VIBRATIONAL RELAXATION STUDIED WITH LIGHT

Ad Lagendijk[1,2]

[1] FOM-Institute for Atomic and Molecular Physics
Kruislaan 407, 1098 SJ Amsterdam, The Netherlands.
[2] Van der Waals-Zeeman Laboratorium, Valckenierstraat 65-67
1018 XE Amsterdam, The Netherlands.

1 ABSTRACT

In this lecture I will deal with optical studies of vibrational dynamics. I will address the question of what (minimal number of) parameters characterize vibrational motion and under what conditions the relaxation-time approximation is valid. Within this assumption the vibrational dynamics are represented by, besides the eigenfrequency, the decay constants T_2 and T_1. More of these constants will be needed when different oscillators are coupled.

Following we will consider the extraction of these quantities from optical experiments. Relevant optical techniques are linear and nonlinear absorption spectroscopy, spontaneous Raman scattering and forced or "coherent" Raman scattering. The complications of strong coupling of the oscillators to the light field, in absorption spectroscopy referred to as polariton formation, will be dealt with.

It will be demonstrated that the conventional response theory that is employed to calculate line shapes has serious shortcomings. I refer to these "Kubo-type" response principles as to "one-way response theory" in contrast to, what I have coined, "two-ways response theory".

2 INTRODUCTION

When one considers nowadays the many scientific papers in natural science one might get confused and a little desperate. The fact that a major fraction of these writings deals with optics and lasers, that is to say our field, adds to this discomfort. One should not be surprised that there are numerous publications on scientific applications of lasers, and in particular on *non*linear optics, because many very clever scientific workers have bought many lasers and all the necessary auxiliary equipment. Optics is in fashion. Almost any physicist, and chemist for that matter, that respects himself is in some way associated with research in which lasers and nonlinear optics are involved. Whatever he will find or not find, he will write a publication on it. No more microwave spectroscopy, no more magnetic resonance spectroscopy, no more acoustics, no, just lasers. The problem with lasers is that they work on optical principles. Suddenly everybody realizes that before one can embark in this trendy field on nonlinear optics one has to learn something about this dull field of linear optics. Linear optics has long been abandoned as a basic subject, not only in (Dutch) highschools, but also in (Dutch) curricula for university degrees in physics. Many scientists working in the field of nonlinear optics feel the need to reinvent the wheel, not realizing that a great deal of the formalism for linear and nonlinear optical response was developed much earlier for magnetic resonance in the 1960's. Famous names associated with major contributions during that period are for instance P.W. Anderson, N. Bloembergen, R. Kubo, and A.G. Redfield. When somebody wants to learn the subject my advice is: just buy and read a good book on magnetic resonance like C.P. Slichter's "Principles of Magnetic Resonance".[1] The problem with the field of nonlinear optics is that the subject is of enormous extension. Is it a technique, or is it a phenomenon, or a class of phenomena? Should one discuss the possible applications of the technique, or should one look at purely optical phenomenon like nonlinear light propagation? Books on nonlinear optics necessarily only describe a tiny part of the field. Nonlinear optics is also often approached from the viewpoint of

an engineer. For an engineer relaxation times like the well-known, and for us crucial, T_2 and T_1, are just some parameters determining the quality factor of a lumped circuit, in the same way as for instance the reflectivity of a mirror does, but for a solid state physicist these relaxation times differ like day and night from the imperfections of a mirror.

Let us first find out what type of physics you are interested in and let us see if nonlinear optics has a role to play in satisfying your curiosity about nature. The emphasis throughout this lecture will be: how can we *use* optical principles, and if necessary nonlinear optical experiments, to learn more about our material.

Optical studies of vibrational dynamics are numerous. This is not a review paper, so I cannot give proper credit to the vast and excellent work carried out by many colleagues. Sometimes, for educational purposes, I will need and discuss some specific examples. In the majority of cases these demonstrations will be drawn from our own work. This implies nothing whatsoever about (my opinion on) the quality of other work, but is purely done on the basis of convenience (for me).

3 TYPE OF PHYSICS

Before we switch on our laser, let us first discuss what type of information we want to obtain from our experiment. This is a very important contemplation as otherwise we get lost in the tremendous complications of laser systems and nonlinear optical phenomena. A laser is a very bright source of light and as such able to introduce all kinds of nonlinear couplings and effects, many of which are a nuisance and are unwanted. Nonlinearity is found in many branches of physics besides optics, and its presence invariably means complications. If you can avoid nonlinearities when gathering your information, please do.

I assume you are interested in some "system", consisting of many particles or degrees of freedom. This corpus is a large collection of atoms or molecules if you are a solid-state physicist and a smaller number of atoms if you are a molecular physicist. What quantities are of interest? When do you "understand" the behavior of your crowd of particles? For any system one can define many, often even an infinite number of, observables and one cannot be interested in all of them, leave alone to know everything about them. Understanding implies among other things characterization and prediction of the behavior of your object using a minimal number of observables. Consequently we will have to put some limitation on our ambition for knowledge of everything.

This lecture is about vibrational dynamics, so apparently we want to understand more about the displacement of atoms around their equilibrium positions. Quantities of concern could be for instance simple expectation values like $\ll u^2 \gg$, the mean-squared displacement of an atom at a particular site, but also the more informative functions describing spatial correlations between displacements. Examples of these (static) correlation functions are: $\ll \mathbf{u}(\mathbf{r}_1)\mathbf{u}(\mathbf{r}_2) \gg$, $\ll \mathbf{u}(\mathbf{r}_1)\mathbf{u}(\mathbf{r}_2)\mathbf{u}(\mathbf{r}_3) \gg$, $\ll \mathbf{u}(\mathbf{r}_1)\mathbf{u}(\mathbf{r}_2)\mathbf{u}(\mathbf{r}_3)\mathbf{u}(\mathbf{r}_4) \gg$ and higher-order functions. The expectation value $\ll \cdots \gg$ denotes an averaging over some relevant distribution function, like the quantum-mechanical expectation value over the ground state, or a thermal average (quantum or classical). Sometimes we would like to Fourier transform our spatial correlation functions, to take maximal advantage

of translational symmetry, and wavevectors will become the relevant variables like for instance: $\ll \mathbf{u}(\mathbf{p}_1)\mathbf{u}(\mathbf{p}_2) \gg$. In case of translational symmetry this function is proportional to $\delta(\mathbf{p}_1+\mathbf{p}_2)$. We have used the label $\mathbf{p}$ rather than $\mathbf{k}$ for the momentum variable as later $\mathbf{k}$ will be used for light propagation.

All expectation values are equilibrium expectation values, and they are only determined by the, time-independent, Hamiltonian of your system. We call the correlation functions discussed so far *static*, because they do not depend on time. One can, and for the study of vibrational *dynamics* one should, also be interested in *dynamic* correlation functions, that depend on time (or frequency). An important dynamic correlation function is

$$\ll \mathbf{u}(\mathbf{r}_1, t_1)\mathbf{u}(\mathbf{r}_2, t_2) \gg, \tag{1}$$

which describes correlations in both spatial and temporal variables. This dynamic correlation function can also be generalized to situations where more displacement operators are involved in the correlation function. Note that these correlation functions are again defined in terms of equilibrium expectation values, and time is only a parameter. The distribution function over which the averaging is performed, does not depend on time. Fourier transforming over time is of course very frequently performed as nature looks commonly much simpler in the frequency domain than in the time domain. Fourier transforming the lowest-order dynamic displacement correlation function (1) with respect to space and time defines $\sigma(\mathbf{p}, \Omega)$,

$$\sigma(\mathbf{p}, \Omega)\delta(\mathbf{p}_1 + \mathbf{p}_2) = \int \mathrm{e}^{i\Omega t} \ll \mathbf{u}(\mathbf{p}_1, t_1)\mathbf{u}(\mathbf{p}_2, t_1 + t) \gg dt\,, \tag{2}$$

where the delta function comes from the assumption of spatial translational symmetry. Integrals without explicit indication of the domain of integration are to be taken from $[-\infty, +\infty]$. I have chosen to use Ω rather than ω for frequency as ω is often used to describe the deviation from the central frequency (bandwidth) of incoming light.

Up to now I have mainly considered correlation functions containing displacement operators. This is a professional deformation of workers in optics. The coupling between light and our type of matter is primarily through the displacement operator. People studying vibrational dynamics with neutron scattering always consider correlation function of *density* operators. The obvious reason is here that the interaction between neutrons and atoms is through a very local potential (often represented as a delta-function pseudo-potential). Using the first Born approximation to describe the scattering of neutrons off nuclei is an excellent approach and immediately leads to the study of correlation functions of the type

$$\ll \rho(\mathbf{r}_1, t_1)\rho(\mathbf{r}_2, t_2) \gg, \tag{3}$$

in which ρ is the density operator: $\rho(\mathbf{r}) \equiv \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i)$, where N is the number of particles. A rather famous function is the double (space and time) Fourier transform of correlation function (3), which after performing the spatial Fourier transform explicitly looks like

$$S(\mathbf{p}, \Omega)\delta(\mathbf{p}_1 + \mathbf{p}_2) = \int \mathrm{e}^{i\Omega t} \ll \rho(\mathbf{p}_1, t_1)\rho(\mathbf{p}_2, t_2) \gg dt\,, \tag{4}$$

This function $S(\mathbf{p}, \Omega)$ is the most important function describing neutron scattering (where it is better known as $S(\mathbf{k}, \omega)$, and sometimes referred to as the "scattering law"). Under mild conditions, to be discussed later, the correlation function $S(\mathbf{p}, \Omega)$

has a Lorentzian shape[2] (as a function of Ω) and we call its center, $\Omega_0(\mathbf{p})$, the frequency, and the width of the line, $\tau(\mathbf{p})$, we refer to as the damping of the oscillator. The displacement correlation function $\sigma(\mathbf{p}, \Omega)$ will have approximately the same shape as $S(\mathbf{p}, \Omega)$.[3]

In many cases higher-order correlation functions can be decoupled into lower-order correlation functions. This alleviates the task of characterizing your system as only some lower-order correlation functions need to be known under these circumstances. One cannot in all seriousness claim that a correlation function consisting of a product of thirty-six displacement operators will contain much different information from one holding thirty-eight or thirty-four displacement operators. Many of these operators will fluctuate (almost) independently and can, besides some exceptional pathological cases, safely be decoupled into a product of lower-order correlation functions. The correlation functions discussed so far are intrinsic to your system, that is to say these correlation functions are all defined in the absence of the probes. I have indicated that information can be obtained about these intrinsic correlation functions by coupling your system to a probe and perform an experiment. This is only true for investigations that can be described as linear-response experiments. We will discuss this question in more detail in Section 8. In the philosophy of linear response the influence of the probe is as weak as possible. When using bright light sources linear-response concept is often not valid any more. The good thing is that a stronger coupling between probe and system will potentially lead to a situation where more knowledge can be accumulated. The bad news is that the correlation functions being measured are much more complicated and not any longer intrinsic to your system as the separation of probe and system is not longer possible.

I want to stress that if you are a condensed matter physicist who is interested in vibrational dynamics you do not want to know everything about chirps, bandwidth-limited pulses, regenerative gain, amplified stimulated emission, unstable cavities, mode locking etc. These effects and phenomena are indeed very important, but for you only to the extent that they enable you to study your vibrational system. If you like to study all these nonlinear optical effects in great detail and just for the sake of themselves, fine, but do not call yourself a condensed-matter physicist. A butcher should sell meat, and not bread. Many specialists in vibrational dynamics do not know anything about lasers but it is of crucial importance to keep on communicating with them. By using too much "optical jargon" it is quite simple to alienate oneself from this community. A condensed-matter physicist should have an understanding of the general concepts and methods of calculation of correlation functions for many-body systems. He should know about Green's functions for phonons and electrons. Only if you speak that language, and present your new results using that terminology your discoveries will get the full attention and appreciation of the condensed-matter community at large.

4 HOW COMPLICATED IS YOUR SYSTEM?

If your system is not complicated it is not interesting because many other people will already have studied it in great detail, even before you were born. If a system has only a "discrete" spectrum, containing a finite or at most a countable infinity of spectral points, the dynamics are often solved trivially and uninteresting. In the systems you

will be interested in this is never the case. You will always have to deal with a continuous spectrum and in many circumstances you also will have to handle a many-body system (which is the physicist understatement for an infinite-body system). When I explain to colleagues that they work with an infinite-body system they often look very surprised. In their view they only study a few oscillators having some irreversible behavior (for instance expressed by a dephasing time). But they do not realize that in order to obtain irreversible behavior from a microscopic point of view one has to couple the oscillator to degrees of freedom having a continuous spectrum, frequently a many-body system. Otherwise the system could be described with a finite number of modes and no irreversible behavior could be established.

In all situations a large number, and in many cases an infinite number of operators and degrees of freedom is involved. We want to understand our system, and we will come a long way in comprehension when we determine the behavior of some simple correlation functions. One always speaks of the *relevant* or *slow* variables in contrast to the fast or irrelevant variables. The influence of the fast variables can usually be taken into account as a (generalized) damping force on the slow variables. One looks for the slow variables and treats them in great detail and the fast variables are taken into account only in an approximate way as one expects the details of their dynamics not to be very important. Conserved variables always belong to the collection of slow variables but in many cases there is quite some ambiguity about which variables are allowed to be classified as slow and which not. Sometimes it is a matter of taste and often one really does not know whether in a theory all slow variables have been taken into account properly. When an important slow variable has been overlooked we are frequently warned because when this has happened the damping forces show unphysical behavior (like a very strange resonance structure) signalling an unjustified neglect of a slow variable. There are several systematic approaches to incorporate more and more, less and less slow, variables in a consistent (and hopefully convergent) procedure and one such a method is the Mori method.[4,5]

4.1 Introducing T_2 and T_1

We will outline the general conditions under which the temporal behavior of a system can be described by a single (or just a few) relaxation time(s) and the frequency response can be described by a single (or just a few) Lorentzian line(s). This is very important as the description of a complicated system in these simple terms is a tremendous simplification and a big step towards understanding the system. The Hamiltonian of a system, H_{sys} can very often be partitioned in two commuting parts and a third coupling or interaction term. We will call them subsystem, bath and interaction term:

$$H_{sys} = H_{subsystem} + H_{interaction} + H_{bath} \equiv H_{sub} + H_{int} + H_{bath}. \tag{5}$$

We suppose that H_{sub} is a very simple system, like consisting of only a few bound states or just an undamped harmonic oscillator, and we assume H_{bath} to fluctuate on very short time scales. If this is not true we will switch their roles and call the bath the system and vice versa. If this does not work your system is either trivial or very complicated. Skipping the trivial case the failure of a partitioning like expression (5) can only be due if both H_{sub} and H_{bath} are nontrivial many-body systems. But in that case you can probably forget all about their coupling because they have enough problems to take care of their own dynamics and will have no time to influence each other. Under these circumstances you can disregard the "bath" and can continue with H_{sub} which

you should try to partition in its own subsystem and bath as indicated in expression (5). I hope by now that I have convinced you that partitioning (5) is indeed very general and poses hardly any limitation on the systems you can describe. I expect the reader also to recognize that partitioning (5) complies with the principle of separation of the variables into slow and fast variables. The subsystem takes care of the slow variables and the rest of the fast variables. In some approaches $H_{int} + H_{bath}$ are taken together into one term in which the bath variables do not appear as operators, but as classical fields having an explicit time dependence. This more limited scheme is often the basis of stochastic methods. The line of attack that we propose to deal with H_{sys} can easily be adjusted so as to apply to these more limited cases also. In the following I will give one explicit example of a stochastic model and indicate the parallel with the approach based on Hamiltonian dynamics.

We assume that our system can be depicted as a few levels in interaction with a continuum. In principle one could define dephasing for a single level, but the situation is much clearer for a two-level system. Let us suppose we have two levels, the excited state $\mid e >$ and the ground state $\mid g >$ both interacting with a continuum. The total wavefunction of the compound system can then be described as $\mid i > \otimes \mid \{bath\} >$, where a symbolic notation for the bath wavefunctions has been introduced and $\mid i >$ stands for either the excited or the ground state. Due to the presence of a high density of continuum states the state $\mid i > \otimes \mid \{bath\} >$ will be (almost) degenerate with many other states which differ only in their bath states. After the system is excited in one of these compound states it will go through many of the other compound states almost degenerate with it, because the interaction term will couple them and will take care of the energy balance. After some characteristic time the phase of this initial state will not be related anymore to the phase at a later time. The typical time associated with this dephasing mechanism is called the pure dephasing time T_{deph} and is best described to be associated with quasi-elastic processes. Elastic because the subsystem component of the wavefunction is unaltered and remains $\mid i >$, and quasi because the coupling energy takes care of small, but not vanishing, energy mismatches.

A second dynamic process can be visualized which could be classified as inelastic. Let us again discuss the case of two levels for the time being: $\mid e >$ and the subsystem ground state $\mid g >$. Both these levels will be coupled to the continuum and would be dephased by it. However the state $\mid e > \otimes \mid \{bath\} >$ can now also be degenerate with a state of the type $\mid g > \otimes \mid \{bath\} >$ if the right continuum states are excited. The state $\mid e > \otimes \mid \{bath\} >$ will decay into these states but after this kinetic process the excited compound state has ended, in contrast to a pure dephasing process, in a *different* subsystem state. The bath will take care of extreme fast de-excitation of the excited continuum states and we will wind up with the subsystem in the ground state. This mechanism represents an inelastic process and the characteristic time will be called $T_{1,eg}$. The reader might complain that this inelastic process is also, in a way, described as a dephasing process, albeit one in which the subsystem component has changed. This is a very fundamental problem with present day (quantum) physics. Building one's models using hermitian Hamilton operators will always lead to energy conservation and the inability to formally describe inelastic processes. The only technique to represent inelastic behavior is by relating it to dephasing into a larger space.

The method to obtain the simplest possible theoretical expressions for the relaxation times from Hamiltonian (5) goes under many names, my favorite is *the weak-coupling*

approach. All these methods can be viewed as variations on the theme of "expansion in the interaction Hamiltonian H_{int}". I will estimate higher-order corrections later but will now indicate the lowest-order result. The dephasing time is measured when a spectroscopic measurement which induces transitions and for this reason the dephasing time is a measure for the *difference* in dephasing of the two levels. The answer, correct to lowest order in H_{int}, for T_{deph} can also be found by applying Fermi's Golden Rule, and the result is given by

$$\frac{1}{T_{deph}} = E_{int}^2 \rho(E=0), \tag{6}$$

in which E_{int} is a typical energy scale for H_{int} (or difference in diagonal matrix elements of H_{int} between excited and ground states), and where $\rho(E)$ is the density of states of the continuum or bath at energy E. I cannot be more specific about the precise magnitude of E_{int} unless I would treat in much more detail the exact form of the interaction Hamiltonian. It is important to realize that E_{int} is not at all influenced or determined by H_{bath}. People familiar with motional-narrowing theory in magnetic resonance know that E_{int}^2 is also equal to the second moment of a relevant operator of the subsystem (the relevant operator is the simplest operator of the subsystem that has matrix elements between the ground state and the excited state).[1] This second moment is not determined by the bath Hamiltonian because this Hamiltonian commutes with any operator of the Hilbert space of H_{sub}. The density of bath states has as dimension *per energy* so it denotes a time scale. For this reason many people refer to it as $\tau_c(E=0)$, the correlation time of the bath. In the latter notation the dependence on the frequency E is often dropped. With the inverse of E_{int} one can also associate a time scale and this could be called the time scale of the interaction term.

Knowing the expression for the dephasing times makes it rather simple to anticipate the expression for the inelastic time $T_{1,eg}$:

$$\frac{1}{T_{1,eg}} = E_{int}^2 \rho(E=E_e-E_g) \equiv E_{int}^2 \tau_c(E=E_e-E_g). \tag{7}$$

As expected in this case the density of states has to be evaluated at the energy difference between the two levels as the energy has to be supplied or received somewhere. Of course $T_{1,eg}$ and $T_{1,ge}$ are connected through detailed balance in case of thermal equilibrium:

$$\frac{T_{1,eg}}{T_{1,ge}} = e^{(E_g-E_e)/k_BT}. \tag{8}$$

This relation is essentially a condition on the density of states of the continuum. The magnitude of E_{int} appearing in equation (7) for the inelastic rate constants can differ by factors of order one from the factor appearing in equation (6) for the pure dephasing kinetics.

Up to now we have discussed the relaxation times in terms of levels. In a many-level system each two levels have their own dephasing times and all levels are potentially connected through inelastic channels. The relaxation times then really become matrices where the indices denote the levels.[6] Experiments often measure lumped combinations of these times and for these reason some linear combinations of these times are well-known. When one measures the decay of the population *difference* between two levels i and j the decay time, defined as T_1, would be given by the parallel sum , $1/T_1 \equiv 1/T_{1,ij} + 1/T_{1,ji}$.

For a multilevel system it makes more sense not to combine them in lumped relaxation times. In a spectroscopic measurement, like optical absorption, the linewidth is defined to be $1/T_2$. The pure dephasing process gives rise to an energy uncertainty of $1/T_{deph}$, and the inelastic process gives rise to an energy uncertainty of $1/(2T_1)$. These processes are independent and add, and the linewidth is given by

$$\frac{1}{T_2} = \frac{1}{T_{deph}} + \frac{1}{2T_1}. \tag{9}$$

The factor of two appearing in this equation can be explained in the following way. Dephasing is a property of the probability amplitude (or wavefunction or vibrational amplitude) and population dynamics is connected with the amplitude *squared*.

I have presented enough information to make it possible to indicate how higher-order corrections to the weak-coupling result would look like. Higher-order corrections to the relaxation times can be visualized by writing expression (6) for T_{deph} in a different way,

$$\frac{1}{T_{deph}} = E_{int}\left\{E_{int}\rho(E=0)\right\} = E_{int}\left(E_{int}\tau_c\right). \tag{10}$$

The dimensionless parameter $E_{int}\tau_c$ is the expansion parameter. In a theory incorporating higher-order corrections to the weak-coupling theory higher-order powers of $E_{int}\tau_c$ would be added to the term between brackets in equation (10). The expansion parameter must be small for our approach to be valid and we see that the request is that the bath should be much faster than the time connected with interaction Hamiltonian E_{int}^{-1}. For many types of baths this will be an excellent approximation and we have done a superb job. We are also able to understand the term motional narrowing. If the bath would be extremely slow the dynamics of the subsystem would be totally determined by E_{int}, and the linewidths and inverse decay times would be of the order of E_{int}, so the width is narrowed by the motion of the bath by a factor $E_{int}\tau_c$.

In the foregoing it was assumed that the continuum could supply or absorb the necessary energy for the inelastic process to go. Here a practical difference between optics and magnetic resonance shows up. In the standard magnetic resonance techniques, like NMR and ESR, the spin energies are usually so small that $\tau_c(E=0) \approx \tau_c(E_e - E_g)$. In that case the quasi-elastic dephasing times and the inelastic times are approximately equal. In optics the situation is very often the other way around. The energies are usually much larger than k_BT and for that reason the T_1's are usually orders of magnitude larger than the dephasing times. The situation could even be more dramatic since it is possible that the density of states is zero at the required energy. This is because the degrees of freedom of the bath have a certain energy scale and for much larger energies than this energy scale there are less and less states and beyond the edge of the density states there will be no states at all. Under these conditions the energy gap is too large and the continuum cannot absorb or supply the necessary energy. In the language of quantum mechanics one would say that one quantum of the bath is not enough to bridge the energy gap and more quanta are necessary, until we have enough of them to span the energy gap. It is clear that we have to go to higher-order perturbation theory with the interaction as perturbation in order to obtain a finite transition probability. There are in principle two ways of doing this. We could use the same interaction energy and go to higher-order in weak-coupling theory but this is very cumbersome. A simpler way to go about is to realize that in the Hamiltonian (5) the interaction term is also usually a first term in an expansion in bath coordinates. If we would allow higher-order

term there we could still do simple weak-coupling theory to obtain expression for the linewidths.[7–9] Let me be more specific. Suppose we consider some subsystem interacting with its environment. We could symbolically denote the environment with some lumped variable $\{R\}$. The coordinates of our simple subsystem will be represented by the coordinate q. We can formally define the "displacement" δR of the bath operators by

$$R \equiv R_0 + \delta R, \tag{11}$$

where R_0 is some equilibrium or reference set of coordinates for the bath. Let us imagine some adiabatic approach (for instance by assuming for the time being that the masses of the degrees of freedom of the bath are infinite). We can then calculate the energy of the total system. This energy is a function of $\{R\}$ and q. Assuming now that we can expand in small bath displacements we will have the following terms (in symbolic notation !!)

$$\begin{aligned} U(\delta\{R\},q) \quad - \quad U(\delta\{R\}=0,q) = & \left(\frac{\partial U(\delta\{R\},q)}{\partial \delta\{R\}}\right)_{\delta\{R\}=0} \delta\{R\} + \\ & \frac{1}{2!}\left(\frac{\partial^2 U(\delta\{R\},q)}{\partial \delta^2\{R\}}\right)_{\delta\{R\}=0} \delta^2\{R\} + \cdots \quad . \end{aligned} \tag{12}$$

All the terms on the rhs. of this expansion will give rise to coupling between bath and subsystem as they depend on both the dynamic coordinates q and $\delta\{R\}$. In the higher-order terms successively more bath operators (or bath quanta) are involved. If we use weak-coupling theory for each separate coupling term we would find that the successive densities of states we would need in the expression for the linewidths (See for instance equation (6)) consists of a compound density of states which can usually simply be described as a (convolution) product of the single-excitation density of bath states. For the first term in the rhs. of the equation above we need the single-excitation density of bath states and for the second term we would need the convolution of two single-excitation density of states, approximately twice as wide in the frequency domain as the single-excitation density of states. If we would go to higher order we would be able to bridge any energy gap. If this is after N terms we would describe the vibrational relaxation in terms of a N-excitation (or N-phonon) process. One calls this the energy-gap law as one expects an exponential decrease in magnitude of the successive terms in expansion (12).[10] The sharp drop for each next order implies that if the process goes for a certain N, one does not need to take into account additional higher-order processes as their contributions will be exponentially smaller.

In multilevel systems some new and interesting phenomena can occur. We will use a three-level system as a generic example: $\mid i >, \mid j >,$ and $\mid g >$ presented in order of decreasing subsystem energy. Excitation to a compound $\mid i > \otimes \mid \{bath\} >$ can be followed by direct decay to the ground state but it can also decay through the intermediate excited compound state $\mid j > \otimes \mid \{bath\} >$. Important factors determining the branching ratio's of the two channels are the respective energy gaps and the magnitudes of the matrix elements of the interaction Hamiltonians. This new situation can again simply be handled in the weak-coupling situation. A system having two excited levels can be described with the two coordinates q_1 and q_2. The potential energy of this system will be of the form $U(q_1, q_2, \{R\})$. The actual dependence of this energy on $\{\delta R\}$ will determine what type of bath excitations will drive the transition. An expansion of this energy in small deviations $\{\delta R\}$ will give rise to terms like $q_1 q_2 \delta\{R\}$,

a bath-induced harmonic coupling between normal modes, but also higher-order terms like bath-induced anharmonicity would show up in this expansion and would provide the right matrix elements to allow for a path in which a low-lying excited subsystem level gets populated before the transport to the ground state takes place.

In addition to introducing new inelastic channels a third level can also have interesting effects on linewidths and line positions. These two quantities are closely related as lineshifts and linewidths are given by respectively the sine and cosine transform of the same dynamic correlation function. From motional narrowing theory in magnetic resonance it is known that if an absorption spectrum consist of two narrow separate lines and a kinetic process exist connecting the two transitions that is faster than the inverse of the separation in frequency between the two lines one "motionally averaged" line will appear at some weighted average position.[11] The weight depends on the occupation probability of the initial states of the two separate transitions. When these initial states differ in energy, a Boltzmann factor will determine their occupation probability and so also the position of the averaged line. Changing the temperature will then change the line position. This is a situation often encountered in optics where a high-energy mode is anharmonically coupled to a low-energy mode that has a fast T_1. The fast occupation dynamics of the low-energy mode (usually referred to as the "exchanging mode") will pull on the frequency of the high-energy mode. The existence of an fast exchanging mode will also have effects on the linewidth. The structure of the contribution to the linewidth must again be of the form $E_{int}^2\tau_c$. E_{int} will be of the order of the anharmonic coupling between the two modes and τ_c will be the correlation time of the kinetic process weighted by a Boltzmann factor containing the energy of the exchanging mode.[12] Under these circumstances measurement of the lineposition and width of a certain oscillator as a function of temperature could give information on the energy of a different mode. In Figure (1) I depict the the linewidth of the CN stretch vibration 2095 cm^{-1} in a ferroelectric crystal as a function of temperature, data obtained with spontaneous and forced Raman scattering. From the slope one infers that this stretch vibration is anharmonically coupled to a mode at 346 cm^{-1}.[13]

4.2 Relaxation of an Oscillator

We will discuss a model that is more phenomenological than our previous example that existed of a few low-lying quantum-mechanical levels. The new model we introduce for the description of a system coupled to a continuum is the very popular harmonic oscillator. An oscillator is a generic model used in physics from degrees of freedom having energies in the range of milli-electron-volts (nuclear magnetism) up to Giga-electron-volts (high-energy physics). In our case we specifically mean the displacement of an atom around its equilibrium position, but much of what we will discuss has a much wider applicability. A major advantage of the oscillator model is that the classical interpretation hardly differs from a quantum-mechanical one. We will assume that our degree of freedom has some generalized coordinate q (we will use q rather than u to comply with a convention popular in the optical literature). In an experiment one measures a collection of these oscillators and we define the "coherent" amplitude

$$Q \equiv \sum_{i=1}^{N} q_i. \tag{13}$$

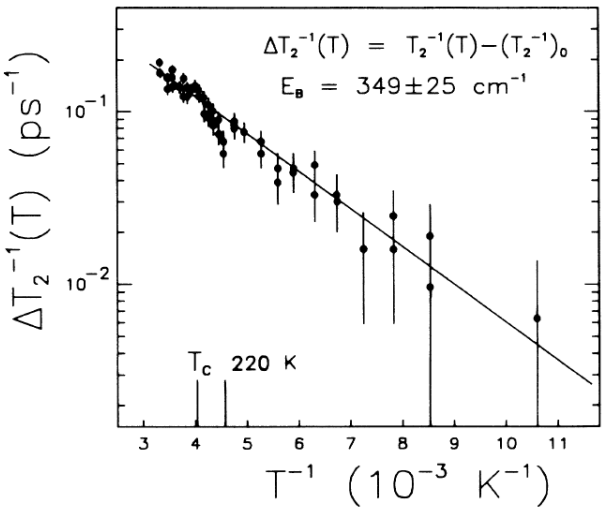

Figure 1. *Logarithmic plot of the temperature dependent contribution to the Raman linewidth of the CN stretching mode in the ferroelectric crystal* $K_4Fe(CN)_6 \cdot 3H_2O$. *The solid line is a fit to exponential behavior which yields an activation energy of* $349 \pm 25 cm^{-1}$. *The interpretation is that the CN mode at* $2095\ cm^{-1}$ *is anharmonically couple to a water mode having this low energy. After Reference (13).*

A careful reader would notice that in the beginning we addressed wavevector-dependent displacement operators of the type $\mathbf{u}(\mathbf{p})$ and now we are only considering displacement operator (13), which is essentially the $p = 0$ component of $\mathbf{u}(\mathbf{p})$. The reason is that due to the discrepancy between atomic length scales and the wavelength of light only the $p = 0$ displacement operator will be probed with light. This is a tremendous disadvantage compared to other techniques like inelastic neutron scattering when investigating vibrational dynamics in the solid state, but less in liquids where (almost) no dispersion (that is $\mathbf{p}$-dependence) exists.

I will treat only a single coordinate Q . It is trivial to extend this into regime of more coupled systems, but I do not want to write out matrices all of the time. The equation of motion of the simple damped harmonic oscillators is

$$\frac{\mathrm{d}^2 Q(t)}{\mathrm{d}t^2} + \frac{2}{T_2}\frac{\mathrm{d}Q(t)}{\mathrm{d}t} + \Omega_0^2 Q(t) = 0. \tag{14}$$

We have neglected a possible distribution ("inhomogeneous linewidth") in the frequencies of the oscillator. The coordinate Q is a generalized coordinate. To emphasize the general character of the present approach I will point out that Q could be the displacement of a lattice vibration (which we have called u earlier), but it could also be the excitation of a bound electron, or it could be the excitation of a magnon, etc. We could even generalize the meaning of Q more. Q could denote some non-equilibrium expectation value introduced by some probe at earlier times and equation (14) would describe the return to equilibrium.[14] Q could also denote some equilibrium displacement-displacement correlation function $\ll Q(0)Q(t) \gg$.

The damping has been introduced in a phenomenological way. In the previous section we have outlined how to justify the concept of damping times in a microscopic way using one of the possible interpretations of Q when we are dealing with a weakly-coupled system. The coordinate Q describes the phase of the oscillators and that is why the damping time contains the total dephasing time T_2. Equation (14) cannot represent population dynamics of the oscillator and so T_1 is not present in it. The extension to include the population dynamics will be discussed later. The solution of equation (14) is simple and requires two initial (or final) boundary conditions i.e. $Q(t=0)$ and $\left(\frac{\mathrm{d}Q(t)}{\mathrm{d}t}\right)_{t=0}$. In the case the interpretation is in terms of correlation functions these initial conditions are determined by the equilibrium ensemble.

I have outlined that equation (14) is a very general equation. Now I want to stress the limitations of this equation. This equation can never be an operator equation. That is to say equation of motion (14) describes the motion of Q, but one cannot infer from the behavior of, for instance, Q^2 from it. The reason is that equation of motion (14) can only be obtained from a microscopic equation after some averaging has been performed. This averaging could have been sneaked in our calculation either because quantum-mechanical matrix-elements have been taken, or because a statistical-mechanical averaging has been carried out, or because an averaging over some stochastic process has been invoked. Equation (14) can never be a microscopic equation, as no microscopic origin of irreversible behavior exists for a finite system. A better practice would have been to use $\langle q \rangle$ and $\langle Q \rangle$ in equation of motion (14) rather than q and Q, to emphasize that some averaging procedure has been introduced.

A simple example will demonstrate the phenomenological character. The dynamical coordinate of our oscillator will be denoted by $\tilde{q}$ (and their sum by $\tilde{Q}$). This coordinate does not represent an averaged coordinate and would for instance appear in the Hamiltonian. The oscillator is coupled to other degrees of freedom. This makes the whole situation into an interacting many-body system. A simple, and ad hoc way, to introduce the rest of the system is to propose a random force, that modulates the frequency,

$$\frac{\mathrm{d}^2\tilde{q}(t)}{\mathrm{d}t^2} + +\Omega_0\left(\Omega_0 + E_{int}\xi(t)\right)\tilde{q}(t) = 0, \tag{15}$$

where ξ is a zero-mean random function with correlation time τ_c. The strength of the random-force has been parametrized with the energy parameter E_{int}. It will turn out that this parameter maps exactly on the E_{int} we have been using extensively in the weak-coupling approach when describing relaxation with a microscopic approach using Hamiltonians. The solution of the stochastic equation is straightforward. Essentially the same expansion in $E_{int}\tau_c$ as in the case of weak-coupling can be inserted. (See van Kampen[15] and his excellent book[16]). The result for the coordinate $\langle\tilde{q}\rangle$, where the brackets mean averaging over the stochastic process, to lowest order in an expansion of $E_{int}\tau_c$ is that one obtains exactly the phenomenological equation of motion (15) when we make the identification: $\langle\tilde{q}\rangle \equiv q$ and $\langle\tilde{q}\rangle \equiv Q$, and the most important result $1/T_2 \equiv (\pi/8)E_{int}^2\tau_c$. I hope the reader sees that this stochastic approach is very analogous to the Hamiltonian approach within the weak-coupling scheme. It is clear that the solution of the Langevin equation for $\langle\tilde{q}^2\rangle$ not simply is given by $\langle\tilde{q}\rangle^2$. A stochastic fluctuation of the frequency is related to pure dephasing. It is also possible to include inelastic channels in this stochastic approach. For instance an anharmonic force proportional to $\tilde{q}^2\xi(t)$ would lead, after the assumption of $\tilde{q}^2 \approx \langle\tilde{q}\rangle\tilde{q}$ to a $T_1 - type$

contribution to the linewidth T_2.

I will now explain why I have emphasized so much the fact that the phenomenological equation of motion can only be obtained aver averaging. Later we will couple this oscillator to an electromagnetic field, and if we would interpret the phenomenological equation as a real dynamical equation, the oscillator cannot absorb light but only scatter light elastically. In the latter case the dephasing time T_2 must correspond to the radiative lifetime of the level as the radiation field would be only reservoir that could cause dephasing of the oscillator. We must allow for the possibility that our oscillators absorb light (that is convert radiation into non-radiative degrees of freedom), and consequently we have to accept that equation of motion (14) has been obtained after averaging over some bath and it describes the motion of the *averaged* amplitude.

Higher-order correlation function are in general notoriously difficult to calculate. Under many circumstances they can be decoupled into a product of lower-order correlation functions, and the only independent correlation functions contains two operators. In the weak-coupling situation we have found that $\ll Q(0)Q(t) \gg$ decays exponentialy with a decay time of T_2. The decoupling of correlation functions is allowed in more general circumstances and certainly in the weak-coupling situation. A decoupling procedure would look like

$$\ll Q(0)Q(t')Q(t'')Q(t''') \gg \approx \ll Q(0)Q(t') \gg \ll Q(t'')Q(t''') \gg + \dots \quad , \tag{16}$$

where the ellipsis stand for the two other additional factorizations. We see that in the case of exponential decay for the $\ll Q(t')Q(t'') \gg$ the (decoupled) higher-order correlations also fall off exponentially.

4.3 Beyond the Relaxation-Time Approximation

Very often $E_{int}\tau_c$ is very small and this is why in the frequency domain simple Lorentz lines and in the time domain simple decay are so often observed. In a way one should also be a little disappointed when confronted with simple exponential decay in a time-resolved experiment as one does not get to know much of the bath as the correlation time τ_c only reflects global behavior of the bath. I call this type of findings about the bath second-hand information (or hear-say witness) If the weak-coupling expansion parameter is not small your line shapes will certainly not be like Lorentzians and you have stumbled on an extremely interesting case. The price you will have to pay when the system allows you to obtain first-hand information is that you have to do a much better job than first-order weak-coupling theory to be able to compare your experimental results with theory.

If the weak-coupling approach does not converge very fast (that is when $E_{int}\tau_c$ is not a small number) one can use the so-called memory function formalism. The central object in this approach is an operator which has many names in different branches of physics. To mention a few: mass operator, or self-energy, or memory function, and it is all the same stuff. We will indicate it in our approach for the oscillator. Equation of motion (14) is of the Markov type. The dynamics of $Q(t)$ only depends on Q, and some of its derivatives, at the same value of t. There is no memory in the system. This

means that a much better and more general equation of motion would look like

$$\frac{\mathrm{d}^2Q(t)}{\mathrm{d}t^2} + \int_0^t \Sigma(t-\tau)\left(\frac{\mathrm{d}Q(t')}{\mathrm{d}t'}\right)_{t'=\tau} d\tau + \Omega_0^2 Q(t) = 0, \tag{17}$$

where the memory function Σ(t) has been introduced. The simple weak-coupling result can be recovered when the memory function is approximated by a delta function. A finite response time of the bath can be taken into account in this way. Very often it can be proved that the *form* of equation (17) is exact. This means of course that all our problems have been sublimated in this memory function. The advantage of this approach is that the memory function is a very convenient vehicle to make approximations for. One of the simplest simplifications that recently got quite popular in optics, but is already known much longer,[17] is to perform a continued-fraction expansion.

An expansion in the coupling does not make sense at all in case the bath is extremely slow. A very important and popular theory developed to deal with this situation for magnetic resonance has never become popular in optics, for reasons beyond my understanding. I am referring to the Kubo and Tomita theory.[18] This theory is essentially a theory based on cumulants. An important aspect of this idea is that it is very easy to obtain either a Gaussian or a Lorentzian lineshape. To obtain the same results in weak-coupling-type theories or memory-function approaches is very involved and only emerges as the solution of a set of difficult self-consistent equations, often referred to as the mode-coupling approach.[4,19] A disadvantage of the Kubo-and-Tomita theory is that an ultra-slow bath always gives a Gaussian lineshape, whereas the correct result should reflect the distribution of the slow bath fields, not necessarily characterized by a Gaussian stochastic process. Another possible method, and alternative for the weak-coupling expansion, would be to use an asymptotic series expansion much like the WKB method by expanding in the kinetic energy of the bath.[20] In a weak-coupling approach an expansion in the interaction Hamiltonian is performed and this is essentially a Born expansion. The asymptotic approach is only expected to be a sensible strategy for a slow bath.

5 PROBE

Once we have decided what kind of information we want to accumulate we will have to figure out how to go about it? Physics is an empirical science, consequently we perform a measurement. That is to say we introduce a probe. Since we know that the conservation of energy is one of the main laws of physics we always write down Hamiltonians or Hamiltonian operators. For our situation the relevant energy operator seems to be

$$H_{world} = H_{system} + H_{probe} \equiv H_{sys} + H_{probe}. \tag{18}$$

This Hamiltonian is not very satisfactory, however, because there is no coupling between system and probe as these terms commute and could treated independently. There must be an additional, *coupling*, term. A coupling term is a part of the Hamiltonian that depends on both the variables of the probe field and the system degrees of freedom. We finally arrive at the Hamiltonian:

$$H_{world} = H_{sys} + H_{probe} + H_{coupling} \equiv H_{sys} + H_{probe} + H_{cou}. \tag{19}$$

You already decided that you were mainly interested in H_{sys}, your hobbyhorse with your favorite particles, and your only motivation to introduce a probe is to get to know more about them. This immediately puts some serious restraints on your probe:

- you must know the probe dynamics in great detail.
- the coupling mechanism must be as simple as possible.

Otherwise you do not know how to interpret your signals. With this in mind it seems that it does not make much sense at all to go to nonlinear optical response. Linear response seems already difficult enough. The whole classic book of Born and Wolf[21] is devoted to *part* of linear optics. To treat nonlinear optics to the same depth one would at least need a multiple of these books. When you start to do nonlinear optics you really open a can of worms. Superficially the form of Hamiltonian (19) is very similar to partitioning (5). Nevertheless there is quite a distinction between the two cases, and for that reason I have called the coupling term here H_{cou} rather than H_{int}. Here H_{sys} stands for the Hamiltonian of the total system which is an interacting many-body system that very likely can be partitioned into form (5). H_{probe} should be as simple as possible, preferably consisting of one, or only a few, well-defined modes. An experimentalist excites this mode and registers at which rate energy is absorbed by H_{sys} from the excited mode of the probe. In this case one looks for the response of a many-body system to the excitation of a simple mode to which it is coupled. In the narrowing Hamiltonian (5) one looks for the reaction of a simple oscillator or simple two-level system on a coupling with an interacting many-body system, the bath. The similarity between the two situations is that, often against the will or even the awareness of the experimentalist, H_{probe} has many more modes than the one being excited by the experimentalist. An example of this situation is the existence of radiative lifetimes that find their origin in the same coupling to the electromagnetic field that takes care of the absorption of light in a spectroscopic measurement, albeit that the coupling is to different modes of that field. I will come back to this difficulty in some detail later, when problems with respect to response theory will be dealt with.

An efficient way of getting information on a particular phenomenon is to get your probe in resonance with a relevant degree of freedom of your system. Usually a probe has a certain characteristic length (for instance the wavelength of the probe) and a characteristic time (or inverse frequency of the probe). Your degree of freedom probably can also be characterized by similar parameters. Optimal coupling is achieved when you match them: probe and system are in resonance. It is very clumsy in general to try to get information on nuclear magnetism from x-ray absorption. So if for some reason you are committed to do optical experiments your primary length scale is the wavelength of light and your primary frequency scale is the optical frequency. Other important factors are pulse width and resolution of your probe, which quantities are constrained through a Fourier-transform uncertainty product.

6 CONTROL OF EXPERIMENT

Any physical system contains at least as many as 10^{23} degrees of freedom and usually an infinite number of them. Can you control them all? Can you change them all? No, of course not! How then can you do a reproducible experiment? Well we use some

assumptions about the realization of the microsystems. In the first place the presence of thermal equilibrium takes care of the fact that nature goes through all the possible microstates and you are just measuring the average. This deals with a lot of the uncontrolled parameters. However, almost always there is still some uncontrollable quenched disorder. In solids we talk about random fields, random strains etc, in liquids we talk about slow density fluctuations. It always means that there are non-thermal degrees of freedom. To deal with this situation we assume that we know the kind of stochastic law that the field is described by. We either average in our theory over all these realizations or we pick a typical realization depending on what is better in agreement with the way the experiment is being done (not of course which fits the experimental results better). In practice we assume a stochastic law that is much simpler than the real situation. Many effects, like interaction between impurities and dislocations, are neglected. Usually so much averaging goes on in a disordered system, that the details of the structure of the disorder do not matter much. The exception to this rule is nowadays the subject of intensive study of electron dynamics in condensed matter physics, but does not need to bother us here.

7 INTERACTION LIGHT AND MATTER

We have decided to do an optical experiment. This introduces some problems with respect to getting the probe in resonance with the scatterer. If you are interested in atoms they have dimensions much less than the wavelength of visible light. So the probe cannot simply be put in resonance as far as the length scales is involved. The only way to getting them into resonance is that if there is an internal degree of freedom that couples to the light. Happily atoms and molecules have internal degrees of freedom and in this way resonance coupling can be achieved by matching the energy of the probe with the energy of the internal degrees of freedom.

The, linear and nonlinear, interaction of light with vibrational degrees of freedom of matter, is always treated within, what I would call, a Kubo-response formalism.[22] Not widely recognized is that this framework is very limited, not at all exact, and not at all as general as (formal) scattering theory.[23] The general shortcomings and break-down of Kubo's response theory is probably only fully appreciated in mesoscopic condensed mater physics, mainly due to the pioneering work of Landauer.[24] A major advantage of response theory is its simplicity, clearly demonstrated by the fact that no precise description of the probe is required in Kubo's response theory. This should already warn us about the limitations of this procedure as we would expect that the precise form of the probe Hamiltonian should matter. To emphasize the limitations of Kubo's response theory I will refer to it as "one-way response theory" in contrast to what I will call, and explain later, "two-ways response theory". In scattering theory the whole Hamiltonian is necessary and the precise forms of all the coupling terms are essential. This demand immediately points out that scattering theory is much more general than Kubo's one-way response theory.

One-way response doctrine is explicitly, or implicitly, used in almost all books on nonlinear optics. To emphasize the limitations of one-way response theory, I will describe in some detailed the Hamiltonian of a radiation field and how formal scattering theory deals with this problem. This presentation is not meant to be an introduction to scat-

tering theory and the equations presented are not meant to be understood in detail, and their consequences will not be dealt with in this paper. I will only give you a flavor of the limitations of one-way response theory.

Let us look at the Hamiltonian describing the probe dynamics, H_{probe}, which in this case is given by the radiation Hamiltonian H_{rad},[25,26]

$$H_{rad} = \frac{\epsilon_0}{2}\left(\frac{\partial \mathbf{A}}{\partial \mathrm{t}}\right)^2 + \left(\frac{\nabla \times \mathbf{A}}{2\mu_0}\right)^2 - \frac{\epsilon_0}{2}(\nabla\phi)^{\mathbf{2}}, \tag{20}$$

in which $\mathbf{A}$ and ϕ are the vector potential and the scalar potential of the electromagnetic field respectively. In equation (20) the coulomb gauge $\nabla \cdot \mathbf{A} = 0$ has been used. All the equations of motion of the free field could be obtained from this Hamiltonian using a Lagrangian formalism.[27]

The interaction with an atom at the origin of nuclear charge Z is through the charge and current densities,[26]

$$\sigma(\mathbf{r}) = -e\sum_j \delta(\mathbf{r} - \mathbf{r}_j) + Ze\delta(\mathbf{r}), \tag{21}$$

$$\mathbf{J}(\mathbf{r}) = -e\sum_j e\dot{\mathbf{r}}_j\delta(\mathbf{r} - \mathbf{r}_j). \tag{22}$$

In optical studies the polarization (density) $\mathbf{P}(\mathbf{r})$ and magnetization (density) $\mathbf{M}(\mathbf{r})$ are useful variables, and they are defined by

$$\sigma(\mathbf{r}) = -\nabla \cdot \mathbf{P}(\mathbf{r}), \tag{23}$$

$$\mathbf{J}(\mathbf{r}) = \dot{\mathbf{P}}(\mathbf{r}) + \nabla \times \mathbf{M}(\mathbf{r}). \tag{24}$$

The above set of equations are not simple and I will discuss later why you will hardly ever find for instance Hamiltonian (20) featuring in a text on optical studies of nature. I have treated the field classically which is a very good starting point. Only for experiments performed with low light intensities and in which higher-order electric-field correlation functions are measured, the need arises to introduce the complication of a quantum-mechanical description of light.[28] As the matter part still has to described quantum-mechanically for vibrational degrees of freedom, one calls this treatment "the semi-classical approach".

A tremendous simplification can be obtained in the coupling Hamiltonian when one realizes that the wavelength of light is much larger than the dimensions of the atom. This calls for an expansion in the ratio of these two length scales. To this end one calculates the potential energy of an atom in an electromagnetic field.[26] We only need the transverse part of the electric field as the longitudinal part is simply be described by electrostatics. This interaction energy is

$$H_{cou} = -\int \mathbf{E}_T(\mathbf{r}) \cdot \mathbf{P}(\mathbf{r}) d\mathbf{r}, \tag{25}$$

in which the subscript T refers to the transverse part.

An expansion in atomic length scale over wavelength amounts to an expansion of the polarization in multipoles and only the first term, containing the dipole moment μ is retained in the "dipole approximation",

$$\mathbf{P}(\mathbf{r}) = -e\sum_j \mathbf{r}_j\delta(\mathbf{r}) \equiv \mu\delta(\mathbf{r}), \tag{26}$$

and $\mathbf{E}_T(\mathbf{r})$ will be expanded around $\mathbf{r} = 0$, the center of the atom. Substituting these expansions into expression (25) gives as first term,

$$H_{dip} = -\mu \cdot \mathbf{E}_T(0). \tag{27}$$

So our simplified Hamiltonian describing the probing of our system with radiation looks like

$$H_{world} = H_{sys} + H_{dip} + H_{rad} \ , \tag{28}$$

where we have summed over all atoms. This result should bring about a big disappointment. We have been able to simplify the coupling to the electromagnetic field considerably without loosing any accuracy, but we are still left with the complicated radiation Hamiltonian featuring in H_{world} and so the whole system is still very complicated. In the one-way response theory, to be discussed in the next section, the radiation Hamiltonian will be absent.

The only way I know to handle Hamiltonian (28) properly is scattering theory. In a scattering formalism one describes the scattering of a wave from a potential. To use this approach for our case we have to identify H_{cou} with this potential. In scattering theory the wavefunction is partitioned in a scattered and an unscattered wave,

$$\mathbf{\Psi} = \mathbf{\Psi}_{incoming} + \mathbf{\Psi}_{scattered} \equiv \mathbf{\Psi}_{inc} + \mathbf{\Psi}_{sca}, \tag{29}$$

In scattering theory the solution is obtained by introducing the transition matrix T,

$$\mathbf{\Psi} = (1 + \mathbf{G}_0\mathbf{T})\mathbf{\Psi}_{inc}, \tag{30}$$

where $\mathbf{G}_0$ is the Green's function of the system without scatterers (that is for $H = H_{sys} + H_{rad}$). So $\mathbf{G_0}$ describes the propagation of the system and the free electromagnetic field. The T-matrix takes care of the coupling between them.[23] We will deal in more detail with these concepts in Section (9.1). Obtaining the full T-matrix amounts to obtaining an exact solution, so very often approximations are used, and one such an approximation is an expansion in the potential (Born series), another one is the (asymptotic) expansion in the kinetic energy (WKB method). Retaining only the first term in the, not necessarily converging, Born expansion of the T-matrix is labeled the first Born approximation.

A simplification very often introduced in scattering theory is to mimic the vector field by a scalar field. The justification for this is beyond the scope of the present lecture, but it is quite clear that all polarization phenomena are thrown away when adopting this simplification. The wave equation for the free field for scalar waves looks like

$$\left(\varepsilon(\mathbf{r})\frac{\partial^2}{\partial t^2} - \nabla\right)\Psi(\mathbf{r}, t) = 0\,, \tag{31}$$

in which $\varepsilon = 1/c^2$ is the dielectric constant.

I emphasize again that this is not an article on scattering theory and the scattering equations are only presented here to point out the shortcomings of conventional response theory.

8 (NON)LINEAR RESPONSE THEORY

Using the first term of the Born expansion of the T-matrix, for both the loss and the gain contribution to a a transition probability provides identical results as obtained by the familiar Fermi's Golden Rule. Fermi's Golden Rule is equivalent to Kubo's one way response theory. In a way one-way response theory is the reformulation of Fermi's Golden Rule in the time domain. If the exact answer is contained in the T-matrix, why does, presumably rigorous, one-way linear response theory only gives the first term of the Born expansion of the T-matrix? The answer is that standard one-way linear-response and one-way nonlinear response theory are very incomplete. It comes about when Hamiltonians (19) and (28) are treated as

$$H_{world} = H_{sys} + H'_{cou}. \tag{32}$$

H'_{cou} differs in a principal way from H_{cou}. In the original Hamiltonian the coupling term contained the field degrees of freedom (for instance the electric field) as dynamic variables, the dynamics of which was controlled by the probe Hamiltonian. In H'_{cou} these dynamic degrees of freedom are replaced by massive, adiabatic fields, the magnitude of which is described and controlled by the experimentalist. For our optical example the coupling Hamiltonian given by the dipole approximation H_{dip} is replaced by H'_{dip},

$$H'_{dip} = -\mu \cdot \mathbf{E}_{external}(0), \tag{33}$$

according to the one-way response dogma. The dynamic electric field $\mathbf{E}_T$ has now turned into an external field that does not seem to care about the response of the system. The probe dynamics has been turned off. This is a serious flaw of this approach.

In one-way optical response approach one seeks next an expansion in the response of the system in powers of the applied field. This scheme is possible because the field has been converted from a dynamic variable into a parameter. Collecting powers of the electric field in the response defines linear response and nonlinear response, expressed in terms of the linear and, notorious, nonlinear susceptibilities χ_n. When adhering to this strategy one encounters the following symbolic expansion for the polarization density,

$$\mathbf{P} = \chi_1 \mathbf{E} + \chi_2 \mathbf{EE} + \chi_2 \mathbf{EEE} + \cdots, \tag{34}$$

to be found in any book on nonlinear optics. The expressions given there for these susceptibilities are invariably the result of one-way response theory and thus incomplete because the response formalism does not allow, amongst other things, for a back reaction of the system to the field.

Which effects are not included in one-way response methodology? I will give some examples. Any system being probed is coupled to many more modes of the probe field than the one being excited by the experimentalist. In case of optical spectroscopy, the propagation from one mode of the probe field to another field mode is exactly what we call light scattering. It follows that all forms of probe field scattering are neglected in one-way response theory. If one would be a purist one could even maintain that the influence of absorption on the propagation of the probe field is not contained in one-way response theory. Changes in the index of refraction and polariton behavior are not described by one-way response theory. Cavity pulling is the effect, well-known to occur in high-Q (Q=quality factor) microwave and optical cavities, that the eigenfrequencies of system and (probe) cavity perturb each other. Cavity pulling cannot be described

by one-way response theory. The depletion of an incoming field is also not included in the one-way response doctrine. Some of these shortcomings that I mentioned can be treated, and have been treated, in a phenomenological way. One should realize that these phenomenological approaches are essentially patches to one-way response concepts.

One-way linear response ideology does not describe all the linear response to the field, but it only describes the effects that are linear in the coupling constant. A simple additional example should convince the reader. If low-intensity light is multiply elastically scattered in an inhomogeneous medium, like white paper, and then absorbed in the medium, this effect is linear in the incoming field. This manifestly linear absorption is not described by one-way linear response principles as the scattering is not contained in that approach. The failure of the one-way response concept is even worse in the quantum case which I am not going to examine.

Under what conditions can one use the one-way response philosophy? The material should be transparent for all the optical frequencies being used or generated. The quality factor of the probe cavity should be very small and the mass, or energy capacity, of the probe should be very large. In addition one should introduce some ad hoc patches like wave-vector and frequency dependent "local-field" corrections, that will account for effects like the change in index of refraction due to the presence of matter. This approach certainly does not deserve a price on a beauty contest but that is what is being done all the time. Given these conditions it is simple to contemplate situations where the one-way response hypothesis fails dramatically. A correct theory, which includes back reaction to the field, I would refer to as a "two-ways response theory". It seems obvious to me that such a theory will involve the full T-matrix. For non-propagating fields, like a DC electric field, two-ways response theory amounts to correctly accounting for the boundary conditions under which the experiment is being performed. Application of one-way response theory to these cases assumes implicitly one particular boundary condition out of the many possible under which the experiment can be performed. If one deals with nonlinear optics and one considers the generation of more than one field the reaction of all these fields to each other should be included, and I would refer to this as "all-ways response theory". An example where implicitly all-ways response theory has been used is the case of four-wave mixing in a disordered medium.[29]

9 DIRECTLY DRIVEN OSCILLATOR

Let us make a connection between the coupling Hamiltonian and our vibrational degree of freedom. This degree of freedom has some (generalized) coordinate q_j. The j^{th} oscillator is coupled to the electromagnetic field through its dipole moment. We assume the dipole moment to be a function of the coordinate q_j, and perform a Taylor expansion,

$$\mu_j = \mu(q_j = 0) + \left(\frac{\partial \mu_j}{\partial q_j}\right)_{q_j=0} q_j + \ \cdots . \tag{35}$$

When we have N identical oscillators the total dipole moment is represented by

$$\mathbf{p} = \sum_j^N \mu_j. \tag{36}$$

The interaction of the oscillators with the optical field is given by $E_{cou} = -\mathbf{p}\cdot\mathbf{E}$, and by using expansion (35) we find that there is a contribution proportional to q_j. As minus the gradient of an interaction represents a force, this term gives rise to a total force on the oscillators,

$$\mathbf{F} = N\mathbf{E}\cdot\left(\frac{\partial\mu}{\partial q}\right)_{q=0} . \tag{37}$$

Apparently we can drive an oscillator directly by an optical field only if its dipole moment changes during vibration. Our original equation of motion for the coherent amplitude Q of the vibrator did not contain an external force. Extending the equation of motion with an external driving force in the equation of motion results in

$$\frac{\mathrm{d}^2Q(t)}{\mathrm{d}t^2} + \frac{2}{T_2}\frac{\mathrm{d}Q(t)}{\mathrm{d}t} + \Omega_0^2 Q(t) = \left(\frac{N}{m}\right)\mathbf{E}\cdot\left(\frac{\partial\mu}{\partial q}\right)_{q=0} . \tag{38}$$

The solution of this equation of motion contains two classes of responses. The first type is the transient response and the second is the force-induced reaction. The transient response will have died away after the elapsed time exceeds several damping times and the driven reaction is the only lasting effect. This driven motion has two components, one in phase with the driving force and one 90^o out of phase with the driving field. The simplest and most convenient way to represent these results is to introduce a Fourier-Laplace transform of a time-dependent function $g(t)$,

$$\begin{aligned} g(z) &\equiv \int_0^\infty g(t)\mathrm{e}^{izt}, \\ z &= \omega + i\epsilon, \epsilon > 0 \ . \end{aligned} \tag{39}$$

The solution of equation (38) formulated with the help of this transform is

$$Q(z) = \left(\frac{1}{\Omega_0^2 - z^2 - iz2/T_2}\right)\left(\frac{N}{m}\right)(\partial\mu/\partial q)_{q=0}\cdot\mathbf{E}(z) \ , \tag{40}$$

$$= (1/2) \qquad (\Omega_0^2 - 1/T_2^2)^{-1/2}\left(\frac{1}{z-z_+} - \frac{1}{z-z_-}\right)\left(\frac{N}{m}\right)(\partial\mu/\partial q)_{q=0}\cdot\mathbf{E}(z), \tag{41}$$

in which

$$\begin{aligned} z_\pm &= [\pm(\Omega_0^2 - 1/T_2^2)^{1/2} - i/T_2], \\ &\approx (\pm\Omega_0 - i/T_2). \end{aligned} \tag{42}$$

We will follow closely Loudon[26] and try to make our results look as much as possible equivalent to case of absorption of visible light by atoms. Using Eqs. (36) and (40) shows that we will have an induced polarization ($\equiv$ dipole moment per volume V),

$$\mathbf{P}(z) = (\partial\mu/\partial q)_{q=0}Q(z)/V \equiv \epsilon_0\chi_1(z)\mathbf{E}(z). \tag{43}$$

The definition for the linear, second-rank tensor, susceptibility implies

$$\chi_1(z) = \left(\frac{N}{mV\epsilon_0}\right)(\partial\mu/\partial q)_{q=0}^2\left(\frac{1}{\Omega_0^2 - z^2 - iz2/T_2}\right). \tag{44}$$

To compare the magnitude of this susceptibility with other optical transition probabilities one could introduce the oscillator strength

$$\mathbf{f} \equiv f_0(\partial\mu/\partial q)_{q=0}^2/e^2, \tag{45}$$

where f_0 is introduced to allow for the fact that not all oscillators have frequency Ω_0 and f_0 represents the fraction that does. The oscillator strength can, and should, be given a full quantum-mechanical interpretation. In principle $\mathbf{f}$ is a second-rank tensor. From now on we will drop the vector and tensor notation in this section as different combinations of cartesian components can easily be traced back. We rewrite the expression for the linear susceptibility by using the definition of the oscillator strength and arrive at

$$\chi_1(z) = \left(\frac{fNe^2}{m\epsilon_0 V\Omega_0^2}\right)\left(\frac{\Omega_0^2}{\Omega_0^2 - z^2 - iz2/T_2}\right). \tag{46}$$

$$\equiv S\frac{\Omega_0^2}{\Omega_0^2 - z^2 - iz2/T_2}, \tag{47}$$

where I have defined, following Loudon's treatment[26] the dimensionless parameter S, $S \equiv (fNe^2)/(mV\epsilon_0\Omega_0^2)$. A good estimate of the size of S will be derived later, here it suffizes to say that the answer is that S is approximately equal to the difference between the dielectric constant at a frequency much lower and the dielectric constant at a frequency much higher than the eigenfrequency of the oscillators. The, somewhat artificial, introduction of the oscilator strength has the virtue that a number of our following equations apply too many other situations, besides the case of *vibrational* degrees of freedom coupled to an electromagnetic field, as well. Only when actual numbers have to be computed differences will appear and the oscillator strength has to be evaluated for each instance separately.

It is obvious that the susceptibility and the induced displacement have two resonances, situated at $z \approx \pm\Omega_0 + i\epsilon$. Concentrating on the resonance at positive frequency (that is neglecting the counter-rotating term) yields

$$\chi_1(z \approx \Omega_0 + i\epsilon) = \left(\frac{S\pi\Omega_0}{2}\right)\left(\frac{1}{\pi}\frac{1}{\Omega_0 - z - i/T_2}\right), \tag{48}$$

$$\equiv S\frac{\pi\Omega_0}{2}L(z;\Omega_0,T_2), \tag{49}$$

where I have defined the normalized "complex" Lorentzian $L(z;\Omega_0,T_2)$. The linear susceptibility χ_1 can be partitioned in a real and an imaginary part according to

$$\chi_1(z = \Omega + i\epsilon) \equiv \chi_1'(\Omega) + i\chi_1''(\Omega). \tag{50}$$

In the following we will indicate $\Omega + i\epsilon$ by Ω^+. (The use of the infinitesimal complex number ϵ is just a mathematical trick to write the sine and cosine transform of a time-dependent function in a convenient and compact way in one formula.) If we would use as external drive for our system of harmonic oscillators a harmonic driving force, $F(t) \propto \mathbf{E}(t) = Re(\mathbf{E}_0 e^{i\Omega t}) \rightarrow F(z) \propto i/(z + \Omega)$, the real and imaginary part would give the in-phase and out-phase component, respectively. We notice that the out-phase component χ_1'' is proportional to a normalized Lorentzian lineshape,

$$L''(\Omega;\Omega_0,T_2) \equiv \frac{1}{\pi T_2}\frac{1}{(\Omega_0 - \Omega)^2 + (1/T_2)^2}, \tag{51}$$

and the in-phase component χ_1' is proportional to, what is sometimes called, a "dispersive Lorentzian"

$$L'(\Omega;\Omega_0,T_2) \equiv \frac{1}{\pi}\frac{(\Omega_0 - \Omega)}{(\Omega_0 - \Omega)^2 + (1/T_2)^2}. \tag{52}$$

Standard arguments based on energy flow, show that the rate at which energy (per volume and cycle-averaged) is absorbed from a time-varying electric field with frequency Ω is[30]

$$\frac{\partial U}{\partial t} =\ll \mathbf{E}(t)\cdot\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t}\gg\propto \chi''(\Omega). \tag{53}$$

This outcome is typically a fruit of one-way response theory, and in the next section I will point out that it only holds under limiting conditions.

According to equation (53) the observation of absorption out of a harmonic field as a function of the frequency of this field supplies us, experimentalists, with both Ω_0 and T_2. In addition to these parameters we also would like to infer T_1 from our investigations. T_1 relates to population dynamics which are expected to become significant only at higher light intensities. Up to now there is no nonlinear response to the external field. A damped harmonic oscillator does not have any nonlinear response. Consequently there is also no saturation, which seems rather unphysical. The only way to encompass nonlinear response in a model for a classical oscillator is to introduce anharmonicities, but anharmonicities are notoriously difficult to deal with. Happily if we use a quantum-mechanical picture the situation becomes simpler. In practice energy scales and temperatures are such that in the study of vibrational dynamics it is hardly ever allowed to envisage the dynamics from the standpoint of classical (statistical) mechanics anyway. The simplest quantum-mechanical picture of an anharmonic oscillator is to limit oneself to the two lowest levels of the oscillator. This means we are discussing as a model for our oscillator a two-level system. A two-level system can be considered the embody the "ultimate anharmonic oscillator". The saturation behavior can now be incorporated by explicitly discussing the population dynamics between the two levels. This can be done in the following way. The proportionality factor N in the force on the oscillator will be interpreted as the (dynamic) difference in population between excited and ground state, $\Delta N(t)$. An additional equation for the population dynamics is then easily established. We arrive at[30]

$$\frac{\mathrm{d}^2Q(t)}{\mathrm{d}t^2}+\frac{2}{T_2}\frac{\mathrm{d}Q(t)}{\mathrm{d}t}+\Omega_0^2Q(t) = \frac{\Delta N(t)}{m}\mathbf{E}\cdot\left(\frac{\partial\mu}{\partial q}\right)_{q=0}, \tag{54}$$

$$\frac{\mathrm{d}\Delta N(t)}{\mathrm{d}t}+\frac{\Delta N(t)-\Delta N_0}{T_1} = \frac{-2}{\hbar\Omega_0}\mathbf{E}(t)\cdot\frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t}, \tag{55}$$

in which ΔN_0 is the Boltzmann equilibrium population difference. In the rhs. of equation (55) the rate is presented at which energy (in units $\hbar\Omega_0$) is absorbed from the external electric field. Per transition the population difference changes by a factor of two and the negative sign comes from the fact that an increasing field intensity will lead to a decreasing population difference. The steady-state solution to both equations can best be obtained by introducing as ansatz, $Q(t)=Q_0(t)\mathrm{e}^{i\Omega_0 t}$, and keeping only the lowest-order time derivatives. In this way the effect of the counter-rotating field has been eliminated. Requesting that $\frac{\mathrm{d}\Delta N(t)}{\mathrm{d}t}=\frac{\mathrm{d}Q_0(t)}{\mathrm{d}t}=0$ gives as solution of Eqs. (54) and (55)

$$\mathbf{P}(z)\equiv\epsilon_0\chi(z)\mathbf{E}(z), \tag{56}$$

in which the full nonlinear susceptibility

$$\chi(z\approx\Omega_0^+)=\left(\frac{S\Omega_0}{2}\right)\left(\frac{1}{\Omega_0-z-i/T_2+\frac{\Delta^2T_1/(2T_2)}{\Omega_0-z+i/T_2}}\right), \tag{57}$$

in which Δ is the Rabi frequency, $\Delta \equiv \frac{(\partial\mu/\partial q)_{q=0}E_0}{2\hbar}$. It is evident that our coupled system shows nonlinear behavior and saturation for large Rabi frequencies. We notice that from the saturation behavior we can determine both T_2 and T_1, and these were the relaxation parameters we were after.

Eqs. (54) and (55) with solution (57) are well-known in magnetic resonance, where they are called the Bloch equations. To make contact with the formalism of nonlinear susceptibilities used in optics, we will apply a simple mathematical trick. The Rabi frequency is located in the denominator of solution (57). The Bloch solution can be expanded in a power series of the Rabi frequency. We will rewrite the expression for the susceptibility as

$$\chi(z \approx \Omega_0^+) = \left(\frac{S\Omega_0}{2}\right)\left(\frac{1}{\Omega_0 - z - i/T_2}\right)\left(\frac{1}{1 + \frac{\Delta^2 T_1/(2T_2)}{(\Omega_0 - z)^2 + 1/T_2^2)}}\right). \tag{58}$$

The last, dimensionless, factor on the rhs. of equation (58) is of the form $1/(1+\alpha)$, and can be expanded in powers of α when α is small. I hope that the reader recognizes that this indeed generates an odd power series in E_0 with the successive terms represent χ_1, χ_3, and higher-order susceptibilities. This procedure makes immediately transparent that for saturation we need *all* the nonlinear susceptibilities, up to infinite order. In addition we come to the somewhat surprising conclusion from our expansion that if the Rabi frequency is very large the series expansion of the polarization density in nonlinear susceptibilities does not even converge.

I have stated that higher-order correlation functions of the operator Q could be decoupled in the weak-coupling scheme. The existence of decoupling within the weak-coupling situation could already be concluded from the solution to the Bloch equations. We have solved the Bloch equations for all powers of the electric field. When we expanded this solution in powers of the electric field we found *explicit* answers for all the nonlinear susceptibilities χ_n. A formalism calculating these susceptibilities from one-way linear response theory demonstrates that the higher-order nonlinear susceptibilities are represented by higher-order correlation functions of the dipole-moment operator, and through Taylor expansion (35), by higher-order correlation functions of the displacement operator Q. Apparently we have calculated all these higher-order correlation functions, and this is only possible if somewhere the decoupling approximation has crept inside. The answer is that the Bloch equations (54) already imply a decoupling of the dynamic correlation functions, as first the averaging over the bath has been introduced and then the influence of the electric field. A proper, but much more dificult, way would have been to reverse the order of these two actions.

From the experimental observation of saturation behavior of the susceptibility of the vibrational degrees of freedom the eigenfrequency and the relaxation parameters can be obtained for infrared active oscillators. Such experiments could be realized in the frequency as well as the time domain. Given the fact that high light intensities are easier to come by with pulsed light sources, the experiments are commonly carried out in the time domain. The simplest implementation is a one-color two-pulse infrared experiment, in which the first, intense, infrared *pump* pulse saturates the transition and the increased transmission of a delayed, possibly weak, *probe* pulse monitors the recovery of the population difference. For this purpose one needs a tunable, and powerful, pulsed infrared source. One possibility, and one of the most convenient realizations, is

to use parametric down conversion in nonlinear optical crystals of the light beam at the fundamental frequency of a Nd-Yag laser.[31] Handicaps of this technique are the rather small maximum wavelengths of about 4.5 μ and the rather long pulse widths of about 20 picoseconds.

As explained earlier vibrational dephasing times are under normal conditions (temperatures much larger than $4K$) much shorter than inelastic times. When monitoring a population difference on its characteristic T_1 time scale all coherence has been lost and the dynamics of the coherent amplitude $Q(t)$ can be neglected. In such experiments the rate processes to deal with are only the population and depopulation kinetics of the various levels. Many types of experiments have been conceived to study the kinetics of level occupation and it is absolutely impossible to mention even a fraction of them here, leave alone review or explain them. I will make one exception and, as explained in the introduction, work of our own group is involved. Bakker has performed an extensive study of the dynamics C-H stretch vibrations in various liquids using infrared saturation spectroscopy.[32] An interesting feature has been discovered in these studies. In many situations the return of a transferred population to the ground state was found to evolve via an intermediate excited vibrational state of the same molecule. This mechanism is referred to as IVR (Intra-molecular Vibrational Relaxation). Subsequently this excited low-energy mode (or modes) redistributes its energy among the many surrounding liquid molecules. This decay mechanism is referred to as IET (Intermolecular Energy Transfer). The time constants associated with IVR and IET could both often easily be obtained when monitoring the transmission of the probe beam as a function of the time delay because this intermediate vibrational state of the molecule has a transition probability for the probe pulse differing from the transition probability of the molecule in the ground state. This can most easily be interpreted in terms of a shift of the absorption band induced by anharmonicities. Depending on the sign of the change in spectral line position and on the frequency of the infrared light with respect to the center of the absorption band, the probe beam can be pushed more into or more out of resonance. If the mode is driven more into resonance one observes the amusing feature that during the decay of the non-equilibrium population transmission of the probe can temporarily be *less* than the long-time equilibrium situation. See Figure (2).

Many more sophisticated nonlinear optical techniques, besides simple saturation spectroscopy, have been devised in order to extract the same relaxation parameters. I have not treated the complication of inhomogeneous broadening. Some of the nonlinear techniques, like photon-echo spectroscopy, allow one to get rid of the complications of inhomogeneous broadening, and admit a determination of the pure relaxation parameters. The coupling of light to an infrared active oscillator is usually strong. This explains why many nonlinear techniques can be and have been applied with great success. Sometimes the coupling to the light field is so strong that the propagation of the light is seriously influenced, an effect not taken into account yet.

9.1 Polaritons

The coupling of light to a collection of oscillators can have a profound effect on the propagation of light. We call this polariton behavior. Polariton action is connected to the fact that infrared active oscillators radiate light back into the system. Such a phenomenon will never be included in a one-way response theory and indeed this back

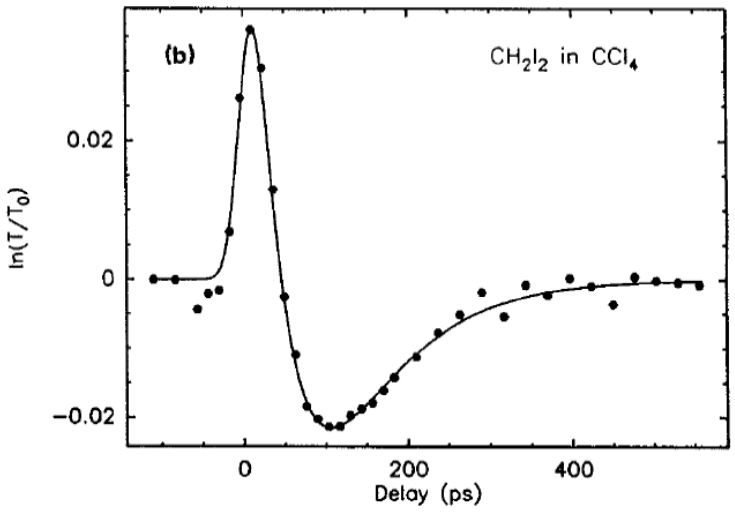

Figure 2. *Relative transmission,* $\ln(T/T_0)$*, of an infrared probe pulse as a function of the delay between probe and pump pulse for the* 3065 cm^{-1} *asymmetric* CH_2 *stretch vibration of* CH_2I_2 *dissolved in* CCl_4*. Pump and probe have the same frequency spectrum and are tuned below the maximum of the absorption band. Clearly two kinetic recovery proceses can be identified. After Reference (32).*

reaction to the field was neglected when deriving the one-way response formula (53). The complication of the presence of polaritons has to be fully appreciated by the experimentalist as an ill-understood propagation behavior would lead to erroneous conclusions about the susceptibility of the oscillators, and would lead to wrong deductions about the relaxation time T_2. There is quite some confusion about what a polariton essentially is. Any resonance structure in the index of refraction should be called polariton behavior, in my opinion. Any coupling to an infrared active oscillator will lead to polariton behavior. Sometimes this coupling is described as resulting in a hybridized mode in which a photon is coupled to another quantized field, for instance a phonon, and the new mixed modes are called polaritons. This representation gives the impression that the concept of a polariton finds its origin in quantum mechanics. However, the classical Maxwell equations condone polariton behavior. If one is only interested in the transmission of light and not in the propagation of the material excitation (which is often the case when the dipole approximation is valid) one can integrate out the dynamic material degrees of freedom, and their behavior is lumped into parameters like the index of refraction and the mean free path for light. At this level the microscopic origin of these parameters, either of classical or of quantum-mechanical nature, is of no concern any longer and maybe the term polarization wave would be more appropriate. A very intimate connection exists between polariton behavior and scattering theory, that is very important but hardly ever emphasized, so I will do it here. The close relationship with scattering theory is usually overlooked, as one apparently does not realize that, from the view point of the amplitude of light, there is no distinction between elastic scatterers and real absorbers. Only the sum of these two effects, elastic scattering and real absorption, counts, and this combined effect is called extinction.

The key feature of polariton theory is the fact that the presence of infrared active degrees of freedom influence the dielectric constant of the medium. The best method to incorporate this effect is along the lines of scattering theory, but it is not necessary. We could deduce the change in index of refraction directly from our equation-of-motion approach. A major disadvantage is that in this scheme all the implicit assumptions remain hidden, whereas they have to be made manifest in the scattering approach in a very elegant manner. In addition the partitioning of extinction in elastic scattering and real absorption follows most naturally and harmoniously from a full scattering approach. We will review some multiple-scattering theory as all polariton effects can be classified as multiple-scattering phenomena. Nowadays, there is a tremendous revival of in interest in multiple-scattering events and the realization of these analogies will be beneficial to all of us.[33] For simplicity we will treat the electromagnetic waves in the following as scalar waves with amplitude Ψ.

An important aspect of wave dynamics and propagation is Huygens' principle. The mathematical expression for this principle is couched in a function called a Green's function, which in free space is essentially a spherical wave. The free (unperturbed by scatterers) Green's function in real space only depends, besides on the frequency of the wave, on the *difference* of two spatial variables,

$$G_0(\mathbf{r}_1, \mathbf{r}_2; \Omega^+) = -\frac{e^{ik_0^+|\mathbf{r}_1-\mathbf{r}_2|}}{4\pi \mid \mathbf{r}_1 - \mathbf{r}_2 \mid}, \tag{59}$$

where $k_0 = \Omega/c$. Free space is homogeneous and momentum is conserved, which can be fully exploited by going to wavevector space,

$$G_0(\mathbf{p}; \Omega^+) = -\frac{1}{p^2 - (\Omega^+/c)^2}. \tag{60}$$

Let us first look at what one scatterer does. As always we partition the wavefunction in a scattering geometry in an incoming and a scattered wave

$$\mathbf{\Psi} = \mathbf{\Psi}_{inc} + \mathbf{\Psi}_{sca}. \tag{61}$$

The simplest possible scatterer is an isotropic point scatterer. The concept of an isotropic point scatterer in a wave equation is not without mathematical difficulties,[34] but we will not be sidetracked here by requesting too much mathematical rigor. Part of an incoming plane wave, with direction given by the unit vector $\hat{\mathbf{k}}$, will be scattered by the jth point scatterer located at $\mathbf{r}_j$ into a spherical wave and the result is that

$$\Psi(\mathbf{r}; \Omega^+) = e^{ik_0^+ \hat{\mathbf{k}}\cdot\mathbf{r}} + \frac{f_j(\Omega^+) e^{ik_0|\mathbf{r}-\mathbf{r}_j|}}{\mid \mathbf{r} - \mathbf{r}_j \mid}, \tag{62}$$

in which the, complex-valued, single-particle scattering length or scattering amplitude f has been specified. The cross-section for scattering is given by

$$\sigma_{sca} = 4\pi \mid f(\Omega^+) \mid^2 \tag{63}$$

The calculation of the flux or the intensity of a wave involves the product of two wave amplitudes. We will only do this in a symbolic way

$$\mid \mathbf{\Psi} \mid^2 - \mid \mathbf{\Psi}_{inc} \mid^2 = \mid \mathbf{\Psi}_{sca} \mid^2 + (\mathbf{\Psi}_{sca}\mathbf{\Psi}^*_{inc} + c.c.). \tag{64}$$

The first term on the rhs. of this equation containing a product of two scattering amplitudes describes the scattered intensity. The left side contains the total outcoming

intensity minus all the incoming intensity, and a nonvanishing negative number would indicate that the scatterer has absorbed part of the radiation. Apparently (minus) the second term on the rhs. of equation (64), characterizing the interference between scattered wave and incoming wave, accounts for all the intensity being taken away from the incoming beam. This describes the extinction. If there is no real absorption and all the extinguished intensity reappears as scattering intensity the lhs. of equation (64) vanishes. This puts a constraint on the scattering amplitudes

$$\mathrm{Im}f(\Omega^+) = k_0 \mid f(\Omega^+) \mid^2, \tag{65}$$

which is known as the optical theorem,[35] taking this simple form because of the assumed isotropic character of the scattering.

Formally we would like to describe the scattering process as scattering of the incoming wave from $\mathbf{r}'$ to $\mathbf{r}''$ and then propagation from $\mathbf{r}''$ to $\mathbf{r}$. Integration over all the primed coordinates would generate the scattered light at $\mathbf{r}$. We arrive at this by defining a single-particle transition matrix $t(\mathbf{r}, \mathbf{r}')$ through

$$\Psi(\mathbf{r}; \Omega^+) = \Psi_{inc}(\mathbf{r}; \Omega^+) + \int \cdots \int G_0(\mathbf{r}, \mathbf{r}''; \Omega^+) t_j(\mathbf{r}'', \mathbf{r}'; \Omega^+) \Psi_{inc}(\mathbf{r}') \mathrm{d}\mathbf{r}' \mathrm{d}\mathbf{r}''. \tag{66}$$

It is evident that our original result (62) can be recovered by using

$$t_j(\mathbf{r}', \mathbf{r}; \Omega^+) = -4\pi f_j(\Omega^+) \delta(\mathbf{r} - \mathbf{r}_j) \delta(\mathbf{r}' - \mathbf{r}_j), \tag{67}$$

and by using as the incoming wave the plane wave defined in equation (62). The two delta functions in equation (67) illustrate convincingly the zero-range of the scattering potential. From now on we will use only a matrix notation in real space as the number of convolution integrals would be to large. This matrix notation implies integration over all dummy spatial variables. The rewriting in symbolic, or matrix, notation of equation (66)

$$\mathbf{\Psi} = (1 + \mathbf{G}_0 \mathbf{t}_j) \mathbf{\Psi}_{inc}, \tag{68}$$

has a much simpler look indeed. A similar equation was alread presented in Section (7), compare equation (30), where we were dealing with the interaction between light and matter.

What happens when we have many scattering centers? All waves can be scattered once, twice up to an infinite number of times. All these events will be contained in the many-particle T-matrix. In the same symbolic notation we define

$$\mathbf{\Psi} = (1 + \mathbf{G}_0 \mathbf{T}) \mathbf{\Psi}_{inc}. \tag{69}$$

All we have to do is to find the full T-matrix, which is unfortunately a horrendous undertaking. In order to make progress some dramatic assumption has to be advanced. All single-scattering events can symbolically be represented by $\sum_j \mathbf{t}_j$, and the double events by $\sum_j \sum_k \mathbf{t}_j \mathbf{G}_0 \mathbf{t}_k$ etc. The total T-matrix is given by

$$\mathbf{T} \;=\; \sum_j \mathbf{t}_j + \sum_j \sum_k \mathbf{t}_j \mathbf{G}_0 \mathbf{t}_k + \sum_j \sum_k \sum_l \mathbf{t}_j \mathbf{G}_0 \mathbf{t}_k \mathbf{G}_0 \mathbf{t}_l + \cdots \quad . \tag{70}$$

The scattering centers are located at random positions and we have to average over all positions (indicated by $\langle \cdots \rangle$). This was discussed in Section (6). To take the mean over all particle positions becomes very cumbersome because of the existence of strong

correlations when in the products a particular particle appears more than once. The simplest possible hypothesis is to request that we retain only those contributions in the averaging of the terms in summation (70) of the total T-matrix where *no specific particle occurs more than once*. With this sledgehammer approach we find

$$\langle \mathbf{T} \rangle \approx \langle \sum_j \mathbf{t}_j \rangle + \langle \sum_j t_j \rangle \mathbf{G}_0 \langle \sum_k t_k \rangle + \langle \sum_j t_j \rangle \mathbf{G}_0 \langle \sum_k t_k \rangle \mathbf{G}_0 \langle \sum_l t_l \rangle + \cdots \quad . \tag{71}$$

All recurrent scattering events are ignored, and we have called this assumption the "independent scattering approximation". It seems that with this supposition this series can easily be summed as it has the looks of a geometric series with $\langle \sum_j \mathbf{t}_j \rangle$ as the first term and $\langle \sum_j \mathbf{t}_j \rangle \mathbf{G}_0$ as the multiplicative factor. Now our symbolic notation has tricked us because, if we would restore the spatial coordinates we would still find a number of convolution integrals. Happily, this cluttering can easily be entangled by proceeding to wavevector space, where due to translational symmetry, all convolution integrals will become simple products. The averaged Green's function will then be

$$G(\mathbf{p};\Omega^+) = \frac{\frac{-1}{p^2-(\Omega^+/c)^2}}{1+\frac{(N/V)t(\Omega^+)}{p^2-(\Omega^+/c)^2}}, \tag{72}$$

$$= \frac{-1}{p^2-(\Omega/c)^2+(N/V)t(\Omega^+)}. \tag{73}$$

where the first, awkwardly looking, equation has been presented for educational reasons, because it demonstrates lucidly that we are really dealing with summing a geometric series. The t-matrix element in this equation is factually $t(\mathbf{p},\mathbf{p};\Omega^+)$ but the momentum variables have been been dropped as for a point scatterer there is no momentum dependence. The optical theorem can now be rephrased in terms of the t-matrix,

$$\text{Im}\, t(\Omega^+) = -\frac{k_0}{4\pi}\, |\, t(\Omega^+)\,|^2 \,. \tag{74}$$

If not all the scattering is elastic this relation does not hold. The ratio of these two quantities (rhs. over lhs.) is called the albedo a , and is a number between one and zero. When the albedo equals one the scattering cross-section equals the total cross-section. A vanishing albedo implies a total cross-section equal to the cross-section for absorption and a vanishing cross-section for scattering. To give an example, for white paint the albedo is about 0.9999 for visible light, whereas I will show that for our oscillators it is much smaller for infrared light. Equation (73) shows immediately that one can define an effective, complex valued, dielectric constant $\varepsilon(\Omega^+)$,

$$G(\mathbf{p};\Omega^+) = \frac{-1}{p^2-\varepsilon(\Omega^+)(\Omega/c)^2}, \tag{75}$$

where

$$\varepsilon(\Omega^+) \equiv 1-(N/V)(c/\Omega)^2 t(\Omega^+). \tag{76}$$

This equation should be compared with empty-space equation (59), that can also be obtained from equation (75) by setting in equation (76) the density of scatterers, N/V, equal to zero. The implications of a complex valued dielectric constant become clear when we go back to real space. To facilitate the process of spatial Fourier transforming we define the effective, complex valued, index of refraction n,

$$n^2(\Omega^+) \equiv \varepsilon(\Omega^+) \equiv \varepsilon'(\Omega) + i\varepsilon''(\Omega), \tag{77}$$

$$n(\Omega^+) \equiv \eta(\Omega) + i\kappa(\Omega). \tag{78}$$

To avoid too many formulas that look clumsy and complex the frequency arguments will often be dropped from the real and imaginary parts of the dielectric constant ϵ and from the index of refraction n. Definitions (77) and (78) imply

$$\varepsilon' = \eta^2 - \kappa^2, \tag{79}$$
$$\varepsilon'' = 2\eta\kappa. \tag{80}$$

The propagation of light in real space is represented by the real-space Fourier transform of equation (75)

$$G_0(\mathbf{r}_1, \mathbf{r}_2; \Omega^+) = -\frac{e^{i\eta k_0^+ |\mathbf{r}_1 - \mathbf{r}_2|} e^{-\kappa k_0^+ |\mathbf{r}_1 - \mathbf{r}_2|}}{4\pi \mid \mathbf{r}_1 - \mathbf{r}_2 \mid}. \tag{81}$$

This very sensible outcome represents a damped wave with a renormalized wavevector

$$k \equiv k' + ik'' \equiv k_0(\eta + i\kappa). \tag{82}$$

The imaginary part κ of the index of refraction reflects the exponential decay in length and its effect is usually described in terms of a mean free path $\ell_{ext} \equiv 2/(\kappa k_0)$, where we have added the subscript to emphasize that we are monitoring the extinction rather than the scattering.

Let us explicitly find the expressions for η and κ in terms of ε' and ε''. Using

$$\eta = \frac{1}{\sqrt{2}} \sqrt{\eta^2 - \kappa^2 + \sqrt{(\eta^2 - \kappa^2)^2 + 4\eta\kappa)}} \quad , \tag{83}$$

we find that

$$\eta = \frac{1}{\sqrt{2}} \sqrt{\varepsilon' + \sqrt{\varepsilon'^2 + \varepsilon''^2}} \quad , \tag{84}$$
$$\kappa = \varepsilon''/(2\eta). \tag{85}$$

Given the complex index of refraction allows us to define the phase velovity v_p as

$$ck'/k_0 \equiv v_p = c/\eta. \tag{86}$$

For pulse propagation the group velocity is relevant,

$$v_g = \frac{\partial \Omega}{\partial k'} = 1/(\partial k'/\partial \Omega), \tag{87}$$
$$v_g = \frac{1}{\eta + \Omega(\partial \eta/\partial \Omega)}. \tag{88}$$

Infrared active oscillators scatter and possibly absorb light, so each of them can be described by a single-particle t-matrix. We have not treated our oscillators within the framework of t-matrices, but we have used an approach based on an explicit equation of motion of the vibrators. We should by now have become very curious about what t-matrix applies actually to our oscillator. This is a crucial step in making the connection between standard polariton theory and scattering theory. The relationship can be found

easily though the use of the well-known link between the dielectric constant and the linear susceptibility,

$$\begin{aligned}
\varepsilon(z) &= 1+\chi_1(z), && (89)\\
&= 1+S\left(\frac{\Omega_0^2}{\Omega_0^2-z^2-iz2/T_2}\right), && (90)\\
&= 1+S\left(\frac{(\Omega_0^2)(\Omega_0^2-z^2)}{(\Omega_0^2-z^2)^2+z^2 4/T_2^2}\right)+iS\left(\frac{(\Omega_0^2)z2/T_2}{(\Omega_0^2-z^2)^2+z^2 4/T_2^2}\right), && (91)\\
&= \varepsilon'_\infty+(\varepsilon'_0-\varepsilon'_\infty)\left(\frac{\Omega_0^2}{\Omega_0^2-z^2-iz2/T_2}\right), && (92)
\end{aligned}$$

where the last equation has been presented to make contact with a convention that is popular in the solid state physics community, and corroborates our earlier identification of S with dielectric constants. From relations (76) and (89) we conclude that the t-matrix of an infrared active oscillator

$$\begin{aligned}
t(\Omega^+) &= -(V/N)\chi_1(\Omega^+)(\Omega/c)^2, && (93)\\
&= -\left(\frac{fe^2}{m\epsilon_0 c^2}\right)\left(\frac{\Omega^2}{\Omega_0^2-\Omega^2-i2\Omega/T_2}\right), && (94)\\
&\equiv -4\pi r_o\left(\frac{\Omega^2}{\Omega_0^2-\Omega^2-i2\Omega/T_2}\right), && (95)\\
t(\Omega^+\approx\Omega_0^+) &= -2\pi r_o\left(\frac{\Omega}{\Omega_0-\Omega-i/T_2}\right), && (96)
\end{aligned}$$

in which the length r_o has been introduced. Our search path for an expression for the t-matrix really has been a round about. I have preferred this approach over the one in which I would have solved the scattering of light of one oscilator directly. Of course our "derivation" of the t-matrix does not depend on whether or not relation (93) holds for any density since the identification may be made for arbitrarily small density.

The cross-section for scattering of the oscillator is given by

$$\sigma_{sca}=(1/4\pi)\,|\,t(\Omega^+)\,|^2=4\pi r_o^2\left(\frac{\Omega^4}{(\Omega_0^2-\Omega^2)^2+4\Omega^2/T_2^2}\right) \qquad (97)$$

With these results we can find out what the albedo is of our oscillator.

$$a\equiv\frac{-\frac{k_0}{4\pi}\,|\,t(\Omega^+)\,|^2}{\mathrm{Im}t(\Omega^+)}\approx\frac{r_o k_0\Omega T_2}{2}, \qquad (98)$$

in the neighborhood of the resonance. Before we will estimate this number we will ask ourselves when is the albedo equal to one and are we talking about elastic scattering. The answer is of course that in that case the only dephasing comes from the radiation field and T_2 will be the radiation time T_{rad},

$$T_{rad}=2/(r_o k_0\Omega). \qquad (99)$$

Knowing that radiation times in the visible are varying roughly from nano to microseconds, and realizing that the scaling is with Ω^3 (from the Einstein coefficient), we estimate these times in the infrared to vay between micro and milliseconds. Given dephasing times of the order of picoseconds we find that the albedo roughly varying between 10^{-6} and 10^{-9} . Consequently all scattering by our oscillators is absorption

and the elastic scattering can safely be neglected.

The relation between effective dielectric constant and effective dielectric index of refraction has some important complications. This can best be seen when there is no dephasing (infinite T_2). Under this (unphysical) condition $\varepsilon'' = 0$ and

$$\eta = \left[1 + S\left(\frac{\Omega_0^2}{\Omega_0^2 - \Omega^2}\right)\right]^{1/2}, \quad \text{if } \varepsilon'' = 0. \tag{100}$$

When we look at a frequency slightly larger than the resonance frequency we will find that we cannot take the square root as the argument will be negative. There is a large counteracting polarization as all the oscillators lag the optical force and they interfere destructively with the incoming field. There will be total reflection described by an evanescent wave. The decay length of this evanescent wave is related to fully imaginary root of equation (100). For very large frequency the argument of the square root will be positive again. There will be a "stopband" that is located at $\Omega_0 < \Omega < \Omega_0(\varepsilon_0'/\varepsilon_\infty')$.

If the damping is not zero, the situation becomes more intricate. We now define the dimensionless "polariton parameter", $\mathcal{P}$,

$$\mathcal{P} \equiv 8S\Omega_0 T_2 \quad . \tag{101}$$

The factor of eight is for cosmetic reasons. The smaller $\mathcal{P}$ the lesser the polariton behavior. An expansion in small $\mathcal{P}$ gives.

$$\eta \approx 1 + \frac{S}{4}\frac{\Omega_0(\Omega_0 - \Omega)}{(\Omega_0 - \Omega)^2 + 1/T_2^2}, \tag{102}$$

$$\kappa \approx \frac{S}{4}\frac{\Omega_0/T_2}{(\Omega_0 - \Omega)^2 + 1/T_2^2}, \tag{103}$$

$$1/\ell_{ext} = 2\kappa k_0 \approx \frac{S}{2}\frac{\Omega_0^2/(cT_2)}{(\Omega_0 - \Omega)^2 + 1/T_2^2} \tag{104}$$

The smaller the damping or the higher the density of scatterers the more pronounced the polariton behavior is. Our previous example of the stopband for a dampingless system corresponds to $\mathcal{P} = \infty$. In contrast to the case of $\mathcal{P} = \infty$ we will always have a solution for both η and κ when $\mathcal{P}$ is finite and no real stopband seems to occur, as there is always some propagation going on, characterized by a real η. However we should still define the stopband to occur when $\varepsilon' < 0$, implying η to become smaller that κ. In this situation the wave is overdamped (the damping length is shorter than the wavelength divided by 2π), and indeed no substantial propagation is going on. We have defined $\mathcal{P}$ in such a way that $\varepsilon' < 0$ corresponds to $\mathcal{P} > 2$.

When $\mathcal{P}$ is not small the dispersion relation of light has a funny look as it displays an S-curve. Figure (3). Here the frequency has been plotted as a function η.

I will now show that the expansion parameter $\mathcal{P}$ is equivalent too the much more famous expansion parameter $(k'\ell_{ext})^{-1}$ in multiple scattering. A stopband exists for $\mathcal{P} > 2$, but in terms of η and κ it occurs when $\eta < \kappa$ or when $\eta/(2\kappa) \equiv k'\ell_{ext} < 0.5$. Indeed $\mathcal{P}$ and $(k'\ell_{ext})^{-1}$ are equivalent and this identification makes an interesting connection with multiple-scattering theory. In localization of waves one says that it occurs when $k'\ell_{sca} < 1$, where $\ell_{sca} \equiv (1/a)\ell_{ext}$ There are several points to be made here. Localization is concerned with transport properties rather than amplitude properties like

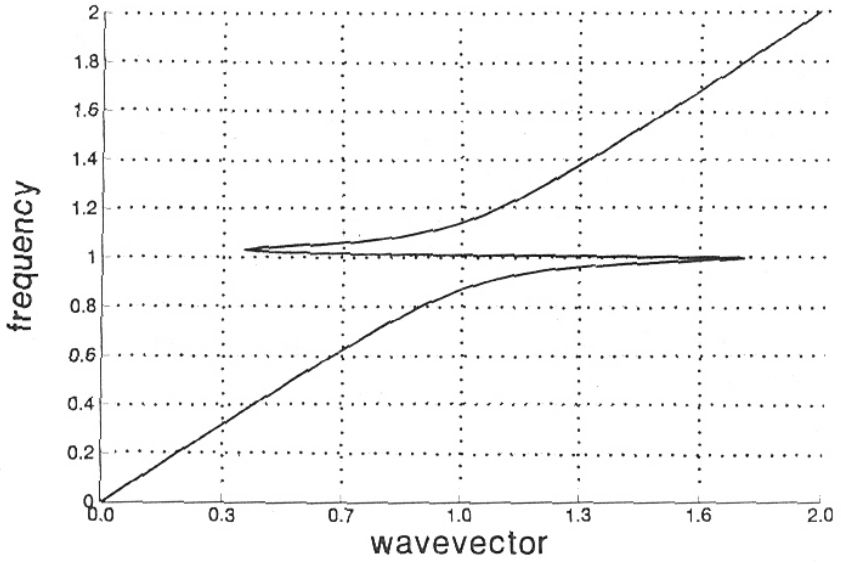

Figure 3. *Dispersion relation for polariton. The following parameters have been used:* $\Omega_0 = 1$, $S = 0.075$, $(1/T_2) = 0.01$. *The wavevectors are given in units of* Ω_0/c.

dielectric constants. Localization only occurs when real absorption is very small, and consequently it is not relevant for the present case of polaritons where all interaction is absorption rather than scattering, and the scattering mean free path is much longer than the extinction mean free path .

How good is our bulldozer approach (71), the independent scattering approximation, to the total T-matrix. It can be shown that here again $(k'\ell_{ext})^{-1}$ is the expansion parameter. So for strong polariton behavior we should improve our basic assumption and come up with a much better treatment of the T-matrix than our bulldozer ansatz. We call these effects dependent-scattering effects: the scattering properties of one center depend on the presence of other scatterers, an effect which is difficult to deal with. In principle it is not even possible anymore to *define* an effective dielectric constant as the T-matrix is going to be momentum dependent. We have performed some dependent-scattering corrections[36] for elastic scattering and large corrections at $k'\ell_{sca} \sim 1$ have been found. We also see the large advantage in relating polariton behavior to multiple scattering. In our equation of motion approach we never realized that there was a serious assumption being made. Where did we make our equivalent sledgehammer assumption in the equation of motion? This happened of course when we assumed the force to be proportional to the number of particles. Here the independent-scattering approximation has sneaked in. It is obvious that improving on our equation of motion is very difficult, whereas the t-matrix formalism is ideally suited for this.

If a transition is infrared active either through scattering or absorption one must be careful as the coupling is always strong and the propagation is always perturbed. Much of the misery will occur at the surface as the light propagation is seriously hampered. One way to circumvent this suffering is to try to get $\mathcal{P}$ as small as possible by using a low concentration of infrared active material in an otherwise inactive solvent or host. In that case equation (104) can be used. This still amounts to an exponentially attenuated

beam (Lambert-Beer law) $I \propto \mathrm{e}^{-L/\ell_{ext}}$, L being the length of the sample. The only way to recover the one-way response result (53) is to request in addition that $L << \ell_{ext}$. I hope that the reader by now has appreciated the limited value of one-way response theory.

It is quite clear that when the polariton parameter is large it is quite involved to obtain the relaxation parameters from experiment as the full polariton behavior has to be accounted for.

10 RAMAN-ACTIVE OSCILLATORS

Up to now we have assumed that we could indeed excite our oscillators directly with a single time-varying electric field. Our oscillator was infrared active. This has advantages and disadvantages. The drawback is that the transition must be allowed and secondly you must have light at the right frequency. For the majority of oscillators this means light in the infrared region. Fields in the visible are much easier to come by than in the infrared.

Many oscillators are not infrared active, but they are polarizable. That is they can scatter light (Raman and Rayleigh). The theory of spontaneous Raman scattering belongs to quantum scattering theory proper. If the scattering is dealt with in the, for this case of a weak scattering potential excellent, first Born approximation it is found that the scattered intensity $S(\mathbf{p}, E)$ is described by the polarizability-polarizability correlation function

$$S(\mathbf{p}, \Omega)\delta(\mathbf{p}_1 + \mathbf{p}_2) = \int \mathrm{e}^{i\Omega t} \ll \alpha(\mathbf{p}_1, t_1)\alpha(\mathbf{p}_2, t_1 + t) \gg dt\,, \tag{105}$$

The polarizability, which is a second-rank tensor, is assumed to be a function of the coordinate q_j of the j^{th} oscillator,

$$\alpha_j = \alpha(q_j = 0) + \left(\frac{\partial \alpha_j}{\partial q_j}\right)_{q_j=0} q_j + \cdots \quad . \tag{106}$$

From this Taylor expansion the contribution of the oscillators to the polarizability correlation function can be evaluated. Apparently the polarizability-polarizability correlation function can be related to $\langle Q(0)Q(t)\rangle$. If for the oscillators the relaxation-time approximation is valid, equation of motion (14), that is the equation of motion *without* external force, can be used to evaluate this correlation function. One finds two Lorentzian lines, one at the Stokes and one at the Anti-Stokes frequency, their frequency shift given by the eigenfrequency Ω_0 of the oscillator and their width by $1/T_2$. The temperature can be deduced from the Boltzmann factor determining the ratio of the two intensities. This property can be deduced from time-symmetry properties of the polarizability-polarizability correlation function.[5] The relaxation time and frequency can be obtained. I will discuss later the difference between the information obtained from this technique of spontaneous Raman scattering and the socalled "coherent Raman techniques".

In nonlinear Raman experiments the oscillators are *driven* by the Raman effect and that effect can be described classically. In this case the oscillator can be excited by two

frequencies. Let us look how this works.

The interaction of this oscillator with an external field $\mathbf{E}$ is given by

$$U = -\frac{1}{2}\mathbf{E}\cdot\alpha\cdot\mathbf{E} \quad . \tag{107}$$

Using expansion (106) in the above equation we find a term proportional to q_j. Completely analogous to the case of the infrared active oscillator this gives rise to the following equation of motion for a Raman-active oscillator

$$\frac{\mathrm{d}^2Q(t)}{\mathrm{d}t^2} + \frac{2}{T_2}\frac{\mathrm{d}Q(t)}{\mathrm{d}t} + \Omega_0^2 Q(t) = \left(\frac{N}{m}\right)\mathbf{E}\cdot\left(\frac{\partial\alpha}{\partial q}\right)_{q=0}\cdot\mathbf{E} \ . \tag{108}$$

So you now see this oscillator is not driven by one field, but by two fields. This makes many more effects possible. Suppose now that our incoming field has two frequency components: Ω_1 and Ω_2, $\Omega_1 > \Omega_2$. The field looks like

$$\mathbf{E} = Re(\mathbf{A}e^{i\Omega_1 t} + \mathbf{B}e^{i\Omega_2 t}). \tag{109}$$

The bilinear driving field in the equation of motion (108) is now a combination of two optical fields. This field has four possible frequencies, $\pm\Omega_1 \pm \Omega_2$. This indicates that the induced coordinate q has also these four contributions,

$$q(t) \propto Re\left(e^{(\pm\Omega_1\pm\Omega_2)t}\right). \tag{110}$$

From equation (106) we perceive that this implies that the polarizability has the same time components

$$\alpha(t) \propto Re\left(e^{(\pm\Omega_1\pm\Omega_2)t}\right). \tag{111}$$

The polarization is the polarizability times the electric field and has therefore the following components

$$\mathbf{P}(t) \propto Re\left(e^{(\pm\Omega_1\pm\Omega_2\pm\Omega_{1,2})t}\right), \tag{112}$$

which should be considered a χ_3 effect as the product of three electric fields, albeit not all at the same frequencies, is involved. In experiments with the two frequencies in the visible one always picks the sign combination that causes the polarizability to be forced to oscillate at the frequency $\pm \mid \Omega_1 - \Omega_2 \mid$. This combination allows for a resonant driving of the oscillator when this frequency difference fits the transition frequency. Two final possible sign combinations are very popular for the study of oscillators. One option for the signs implies that the frequency $2\Omega_1 - \Omega_2$ arises in the polarization. This is a polarization at a new frequency, and consequently gives rise to radiation at a new frequency. One could say that the experimentalist supplies the laser field and the Stokes field and the system returns the Anti-Stokes field. For this reason this techniques is called CARS, Coherent Anti Stokes Raman Scattering. This method requires phase matching as the wavevector and the frequency of the polarization are determined by the two incoming fields. This polarization field can generate propagating light only if wavevector and frequency are matched. This can be fulfilled for instance by choosing a particular angle between the two incoming beams. The other useful sign combination, besides the combination relevant for CARS, for the induced polarization density, returns an induced polarization at either Ω_1 or Ω_2. At those frequencies we already have our incoming lasers and consequently they are then amplified or attenuated due to the

driven nonlinear polarization. This is called stimulated Raman scattering and in the form when using two laser beams it is called Raman gain (and loss). Raman gain is self-phasematched. All these experiments can both be performed in the time domain and the frequency domain. In the frequency domain one of the two incoming laser beams will have its frequency scanned such that the difference fits the eigenfrequency of the oscillator. In the time-resolved version pump and probe pulses have to be applied. The probe pulse can have two pulsed incoming laser beams and the probe pulse one (CARS) or two (Raman Gain). If the bandwidth of one pulse is large enough to find the necessary combination $\Omega_1 - \Omega_2$ within the frequency content of one pulse, a single pump and a single probe pulse will be enough (impulsive Raman scattering[37]). The time-resolved CARS experiments will be discussed in detail in the lectures by Professor Bron.

What information can be obtained from these forced Raman spectroscopies? In as much as our equation of motion of the oscillator is valid, irrespective of the type of Raman technique or force being used force (like CARS, Raman Gain, Impulsive Scattering or else) we can measure both the frequency Ω_0 and the dephasing time T_2. Different techniques measure the *same* parameters in different ways. Analogously to the infrared active oscillators we could define a susceptibility associated with our driving polarization. This will be a χ_3 susceptibility. The fact that this is a fourth-rank tensor implies that its definition and tracing will request a lot of bookkeeping. The various techniques will measure different parts (real part, imaginary part or absolute value squared) of this nonlinear suceptibility, but their result pertains basically to the same information.

Analogously to the situation of infrared active oscillators there is no saturation and T_1 cannot be obtained. Of course I could again introduce the equation for the population dynamics and find the higher-order Raman susceptibilities. It is clear that the observation of forced Raman polarization beyond the lowest-order $\boldsymbol{\chi_3}$ would in principle give information on T_1. I will not discuss the extension with the population dynamics because the "Raman force" is quite weak and in practice in condensed matter it is very difficult to saturate transitions in this way. One practical reason is that there will always be some other degree of freedom present (possibly due to impurities) that will be excited directly by the light. This usually means a damage threshold that low that no sizeable Rabi frequencies can be induced through the Raman effect.

11 NEW INFORMATION FROM NONLINEAR OPTICS

Is there any advantage by going to a nonlinear technique rather than using a linear technique? To answer this problem we will have to make a distinction whether or not the expansion of the polarizability in terms of the normal coordinate is allowed (See for instance Eq. (106). Implicitly an adiabatic approximation is introduced when using this expansion. I will not discuss the situation when this adiabatic approach fails, as this condition is much more relevant for electronic degrees of freedom than for vibrational degrees of freedom. So we assume the expansion in the normal coordinate of the polarizability to be valid. The situation is then quite clear with respect to the difference between spontaneous and driven Raman techniques. CARS and other coherent Raman are techniques where an oscillator is *driven* by a combination of electric fields. In such a

technique one measures the balance between *gain* and *loss* and that is why the susceptibility describing this effect is a *commutator* of correlation functions. In spontaneous Raman experiments, which is a quantum-mechanical scattering technique, one measures gain and loss separately: in Raman language they are called Anti-Stokes and Stokes. So in the latter case one measures the two correlation functions corresponding to loss and gain separately. This makes it for instance possible to measure the temperature of one mode by comparing the intensity of the Stokes and Anti-Stokes signal. Consequently there is more information in the spontaneous-Raman experiments. As these experiments are often more easier to perform than the nonlinear Raman experiments it is quite clear what experiments have to be performed first, even if they are not so fancy as the nonlinear Raman experiments.

When linewidths become extremely narrow in the frequency domain it can become quite cumbersome to measure them accurately due to the limitations imposed by the instrumental resolving power. In that case experiments in the time domain become much simpler. Only the driven coherent Raman techniques can be performed in the time domain. This can be an important reason to opt for a nonlinear Raman technique.

As I indicated in the previous section the situation becomes different when the driving Raman fields are so strong that the response can be saturated. In that case the susceptibilities beyond third order have to be taken into account, and much more information can be obtained. There is no anologue of saturation in spontaneous Raman scattering.

12 ACKNOWLEDGEMENTS

I am very grateful to Huib Bakker, Marco Brugmans, Ron Kroon, Paul Planken, Rudolf Sprik, Mark Switser, and Bart van Tiggelen for discussions.

The work described in this paper is part of the research program of the "Stichting voor Fundamenteel Onderzoek van de Materie" (Foundation for Fundamental Research on Matter) and was made possible by financial supoport from the "Nederlandse Organisatie voor Wetenschappelijk Onderzoek" (Netherlands Organization for the Advancement of Research)

References

[1]C.P. Slichter."Principles of Magnetic Resonance," Harper, New York (1965) 1st edition.

[2]the far wings of a spectral line will never have a lorentzian shape as otherwise the moments would diverge

[3]S.W. Lovesey in "Dynamics of Solids and Liquids by Neutron Scattering,", edited by S.W. Lovesey and T. Springer, Springer, Berlin (1977).

[4]H. Mori, Progr. Theor. Phys. **33**, 423 (1965).

[5]D. Forster. "Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions," Benjamin, Reading (1975).

[6]P.A. Madden and R.M. Lynden-Bell, Chem. Phys. lett. **38** 163 (1976)

[7]A. Nitzan, S. Mukamel, and J. Jortner, J. Chem. Phys. **60**, 3929 (1974).

[8]A. Nitzan and J. Jortner, Mol. Phys. **25**, 713 (1973).

[9]A. Nitzan and R.J. Silbey, J. Chem. Phys. **60**, 4070 (1974).

[10]A. Nitzan, S. Mukamel, and J. Jortner, J. Chem. Phys. **63**, 200 (1975).

[11]A. Carrington and A.D. McLachlan "Introduction to Magnetic Resonance, " Harper, New York (1967).

[12]C.B. Harris, R.M. Shelby, and P.A. Cornelius, Phys. Rev. Lett. **38**, 1415 (1977).

[13]R. Kroon, R. Sprik, and A. Lagendijk, Phys. Rev. B **42**, 2785 (1990).

[14]A. Laubereau and W. Kaiser, Rev. Mod. Phys. **50**, 607 (1978).

[15]N.G. van Kampen, Phys. Rep. 24C, 171 (1976).

[16]N.G. van Kampen, "Stochastic Processes in Physics and Chemistry," North-Holland, Amsterdam (1981).

[17]H. Mori, Progr. Theor. Phys. **34**, 399 (1965); M. Dupuis, Progr. Theor. Phys. **37**, 502 (1967)

[18]R. Kubo and K. Tomita, J. Phys. Soc. Jpn. **9**, 888 (1954).

[19]W. Götze and K.H. Michel, in "Dynamical Properties of Solids," edited by G.H.Horton and A.A. Maradudin, North-Holland, Amsterdam (1974).

[20]an outline of such an approach has been given recently in H. Bakker, submitted to J. Chem. Phys.

[21]M. Born and E. Wolf. "Principles of Optics," Pergamon, Oxford 6th edition, 1987).

[22]R. Kubo, J. Phys. Soc. Jpn., **12**, 570 (1957).

[23]C.J. Joachain,"Quantum Collision Theory," North-Holland (1975).

[24]R. Landauer in "Analogies in Optics and Micro Electronics", edited by W. van Haeringen and D. Lenstra, Kluwer, Dordrecht (1990) p. 243.

[25]S.R. de Groot and L.G. Suttorp, "Foundations of Electrodynamics," North-Holland, Amsterdam (1972).

[26]R. Loudon, *The Quantum Theory of Light* (Clarendon, Oxford, 1973) first edition. (The second edition differs considerably from the first edition.)

[27]W.K.H. Panofsky and M. Phillips, "Classical Electricity and Magnetism," Addison-Wesley, Reading (1962); P.M. Morse and H. Feshbach,"Methods of Theoretical Physics," McGraw-Hill, New York (1953).

[28] "Nonclassical Effects in Quantum Optics,", edited by P. Meystre and D.F. Walls, American Institute of Physics, New York (1991).

[29]V.E. Kravtsov, V.I. Yudson, and V.M. Agranovich, Phys. Rev. **B 41**, 2794 (1990).

[30]A.E. Siegman, "Lasers," Oxford, Oxford (1986) .

[31]A. Laubereau, L. Greiter, and W. Kaiser, Appl. Phys. Lett. **25**, 87 (1974);H. Graener and A. Laubereau, Appl. Phys B **29**, 213 (1982).

[32]H.J. Bakker, P.C.M. Planken, and A. Lagendijk, Nature **347**, 745 (1990); H.J. Bakker, P.C.M. Planken, L. Kuipers, and A. Lagendijk, J. Chem Phys. **94**, 1730 (1991); H.J. Bakker, P.C.M. Planken and A. Lagendijk, J. Chem. Phys. **94**, 6007 (1991).

[33] "Scattering and Localization of Classical Waves in Random Media," edited by Ping Sheng, World Scientific, Singapore, (1990).

[34]Th.M. Nieuwenhuizen, A. Lagendijk and B.A. van Tiggelen. Phys. Lett. A. **169**, 191 (1992).

[35]A. Messiah, "Quantum Mechanics," Vols. I and II, North-Holland, Amsterdam (1961); H.C. van de Hulst, "Light Scattering by Small Particles,", Dover, New York (1981).

[36]B.A. van Tiggelen, A. Lagendijk, A. Tip, and G.F. Reiter, Eur. Phys. Lett. **15**, 535 (1991).

[37]Y. Yan and K.A. Nelson, J. Chem. Phys. **87** 6240 (1987).