model with most atomic positions determined has been established. The HAR procedure initiates by converting these XRD-refined positions into input for quantum mechanical (QM) calculations, typically employing density functional theory (DFT) to compute the molecular (or crystalline) wavefunction. The resulting wavefunction undergoes Hirshfeld partitioning, with the outcomes stored in tabular format for subsequent refinement software. In certain implementations, such as the Tonto software [32], the QM calculations and crystallographic refinement are performed within the same computational framework.

The modified atomic scattering factors are then incorporated into crystallographic refinement as detailed in Section 2.6. This iterative process, involving successive wavefunction calculations and refinements, continues until convergence criteria are satisfied. Figure 11 illustrates the complete workflow as implemented in the *NoSpherA2* software within the *OLEX2* package.

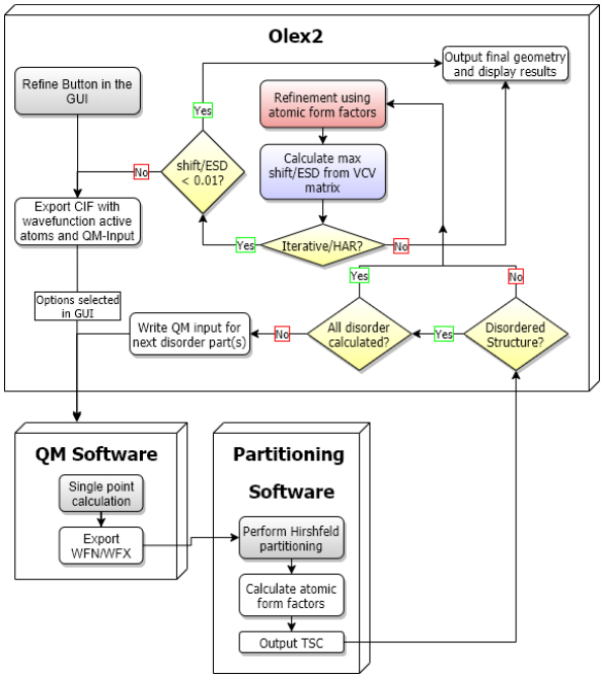


Figure 11 – NoSpherA2 scheme.

It is well known that HAR has some advantages in the precise determination of crystalline structure. Besides that, other capabilities of HAR are currently being explored. In example, as of 2024 Brüx et al [33] showed that HAR refinement may be able to determine spin states via the difference in residual electronic density between refinements that considered different spin states, as shown in Figure 12, and also the difference in the

statistical parameters of the refinements.

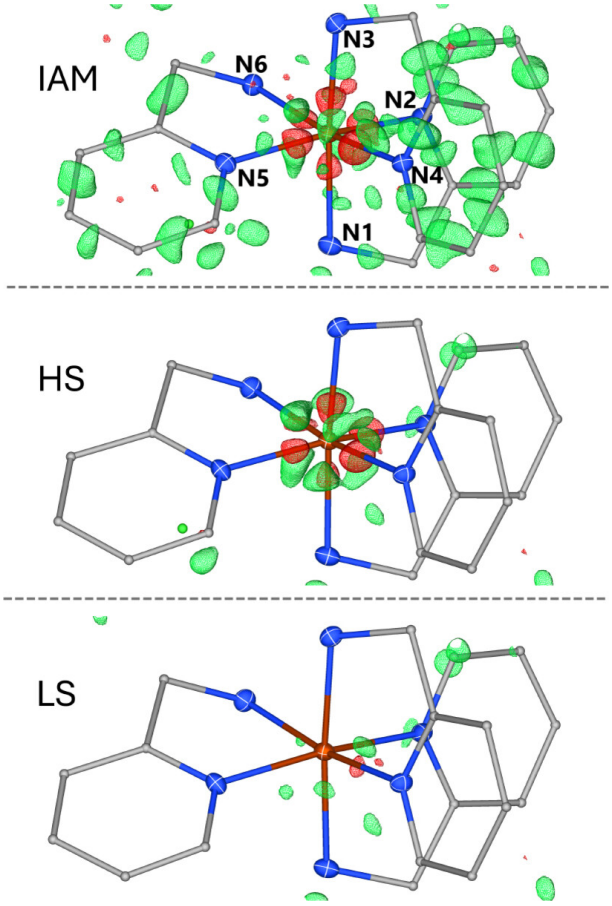


Figure 12 – Difference in refinement from variance in metallic spin multiplicity

The *NoSpherA*2 software represents a significant advancement in quantum mechanical crystallographic refinement, being the first user-friendly implementation that enables complete structure refinement through integration with the *olex*2.*refine* algorithm [27].

2.8 Anomalous dispersion refinement (ADR)

As discussed in Section 2.2, Anomalous dispersion play a significant role in XRD, particularly in experiments involving substantial absorption effects, such as those containing heavy atoms. The atomic scattering factor f_j can be decomposed as shown in Equation 2.41, where f_{0j} represents the elastic scattering component, f'_j the dispersive term, and f''_j the absorptive term:

$$f_j = f_{0j} + f'_j + i f''_j \tag{2.41}$$

The conventional approach to addressing this complexity in XRD involves utilizing tabulated values for f'_j and f''_j. Among the first reasonably accurate compilations was

that by Hönl [34], which served as the foundation for Cromer's subsequent work. This culminated in the 1970 publication describing ab initio calculation and interpolation methods for determining complex f_j values [35]. During the late 1980s and early 1990s, the International Union of Crystallography (IUCr) undertook a comprehensive effort to produce the most accurate tables of f' and f'' parameters for XRD refinement. These efforts resulted in multiple publications, with the definitive values being compiled in *The International Tables for Crystallography*, Volume C [36], which remain the standard reference for most conventional XRD refinements today. For wavelengths not covered by these tables, alternative theoretical sources such as those by Sasaki [37] and Henke [38] are typically employed.

The primary limitation of these approaches stems from their inability to account for the chemical environment of atoms within a structure. While the referenced tables demonstrate good experimental agreement for pure atomic systems in techniques like XAS, their accuracy diminishes when applied to bonded atoms, which represent the majority of structures investigated through XRD.

As demonstrated by Meurer et al. (2022) [39], refinement of anomalous dispersion parameters can yield significantly improved models when sufficient parameter-to-data ratios are available. Their work further illustrates how anomalous dispersion refinement (ADR) enhances atomic identification compared to conventional tabular methods. A notable example from their study reveals how classical refinement misidentifies a Mo atom as Cu due to improper modeling of absorption behavior in specific chemical environments.

These findings highlight the potential of ADR for structural analysis in XRD, particularly for systems containing bonded heavy atoms where substantial variations in f' and f'' may occur.

3 Case study

Two representative cases were selected to demonstrate the methods discussed in Chapter 2. The first case combines conventional crystallographic approaches for geometric analysis with Hirshfeld surface analysis (Section 2.7) to identify significant intermolecular interactions in the crystal packing. Hirshfeld atom refinement (HAR) was applied to both studied structures, significantly improving the reliability of measured distances as evidenced by notable reductions in both R1 and wR2 refinement metrics. The resulting structural data were subsequently used to determine the oxidation states of both the cobalt center electronic density and the juglone organic linker.

The second case utilizes inelastic scattering effects (Sections 2.2 and 2.8) to characterize metal site occupancy in metal-organic frameworks (MOFs). This approach provides a powerful analytical tool that not only identifies metal substitution but also distinguishes framework substitution from pore-confined metal ions, a critical distinction in MOF characterization.

3.1 Objectives

General Objectives

For both cases, a shared objective was to apply **quantum crystallography** techniques in the analysis of the structures and their properties. Another objective of the work was to provide the reader with a review of basic crystallographic concepts and of XRD-related techniques and advancements.

Specific objectives

- **Synthesize** both $[Co(Jug)_2(4\text{-}CN\text{-}py)_2]$ and $[Co(Jug)_2(py)_2]$
- Determine the **structure** of both compounds
- Analyze the **packing** of both structures
- Determine the **oxidation state** of the cobalt atoms
- Analyze the **coordination** of the juglone moiety with the metallic center
- **Synthesize** bimetalic MM-MOF-74
- Determine the **structure** of the bimetallic MOF with SCXRD
- Determine the **metallic occupation** with ADR analysis.

3.2 Cobalt-Juglone Compounds

Substituted p-quinones represent a class of redox-active molecules that play crucial roles in various biological processes, including respiration, photosynthesis, and enzymatic transformations [40, 41]. These cofactors exhibit reversible one- and two-electron transfer [42], as well as proton and proton-coupled electron transfer under suitable conditions, forming semiquinone ($pSQ^{\bullet -}$) and hydroquinone (pHQ) species [43]. Among these compounds, juglone (5-hydroxy-1,4-naphthoquinone) [44] emerges as particularly noteworthy. With the chemical formula $C_{10}H_6O_3$ and a molar mass of 174.16 g/mol, juglone consists of a bicyclic structure featuring a hydroxylated aromatic ring at the 5-position and a quinone moiety with oxygen atoms at positions 1 and 4. Although traditionally recognized for its applications as a natural dye and herbicide, several studies have investigated juglone's redox activity [45] and metal-complexing capabilities. However, most of these investigations have focused on non-structural characterization techniques such as thermometric analysis and infrared spectroscopy [46–48].

Metal-organic compounds, in turn, serve as excellent model systems for developing novel crystallographic techniques. These materials present some advantages for methodological exploration: (1) their spin and oxidation states offer potential for future applications; and (2) their supramolecular architectures often exhibit complex packing interactions that are not immediately obvious.

3.2.1 Synthetic route

All solvent and reactant manipulations were conducted over the inert gas atmosphere. The solvents were dried and degassed. The reactions were performed using Schlenk techniques to avoid contact with air.

A solution of dicobalt octacarbonyl (0.68 g, 2 mmol) in 15 mL of anhydrous, benzene was transferred using a cannula into a Schlenk flask, where it was mixed with a solution of (5-Hydroxy-1,4-naphthoquinone) (juglone- jug) (1.40 g, 8 mmol) in a blend of 10 mL hexane and 0.7 mL triethylamine, while stirring. It was observed bubbles from the mixture. The mixture was left to decant overnight. The supernatant solution was subsequently removed with a cannula. The reddish solid was left to sit within the Schlenk flask.

3.2.1.1 $[Co(Jug)_2(4\text{-}CN\text{-}py)_2]$

0.162 g of the reddish solid was dissolved in 20 mL of benzene. A solution of 4-cyanopyridine (0.083 g, 0.8 mmol) in 20 mL of benzene was introduced to the flask and stirred for one hour. Over the course of six months, as the benzene gradually evaporated,

small single crystals of $[Co(Jug)_2(4\text{-}CN\text{-}py)_2] \bullet [C_6H_6]$ (**1**) suitable for X-ray diffraction studies formed.

3.2.1.2 $[Co(Jug)_2Py_2]$

18 μL (0.22 mmol) of pyridine was dissolved in 4 mL of n-hexane inside a Schlenk flask under a positive argon atmosphere, and then slowly transferred using a cannula to another Schlenk flask containing 0.1859 g of the reddish solid, dissolved in 13 mL of benzene. The mixture was covered with a thin layer of aluminum foil and stirred for 1 hour. After resting for 24 hours to allow decantation, a small amount of monocrystalline $[Co(jug)_2(py)_2]$ (**2**) sample, suitable for X-ray diffraction studies, was formed.

3.2.2 Results and discussion

Single crystals of $[Co(Jug)_2(4\text{-}CN\text{-}py)_2] \bullet [C_6H_6]$ (**1**) were investigated by single-crystal X-ray diffraction at 100 K using 14 keV radiation at the MX2 beamline of the Brazilian Light Source (LNLS). $[Co(Jug)_2(py)_2]$ (**2**) was studied at 170 K using MoKα radiation on a GEMINI diffractometer. Data integration and intensity scaling were performed using CrysAlisPro 1.171.42.68a [49]. Both structures were solved by intrinsic phasing, with final unit cell parameters determined from refinement of all measured reflection positions.

For compound **2**, an analytical numeric absorption correction was applied using a multifaceted crystal model based on Clark and Reid's formalism [50], followed by empirical absorption correction via spherical harmonics implemented in the SCALE3 ABSPACK algorithm [49]. Structure refinements were carried out using *Olex*2.*refine* [26] with least-squares minimization through the Gauss-Newton algorithm. Both structures underwent Hirshfeld atom refinement (Section 2.7) employing the def2-SVP basis set and B3LYP level of theory.

All non-hydrogen atoms in compound **1** were unambiguously located in consecutive difference Fourier maps. For compound **2**, the carbonyl oxygen and its symmetry-equivalent hydrogen atom were refined with constrained site occupancy. Two benzene solvate molecules were identified and modeled in **1**. Hydrogen atoms were placed in idealized positions (C-H = 0.97 Å, $U_{iso}(H) = 1.2U_{eq}(C)$) [51] based on neutron diffraction tables, consistent with HAR requirements, and refined using a riding model. All non-hydrogen atoms were refined anisotropically.

Crystallographic details, including data collection parameters and refinement statistics, are compiled in Table 3. Supplementary crystallographic data have been deposited at the Cambridge Crystallographic Data Centre as CCDC 2371028 and 2371029, available at `www.ccdc.cam.ac.uk/data_request/cif`.

3.2.2.1 $[Co(Jug)_2(4\text{-CN-py})_2] \cdot [C_6H_6]$ structure

Compound **1** crystallizes in the $P\bar{1}$ space group, containing two half-independent $[Co(Jug)_2(4\text{-CN-py})_2]$ molecules in the asymmetric unit (Z = 2). Figure 13 presents the molecular structure and atomic labeling scheme for compound **1**. An overlay of the two molecules, generated using the Mercury software [52], reveals no significant structural differences (RMSD = 0.07 Å).

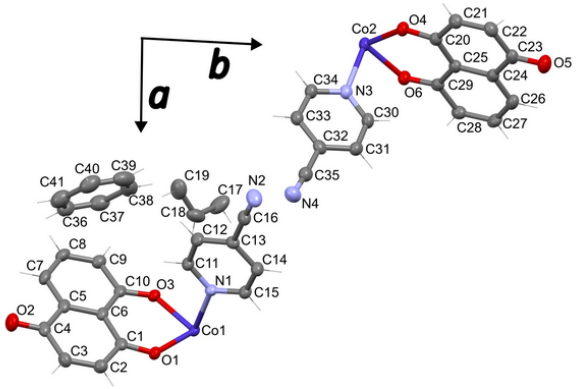


Figure 13 – Asymmetric unit of compound **1**. Color scheme: Carbon (grey), Oxygen (red), Nitrogen (light blue), Cobalt (dark blue). Ellipsoids represent 50% probability surfaces.

The juglone ligands coordinate the cobalt ions in a trans configuration, with each metal center located at an inversion center. The coordination sphere is completed by 4-CN-pyridine ligands, forming a nearly perfect octahedral geometry. Analysis using the *SHAPE* software [53] indicates minimal deviations from ideal octahedral geometry (0.212 Å and 0.285 Å for the two independent molecules). X-ray diffraction analysis successfully distinguishes between C-O single bonds (C10-O3 and C29-O6, 1.293(3) Å) and C=O double bonds (C1-O1, C4-O2, C20-O4, and C23-O5, 1.24(2) Å). These bond distances agree well with those observed in non-coordinated juglone molecules [44].

The absence of symmetry relations between the two independent moieties within the unit cell appears to originate from positional differences in the solvated benzene rings relative to their associated molecules, as illustrated in Figure 14. Notably, benzene B (situated at the inversion center) in Figure 14 is stabilized by two distinct interaction types: (1) T-shaped CH-π interactions with symmetry-related benzene molecules A and A', and (2) two opposing C_{Bz}-H...O_{jug} hydrogen bonds (3.312(3) Å and 3.586(4) Å) with symmetry-related **1** molecules

In the crystal structure, compound **1** exhibits multiple intermolecular interactions: (1) π-stacking between 4-CN-py and juglone aromatic rings (C-π distance = 3.387(3) Å), (2) π-stacking between juglone-juglone aromatic rings (O-C distances = 3.129(3) Å and 3.558(3) Å), as shown in Figure **??**, and (3) O_{jug}–H $\cdot\cdot\cdot$ C_{py} hydrogen bonds (3.13(3) Å). Additionally, the CN-Py plane displays a significant twist relative to the

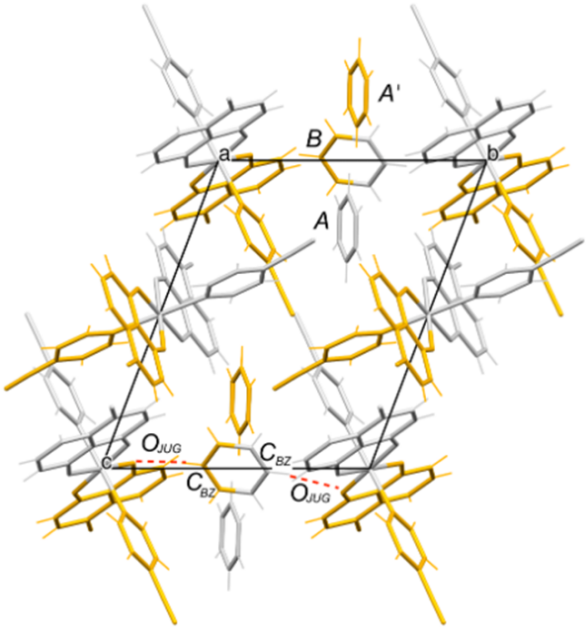


Figure 14 – Unit cell and molecular structure of benzene solvated **1** crystal at 100 K indicating the relative orientation of the benzene solvate molecules. Symmetry related moieties indicated by gray and yellow.

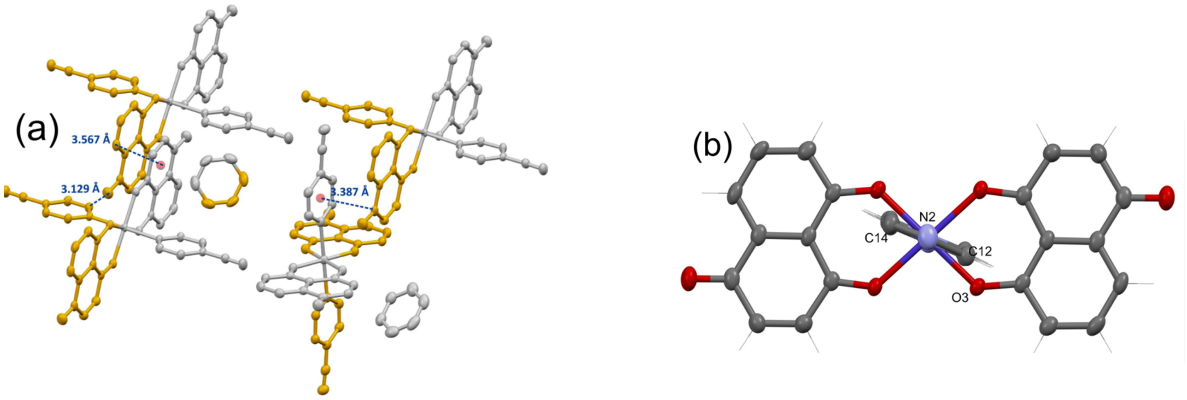


Figure 15 – (a)Supramolecular interactions of **1**.(b)Relative orientation of the CN-Py molecule in relation to the juglone plane in **1**. Symmetry related moieties indicated by gray and yellow. Black atoms are Carbon, light blue is Nitrogen, dark blue is Cobalt and red is Oxygen. Atoms are represented by ellipsoids with 50% probability

juglone plane, orienting the C11-H11, C15-H15, C30-H30, and C34-H34 bonds of each symmetry-independent pyridine molecule toward the single-bonded oxygen atoms O3, $O3^{i}$ (i = 1-x,2-y,2-z), O6, and $O3^{ii}$ (ii = 2-x,y,1-z), as illustrated in Figure **??**.

3.2.2.2 $Co(Jug)_2(py)_2$structure

Compound **2** crystallizes in the triclinic $P\bar{1}$ space group with half a $Co(Jug)_2(py)_2$ molecule in the asymmetric unit (Z=1). Figure 16 displays the molecular structure and

atomic labeling scheme for the asymmetric unit of **2**.

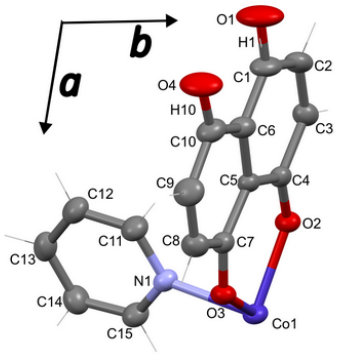


Figure 16 – Asymmetric unit of compound **2**. Color scheme: Carbon (grey), Oxygen (red), Nitrogen (light blue), Cobalt (dark blue). Ellipsoids represent 50% probability surfaces.

Similar to **1**, the juglone ligands coordinate the cobalt ion through both the deprotonated oxygen at position 5 of the hydroxylated aromatic ring and the carbonyl oxygen at position 1 of the quinone ring. The coordination sphere is completed by two pyridine ligands in a trans configuration. The cobalt center, located at an inversion center, exhibits octahedral geometry with only two independent Co-O and one Co-N bond distances refined. The juglone anion displays two equally probable coordination orientations, resulting in disorder of the position 4 carbonyl oxygen (O1 and O4 in Figure 16).

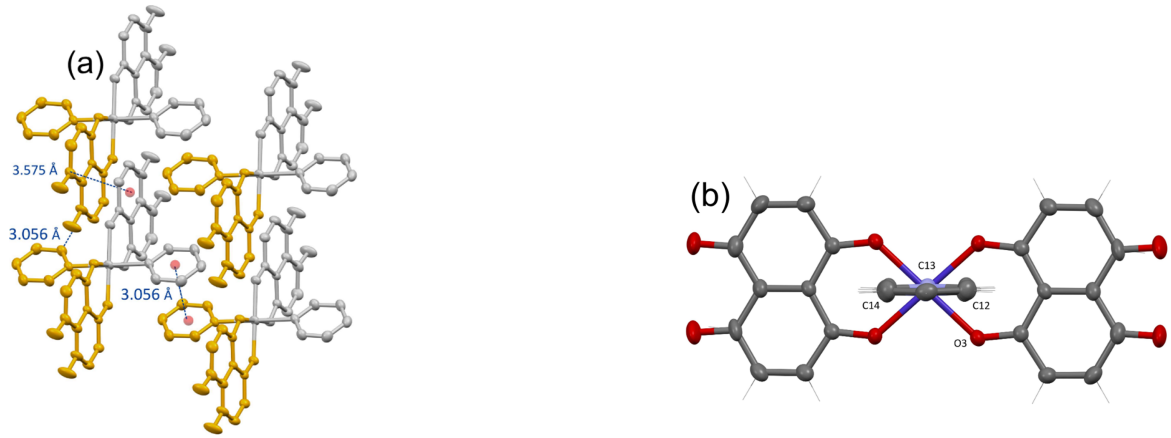
Analysis using SHAPE v2.1 software [53] reveals the octahedral coordination deviates from ideal geometry by 0.212 Å. The C4-O2 and C7-O3 bond distances of 1.271(1) Å and 1.272(1) Å, respectively, exhibit intermediate character between single and double bonds, consistent with the disordered coordination mode.

The packing interactions, analyzed similarly to **1** (Figure 17a), show distance variations attributable to the absence of benzene solvate molecules. Furthermore, the specific angle formed between the pyridine and juglone planes (Figure 17b) provides additional evidence for the two distinct juglone coordination modes, confirming this as a genuine structural feature rather than a geometric artifact.

Crystallographic tables with information on the refinement of both **1** and **2** can be found in Table 3

3.2.3 Hirshfeld Surface Analysis

As discussed in Section 2.7.2, Hirshfeld surface analysis provides valuable insights into intermolecular interactions, substantially improving our understanding of crystallographic packing arrangements [54].

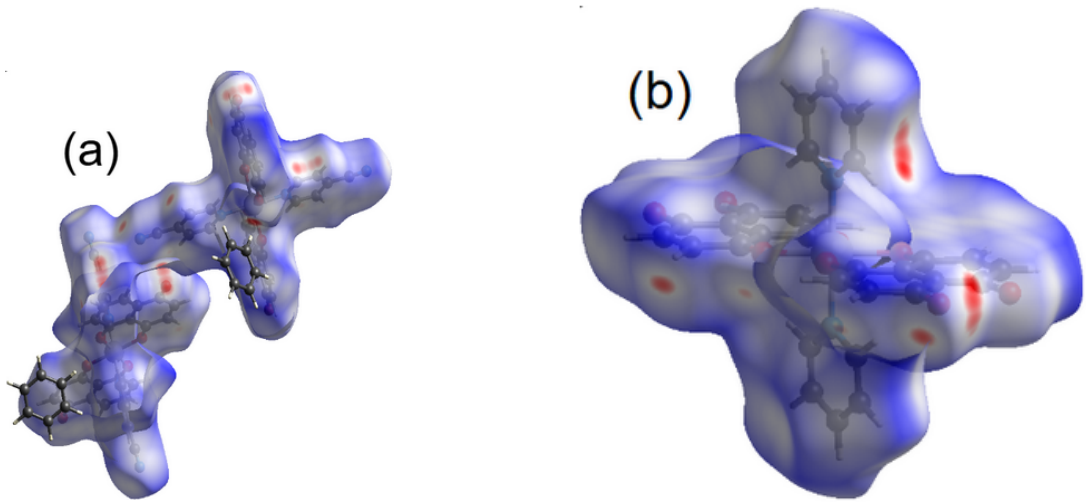


(a) Packing (b) Angle

Figure 17 – (a) π-stacking interactions between and the O_{jug}–*H*...C_{py} in **2**. (b) Relative orientation of the Py molecule. The angle between the juglone plane and a plane demarcated by Co1, O3, and N1 is 41.7(1)/°

The Hirshfeld surface of compound **1** (Figure 18a) exhibits distinct hydrogen-bond acceptor regions (red) surrounding the O1 and O3 oxygen atoms in the primary coordination sphere of Co1. Structural examination of these regions confirms the presence of Ojug · · · H-Cbz interactions between these oxygen atoms and the solvent benzene molecules. Similarly, the Hirshfeld surface of compound **2** (Figure 18b) reveals hydrogen-bond acceptor areas near O1, O4, C14, and C15, with the O4 · · · C14 interaction being particularly evident in the crystal packing.

The 2D fingerprint plot of **1** (Figure 19a) demonstrates a predominance of H · · · H repulsive interactions, as indicated by characteristic peaks at $d_i + d_e \approx 2.1$ Å. In contrast, the fingerprint plot of **2** (Figure 19b) lacks well-defined peaks at specific distances. For compound **2**, the most significant intermolecular interactions involve H...H contacts at slightly longer distances than those observed in **1**, with O · · · H interactions representing the second most prevalent interaction type.



(a) Hirshfeld Surface of compound **1**

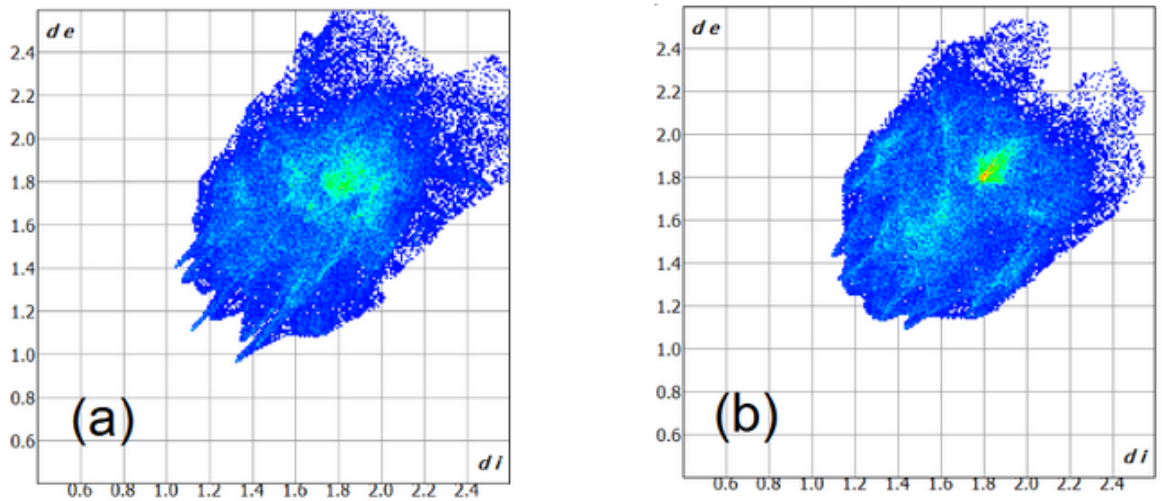
(b) Hirshfeld Surface of compound **2**

Figure 18 – (a) Hirshfeld Surfaces of **1**. The image shows the surface for both molecules of **1** in the unit cell, indicating in red the hydrogen-bond acceptor areas. (b) Hirshfeld Surfaces of **2**. The strong interaction area highlighted in the surface (red) is a strong intermolecular interaction.

Table 3 – *Crystallographic parameters for structures **1** and **2***

Identification code	$Co(Jug)_2(4\text{-}CN\text{-}py)_2$	$Co(Jug)_2(py)_2$
Empirical formula	$C_{41}H_{27}CoN_4O_6$	$C_{30}H_{20.02}CoN_2O_{5.98}$
Formula weight	730.628	563.136
Temperature/K	100.15	170.15
Crystal system	triclinic	triclinic
Space group	P-1	P-1
a/Å	11.214(3)	7.8650(6)
b/Å	11.368(3)	8.1792(6)
c/Å	13.964(4)	10.0582(7)
$\alpha/°$	110.79(2)	105.029(6)
$\beta/°$	90.80(2)	100.691(6)
$\gamma/°$	92.40(2)	97.740(6)
Volume/Å^3	1662.0(8)	602.56(8)
Z	2	1
ρ calcg/cm^3	1.460	1.552
μ/mm^{-1}	1.028	0.764
F(000)	753.3	289.4
Crystal size/mm^3	-	$0.2 \times 0.2 \times 0.1$
Radiation	synchrotron ($\lambda = 0.88563$)	Mo Kα ($\lambda = 0.71073$)
2θ range for data collection/°	4.54 to 61.82	5.38 to 52.74
Index ranges	$-12 \leq h \leq 12$, $-11 \leq k \leq 11$, $-15 \leq l \leq 15$	$-10 \leq h \leq 10$, $-10 \leq k \leq 10$, $-12 \leq l \leq 13$
Reflections collected	11686	5911
Independent reflections	4155 [Rint = 0.0279, Rsigma = 0.0295]	2386 [Rint = 0.0513, Rsigma = 0.0744]
Data/restraints/parameters	4155/0/505	2386/4/224
Goodness-of-fit on F^2	1.069	1.028
Final R indexes [I>=2σ (I)]	R1 = 0.0337, wR2 = 0.0936	R1 = 0.0405, wR2 = 0.0633
Final R indexes [all data]	R1 = 0.0339, wR2 = 0.0939	R1 = 0.0586, wR2 = 0.0732
Largest diff. peak/hole / e Å^{-3}	0.76/-0.25	0.47/-0.38

In this work, Hirshfeld surface analysis served as a complementary tool for structural characterization. Certain significant interactions that might be overlooked in conventional analyses can be identified through this approach, particularly when contact distances fall below the sum of the van der Waals radii of the interacting atoms. The Ojug–H · · · Cpy interaction discussed in the structural analysis section was specifically identified through Hirshfeld surface examination.



(a) 2D Fingerprint plot of compound **1** (b) 2D Fingerprint plot of compound **2**

Figure 19 – (a) 2D Fingerprint plot of Structure **1**. (b) 2D Fingerprint plot of Structure **2**. Distances in Angstroms (Å)

3.2.4 Oxidation State

The oxidation states of both the metal center and juglone ligand are of particular interest, as potential molecular applications require juglone to undergo transitions between oxidation and spin states. To investigate this, a comparative survey was conducted using the Cambridge Structural Database (CSD) [55].

Two distinct types of octahedrally coordinated cobalt compounds were examined. The first type (Scheme 3a) features cobalt centers coordinated to 1,2-dioxolene oxygen atoms, while the second type (Scheme 3b) involves oxygen atoms not bonded to adjacent carbon atoms. Table 5 presents the average Co-N and Co-O bond lengths, along with O-Co-O bond angles, for 195 structures from the CSD. Following the approach of Ribeiro et al. (2016) [56], which correlates cobalt oxidation states with coordination distances, the measured bond lengths in compounds **1** and **2** (2.21(2) Å, 2.03(3) Å, 2.185(2) Å, and 2.05(2) Å, respectively) closely match those observed in 42 Co(II) structures with Scheme 3b coordination.

Analysis of the bidentate coordination angles reveals that Co-(1,2-dioxolene) complexes exhibit smaller angles due to the constrained five-membered ring formation. In

contrast, the six-membered ring in Scheme 3b coordination allows greater flexibility, resulting in larger O-Co-O angles. The observed angles in compounds **1** and **2** show better agreement with Co(II) complexes of type 3b than with the two reported Co(III) analogues. These comparisons strongly suggest a Co(II) oxidation state in both compounds.

The formal charge of juglone was estimated as -1e based on average C-O, C-N, and C-C bond lengths, following Mukherjee's methodology [57]. This charge assignment is consistent with deprotonation of the juglone ligand.

Table 5 – Table summarizing the results of the survey

	Co-O (Å)	**Co-N (Å)**	**O-M-O (°)**	**Structures**	**State**
Scheme 3a	2.04(6)	2.13(6)	80(2)	49	Co(II)
Scheme 3a	1.89(2)	1.96(3)	87(1)	123	Co(III)
Scheme 3b	2.05(3)	2.18(6)	89.1(8)	42	Co(II)
Scheme 3b	1.883(3)	1.946(6)	95.59	2	Co(III)
1	2.03(3)	2.21(2)	86.9(5)	N/A	Co(II)
2	2.05(2)	2.185(2)	85.35(2)	N/A	Co(II)

3.2.5 Case conclusion

We successfully synthesized two novel cobalt-containing coordination complexes with juglone and pyridine-based ancillary ligands: $Co(Jug)_2(4\text{-CN-py})_2$ (**1**) and $Co(Jug)_2(py)_2$ (**2**). These structures demonstrate that negatively charged juglone can bind to Co(II) metal centers in a square planar configuration, facilitating the *trans*-coordination of pyridine ancillary ligands and resulting in octahedral coordination around the cobalt centers, even with nitrogenated ancillary ligands.

A comparison of the structures of compound **1** and compound **2** reveals the ability of juglone compounds to crystallize both with and without solvent in the structure while maintaining a *trans* configuration. The coordination symmetry of juglone appears to depend on the presence of the solvent, explaining why juglone coordinates in an ordered fashion in compound **1** but in a disordered manner in compound **2**. Consequently, information regarding C–O single- and double-bond character is preserved in compound **1** but lost in compound **2**.

Despite the inherent redox-active potential of naphthoquinones and the apparent rotational freedom of the ancillary ligand around its axis, according to X-ray diffraction data none of the juglone complexes examined in this work exhibited redox activity in the studied temperature range.

3.3 Occupancy determination in MM-MOF-74 compounds

Metal-Organic Frameworks (MOFs) represent a prominent class of porous coordination polymers characterized by the strong coordination bonds formed between multifunctional organic linkers and metal ions or polynuclear metal clusters. [58] These interactions yield highly ordered, crystalline structures with structural integrity and tunable porosity. Upon activation, typically through solvent removal, the intrinsic pores of MOFs are evacuated, exposing unsaturated metal sites (UMOs) and other active coordination centers.

The modular nature of MOF synthesis allows for precise control over their physicochemical properties by careful selection of metal nodes (e.g., Zn^{2+}, Cu^{2+}, Zr^{4+}) and organic ligands (e.g., carboxylates, azolates) [59]. The existence of such variety in the selection of SBUs and organic linkers, different pore size and geometry are achievable through chemical engineering, which makes MOF compounds very tunable for specific applications. Furthermore, post-synthetic modification (PSM) of the organic linkers, via functionalization with amine, hydroxyl, or other reactive groups, enables fine-tuning of MOF properties for targeted applications. [60]

Owing to their structural versatility, high surface areas, and tailorable chemical environments, MOFs have emerged as highly promising materials for diverse applications. These include heterogeneous catalysis (e.g., CO_2 reduction, organic transformations), [61] gas storage and separation (e.g., H_2, CH_4, CO_2 capture), [62] sensing, and even drug delivery. Their unique combination of properties positions MOFs at the forefront of advanced functional materials research.

3.3.1 Motivation

MOF-74 stands out due to its open one-dimensional infinite hexagonal pores, which facilitate the emptying process while also allowing molecules inside the pore to interact with both the metal ion and the organic linker [63]. Traditionally M-MOF-74 is composed of di-valent metals (M = Fe^{2+}, Zn^{2+}, Co^{2+}, Cu^{2+}, Ni^{2+}, Mg^{2+} or Mn^{2+}) with the organic linker 2,5-dihydroxyterephthalic ($C_8H_6O_6$). The array of possible metals and the stability of the structure give the MOF-74 great flexibility for applications. The packing of MOF-74, showing the hexagonal one-dimensional pore, can be seen in Figure 20.

The metallic center has a great influence on the possible applications of MOF compounds. Mg-MOF-74, for example, has been studied for potential gas sorption, with a dynamic gas sorption capability of 8.9% reported for the material [64]. Zn-MOF-74 was studied for its capability of immobilizing catalase enzyme (CAT) and α-chymotrypsin enzyme (CHT) successfully in aqueous conditions [65]. This successful immobilization also open doors for applications within the bio-catalyses realm as the MOF work as a

frame that can host other enzymes inside the pore and also grants the enzyme inside it anti-unfolding functions [65]. Co-MOF-74 has been studied for applications in energy storage, showing a reversible charge capability of 526.1 mA h g^{-1} [66].

The incorporation of mixed metals significantly broadens the range of potential applications for metal-organic frameworks (MOFs), as the properties, including chemical affinities, of multiple metals can be combined and fine-tuned for specific applications and environments. NiCo-MOF-74, for example, has been investigated for catalytic applications in wastewater treatment [67].

Characterizing the internal composition of such materials has largely relied on techniques that do not directly assess the atomic-scale structural distribution, such as inductively coupled plasma (ICP) spectroscopy [68], synchrotron-based methods like X-ray absorption spectroscopy (XAS) [69], or neutron diffraction [70]. Scanning electron microscopy (SEM) coupled with energy-dispersive X-ray spectroscopy (EDS) has also been employed as an analytical tool to examine metal substitution in MOFs [71]. However, SEM/EDS cannot distinguish between a mixture of phases and a homogeneous bimetallic phase, nor can it differentiate framework substitutions from metal atoms residing in the pores, whether as ions or unreacted reagents.

The described techniques share one major limitation in the analysis of metal substitution in MOF compounds: the inability to precisely determine the location of metal ions at the structural level. However, XRD-based techniques circumvent this problem by combining spectroscopic features with structural characterization capabilities.

This work reports the synthesis, structural characterization, and chemical analysis of MM-MOF-74 compounds, leveraging the newly developed capability to chemically probe the structure using techniques enabled by laboratory X-ray sources. This approach facilitates structural chemical exploration while improving accessibility.

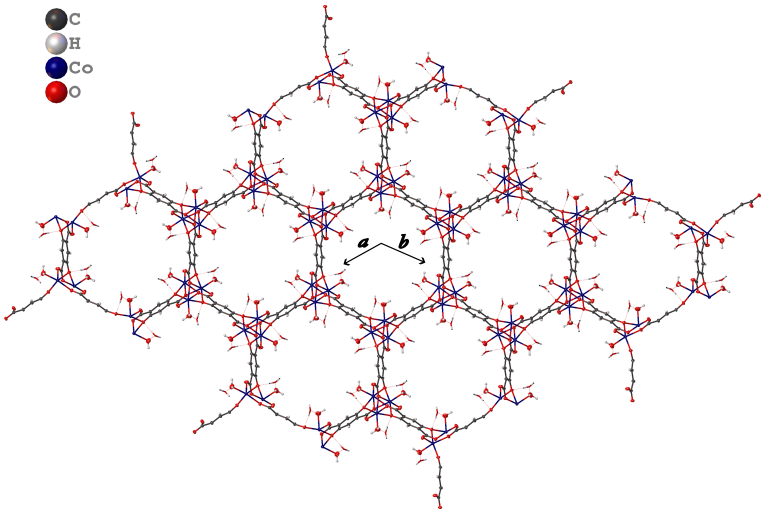


Figure 20 – Co-MOF-74 packing shown with suppressed oxygen atoms inside the pore.

3.3.2 Synthesis

MM-MOF-74 were synthesized via the solvothermal method as described by Juliana Figueiredo [72]. The synthesis proceeds at elevated temperatures within a sealed reactor. This method was selected due to its production of high-quality crystals suitable for single-crystal XRD analysis. The reaction occurs under pressures exceeding atmospheric pressure, which enhances the solubilization of both metal salts and organic linkers.

The procedure begins with the preparation of two separate DMF solutions. The first solution contains 2,5-dihydroxyterephthalic acid at a concentration of 1.26 mmol/L, while the second contains metal nitrates at 2.42 mmol/L for each nitrate salt, resulting in a total metal concentration of 4.84 mmol/L (comparable to conventional monometallic solvothermal MOF-74 syntheses). For this synthesis, both solutions were prepared with 8 mL volumes each, totaling 16 mL upon mixing. Following preparation, the metal solution was gradually added to the organic linker solution under continuous stirring for 10 minutes. Subsequently, 8 μL of water was introduced to the final mixture.

The resulting solution was transferred to a 50 mL Teflon reactor, which was securely sealed to ensure synthetic reproducibility. The reactor was then placed in a preheated oven at 120 °C for 24 hours. Following the reaction period, suitable single crystals (shown in Figure 21) were isolated from the product mixture.

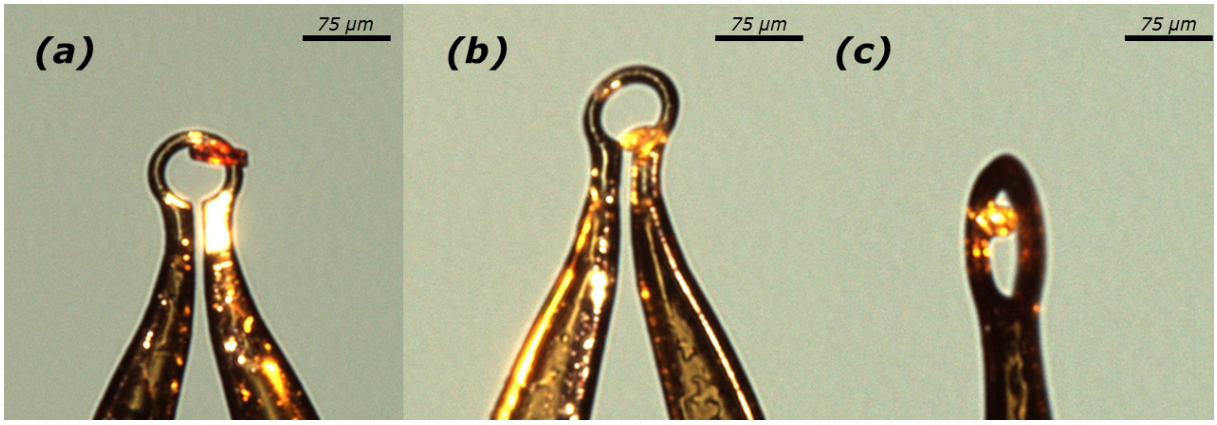


Figure 21 – MM-MOF-74 single crystals measured at the SYNERGY XTAL diffractometer located at the Crystallography Laboratory of UFMG (LabCri). (a) CoMg-MOF-74 (b)ZnMg-MOF-74 (c) CoZn-MOF-74

3.3.3 Methodology

Single crystals of CoMg-MOF-74 (**3**), ZnMg-MOF-74 (**4**), and CoZn-MOF-74 (**5**) were investigated by single-crystal X-ray diffraction at 300 K using CuKα radiation on a SYNERGY diffractometer at room temperature. Data integration and intensity scaling were performed using the *CrysAlisPro* 1.171.43.90 suite [49]. The final unit cell parameters were determined by fitting all measured reflection positions. An analytical numeric absorption

correction was applied using a multifaceted crystal model based on expressions derived by Clark and Reid [50], along with an empirical absorption correction employing spherical harmonics, implemented in the *SCALE*3 *ABSPACK* scaling algorithm [49].

All structures were solved by intrinsic phasing, and refinements were performed using *Olex*2.*refine* through the Gauss-Newton algorithm. All atomic positions in the frameworks of **3**, **4**, and **5** were unambiguously assigned through consecutive difference Fourier maps and refined with anisotropic atomic displacement parameters. Anomalous dispersion was refined according to the method described by Meurer et al. (2022) [39]. Occupancy was determined through two complementary methods: first by modeling as disordered atoms with split occupations, and second by performing anomalous dispersion refinement (ADR) and treating the resulting f'' as a linear combination of the isolated metals' f'' values, as proposed by Casaday et al. (2025) [73]. This approach is mathematically described in Equation 3.1.

$$f''_{\text{refined}}(E) = x f''_{\text{Metal 1}} + (1 - x) f''_{\text{Metal 2}} \tag{3.1}$$

3.3.4 Results and discussion

3.3.4.1 Crystaline Structure

All compounds crystallized in the $R\overline{3}$ space group, with the polymeric unit $M_1C_4O_4H_1$ (M = Co, Zn, Mg) consistently present (Z = 18). Figure 22 illustrates the labeling scheme for water-containing structures in the asymmetric unit. The oxygen atom at the O4 position may belong to either a water molecule or a dimethylformamide (DMF) molecule. When DMF is present, it consistently exhibits disorder and was accordingly modeled with two distinct positions. The symmetry of the polymeric structure restricts analysis of metal substitution distribution within the secondary building unit (SBU), as all metal positions are symmetry-equivalent.

3.3.4.2 Occupation

The occupancy was refined using a disorder model, where each atom was assigned partial occupancy at the crystallographic site with the combined occupancies summing to unity. These results were compared with the occupancy values obtained from Equation 3.1, which was derived by considering a single atom in the model and refining its anomalous scattering parameter. Both sets of results were further validated against chemical composition data obtained from atomic absorption spectroscopy (AAS). This technique was selected for its minimal sensitivity to interference from other atomic species, making it particularly suitable for analyzing metal-organic compounds that may contain unidentified

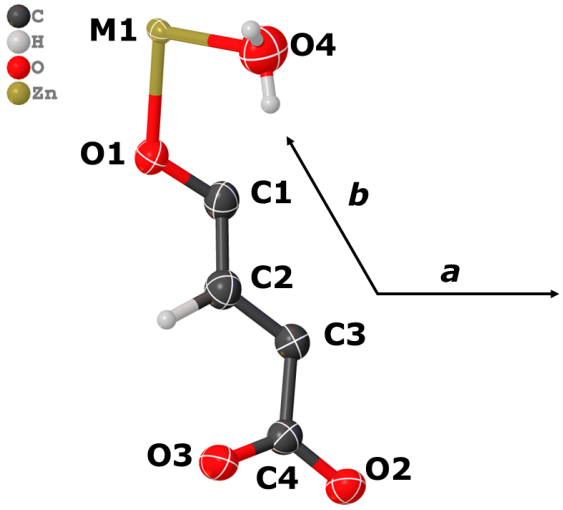


Figure 22 – Assymetric unit and labeling scheme for all MOF compounds. Color scheme: Carbon (black), Hydrogen (white), Oxygen (red) and metal (yellow). For the composition of the image, Zn-MOF-74 was used as reference. Ellipsoids represent 50% probability surfaces.

species within their pores [74]. The metal occupancy values determined by each method are presented in Table 6.

Table 6 – *Metal occupation found from different methods*

	Co:Mg	**Zn:Mg**	**Co:Zn**
Disord. Occu (%)	94:6(1)	61:39(1)	32:68(4)
ADR Occu (%)	86:14(10)	58:42(10)	42:58(10)
AAS Occu (%)	90:10(1)	68:32(10)	51:49(10)

These results demonstrate that refinement of anomalous scattering parameters provides an alternative method for determining metal center occupancy in these compounds. Furthermore, the f" parameter proves reliable even in space groups possessing inversion symmetry. Table 6 reveals that in CoMg-MOF-74, cobalt dominates magnesium within the structure, while other metal combinations exhibit nearly equimolar distributions.

The technique's site-specific nature ensures that the observed substitution occurs exclusively within the framework rather than in pore contents. Notably, for compound **5**, the occupancy determined through anomalous dispersion refinement (ADR) shows better agreement with AAS reference data than conventional disorder modeling. This observation can be partially explained by the proximity of cobalt and zinc atomic numbers. Traditional XRD modeling exhibits limitations when differentiating elements with similar atomic numbers, even when incorporating anomalous scattering parameters. However, ADR proves particularly effective for distinguishing cobalt from zinc in this context, as the experiment employed CuKα radiation (E = 8.04 keV). This energy exceeds the cobalt absorption edge (E = 7.71 keV), resulting in significantly larger f'' values for cobalt compared to zinc ($\Delta f'' \approx 3$ electrons).

Table 7 presents the refined values of f' and f'' for MM-MOF-74, along with reference values from the Brennan & Cowan tables in *The International Tables for Crystallography* [36].

Table 7 – *Refined anomalous parameters summarized*

	CoMg	**ZnMg**	**CoZn**	**Reference values**
Co f' (e)	-2.9(1)	-	0.7(2)	-2.384
Co f" (e)	3.3(1)	-	1.9(2)	3.620
Mg f' (e)	10.0(1)	10.1(1)	-	0.173
Mg f" (e)	4.4(1)	1.9(2)	-	0.178
Zn f' (e)	-	-8.5(1)	-3.1(2)	-1.542
Zn f" (e)	-	0.6(2)	1.5(2)	0.679

The refinement of the f' parameter appears less reliable than that of f'' for bimetallic MOFs. The literature concerning this technique remains underdeveloped, and standardized physical interpretations for certain refinements have not yet been established. Crystallographic data for compounds **3**, **4**, and **5** are presented in Tables 8 and 9.

Table 8 – *Crystallographic parameters for structures **3** and **4***

Identification code	CoMg-MOF-74	ZnMg-MOF-74
Empirical formula	C7H5.33CoNO4.67	C4H5O7Zn
Formula weight	237.107	230.476
Temperature/K	300.65(10)	300.74(10)
Crystal system	trigonal	trigonal
Space group	R-3	R-3
a/Å	26.0553(6)	26.1027(11)
b/Å	26.0553(6)	26.1027(11)
c/Å	6.8563(2)	6.7076(3)
$\alpha/^\circ$	90	90
$\beta/^\circ$	90	90
$\gamma/^\circ$	120	120
Volume/Å^3	4031.00(18)	3957.9(3)
Z	18	18
ρcalc g/cm^3	1.758	1.741
μ/mm-1	15.024	3.997
F(000)	2105.1	2050.1
Crystal size/mm^3	0.12 × 0.04 × 0.02	0.09 × 0.05 × 0.01
Radiation	Cu Kα (λ = 1.54184)	Cu Kα (λ = 1.54184)
2θ range for data collection/$^\circ$	6.78 to 157.92	6.78 to 157.66
Index ranges	-31 ≤ h ≤ 32, -32 ≤ k ≤ 28, -8 ≤ l ≤ 8	-31 ≤ h ≤ 32, -33 ≤ k ≤ 30, -8 ≤ l ≤ 8
Reflections collected	35484	34471
Independent reflections	1892 [Rint = 0.1217, Rsigma = 0.0418]	1866 [Rint = 0.1798, Rsigma = 0.0569]
Data/restraints/parameters	1892/0/150	1866/0/115
Goodness-of-fit on F2	0.993	0.992
Final R indexes (I ≥ 2σ)	R1 = 0.0489, wR2 =0.1372	R1 = 0.0597, wR2 =0.1555
Final R indexes (all data)	R1 = 0.0627, wR2 = 0.1465	R1 = 0.0955, wR2 = 0.1843
Largest diff. peak/hole / e Å^{-3}	0.73/-0.64	0.78/-0.59

Table 9 – *Crystallographic parameters for structures* ***5***

Identification code	CoZn-MOF-74
Empirical formula	C6H1.11CoN0.5O4.56
Formula weight	212.080
Temperature/K	301.11(10)
Crystal system	trigonal
Space group	R-3
a/Å)	26.0788(6)
b/Å	26.0788(6)
c/Å	6.7950(3)
$\alpha/^\circ$	90
$\beta/^\circ$	90
$\gamma/^\circ$	120
Volume/Å^3	4002.2(2)
Z	18
ρcalc g/cm^3	1.584
μ/mm-1	15.037
F(000)	1841.5
Crystal size/mm^3	0.06 × 0.06 × 0.03
Radiation	Cu Kα (λ = 1.54184)
2θ range for data collection/$^\circ$	-32 ≤ h ≤ 33, -31 ≤ k ≤ 31, -8 ≤ l ≤ 8
Reflections collected	35264
Independent reflections	1879 [Rint = 0.1017, Rsigma = 0.0324]
Data/restraints/parameters	1879/0/129
Goodness-of-fit on F2	1.020
Final R indexes (I ≥ 2σ)	R1 = 0.0633, wR2 =0.1952
Final R indexes (all data)	R1 = 0.0749, wR2 = 0.2099
Largest diff. peak/hole / e Å^{-3}	1.39/-1.30

3.3.5 Case conclusion

The synthesis and characterization of compounds **3**, **4**, and **5** were successful, with metallic substitution confirmed by two independent techniques: XRD and AAS.

We demonstrate that anomalous dispersion refinement (ADR) reliably determines metal center occupancy in coordination polymers, even when using radiation both far from (**4**) and near (**3** and **5**) the metal absorption edges. Furthermore, our results reveal that in **3**, cobalt atoms preferentially occupied the crystallographic sites despite equimolar stoichiometry in the synthesis, unlike the other bimetallic compounds which exhibited nearly equal (50-50) metal distributions in this study.

This technique successfully identifies framework-level metal substitutions, thereby complementing conventional characterization methods that previously could not unambiguously localize such substitutions. Notably, ADR provides more accurate results than traditional disorder modeling for distinguishing metals with similar atomic numbers,

highlighting its potential for such systems. These findings suggest that ADR can be generalized to other metal-substituted compounds, establishing it as a promising approach for characterizing metal-organic materials.

4 Conclusion and perspectives

In this work, X-ray diffraction and crystallographic techniques were employed as analytical tools for materials characterization.

We studied two novel compounds $Co(Jug)_2(4\text{-CN-py})_2$ (**1**) and $Co(Jug)_2(py)_2$ (**2**), reporting their synthesis and analyzing their structural features. Hirshfeld surface analysis revealed close contacts within the structures, identifying a strong intermolecular interaction between O4 and C14 in the crystal structure of **1** and its equivalent in **2**. The 2D fingerprint plot of **1** exhibits predominant H $\cdot\cdot\cdot$ H repulsive interactions, evidenced by characteristic peaks at $d_i + d_e \approx 2.1$Å. In contrast to **1**, the 2D fingerprint plot of **2** lacks distinct peaks, indicating increased hydrogen contact distances.

Interatomic distances determined by X-ray analysis confirmed the cobalt oxidation state as Co^{II} in both compounds, with each juglone moiety carrying a formal charge of -1e - a result consistent with the synthetic protocol. Furthermore, the coordination symmetry of juglone was found to depend on solvent presence: compound **1** preserves distinct C-O and C=O bond character, while this differentiation is absent in solvent-free compound **2**.

MOF-74 compounds were investigated using both conventional crystallographic methods and anomalous dispersion refinement (ADR). The structures were successfully solved, and chemical analysis demonstrated ADR's capability to determine metal occupancy in bimetallic systems. The same principle applied to bimetallic MOFs can be readily extended to polymetallic systems, provided sufficient wavelength options are available. At the Laboratory of Crystallography at the Federal University of Minas Gerais (LabCri), two wavelengths (CuKα and MoKα) are routinely accessible. However, synchrotron facilities offer substantially greater wavelength availability, enabling optimization for specific experimental requirements.

This work demonstrates that crystallography remains as powerful as ever and continues to occupy a forefront position in scientific research, now incorporating to conventional single crystal XRD analysis spectroscopic information that was previously accessible only through specialized synchrotron measurements.

As future perspectives, several research directions merit investigation. First, the solvent's role in the crystallization of $Co(Jug)_2(\text{Pyridine-based})_2$ compounds can be more thorough examined, as the presence/absence of solvent is determining to the presence of valence tautomerism in other compounds. Additionally, the electronic and magnetic properties of these compounds could be studied at lower temperatures using techniques such as electron paramagnetic resonance (EPR) as a transition may still happen at lower temperatures.

Second, the same analytical approach applied to bimetallic MOF-74 could be applied to other MOFs and materials, particularly those with superior long-range order, to compare results obtained with CuKα and MoKα radiation. Such studies could be extended to systems with three-metal substitutions, where the technique should, in principle, remain applicable for occupancy determination. A fundamental question concerns why the refinement yields physically reasonable results for one metal center but not for another in the same compound and if that could be explained by inherent limitations of the independent atom model.

Bibliography

[1] Bijelic, Aleksandar and Annette Rompel: *Polyoxometalates: more than a phasing tool in protein crystallography.* ChemTexts, 4(3):10, August 2018, ISSN 2199-3793. https://doi.org/10.1007/s40828-018-0064-1, visited on 2025-09-29. Cited 2 times in pages 9 and 14.

[2] Dolomanov, O. V., L. J. Bourhis, R. J. Gildea, J. a. K. Howard, and H. Puschmann: *OLEX2: a complete structure solution, refinement and analysis program.* Journal of Applied Crystallography, 42(2):339–341, April 2009, ISSN 0021-8898. http://scripts.iucr.org/cgi-bin/paper?kk5042, visited on 2022-08-15, Number: 2 Publisher: International Union of Crystallography. Cited 3 times in pages 9, 16, and 39.

[3] Königlich Preussische Akademie der Wissenschaften zu Berlin: *Abhandlungen der physikalischen Klasse der Königlich-Preussischen Akademie der Wissenschaften.* Berlin : Realschul-Buchhandlung, 1816. http://archive.org/details/abhandlungenderp16akad, visited on 2025-01-10. Cited in page 14.

[4] Miller, William Hallowes: *A Treatise on Crystallography.* For J. & J. J. Deighton, 1839, ISBN 978-0-598-98499-9. Google-Books-ID: MDcAAAAAQAAJ. Cited in page 14.

[5] Laue, Max von: *Theoretical section" from "interference phenomena for x-rays" (1912.* German History Intersections, 1912. Cited in page 14.

[6] Ewald, P. P.: *Laue's Discovery of X-ray Diffraction by Crystals.* In Ewald, P. P. (editor): *Fifty Years of X-Ray Diffraction: Dedicated to the International Union of Crystallography on the Occasion of the Commemoration Meeting in Munich July 1962*, pages 31–56. Springer US, Boston, MA, 1962, ISBN 978-1-4615-9961-6. https://doi.org/10.1007/978-1-4615-9961-6_4, visited on 2025-01-10. Cited in page 14.

[7] Bragg, W. H. and W. L. Bragg: *The Reflection of X-rays by Crystals.* Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 88(605):428–438, July 1913, ISSN 1364-5021, 1471-2946. http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/rspa.1913.0040, visited on 2018-04-23. Cited in page 15.

[8] Furukawa, Hiroyasu, Kyle E. Cordova, Michael O'Keeffe, and Omar M. Yaghi: *The Chemistry and Applications of Metal-Organic Frameworks.* Science, 341(6149):1230444,

August 2013, ISSN 0036-8075, 1095-9203. https://www.science.org/doi/10.1126/science.1230444, visited on 2024-11-07. Cited in page 15.

[9] Einsle, Oliver, Susana L. A. Andrade, Holger Dobbek, Jacques Meyer, and Douglas C. Rees: *Assignment of Individual Metal Redox States in a Metalloprotein by Crystallographic Refinement at Multiple X-ray Wavelengths*. Journal of the American Chemical Society, 129(8):2210–2211, February 2007, ISSN 0002-7863, 1520-5126. http://pubs.acs.org/doi/abs/10.1021/ja067562o, visited on 2016-12-16. Cited in page 15.

[10] Rajashankar, Kanagalaghatta and Zbigniew Dauter: *Data Collection for Crystallographic Structure Determination*. Methods in molecular biology (Clifton, N.J.), 1140:211–237, 2014, ISSN 1064-3745. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7670882/, visited on 2023-09-19. Cited in page 15.

[11] Schoenlein, R. W.: *Generation of Femtosecond Pulses of Synchrotron Radiation*. Science, 287(5461):2237–2240, March 2000, ISSN 00368075, 10959203. http://www.sciencemag.org/cgi/doi/10.1126/science.287.5461.2237, visited on 2019-07-31. Cited in page 15.

[12] Sheldrick, George M.: *SHELXT – Integrated space-group and crystal-structure determination*. Acta Crystallographica Section A Foundations and Advances, 71(1):3–8, January 2015, ISSN 2053-2733. http://scripts.iucr.org/cgi-bin/paper?S2053273314026370, visited on 2016-06-17. Cited 2 times in pages 15 and 37.

[13] Palatinus, Lukáš: *The charge-flipping algorithm in crystallography*. Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials, 69(1):1–16, February 2013, ISSN 2052-5192. http://scripts.iucr.org/cgi-bin/paper?S2052519212051366, visited on 2014-06-26. Cited in page 15.

[14] Crystallography, Online Dictionary of: *Law of rational indices - Online Dictionary of Crystallography*. https://dictionary.iucr.org/Law_of_rational_indices, visited on 2025-01-10. Cited in page 15.

[15] Crystallography, International Union of: *Report of the Executive Committee for 1991*. Acta Crystallographica Section A Foundations of Crystallography, 48(6):922–946, November 1992, ISSN 01087673. https://journals.iucr.org/paper?S0108767392008328, visited on 2025-01-10. Cited in page 15.

[16] Hahn, Theo and Internationale Union für Kristallographie (editors): *International tables for crystallography. Vol. A: Space-group symmetry*. Springer, Dordrecht, 5. ed., reprinted with corrections edition, 2005, ISBN 978-0-7923-6590-7. OCLC: 255316244. Cited 5 times in pages 16, 18, 19, 21, and 25.

[17] Giacovazzo, Carmelo, Hugo Luis Monaco, Gilberto Artioli, Davide Viterbo, Marco Milanesio, Gastone Gilli, Paola Gilli, Giuseppe Zanotti, Giovanni Ferraris, and and Michele Catti Carmelo Giacovazzo: *Fundamentals of Crystallography*. International Union of Crystallography Texts on Crystallography. Oxford University Press, Oxford, New York, third edition, third edition edition, June 2011, ISBN 978-0-19-957365-3. Cited 3 times in pages 23, 31, and 33.

[18] Yano, Junko and Vittal K. Yachandra: *X-ray absorption spectroscopy*. Photosynthesis Research, 102(2-3):241–254, December 2009, ISSN 1573-5079. Cited in page 28.

[19] Als-Nielsen, Jens and Des McMorrow: *Elements of Modern X-ray Physics*. Wiley, 1st edition, March 2011, ISBN 978-0-470-97395-0 978-1-119-99836-5. `https://onlinelibrary.wiley.com/doi/book/10.1002/9781119998365`, visited on 2020-04-28. Cited in page 28.

[20] Sheldrick, G. M.: *Classical direct methods*, 2008. `https://shelx.uni-goettingen.de/tutorials.php`. Cited in page 36.

[21] Sheldrick, George M: *A short history of SHELX*. Acta Crystallogr., A, Found. Crystallogr., 64(Pt 1):112–122, 2008. Cited in page 37.

[22] Sheldrick, G. M. and R. O. Gould: *Structure solution by iterative peaklist optimization and tangent expansion in space group P1*. Acta Crystallographica Section B: Structural Science, 51(4):423–431, August 1995, ISSN 0108-7681. `https://journals.iucr.org/b/issues/1995/04/00/hr0021/`, visited on 2025-01-13, Publisher: International Union of Crystallography. Cited in page 37.

[23] Sheldrick, G. M.: *Shelxt: Integrating space group determination and structure solution*, 2015. `https://shelx.uni-goettingen.de/tutorials.php`. Cited in page 37.

[24] Caliandro, R., B. Carrozzini, G. L. Cascarano, L. De Caro, C. Giacovazzo, and D. Siliqi: *Advances in the free lunch method*. Journal of Applied Crystallography, 40(5):931–937, October 2007, ISSN 0021-8898. `https://journals.iucr.org/j/issues/2007/05/00/he5370/`, visited on 2025-01-13, Publisher: International Union of Crystallography. Cited in page 38.

[25] Sheldrick, G. M.: *Crystal structure refinement with SHELXL*. Acta Crystallographica Section C: Structural Chemistry, 71(1):3–8, January 2015, ISSN 2053-2296. `http://scripts.iucr.org/cgi-bin/paper?fa3356`, visited on 2016-02-18. Cited in page 39.

[26] Bourhis, Luc J., Oleg V. Dolomanov, Richard J. Gildea, Judith A. K. Howard, and Horst Puschmann: *The anatomy of a comprehensive constrained, restrained refinement program for the modern computing environment – Olex2 dissected*. Acta

Crystallographica Section A Foundations and Advances, 71(1):59–75, January 2015, ISSN 2053-2733. https://journals.iucr.org/paper?S2053273314022207, visited on 2025-01-19. Cited 2 times in pages 39 and 48.

[27] Kleemiss, Florian, Oleg V. Dolomanov, Michael Bodensteiner, Norbert Peyerimhoff, Laura Midgley, Luc J. Bourhis, Alessandro Genoni, Lorraine A. Malaspina, Dylan Jayatilaka, John L. Spencer, Fraser White, Bernhard Grundkötter-Stock, Simon Steinhauer, Dieter Lentz, Horst Puschmann, and Simon Grabowsky: *Accurate crystal structures and chemical properties from NoSpherA2*. Chemical Science, 12(5):1675–1692, February 2021, ISSN 2041-6539. https://pubs.rsc.org/en/content/articlelanding/2021/sc/d0sc05526c, visited on 2024-07-30, Publisher: The Royal Society of Chemistry. Cited 3 times in pages 40, 42, and 44.

[28] Woińska, Magdalena, Simon Grabowsky, Paulina M. Dominiak, Krzysztof Woźniak, and Dylan Jayatilaka: *Hydrogen atoms can be located accurately and precisely by x-ray crystallography*. Science Advances, 2(5):e1600192, May 2016, ISSN 2375-2548. https://www.science.org/doi/10.1126/sciadv.1600192, visited on 2024-07-30. Cited in page 40.

[29] Hirshfeld, F. L.: *Bonded-atom fragments for describing molecular charge densities*. Theoretica chimica acta, 44(2):129–138, June 1977, ISSN 1432-2234. https://doi.org/10.1007/BF00549096, visited on 2025-01-29. Cited in page 40.

[30] Spackman, Mark A. and Dylan Jayatilaka: *Hirshfeld surface analysis*. CrystEngComm, 11(1):19–32, 2009, ISSN 1466-8033. https://xlink.rsc.org/?DOI=B818330A, visited on 2024-10-21. Cited in page 41.

[31] Mackenzie, C. F., P. R. Spackman, D. Jayatilaka, and M. A. Spackman: *CrystalExplorer model energies and energy frameworks: extension to metal coordination compounds, organic salts, solvates and open-shell systems*. IUCrJ, 4(5):575–587, September 2017, ISSN 2052-2525. https://journals.iucr.org/m/issues/2017/05/00/lc5090/, visited on 2024-05-20, Publisher: International Union of Crystallography. Cited in page 41.

[32] Jayatilaka, Dylan and Daniel J. Grimwood: *Tonto: A Fortran Based Object-Oriented System for Quantum Chemistry and Crystallography*. In Sloot, Peter M. A., David Abramson, Alexander V. Bogdanov, Yuriy E. Gorbachev, Jack J. Dongarra, and Albert Y. Zomaya (editors): *Computational Science — ICCS 2003*, pages 142–151, Berlin, Heidelberg, 2003. Springer, ISBN 978-3-540-44864-8. Cited in page 43.

[33] Brüx, Daniel, Ben Ebel, Niklas Pelzer, Irmgard Kalf, and Florian Kleemiss: *Experimental Spin State Determination of Iron(II) Complexes by Hirshfeld Atom Refinement*. Chemistry – A European Journal, 31(14):e202404017, 2025, ISSN 1521-3765. https:

//onlinelibrary.wiley.com/doi/abs/10.1002/chem.202404017, visited on 2025-03-10, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/chem.202404017. Cited in page 43.

[34] Hönl, H.: *Zur Dispersionstheorie der Röntgenstrahlen.* Zeitschrift für Physik, 84(1):1–16, January 1933, ISSN 0044-3328. https://doi.org/10.1007/BF01330269, visited on 2025-02-20. Cited in page 45.

[35] Cromer, Don T. and David Liberman: *Relativistic Calculation of Anomalous Scattering Factors for X Rays.* The Journal of Chemical Physics, 53(5):1891–1898, September 1970, ISSN 0021-9606. https://doi.org/10.1063/1.1674266, visited on 2025-02-20. Cited in page 45.

[36] Albinati, A.: *International Tables for Crystallography Volume C: Mathematical, physical and chemical tables.* In Prince, E. (editor): *International Tables for Crystallography*, volume C. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2006. Cited 2 times in pages 45 and 62.

[37] Sasaki, Satoshi.: *Numerical tables of anomalous scattering factors calculated by the Cromer and Liberman's method.* Technical report, National Laboratory for High Energy Physics, Japan, 1989. KEK–88-14 INIS Reference Number: 20081580. Cited in page 45.

[38] Henke, B. L., E. M. Gullikson, and J. C. Davis: *X-Ray Interactions: Photoabsorption, Scattering, Transmission, and Reflection at E = 50-30,000 eV, Z = 1-92.* Atomic Data and Nuclear Data Tables, 54(2):181–342, July 1993, ISSN 0092-640X. https://www.sciencedirect.com/science/article/pii/S0092640X83710132, visited on 2024-07-30. Cited in page 45.

[39] Meurer, Florian, Oleg V. Dolomanov, Christoph Hennig, Norbert Peyerimhoff, Florian Kleemiss, Horst Puschmann, and Michael Bodensteiner: *Refinement of anomalous dispersion correction parameters in single-crystal structure determinations.* IUCrJ, 9(5):604–609, September 2022, ISSN 2052-2525. https://scripts.iucr.org/cgi-bin/paper?S2052252522006844, visited on 2023-09-17. Cited 2 times in pages 45 and 60.

[40] Masek, Anna, Ewa Chrzescijanska, Malgorzata Latos-Brozio, and Marian Zaborski: *Characteristics of juglone (5-hydroxy-1,4,-naphthoquinone) using voltammetry and spectrophotometric methods.* Food Chemistry, 301:125279, December 2019, ISSN 0308-8146. https://www.sciencedirect.com/science/article/pii/S0308814619313895, visited on 2024-10-04. Cited in page 47.

[41] Juang, Yu Pu, Ju Ying Tsai, Wan Lan Gu, Hui Ching Hsu, Chao Lung Lin, Chin Chung Wu, and Pi Hui Liang: *Discovery of 5-Hydroxy-1,4-naphthoquinone (Juglone)*

Derivatives as Dual Effective Agents Targeting Platelet-Cancer Interplay through Protein Disulfide Isomerase Inhibition. Journal of Medicinal Chemistry, 67(5):3626–3642, March 2024, ISSN 0022-2623. https://doi.org/10.1021/acs.jmedchem.3c02107, visited on 2024-05-17, Publisher: American Chemical Society. Cited in page 47.

[42] Munir, Shamsa, Afzal Shah, Abdur Rauf, Amin Badshah, Hidayat Hussain, Zia-ur-Rehman, and Zahoor Ahmad: *Redox behavior of juglone in buffered aq.: Ethanol media.* Comptes Rendus. Chimie, 16(12):1140–1146, July 2013, ISSN 1878-1543. https://comptes-rendus.academie-sciences.fr/chimie/articles/10.1016/j.crci.2013.04.011/, visited on 2024-05-20. Cited in page 47.

[43] Bolton, Judy L. and Tareisha Dunlap: *Formation and Biological Targets of Quinones: Cytotoxic versus Cytoprotective Effects.* Chemical Research in Toxicology, 30(1):13–37, January 2017, ISSN 0893-228X, 1520-5010. https://pubs.acs.org/doi/10.1021/acs.chemrestox.6b00256, visited on 2024-05-17. Cited in page 47.

[44] Cradwick, P. D. and D. Hall: *The crystal structure of juglone.* Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry, 27(7):1468–1470, July 1971, ISSN 0567-7408. https://journals.iucr.org/b/issues/1971/07/00/a08254/, visited on 2024-05-23, Publisher: International Union of Crystallography. Cited 2 times in pages 47 and 49.

[45] Mukherjee, T.: *One-electron reduction of juglone (5-hydroxy-1,4-naphthoquinone): A pulse radiolysis study.* International Journal of Radiation Applications and Instrumentation. Part C. Radiation Physics and Chemistry, 29(6):455–462, January 1987, ISSN 1359-0197. https://www.sciencedirect.com/science/article/pii/1359019787900245, visited on 2021-06-20. Cited in page 47.

[46] Salunke-Gawali, Sunita, Eulália Pereira, Umar Ali Dar, and Sujit Bhand: *Metal complexes of hydroxynaphthoquinones: Lawsone, bis-lawsone, lapachol, plumbagin and juglone.* Journal of Molecular Structure, 1148:435–458, November 2017, ISSN 0022-2860. https://www.sciencedirect.com/science/article/pii/S0022286017309110, visited on 2024-11-07. Cited in page 47.

[47] Pawar, A B, S R Bamane, K D Jadhav, and R G Sarawadekar: *SPECTRAL, THERMAL, X-RAY DIFFRACTION AND ANTIMICROBIAL STUDIES OF SOME BIVALENT METAL CHELATES OF JUGLONE.* Thermochimica Acta, 2012. Cited in page 47.

[48] Sawhney, S. S., Rakesh M. Sati, and Surya K. Chandel: *Kinetic analysis of derivative thermogravimetric data of 5-hydroxy-1,4-naphthoquinone (juglone) and some of its divalent metal chelates.* Thermochimica Acta, 55(3):363–365, June 1982, ISSN 0040-6031.

https://www.sciencedirect.com/science/article/pii/004060318285051X, visited on 2024-11-07. Cited in page 47.

[49] CrysalisPro: *Rigaku Oxford Diffraction, CrysAlisPro Software system*, 2015. Cited 3 times in pages 48, 59, and 60.

[50] Clark, R. C. and J. S. Reid: *The analytical calculation of absorption in multifaceted crystals.* Acta Crystallographica Section A: Foundations of Crystallography, 51(6):887–897, November 1995, ISSN 0108-7673. http://scripts.iucr.org/cgi-bin/paper?sh0068, visited on 2016-02-18. Cited 2 times in pages 48 and 60.

[51] Johnson, C. K.: *Introduction to Thermal- Motion Analysis.* In Ahmed, F. R., S. R. Hall, and C. P. Huber (editors): *Crystallographic Computing Proceedings of an international Summer School*, pages 207–219. Munksgaard, Copenhagen, Ottawa, 1970. Cited in page 48.

[52] Macrae, C. F., I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, and P. A. Wood: *Mercury CSD 2.0 – new features for the visualization and investigation of crystal structures.* Journal of Applied Crystallography, 41(2):466–470, April 2008, ISSN 0021-8898. http://scripts.iucr.org/cgi-bin/paper?dd5037, visited on 2016-02-18. Cited in page 49.

[53] Alvarez, Santiago: *Polyhedra in (inorganic) chemistry.* Dalton Transactions, 13(13):2209, 2005, ISSN 1477-9226, 1477-9234. https://xlink.rsc.org/?DOI=b503582c, visited on 2024-05-23. Cited 2 times in pages 49 and 51.

[54] Rohl, Andrew L., Massimo Moret, Werner Kaminsky, Kacey Claborn, Joshua J. McKinnon, and Bart Kahr: *Hirshfeld Surfaces Identify Inadequacies in Computations of Intermolecular Interactions in Crystals: Pentamorphic 1,8-Dihydroxyanthraquinone.* Crystal Growth & Design, 8(12):4517–4525, December 2008, ISSN 1528-7483. https://doi.org/10.1021/cg8005212, visited on 2024-05-20, Publisher: American Chemical Society. Cited in page 51.

[55] Allen, F. H.: *The Cambridge Structural Database: a quarter of a million crystal structures and rising.* Acta Crystallographica Section B: Structural Science, 58(3):380–388, June 2002, ISSN 0108-7681. http://scripts.iucr.org/cgi-bin/paper?an0610, visited on 2016-02-18. Cited in page 55.

[56] Ribeiro, Marcos A., Daniel E. Stasiw, Philip Pattison, Paul R. Raithby, David A. Shultz, and Carlos B. Pinheiro: *Towards Controlling the Solid State Valence Tautomeric Interconversion Character by Solvation.* Crystal Growth & Design, March 2016, ISSN 1528-7483. http://dx.doi.org/10.1021/acs.cgd.6b00159, visited on 2016-03-19. Cited in page 55.

[57] Mukherjee, Rabindranath: *Assigning Ligand Redox Levels in Complexes of 2-Aminophenolates: Structural Signatures*. Inorganic Chemistry, 59(18):12961–12977, September 2020, ISSN 0020-1669, 1520-510X. https://pubs.acs.org/doi/10.1021/acs.inorgchem.0c00240, visited on 2025-01-07. Cited in page 56.

[58] Rowsell, Jesse L C and Omar M Yaghi: *Metal-organic frameworks: a new class of porous materials*. Microporous and Mesoporous Materials, 2004. Cited in page 57.

[59] Zhou, Hong Cai, Jeffrey R. Long, and Omar M. Yaghi: *Introduction to Metal–Organic Frameworks*. Chemical Reviews, 112(2):673–674, February 2012, ISSN 0009-2665. https://doi.org/10.1021/cr300014x, visited on 2025-07-04, Publisher: American Chemical Society. Cited in page 57.

[60] Chen, Zhijie, Kent O. Kirlikovali, Le Shi, and Omar K. Farha: *Rational design of stable functional metal–organic frameworks*. Materials Horizons, 10(9):3257–3268, 2023, ISSN 2051-6347, 2051-6355. https://xlink.rsc.org/?DOI=D3MH00541K, visited on 2024-11-18. Cited in page 57.

[61] Andrade, Pedro H.M., Hugo Palhares, Christophe Volkringer, Thierry Loiseau, Matthieu Hureau, Eduardo Nunes, and Alain Moissette: *State of the art in visible-light photocatalysis of aqueous pollutants using metal-organic frameworks*. Journal of Photochemistry and Photobiology C: Photochemistry Reviews, 57:100635, December 2023, ISSN 13895567. https://linkinghub.elsevier.com/retrieve/pii/S1389556723000667, visited on 2025-01-29. Cited in page 57.

[62] Oktavian, Rama, Ruben Goeminne, Lawson T. Glasby, Ping Song, Racheal Huynh, Omid Taheri Qazvini, Omid Ghaffari-Nik, Nima Masoumifard, Joan L. Cordiner, Pierre Hovington, Veronique Van Speybroeck, and Peyman Z. Moghadam: *Gas adsorption and framework flexibility of CALF-20 explored via experiments and simulations*. Nature Communications, 15(1):3898, May 2024, ISSN 2041-1723. https://www.nature.com/articles/s41467-024-48136-0, visited on 2025-01-14. Cited in page 57.

[63] Gonzalez, Miguel I., Jarad A. Mason, Eric D. Bloch, Simon J. Teat, Kevin J. Gagnon, Gregory Y. Morrison, Wendy L. Queen, and Jeffrey R. Long: *Structural characterization of framework–gas interactions in the metal–organic framework Co2(dobdc) by in situ single-crystal X-ray diffraction*. Chemical Science, 8(6):4387–4398, May 2017, ISSN 2041-6539. https://pubs.rsc.org/en/content/articlelanding/2017/sc/c7sc00449d, visited on 2019-11-18. Cited in page 57.

[64] Engel, Emile R. and Janet L. Scott: *Advances in the green chemistry of coordination polymer materials*. Green Chemistry, 22(12):3693–3715, June 2020, ISSN 1463-9270. https://pubs.rsc.org/en/content/articlelanding/2020/gc/

d0gc01074j, visited on 2025-03-31, Publisher: The Royal Society of Chemistry. Cited in page 57.

[65] Hsu, Pei Hsiang, Chien Chun Chang, Tsu Hao Wang, Phuc Khanh Lam, Ming Yu Wei, Ching Tien Chen, Chin Yu Chen, Lien Yang Chou, and Fa Kuen Shieh: *Rapid Fabrication of Biocomposites by Encapsulating Enzymes into Zn-MOF-74 via a Mild Water-Based Approach*. ACS Applied Materials & Interfaces, 13(44):52014–52022, November 2021, ISSN 1944-8244. https://doi.org/10.1021/acsami.1c09052, visited on 2025-03-31, Publisher: American Chemical Society. Cited 2 times in pages 57 and 58.

[66] Liao, Yuxing, Chao Li, Xiaobing Lou, Peng Wang, Qi Yang, Ming Shen, and Bingwen Hu: *Highly reversible lithium storage in cobalt 2,5-dioxido-1,4-benzenedicarboxylate metal-organic frameworks boosted by pseudocapacitance*. Journal of Colloid and Interface Science, 506:365–372, November 2017, ISSN 0021-9797. https://www.sciencedirect.com/science/article/pii/S0021979717308317, visited on 2025-07-05. Cited in page 58.

[67] Dong, Zhongtian, Zhiren Zhao, Fengyun Wang, Fenghe Wang, Mingzhu Xia, and Chao Ding: *Innovative application and micro-catalytic mechanism of MOF-74 derived Ni-Co bimetallic oxides in catalytic wet oxidation of high-concentration 2,4-dinitroanisole wastewater*. Applied Catalysis B: Environment and Energy, 367:125082, June 2025, ISSN 09263373. https://linkinghub.elsevier.com/retrieve/pii/S0926337325000657, visited on 2025-03-27. Cited in page 58.

[68] Gao, Ziyu, Lin Liang, Xiao Zhang, Ping Xu, and Jianmin Sun: *Facile One-Pot Synthesis of Zn/Mg-MOF-74 with Unsaturated Coordination Metal Centers for Efficient CO2 Adsorption and Conversion to Cyclic Carbonates*. ACS Applied Materials & Interfaces, 13(51):61334–61345, December 2021, ISSN 1944-8244. https://doi.org/10.1021/acsami.1c20878, visited on 2022-03-23, Publisher: American Chemical Society. Cited in page 58.

[69] Poneti, Giordano, Matteo Mannini, Lorenzo Sorace, Philippe Sainctavit, Marie Anne Arrio, Andrei Rogalev, Fabrice Wilhelm, and Andrea Dei: *X-ray Absorption Spectroscopy as a Probe of Photo- and Thermally Induced Valence Tautomeric Transition in a 1:1 Cobalt-Dioxolene Complex*. ChemPhysChem, 10(12):2090–2095, August 2009, ISSN 14394235, 14397641. http://doi.wiley.com/10.1002/cphc.200900098, visited on 2014-07-22. Cited in page 58.

[70] Rodríguez-Carvajal, Juan: *Recent advances in magnetic structure determination by neutron powder diffraction*. Physica B: Condensed Matter, 192(1-2):55–69, October 1993, ISSN 09214526. https://linkinghub.elsevier.com/retrieve/pii/092145269390108I, visited on 2020-10-24. Cited in page 58.

[71] Montes-Andrés, Helena, Gisela Orcajo, Carmen Martos, Juan A. Botas, and Guillermo Calleja: *Co/Ni mixed-metal expanded IRMOF-74 series and their hydrogen adsorption properties*. International Journal of Hydrogen Energy, 44(33):18205–18213, July 2019, ISSN 0360-3199. https://www.sciencedirect.com/science/article/pii/S036031991931835X, visited on 2022-07-25. Cited in page 58.

[72] Figueiredo, Juliana Assuncao Pereira de: *PRODUÇÃO E CARACTERIZAÇÃO DE REDES METALORGÂNICAS Me-MOF-74*. PhD thesis, Universidade Federal de Minas Gerais, 2024. Cited in page 59.

[73] Casaday, Claire E., Cristin E. Juda, Shao Liang Zheng, Simon J. Teat, Tieyan Chang, Yu Sheng Chen, and Theodore A. Betley: *Occupancy Determination from Resonant X-ray Diffraction*. Inorganic Chemistry, 64(6):3090–3100, February 2025, ISSN 0020-1669, 1520-510X. https://pubs.acs.org/doi/10.1021/acs.inorgchem.4c05423, visited on 2025-03-24. Cited in page 60.

[74] Christian, Gary D: *Atomic absorption spectroscopy for the determination of elements in medical biological samples*. In *Fortschritte der Chemischen Forschung*, pages 77–112. Springer-Verlag, Berlin/Heidelberg, 1980, ISBN 978-3-540-05589-1. http://link.springer.com/10.1007/BFb0051600, visited on 2025-07-08. Cited in page 61.

[75] Hu, Ben Yu-Kuang: *Kramers–Kronig in two lines*. American Journal of Physics, 57(9):821, September 1989, ISSN 0002-9505. https://doi.org/10.1119/1.15901, visited on 2025-02-17. Cited in page 79.

Appendix

APPENDIX A – X-ray Diffraction appendix

A.1 Kramers-Kronig Relations (KKR)

In this appendix we will derive the Kramers-Kronig relations. The route selected will be based on $\chi(\omega)$ being analytical in the upper-half of the complex plane.†

Assume a function $\chi(\omega')$ analytical in the upper-half of the complex plane. It follows that, by Cauchy's residue theorem, the integral over any closed path in the upper half of the complex plane of this function is equal to zero. Also from this, the closed path integral of any path avoiding the point $(\omega' = \omega)$ of the function $\frac{\chi(\omega')}{\omega'-\omega}$ is also equal to zero. One can then construct an integral following a path as shown in Figure 23, and the result of the integration is known a priori.

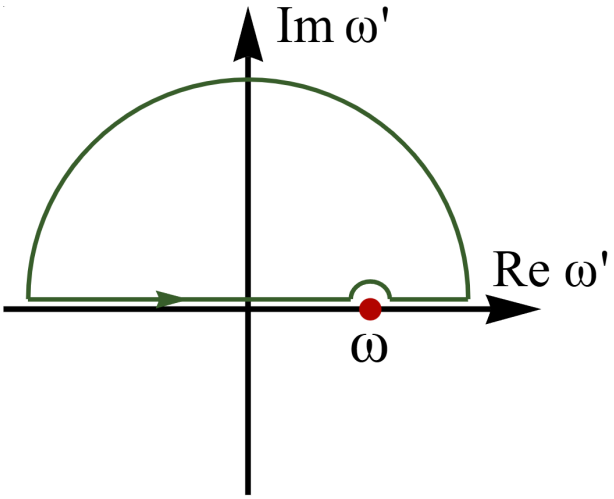


Figure 23 – Contour selected to derive the Kramers-Kronig relations

One can divide the region above in four paths. The first is a semi-circle on the infinite, the second is the region from $-\infty$ to $\omega - \epsilon$, the third is a semi-circle around the singularity at $\omega' = \omega$, and the fourth is the path from $\omega + \epsilon$ to $+\infty$. Since the function $\chi(\omega)$ goes to zero faster than $\frac{1}{|\omega'|}$, the integral over the semi-circle in infinite is zero. Using Cauchy's residue theorem, the integral over the semi-circle around ω is $-\pi i\chi(\omega)$. This process leaves only the integral over the real line, that can be written as:

$$\int_{-\infty}^{\omega-\epsilon} \frac{\chi(\omega')}{\omega' - \omega} d\omega' + \int_{\omega+\epsilon}^{\infty} \frac{\chi(\omega')}{\omega' - \omega} d\omega' \tag{A.1}$$

In the limit $\epsilon \to 0$, equation A.1 is the definition of Cauchy's Principal Value integral.

† Another deduction, explored by Hu [75], presents itself as more intuitive. However the time space equivalent of absorption and scattering is not obviously intuitive, so this route was not selected.

Uniting all those results we arrive at:

$$\chi(\omega) = \frac{1}{i\pi} \fint_{-\infty}^{+\infty} \frac{\chi(\omega')}{\omega' - \omega} d\omega' \tag{A.2}$$

Where $\fint$ represents the principal value integral of the function. One can now open in the real and complex parts of $\chi(\omega)$ as:

$$\chi_1(\omega) + i\chi_2(\omega) = -\frac{i}{\pi} \fint_{-\infty}^{+\infty} \frac{\chi_1(\omega') + i\chi_2(\omega)}{\omega' - \omega} d\omega'$$

$$\chi_1(\omega) = \frac{1}{\pi} \fint_{-\infty}^{+\infty} \frac{\chi_2(\omega')}{\omega' - \omega} d\omega' \tag{A.3}$$

$$\chi_2(\omega) = -\frac{1}{\pi} \fint_{-\infty}^{+\infty} \frac{\chi_1(\omega')}{\omega' - \omega} d\omega' \tag{A.4}$$

Equations A.3 and A.4 are already capable of relating the real and imaginary parts of complex functions. But to arrive at the recognized form of the KKR one needs to consider the physical nature of the functions. Assuming causality, those functions in complex frequency space must be the Fourier transform of positive-defined real time space functions. This restriction means that $\chi(-\omega) = \chi^*(\omega)$. Due to that $\chi_1(\omega)$ is necessarily an even function while $\chi_2(\omega)$ is an odd function. Applying those relations to Equations A.3 and A.4 we arrive at:

$$\chi_1(\omega) = \frac{1}{\pi} \left(\fint_{-\infty}^{+\infty} \frac{\omega' \chi_2(\omega')}{\omega'^2 - \omega^2} d\omega' + \fint_{-\infty}^{+\infty} \frac{\omega \chi_2(\omega')}{\omega'^2 - \omega^2} d\omega' \right)$$

$$\chi_1(\omega) = \frac{2}{\pi} \fint_{0}^{+\infty} \frac{\omega' \chi_2(\omega')}{\omega'^2 - \omega^2} d\omega' \tag{A.5}$$

$$\chi_2(\omega) = -\frac{2\omega}{\pi} \fint_{0}^{+\infty} \frac{\chi_1(\omega')}{\omega'^2 - \omega^2} d\omega' \tag{A.6}$$

Equations A.5 and A.6 are the Kramers-Kronig Relations (KKR), useful in a myriad of fields spanning all across optics and signal processing.

A.2 Structure invariants

Direct Methods try to determine the structure directly from the measured intensities, as seen in Subsection 2.5.2. The absolute value of phases is generally dependent on the origin choice, so it is important for direct methods that we define constructs that are invariant to origin choice and symmetry operations. Let us consider a product of F_h with the form:

$$F_{h_1} F_{h_2} \dots F_{h_n} = |F_{h_1} F_{h_2} \dots F_{h_n}| \exp[i(\varphi_1 + \varphi_2 + \dots + \varphi_n)]$$

Let us find a condition in which the product above does not vary with a origin change. When the origin is changed by a vector $\vec{r}_0$, one can write the new structure Factor

F'_h as:

$$\begin{aligned} F'_h &= \sum_j f_j \exp(2\pi i \vec{h} \cdot (\vec{r}_j - \vec{r}_0)) \\ F'_h &= F_h \exp(2\pi i \vec{h} \cdot \vec{r}_0) \\ F'_h &= |F_h| \exp(\varphi_h - 2\pi \vec{h} \cdot \vec{r}_j) \end{aligned} \tag{A.7}$$

This means that, when combining this new phases, one can expect, from equation A.7, a phase difference of:

$$\Delta\varphi = 2\pi \vec{r}_0 \cdot \sum_n \vec{h}_n \tag{A.8}$$

This variation is, in general, 0 when $\sum_n \vec{h}_n = 0$. Any construction in which $\Delta\varphi = 0$ can be called a structure invariant. Those are extremely important constructions in the theory of direct methods. In principle, any number of Bragg peaks can be selected to respect the invariability criteria, in practice only selections of three (called triplet invariants) and of four (called quartet invariants) peaks are used in most direct method algorithms. The simplest structure invariants are:

- $F_{000} = \sum_j Z_j$ is the number of electrons in the unit cell.
- $F_h F_{-h} = |F_h|^2$, the doublet invariant completely gets rid of the phase information.
- $F_{-h} F_k F_{h-k}$, the so-called triplet invariant has phase $\varphi_{-h} + \varphi_k + \varphi_{h-k}$
- $F_h F_k F_l F_{-h-k-l}$, is the quartet invariant. It was responsible for permitting the use of a greater part of the measured intensities in structure solution.

Another important construct is the idea of structure semi-invariant. These are mathematical constructs that respects $\Delta\varphi = 0$ for origin shifts between points of same symmetry. In general, for a space group possessing the symmetry operator $\mathbf{C} \equiv (\mathbf{R}, \mathbf{T})$, when $\psi = \varphi_{\vec{H}} - \varphi_{\vec{h}} + \varphi_{\vec{h}\mathbf{R}}$ it is possible to find a peak $\vec{h}$ that respects, $\vec{H} - \vec{h} + h\vec{\mathbf{R}} = 0$. When that is the case, ψ can be written:

$$\psi = \varphi_{\vec{H}} - 2\pi \vec{h} \cdot \mathbf{T} \tag{A.9}$$

In general, $\varphi_{\vec{H}}$ and $\mathbf{T}$ depend on the choice of origin, but in points of equal symmetry, $\mathbf{T}$ is constant, since ψ was defined to be invariant to origin selection, $\varphi_{\vec{H}}$ will also be. Equation A.9 defines the structure semi-invariant criteria. For a practical example, in space group $P\bar{1}$ $\varphi_{2\vec{h}}$ is a structure semi-invariant with the form:

$$\psi = \varphi_{2h} - \varphi_h - \varphi_{-h} \tag{A.10}$$

When the origin is conveniently chosen to be on the inversion center, there are eight total points where the origin while still respecting A.10. Structure semi-invariant helps on extending the number of usable reflections in direct method solutions, due to respecting the same probability equations as structure invariants, baring a adequate origin selection.

A.3 Probability Distribution

As discussed in sub-section 2.5.2, the conditional probability distribution of φ_h is of the outmost importance for many direct methods. To construct the probability distribution function, firstly one must construct the normalized structure factor $|E_h|$ as in equation A.11:

$$\begin{aligned} E_h &= \frac{1}{\sqrt{N}} \sum_j \exp(2\pi i \vec{h} \cdot \vec{r}_j) \\ E_h &= \frac{1}{\sqrt{N}} \sum_j \exp(2\pi i \vec{k} \cdot \vec{r}_j) \exp(2\pi i (\vec{h} - \vec{k}) \cdot \vec{r}_j) \\ E_h &= \frac{1}{\sqrt{N}} \sum_j \xi_j \end{aligned} \tag{A.11}$$

Assuming that every point in the unit cell has the same probability of containing an atom, $\vec{r}_j$ is a conjunct of independent variables, this means that ξ_j is also a conjunct of independent variables. The central limit theory can than be applied to such a distribution and we arive at:

$$\langle E_h \rangle = \frac{1}{\sqrt{N}} \sum_j \langle \xi_j \rangle \tag{A.12}$$

And variance given by:

$$\sigma_{\xi_j} = \frac{1}{N} \sum_j \left(\langle \xi_j^2 \rangle - \langle \xi_j \rangle^2 \right) \tag{A.13}$$

Assuming the two exponential in A.11 to be independent we have:

$$\langle \xi_j \rangle = \langle \exp(2\pi i \vec{k} \cdot \vec{r}_j) \rangle \langle \exp(2\pi i h \vec{-} k \cdot \vec{r}_j) \rangle \tag{A.14}$$

We now assume that we know both the intensity and the phase of the two structure factors E_k and E_{h-k} one can write:

$$E_k = \sqrt{N} \langle \exp(2\pi i \vec{k} \cdot \vec{r}) \rangle \tag{A.15}$$

$$\langle \xi_j \rangle = \frac{E_k E_{h-k}}{N} \tag{A.16}$$

Using identity A.16 on equation A.12 one arrives at:

$$\langle E_h \rangle = \frac{E_k E_{h-k}}{\sqrt{N}} \tag{A.17}$$

Writing E_h as $A_h + iB_h$ one arrives at:

$$\langle A_h \rangle = \frac{1}{\sqrt{N}} |E_k E_{h-k}| \cos(\varphi_k + \varphi_{h-k}) \tag{A.18}$$

$$\langle B_h \rangle = \frac{1}{\sqrt{N}} |E_k E_{h-k}| \sin(\varphi_k + \varphi_{h-k}) \tag{A.19}$$

Similarly to equation A.11 one can manipulate equation A.18 to get:

$$\begin{aligned} A_h &= \frac{1}{\sqrt{N}} \sum_j \cos(2\pi(\vec{h} - \vec{k} + \vec{k}) \cdot \vec{r}_j) \\ A_h &= \frac{1}{\sqrt{N}} \sum_j [C_j(k) C_j(h-k) - S_j(k) S_j(h-k)] \\ A_h &= \frac{1}{N} \sum_j \alpha_j \end{aligned} \tag{A.20}$$

Where $C_j(k) = \cos(2\pi \vec{k} \cdot \vec{r}_j)$ and $S_j(k) = \sin(2\pi \vec{k} \cdot \vec{r}_j)$. One can algebraically derive:

$$\begin{aligned} \langle \alpha_j^2 \rangle =& \langle C_j(k)^2 \rangle \langle C_j(h-k)^2 \rangle - 2 \langle C_j(k) \rangle \langle C_j(h-k) \rangle \\ & \langle S_j(k) \rangle \langle S_j(h-k) \rangle + \langle S_j(k)^2 \rangle \langle S_j(h-k)^2 \rangle \end{aligned} \tag{A.21}$$

$$\sigma_{A_h}^2 = \frac{1}{N} \sum_j \left(\langle \alpha_j^2 \rangle - \langle A_h \rangle^2 \right) \tag{A.22}$$

The conditional probability of A_h and B_h given the values for E_k and E_{h-k} is given by the product of the probability for each, given by:

$$P(A_h, B_h | E_h E_{h-k}) = \frac{1}{\pi} \exp\left(- (A_h - \langle A_h \rangle)^2 + (B_h - \langle B_h \rangle)^2 \right) \tag{A.23}$$

Changing to polar coordinates a factor $|E_h| dE_h d\varphi$ appears and we can write:

$$\begin{aligned} P(|E_h|, \varphi_h | E_k, E_{h-k}) = \frac{|E_h|}{\pi} \exp\Bigg(& - [|E_h| \cos\varphi_h - \frac{1}{\sqrt{N}} |E_k E_{h-k}| \cos(\varphi_h - \varphi_{h-k})]^2 + \\ & - [|E_h| \sin\varphi_h - \frac{1}{\sqrt{N}} |E_k E_{h-k}| \sin(\varphi_h + \varphi_{h-k})]^2 \Bigg) \\ P(|E_h|, \varphi_h | E_k, E_{h-k}) = \frac{|E_h|}{\pi} \exp & \left(- |E_h|^2 + \frac{1}{N} |E_k E_{h-k}|^2 \right) \\ & \exp\left(\frac{2}{\sqrt{N}} |E_{-h} E_k E_{h-k}| \cos(\varphi_{-h} - \varphi_k - \varphi_{h-k}) \right) \end{aligned} \tag{A.24}$$

Since $|E_h|$ was measured, and therefore known, Equation A.24 can be re-written by normalizing for the probability of only the unknown phase φ_h. Writing $G_{hk} = \frac{2}{\sqrt{N}} |E_{-h} E_k E_{h-k}|$ and $\Phi_{hk} = \varphi_{-h} - \varphi_k - \varphi_{h-k}$ it is possible to write:

$$P(\varphi_h | G_{hk}) = C \exp(G_{hk} \cos(\Phi_{hk})) \tag{A.25}$$

A.4 Least Squares Problem

Imagine a set of quantities, F, that can be linearly described by a variable set X. It is possible to write a matrix equation for this quantity as follows:

$$F = AX$$

where A is the so-called **design matrix** is assumed to be previously known. If F is a physical quantity a new problem arise, due to measurement limitation a perfect fit is not always possible by our equation and the best equation may be described by introducing an "error" term V at our previous equation as:

$$F = AX + V$$

This error term V is the distance from the measured value to the modeled one by AX, this means that a more natural way of writing the equation above is:

$$V = F - AX \tag{A.26}$$

Up until this point nothing has been said about the modeling process itself, only that it can be written as equation A.26. Now we will concern ourselves with the process of determining a variable set X that best fits the measured set F. Firstly we must assume that the errors found in set V are unbiased. This has two main purposes: the first one is physical as a biased error results from a poorly performed experiment; the second one is mathematical as the linear least squares is not the best estimator for biased samples, the James-Stein estimator is an example of biased estimator. Secondly the mean of V must equal zero and it needs to have finite second moment. If both assertions are true, it can be written that:

$$\langle F \rangle \equiv F^0 = AX$$

If all of the above is true, the Gauss-Markov theorem guarantees that the least-squares estimator gives the best fit. For the linear case another essential property is that the least-squares algorithm always converge to a global minima.

Given a value S given by $S = \sum_i v_i^2$, where $v_i = f_i - f_i^0$, the problem of least squares can be stated as finding the variable set X which minimizes S. This convergence can be achieved by looking at the gradient of S, given by:

$$\frac{\partial S}{\partial x_j} = 2\sum_j v_i \frac{\partial v_i}{\partial x_j} = -2\sum v_i \frac{\partial f_i^0}{\partial x_j} = 0 \tag{A.27}$$

This create an equation system with the same size as the original variable set X, analysis of this equation leads to a closed-form solution that can be written as:

$$(A^T A)^{-1} A^T F = X \tag{A.28}$$

This powerful tool is limited in use to linear problems as it is written. The first limitation is that this identity only applies to the linear least square problem. The second limitation has to to with the lack of restraint in relation to experimental imprecision. The second one can be treated by introducing experimental errors in the mix. It can be noted that our definition of S is equivalent to calculating the product : $S = V^T V$.

It is possible to define the **covariance matrix**, M, that is given by: $M = \langle (F - F^0)(F - F^0)^T \rangle$. The elements in the diagonal of M are given by σ^2, where sigma is the associated imprecision given by the experiment. A new estimator S can be calculated as $S = V^T M^{-1} V$. The values outside the diagonal also depend on the experimental error but also on a correlation parameter, which will not be discussed here.

For functions where the partial derivatives can be calculated the first problem can be addressed as the identity A.28 can be extended by using the Jacobian matrices. First we write the Taylor series of the variable f_i around each variable. This can be written as:

$$f_i(X) \approx \sum_{n,j} \frac{1}{n!} \left(\frac{\partial f_i^{(n)}(a)}{\partial x_j} \right)^0 (x - a)^n \tag{A.29}$$

where n goes from 0 to infinity and j covers each variable. Extracting the two first terms of identity A.29 it can be rearranged as:

$$\Delta f_i = \sum_j \left(\frac{\partial f_i(X^0)}{\partial x_j} \right)^0 \Delta x + v_i \tag{A.30}$$

where $\Delta f_i = f_i(X) - f_i(X^0)$ and $\Delta x = x_j - x_j^0$. v_i is an error parameter that now encloses every higher dimension error associated with the equation. The main changes from A.29 to A.30 are that now it is assumed the expansion around X^0 and the f_i is calculated at the new point. The parallels of equation A.30 and A.26 leads us to generalize the algorithm described above by simply replacing the design matrix A with the Jacobian matrix that assumes its role in A.30, which can be written as:

$$(J_r)_{ij} = \frac{\partial r_i(x^{(n)})}{\partial x_j} \tag{A.31}$$

A.4.1 The G-N and L-M algorithms

The process described above is the Gauss-Newton (G-N) algorithm for least squares problem. For non-linear problems the n-th iteration of x, denoted as $x^{(n)}$, in the G-N algorithm is given by:

$$x^{(n+1)} = x^{(n)} - (J_r^T J_r)^{-1} J_r^T r(x^{(n)}) \tag{A.32}$$

Equation A.32 can be manipulated to obtain:

$$\begin{aligned} A\Delta &= b \\ \Delta &= x^{(n+1)} - x^{(n)} \\ b &= -J_r^T r(x^{(n)}) \\ A &= J_r^T J_r \end{aligned} \tag{A.33}$$

Equation A.33 represents a common matrix equation solvable by various available methods.

The Levenberg-Marquardt algorithm can be derived by modifying Equation A.33 to:

$$(A + \lambda I)\Delta = b \tag{A.34}$$

where λ is known as the Marquadt parameter and no fixed algorithm is known for it's determination.