Colour and Spectroscopy

One should always be in awe of the fact that our eyes can gather so much information in our daily life. With many marvellous functions like automatic focusing or automatic aperture, the most prominent feature of our vision system is its sensation of colours. More strictly speaking, our eyes can resolve light with different wavelengths and thus serve as a spectrometer that works in the "visible" light range of the electromagnetic (EM) spectrum.

The usual way in which we sense the colour of objects is to capture the light which is reflected from the objects into our pupils. We have three different cone cells that each respond to different parts of the visible spectrum [68]. From the combined response, our brains generate a colourful image of our surrounding world. As we shall see below, the main experimental technique, infrared spectroscopy, presented in this thesis is very similar to our vision system except that it has a much broader detection range and a much larger colour space. Therefore, the "colours" "seen" by infrared spectroscopy contain more information than the perceived colours generated by our brains. This chapter contains two parts: how do we measure colours with infrared spectroscopy? and what can we learn from these colours?

3.1. How do we measure colours with infrared spectroscopy?

Colour, in the daily life context, is a perceptual description of the spectra of light. Our human eyes decompose continuous spectra with 3 basis functions, which are the response functions of our “cone” cells. In other words, the spectra of light are mapped onto a 3-dimensional manifold, as other trichromatic models do [69–71]. As the result, much information is lost during this mapping process since three response functions cannot be a complete basis set for random spectrum functions. For example, metamerism, also known as colour
matching, describes the phenomenon that different spectra give the same response in human eyes [68, 72, 73]. The colour matching phenomena becomes prominent in colour vision deficiency, i.e. colour blindness when the colour space is only two- or even one-dimensional. But with the infrared spectroscopy system, one can measure the fine frequency dependence of the intensity of a spectrum and distinguish colours that are the same in common trichromatic systems.

As we all know that we can sense the colour of an object through its reflected or transmitted colour, the spectra of samples can also be measured in the spectrometer with the reflection or transmission configurations. All the measurements presented in this work are performed in the reflection configuration.

The intuitive way to measure the reflectance of samples at certain frequencies is to use monochromatic light with different frequencies and measure how much light is reflected at each frequency. Then we could obtain the reflectance spectra at discrete frequency points. Fourier-transform infrared spectroscopy (FTIR) uses polychromatic light sources integrated into a Michelson interferometer as shown in Fig. 3.1. This enables us to measure the reflected light from a sample with a broad spectrum at once as we will now explain [74].

The light source emits light with the spectral composition:

$$
\sum_{\omega} E_0(\omega)e^{-i\omega t+ik \cdot r+\phi}
$$

Figure 3.1: The schematics of the Fourier-transform spectrometer. One of the mirrors at the end of the arms of the interferometer is fixed (Arm 1) while the other can move (Arm 2).
where the field of the light is decomposed in Fourier components with $\omega$ as the frequency of the light, $k$ as the wavenumber and $\phi$ as a constant phase factor. The optical path difference $x$ between the two arms of the interferometer can be controlled by moving one of the mirrors. To measure the complete spectrum of the source, the mirror would have to move from $x = 0$ to $x = \infty$. In practice, this can be done over a finite path difference. The field $E(x)$ at the output of the interferometer is the result of the interference of the beams at the beam splitter:

$$E(x) = \sum_{\omega} E_0(\omega)e^{-i\omega t}(e^{ik\cdot r_1} + e^{ik\cdot(r_1+x)})$$

(3.2)

in which $r_1$ is the optical path of Arm 1 and $x$ is the optical path difference between Arm 2 and Arm 1. The intensity $I(x)$ is the modulus of $E(x)$ and has maxima and minima from constructive or destructive interference according to:

$$I(x) = \sum_{\omega} I(\omega)(1 + \cos(k \cdot x))$$

(3.3)

The momentum $k$ and frequency $\omega$ can be expressed in wavenumber by $k = 2\pi\nu$ and $\omega = 2\pi c \nu$, and The non-oscillating part can be rewritten by the intensity with zero optical path difference $I(x = 0)$ when all the wavelength components contribute constructively:

$$I(x) = \sum_{\nu} I(\nu)\cos(2\pi\nu x) + \frac{1}{2}I(0)$$

(3.4)

Therefore, the spectra $I(\nu)$ could be obtained by doing the inverse Fourier transform as shown in Fig. 3.2. In experiments, we collect the data at different optical path differences while the mirror is moving:

$$I(\nu) = \int_{-\infty}^{\infty} dx[I(x) - \frac{1}{2}I(0)]e^{i2\pi\nu x}$$

(3.5)

In order to obtain the reflectivity of samples, one should at least do two measurements: one for measuring the signal reflected from the sample, the other for measuring the signal reflected from a perfect mirror that is placed at the same position and has the same geometry of the sample. The ratio between the two signals gives the reflectivity of the samples. Starting from this one can explore other optical response functions using the Fresnel equations. The data discussed in this thesis were obtained from a sample measurement and a reference measurement using metal evaporated \textit{in situ} on the sample as reference [76].

We use combinations of detectors and beam splitters in order to obtain the spectra over a wide frequency range. Different light sources are used for

\footnote{In a realistic measurement, the infinity as the integral limits cannot be achieved. The maximum moving distance $L$ of the mirror applies the limits from $-\frac{L}{2}$ to $\frac{L}{2}$ to the inverse Fourier transform, where the apodisation is needed in the signal processing. See detailed discussions in [75].}
energy windows between 6 meV and 4 eV. Different metals are used to serve as reference mirrors at particular spectral ranges [76]. Typically, we use Au, Ag and Al as reference metals. Fig. 3.3 shows the reflectivity for each of these metals. In the infrared range, Au has the highest reflectivity. However, its plasma edge is only 1 eV above which it can no longer be used. Between 0.7 eV and 2 eV, we can instead use Ag, while for the UV range (2.5 eV ≥ 4 eV) we use Al. Note that the reflectivity of Al is only 90% in the UV part as the result of an interband transition at 1.25 eV. One needs to make corrections for this when computing the reflectivity of a sample.

Apart from the Fourier-transform infrared spectrometer introduced above, our system also includes a helium-cooled cryostat that enables us to continuously perform the reflectivity measurement over a broad temperature range. To give an overview of the experimental procedure, the measurements are designed to be able to run through cooling and warming cycles between 14 K and 400 K. In each cycle we used the constant heating or cooling rates. An infrared spectrum was recorded every minute during the temperature cycles which will
be repeated two to three times to increase the signal-to-noise ratio and ensure reproducibility.

3.2. WHAT CAN WE LEARN FROM COLOURS?

For the reflectivity measurements concerned in this thesis, the most important formula, which connects the macroscopic experimental quantity $R(\omega)$ to microscopic properties, is the Fresnel equation in the normal incidence case [38]:

$$\hat{r}(\omega) = \frac{1 - \sqrt{\varepsilon(\omega)}}{1 + \sqrt{\varepsilon(\omega)}}$$

(3.6)

Here $\hat{r}(\omega)$ is the complex reflectivity. Its modulus square $R(\omega) = |\hat{r}(\omega)|^2$ is the absolute reflectivity defined by the reflectance of the infinitely large sample with infinite thickness, and $\varepsilon(\omega)$ is the dielectric function that characterises the electronic response of the material to the surrounding EM field.

3.2.1. FITTING METHODS AND THE CLASSICAL MODEL

One should notice that the dielectric function $\varepsilon(\omega)$ has both a real and imaginary part. Although these two parts are related via the Kramers-Kronig relations [77], data obtained within finite energy range makes it impossible to exactly map from one real function $R(\omega)$ to a complex function $\varepsilon_1(\omega) + i \varepsilon_2(\omega)$. In order to calculate $\varepsilon(\omega)$ as well as other optical functions, there are various practical methods that are commonly used [78].

**Drude-Lorentz model**

One method is to use models of the dielectric function $\varepsilon(\omega)$ and optimise parameters through the fitting of the reflectivity data using Eq. 3.6, such as the classical Drude-Lorentz model, Fano lineshape for asymmetric phonon peaks [79] or the Zimmermann model for the optical response of the BCS superconductor [48]. The fitting parameters sometimes can be regarded as the approximation of certain physical parameters, such as the plasma frequency $\omega_p$ in the Drude model or the line width $\gamma$ of a phonon peak. The model functions are by definition Kramers-Kronig compatible.

Among all the models used, the Lorentz model is at the heart of fitting optical data because of its simplicity and clear physical picture. We consider a single electron moving in the time-dependent electric field $E_0 e^{i\omega t}$, which follows the general form of the equation of motion:

$$m \frac{d^2x(t)}{dt^2} - m \gamma \frac{dx(t)}{dt} + m \omega_0^2 x(t) = -eE(t)$$

(3.7)

The second term is the scattering term characterised by the scattering rate $\gamma$, and the third term specifies the eigenfrequency $\omega_0$ of electrons. Here we have assumed that the electrons are bound to ionic cores and that the atomic
potential can be approximated by a harmonic oscillator. The solution of this equation can be easily obtained by the Fourier transform:

\[ x(\omega) = -\frac{eE_0}{m} \frac{e^{i\omega t}}{\omega_0^2 - \omega^2 - i\gamma \omega} \]  

(3.8)

The dynamics of electrons determine the electronic response to the applied EM field, such as the polarisation or the charge current induced by the applied field. In linear response, the polarisation or other physical quantities are taken to be proportional to the applied field:

\[ P = \chi E \]
\[ J = \sigma E \]  

(3.9)

The response function is defined to be the coefficient of proportionality. For example, the dielectric function is defined as the response function in terms of electric displacement \( D \) to the external electric field \( E \):

\[ \varepsilon = \frac{4\pi D}{E} = \frac{E + 4\pi P}{E} \]  

(3.10)

Here we assume that the material is isotropic so that \( \varepsilon \) is a scalar. The polarisation \( P \) is proportional to the displacement of charge relative to the ion core. Averaged over the volume of the sample, this gives:

\[ P = -\frac{eNx}{V} \]  

(3.11)

Combining the equations above, we obtain the classical Lorentz oscillator:

\[ \varepsilon(\omega) = 1 + \frac{4\pi e^2 N}{mV} \frac{1}{\omega_0^2 - \omega^2 - i\gamma \omega} \]  

(3.12)

For free electrons, there is no restoring force and thus \( \omega_0 = 0 \). In this case, the Lorentz oscillator reduces to the Drude response:

\[ \varepsilon(\omega) = 1 - \frac{4\pi e^2 N}{mV} \frac{1}{\omega^2 + i\gamma \omega} \]  

(3.13)

In the case of free electrons, the prefactor \( 4\pi e^2 N/mV \) is the so-called plasma frequency \( \omega_p^2 \). The general Drude-Lorentz model considers a free charge response and a bound charge response:

\[ \varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma \omega} + \sum_j \frac{f_j^2}{\omega_0^2,j - \omega^2 - i\gamma_j \omega} \]  

(3.14)

The optical conductivity \( \sigma(\omega) \) of the Drude-Lorentz model can also be obtained knowing that \( \sigma(\omega) \) is the response function in terms of the current density \( J \) with respect to the field \( E \):

\[ \sigma(t) = \frac{J(t)}{E(t)} = -\frac{eN \dot{x}}{E(t)} \]  

(3.15)
3.2. What can we learn from colours?

\( \dot{x} \) is the time derivative of the electron coordinate. After the Fourier transform and by using Eq. 3.8, we obtain the relation between \( \varepsilon(\omega) \) and \( \sigma(\omega) \):

\[
\varepsilon(\omega) = 1 + i \frac{4\pi \sigma(\omega)}{\omega} \quad (3.16)
\]

The correct way of interpreting the Lorentz oscillators is in terms of interband transitions, while the Drude term describes the intraband scattering of electrons. We shall go in-depth on the rigorous description using quantum theory in the next section.

**VARIATIONAL DIELECTRIC FUNCTION METHOD**

A more general approach to fit the optical data is to extract the complex response functions without assumptions of specific models. For example, one can directly apply Kramers-Kronig relations to the reflectivity data obtained within a finite frequency range [80]. The measured \( R(\omega) \) as the modulus square of the complex reflectivity \( |r|e^{i\theta} \) can be related with the phase angle \( \theta \) via the Kramers-Kronig relations:

\[
\theta(\omega) = \frac{2\omega}{\pi} \int_{0}^{\infty} \frac{\ln R(\omega')}{\omega'^2 - \omega^2} d\omega' \quad (3.17)
\]

The limitation is that we only have \( R(\omega) \) on a finite domain \([\omega_{\text{min}}, \omega_{\text{max}}]\). There are ways we can use to circumvent this limitation. For example, we can use some kind of reasonable estimation of the reflectivity in the low energy and the high energy limit which comes from the Drude-Lorentz model fitted beforehand.

An alternate approach was proposed in [81], which is a model-independent method to calculate the optical response functions using Kramers-Kronig constrained variational dielectric function method. The basic idea of this method is to use a set of approximation Lorentzian functions as the basis functions to fit Eq. 3.6. It can be regarded as an approximate of the Drude-Lorentz model which requires \( \sim O(N) \) oscillators\(^2\) (\( N \) is the number of data points). In the optical spectra fitting software RefFit, the basis is chosen to be a set of local triangular functions around each frequency mesh point as plotted in Fig. 3.4.

We start from an obtained Drude-Lorentz model that already has a reasonable agreement with the original data. The dielectric function is then the linear combination of the basis functions plus the Drude-Lorentz model which is kept fixed during the variational fitting:

\[
\varepsilon(\omega) = \sum_{i} A_{i} e^{i \Delta(\omega)} + \varepsilon_{DL}(\omega) \quad (3.18)
\]

One can vary the parameter set \( \{A_{i}\} \) to minimise the difference between the data and the fitting function, which is the \( \chi^{2} \) term in the least squared method.

\(^2\)More precisely, the number of oscillators in the model is \( N/2 \).
As the result, even the noise of the data is fitted by the variational dielectric functions as well. The resulting optical response functions are model-independent and satisfy Kramers-Kronig relations.

The Hagen-Rubens relation
The Hagen-Rubens relation can be used to extract the DC resistivity directly from the reflectivity data. We shall show that the free charge response gives a very unique asymptotic behaviour in the regime where $\omega \ll \gamma$.

From Eq. 3.13, one could find the low energy response of the Drude model in the limit $\omega \ll \gamma$:

$$\varepsilon_1(\omega) = 1 - \frac{\omega_p^2}{\gamma^2}$$

$$\varepsilon_2(\omega) = \frac{\omega_p^2}{\gamma} \frac{1}{\omega}$$  \hspace{1cm} (3.19)

With the aid of the Fresnel Equation Eq. 3.6, one could arrive [38]3:

$$R \approx 1 - 2 \sqrt{\frac{2}{\varepsilon_2(\omega)}} = 1 - 2 \sqrt{\frac{2\omega\gamma}{\omega_p^2}}$$  \hspace{1cm} (3.20)

realising that the dc conductivity of the free electrons is given by Eq. 3.13 and Eq. 3.16:

$$\sigma(\omega = 0) = \frac{\omega_p^2}{4\pi\gamma}$$  \hspace{1cm} (3.21)

The low-energy asymptotic form of reflectivity is written:

$$R \approx 1 - \sqrt{\frac{2\omega}{\pi\sigma_0}}$$  \hspace{1cm} (3.22)

3The detailed derivation can be found in Chap. 5.1 of [38].
Thus the Hagen-Rubens relation serves as the verification of the metallicity of the materials, which shows the linearity of \( R \) vs \( \omega^2 \) and eventually extrapolates to 1 at zero frequency. Furthermore, its derivative \( \frac{\partial R}{\partial \omega^2} \) also provides a reasonable estimation of the dc conductivity from the reflectivity data collected at finite frequency range. An application of the Hagen-Rubens relation applied to a real material can be found in Chap. 4.

### 3.2.2. One more step: what can we learn from the optical functions?

We have discussed how to extract the optical functions from the reflectivity spectra of materials. In this section, we shall explore several aspects of optical functions.

#### Quantum theory of linear response functions

As we already introduced the Drude-Lorentz model describing the dielectric function of materials, the classical picture considers that the electrons oscillate in response to the applied field. A quantum description of the optical response considers the excitation of electron-hole pairs by the incoming photon. These processes can be described by evaluating the polarization bubble (see section 2.1.2), which can be connected to the macroscopic optical functions in the Kubo formalism.

The Kubo formalism evaluates the time evolution of the observables of a many-particle system that responds to the applied field linearly [82, 83]. The interaction between the field and the system can be described by the coupling of the field \( f \) and generalised coordinate \( \dot{B} \) so that \( H_{\text{int}} = \dot{B}(t)f(t) \). Then the related observables of the system \( \dot{A} \) is given by:

\[
\langle \dot{A}(t) \rangle = \langle \dot{A}(0) \rangle - i \int_0^t dt' \langle [\dot{A}(t), \dot{B}(t')] \rangle f(t')
\]  

(3.23)

where \((...)\) represents the thermal average in the canonical ensemble. Furthermore, the coefficient of \( f(t') \) is defined as the response function or susceptibility as in Eq. 3.9:

\[
\chi_{AB}(t - t') = -i \theta(t - t') \langle [\dot{A}(t), \dot{B}(t')] \rangle
\]  

(3.24)

in which the step function \( \theta(t - t') \) guarantees causality of that field acts first at \( t' \) then the system responds at \( t \).

As mentioned before, the optical conductivity \( \sigma(\omega, q) \) is the susceptibility of the current density \( j \) in response to the applied electric field \( E \). Now in order to evaluate \( \sigma(\omega, q) \) using the Kubo formula Eq. 3.24, one should obtain the form of the current density \( j \) and the related interacting Hamiltonian \( \dot{B}f \) where the current couples with the field.

The current density operator of the electron can be obtained by its definition
\[
\mathbf{j} = -e \mathbf{v} n = -e \frac{\mathbf{p} + e \mathbf{A}}{m} \sum_{\sigma} \psi_{\sigma}^{\dagger} \psi_{\sigma} \\
= -e \frac{i \hbar}{2m} \sum_{\sigma} [(\nabla \psi_{\sigma}^{\dagger})_{\sigma} \psi_{\sigma} - \psi_{\sigma}^{\dagger} (\nabla \psi_{\sigma})] - e^2 \frac{\mathbf{A}}{m} \sum_{\sigma} \psi_{\sigma}^{\dagger} \psi_{\sigma} \tag{3.25}
\]

The two parts of the current operator are defined as paramagnetic current and diamagnetic current \[84\]. The thermal average of the diamagnetic current \( \langle \mathbf{j}_d \rangle \) can be obtained directly in the canonical ensemble:

\[
\langle \mathbf{j}_d \rangle = -e^2 \frac{\mathbf{A}}{m} \sum_{\sigma} \langle \psi_{\sigma}^{\dagger} \psi_{\sigma} \rangle = - \frac{e^2 N}{mV} \mathbf{A} \tag{3.26}
\]

The density of the electrons \( N/V \) is the ensemble average of the density operator.

In order to know how the current couples with the field, we can use the Hamiltonian of particles in the EM field subtracted by the free particle Hamiltonian \[84\]:

\[
H_{\text{int}} = \int d\mathbf{x} \sum_{\sigma} \left[ \frac{\mathbf{p} + e \mathbf{A}(t, \mathbf{x})}{2m} - e\varphi(t, \mathbf{x}) \right] \psi_{\sigma}(\mathbf{x}) \psi_{\sigma}^{\dagger}(\mathbf{x}) - \int d\mathbf{x} \sum_{\sigma} \frac{\mathbf{p}_{\sigma}^{2}}{2m} \psi_{\sigma}(\mathbf{x}) \psi_{\sigma}^{\dagger}(\mathbf{x}) \tag{3.27}
\]

Here we choose the gauge where the scalar potential \( \varphi = 0 \) so that the \( e\varphi(t, \mathbf{x}) \) term vanishes \[84\]. The first order terms in \( H_{\text{int}} \) containing \( \mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p} \) can be rewritten:

\[
H_{\text{int}} = e \int d\mathbf{x} \frac{i \hbar}{2m} \mathbf{A}(t, \mathbf{x}) \cdot \sum_{\sigma} \left[ (\nabla \psi_{\sigma}^{\dagger})_{\sigma} \psi_{\sigma} - \psi_{\sigma}^{\dagger} (\nabla \psi_{\sigma}) \right] + \frac{e^2}{2m} \mathbf{A}^2(t, \mathbf{x}) \psi_{\sigma}(\mathbf{x}) \psi_{\sigma}^{\dagger}(\mathbf{x}) \tag{3.28}
\]

As momentum \( \mathbf{p} + e \mathbf{A} \) and \( \mathbf{x} \) are canonical conjugates, the current density \( \mathbf{j} \) should follow the Hamilton equation:

\[
\mathbf{j} \propto \frac{\delta H}{\delta \mathbf{A}} \tag{3.29}
\]

Then we combine Eq. 3.25, 3.28 and 3.29 to determine the specific form of the interaction term:

\[
H_{\text{int}} = - \int d\mathbf{x} \mathbf{A}(t, \mathbf{x}) \cdot \mathbf{j}(\mathbf{x}) \tag{3.30}
\]

Here the vector potential \( \mathbf{A}(t, \mathbf{x}) \) is the field \( f \) and current density \( \mathbf{j}(\mathbf{x}) \) is the generalised coordinate \( B \) in Eq. 3.23. Now we can write down the Kubo formula for the paramagnetic current according to Eq. 3.23 and Eq. 3.24 \[85\]:

\[
\delta \langle \mathbf{j}_{\mu}(t, \mathbf{x}) \rangle = \sum_{\nu} \int d\mathbf{x}' dt' \chi_{\mu\nu}(t - t', \mathbf{x} - \mathbf{x}') \mathbf{A}_{\nu}(t', \mathbf{x}') \tag{3.31}
\]
where the response function $\chi_{\mu\nu}(t - t', x - x')$ is defined in Eq. 3.24:

$$
\chi_{\mu\nu}(t - t', x - x') = -i\theta(t - t')\langle [j_\mu^p(t, x), j_\nu^p(t', x')] \rangle
$$

(3.32)

$\chi_{\mu\nu}(t - t', x - x')$ is the current-current correlation function. Its expectation value is the measure of the response of the charge current to the vector potential $A$. This lead to the more familiar optical conductivity which is the current responding to the electric field $E$. Knowing that the relation between them is:

$$
E = -\frac{\partial A}{\partial t} = i\omega A
$$

(3.33)

we can obtain the response function of $j$ and $E$:

$$
\delta\langle j_\mu(t, x) \rangle = i\frac{1}{\omega} \sum_\nu \int dx' dt' \chi_{\mu\nu}(t - t', x - x') E_\mu(t', x') + i \frac{e^2 N}{\omega m V} E_\mu(t, x)
$$

(3.34)

The optical conductivity is the coefficient of the electric field:

$$
\sigma_{\mu,\nu}(t, x - x') = \frac{1}{\omega} \theta(t) \langle [j_\mu^p(t, x), j_\nu^p(0, x')] \rangle + i \frac{e^2 N}{\omega m V} \delta_{\mu,\nu}
$$

(3.35)

The first term can be evaluated in the canonical ensemble:

$$
\langle [j_\mu^p(t, x), j_\nu^p(0, x')] \rangle = \sum_n \frac{e^{-\beta E_n}}{Z} \langle n | j_\mu^p(t, x) j_\nu^p(0, x') | n \rangle
$$

(3.36)

By taking $t > 0$ compulsorily, we neglect the step function $\theta(t)$. In the interaction picture, we know:

$$
j_\mu^p(t, x) = e^{iH_0 t} j_\mu^p(0, x) e^{-iH_0 t}
$$

(3.37)

Thus, the first term of the optical conductivity in Eq. 3.35 can be obtained [85]:

$$
\sigma_{\mu,\nu}^p(\omega, q) = i \frac{1}{\omega} \sum_{\alpha,\beta} (n_\alpha - n_\beta) \frac{\langle \alpha | j_\mu^p(q) | \beta \rangle \langle \beta | j_\nu^p(-q) | \alpha \rangle}{\omega + \epsilon_\alpha - \epsilon_\beta + i\delta}
$$

(3.38)

Here $\alpha$ and $\beta$ are dummy indices of the quantum states, and $n_{\alpha,\beta}$ is the Fermi-Dirac distribution. Conventionally, the charge current density operator $j_\mu^p$ is represented in the form of the first quantisation using the momentum operator $p_\mu$ [85]:

$$
\hat{j}_\mu^p = -\frac{e}{2m} (p_\mu e^{-iq.r} + e^{-iq.r} p_\mu)
$$

(3.39)

In optics, the measurement is done in the $q \to 0$ limit, so that $\sigma(\omega, q = 0)$ can be even more simplified using Eq. 3.39 [85]:

$$
\sigma_{\mu,\nu}^p(\omega, q = 0) = \frac{ie^2}{\omega m^2 V} \sum_{\alpha,\beta} (n_\alpha - n_\beta) \frac{\langle \alpha | p_\mu | \beta \rangle \langle \beta | p_\nu | \alpha \rangle}{\omega + \epsilon_\alpha - \epsilon_\beta + i\delta}
$$

(3.40)
One could rewrite its denominator in order to cancel the diamagnetic term in Eq. 3.35:

\[
\frac{1}{\omega + i\delta + \omega_{\alpha\beta}} = \frac{1}{\omega_{\alpha\beta}} \left( 1 - \frac{\omega + i\delta}{\omega + i\delta + \omega_{\alpha\beta}} \right)
\]  

(3.41)

where \( \omega_{\alpha\beta} = \epsilon_{\alpha} - \epsilon_{\beta} \). The first part of the paramagnetic \( \sigma(\omega) \) in Eq. 3.40 is:

\[
\frac{ie^2}{\omega m^2 V} \sum_{\alpha,\beta} (n_{\alpha} - n_{\beta}) \frac{\langle \alpha | p_{\mu} | \beta \rangle \langle \beta | p_{\nu} | \alpha \rangle}{\omega_{\alpha\beta}} = -\frac{i e^2 N}{\omega m V} \delta_{\mu\nu}
\]  

(3.42)

We use the f-sum rule to obtain this equality, which will be discussed in the next section. Thus, one could see that the diamagnetic term in Eq. 3.35 is canceled out. The optical conductivity is then:

\[
\sigma_{\mu\nu}(\omega, q = 0) = -\frac{ie^2}{\omega m^2 V} \sum_{\alpha,\beta} (n_{\alpha} - n_{\beta}) \frac{\langle \alpha | p_{\mu} | \beta \rangle \langle \beta | p_{\nu} | \alpha \rangle}{(\omega + \omega_{\alpha\beta} + i\delta)\omega_{\alpha\beta}}
\]  

(3.43)

In the \( q \rightarrow 0 \) limit, the matrix element is denoted with a simpler form:

\[
\langle \alpha | p_{\mu} | \beta \rangle = m v_{\mu}^{\alpha\beta}
\]  

(3.44)

We could re-index \( \alpha \) and \( \beta \) so that \( \omega_{\alpha\beta} \) is always positive. Then the summation in \( \sigma(\omega) \) can be rewritten as:

\[
\sigma_{\mu\nu}(\omega, q = 0) = -\frac{ie^2}{V} \sum_{\alpha,\beta} \frac{n_{\alpha} - n_{\beta}}{\omega_{\alpha\beta}} \left[ \frac{v_{\mu}^{\alpha\beta} v_{\nu}^{\beta\alpha}}{(\omega + \omega_{\alpha\beta} + i\delta)} + \frac{v_{\mu}^{\alpha\beta} v_{\nu}^{\beta\alpha}}{(\omega - \omega_{\alpha\beta} + i\delta)} \right]
\]  

(3.45)

One could show that this expression leads to the same form of the Lorentz model discussed before in Eq. 3.14:

\[
\sigma_{\mu\nu}(\omega) = \frac{i\omega}{4\pi} \sum_{\alpha,\beta} \frac{\Omega_{\alpha\beta}^2}{\omega(\omega + 2i\delta) - \omega_{\alpha\beta}^2}
\]  

(3.46)

with the oscillator strength \( \Omega_{\alpha\beta}^2 \) defined as:

\[
\Omega_{\alpha\beta}^2 = -\frac{8\pi e^2 v_{\mu}^{\alpha\beta} v_{\nu}^{\beta\alpha} (n_{\alpha} - n_{\beta})}{\omega_{\alpha\beta} V}
\]  

(3.47)

Interestingly, the quantum theory of the optical response function gives the same form as the classical model does.

**Sum Rules**

As discussed in the previous section, we can extract information from the optical response functions, such as the approximate energy of optical transitions and corresponding spectral weight encoding the joint density of states and the matrix element of optical transitions. Apart from this, we can also extract
information that does not rely on the quality of the modeling of real data: sum rules. The most fundamental sum rule is the $f$-sum rule, which is also known as the Thomas-Reich-Kuhn sum rule \([86, 87]\). It states that the integral of $\sigma_1(\omega)$ over the entire spectral range gives a constant that is proportional to the total density of electrons. The real part of $\sigma(\omega)$ can be obtained from Eq. 3.43:

$$\sigma_1(\omega) = -\frac{e^2 \delta}{m^2 V} \sum_{\alpha, \beta} \frac{(n_\alpha - n_\beta) \langle \alpha | p | \beta \rangle \langle \beta | p | \alpha \rangle}{\omega_{\alpha \beta}(\omega + \omega_{\alpha \beta})^2 + \delta^2}$$  (3.48)

The spectral weight $W$ is defined as the integral of $\sigma_1(\omega)$:

$$W = \int_0^\infty d\omega \sigma_1(\omega) = -\frac{e^2 \delta}{m^2 V} \int_0^\infty d\omega \sum_{\alpha, \beta} \frac{(n_\alpha - n_\beta) \langle \alpha | p | \beta \rangle \langle \beta | p | \alpha \rangle}{\omega_{\alpha \beta}(\omega + \omega_{\alpha \beta})^2 + \delta^2}$$  (3.49)

Since the integrated function $\sigma_1(\omega)$ is even with respect to $\omega$, we can first work out the integral part from $-\infty$ to $\infty$:

$$\frac{\delta}{\omega_{\alpha \beta}} \int_{-\infty}^\infty d\omega \frac{1}{(\omega + \omega_{\alpha \beta})^2 + \delta^2} = \frac{\pi}{\omega_{\alpha \beta}}$$  (3.50)

Then we deal with the summation part. First we need to change the momentum operator into the coordinate basis. We can assume a simple form of Hamiltonian:

$$H_0 = \frac{p^2}{2m} + V(x)$$  (3.51)

The momentum operator in the matrix element could be represented by the commutator:

$$p = \frac{im}{\hbar} [H_0, x]$$  (3.52)

Therefore the summation part could be rewritten as:

$$\sum_{\alpha, \beta} \frac{\langle \alpha | p | \beta \rangle \langle \beta | p | \alpha \rangle (n_\alpha - n_\beta)}{\omega_{\alpha \beta}} = -\frac{m^2}{\hbar^2} \sum_{\alpha, \beta} \frac{|\langle \alpha | [H_0, x] | \beta \rangle|^2 (n_\alpha - n_\beta)}{\omega_{\alpha \beta}}$$

$$= -\frac{m^2}{\hbar^2} \sum_{\alpha, \beta} \frac{|\langle \alpha | H_0 x - x H_0 | \beta \rangle|^2 (n_\alpha - n_\beta)}{\omega_{\alpha \beta}}$$

$$= \frac{m^2}{\hbar^2} \sum_{\alpha, \beta} |\langle \alpha | x | \beta \rangle|^2 \omega_{\alpha \beta} (n_\alpha - n_\beta)$$  (3.53)

In order to calculate this summation, we now first evaluate the matrix element of the commutator $[x, [H_0, x]]$ \([85]\):

$$\langle \alpha | [x, [H_0, x]] | \alpha \rangle = \sum_m (\langle \alpha | x | m \rangle \langle m | [H_0, x] | \alpha \rangle - \langle \alpha | [H_0, x] | m \rangle \langle m | x | \alpha \rangle)$$

$$= 2 \sum_m (\langle \alpha | x | m \rangle \langle m | x | \alpha \rangle (E_m - E_\alpha))$$  (3.54)
On the other hand, the matrix element of the commutator $[x, [H_0, x]]$ can be directly calculated since:

$$[x, [H_0, x]] = \frac{1}{2m} [x, [p^2, x]] = \frac{1}{2m} [x, p \frac{2\hbar}{i}] = \frac{\hbar^2}{m} \tag{3.55}$$

Thus we obtain the equality by combining Eq. 3.54 and 3.55:

$$\frac{\hbar^2}{2m} = \sum_{m} (\langle \alpha | x | m \rangle \langle m | x | \alpha \rangle (E_m - E_\alpha)) \tag{3.56}$$

Notice that the right side of the equation has the same form of the result of Eq. 3.53. We could break down the result of Eq. 3.53:

$$\frac{m^2}{\hbar^2} \sum_{\alpha, \beta} |\langle \alpha | x | \beta \rangle|^2 \omega_{\alpha \beta} (n_\alpha - n_\beta) = -\frac{m^2}{\hbar^2} \sum_{\alpha} n_\alpha \sum_{\beta} |\langle \alpha | x | \beta \rangle|^2 (E_\beta - E_\alpha)$$

$$- \frac{m^2}{\hbar^2} \sum_{\beta} n_\beta \sum_{\alpha} |\langle \alpha | x | \beta \rangle|^2 (E_\alpha - E_\beta)$$

$$= -Nm \tag{3.57}$$

Thus the integral of $\sigma_1(\omega)$ is obtained:

$$\int_{-\infty}^{\infty} d\omega \sigma_1(\omega) = \frac{\pi e^2}{mV} N = \frac{\omega_p^2}{4} \tag{3.58}$$

The spectral weight which integrates starting from $\omega = 0$ is, therefore:

$$W = \frac{\omega_p^2}{8} \tag{3.59}$$

where the plasma frequency is defined in the Section 3.2.1: $\omega_p^2 = 4\pi e^2 N/mV$.

The sum rule states that the integral of $\sigma_1(\omega)$ tells us how many electrons are involved in the optical transitions. The total density of electrons in a material would be recovered if one could integrate $\sigma_1(\omega)$ up to infinitely large energy. A very nice example to illustrate this is aluminium, whose optical response function is measured up to $\sim 10^4$ eV [88, 89]. As shown in Fig. 3.5, the optical response at different spectral ranges roughly corresponds to electrons in different core levels of Al. The spectral weight up to $10^5$ eV recovers the electron density of Al, that is 13 electrons per Al atom.

The conservation of the spectral weight is so powerful since the re-organisation of the electronic structure only re-distributes the total spectral weight. By analysing the spectral weight redistribution, one obtains information on the reconstruction of the electronic bands if a phase transition occurs in the material.
3.2. What can we learn from colours?

In this case, another sum rule is frequently adopted in the analysis: the Ferrell-Glover-Tinkham (FGT) sum rule [90, 91]. It specifically works for the spontaneous symmetry breaking phases, where spontaneous symmetry breaking gives rise to a zero energy collective excitation. The collective excitation would take part of the spectral weight to zero energy, such as the superfluid density of the superconducting phase, or to the low energy range if it is pinned, such as the sliding mode of the CDW phase. In either case, the energy range is normally too low to be detected directly with FTIR spectroscopy. Therefore, part of the spectral weight appears to be missing in the symmetry-breaking states. We can, however, compare the spectral weight measured in the normal state and the symmetry-breaking state, and the obtained difference should be attributed to the spectral weight of the collective excitation:

\[
W_{CE} = \int_{0+}^{\Omega} \left[ \sigma_{1,N}(\omega,T) - \sigma_{1,SB}(\omega,T) \right] d\omega.
\]  

(3.60)

\( \sigma_{1,N}(\omega,T) \) is the normal state optical conductivity and \( \sigma_{1,SB}(\omega,T) \) the optical conductivity of the symmetry-breaking state with \( \Omega \) is a high energy cut-off. Note that the integral starts at a lower cutoff energy \( 0+ \), indicating that the \( \omega = 0 \) contribution is not included, where the collective excitation is. The pinning and scattering of the collective mode would shift and broaden the sliding mode spectral weight to the finite energy range, but beyond our detecting limit and the FGT sum rule can still be applied. More discussion of the FGT sum rule is in Chap. 4 where we use the FGT sum rule to estimate the spectral weight of the sliding mode in VSe\(_2\). The total spectral weight is conserved even in the symmetry-breaking phases if the spectral weight of the collective excitation is also counted.

The spectral weight contributed by the lattice vibrations also contains useful information. For the optical phonons, the spectral weight is expected to be...
proportional to the charge carried by the corresponding vibration mode [88]:

\[
\int_0^\infty d\omega \sigma_1^{ph}(\omega) = \sum_j \frac{\pi e_j^2 n_j}{2m_j} \tag{3.61}
\]

So each IR-active vibration mode can be related to the charge density \( n_j \) and the effective mass \( m_j \). The difficulty of interpreting the spectral weight of optical phonons is that it often involves more than one nucleus, which makes it subtle to properly evaluate the mass and the charge density of the vibration mode. We first consider the simplest case of an IR-active vibration mode involving two nuclei with opposite charges in the center of mass motion. The reduced mass of this vibration is straightforward:

\[
\frac{1}{m_r} = \frac{1}{m_1} + \frac{1}{m_2} \tag{3.62}
\]

Therefore, the spectral weight of this diatomic vibration mode is:

\[
\int_0^\infty d\omega \sigma_1^{ph}(\omega) = \frac{\pi Z^2 e^2 n}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \tag{3.63}
\]

Figure 3.6: The optical conductivity (left) of the optical phonon in MgO. The effective charge \( n_{\text{eff}} \) is defined as the deviation from the nominal valence of the ions Mg\(^{2+}\) and O\(^{-2}\). \( n_{\text{eff}} = 1 \) is recovered after integrating above the optical phonon. The figure is taken from [88]

The real materials whose optical phonon is described by the above perfectly polarised diatomic vibration mode would be diatomic ionic insulators, such as MgO. We define the effective charge \( n_{\text{eff}} \) [88]:

\[
n_{\text{eff}} = \frac{2V}{\pi(Ze)^2(m_{\text{Mg}}^{-1} + m_{\text{O}}^{-1})} \int_0^\infty d\omega \sigma_1^{ph}(\omega) \tag{3.64}
\]

Here \( Z \) is the nominal valence of 2 for Mg\(^{2+}\) and O\(^{-2}\), and \( V \) is the average volume of a single diatomic pair. \( n_{\text{eff}} = 1 \) means that the diatomic model perfectly describes the lattice vibration spectral weight.
As for materials with more than 2 types of atoms in the unit cell, the simple form of the effective mass in Eq. 3.62 is not applicable. Instead, one should use the eigenmodes of the multi-atom unit cell for the optical phonons, and calculate the reduced mass for each mode.

![Figure 3.7: The optical conductivity $\sigma_1(\omega)$ and the effective charge $n_{\text{eff}}$ of the optical phonon in AgCrSe$_2$ at low temperature. $n_{\text{eff}}$ is calculated using Eq. 3.61.](image)

The interaction also modifies the phonon-associated spectral weight. As mentioned above, the types of chemical bonds determine the effective charges associated with the dipole moment. FeSi [79], SrTiO$_3$ [92] and AgCrSe$_2$ [3] are examples where $n_{\text{eff}} > 1$ was found. Thus there is more spectral weight associated with the optical phonons, which is called the charged phonon effect [93]. As an example, in AgCrSe$_2$ the charged effective charge is much smaller than in SrTiO$_3$ or FeSi, which only deviates from 1 for 10% as shown in Fig. 3.7. It is still an open question whether this suggests a weak charged phonon effect or it is due to the systematic error of calculating the reduced mass.