Detailed Derivation of the Inversion Condition

Advanced

As in the "short-cut" derivation of the inversion condition, we look at the number of stimulated emission and fundamental absorption processes per time and volume unit (i.e., the rate density). We use the same symbols, however, with some additional indices if necessary. We have:

- $R_{fa}(E_1, E_2) =$ rate for **fundamental absorption** of a photon by an electron sitting in the valence band in the energy interval defined by E_1 and $E_1 + \Delta E_1$ that is then moved "up" into the conduction band to the energy interval $E_2, E_2 + \Delta E_2$ and under the condition that the wave vector of the electron (in a reduced band diagram) does not change.

- $R_{se}(E_1, E_2) =$ rate for **stimulated emission** of a photon by a photon from an electron sitting in the conduction band in the energy interval defined by E_2 and $E_2 + \Delta E_2$ that is then moved "down" into the valence band to the energy interval $E_1, E_1 + \Delta E_1$ and under the condition that the wave vector of the electron (in a reduced band diagram) does not change.

We now define the rates *relative to an energy interval* because we will see that the photon does not need to have *exactly* the energy $h \cdot \nu = E_2 - E_1$ to induce transitions – this follows from **Heisenberg's** uncertainty relations.

- As before, the rates must be proportional to the number of electrons available for the interaction with a photon *and* the number of empty states available for the electrons to occupy after the interaction

At the respective energies, the density in some small energy interval ΔE (ideally being zero) around some energy value E_2 or E_1 in the conduction or valence band, is always given by the density of states at the chosen energy times probability of occupations times small energy interval, i.e.

$D_C(E_2) \cdot \Delta E \cdot [f(E_2, E_F^e, T)] =$	(differential) density of electrons close to the conduction band edge in the energy interval $E_2, E_2 + \Delta E$
$D_V(E_1) \cdot \Delta E \cdot [1 - f(E_1, E_F^h, T)] =$	(differential) density of holes close to the valence band edge in the energy interval $E_1, E_1 + \Delta E$

- The density of states $D_{C, V}$ are functions of $E_{2, 1}$, of course.

Since we use the quasi Fermi energies, we have to use *two* different Fermi distributions for the two quasi Fermi energies; we have $f(E, E_F^e, T)$ and $f(E, E_F^h, T)$

- However, for a full appreciation of the symmetries in the problem, it may be advantageous to define alltogether *four* Fermi distributions, two of which we then will really use as already pointed out when quasi Fermi energies were introduced.

- Besides $f_e(E, E_F^e, T)$, giving the *direct* probability of finding electrons in the conduction band, we simply define the *direct* probability for finding holes in the valence band as $f_h(E, E_F^h, T) = 1 - f_e(E, E_F^h, T)$. Watch out for the indizes h and e in this!

- Next we simply agree to use *only* f_h if we discuss what is going on in the valence band, and *only* f_e for the conduction band.

Using this convention, the equations from above, augmented by two more equations to account for all densities, we have

$D_C(E_2) \cdot \Delta E \cdot [f_e(E_2, E_F^e, T)] =$	(differential) density of *electrons* close to the *conduction* band edge in the energy interval $E_2, E_2 + \Delta E$
$D_V(E_1) \cdot \Delta E \cdot f_h(E_1, E_F^h, T) =$	(differential) density of *holes* close to the *valence* band edge in the energy interval $E_1, E_1 + \Delta E$
$D_V(E_1) \cdot \Delta E \cdot [1 - f_h(E_1, E_F^h, T)]$	(differential) density of *electrons* close to the *valence* band edge in the energy interval $E_1, E_1 + \Delta E$
$D_C(E_2) \cdot \Delta E \cdot [1 - f_e(E_2, E_F^e, T)]$	(differential) density of *empty states* close to the *conduction* band edge in the energy interval $E_2, E_2 + \Delta E$

- This should be clear, simply recall that $f_h(E_1, E_F{}^h, T)$ is the *direct* probability of finding *holes* in $E_1 \Delta E_1$; we therefore must take $(1 - f_h(E_1, E_F{}^h, T)$ for the probability of finding *electrons* if we *only* use $f_h(E_1, E_F{}^h, T)$ in the valence band.

▶ This allows us to formulate the electron and hole density part of the rates R we are trying to calculate.

- These rates are also proportional to the number of photons with the energy $h\nu \approx E_2 - E_1$, i.e., to $u(\nu)$. We use "$\approx$" instead of "$=$", because one of the goals of this module is to see how far off the photon energy could be from the band gap energy while still giving some effects.

- Using the proportionality constants A_{fa} and A_{se}, we now can formulate the same equations as we had before, but in a more precise and general way (omitting the temperature to save space):

$$R_{fa}(E_1, E_2) = \left(A_{fa} \right) \cdot \left(D_V(E_1) \cdot \Delta E_1 \cdot [1 - f_h(E_1, E_F{}^h)] \right) \cdot \left(D_C(E_2) \cdot \Delta E_2 \cdot [1 - f_e(E_2, E_F{}^e)] \right) \cdot \left(u(\nu) \right)$$

$$R_{se}(E_1, E_2) = \left(A_{se} \right) \cdot \left(D_V(E_1) \cdot \Delta E_1 \cdot f_h(E_1, E_F{}^h) \right) \cdot \left(D_C(E_2) \cdot \Delta E_2 \cdot f_e(E_2, E_F{}^e) \right) \cdot \left(u(\nu) \right)$$

- The terms in the large brackets are (from left to right)

  - The proportionality coefficients or Einstein coefficients
  - The density of unoccupied states (= density of holes) available for the electrons at the end point of the transition.
  - The density of electrons present for interactions with the photons at the energy interval considered, i.e. the occupied states.
  - The density of photons available for the processes considered

- This kind of relation for a transition probability of a particle from one state to another state is an example of an universal principle known as "**Fermi's golden rule**". While it is very easy to grasp the way it is presented here, it is actually a bit strange: What the electrons do depends on the density of unoccupied energy levels *somewhere and sometime else*. How do the electrons "*know*" if there is some free space "down there"? More to that in another advanced module.

▶ **Einstein** showed that $A_{se} = A_{fa} = A = $ **Einstein coefficient** for fundamental absorption under all circumstances. We will derive this relation in another modul – it is not so easy (Einstein was a great physicist indeed).

▶ We will now calculate $R^{net}{}_{se}$, the *net* rate of stimulated emission, which is

$$R^{net}{}_{se} = R_{se} - R_{fa}$$

- Using the formula from above and gong back to the regular Fermi distribution, we obtain

$$R^{net}{}_{se} = A \cdot u(\nu) \cdot D_C(E_2) \cdot D_V(E_1) \cdot \left(f(E_2, E_F{}^e) + f(E_1, E_F{}^h) - 1 \right) \cdot \Delta E_1 \cdot \Delta E_2$$

▶ We now must look more closely at the $\Delta E_{1,2}$ terms and consider their physical significance.

- The key point is that we are looking at electron (or hole) states that are very short lived. Consider the electron that was knocked into the state with the energy E_2 in the conduction band (or the hole state left at E_1). They will disappear very quickly (within a time given by the dielectric relaxation time of the system, which is around 10^{-12} s for typical optoelectronic semiconductors).

- In other words, the electron in the conduction band will very quickly thermalize by interactions with phonons and settle at energies close to E_C.

▶ Heisenberg's uncertainty principle for energies states that

$$\Delta E_2 \cdot \Delta \tau \geq h$$

- This means that for electrons occupying some state E_2 for some small average (= uncertain individual) time $\Delta \tau$, we have an uncertainty in the precise value of the energy. It is not given precisely, but only within some uncertainty $\Delta E = h/\Delta \tau$.

- More precisely, things can happen even if the energy available is not precisely what is needed - all the it is required that it happens so fast (small $\Delta\tau$) that the difference in energy is covered by the resulting uncertainty. In other words, there is a certain likelihood that transitions may occur for not exactly matching energies, too

- In yet other words: A photon with the precise energy of $h \cdot \nu$ may create electron–hole pairs with energies of $h \cdot \nu \pm \Delta E$ with some probability that depends on ΔE and $\Delta\tau$.

- This means that there is a certain probability distribution $L(E_2, E_1, \tau_C, \tau_V)$ for obtaining a transition even if the photon does not have the exactly right energy $h\nu = E_2 - E_1$, but deviates by some ΔE.

- We cannot calculate that distribution at this point, but we can be pretty sure that it must be something as shown in the picture on the right.

- In total, we have the possibility that photons somewhat off the proper energy still may induce a transition. The probability, however, will go down rapidly as the deviation ΔE is increased or τ increases.

- The reverse then is also true: Photons with only *one* well defined energy $h\nu$ will cause transitions not only between energy states $E_2 - E_1 = h\nu$, but also between states with an energy $(E_2 - E_1) \pm \Delta E$ and the probability for some ΔE is given by the **L** curve.
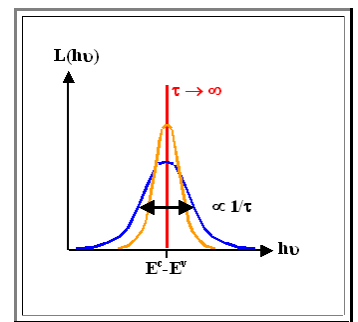
The *total* transition rates of electrons in response to photons with some density $u(\nu)$ at the frequency ν is thus obtained by integrating over all possible transitions weighted by the probability function $L(h\nu)$. If we include this into the formula for the net emission rate we obtain

$$R^{net}_{se} = \int\int A \cdot u(\nu) \cdot D_C(E_2) \cdot D_V(E_1) \cdot \left(f(E_2, E_F^{e}) + f(E_1, E_F^{h}) - 1 \right) \cdot \left(L(E_2, E_1, \tau_C, \tau_V) \right) \cdot dE_1 \cdot dE_2$$

- and we have to integrate over the valence and conduction band.

Now we reached the end of our tether. Because we neither know $L(E_2, E_1, \tau_C, \tau_V)$, nor could we easily do the integration if we would know it (it must be, after all, some bell-shaped kind of probability distribution).

- We thus introduce the first approximation into our so far rather rigid derivations: We assume a δ-function for **L**; in other words, we neglect the energy uncertainties and allow only transitions with identical energies to occur. This does *not* mean that the energy levels are now fixed, only that the energy level you reach from some level is fixed.

One integration now becomes trivial. We integrate over the conduction band, rewrite E_2 as $E_1 + h\nu$ (we can do that now because we now have a sharp level in the conduction band), and obtain

$$R^{net}_{se} = A \cdot u(\nu) \cdot \int D_C(E_1 + h\nu) \cdot D_V(E_1) \cdot \left(f(E_1 + h\nu, E_F^{e}) + f(E_1, E_F^{h}) - 1 \right) \cdot dE_1$$

Far simpler, but his equation still contains an integration over the density of states, which is not so easy to do even in the simplest approximations.

- In other words: We still allow transitions from *different* energy states to the corresponding states $h\nu$ larger in energy (and always at the same wave vector k). So we seem to be stuck again.

Well, not completely so, because the last equation also contains the *inversion condition* in a slightly veiled form. Let's see how this can be verified:

- For inversion we defined that $R_{se} \geq R_{fa}$, and that implies that $R^{net} = R_{se} - R_{fa}$ must be zero or *positive*.

- Since A, $u(\nu)$, and $D_{C,V}$ are *always* positive, $R^{net} \geq 0$ only applies if

$$
\begin{aligned}
f(E_1 + h\nu, E_F^{e}) + f(E_1, E_F^{h}) - 1 &\geq 0 \\
f(E_1 + h\nu, E_F^{e}) &\geq 1 - f(E_1, E_F^{h})
\end{aligned}
$$

This is a condition for *any* combination of two energy levels from the valence and conduction band with energy difference $h\nu$.

- Let's see what that means by looking at a *special* combination of energies: $E_1 = E_V$, $E_2 = E_C$, $h\nu = E_g$. Then, $f(E_1 + h\nu, E_F^e) = f(E_2, E_F^e)$ is the probability that we find electrons at the conduction band edge, and $1 - f(E_1, E_F^h) = 1 - f(E_V, E_F^h)$ is the probability that the valence band edge states are *not occupied by holes* and therefore occupied by electrons.

- The inversion condition as expressed in the equation above then states that stimulated emission *only* surpasses fundamental absorption if it is more likely to find an electron at the conduction band edge than at the valence band edge.

- This statement applies to any pair of energies. Inversion occurs as soon as it is more likely to find an electron at some energy in the conduction band, than at this energy minus $h\nu$ in the valence band.

Spelling out the Fermi distribution functions,

$$f(E_1 + h\nu, E_F^e) = \frac{1}{1 + \exp\dfrac{(E_1 + h\nu) - E_F^e}{kT}}$$

- and so on, we find once more as the inversion condition

$$E_F^e - E_F^h \geq h\nu$$

What did we gain by this procedure? Quite a lot, because we now know where we cut corners. We know how we can calculate rates for stimulated emission and fundamental absorption with sufficient precision – all we have to do is some numerical work.

- But actually, what we really want to do is to make a laser diode and not to calculate emission rates. Making an operating Laser diode entails that we have some *specific volume* where we need to have inversion conditions. Then we send some photons into this volume to start stimulated emission, and, since this would deplete the electron population in the conduction band, we also inject electrons into the conduction band (and, of course, holes in the valence band), with the aim of achieving some steady state in the output of light produced by stimulated emission!

- In other words, we must look at the *dynamics* of inversion, at how all quantities involved *change with time*. And this can be done by working with the formulas derived here – but in a different module.