

NTNU 2015

Project Notes

Jørgen Vågan
Supervisor: Ingve Simonsen

September 19, 2015

Abstract

Abstract.. abstract.

1 ADDITIONAL THEORY

1.1 THE DRUDE MODEL [1, p. ?]

REFERENCES

- [1] Jensen B. The quantum extension of the Drude-Zener theory in polar Semiconductors. Handbook of optical constants of Solids, Five-Volume 1997 (1985??)(9)<-it's chapter 9;169-170

1.2 COMPLEX PERMITTIVITY AND REFRACTIVE INDEX (FROM WIKIPEDIA) DON'T USE, LACKING REFERENCES!

Complex electric permittivity:

$$\hat{\epsilon}_r(\omega) = \frac{\hat{\epsilon}(\omega)}{\epsilon_0} \quad (1.1)$$

where

$$\hat{\epsilon}_r(\omega) = \epsilon_r(\omega) + i\tilde{\epsilon}_r(\omega) \quad (1.2)$$

$$= \epsilon_r(\omega) + i\frac{\sigma}{\omega\epsilon_0} \quad (1.3)$$

The complex refractive index \hat{n} is given by

$$\hat{n} = \sqrt{\hat{\epsilon}_r}, \quad (1.4)$$

when the magnetic properties are neglected ($\mu_r = 1$). From this, an expression for the complex refractive index $\hat{n} = n - i\kappa$ can be found:

$$\hat{\epsilon}_r = \hat{n}^2 \quad (1.5)$$

$$\epsilon_r + i\tilde{\epsilon}_r = (n + i\kappa)^2 \quad (1.6)$$

$$\epsilon_r + i\tilde{\epsilon}_r = n^2 - \kappa^2 + i2n\kappa \quad (1.7)$$

giving

$$\epsilon_r = n^2 - \kappa^2 \quad \tilde{\epsilon}_r = 2n\kappa. \quad (1.8)$$

Taking the absolute value or modulus of the relative permittivity

$$|\hat{\epsilon}_r| = \sqrt{\epsilon_r^2 + \tilde{\epsilon}_r^2} \quad (1.9)$$

$$|\hat{\epsilon}_r| = \sqrt{(n^2 - \kappa^2)^2 + (2n\kappa)^2} \quad (1.10)$$

$$|\hat{\epsilon}_r|^2 = (n^4 - 2n^2\kappa^2 + \kappa^4) + 4n^2\kappa^2 \quad (1.11)$$

$$|\hat{\epsilon}_r|^2 = n^4 + 2n^2\kappa^2 + \kappa^4 \quad (1.12)$$

$$|\hat{\epsilon}_r|^2 = (n^2 + \kappa^2)^2 \quad (1.13)$$

$$|\hat{\epsilon}_r| = n^2 + \kappa^2 \quad (1.14)$$

and adding or subtracting the real part of the permittivity, gives

$$|\hat{\epsilon}_r| + \epsilon_r = (n^2 + \kappa^2) + (n^2 - \kappa^2) = 2n^2 \quad (1.15)$$

$$|\hat{\epsilon}_r| - \epsilon_r = (n^2 + \kappa^2) - (n^2 - \kappa^2) = 2\kappa^2. \quad (1.16)$$

Reformulating the expression gives the real and imaginary parts of \hat{n}

$$n = \sqrt{\frac{|\hat{\epsilon}_r| + \epsilon_r}{2}} = \sqrt{\frac{|\hat{\epsilon}| + \epsilon}{2\epsilon_0}} \quad (1.17)$$

$$\kappa = \sqrt{\frac{|\hat{\epsilon}_r| - \epsilon_r}{2}} = \sqrt{\frac{|\hat{\epsilon}| - \epsilon}{2\epsilon_0}} \quad (1.18)$$

2 BASIC THEORY

2.1 THE ELECTRIC POTENTIAL, LAPLACE'S EQUATION AND THE UNIQUENESS THEOREM

The usual task of electrostatics is to compute the electric field \mathbf{E} given a stationary electric charge distribution $\rho_E(\mathbf{r}')$

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} \rho(\mathbf{r}') d\mathbf{r}'. \quad (2.1)$$

The notation for Eq.(2.1) can be shown in Figure 2.1 and $\epsilon_0 = 8.85 \cdot 10^{-12} \text{C}^2/\text{Nm}^2$ is the permittivity of free space [?, . 58-62]. However, it is usually simpler to calculate the electric potential Ψ

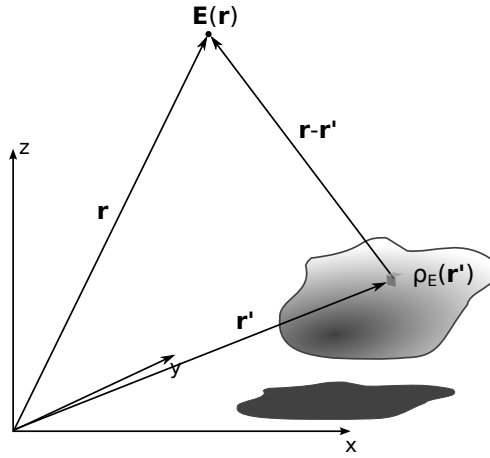


Figure 2.1: The electric field $\mathbf{E}(\mathbf{r})$ (at position \mathbf{r}) due to the charge distribution ρ_E located at \mathbf{r}' .

$$\Psi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{1}{|\mathbf{r} - \mathbf{r}'|} \rho(\mathbf{r}') d\mathbf{r}' \quad (2.2)$$

first and then calculate the electric field from

$$\mathbf{E} = -\nabla\Psi. \quad (2.3)$$

This might in some situations, where do not necessarily know ρ but only the total amount of charge, also be to tough to handle analytically. In situations like these it is better to use Poisson's equation

$$\nabla^2\Psi = -\frac{1}{\epsilon_0} \rho, \quad (2.4)$$

which together with appropriate boundary conditions, is equivalent to Eq.(2.2). Very often, we are interested in finding the potential containing no charge (because the charge is located on the outside of our region of interest. In such cases Eq. (2.4) reduces to Laplace's equation ([?], p.110-111)

$$\nabla^2\Psi = 0. \quad (2.5)$$

According to the *Uniqueness Theorems*, the solution to Laplace's equation is uniquely determined in some volume if the potential is specified on the boundary of the volume. This easily extends to Poisson's equation by further requiring, in addition to the potential on the boundary, that the charge distribution throughout the region is known.

When considering conductors, charge are allowed to move freely and might start to rearrange themselves, leading to the *Second uniqueness theorem*, which states that the potential in a given volume, surrounded by conductors is uniquely determined if the total charge on each conductor is given.

The uniqueness theorem grants an enlarged mathematical freedom in the approach of finding the potential of a region of space. This is because the boundary uniquely determines the potential in the region enclosed region and any approach giving the correct boundary conditions would give you the correct potential function through Laplace's equation Eq. (2.5). This allows the use of tricks, like for example the classical *method of images* ([?], p.116-121).

2.2 POLARIZABILITY

When a neutral atom is placed in an electric field E , the field tries to rip the atom apart by pushing the nucleus in the direction of the field and the electrons in the opposite direction. Because of the attraction between the positive and negative charge within the atom, an equilibrium displacement of the electrons compared to the nucleus is achieved, leaving the atom polarized and giving it a dipole moment. The dipole moment can be approximated by

$$\mathbf{p} = \alpha \mathbf{E}, \quad (2.6)$$

where α is the atomic polarizability and may depend on the detailed structure of the atom. For more complicated situations, like an asymmetrical molecule, the gained dipole moment of the molecule does not necessarily have to be in the same direction as the applied electric field. In such a case, the scalar polarizability in the expression above is replaced by a polarizability tensor

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{bmatrix}. \quad (2.7)$$

In this way, an applied electric field induces many dipole moments in a material. In addition, any polar molecules will be subject to a torque, aligning it to the direction of the field. These two mechanisms leads to the polarization \mathbf{P} of the material

$$\mathbf{P} = \text{dipole moment per unit volume} = \epsilon_0 \chi_e \mathbf{E}. \quad (2.8)$$

In the above expression, there has been assumed a linear dielectric media, where χ_e is the electric susceptibility and depends on the microscopic structure of the material, in addition to the external temperature ([?], p.160-166, 179)

2.3 PLASMONS

2.4 THE DRUDE MODEL [1]

Defining the differences between the electrical properties of metals, semiconductors and insulators is not a trivial matter. To put it simple, metals are good conductors, while semiconductors and insulators are not. However, semiconductors such as silicon conduct electricity relatively well, so the picture is not quite that simple.

A simple model, with fundamental importance regarding electrical conductivity, was suggested by P. Drude in an attempt to explain the observed properties of metals. It is assumed that the electrons in a solid behave like a classical gas and do not interact with each other whatsoever. Coloumb interaction is also neglected. This is known as the *independent electron approximation*. The electron gas can be viewed as negatively charged particles bouncing about immobile positively charged ion cores. The only form of interaction in this model are instantaneous collision between the electrons and ions; see Figure 2.2. The simplification of removing the Coloumb ion-electron interaction is called the *free electron approximation*. The electrons reach thermal equilibrium with the lattice through the collisions, giving them the kinetic energy:

$$\frac{1}{2} m_e v_t^2 = \frac{3}{2} k_B T. \quad (2.9)$$

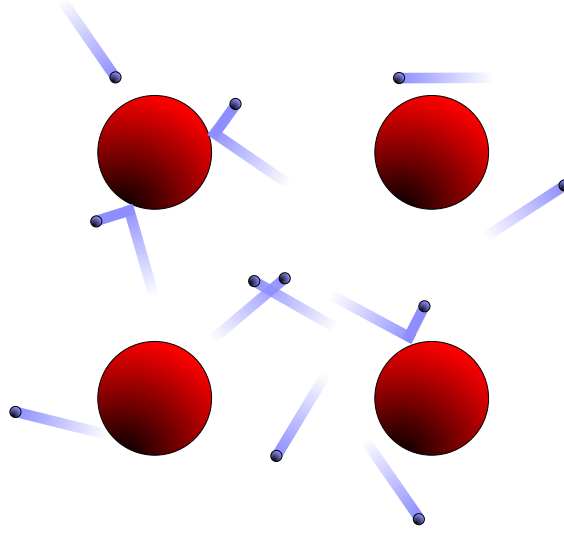


Figure 2.2: The Drude model. A non-interacting electron gas (blue) is swirling around and colliding with stationary positive ion cores (red). Since all Coloumb forces are neglected together with electron-electron collision, this ion-electron interaction is the only one being considered.

The average time between two collisions is called the relaxation time τ , with the corresponding mean free path defined as $\lambda = \tau v_t$. The probability for a electron to collide per unit time is assumed to be $1/\tau$.

It is necessary to determine the conduction electron density $\rho_{c,e}$ in order to describe properties such as electrical conductance. Assuming that every atom contributes every electron in its outermost shell N_e , e.g. that alkali metals contribute $N_e = 1$ and that earch alkaline earth metals contribute $N_e = 2$ etc. The number of atoms $N_a = \rho_m / m_a$, where ρ_m is the density of the metal in $[\text{kg}/\text{m}^{-3}]$ and m_a is the atomic mass in $[\text{kg}]$ (mass per atom).

2.5 CONDUCTIVITY; THE DRUDE MODEL

Subjecting the drude metal of an electric field \mathbf{E} will lead to a drift of the electrons

$$\frac{d\mathbf{v}}{dt} m_e = -e\mathbf{E}, \quad (2.10)$$

where e is the elementary charge. The solution is given by

$$\mathbf{v}(t) = \frac{-e\mathbf{E}t}{m_e}. \quad (2.11)$$

Assuming that the drift montion is destroyed in a collision, the average drift speed of the electrons become

$$\bar{\mathbf{v}}(t) = \frac{-e\mathbf{E}\tau}{m_e}, \quad (2.12)$$

which is very slow compared to the thermal movement of the electrons v_t .

Considering an area A perpendicular to the electric field, the amount of charge passing through the area is $-en_e|\bar{\mathbf{v}}|A$ and the resulting current density would be

$$\mathbf{j} = n\bar{\mathbf{v}}(-e) = \frac{n_e e^2 \tau}{m_e} \mathbf{E} \equiv \sigma \mathbf{E} = \mathbf{E} / \rho. \quad (2.13)$$

This is the familiar Ohm's law, with the conductivity

$$\sigma = \frac{n_e e^2 \tau}{m_e} \quad (2.14)$$

and the resistivity ρ . The e^2 dependence comes from the size of the pulling force due to the electric field, i.e. how fast the particles move, and the other is how the same charge, now moving, defines the current. This also means that we would get the same result for carriers with charge e instead of $-e$, which would be the case for semiconductors with positive holes.

2.6 OPTICAL REFLECTIVITY OF DRUDE METALS

The optical properties of materials are described by the complex refractive index $N(\omega)$ or the dielectric constant $\epsilon(\omega)$. To explain the reflectivity of metals we can consider an electron in an electromagnetic field induced by an optical incident wave with wavenumber $q = 2\pi N/\lambda_0$, where λ_0 is the wavelength of the wave in free space.

If ω is low, we basically retain the DC behavior. However, if ω is so high that $1/\omega \ll \tau$, the electrons does not manage to react fast enough and the collisions with the ions altogether can be ignored (this is fulfilled for optical frequencies when $\tau = 10^{-14}$ s). The electrons can now be treated as completely free, and a single electron follows

$$m_e \frac{d^2 x(t)}{dt^2} = -eE(t) = -eE e^{-i\omega t}. \quad (2.15)$$

A good ansatz for the solution is setting $x(t) = x e^{-i\omega t}$, which mean that the electron is oscillating along with the electric field. This gives an amplitude of

$$x = \frac{eR}{m_e \omega^2}. \quad (2.16)$$

The corresponding polarization due to the dipole moment of $-ex$ for a solid with conduction electron density n_e is

$$P = -n_e ex = -\frac{n_e e^2 E}{m_e \omega^2}. \quad (2.17)$$

From the constitutive relation

$$D = \epsilon \epsilon_0 E = \epsilon_0 E + P, \quad (2.18)$$

we get that

$$\epsilon = 1 + \frac{P}{\epsilon_0 E} = 1 - \frac{n_e e^2}{\epsilon m_e \omega^2} \equiv 1 - \text{frac} \omega_p^2 \omega^2, \quad (2.19)$$

with the plasma frequency ω_p defined as

$$\omega_p = \frac{n_e e^2}{m_e \epsilon_0}. \quad (2.20)$$

We can distinguish between two different cases and resulting outcome can be seen from the the electromagnetic plane wave

$$E(z, t) = E_0 e^{i(2\pi N z / \lambda_0 - \omega t)}. \quad (2.21)$$

In case 1) $\omega < \omega_p$ and ϵ is a real and negative, $\epsilon = \Re\{\epsilon\} < 0$. Therefore, $N = \sqrt{\epsilon}$ is purely imaginary, and the wave penetrating the solid is exponentially damped. Because Eq.(2.15) contain no inelastic properties, nothing is absorbed. The light that is not transmitted into the medium must therefore, due to energy conservation, be reflected back. For $\omega > \omega_p$, the dielectric constant is real and positive, $\epsilon = \Re\{\epsilon\} > 0$, and so is, followingly, the index of refraction. The result is a plane wave that propagates into the metal. This explains why metals are so reflective, or shiny. They are reflective for low-frequency light, but transparent for high-frequency light. The transition happens at the plasma frequency, which can be calculated solely from the conduction electron density of the metal. For most metals, the plasma frequency is in the far UV region, making them reflective in the visible range. Frequently, the plasma energy $\hbar\omega_p$ is used instead of the plasma frequency ω_p .

2.7 SHORTCOMINGS OF THE DRUDE MODEL

The questionable assumptions of the Drude model is the removal of the electron-electron interaction together with all the Coulomb forces. In addition, the assumption of treating the electrons as particles is not justified due to the fact that their de Broglie wavelength, in the case of thermal electrons, is in the order of nanometers. The assumption would however only satisfy electrons moving in structures much larger than the de Broglie wavelength.

The resulting conductivity of the model is not high enough at low temperatures, and is due to the assumption of a fixed mean free path, given by the atomic spacing. Apparently, at low temperature, the electrons manage to sneak past the other electrons and ions.

Also, the conductivity of alloys, in which impurities drastically reduce the conductivity, finds no justification in the Drude model.

2.8 DIELECTRICS

2.9 MICROSCOPIC POLARIZATION

There are several mechanisms causing microscopic electric dipole moments that lead to macroscopic polarization. E.g. displacement of electronic clouds and core, opposite displacement of the ions in a solid or orientation of permanent dipoles such as water molecules, called orientational polarization.

2.10 THE LOCAL FIELD

To calculate the microscopic polarizability α of the atoms making up the solid, start with the constitutive relation and assume that \mathbf{P} can be written as the total dipole moment per unit volume

$$\mathbf{P} = (\epsilon - 1)\epsilon_0 \mathbf{E} = \frac{N}{V} \mathbf{p} = \frac{N}{V} \alpha \mathbf{E}. \quad (2.22)$$

Because a microscopic dipole within the solid does not simply feel the average electric field \mathbf{E} but a microscopic local and stronger electric field \mathbf{E}_{loc} , the polarizability can not be correctly calculated from the above expression. Without derivation, the approximate local field can be given by

$$\mathbf{E}_{loc} = \frac{1}{3}(\epsilon + 2)\mathbf{E}. \quad (2.23)$$

So, with

$$\mathbf{P} = \frac{N}{V} \alpha \mathbf{E}_{loc} \quad (2.24)$$

and approximating the polarization \mathbf{P} with Eq.(2.22), we get the so-called Clausius-Mossotti relation

$$\alpha = \frac{\epsilon - 1}{\epsilon + 2} \frac{3\epsilon_0 V}{N}, \quad (2.25)$$

relating the atomic polarizability to the dielectric constant.

2.11 FREQUENCY DEPENDENCE OF THE DIELECTRIC CONSTANT

The frequency dependent permittivity $\epsilon(\omega)$ is usually called the dielectric function. For insulators, $\epsilon(\omega)$ is complex and energy can be resonantly transferred to the solid for certain frequencies. $\epsilon(\omega)$ implies a frequency dependence of the refractive index $N(\omega)$. Most of the frequency dependence can be explained by a simple idea combined with knowledge about the polarization mechanisms and how the different polarization mechanisms manages to keep up with the oscillating electric field. E.g. orientation polarization and ionic polarization does not manage to oscillate fast enough at higher frequencies, while the atomic polarization will, see Figure 2.3.

For a quantitative description of the frequencydependence of ϵ we can consider a simplified version of ionic vibration. Light can couple to optical phonons, e.g in ionic crystals where the phonons correspond to an out-of-phase vibration of the positive and negative ions in the unit cell. These vibrations can be approximated by independent

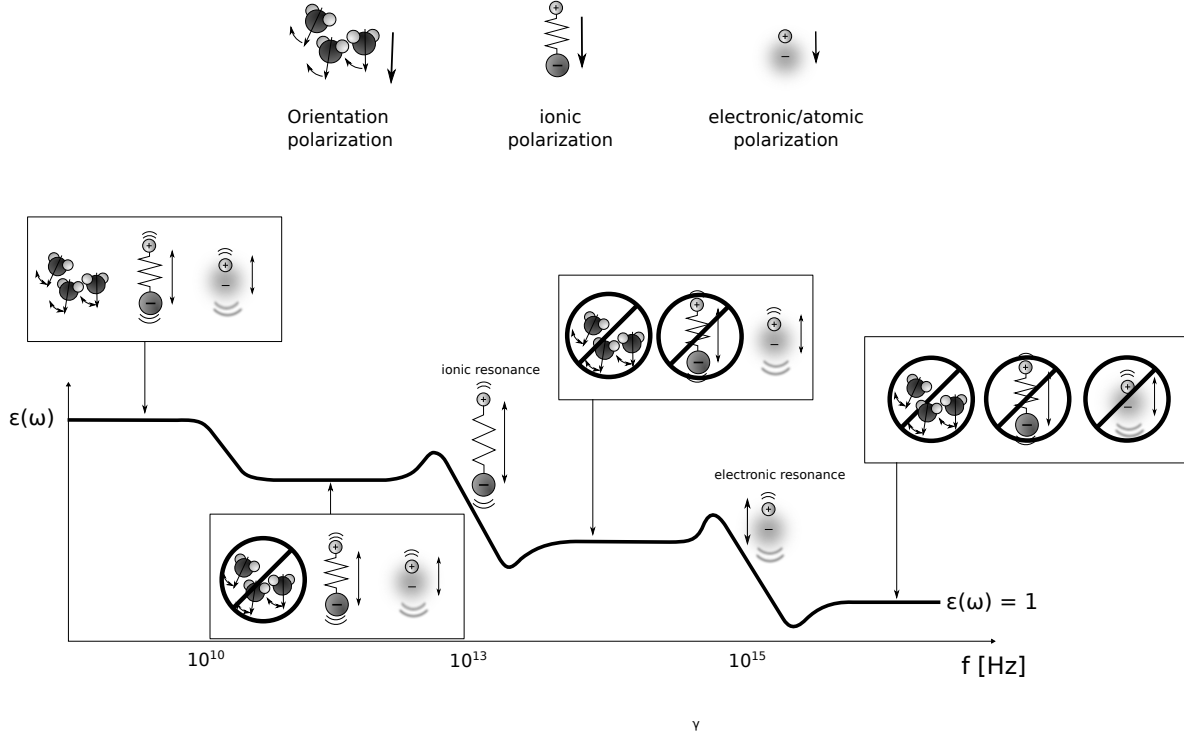


Figure 2.3: The contribution of orientation, ionic and electronic polarization. As the frequency of the applied electric field increases, the different polarization mechanisms fail to remain in step with the field when above a characteristic frequency. At sufficiently high frequencies the material no longer manages to polarize and the dielectric constant drops to 1, corresponding to the permittivity of free space. Figure is adapted from [2].

harmonic oscillators driven by an electric field $E e^{-i\omega t}$, with one such oscillator per unit cell of the crystal. Each oscillator have a resonant frequency of $\omega_o = (2\gamma/M)^{1/2}$, where γ is the force constant and M is the reduced mass of the two ions. The motion is damped by a term proportional to the velocity $\eta dx/dt$ and represents the excitation of other vibrations in the material, due to the large displacement. The resulting equation of motion is that of a driven harmonic oscillator with damping

$$\frac{d^2x}{dt^2} + \eta \frac{dx}{dt} + \omega_o^2 x = \frac{eE}{M} e^{-i\omega t}. \quad (2.26)$$

A good ansatz for the solution is

$$x(t) = A e^{-i\omega t}, \quad (2.27)$$

resulting in the amplitude

$$A = \frac{eE}{M} \frac{1}{\omega_o^2 - \omega^2 - i\eta\omega} \quad (2.28)$$

$$= \frac{eE}{M} \left[\frac{\omega_o^2 - \omega^2}{(\omega_o^2 - \omega^2)^2 + \eta^2\omega^2} + \frac{i\eta\omega}{(\omega_o^2 - \omega^2)^2 + \eta^2\omega^2} \right] \quad (2.29)$$

Using this as the ionic vibration, we can calculate the total polarization for a crystal with N unit cells and volume V . Considering only one type of ions with a density N/V and effective atomic polarizability α , assuming both ionic

and atomic polarization, $P_i(\omega)$, $P_a(\omega)$, the result reads

$$P(\omega) = P_i(\omega) + P_a(\omega) = \frac{N}{V} e A(\omega) e^{-i\omega t} + \frac{N}{V} \alpha E e^{-i\omega t}. \quad (2.30)$$

When dealing with two types of ions like in a NaCl crystal, the different polarizabilities can be taken care of by a suitable definition of α . The resulting dielectric function can be calculated from

$$\epsilon_0 \epsilon(\omega) E(\omega) = P(\omega) + \epsilon_0 E(\omega), \quad (2.31)$$

giving

$$\epsilon(\omega) = \frac{P(\omega)}{\epsilon_0 E e^{-i\omega t}} + 1 \quad (2.32)$$

$$= \frac{NeA(\omega)}{V\epsilon_0} + \frac{N\alpha}{V\epsilon_0} + 1 \quad (2.33)$$

$$= \frac{NeA(\omega)}{V\epsilon_0} + \epsilon_{opt}. \quad (2.34)$$

Here, ϵ_{opt} is the high frequency or optical limit, where the fields move too quickly for the ions to respond and $P_i(\omega) = 0$, i.e.

$$\epsilon_{opt} = \lim_{\omega \rightarrow \infty} \epsilon(\omega) = \frac{N\alpha}{V\epsilon_0} + 1. \quad (2.35)$$

Plugging in the expression for $A(\omega)$ the real and complex values of the dielectric function $\epsilon(\omega) = \epsilon_{re}(\omega) + i\epsilon_{im}(\omega)$ may be written as

$$\epsilon_{re}(\omega) = \frac{Ne^2}{V\epsilon_0 M} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \eta^2 \omega^2} + \epsilon_{opt}, \quad \epsilon_{im}(\omega) = \frac{Ne^2}{V\epsilon_0 M} \frac{\eta \omega}{(\omega_0^2 - \omega^2)^2 + \eta^2 \omega^2} \quad (2.36)$$

The behaviour of the result is shown in Figure 2.4. The real part $\epsilon_{re}(\omega)$ is almost constant away from the resonance

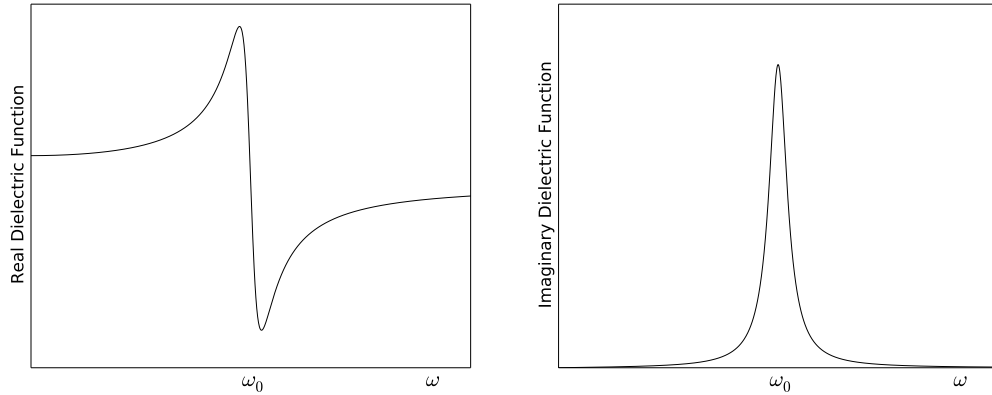


Figure 2.4: The dielectric function of an ionic crystal approximated by a driven harmonic oscillator with damping. The left and right figures show the behavior of the real and imaginary dielectric function close to the resonance frequency ω_0 .

frequency, but its value is higher at lower frequencies due to the loss of the contribution from the ionic polarization. The imaginary part $\epsilon_{im}(\omega)$ is however zero everywhere except at the vicinity of the resonant frequency, where it shows a peak with a width given by the damping coefficient η .

To understand the meaning of $\epsilon_{im}(\omega)$ and that the width of its resonance peak is connected to the damping coefficient η , one can consider the energy dissipation in the system. The instantaneous electrical power dissipated per unit volume is given by

$$P(t) = j(t)E(t) = j(t)E e^{-i\omega t} \quad (2.37)$$

where $j(t)$ is the current density. In an insulator, there are no free currents, only polarization currents

$$j(t) = -\frac{\partial D}{\partial t} = -\frac{\partial}{\partial t} \epsilon \epsilon_0 E e^{-i\omega t} = i\omega \epsilon \epsilon_0 E e^{-i\omega t} \quad (2.38)$$

The average dissipated power is found by averaging over one cycle $T = 2\pi/\omega$

$$P = \frac{1}{T} \int_0^T E(t) j(t) dt. \quad (2.39)$$

If ϵ is purely imaginary, $j(t)$ is out of phase with $E(t)$ and their product will always give a nonzero negative value, $-\epsilon_0 \epsilon_{3m}(\omega) E^2$. On the other hand, if ϵ is purely real, the phase shift will be $\pi/2$ and the integral will give $P = 0$. $\epsilon_{3m}(\omega)$ is therefore a measure of the energy dissipation of the electric field due to the solid, and is obviously highest at the resonance.

The discussion explains some of the optical behaviour of the material. This is however not the entire picture. E.g. the frequency dependence in the visible and UV region is not explained here. These effects are due to the valence electrons, which would need a quantum mechanical description of the electronic structure of the solid. From the band structure of solids, a qualitative understanding is obtainable. Figure 2.5 shows regions of high photon absorption due to the excitation of electrons, and how it is given by the imaginary dielectric function $\epsilon_{3m}(\omega) \epsilon_i(\omega)$. Note that the photons do not have enough energy to change the electrons wave vector

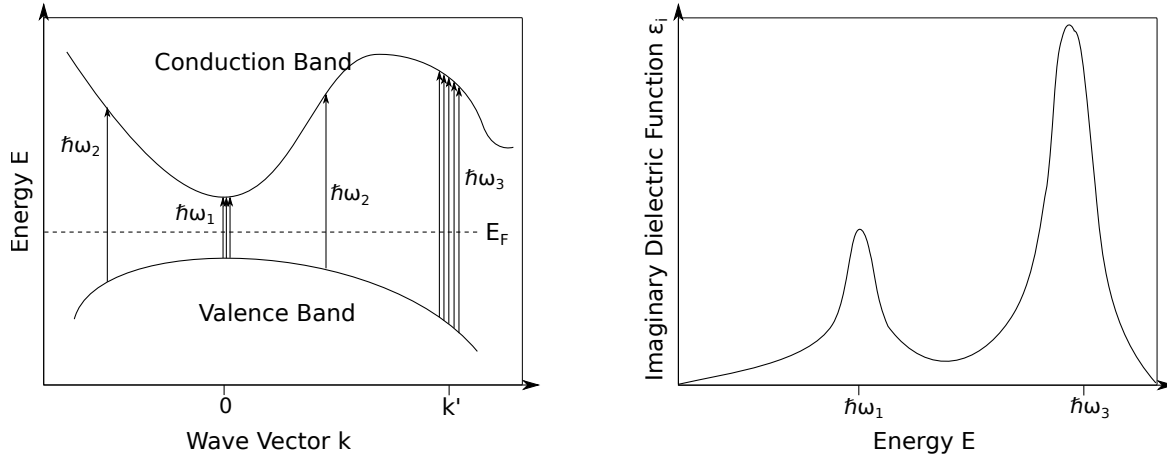


Figure 2.5: The left figure shows the photon-induced transitions between occupied and unoccupied states in the band structure of a solid. E_F is the fermi energy. In the regions where the valence band and conduction bands are parallel, a certain photon energy can excite several states. Here, the parallel regions are located at $\mathbf{k} = 0$ and $\mathbf{k} = \mathbf{k}'$ and result in higher transitions density. These transitions correspond to absorption of the electromagnetic wave given by the imaginary part of the dielectric function ϵ_i . The resulting resonances in ϵ_i due to the absorption of the $\hbar\omega_1$ and $\hbar\omega_2$ transitions are depicted in the right figure.

REFERENCES

- [1] Hofmann P. Solid State Physics, An Introduction. Wiley-VCH 2008; p.71-74,76-81

2.12 COMPLEX PERMITTIVITY AND INDEX OF REFRACTION [1, P. 169-170]

The complex dielectric constant or relative permittivity $\hat{\epsilon}_r = \epsilon_r + i\tilde{\epsilon}_r$ of a material, is a measure of the material's response subject to an electromagnetic field. Here ϵ_r and $\tilde{\epsilon}_r$ denotes the real and imaginary components, respectively. The relative permittivity is related to the square of the refractive index,

$$N^2 = \hat{\epsilon}_r \quad (2.40)$$

which determines the optical properties of a given material. It is like the dielectric constant complex

$$N = n + ik. \quad (2.41)$$

The real and imaginary refractive indices are denoted by n and k , respectively. The choice of sign convention, i.e. using $n + ik$ rather than $n - ik$, is determined by the choice of the sign in the plane wave solution, $\exp i(\mathbf{q} \cdot \mathbf{r} - \omega t)$, of Maxwell's equations. From Eq.(6.3)

$$\hat{\epsilon}_r = N^2 \quad (2.42)$$

$$\epsilon_r + i\tilde{\epsilon}_r = (n + ik)^2 \quad (2.43)$$

$$\epsilon_r + i\tilde{\epsilon}_r = n^2 - k^2 + i2nk, \quad (2.44)$$

and with some simple comparison, the components of the permittivity can be expressed as

$$\epsilon_r = n^2 - k^2 \quad \tilde{\epsilon}_r = 2nk. \quad (2.45)$$

Taking the absolute value or modulus,

$$|\hat{\epsilon}_r| = \sqrt{\epsilon_r^2 + \tilde{\epsilon}_r^2} \quad (2.46)$$

$$|\hat{\epsilon}_r| = \sqrt{(n^2 - k^2)^2 + (2nk)^2} \quad (2.47)$$

$$|\hat{\epsilon}_r|^2 = n^4 + 2n^2k^2 + k^4 \quad (2.48)$$

$$|\hat{\epsilon}_r|^2 = (n^2 + k^2)^2 \quad (2.49)$$

$$|\hat{\epsilon}_r| = n^2 + k^2, \quad (2.50)$$

and putting it all together, gives the real and imaginary parts of N expressed through the relative permittivity

$$n = \left(\frac{|\hat{\epsilon}_r| + \epsilon_r}{2} \right)^{\frac{1}{2}} \quad k = \left(\frac{|\hat{\epsilon}_r| - \epsilon_r}{2} \right)^{\frac{1}{2}} \quad (2.51)$$

Considering a plane wave with a complex wave vector $\hat{q} = q + i\tilde{q}$ moving in the material [?, p. 402]

$$e^{i(\hat{q}z - \omega t)} = e^{-\tilde{q}z} e^{iqz - \omega t}, \quad (2.52)$$

one sees that the wave is attenuated. The quantity $\alpha \equiv 2\tilde{q}$ is called the absorption coefficient and is proportional to the optical conductivity σ , to $\tilde{\epsilon}$, and to k :

$$n\alpha = \frac{4\pi\sigma}{c} = \frac{\omega}{c} \tilde{\epsilon} \quad (2.53)$$

$$\Downarrow \quad (2.54)$$

$$\frac{\alpha}{2} = \frac{\omega}{c} k = \frac{1}{\delta} \quad (2.55)$$

k is usually called the extinction coefficient and is essentially the ratio of the free-space wave frequency ω to the skin depth δ .

n and k can be found experimentally by measuring the reflectivity R of a bulk, opaque sample, in addition to the transmittance T of a slab, which are given in terms of n and k as

$$R = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2} \quad (2.56)$$

$$T = \frac{(1-R)^2 e^{-2\omega kd/c}}{1 - R^2 e^{-4\omega kd/c}}, \quad (2.57)$$

where d is the sample thickness. The slab multiple-reflection effects are averaged, so that interface fringes are not resolved.

REFERENCES

- [1] Jensen B. The quantum extension of the Drude-Zener theory in polar Semiconductors. Handbook of optical constants of Solids, Five-Volume 1997 (1985??)(9)<-it's chapter 9;169-170
- [2] . Dissemination of IT for the promotion of Materials Science (2000) Variation of the dielectric constant in alternating fields. University of Cambridge 2004: <http://www.doitpoms.ac.uk/tlplib/dielectrics/variation.php> (24. August 2015).

3 PLASMONS

Notes from Justin White <http://large.stanford.edu/courses/2007/ap272/white1/>

Surface Plasmon Polaritons

- Surface Plasmon polaritons are collective longitudinal oscillations of electrons near a material surface, strongly coupled to an electromagnetic wave.
- Both bulk and surface plasmons have associated EM-waves, and can consequently be described by Maxwell's equations.
- the coherent oscillations of electron motion can be encapsulated in the dielectric constant of the material.
- The basic form of the bulk and surface plasmon solutions are shown below:

$$E_{bulk} = E_0 e^{k_x x - \omega t} \quad (3.1)$$

$$E_{spp} = E_0 e^{-\kappa|z|} e^{k_x x - \omega t} \quad (3.2)$$

The bulk plasmons are associated with purely transverse EM waves ($E \perp k$ and $B \perp k$) and can only exist for $\omega < \omega_p$. For $\omega > \omega_p$: the wave-vector for bulk plasmons becomes imaginary, giving an exponentially decaying wave, instead of a propagating wave. It is for this reason that most metals are highly reflective for visible light ($\omega < \omega_p$), but transparent for ultraviolet light ($\omega > \omega_p$)

- Surface plasmons have an associated EM wave with both transverse and longitudinal field components. Such waves can only be excited at the interface between a conductor and dielectric, and are tightly bound to the surface. The field reaches their maximum at the interface ($z = 0$), and exponentially decay away from the surface.
- The wave-vector of the surface plasmon mode k_{spp} always lies to the right of the free space wave-vector k_0 (in the dispersion diagram/relation), such that $\lambda_{spp} < \lambda_0$. This makes it impossible to directly launch a surface plasmon wave by illumination with free-space radiation, because the free-space photons simply do not have enough momentum to excite the surface plasmon.

As ω increases, k_{spp} gets larger and larger, moving further away from k_0 (making it harder and harder for light to excite the surface plasmons). As k_{spp} increases, the surface plasmon wave is more tightly bound to the surface. This process has an upper limit of ω_{sp} , the surface plasmon resonant frequency, which occurs when the dielectric constant of the metal and the dielectric have the same magnitude but opposite signs.

• Excitation of Surface Plasmons

High energy electrons that bombard a thin metallic film can launch surface plasmons and a surface plasmon of a whole range of wavelengths can be excited. However, only plasmons far along the dispersion curve, where k_{spp} is largest are generally excited.

As mentioned previously, direct excitation of surface plasmons by free-space photons is not achievable because k_{spp} is always greater than k_0 ; this can be seen from the dispersion relation, where the surface plasmon dispersion relation always lies to the right of the free space dispersion curve.

This can be overcome by back-side illumination through a material with a higher index of refraction n , where the far field radiation has a larger wave-vector ($k = nk_0$) (like done in a Kretschmann-Raether coupler) [5]. A surface plasmon will be efficiently excited when

$$k_{\parallel} = nk_0 \sin \theta = k_{spp} \quad (3.3)$$

A more general approach to launch surface plasmons with light is the use of structured surfaces that are able to impart momentum on the photon, enabling it to couple to the surface plasmon mode. Anything from a single sub-wavelength disk or slit, to rectangular or sinusoidal diffraction gratings are used for this type of coupling. **A thorough overview of surface plasmon coupling and patterned and rough surfaces is given by Raether [6]**

- Appendix: Derivation of Bulk and Surface Plasmons (see article).

Other Nice sources:

- Really nice article on Plasmons:
<http://nanocomposix.com/pages/plasmonics>

- Article:
<http://cdn.intechopen.com/pdfs-wm/44351.pdf>
- Surface Plasmon Theory(book?):
<https://www.physik.hu-berlin.de/de/nano/lehre/Gastvorlesung>
- Chemistry-blog:
<http://www.chemistry-blog.com/2007/03/19/plasmonics-part-ii/>
- Mie theory:
http://www.orc.soton.ac.uk/publications/theses/1460T_lnn/1460T_lnn_03.pdf

From Wikipedia

- **What are Plasmons?**

A plasmon is a quantum of plasma oscillation (quasiparticle from the quantization of plasma oscillations). Plasmons are collective (a discrete number) oscillations of the free electron gas density. Plasmons can also couple with a photon to create another quasiparticle called a plasma polariton (electromagnetic wave - electric/magnetic dipole-carrying excitation - coupling).

Plasmons can be described as oscillations of free electron density with respect to fixed positive ions in a metal. Imagine a cube of metal placed in an external electric field pointing to the right. Electrons will move to the left side and uncover positive ions on the right side. The electrons will continue moving left until they cancel the field inside the metal. Removing the field will make the electrons move back by their mutual repulsion and attraction to the ions, leaving the electrons to oscillate back and forth, at the **plasma frequency**, in a so called plasma oscillation.

- **Plasma Oscillation, aka "Langmuir waves"**

Rapid oscillations of electron density in conducting media such as plasmas or metals. The oscillations can be described as an instability in the dielectric function of a free electron gas. The frequency depends weakly on the wavelength of the oscillation.

'Cold' electrons (plasma oscillations)

If the thermal motion of the electrons is ignored and assuming infinite ion mass, the charge density oscillates at the plasma frequency

$$\omega_{pe} = \sqrt{\frac{n_e e^2}{m^* \epsilon_0}}, [\text{rad/s}] \text{ (SI-units)} \quad (3.4)$$

$$\omega_{pe} = \sqrt{\frac{4\pi n_e e^2}{m^*}}, (\text{cgs-units}), \quad (3.5)$$

where n_e is the number density of electrons, e is the electric charge, m^* is the effective mass of the electron and ϵ_0 is the permittivity of free space. Since the frequency is independent of the wavelength, these oscillations have an infinite phase velocity and zero group velocity. Note in addition that, if m^* is the electron mass m_e , the plasma frequency ω_{pe} depends only on the physical constants and concentration of electrons n_e . The numeric expression is:

$$f_{pe} = \frac{\omega_{pe}}{2\pi} \approx 8980 \sqrt{n_e}, [\text{Hz}] \quad (3.6)$$

with n_e in $[\text{cm}^{-3}]$

'Warm' electrons (plasma oscillations)

When the effects of the electron thermal speed $v_{e,th} = \sqrt{\frac{k_B T_e}{m_e}}$ are taken into account, the electron pressure acts as an additional restoring force and the oscillations propagate with frequency and wavenumber related by the longitudinal Langmuir wave:

$$\omega^2 = \omega_{pe}^2 + \frac{3k_B T_e}{m_e} k^2 = \omega_{pe}^2 + 3k^2 v_{e,th}^2 \quad (3.7)$$

called the 'Bohm-Gross dispersion relation'. If the spatial scale is large compared to the Debye length (measure of a charge carrier's net electrostatic effect in solution, and how far those electrostatic effects persist), the oscillations are only weakly modified by the pressure term, but at small scales the pressure term dominates and the waves become dispersionless with a speed of $\sqrt{3}v_{e,th}$. For such waves, however, the electron thermal speed is comparable to the phase velocity, i.e.

$$v \sim v_{p,th} \equiv \frac{\omega}{k}, \quad (3.8)$$

so the plasma waves can accelerate electrons that are moving with speed nearly equal to the phase velocity of the wave. This process often leads to a form of collisionless damping called Landau damping. Consequently, the large- k portion in the dispersion relation is difficult to observe and seldom of consequence.

In metal or semiconductor, the effect of the ions periodic potential must be taken into account. This is usually done by using the electrons effective mass in place of m .

- **Role of Plasmons**

Plasmons play a large role in the optical properties of metals. Light of frequencies below the plasma frequency is reflected, because the electrons in the metal screen the electric field of the light. Light of frequencies above the plasma frequency is transmitted, because the electrons cannot respond fast enough to screen it. In most metals, the plasma frequency is in the ultraviolet, making them shiny (reflective) in the visible range. In semiconductors, the valence electron plasma frequency is usually in the deep ultraviolet, which is why they are reflective.

The plasmon energy can often be estimated in the free electron model as

$$E = \hbar \sqrt{\frac{n_e e^2}{m^* \epsilon_0}} = \hbar \omega_p \quad (3.9)$$

- **Surface Plasmons (SPs)**

Surface plasmons (plasmons at the interface of two materials) interact strongly with light, resulting in a polariton (usually occurs at metal or doped dielectric interface, which both have small $\text{Im}(\epsilon) > 0$ and big $\text{Re}(\epsilon) < 0$). These surface electron oscillations can exist at the interface between any two materials where the real part of the dielectric function changes sign across the interface (e.g. a metal-dielectric interface like metal sheet in air).

SPs have lower energy than **bulk (or volume)** plasmons, which quantise the longitudinal electron oscillations about positive ion cores within the bulk of an electron gas (or plasma).

The charge motion in a surface plasmon always creates electromagnetic fields outside (as well as inside) the metal. The total excitation, including both the charge motion and associated electromagnetic field, is called either a **surface plasmon polariton** at a planar interface, or a **localized surface plasmon** for the closed surface of a small particle.

Surface Plasmon polaritons can be excited by electrons or photons. In the case of photons, it cannot be done directly, but requires a prism, or a grating, or a defect on the metal surface. ??? Or like truncated spheres on granular films???

At low frequency an SPP approaches the dispersion relation in free space $\omega = ck$. At high frequency, the dispersion relation reaches an asymptotic limit called the "surface plasma frequency".

As an SPP propagates along the surface, it loses energy to the metal due to absorption and due to scattering into free-space or into other directions. The electric field falls off evanescently perpendicular to the metal surface. At low frequencies, the SPP penetration depth into the metal is commonly approximated using the 'skin depth formula'. In the dielectric, the field will fall off far more slowly. SPPs are very sensitive to slight perturbations within the skin depth and because of this, SPPs are often used to probe inhomogeneities of a surface.

Surface plasmons have been used to control colors of materials and is possible since controlling the particle's shape and size determines the types of surface plasmons that can couple to it and propagate across it. This in turn controls the interaction of light with the surface. These effects are illustrated by the historic *stained glass* which adorn medieval cathedrals. In this case, the color is given by metal nanoparticles of a fixed size which interact with the optical field to give the glass its vibrant color. To produce optical range surface plasmons effects involves producing surfaces which have features $< 400\text{nm}$.

Surface plasmons are very sensitive to the properties of the materials on which they propagate.

- **Surface Plasmons Resonance (SPR)**

Surface plasmon resonance is the resonant oscillation of conduction electrons at the interface between a negative and positive permittivity material stimulated by incident light. The resonance condition is established when the frequency of incident photons matches the natural frequency of surface electrons oscillating against the restoring force of positive nuclei.

Surface plasmon polaritons are surface electromagnetic waves that propagate in a direction parallel to the metal/dielectric (or metal/vacuum) interface. Since the wave is on the boundary of the metal and the external medium, these oscillations are very sensitive to any change of this boundary, such as adsorption of molecules to the metal surface.

To describe the existence and properties of surface plasmon polaritons, one can choose from various models, e.g. the **Drude Model**. The simplest way to approach the problem is to treat each material as a homogeneous continuum, described by a frequency-dependent relative permittivity between the external medium and the surface (this is a complex dielectric function). In order for the terms that describe the electronic surface plasmons to exist, the real part of the dielectric constant of the metal must be negative and its magnitude must be greater than that of the dielectric. This condition is met in the infrared-visible wavelength region for air/metal and water/metal interfaces (where the real dielectric constant of a metal is negative and that of air or water is positive).

Localized SPRs (LSPRs) are collective charge oscillations in metallic nanoparticles that are excited by light. They exhibit enhanced near-field amplitude at the resonance wavelength. This field is highly localized at the nanoparticle and decays rapidly away from the nanoparticle/dielectric interface into the dielectric background, though far-field scattering by the particle is also enhanced by the resonance. Light intensity enhancement is a very important aspect of LSPRs and localization means the LSPR has very high spatial resolution (subwavelength), limited only by the size of nanoparticles. Because of the enhanced field amplitude, effects that depend on the amplitude such as magneto-optical effect are also enhanced by LSPRs.

In order to excite surface plasmons in a resonant manner, one can use an electron or light beam (visible and infrared are typical). The incoming beam has to match its momentum to that of the plasmon. With p-polarization this is possible by passing the light through a block of glass to increase the wavenumber (and the momentum) and achieve the resonance at a given wavelength and angle. s-polarized light however cannot excite electronic surface plasmons.

When the surface plasmon wave interacts with a local particle or irregularity, such as a rough surface, part of the energy can be re-emitted as light. This emitted light can be detected behind the metal film from various directions.

- **The Drude Model**

Treats the behavior of electrons in a solid like a pinball machine. The electrons are small light balls in a sea of static, positively charged ions. The only form of action instantaneous collisions.

- **Mie Scattering**

Mie theory is sometimes used for the collection of methods and solutions to Maxwell's equations for scattering, by e.g. using geometries where one can write separate equations for the radial and angular dependence of solutions. More broadly, "Mie scattering" suggests situations where the size of the scattering particles is comparable to the wavelength of the light, rather than much smaller or much larger.

REFERENCES

[1]

4 ARTICLE NOTES

4.1 THERMOCHROMISM

[2] Intelligent Thermochromic Windows

The use of air-conditioning systems to maintain comfortable working and living environments has become more common [1]. This leads to an increase in the use of electricity and a concurrent increase in carbon dioxide emissions and other atmospheric pollutants formed in the electricity generation process. A self-propagating cycle results, in which global warming due to increases in these greenhouse gases necessitates the increased use of air conditioning systems. Technology is thus required that can reduce the use of air conditioning commercial and residential buildings to help break this cycle.

(...) window coatings can reduce cooling costs or heating requirements [2]. Using thermochromic coatings as intelligent window coatings [1-7], which change their optical properties with temperature; usually related to a structural phase change on passing through a critical temperature T_c . Thermochromic coatings would be applicable to climates where there are extreme changes in temperature over the year, for example, central and northern Europe, Japan, the United States, and Canada, which have hot summers and cold winters.

Vanadium(IV) oxide; transition temperature $T_c = 68^\circ\text{C}$; visually and infrared transparent below $T_c \rightarrow$ solar radiation passes through, keeping the interior warm. Below T_c it becomes infrared reflective and preventing excessive heating, while remaining visually transparent.

Critical temperature for vanadium is too high, but this can be lowered to 25°C using dopants ([9]), most efficiently with tungsten (loading of only 2 atom percent required), in thin films prepared by physical vapor deposition methods [10] and sol-gel spin or dip coating [11].

problem: low luminous transmittance of the glazing VO_2 film [10-13]. (could be solved with doping [4] or anti-reflective coating ([12])). Also one needs a method where the thin films of the material can be applied cheaply and efficiently to the glass ([17]).

(p.394) Discussion of MST (metal-to-semiconductor transition) of VO_2 and structure changes through the MST. Discussion involves structure figures.

Goodenough proposes antiferroelectric transition being the driving force for the MST in VO_2 . \rightarrow two transition temperatures: one due to antiferroelectric distortion and one due to the crystallographic distortion.

The next paragraph explains how doping of various elements varies the MST temperature. The most effective dopant in reducing the temperature is Tungsten (additional info about tungsten and after that it considers other dopants).

Thinner thickness, stress and strain can also reduce the thermochromic transition temperature.

A little bit on VO_2 thin film durability?

Methods of preparing Pure and doped Vanadium(IV) Oxide Films:

Sol-Gel Method: forming thin films by dip- or spin-coating substrates with solutions of metal alkoxides.

PVD Method: energetically removing atoms/molecules under reduced pressure conditions, then to react with seed gas.

CVD method: chemical vapor deposition, in particular atmospheric pressure CVD (APCVD).

APCVD: deposit thin solid films from gaseous precursors onto a suitable substrate. (+Pictorial representation) (and more on APCVD).

Comparison of the above methods.

3 atom percent tungsten(VI) \rightarrow transition temperature reduced to 5°C .

1.9 percent $\rightarrow 29^\circ\text{C}$.

transition temperature decreases linearly with tungsten atom percent incorporation (Figure).

Summary: intelligent TC glass with desired switch temperature ($25 - 30^{\circ}\text{C}$), obtainable using APCVD. Most of the problems regarding commercial use are solvable. Market in household, offices, factories and space exploration.

[3] Thermochromism in Commercial Products

- Thermochromic liquid crystals: Periodicity between layers, PITCH, and constructive interference!. TC liquid crystals can have a versatile range of colors and useful color changes between -30 and 120°C , often with very high temperature sensitivity. TC liquid crystals are only useful when they are in the liquid crystalline phase, which is a meso-phase (an intermediate phase of matter) between an isotropic liquid (high temperature) and crystalline solid (low temperature), which restricts the temperature range of their applicability.
- **Microencapsulation**: Defined as the coating of small solid particles, liquid droplets, or gas bubbles with a thin film or coating or shell material, and typical particle sizes are 1 to $1000\text{ }\mu\text{m}$ ([20] Kirk-Othmer Encyclopedia of Chem. Tech. 4thEd). For three component organic mixtures, particle sizes are $< 50\text{ }\mu\text{m}$. **Micro encapsulation allows the additional advantage of combinations of several narrow color ranges**, and very sharp color changes, as well as protection of the coloring agent from the environment ([16] Nakasuji med flere. Chem.Abs.). **complex coacervation** and **interfacial polymerization** \rightarrow processes to microencapsulate thermochromic materials! Also described! Nice to include if I use thin layer on my granular film!
- Smart window candidates: Fe_3O_4 , FeSi_4 , NbO_2 , NiS , Ti_2O_3 and VO_2 , which owe their temperature change to a semiconductor-to-metallic state transition (aka **Mott transition temperature**)

[4] A Qualitative Description of Thermochromism in Color Measurements

- Wyszecki and Stiles stated ([2] color science), for TC transmitting filters, that the spectral transmittance curve at a given wavelength with a large positive slope usually decreases with increasing temperature. As a rule: steeper(positive) slope \rightarrow greater temperature effect!
The curve with negative slope is of minor importance, but often causes transmittance increase for increasing temperature (if it is important).

Neutral samples (gray, white, black) did not exhibit TC, because their spectral reflectance curves have a small or no slope.

- Assuming nonfluorescent, linear material which is "nice" with respect to polarization effects, then

$$1 = R_\lambda + A_\lambda + T_\lambda$$

where, R_λ is the reflectance, A_λ is the absorbance and T_λ is the transmittance of the sample.

Considering **transmitting samples**, the intensity I_λ transmitted through a sample of thickness d is given by

$$I_\lambda = I_{0\lambda} e^{-\mu_\lambda d}$$

where μ_λ is the absorption coefficient of the sample at wavelength λ . The optical density D is then given by:

$$D = -\log(T_\lambda) = -\log\left(\frac{I_\lambda}{I_{0\lambda}}\right) = \mu_\lambda d \log e = \frac{\mu_\lambda}{\ln 10}$$

For **opaque samples** there is no transmittance, but the optical density can be calculated from the reflected intensity $I_{R\lambda}$:

$$D_\lambda = -\log\left(\frac{I_{R\lambda}}{I_{0\lambda}}\right).$$

The light reflected from the material is also exponentially attenuated, for opaque materials. Thus, $D = \frac{\mu_\lambda}{\ln 10}$ holds for reflected light if d is the distance the reflected light has passed in the material.

[?] A Qualitative Description of Thermochromism in Color Measurements

-

[?] Performance, materials and coating technologies of thermochromic thin films on smart windows

- A significant amount of energy is consumed to maintain thermal comfort in buildings, a huge portion which is lost through windows. smart windows obtained by thin films is the solution. The touchstone of performance is the change in visible and infra-red transmission and reflectance!
- A significant amount of the energy consumption in buildings are mainly due to HVAC (heating, ventilation and air conditioning) devices, used to obtain thermal comfort, The building energy consumption in developed countries accounts for 20-40% of the total energy use. (including further details of US and China energy consumption). The building energy consumption is even more dominant in hot and humid regions, using one-third to half of the electricity produced in some countries. Energy related carbon dioxide emission. Motivates energy saving measures to reduce building energy losses and CO₂ emissions.
- Two approaches to increase energy efficiency (7-10)
 - Active strategies: improving HVAC systems and building lighting.
 - Passive strategies: improving the thermal properties of the building envelope (elements separating the indoor from outdoor), i.e. thermal insulation to wall, cool coatings on roofs and coated window glazings.
- Windows are known as one of the most energy inefficient components of buildings.(11)
Improving the thermal performance of windows will result in reduced electricity costs and less greenhouse gas emissions.
In addition to controlling transmitted IR radiation an ideal window should be capable of sufficient transmission of visible light(12).
Improving glazing characteristics of windows such as thermal transmittance and solar parameters is the most important criterion to be considered in building windows standards (14).
- International and local standards related to energy and lighting performance of windows TABLE.
- **Smart windows** defined as the type of windows that partially block the unwanted solar radiation. The energy performance can be improved by increasing heat gain in cold weather and decreasing it in hot weather by adopting windows radiative and thermal properties dynamically (25). Adding a controllable absorptive layer on the surface of the glass can change the optical properties of the glass by controlling the incident solar heat flux(26). Therefore, smart windows lead to reduced HVAC energy consumption, size and electric demand of the building (11,27,28).
- (Low emissivity (low-E) coatings are spectrally selective films that are aimed to let the visible light pass through and block the IR and UV-wavelengths which generally create heating(10). Typically, there are two types of these coatings: the tin oxide based hard coating and the silver based soft coating with higher IR reflectance and lower transmittance. The visible transmittance of hard coatings can be boosted with anti-reflecting silicon dioxide (29).)
- **The switchable reflective devices** (also called dynamic tintable windows) are categorized into **passive-** and **active systems**:
 - Passive devices: the switching process is activated automatically in accordance with the environmental conditions, e.g. temperature and heat in thermochromic windows.
 - Active systems: Require an external triggering mechanism to perform the modulation. For instance, electricity is the actuating signal in electrochromic windows. The active switchable glazing systems offer supplementary options compared to the passive systems whereas their dependency on power supply and wiring should be reckoned with as a drawback.
- Chromic material, liquid crystals and suspended particle windows are the three most common active controlled intelligent windows (11). (Chromic materials = electrochromic(active), gasochromic(active), photochromic and thermochromic.
- Providing a see-through mode is a must in any application.
- (p.356) The technology using liquid crystals in intelligent windows is called Polymer. Dispersed liquid crystals (PDLC).

- Electrochromic windows and thermochromic windows demand the lowest cooling energy, where the former require less energy for lighting than the latter (69). **Figure 1 (24) (nice figure comparing TC to the other chromic glazings together with clear glass, tinted glass and reflective glass) .**
- **Thermochromic Windows: ALSO NOTE PAGE 357! ALOT OF NICE FIGURES!!!**
 Word originates from the Greek roots: "thermos" meaning warm or hot; and "Chroma" which means color. Generally TC materials change color in response to temperature variations.
 The TC thin film is initially in its monoclinic state(cold state) at lower temperatures (usually room temperature). Monoclinic materials behave as semiconductors, less reflective especially in the near-IR (NIR) radiation. As the temperature becomes higher than a certain point, the TC material changes its nature from monoclinic to rutile state(hot state), where the material acts like a semi-metal, reflecting a wide range of solar radiation (76). **FIGURE 3** The transition is called **metal to semiconductor transition (MST)**.
- **Figure 4**, The majority of the heat gain in solar spectrum takes place at NIR range (800-1200 nm) (78-80). The red line(line 1) indicates the transmittance of a perfect TCW in cold state. Visible light should be transmitted and NIR should be reflected. Long wave radiation is also reflected back to indoor. This transmittance approach leads to reduction of solar heat gain and is apt in nearly all climates.
 The blue line(line 2) indicates the transmittance of a perfect TCW in its hot state. Visible and near infrared radiation are transmitted, while long-wave infrared is reflected to inside. This transmittance mode is suitable in low temperature climates where solar heat gain is desired. Therefore, in high temperatures, TCW reduce NIR and far-IR transmittance, while in low temperatures they allow these parts of solar adiation to pass (82), (Figure 5).
 The MST is fully reversible, co-occurred with large variations in electrical and optical properties in NIR range (83). The MST temperature should decrease to near the ambient temperature. Doping metal ions into the lattice of TC materials can alter the transition temp(84,85). The size and charge (84,86,87) of dopant ion, film's strain (88,89) as well as the variations in electron carrier density are the determinant factors prevailing on the fall or rise of the transition temperature (90).
- The **Ideal spectral behavior of TCW** is presented in **Table 3**. The visible transmission and reflectance should be equal on both sides of transition, while the infra-red variations are from 0% to 65%. The change in transmittance ($\Delta T\%$) and reflectance ($\Delta R\%$) can be formulated as (92):

$$\Delta T\% = (T_{hot} - T_{cold}) \cdot 100$$

$$\Delta R\% = (R_{hot} - R_{cold}) \cdot 100$$

where hot and cold denotes transmittance/reflectance at the hot and cold state respectively.

- The most common TC material in TCWs is pure **vanadium dioxide**, with a transition temperature of 68°C which should be decreased to ambient temperature for practical use.
 The most critical weakness of VO₂ coatings is their low transmittance in the visible range. Many studies have reported values between 40% and 50%, which is well below the acceptable value of 60% (93, 94). **Table 4** shows the reported values of transmittance and reflectance in the visible and IR range for VO₂.
- Low energy-saving efficiency also limits the application of VO₂ coatings. The change in transmittance before and after the transmission temperature T_t , at 2500nm, is known as the **switching efficiency η_T** and is the benchmark of energy-saving efficiency. (?why? because of lighting?). This value is influenced by doping(107,108), microstructure(80,95,109-111), **and film thickness (80,88)**. The most paramount factor among them is film thickness that affects the switching efficiency most significantly. However, increasing the film thickness has an adverse effect on T_{vis} . As observed from table 4, **the ideal film thickness is between 40 and 80 nm**.
- Crucial Steps to overcome the limited application of TWCs:
 - Suitable doping(reducing T_t and improving T_{vis})
 - Appropriate Coating Technology
 - Adding efficient anti-reflecting coating(to increase T_{vis}) (read next title in article)
 - Reducing coating costs

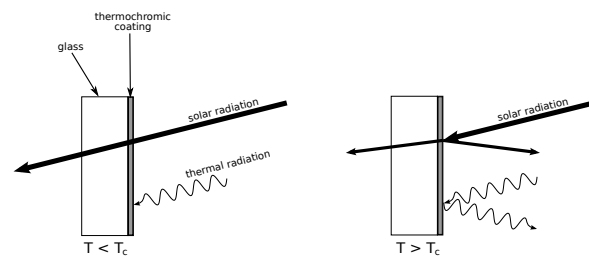


Figure 4.1: Schematic demonstration of the application of thermochromic materials to advanced window glazing [8]. In the article it is used as a pictorial representation of how vanadium(IV) oxide thin film will work as an intelligent window.

REFERENCES

- [1] D. Bedeaux, J. Vlieger, *Optical Properties of Surfaces*, Imperial College Press, London, 2001
- [2] Parkin IP, Manning TD. *Intelligent thermochromic windows*, Journal of Chemical Education, London 2006;83(3):393. ?I DON'T KNOW WHAT 393 is!? what is it?
- [3] White MA, LeBlanc M. *Thermochromism in Commercial Products*, Journal of Chemical Education, Canada, September 1999;76(9) ?IS THIS RIGHT? IS SOMETHING WRONG? DO I MISS SOMETHING?!
- [4] Hiltunen J, Silfsten P, Jaaskelainen T, Parkkinen JPS, *A Qualitative Description of Thermochromism in Color Measurements*, ?????????????, ?????????????

5 BOOK NOTES

6 HANDBOOK OF OPTICAL CONSTANTS OF SOLIDS; EDWARD D. PALIK

Institute for physical Science and Technology, University of Maryland
Academic Press 1998, 1985.

6.1 BULK AND THIN-FILM EFFECTS; EFFECTIVE-MEDIUM THEORY; p.104

- p. 105:
In this discussion, we assume that the characteristic dimensions of the microstructure are large enough ($> \lambda \sim 10 - 20\text{\AA}$) so that the individual regions retain essentially their bulk dielectric responses, but small ($> \lambda \sim 0.1-0.2 \lambda$) compared to the wavelength of light. Then, the macroscopic \mathbf{E} and \mathbf{H} fields of Maxwell's equations will not vary appreciably over any single region, and quasistatic theory can be used. This avoids complications due to scattering and retardation effects that are dominant in macroscopically inhomogeneous systems [1].

The dielectric functions is obtained from the macroscopic average electric field \mathbf{E} and polarization \mathbf{P} according to

$$\mathbf{D} = \epsilon \mathbf{E} = \mathbf{E} + 4\pi \mathbf{P} \quad (6.1)$$

$$\mathbf{P} = \frac{1}{V} \sum q_i \Delta \mathbf{x}_i, \quad (6.2)$$

where $\Delta \mathbf{x}_i$ is the displacement of the charge q_i under the action of the local field at q_i . It is the appearance of the volume normalizing factor in the latter equation that is responsible for the sensitivity of ϵ to density.

•

REFERENCES

- [1] Beckmann P. The depolarization of electromagnetic waves. Golem, Boulder, Colorado, 1968.

6.2 JENSEN B: THE QUANTUM EXTENSION OF THE DRUDE-ZENER THEORY IN POLAR SEMICONDUCTORS; p.169-188

p.169-170:

Introduction

The classical Drude model for the complex dielectric constant of a semi-conductor can be used to extract the mobility and the free-carrier density n_e from an analysis of the reflectivity and transmittance data in the far infrared (1-4) [1, 2, 3, 4, 5]. The dielectric constant ϵ is the square of the complex refractive index, which determines the optical properties of a given material. One has

$$\epsilon = \epsilon_1 - i\epsilon_2 = N^2, \quad (6.3)$$

where the real and imaginary parts of the complex parts of the complex dielectric constant ϵ_1 and ϵ_2 are functions of the complex refractive index N as

$$N = n - ik \quad (6.4)$$

$$\epsilon = n^2 - k^2 \quad (6.5)$$

$$\epsilon = 2nk = \frac{4\pi\sigma}{\omega} \quad (6.6)$$

The choice of $n - ik$ rather than $n + ik$ is determined by the original use of $\exp i(\omega t - \mathbf{q} \cdot \mathbf{r})$ in the assumed plane-wave solution of Maxwell's equations. In Eqs (6.3)-(6.6), n is the real part of the complex refractive index, k the imaginary part or extinction coefficient, and σ the optical conductivity. The absorption coefficient α is proportional to σ , to ε_2 , and to k :

$$n\alpha = \frac{4\pi\sigma}{c} = \frac{\omega}{c}\varepsilon_2 \quad (6.7)$$

$$\frac{\alpha}{2} = \frac{\omega}{c}k = \frac{1}{\delta} \quad (6.8)$$

The extinction coefficient k is essentially the ratio of the free-space wavelength of light of frequency ω to the skin depth δ .

The Drude theory gives the free-carrier contribution to ε_1 and ε_2 in terms of the plasma frequency $\bar{\omega}_p$ and the electron scattering time τ as

$$\varepsilon_1 = \varepsilon_\infty \frac{1 - \bar{\omega}_p^2}{\omega^2\eta} \quad (6.9)$$

$$(6.10)$$

$$\varepsilon_2 = \frac{\omega_p^2}{\omega^2\eta} \frac{1}{\omega\tau} \quad (6.11)$$

where ε_∞ is the high-frequency lattice dielectric constant. The real and imaginary parts of the complex refractive index are obtained from ε_1 and ε_2 . One has

$$\varepsilon = \sqrt{\varepsilon_1^2 + \varepsilon_2^2} = n^2 + k^2, \quad (6.12)$$

$$n = \sqrt{\frac{\varepsilon + \varepsilon_1}{2}}, \quad (6.13)$$

$$k = \frac{\varepsilon_2}{2n} = \sqrt{\frac{\varepsilon - \varepsilon_1}{2}}. \quad (6.14)$$

Experimentally, n and k are found from measurements of the reflectivity R of a bulk, opaque sample and the transmittance T of a slab, which are given in terms of n and k as

$$R = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2} \quad (6.15)$$

$$T = \frac{(1-R)^2 e^{-2\omega kd/c}}{1 - R^2 e^{-4\omega kd/c}}, \quad (6.16)$$

where d is the sample thickness. For the slab multiple-reflection effects are averaged, so that interface fringes are not resolved.

p.171:

In the far infrared, for photon energies small compared with $k_0 T$ (k_0 Boltzmann's const.) and with the energy $\hbar\omega_Q$ of the phonon involved in the scattering, the quantum result reduces to the λ^2 dependence given by the Drude Theory, and the quasi classical Boltzmann transport equation (1-3). The departures from the Drude theory at high frequencies are associated mainly with k rather than n , and hence, the transmission is affected more than the reflectivity. The latter depends on k in the region of the reflectivity minimum, where $n \approx 1$, but is determined essentially by n over the region of the absorption spectrum for which $n > k$, which is the region where departures from the Drude theory would occur.

(...)

The response of electrons to a driving field may be followed from the quasi-classical limit of small ω to the quantum limit that occurs when $\hbar\omega$ is no longer small compared with characteristic energies of the system. In this case, a generalized Boltzmann equation is obtained that reduces to the quasi-classical Boltzmann transport equation when the electron wave vector q tends to zero and ω is small (14-17) = [6, 7, 8, 9]. When ω is appreciable,

one obtains, under certain conditions, a solution of the Boltzmann equation in terms of a frequency-dependent relaxation time. This relaxation rate, which has been tabulated as a function of frequency and carrier concentration for various materials (18-20) = [10, 11, 12], can be used in the usual expression of the classical Drude theory to obtain the quantum result. In particular, the low-frequency $\hbar\omega \simeq k_0 T$ limit gives a good estimate for the dc mobility as a function of carrier concentration. At high frequencies, in lightly doped materials in which polar scattering dominates, $n\alpha$ is proportional to λ^3 and ε_2 and k are proportional to λ^4 rather than λ^3 . The real part of the dielectric constant is given approximately by the Drude-theory expression and $n \simeq \sqrt{\varepsilon_\infty}$ for $\bar{\omega}_p \ll \omega \ll G/\hbar$, where G/\hbar is the frequency of the fundamental absorption edge and G is the direct-band-gap energy of the semiconductor. As ω approaches G/\hbar there is a small quantum-mechanical correction to ε_1 and hence to n . A summary of the results of the quantum theory is given in Section II.

p.176:

Comparison with experimental data

A calculation of ε_1 appropriate to electrons in polar semiconductors with the band structure of the Kane theory and based on the quantum density-matrix equation of motion yields a high frequency modification to Eq. (6.10). For $\omega\tau \gg 1$ and $X < 0.1$, ($X = \hbar\omega/G$) one obtains (18,22)

$$\varepsilon_1 = \left[\frac{\varepsilon_\infty}{1-X} \right] [1 - (X/\varepsilon_\infty) - \bar{\omega}_p^2/\omega^2] \quad (6.17)$$

$$= \left[\frac{\varepsilon_\infty}{1-X} \right] [1 - \bar{\omega}_p^2/\omega^2] \quad (6.18)$$

$$= \varepsilon_\infty(1 - \bar{\omega}_p^2/\omega^2), \quad X \ll 1 \quad (6.19)$$

We note that $1/\varepsilon_\infty < \sim 0.1$ and hence $X/\varepsilon \ll 1$, for compounds we consider, and this term can be neglected. In the limit $X \ll 1$, the quasi-classical high-frequency Drude result is recovered, as required. For $X \sim 0.1$, there is a high-frequency correction given by Eq. (some equation), which is used to calculate the numerical values of $\varepsilon_1 = n^2 - k^2$. The major modification of the classical result is dispersion in n as one approaches the fundamental absorption edge [13].

REFERENCES

- [1] Palik ED, Hold RT. Nondestructive evaluation of semiconductor materials and devices. Plenum 1979 (Chap. 7)
- [2] Holm RT, Gibson, Palik ED, J. Appl.Phys. 48,212 (1977)
- [3] Perkowitz S. J.Phys.Chem.Solids 32, 2267 (1971)
- [4] Perkowitz S, Thorland RH, Phys.Rev. B9, 545 (1974)
- [5] Fan HY. Semiconductors and semimetals. (R.K. Willardson and A.C. Beer eds.), Vol.3, Academic Press, New York 1967
- [6] Jensen B. Ann. Phys. 95,229 (1975).
- [7] Price PJ. IBM J. Res. Dev. 10,395(1966)
- [8] Argyres PN, J. Phys.Chem.Solids 19, 66 (1961)
- [9] Kohn W, Luttinger JM. Phys. Rev. 108, 590 (1957) Kohn W, Luttinger JM. Phys.Rev.109,1892(1958)
- [10] Jensen B. Phys. StatusSolidi. 86, 291 (1978); Jensen B. SolidState commun. 24, 853 (1977)
- [11] Jensen B. J. Appl Phys. 50, 5800 (1979)
- [12] Jensen B. Laser Induced damage in optical materials; 1980 (H.E. Bennet, A.J. Glass, A.H. Guenther, and B.D.Newman, eds.), p.416, National bureau of standards special publication 620, Boulder, Colorado, 1981.
- [13] Jensen B, IEEE J. Quantum Electron. QE-18, 1361 (1982); Jensen B, Torabi A, IEEE J.Quantum Electron. QE-19, 448-457,877-882,1362-1365 (1983); Jensen N, Torabi A, J.Appl.Phys. 54,2030-2035,3623-3625,5945-5949 (1983).

6.3 SHASHANKA S. MITRA

OPTICAL PROPERTIES OF NONMETALLIC SOLIDS FOR PHOTON ENERGIES BELOW THE FUNDAMENTAL BAND GAP

Department of electrical engineering, University of Rhode Island.

p.263-267:

Infrared Dispersion by plasmons

In the preceeding discussion of absorption of infrared radiation by phonons in a solid, it was tacitly assumed that the solid was an insulator. This assumption does not hold well for a narrow-gap semiconductor at ordinary temperatures or for a doped semiconductor with a partially filled conduction or valence band. **The collective excitation of this free-carrier electron gas(plasma) in such a crystal will modify the infrared absorption by phonons,** as discussed earlier. The dispersion mechanism through which electromagnetic radiation interacts with a solid should include the contribution of free-charge carriers in solids for which their numbers are significant, in addition to contributions from bound electrons and phonons. The dielectric response function of such a solid can now be written as

$$\epsilon = 1 + 4\pi(\chi_{BE} + \chi_L + \chi_{FC}), \quad (6.20)$$

where χ_{BE} , χ_L and χ_{FC} , respectively represent the bound electron, lattice, and free-carrier contributions to the electrical susceptibility. For the spectral region of interest, we are not concerned with the bound-electron dispersion; thus this term, as usual, will be represented by a dispersion-free, high-frequency dielectric-constant term

$$\epsilon_{\infty} = 1 + 4\pi\chi_{BE}. \quad (6.21)$$

An approximate expression for the dielectric response function for a free-electron gas in a solid can be obtained from the classical Drude model in which a free electron of effective mass m^* and charge e is displaced by an amount \mathbf{x} as a result of interaction with the electric field \mathbf{E} , with the equation of motion

$$m^* \ddot{\mathbf{x}} + m^* \gamma \dot{\mathbf{x}} = e \mathbf{E}_0 e^{-i\omega t} \quad (6.22)$$

The damping term proportional to the velocity obviously represents the electron-phonon scattering in a phenomenological manner. Solving for \mathbf{x} , one obtains

$$\mathbf{x} = -e \mathbf{E} / m^* \omega(\omega + i\gamma) \quad (6.23)$$

The polarization, defined as the electric-dipole moment per unit volume, is given by

$$\mathbf{P} = N e \mathbf{x}, \quad (6.24)$$

where N is the carrier concentration. Recalling that

$$\mathbf{P} = [(\epsilon - 1)/4\pi] \mathbf{E}, \quad (6.25)$$

one readily obtains

$$\epsilon_{FC} = \epsilon_{\infty} - \frac{4\pi N e^2}{m^* \omega(\omega + i\gamma)} \quad (6.26)$$

for the dielectric response function due to a single-component plasma. In terms of the plasma frequency defined as

$$\omega_p = \sqrt{\frac{4\pi N e^2}{\epsilon_{\infty} m^*}} \quad (6.27)$$

ϵ_{FC} becomes

$$\epsilon_{FC} = \epsilon_{\infty} [1 - (\omega_p^2 / \omega(\omega + i\gamma))]. \quad (6.28)$$

For a number of semiconductors there exists a region of the infrared spectrum in which the free-carrier contributions dominate, and both the bound electron and lattice contributions to dielectric response function are negligible. For a semiconductor, such a situation prevails in a region of the infrared spectrum in which $\omega_g \gg \omega_p \gg \omega_{TO}$,

where $\hbar\omega_g$ is the electronic band gap. The analysis of optical response is also simpler in a region in which the reflection coefficient is independent of carrier scattering. For such a situation to prevail, one usually finds $\gamma \ll \omega$ and $\epsilon_1 \gg \epsilon_2$. In a few semiconductors for which all these conditions are met, the reflectivity spectrum in the appropriate ($\sim \omega_p$) infrared region becomes a function of carrier concentration. Here PbTe satisfies the preceding specified conditions, and its infrared reflection spectrum (117) as a function of hole carrier concentration is shown in Fig. (some figure). Since this case satisfies the preceding outlined conditions, the reflectivity minimum is a measure of the effective mass

$$m^* \simeq \frac{1}{(\epsilon_\infty - 1)} \frac{Ne^2}{\pi c^2} \lambda_{min}^2, \quad (6.29)$$

and the real and imaginary part of ϵ are given by

$$\epsilon_1 = \epsilon_\infty - \frac{e^2}{\pi c^2} \left(\frac{N}{m^*} \right) \lambda^2 \quad \text{and} \quad \epsilon_2 = \frac{e^2}{2\pi c^3} \left(\frac{N}{m^*} \right) \gamma \lambda^3. \quad (6.30)$$

The λ^2 dependence of ϵ_1 is confirmed experimentally (117).

For most polar semiconductors, however, lattice dispersion contributes significantly. For such cases,

$$\epsilon(\omega) = \epsilon_\infty \left[1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)} \right] + \frac{(\epsilon_0 - \epsilon_\infty)\omega_0^2}{\omega_0^2 - \omega^2} \quad (6.31)$$

where the last term represents a single undamped oscillator representing the long-wavelength optical phonons. It may be recalled from Subsection B that $\omega_{TO} = \omega_0$ was identified with the maximum in ϵ , whereas ω_{LO} with its zero. Because of the first term, the second condition is now altered: ie, ω_{TO} can still be identified with the maximum in ϵ but not ω_{LO} with its minimum. Thus for a polar semiconductor with an appreciable carrier concentration, the resonance corresponding to the long-wavelength LO frequency will no longer correspond to the real LO phonon frequency. In the limit of negligible damping for both electrons and phonons, the solution of Eq.(6.31) is given by

$$\omega_\pm^2 = \frac{1}{2}(\omega_{LO}^2 + \omega_p^2) \pm \frac{1}{2}\sqrt{(\omega_{LO}^2 + \omega_p^2)^2 - 4\omega_p^2\omega_{TO}^2} \quad (6.32)$$

where ω_+ and ω_- represent new resonances above and below ω_{LO} .

The plasmon-phonon interaction obviously will be modified by the inclusion of both plasmon and phonon damping terms. (...)

(?Jeg tror LO=longitudinal oscillator og TO=transversal oscillator?)

REFERENCES

[1]

7 OPTICAL PROPERTIES OF SURFACES NOTES

- ρ is the number of island per unit of surface area.
- α_{\parallel}^{10} : (average) parallel quadrupole polarizability
- α_{\perp}^{10} : (average) normal quadrupole polarizability
-

$$\tau = -\rho\alpha_{\parallel}^{10} \quad \text{and} \quad \delta = -\rho[\alpha_{\perp}^{10} + \alpha_{\parallel}^{10}]/\epsilon_a \quad (7.1)$$

ϵ_a is the ambient/surrounding dielectric constant.

- an extremely important difference between an interface and the bulk of the adjacent phases is the inherent asymmetry of the interface in its response to fields normal to the surface or parallel to the interface. This is eminently clear for a thin metal film between two dielectric media. If one applies an electric field along the film this results in a current while an electric field normal to the film does not lead to a current. (p.7)
- excess quantities which not affect the boundary conditions have no relevance for the reflection and transmission of light by the interface.
- The resistance of the sublayers are put in series and add up for an electric field perpendicular to the surface. This is very different from the case described above for an electric field parallel to the layer where the conductivities of the sublayers were put in parallel.

If the electric field is parallel to the layer, one finds an important contribution from the layer if its conductivity is large compared to the surrounding medium.

For a field orthogonal to the layer, however, one finds an important contribution if the conductivity of the layer is much smaller than the conductivity of the surrounding medium. It is this very characteristic difference in the response of the layer to fields orthogonal and parallel to the layer which is the origin of many of the interesting electromagnetic properties of the layer. (p.10)

- If the electric field is parallel to the layer the dielectric constants of the sub-layers are so to say put in parallel. If the electric field is normal to the layer, however, they are put in series. Furthermore one finds that, if the dielectric constant of the layer is much larger than the dielectric constants of the surrounding homogeneous media, the constitutive coefficient γ_e is much smaller than the coefficient β_e and dominates the behaviour of the layer. If the dielectric constant of the layer is much smaller than the dielectric constants of the surrounding homogeneous media, the constitutive coefficient β_e is much larger than the coefficient γ_e and dominates the behaviour of the layer. It is again this very characteristic difference in the response of the layer in the directions orthogonal and parallel to the layer which is the origin of many of the interesting electromagnetic properties of the layer. (p.13)
- the jump in the extrapolated fields are given in terms of the total excess of $D_{ex,\parallel}, E_{ex,z}, B_{ex,\parallel}, H_{ex,z}, I_{ex,\parallel}$ and ρ_{ex} . Notice that it is however not affected by the total excesses of $D_{ex,z}, E_{ex,\parallel}, B_{ex,z}, H_{ex,\parallel}, I_{ex,z}$ and ρ_{ex} ; these total excesses therefore have no effect on the reflection and transmission amplitudes and may as such be neglected in the description of the optical properties of the interface. (I think this is for $k_z = 0$). (p.16)
- p.21 - 72. NOT READ
- If one consider a single island surrounded by the ambient the electro-magnetic response may be characterized by dipole, quadrupole and higher order multipole polarizabilities. If this particle is brought close to the substrate the multipole polarizabilities are modified, due to an induced charge distribution on the surface of the substrate, and become dependent on the distance to the substrate. (p.73)
- neighboring island interaction also changes the polarizabilities of the island and result in polarizabilities parallel and orthogonal to the surface which are unequal, even if the islands are spherical. As a consequence the film is anisotropic in its reaction to fields parallel versus fields normal to the surface. Further contribution to this anisotropy is achieved if the particle is not symmetric. (p.73)
- The resulting dipole polarizabilities of the islands, parallel and orthogonal to the surface, are directly related to the constitutive coefficients γ_e and β_e , respectively. The susceptibilities δ_e and τ account for the fact that the dipoles are situated at a finite distance from the surface of the substrate. It follows in the context of the polarizable dipole model that β_e and γ_e are proportional to the typical diameter of the islands times the coverage. The coverage is the surface area of the particles as viewed from above divided by the total surface

area. For identical spheres the coverage is equal to $\pi r^2 \cdot N_{islands}/A$. Furthermore it follows that δ_e and τ are proportional to the diameter times the distance of the center of the islands to the substrate times the coverage.

In general, and not only in the polarizable dipole model, one may in fact use the ratio of the coefficients δ_e and τ with the coefficients γ_e and β_e as a rough measure for the distance of the center of the islands to the substrate. (p.74)

- the first term (A_{lm} term) is due to sources located in the region $r < a$. The second term (B_{lm} term) is due to sources in the region $r > b$ and may be identified as the incident field due to external sources. The first term is due to the charge distribution induced in the island. The island has no net charge \rightarrow the $l = 0$ contribution in the first term is zero. A_{lm} gives a contribution to the potential which decays as r^{-l-1} , and may be identified (apart from a numerical constant) as an amplitude of the l 'th order multipole field. The amplitude of these multipole fields are given in terms of the amplitudes of the incident field using a polarizability matrix:

$$A_{lm} = - \sum_{l', m'}' \alpha_{lm, l' m'} B_{l' m'} \quad \text{for } l \neq 0 \quad (7.2)$$

where the prime over sum indicates that $l' \neq 0$.

- polarizable dipole model: all amplitudes with l and $l' \neq 1$ is set to zero.
- polarizable quadrupole model: all amplitudes with l or $l' > 2$ is set to zero.

(p.79)

- s. 80 viser en del intuitiv forklaring av hvordan polariseringen påvirker feltet i øvre og nedre medium. Kan hjelpe å forstå hvordan de første A_{lm} amplitudene representerer dipolbidraget.
- For particles with symmetry axis normal to the surface of the substrate (A characterizes the strength of the reflected dipole):

$$\alpha_{\parallel}(0) = [1 + A\alpha_{\parallel}]^{-1} \alpha_{\parallel} \quad (7.3)$$

$$\alpha_{\perp}(0) = [1 + 2A\alpha_{\perp}]^{-1} \alpha_{\perp} \quad (7.4)$$

If there are two different polarizabilities parallel to the surface the above equation is valid for both of them. Furthermore, regarding the anisotropy in the interaction with the substrate, even if the particle is spherical, so that $\alpha_{\parallel} = \alpha_{\perp}$, the different response to fields along the surface and normal to the surface results in $\alpha_{\parallel}(0) \neq \alpha_{\perp}(0)$. (p.81).

If one covers the substrate with a low density of identical islands, which have a rotational symmetry around the normal of the surface, one finds for the first order interfacial susceptibilities

$$\gamma_e(d) = \rho \alpha_{\parallel}(0) = \rho [1 + A\alpha_{\parallel}]^{-1} \alpha_{\parallel} \quad (7.5)$$

$$\beta(d)_e = \rho \epsilon_a^{-2} \alpha_{\perp}(0) = \rho \epsilon_a^{-2} [1 + 2A\alpha_{\perp}]^{-1} \alpha_{\perp}, \quad (7.6)$$

where ρ is the number of particles per unit of surface area. The argument d indicates that the location of the dividing surface is chosen to be the $z = d$ surface, which in this case coincides with the surface of the substrate. One important thing: in writing these formulae the polarization due to the dipoles is in fact taken to be located at the surface of the substrate. The origin of τ and δ_e is related to a proper choice of this location.

- The reason, that it is also in the dipole model important to take finite values for δ_e and τ , is related to the occurrence of phase factors. Light reflected from the film can either be reflected directly from the island film or be transmitted and subsequently reflected by the surface of the substrate. It is clear that the phase difference between these two contributions can be important. The analysis on p. 83 and the resulting finite value of δ_e and τ account for the phase differences between the actual location of the islands and the surface of the substrate to linear order. (p.83)
- δ_e and τ have relatively small modification of the optical properties, but contain nevertheless interesting new information. The coefficients γ_e and β_e , which usually have a much larger influence, are a measure of the amount of material deposited on the surface of the substrate. They are therefore a measure of what one usually calls the weight thickness.

- (p.95)
Incident e field E_0 and the corresponding potential $\psi_0(\mathbf{r}) = -E_0 \cdot \mathbf{r}$ is given on .
- Appendix A (p. 106): The derivation of the surface constitutive coefficients $\gamma_e(d)$ and $\beta_e(d)$. for an identical particle, island film, in the low density limit

Instead of dipoles, small dielectric spheres are considered, located a distance $d \gg R$ above the substrate (substrate is located at $z = d$). Excess fields are introduced and the total field in the ambient is the sum of the external field, the dipole fields of the spheres (with centers in $(\mathbf{R}_{i,\parallel}, 0)$) and the image dipoles (with centers in $(\mathbf{R}_{i,\parallel}, 2d)$).

- (p.117)
When the islands are brought close to the substrate, one must also account for the modification of their dipole and quadrupole polarizabilities due to the interaction with higher order multipoles in the substrate. It is sufficient to calculate only the modification of these polarizabilities, as long as one is interested in the constitutive coefficients γ_e , β_e , δ_e and τ .

- (p.118)
The dipole and the quadrupole approximation is in general inadequate for islands which are flat, i.e. islands for which the distance of the center to the surface of the substrate is smaller than 30% of the linear diameter along the surface.

- p.173

- p.204

(I think this is for the dipole + quadrupole expansion.. maybe)

The calculations were done for scaled densities $\hat{\rho} \equiv \rho(2R)^2 = (2R/L)^2 = (2\hat{R})^2$ of 0 (no interaction), 0.2, 0.4, and 0.8. L is the lattice constant and $\hat{R} \equiv R/L$. The reason to go to a higher order in the multipole expansion in the interaction with the substrate is that, close to touching the substrate, the higher order multipoles start to give corrections of about 5 to 10%. As is analyzed in great detail by Haarmans and Bedeaux (49), **higher order multipole corrections for the interaction along the substrate become important for $\hat{\rho} > 0.9$. They use as density parameter the coverage, which is equal to $(\pi/4)\hat{\rho}$ for the square array.** (...) The interaction along the substrate has a substantial effect on the size of the polarizabilities and the resulting constitutive coefficients. The polarizability along the substrate approximately doubles from $\hat{\rho} = 0$ to 0.8 (for triangular lattice?), while the polarizability normal to the surface decreases by a factor of 2. The resonance frequency shifts a little bit down for the polarizability parallel and up for the polarizability normal to the surface. The contributions due to the quadrupole polarizabilities to the coefficients $\hat{\delta}_e$ and $\hat{\tau}$ are small compared to the contributions due to the dipole polarizabilities.

p.209

These coefficients are therefore mainly due to the shift of the induced dipole to the surface of the substrate.

- p.209-210.

Here (also something about square lattice), they talk about that the distance between the islands must be so far apart to give reasonable results. It also seems like they're saying that the polarizabilities and constitutive coefficients of OBLATE SPHEROIDS do not depend very much on the coverage. The reason is that the center of the OBLATE SPHEROIDS remain relatively far apart, even for high densities.

For PROLATE SPHEROIDS the dependence on the coverage becomes dramatic, reason being that they approach each other for higher coverage to distances smaller than the long axis. It is to be expected that these results are unreliable for the combination of a high density and a small axial ratio

If one would set a criterium that the distance between the particles should be larger than half the long axis, one finds that for densities 0.0, 0.2, 0.4 and 0.8 the maximum values of $1/(1+AR)$ are 1.00, 0.69, 0.61 and 0.53 respectively. It follows that all the new maxima appearing for larger values of $\hat{\rho}$ and $1/(1+AR)$ are outside the range of reliability of the dipole approximation for the interaction along the substrate. Since the shape of the particles in island films is usually oblate rather than prolate, the dipole approximation to describe the interaction between the particles will in practice suffice.

8 COLOR NOTES FROM WIKIPEDIA

- Humans tend to perceive light within the green parts of the spectrum as brighter than red or blue light of equal power. The luminosity function that describes the perceived brightnesses of different wavelengths is thus roughly analogous to the spectral sensitivity of M cones.
- **With three color appearance parameters- colorfulness(or chroma or saturation), lightness (or brightness), and hue, any color can be described.**
- **Hue** is one of the main properties of a color. There are six (unique) hues: red, orange, yellow, green, blue and purple. Hue is described as the degree to which a stimulus can be described as similar to or different from stimuli that are described by the unique hues?
- **Colorfulness** is the visual sensation according to which the perceived color of an area appears to be more or less chromatic. A highly colorful stimulus is vivid and intense, while a less colorful stimulus appears more muted, closer to gray. With no colorfulness at all, a color is a "neutral" gray (an image with no colorfulness in any of its color is called grayscale).
- **Chroma** is the colorfulness relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting. (therefore) Chroma should not be confused with colorfulness!
- Saturation is the colorfulness of a color relative to its own brightness.
- **Lightness** (aka value or tone) is a representation of variation in the perception of a color or color space's brightness. Lightness is a relative term. Lightness means Brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting. Lightness should not be confused with brightness.
- **Brightness** is an attribute of visual perception in which a source appears to be radiating or reflecting light. Brightness is the perception elicited by the luminance (how much luminous power will be detected by an eye looking at e.g. a surface from a particular angle of view) of a visual target. This is a subjective property of an object being observed and one of the color appearance parameters of color appearance models. Brightness refers to an absolute term and should not be confused with Lightness.

In RGB color space, brightness can be thought of as the arithmetic mean μ of the red, green and blue color coordinates (although some of the three components make the light seem brighter than others).

$$\mu = \frac{R + G + B}{3} \quad (8.1)$$

Brightness is also a color coordinate in the HSB or HSV color space (Hue, Saturation, and Brightness/Value).

9 COLOR NOTES REGARDING COLORPY

source: <http://markkness.net/colorpy/ColorPy.html>

How I installed it:

- download zipped folder and unzip it using:
`gunzip -c colorpy-0.1.0.tar.gz | tar xf`

The unzipped folder can be put anywhere (doesn't matter). Go to it (`cd path/colorpy-0.1.0`). Personally, I ran the following command (don't know if you need administrator privileges, but it seemed like it was required):

```
sudo python setup.py install
```

After than I made a folder "TEST", went into it, started python and ran the following commands to test that the program/module worked correctly:

```
# running test cases:
```

```
import colorpy.test
colorpy.test.test()
```

```
# generate sample figures:
```

```
import colorpy.figures
colorpy.figures.figures()
```

Fundamentals:

- COLORPY generally uses wavelengths measured in nanometers, otherwise typical metric units are used.
- The description of the spectra, COLORPY uses two-dimensional NumPy arrays: column 1: wavelength[nm], column 2: light intensity at the wavelength.
- Color values are represented as three-component NumPy vectors (1D-arrays). Typically, these are vectors of floats, except for irgb colors, which are arrays of integers in the range 0-255.
- COLORPY can provide a blank spectrum array, via `colorpy.ciexyz.empty_spectrum()`, which have rows for each wavelength from 360nm to 830nm, at 1nm increments.
- To go from the light spectra to a color (a 3D quantity), we use a set of three matching functions

$$X = \int I(\lambda) \cdot CIE-X(\lambda) d\lambda \quad (9.1)$$

$$Y = \int I(\lambda) \cdot CIE-Y(\lambda) d\lambda \quad (9.2)$$

$$Z = \int I(\lambda) \cdot CIE-Z(\lambda) d\lambda \quad (9.3)$$

- The Y-matching function corresponds exactly to the luminous efficiency of the eye - the eye's response to light of constant luminance (these facts are some of the reasons that make this particular set of matching functions so useful).
- Rescaling XYZ so that their sum is 1.0, we get the values xyz (i.e. $x + y + z = 1$). The chromaticity is given by x and y (z can be reconstructed as $z = 1.0 - x - y$). It is also common to specify colors with the chromaticity (x,y), as well as the total brightness (Y). Occasionally, one also wants to scale XYZ colors so that the resulting Y value is 1.0.
- Here are some 'constructor'-like functions for creating the colors (e.g. XYZ), which are three-component vectors.

```
colorpy.colormodels.xyz_color(x,y,z=None)
colorpy.colormodels.xyz_normalize(xyz)
colorpy.colormodels.xyz_color_from_xyY(x,y,Y)
colorpy.colormodels.xyz_normalize_Y1(xyz)
```

Note that color types are generally specified in ColorPy with lower case letters, as this is more readable. The user must keep track of the particular normalization that applies in each situation.

Fundamentals - converting XYZ colors to RGB colors which we can draw on the computer:

- To convert we can use

```
# Here, xyz is the XYZ color vector:
colorpy.colormodels.irgb_from_xyz(xyz) # Returns [int,int,int] (int in [0,255])
colorpy.colormodels.irgb_string_from_xyz(xyz) # Returns hex string, e.g. '#FF0000' for red.
```

- Note that there are several subtleties and approximations which are important to understand what is happening in the code above.

- the XYZ color is converted to a linear RGB color (represented as floats in the range 0.0-1.0), meaning that the light intensity is proportional to the numerical color values, and is done by a 3x3 element array. The values of the array depends on physical display in question, so the conversion matrix cannot apply to all displays/monitors. Fortunately, we can assume a specification of monitor chromaticities (which are a part of the sRGB standard) which are likely to be close to most displays. COLORPY uses this assumption by default, although you can change the assumed monitor chromaticities to nearly anything you like.

- The RGB colors can be out of range:

1. greater than 1.0, meaning the color is too bright for the display.
2. negative value, meaning that the color is too saturated and vivid for the display.

To put these values between 0.0-1.0 that can actually be displayed is known as color clipping: The first is solved by rescaling such that the maximum value is 1.0, making the color less bright without changing the chromaticity. The second change required some change in chromaticity. By default, COLORPY adds just enough white to the color to make all the components non-negative. (There is also the potential to develop better clipping functions.)

- The next subtlety is that the intensity of colors on the display is not simply proportional to the color values given to the hardware. This situation is known as the 'gamma correction'. We rely on the sRGB standard what to do. The standard assumes a physical 'gamma' exponent of about 2.2., and COLORPY applies this correction by default (can be changed if you like).
- The final step is to convert the RGB components from 0.0-1.0 to 0-255 (done with simple scaling and rounding).

- Summarizing these conversions:

- The function

```
colorpy.colormodels.rgb_from_xyz(xyz)
```

converts an XYZ color to linear RGB color in range 0.0-1.0. Can be out of range or negative, and can not be directly passed to drawing functions.

- The function

```
colorpy.colormodels.irgb_from_rgb(rgb)
```

converts a linear RGB in the range 0.0-1.0 to a displayable irgb color, definitely in the range 0-255. Color clipping may be applied (intensity and chromaticity), and gamma correction is accounted for. This result can be passed to drawing functions.

- To plot the visible spectrum:

```
colorpy.plots.visible_spectrum_plot()
```

To plot the RGB values for the pure spectral lines (or visible spectrum?), use:

```
colorpy.plots.color_vs_param_plot(param_list, rgb_colors, title,
    filename, tight=False, plotfunc=pylab.plot, xlabel='param', ylabel='RGB Color')
```

In this case, the parameter list is wavelength and we also include a list of RGB colors. These must be of the same size. This is a very handy function, useful for many other plots besides this one.

- CIE chromaticity diagram of the visible gamut:

```
colorpy.plots.shark_fin_plot()
```

- 'Patch' plot:

```
colorpy.plots.xyz_patch_plot(xyz_colors, color_names,  
                             title, filename, patch_gap=0.05, num_across=6)
```

`xyz_colors` and `color_names` are two lists (which must be of same size). You can pass 'None' for the second list to skip the labels. (There are also optional arguments to fine-tune the size and arrangement of the patches.)

There is also a similar function:

```
colorpy.plots.rgb_patch_plot()
```

when you have known RGB values that you want to draw.

- To plot spectrums like black body spectrums, use

```
colopy.plots.spectrum_plot(spectrum, title, filename, xlabel, ylabel)
```

Here, `spectrum` is a numpy array.

COLORPY allows us to plot the resulting color vs. temperature, while also showing a plot of the rgb color values:

```
colopy.plots.color_vs_param_plot(param_list, rgb_colors,  
                                 title, filename, tight, plotfunc, xlabel, ylabel)
```

The arguments `tight`, `plotfunc`, `xlabel`, `ylabel` are optional. (In the example, he used `tight=True`, `plotfunc=pylab.semilogy` to obtain the semilog plot, which is needed for the very large range of color values covered).