Cambridge University Press & Assessment
978-1-107-03355-9 — Theory of Inelastic Scattering and Absorption of X-rays
Michel van Veenendaal
Excerpt
More Information

1

Introduction

Although almost all quantum mechanics textbooks consider absorption and emission, the discussion is usually limited to hydrogen-like atoms. This gives a somewhat limited view of the process. This book deals with aborption of high-energy X-rays. Let us look at a particular example to demonstrate the concepts that we will be dealing with throughout the book. Figure 1.1 shows a calculation of the L-edge of a divalent cobalt ion in a solid. In X-ray terminology, L-edge stands for the excitation of an electron from a $2p$ orbital into a $3d$ one. In fact, the calculated spectrum resembles closely the experimental X-ray absorption on CoO. Simply by looking at this spectrum, a number of salient features are apparent that will hopefully become clearer throughout the book.

First, one can start by asking the very basic question: what is a spectrum? Apparently, it is the absorption intensity as a function of energy. This implies that energy is a good quantum number. Therefore, in an absorption process energy is conserved. We therefore have to understand why certain quantities are conserved. We shall see that this is inherently related to the symmetry properties of the system. It is important to understand these properties since they tell us what quantities are exchanged between the incoming photon field and the material. By studying the changes in the photon field (in this case, how many photons are absorbed by the material), we can learn something about the material. We shall see that energy is not the only quantity that can be exchanged between the photons and the atoms. Linear and angular momentum are other examples of quantities that can be conserved in a spectroscopy experiment.

We already identified the spectrum in Fig. 1.1 as arising from a $2p \rightarrow 3d$ transition. This implies that the basic atomic structure is still valid. We shall therefore review some aspects of the hydrogen-like atom in Chapter 2. The absolute energy scale for this transition is of the order of hundreds of electronvolts. This is directly related to the large binding energy of the $2p$ electron and the reason why X-rays are necessary to excite electrons from core-levels into the empty $3d$

Cambridge University Press & Assessment
978-1-107-03355-9 — Theory of Inelastic Scattering and Absorption of X-rays
Michel van Veenendaal
Excerpt
More Information

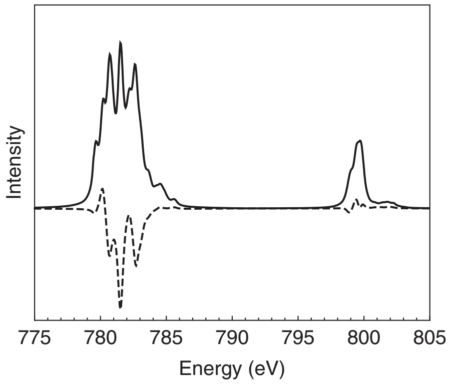

Figure 1.1 The isotropic (solid) and circular dichroic (dashed) spectra of divalent cobalt in an octahedral crystal field.

states. This is significantly larger than the 13.6 eV binding energy for the hydrogen atom. However, excitations within the $3d$ orbital are only of the order of a few electronvolts. These transitions can be observed using visible light. We shall not look at absorption in the visible region which is the subject of many other textbooks. However, we shall touch upon these transitions when discussing resonant inelastic X-ray scattering.

Another important aspect of spectroscopy is the strengths of the transitions between the initial and final states. This gives rise to the different intensities observed in the spectrum in Fig. 1.1. This is a surprisingly difficult topic. L-edges are dipolar transitions due to the $\mathbf{p} \cdot \mathbf{A}$ interaction, where $\mathbf{p}$ is the momentum of the electron that is excited and $\mathbf{A}$ the vector potential describing the X-ray field. Therefore, the different components of the vector potential couple to different directions of the momentum of the electron. For dipolar transitions, there are three components. However, in addition the p and d orbitals are six- and tenfold degenerate, respectively. This gives a total of 180 matrix elements for the single-particle transitions in the dipole limit. In addition, the transitions matrix elements have radial and angular components. The latter can often be evaluated analyically and have many symmetry properties. To gain insight into these aspects, we look at the underlying tensor algebra in Chapter 3.

Looking again at the spectrum, we see that there are in fact two edges, known as the L_3- and L_2-edges. These correspond respectively to the $j = 3/2$ and $1/2$ total angular momenta of the $2p$ core-level. This is another major difference with the hydrogen atom. In contrast to the hydrogen atom where the spin–orbit interaction is less than a milli-electronvolt, the spin–orbit interaction strength can be tens to hundreds of electronvolts for the core-levels of heavy atoms/ions. This has another important consequence. Whereas for the hydrogen atom the spin–orbit interaction can generally be treated within perturbation theory, it is often the dominant interaction for X-ray absorption final states. This implies that the orbital and spin moments are no longer good quantum numbers.

A more detailed look at the spectrum in Fig. 1.1 reveals a significant amount of fine structure. In the hydrogen atom, the fine structure is due to the interaction of the electron with the nucleus. However, for many-electron atoms, the additional structure is for a large part due to the Coulomb interaction between the electrons in the open shells (the closed shells only give rise to a shift in the energy). For the L-edges this means the $2p$–$3d$ and $3d$–$3d$ Coulomb interactions. The Coulomb interaction splits the spectrum into distinct Coulomb multiplets. An aspect of Coulomb multiplets is Hund's rule that tells us which is the lowest Coulomb multiplet in the ground state. This brings us into the complex realm of many-body physics. Let us consider the ground state. For $3d$ transition-metal ions, we have n electrons in the $3d$ shell (with $n = 7$ for divalent cobalt). We can denote these as $3d^n$ configurations (implicitly assuming that all the other shells are completely full or empty). In the absence of electron–electron interactions, this is simply a problem of n independent electrons in the $3d$ shell. However, with Coulomb interactions, different states are no longer independent and we have to deal with $N = 10!/n!(10 - n)!$ states, which are the number of different ways of putting n electrons into the ten $3d$ orbitals. For example, for divalent cobalt ($n = 7$), $N = 120$. Although analytical results can be obtained, we also need to consider the numerical evaluation of the problem. This will be done in Chapter 5.

As mentioned above, the spectrum in Fig. 1.1 closely resembles that of a divalent cobalt ion in cobalt oxide (CoO). Therefore, the spectrum of a divalent cobalt ion in a solid is different from that of a free divalent cobalt ion. There is a variety of ways to include solid-state effects. Unfortunately, due to the presence of strong many-body effects, it is generally not possible to consider the entire solid. One therefore often resorts to the incorporation of solid-state effects in an approximate way, either by including crystal/ligand field effects in an ionic model or by evaluating the spectral lineshape for small clusters. The presence of solid-state effects also lowers the symmetry of the problem, which is discussed in Chapter 4.

Another peculiar aspect of the spectrum in Fig. 1.1 is the unusual intensity ratio of the two spin–orbit split edges. Naively, one would expect the intensities to be proportional to the $2j + 1$ degeneracies of the $2p_j$ core-levels. In this case, there should be a 2:1 intensity ratio between the L_3- and L_2-edges. Some aspects of this change can be evaluated analytically. By integrating over the absorption intensity (either the total spectrum or the edges separately), the details of the final states can be removed and one can relate the integrated intensities to ground-state properties. These are generally called X-ray absorption sum rules, which are derived in Chapter 6. This chapter also discusses a number of effects that occur when including the effects of polarization. Figure 1.1 shows the calculated difference between left and right circularly polarized light. This is known as X-ray magnetic circular dichroism. Since we are measuring/calculating a difference, the intensity can also be negative. In addition, the integrated intensity can also be nonzero. Using sum rules we can relate this quantity to ground-state expectation values of the orbital and spin moments. X-ray circular dichroism is therefore sensitive to magnetism. Note that we are only able to observe this if the material is ferromagnetic (which CoO is not), otherwise the dichroism intensities from sites with opposite moments cancel.

Although there is a strong interaction between the core-hole (created by exciting a $2p$ electron in our example) and the valence shell, its effects on the X-ray absorption spectrum are often small. This is because the excited electron is still on the same ion and therefore the charge has not changed. However, when exciting further above the edge, the ion becomes effectively more positively charged. This results in a flow of electrons to that site in an attempt to screen the additional positive charge. These screening effects are important for the understanding of the lineshape of X-ray spectroscopy and in particular X-ray photoelectron spectroscopy and (K-edge) resonant inelastic X-ray scattering. They lead to the presence of satellite features close to the main lines. The inclusion of such screening effects is a complex many-body problem, which is described in Chapter 7.

Chapter 8 describes the trends in the X-ray absorption lineshape for various elements and the effects of different interactions, such as spin–orbit coupling, crystal fields, and Coulomb interactions. The focus is on general trends and creating a better understanding. The spectra are evaluated numerically. There are good books, review articles, and papers that provide additional spectra.

The final chapter focuses on inelastic X-ray scattering both on and off resonance. Here, one looks at the inelastic scattering of an X-ray from a molecule or a solid. Inelastic X-ray scattering is directly due to the two-photon $\mathbf{A}^2$ term. Resonant inelastic X-ray scattering is a two-step process, where the absorption is followed by a radiative decay. One subsequently measures the energy and sometimes momentum/polarization of the outgoing X-ray. Although the decay can occur between two core-levels, we focus here on decay between the two shells that were involved in the absorption. This leads to final-state excitations of the order of several electronvolts. Since we can measure the scattering angle of the X-rays, we are also able to determine the momentum dependence of the excitations. This allows us to measure spin and orbital excitations, charge-transfer excitations, local crystal-field excitation, etc.

Cambridge University Press & Assessment
978-1-107-03355-9 — Theory of Inelastic Scattering and Absorption of X-rays
Michel van Veenendaal
Excerpt
More Information

2

Background

This chapter reviews several basic concepts from quantum mechanics that are needed for the interpretation of spectroscopy. We start with a review of conservation laws in quantum mechanics. Essentially, one can view a spectroscopy experiment as an exchange of physical quantities, such as energy, momentum, and angular momentum between the photon field and the system that we want to study. We then review the results for the solution of the Schrödinger equation for hydrogen-like atoms, which is an independent electron solution. Many-body wave functions for many-electron atoms are then constructed from the single-electron solutions.

2.1 Symmetry and conservation laws

The purpose of spectroscopy is to extract information from a particular system through its interaction with electromagnetic radiation. Obviously, this is a complex problem that includes many different aspects. First, there is the system that we are trying to study, which can, for example, be an atom, a molecule, or a solid. One of the main reasons why spectroscopy is complicated is because it directly probes the electronic structure. The interpretation of spectroscopy therefore generally requires the development of a model of the electronic structure of the system that we are studying. The electronic structure of the system is described by the Schrödinger equation that describes how the many-particle wave function Ψ depends on time and space,

$$H\Psi = i\hbar \frac{\partial \Psi}{\partial t}, \tag{2.1}$$

where H is the Hamiltonian of the system,

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + \sum_i H_1(\mathbf{r}_i) + \sum_{i,j} H_2(\mathbf{r}_i, \mathbf{r}_j), \tag{2.2}$$

Cambridge University Press & Assessment
978-1-107-03355-9 — Theory of Inelastic Scattering and Absorption of X-rays
Michel van Veenendaal
Excerpt
More Information

and where the summations go over the particles in the system. The first term is the kinetic energy where m_i is the mass of the particle and $\mathbf{p}_i$ its momentum; H_1 describes the potential landscape with $\mathbf{r}_i$ the position of the particle; H_2 describes the interaction between the particles. A successful theoretical interpretation of a spectroscopy experiment generally requires an understanding of the electronic structure of the system. This model is then used to calculate a spectrum, which is then compared to experimental results. This should be contrasted with many structural probes. For example, for X-ray diffraction, interpreting the data generally involves an inversion problem from reciprocal (momentum) to real space.

However, before we can even think about studying the electronic structure, we have to understand what effect the spectroscopy experiment has on the system. One can view a spectroscopy experiment as the exchange of physical quantities, such as energy, momentum, and/or angular momentum, between the photon field and the system. The reason this allows us to learn something about the system is that many of these quantities are (often) conserved. For example, when an X-ray scatters inelastically in a solid, the energy and momentum gained by the system equal those lost by the X-ray. From the probability of this scattering process for a particular energy and momentum, we can extract information regarding the electronic, magnetic, and structural properties of the material. It is therefore important to understand these conservation laws and how they arise. When conservation laws become more complex, they are often called selection rules. This is the convention, for example, for conservation of angular momentum. Both terms shall be used throughout the book.

The underlying reason why certain observables are constants of motion is the symmetry of the physical system under consideration. The treatment of conservation and symmetry in quantum mechanics differs from that in classical mechanics due to the involvement of operators and wave functions. If an operator is a constant of motion, its expectation value does not change as a function of time. The time dependence of the wave function

$$|\Psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\Psi\rangle \tag{2.3}$$

follows directly from the Schrödinger equation (2.1). We can move the time dependence from the wave function to the operator via

$$\langle\Psi(t)|\mathrm{A}|\Psi(t)\rangle = \langle\Psi|e^{\frac{i}{\hbar}Ht}\mathrm{A}e^{-\frac{i}{\hbar}Ht}|\Psi\rangle = \langle\Psi|\mathrm{A}(t)|\Psi\rangle, \tag{2.4}$$

using roman letters for operators working in Hilbert space, by defining the Heisenberg operator

$$A(t) = e^{\frac{i}{\hbar}Ht}Ae^{-\frac{i}{\hbar}Ht}. \tag{2.5}$$

The time dependence of the operator $A(t)$ is given by

$$\frac{\partial A(t)}{\partial t} = \frac{i}{\hbar}[H, A(t)], \tag{2.6}$$

where the commutation relation is defined as

$$[H, A(t)] = HA(t) - A(t)H. \tag{2.7}$$

Now if $A(t)$ is a constant as a function of time, then its derivative with respect to time is zero, implying that A and H commute with each other. This commutation relation has important consequences. Let us denote the eigenfunctions of H as φ_k with eigenenergies E_k,

$$H\varphi_k = E_k\varphi_k, \tag{2.8}$$

where k is a generic index labeling the eigenstates. Now let us operate the Hamiltonian on $A\varphi_k$:

$$H(A\varphi_k) = AH\varphi_k = AE_k\varphi_k = E_k(A\varphi_k), \tag{2.9}$$

where use has been made of the commutation relation $[H, A] = 0$. The above equation implies that $A\varphi_k$ is also an eigenfunction of H with eigenvalue E_k. The vector $A\varphi_k$ must therefore be proportional to φ_k; hence φ_k is also an eigenfunction of A.

Now we must ask ourselves: what makes an operator a constant of motion? An operator becomes a constant of motion if it is related to the symmetry of the system under consideration. Often the action of the operator corresponds to a symmetry operation of the system. Let us consider two very important examples.

2.1.1 Invariance in time

An operator is a constant of motion if it commutes with the Hamiltonian. Since an operator commutes with itself, the Hamiltonian is a constant of motion, and its eigenvalues (i.e. energy) are good quantum numbers. This directly implies that energy is conserved. Energy is obviously a key variable in spectroscopy. Let us define the operator

$$U(t) = e^{-\frac{i}{\hbar}Ht}. \tag{2.10}$$

$U(t)$ is known as a unitary transformation. From the solution of the time-dependent part of the Schrödinger equation (see Eq. (2.3)), it directly follows that this operator causes a displacement in time of the wave function

$$U(\Delta t)\Psi(t) = \Psi(t + \Delta t). \tag{2.11}$$

Cambridge University Press & Assessment
978-1-107-03355-9 — Theory of Inelastic Scattering and Absorption of X-rays
Michel van Veenendaal
Excerpt
More Information

The invariance of the Hamiltonian in time therefore leads to the law of conservation of energy. We now look at conservation laws that follow from symmetries in the spatial coordinates.

2.1.2 Translational symmetry

Translational symmetry imposes restrictions on the Hamiltonian in Eq. (2.2), since it cannot directly depend on the position $\mathbf{r}$ in space. This implies that $H_1(\mathbf{r})$ can only be constant. However, we can still have interactions between particles since these generally only depend on the relative position $\mathbf{r}_i - \mathbf{r}_j$ of the particles, which is independent of the absolute position. In classical mechanics, a translation of the coordinates can easily be made by adding a translation vector $\boldsymbol{\tau}$ to the position vector, giving $\mathbf{r}' = \mathbf{r} + \boldsymbol{\tau}$. In quantum mechanics, one needs to find the operator $U(\boldsymbol{\tau})$ that gives a translation of the wave function in space. Therefore, operating $U(\boldsymbol{\tau})$ on a one-particle time-independent wave function $\psi(\mathbf{r})$ should give

$$U(\boldsymbol{\tau})\psi(\mathbf{r}) = \psi(\mathbf{r} + \boldsymbol{\tau}). \tag{2.12}$$

The right-hand side can be expanded in a Taylor series and subsequently re-summed as

$$\psi(\mathbf{r} + \boldsymbol{\tau}) = \psi(\mathbf{r}) + (\boldsymbol{\tau} \cdot \nabla)\psi + \frac{1}{2!}(\boldsymbol{\tau} \cdot \nabla)^2\psi + \cdots = e^{\boldsymbol{\tau}\nabla}\psi(\mathbf{r}). \tag{2.13}$$

We can always add some constant factors and define $\mathbf{p} = -i\hbar\nabla$ where $\hbar = h/2\pi$ with h as Planck's constant. Recognizing $\mathbf{p}$ as the operator for the momentum, we can write the translation operator as

$$U(\boldsymbol{\tau}) = \exp\left(\frac{i}{\hbar}\mathbf{p} \cdot \boldsymbol{\tau}\right). \tag{2.14}$$

Obviously, if we can shift everything by $\boldsymbol{\tau}$, we should be able to shift it back again with $-\boldsymbol{\tau}$. Therefore, the inverse of U is given by

$$U^{-1}(\boldsymbol{\tau}) = U(-\boldsymbol{\tau}) = e^{-\frac{i}{\hbar}\mathbf{p}\cdot\boldsymbol{\tau}} = (e^{\frac{i}{\hbar}\mathbf{p}\cdot\boldsymbol{\tau}})^{\dagger} = U^{\dagger}(\boldsymbol{\tau}), \tag{2.15}$$

where † stands for the Hermitian conjugate. The last two steps are only possible if the momentum is a Hermitian operator, i.e. $\mathbf{p}^{\dagger} = \mathbf{p}$. In this case, the transformation is unitary meaning that the inverse of an operator is equal to its Hermitian conjugate: $U^{-1} = U^{\dagger}$. In addition, we have $U^{\dagger}U = U^{-1}U = 1$. The terms Hermitian and unitary might at first sight appear as purely mathematical concepts, but they have important physical consequences, which we shall now discuss.

The Hermitian (or adjoint) of an operator A is defined such that its expectation value for any wave function ψ is

$$\int d\mathbf{r}\psi^* A\psi = \int d\mathbf{r}(A^\dagger\psi)^*\psi. \tag{2.16}$$

For the momentum operator, the Hermitian can be obtained via integration by parts:

$$\int d\mathbf{r}\psi^*(-i\hbar\nabla)\psi = \left[-i\hbar|\psi|^2\right] + \int d\mathbf{r}(i\hbar\nabla)\psi^*\psi = \int d\mathbf{r}\,(-i\hbar\nabla\psi)^*\,\psi,$$

where the integrand vanishes as long as the integration volume is large enough that the wave function is zero on its surface. We see from the definition in Eq. (2.16) that the momentum operator is Hermitian. What does that physically imply? Let us assume that ψ is an eigenfunction of a Hermitian operator A with eigenvalue q, i.e. $A\varphi = q\varphi$. Since $A^\dagger = A$, Eq. (2.16) directly gives that $q = q^*$. This implies that the eigenvalue q is real, which is exactly what we expect for a physical quantity. If q was complex, the transformations would no longer be unitary but evanescent (or worse, divergent). The commutation relation between U and H can be combined with the unitarity relation $U^\dagger U = 1$ to give

$$UH = HU \Longrightarrow U^\dagger UH = U^\dagger HU \Longrightarrow H = U^\dagger HU. \tag{2.17}$$

This implies that the Hamiltonian is invariant under the unitary transformation $U^\dagger HU$.

We have found above that the momentum operator $\mathbf{p} = -i\hbar\nabla$ is a proper symmetry operator for a system with translational symmetry. Since $\mathbf{p}$ commutes with the Hamiltonian, the one-particle eigenfunctions of $\mathbf{p}$ are also eigenfunctions of the Hamiltonian. Since from the de Broglie relationship, we know that the eigenvalues of $\mathbf{p}$ are equal to $\hbar\mathbf{k}$, where $\mathbf{k}$ is the wave vector, the eigenvectors can be found from

$$\mathbf{p}\varphi_{\mathbf{k}}(\mathbf{r}) = -i\hbar\nabla\varphi_{\mathbf{k}}(\mathbf{r}) = \hbar\mathbf{k}\varphi_{\mathbf{k}}(\mathbf{r}), \tag{2.18}$$

which gives

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.19}$$

where $\mathcal{V}$ is the volume of the system ensuring normalization of the wave function. Note that the entire wave function is determined by the symmetry of the problem.

The eigenfunctions in free space have great advantages and are often used even when they are not true eigenfunctions. In this case, they are known as basisfunctions. We shall use $\varphi_{\mathbf{k}}(\mathbf{r})$ when using the plane waves as basisfunctions. An arbitrary wave function can always be expressed in terms of a complete set of

Cambridge University Press & Assessment
978-1-107-03355-9 — Theory of Inelastic Scattering and Absorption of X-rays
Michel van Veenendaal
Excerpt
More Information

basisfunctions

$$\psi(\mathbf{r}) = \sum_{\mathbf{k}} c^{\mathbf{k}} \varphi_{\mathbf{k}}(\mathbf{r}), \tag{2.20}$$

where $c^{\mathbf{k}}$ is a coefficient giving the amount of the basisfunction $\varphi_{\mathbf{k}}(\mathbf{r})$ in the wave function. Since we have chosen a finite normalization volume $\mathcal{V}$, the **k** values are discretized and we use a summation instead of an integral. The basisfunctions satisfy orthonormality via

$$\int d\mathbf{r} \varphi^*_{\mathbf{k}'}(\mathbf{r})\varphi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\mathcal{V}} \int d\mathbf{r} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} = \delta_{\mathbf{k},\mathbf{k}'}, \tag{2.21}$$

using the definition of the δ function. However, once we have established this relationship, there is no need to perform an integration over space every time we need to determine if two basisfunctions are orthogonal. Everything is determined by the **k** values. This leads to the concept of function spaces, where one no longer keeps track of the detailed spatial dependence of the wave function, but only of how much of each basisfunction is present in the wave function. The basisfunctions are then unit vectors $|\mathbf{k}\rangle$ in an infinite-dimensional function space, also known as a Hilbert space,

$$|\mathbf{k}\rangle = \begin{pmatrix} \cdot \\ 0 \\ 1 \\ 0 \\ \cdot \end{pmatrix} \leftarrow \mathbf{k} \quad \text{and} \quad \langle \mathbf{k}| = (|\mathbf{k}\rangle)^\dagger = (\ldots, 0, 1, 0, \ldots), \tag{2.22}$$

where the components are zero except for the element that we have chosen to correspond to **k**. The right-hand side of the equation gives the Hermitian conjugate $\langle\mathbf{k}|$ of a vector. How to set up the unit vectors is our choice, as long as we do it consistently. The unit vectors satisfy the exact same orthonormality relationship as in Eq. (2.21),

$$\langle \mathbf{k}'|\mathbf{k}\rangle = \delta_{\mathbf{k},\mathbf{k}'}, \tag{2.23}$$

without the need to perform an integration. In addition, from Eq. (2.18), we also know that the basisfunctions are eigenfunctions of the momentum operator and therefore

$$\mathbf{p}|\mathbf{k}\rangle = \hbar\mathbf{k}|\mathbf{k}\rangle. \tag{2.24}$$

Therefore all the properties of the basisfunctions have now become numbers, known as matrix elements. However, in order to determine these matrix elements, we generally need the spatial dependence of the wave functions, although often