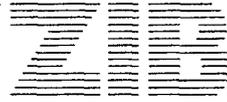


Konrad-Zuse-Zentrum für Informationstechnik Berlin



CHRISTOF SCHÜTTE MICHAEL WULKOW

Quantum Theory with Discrete Spectra
and
Countable Systems of Differential Equations

—

A Numerical Treatment of Infrared Spectroscopy

Quantum Theory with Discrete Spectra and Countable Systems of Differential Equations – A Numerical Treatment of Infrared Spectroscopy

CHRISTOF SCHÜTTE MICHAEL WULKOW

*Konrad-Zuse-Zentrum für Informationstechnik Berlin, Heilbronner Strasse 10,
D-1000 Berlin 31, Germany*

April 1992

ABSTRACT

Models for occupation dynamics in discrete quantum systems lead to large or even infinite systems of ordinary differential equations. Some new mathematical techniques, developed for the simulation of chemical processes, make a numerical solution of countable systems of ordinary differential equations possible. Both, a basic physical concept for the construction of such systems and the structure of the numerical tools for solving them are presented. These conceptual aspects are illustrated by a simulation of an occupation process from spectroscopy. In this example the structures of rotation spectra observed in infrared spectroscopy are explained and some possibilities for an extension of the model are shown.

CONTENTS

1	INTRODUCTION	1
2	RIGID ROTATORS IN AN ELECTROMAGNETIC FIELD	2
3	COUNTABLE SYSTEMS OF ORDINARY DIFFERENTIAL EQUATIONS	12
4	NUMERICAL RESULTS	18
	REFERENCES	23

1 INTRODUCTION

A large class of problems from chemistry or physics leads to countable systems of ordinary differential equations (CODEs). Examples are polymerization processes in chemistry, where single molecules are linked together to long polymer chains, or the occupation dynamic of the interaction between (infinite) many states in a quantum system. In both cases the mathematical models consists of a (infinite) sequence of single ordinary differential equations.

Experiences made in polymer chemistry ([12], [22], [23], [24]) have shown the solution of such CODEs to be a very difficult task. In this context theoretical and numerical tools were developed, which have allowed the efficient (numerical) solution of the considered problems ([7], [1], [2]).

This paper is meant as an encouragement for physicists to consider a countable system not only as an abstract description, but as an really tractable mathematical object. In particular we will demonstrate, that by a careful numerical treatment of CODEs interesting insights into physical processes can be gained.

A fundamental physical reason for the importance of countable systems for the modeling of microscopic processes can be indicated as follows. Let $\mathcal{H} = \mathcal{H}_0 + D$ be the Hamiltonian of physical systems, where \mathcal{H}_0 is the Hamiltonian of the free systems and D that of an external disturbing field. Hence \mathcal{H} , \mathcal{H}_0 , D are self-adjoint operators in a separable Hilbert space H with scalar product $\langle \cdot, \cdot \rangle$. Consider the spectrum of \mathcal{H} to be discrete with eigenvalues $E_n, n \in N$. Let ϕ_n be the eigenvector to E_n , so that $(\phi_n)_{n \in N}$ is an orthonormal basis in H . Besides let

$$D_{l,n} = \langle \phi_l, D\phi_n \rangle$$

be the matrix elements of D .

The evolution of an ensemble of the described systems is given by Schrödinger's equation

$$(1.1) \quad \mathcal{H}\psi(t) = i\hbar \partial_t \psi(t), \quad \psi(t) \in H \quad \forall t \in R$$

with $\psi(0) = \psi_0$.

If we expand $\psi(t)$ in the orthonormal basis given by

$$\Phi_n := e^{-i\frac{E_n}{\hbar}t} \phi_n$$

as

$$\psi(t) = \sum_{n \in N} u_n(t) \Phi_n$$

we can insert this into (1.1) and take the scalar product with (Φ_l) at both sides of the resulting equation. The result is an initial value problem of the form:

$$\begin{aligned}\partial_t u_l &= -\frac{i}{\hbar} \sum_{n \in N} D_{l,n} u_n, \quad l \in N \\ u_l(0) &= \langle \Phi_l, \psi(0) \rangle\end{aligned}$$

which is a countable system of (complex) differential equations, where the disturbance D gives the coupling of the single states $l \in N$. This scheme emerges in all quantum theoretical problems in the presence of discrete spectra. In Section 2 we will take explicitly this to discuss the interaction between rigid rotator and an electromagnetic field. On this basis we can explain the emission spectra observed in *Infrared spectroscopy* of rigid rotator. This frequent appearance of CODEs produces the question how to treat them. Looking at countable systems as at finite systems of ordinary differential equations (ODEs) first the difficulty of transferring the well known theoretical background like uniqueness theorems or numerical treatment to CODEs can be seen. Moreover it turns out that this view also isn't an appropriate description for realistic examples. Therefore in Section 3 we will consider a CODE as an evolution problem in a sequence space, which motivates numerical approximations of its solution and discuss several numerical methods for CODEs. Finally in Section 4 some interesting numerical results for the rotator example are presented.

2 RIGID ROTATORS IN AN ELECTROMAGNETIC FIELD

Infrared spectroscopy is one of the most useful experimental tools for getting information about the structure of molecules measuring their emission spectrum after stimulation with e.g. infrared light (for a short introduction see [19]). As a simple example let us take a linear diatomic molecule with permanent dipole moment like CO . Observing the infrared part of the emission spectrum of a CO molecule crystal we find two characteristic groups of equidistant peaks (called branches) like those shown in figure (2.1) (from [4] p.82).

They are called *rotation spectra* and can be identified as the results of transitions between neighboring energetic eigenstates of the molecule's rotation and vibration: Near each eigenstate v of vibration there are grouped all the eigenstates n_v of rotation. The energy gaps between the states of rotation are much smaller as those between the ones of vibration. The transitions take place between the rotational states of the neighboring vibration states and effect the two groups of peaks on the left and the right of the point O in the center of the spectrum (see figure (2.2)).

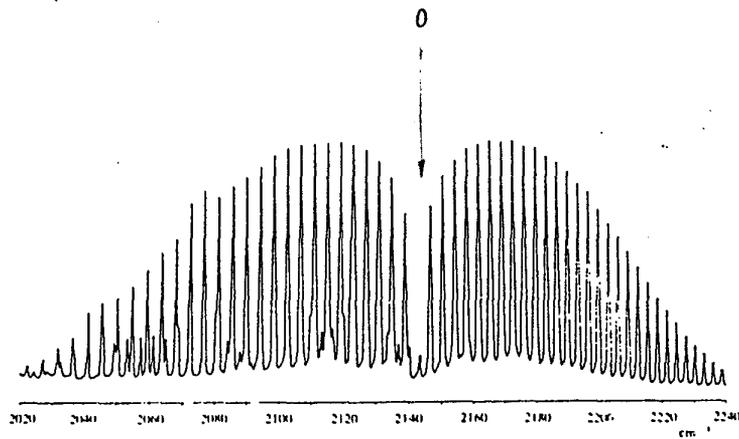
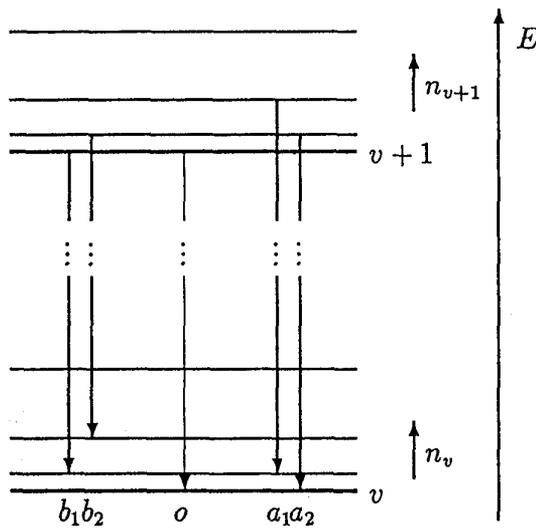


Figure 2.1: Example: Part of the infrared spectrum of CO



Transitions a_1, a_2 and b_1, b_2 take place, o not. The energetic distances are
 $\Delta E_{a_1} - \Delta E_o = \Delta E_o - \Delta E_{b_1}$.
 o would lead to a peak at O in the spectra.
 $\delta := \Delta E_{a_2} - \Delta E_{a_1} = \Delta E_{b_1} - \Delta E_{b_2}$
 is the same in both cases

Figure 2.2: Examples for transitions leading to vibration-rotation-spectra

We are interested only in a description of the structure of pure rotational spectra, which outlines are exactly the same as that of one of the branches in figure (2.1). Therefore we neglect the vibrations and regard the internal transitions in one rotation state system n_v .

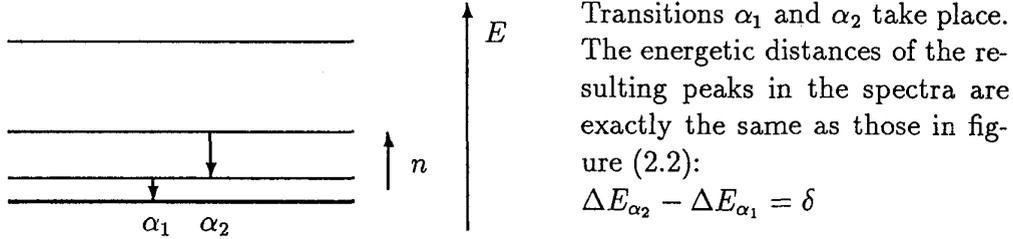


Figure 2.3: Examples for internal transitions in a single rotational state system

Hence we will deal with a free (rigid) rotator without vibrations. If we take the Hamiltonian of such a free *rigid rotator*

$$(2.1) \quad \mathcal{H}_R = \frac{1}{2I} L^2$$

(where I stands for the momentum of inertia of the molecule and L for the angular momentum operator), we obtain the energetic eigenvalues to be

$$(2.2) \quad E_n = \frac{\hbar^2}{2I} n(n+1), \quad n \in N_0$$

where E_n is $2n + 1$ times multiple. Therefore the energetic distance between the eigenstates is

$$\Delta E_n := E_{n+1} - E_n = \frac{\hbar^2}{I}(n+1)$$

Hence the distance between the resulting peaks in the spectra is constant ($\Delta E_{n+1} - \Delta E_n = \frac{\hbar^2}{I}$) and gives the momentum of inertia of the molecule. The specific outline of the intensities of the peaks shown in figure (2.1) is characteristic for all rotation spectra in Infrared spectroscopy. In this section a quantum field theoretical analysis of the dynamics of the interaction between the eigenstates of a rigid rotator will be presented. We have tried to make the following theory "structurally readable", because we want to show the generality of the concept. For physicists there may be some well known parts. Based on this in Section 4 we're able to explain the outlines of rotation spectra solving a CODE .

The Hamiltonian of the interaction: To describe the interaction between a system consisting of charged particle (with charge e) in a given (molecular) potential and an external electromagnetic wave with vector potential \mathbf{A} we have to use quantum field theoretical tools. The free system is characterized by his Hamiltonian

$$\mathcal{H}_{free} = \frac{1}{2m} \mathbf{p}^2 + V(x)$$

with scalar potential V and the impulse operator \mathbf{p} and supposed discrete spectrum. Being in interaction with the electromagnetic field the system must be described using a fermion field operator representing the charge density. The classical theory of charges in electromagnetic fields teaches us how to construct the Hamiltonian of the whole system using the free Hamiltonian, the electromagnetic field operator \mathbf{A} and fermion operator Ψ :

$$(2.3) \quad \mathcal{H} = \int \Psi^\dagger(x) \left(\frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 + V(x) \right) \Psi(x) dx + \mathcal{H}_L$$

where \mathcal{H}_L is the Hamiltonian of a free electromagnetic field after quantization. Now we have to make all the common steps (e.g. see [16]) towards a creation and annihilation operator calculus:

- make Coulomb gauge and cancel the terms quadratic in \mathbf{A} , so that \mathcal{H} looks like

$$\begin{aligned} \mathcal{H} &= \int \Psi^\dagger(x) \left(\frac{1}{2m} \mathbf{p}^2 + V(x) \right) \Psi(x) dx + \mathcal{H}_L \\ &\quad + \int \Psi^\dagger(x) \left(-\frac{e}{m} \mathbf{A} \mathbf{p} \right) \Psi(x) dx \end{aligned}$$

- quantization of the vector potential \mathbf{A} after expansion in plane waves

$$\mathbf{A} = \sum_{k,j} \sqrt{\frac{2\pi\hbar}{\mathcal{V}\omega_k}} \left(e_{k,j} e^{ikx} \mathbf{b}_{k,j} + e_{k,j} e^{-ikx} \mathbf{b}_{k,j}^\dagger \right)$$

where k denotes states in the first Brillouin zone, ω_k are the frequencies of the single plane waves, $e_{k,j}$ the two directions of polarization, $\mathbf{b}_{k,j}^\dagger$ and $\mathbf{b}_{k,j}$ the creation and annihilation operators, and $\mathcal{V} < \infty$ is the used field volume. After this step \mathcal{H}_L looks

$$\mathcal{H}_L = \sum_{k,j} \hbar\omega_k \mathbf{b}_{k,j}^\dagger \mathbf{b}_{k,j}$$

- expanding Ψ in the orthonormal basis $(\phi_n)_{n \in N}$ of eigenvectors to \mathcal{H}_{free}

$$\begin{aligned}\Psi(x) &= \sum_{n \in N} \phi_n(x) \mathbf{a}_n \\ \Psi^+(x) &= \sum_{n \in N} \phi_n^*(x) \mathbf{a}_n^+\end{aligned}$$

- defining the commutators between \mathbf{A} , its canonical conjugated impulse $\mathbf{\Pi}$ and Ψ for getting the commutator rules of the creation and annihilation operators where all $\mathbf{a}_n, \mathbf{a}_n^+, \mathbf{b}_{k,j}, \mathbf{b}_{k,j}^+$ commutes with exception of:

$$[\mathbf{a}_n, \mathbf{a}_l^+]_+ = \mathbf{a}_n \mathbf{a}_l^+ + \mathbf{a}_l^+ \mathbf{a}_n = \delta_{n,l} \quad \text{and} \quad [\mathbf{b}_{k,j}, \mathbf{b}_{k',j'}^+]_- = \delta_{(k,j),(k',j')}$$

Finally we're able to transform \mathcal{H} (for details see [16], p. 294-298) into

$$\begin{aligned}\mathcal{H} &= \sum_n E_n \mathbf{a}_n \mathbf{a}_n^+ + \sum_{k,j} \hbar \omega_k \mathbf{b}_{k,j}^+ \mathbf{b}_{k,j} \\ &\quad + \hbar \sum_{n,l,k,j} (g_{n,l,k,j} \mathbf{a}_n^+ \mathbf{a}_l \mathbf{b}_{k,j} + g_{n,l,-k,j} \mathbf{a}_n^+ \mathbf{a}_l \mathbf{b}_{k,j}^+)\end{aligned}$$

with the *g-coefficients*

$$(2.4) \quad g_{n,l,k,j} = -\frac{e}{m} \sqrt{\frac{2\pi}{V \hbar \omega_k}} \int \phi_n^*(x) e_{k,j} e^{ikx} \mathbf{p} \phi_l(x) dx$$

The g-coefficients as matrix elements of the dipole operator: Our free systems are single molecules (length parameter d), the electromagnetic frequencies of our interest are infrared ones with k -values fulfilling $kd \ll 1$. Hence in that space region where $\Phi_n(x)$ doesn't vanish it is $e^{ikx} \approx 1$. These circumstances allow us to write approximately

$$\int \phi_n^*(x) e^{ikx} \mathbf{p} \phi_l(x) dx = \int \phi_n^*(x) \mathbf{p} \phi_l(x) dx = \langle \phi_n, \mathbf{p} \phi_l \rangle$$

Besides the basic rule $[\mathbf{p}, x]_- = \frac{\hbar}{i} 1$ gives us the commutator of \mathcal{H}_{free} and the position operator x as

$$p = i \frac{m}{\hbar} [\mathcal{H}_{free}, x]_- .$$

Using $\mathcal{H}_{free} \phi_n = E_n \phi_n$ we get with 2.4

$$(2.5) \quad g_{n,l,k,j} = -i \sqrt{\frac{2\pi e^2}{V \hbar \omega_k}} \frac{E_l - E_n}{\hbar} \langle \phi_n, x \phi_l \rangle e_{k,j} .$$

Disturbation theory of first order. As in Section 1 we can understand \mathcal{H} as consisting of the Hamiltonian

$$\mathcal{H}_0 = \sum_n E_n \mathbf{a}_n \mathbf{a}_n^\dagger + \sum_{k,j} \hbar \omega_k \mathbf{b}_{k,j}^\dagger \mathbf{b}_{k,j}$$

of the free (interactionless) systems and the disturbance (interaction) operator

$$D = \hbar \sum_{n,l,k,j} (g_{n,l,k,j} \mathbf{a}_n^\dagger \mathbf{a}_l \mathbf{b}_{k,j} + g_{n,l,-k,j} \mathbf{a}_n^\dagger \mathbf{a}_l \mathbf{b}_{k,j}^\dagger)$$

The theory of creation operators gives us the eigenstates to \mathcal{H}_0

$$\Phi_{n,(k,j)^M} := \frac{1}{\sqrt{M!}} \mathbf{a}_n^\dagger (\mathbf{b}_{k,j}^\dagger)^M \Phi_0$$

using the so called vacuum state Φ_0 , which fulfills the rules

$$\mathbf{a}_n \Phi_0 = \mathbf{b}_{k,j} \Phi_0 = 0.$$

We know how to interpret the $\Phi_{n,(k,j)^M}$ as those states of the ensemble, where each single system is in free-system-state n with M photons of frequency ω_k and polarization j .

The evolution of an ensemble of such systems is given by *Schrödinger's equation* (1.1) which can be transformed by using the evolution operator

$$U = e^{-\frac{i}{\hbar} \mathcal{H}_0 t}$$

for going into the *interaction image*

$$\tilde{\psi} := U^{-1} \psi \quad \text{and} \quad \tilde{D} = U^{-1} D U$$

Now Schrödinger's equation looks

$$\partial_t \tilde{\psi}(t) = \frac{1}{i\hbar} \tilde{D}(t) \tilde{\psi}(t) \quad \text{with} \quad \tilde{\psi}(0) = \psi_0$$

with

$$\begin{aligned} \tilde{D} = \hbar \sum_{n,l,k,j} & \left(g_{n,l,k,j} e^{\frac{i}{\hbar}(E_n - E_l - \hbar\omega_k)t} \mathbf{a}_n^\dagger \mathbf{a}_l \mathbf{b}_{k,j} \right. \\ & \left. + g_{n,l,-k,j} e^{\frac{i}{\hbar}(E_n - E_l + \hbar\omega_k)t} \mathbf{a}_n^\dagger \mathbf{a}_l \mathbf{b}_{k,j}^\dagger \right). \end{aligned}$$

Now we are interested in that cases where transitions from a state $\tilde{\psi}(0) = \Phi_{n,(k,j)^M}$ to another state $\Phi_{l,(k,j)^P}$ take place. The relative number of systems in the ensemble which have done this transition after time t is proportional to

$$\left| \langle \Phi_{l,(k,j)^P}, \tilde{\psi}(t) \rangle \right|^2 = \frac{1}{\hbar^2} \left| \langle \Phi_{l,(k,j)^P}, \int_0^t \tilde{D}(\tau) \tilde{\psi}(\tau) d\tau \rangle \right|^2$$

In cases where this is non-vanishing we have to deal with \tilde{D} to compute the transition probabilities $\frac{d}{dt} |\langle \Phi_{l,(k,j)^P}, \tilde{\psi}(t) \rangle|^2$. This task is exactly equal-structured to the well-known problem of the interaction between electrons and phonons as discussed e.g. in [16] p. 209–217. By means of disturbance theory of first order it can be shown, that in this approximation three kinds of transitions take place:

1. *Spontaneous emission of a photon:* The first part of \tilde{D} effects

$$\mathbf{a}_n^\dagger \mathbf{a}_l \mathbf{b}_{k,j}^\dagger \Phi_{l,(k,j)^0} = \Phi_{n,(k,j)^1}$$

which leads to a real transition process from $\Phi_{l,(k,j)^0}$ to $\Phi_{n,(k,j)^1}$ if the energy conservation

$$\hbar\omega_k = E_l - E_n$$

holds. This is a consequence of the appearance of $e^{\frac{i}{\hbar}(E_n - E_l + \hbar\omega_k)t}$ in \tilde{D} (for details see [16]). Its mean transition probability per unit time is

$$(2.6) \quad \Gamma_{(k,j)^0 \rightarrow (k,j)^1}^{l \rightarrow n} = |g_{l,n,k,j}|^2$$

2. *Induced emission of a photon:* The first part of \tilde{D} effects

$$\mathbf{a}_n^\dagger \mathbf{a}_l \mathbf{b}_{k,j}^\dagger \Phi_{l,(k,j)^M} = \sqrt{M+1} \Phi_{n,(k,j)^{M+1}}$$

describing a real transition if already the energy conservation rule is fulfilled. Its mean transition probability per unit time is

$$(2.7) \quad \Gamma_{(k,j)^M \rightarrow (k,j)^{M+1}}^{l \rightarrow n} = (M+1) |g_{l,n,k,j}|^2$$

3. *Induced absorption of a photon:* This time the second part of \tilde{D} is responsible (notice the commutator rules):

$$\mathbf{a}_n^\dagger \mathbf{a}_l \mathbf{b}_{k,j} \Phi_{l,(k,j)^M} = \sqrt{M} \Phi_{n,(k,j)^{M-1}}$$

describing a real transition iff

$$\hbar\omega_k = E_n - E_l$$

is fulfilled. Its mean transition probability per unit is given by

$$(2.8) \quad \Gamma_{(k,j)^M \rightarrow (k,j)^{M-1}}^{l \rightarrow n} = M |g_{l,n,k,j}|^2$$

Drawing the consequences of the last pages we have found that all systems, which interact with light and fulfill (2.3) and all other rather technical assumptions, only show the named first order transitions. The mean probabilities of these transitions can be calculated using (2.5) for the g-coefficients. These results are valid for a large class of problems, in particular for most of the spectroscopic ones.

Transition probabilities in the rotator example: Going back to the interaction of free rotators in an external electromagnetic field we find the particular form of \mathcal{H}_{free} as

$$\mathcal{H}_{free} = \mathcal{H}_R = \frac{1}{2I} L^2$$

because we can limit our description to the problem of the rotation of the electron density effecting the dipole. Thus we neglect the influences of the electromagnetic field in all other properties of the molecules, in particular those in their electronic structure. We know the eigenstates of \mathcal{H}_{free} from (2.2) and the eigenvectors $\phi_{n,m}$ as those from L^2 . The $\phi_{n,m}$, $m \in \{-n, \dots, n\}$ play the role of the ϕ_n above, the double index reflects the multiplicity of E_n .

Besides we can make the assumption that for all single rigid rotators the polarization of the used electromagnetic field is parallel to the rotation plane (e.g. in a CO molecule crystal). The task of calculating the matrix elements of the vertical component x_{\perp} of the dipole operator has been solved for us in [20] p. 697. The element $\langle \phi_{n,m} | x_{\perp} | \phi_{n',m'} \rangle$ does *not* vanish only if

$$(2.9) \quad |n-n'| = 1 \quad \wedge \quad m = m'$$

so that only following elements are of interest:

$$(2.10) \quad \langle \phi_{n,m} | x_{\perp} | \phi_{n-1,m'} \rangle = \sqrt{\frac{n^2 - m^2}{4n^2 - 1}}$$

Moreover, because of (2.9) we can combine the $(2n+1)$ states (n,m) to one state n . This allows us to compute the transition probabilities $\Gamma_{(k,j)^M \rightarrow (k,j)^{M\pm 1}}^{n \rightarrow n\pm 1}$ using equations (2.5), (2.6), (2.7), (2.8) and in addition the energy conservation rules. Only the following probabilities are nonvanishing:

$$\begin{aligned} \Gamma_{(k,j)^0 \rightarrow (k,j)^1}^{n \rightarrow n-1} &= \alpha \Delta E_{n-1} \sum_{m=-n}^n \frac{n^2 - m^2}{4n^2 - 1} \\ \Gamma_{(k,j)^{M_{n-1}} \rightarrow (k,j)^{M_{n-1}+1}}^{n \rightarrow n-1} &= \alpha \Delta E_{n-1} (M_{n-1} + 1) \sum_{m=-n}^n \frac{n^2 - m^2}{4n^2 - 1} \\ \Gamma_{(k,j)^{M_n} \rightarrow (k,j)^{M_n-1}}^{n \rightarrow n+1} &= \alpha \Delta E_n M_n \sum_{m=-(n+1)}^{n+1} \frac{(n+1)^2 - m^2}{4(n+1)^2 - 1} \end{aligned}$$

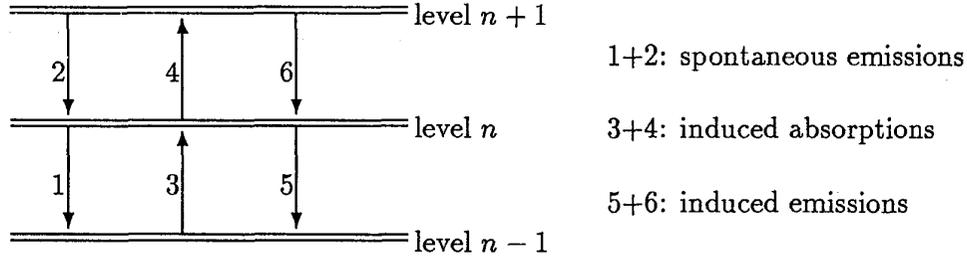
with a constant α which is independent from n, M_n and the photon numbers M_n for photons of frequencies $\omega_n = \frac{1}{\hbar} \Delta E_n$. Together with (2.2) the summation gives:

$$(2.11) \quad \Gamma_{(k,j)^0 \rightarrow (k,j)^1}^{n \rightarrow n-1} = \beta n^2$$

$$(2.12) \quad \Gamma_{(k,j)^{M_{n-1}} \rightarrow (k,j)^{M_{n-1}+1}}^{n \rightarrow n-1} = \beta (M_{n-1} + 1) n^2$$

$$(2.13) \quad \Gamma_{(k,j)^{M_n} \rightarrow (k,j)^{M_n-1}}^{n \rightarrow n+1} = \beta M_n (n + 1)^2$$

The occupation dynamics model: Now we have computed the probabilities of the mean transition processes. We're leaving now the level of quantum field theory and will build a quasi classical occupation dynamical model. Let $u_n(t)$, $n \in N$ be the occupation of state n at time t , that means the number of rigid rotators in a crystal, which are in the state n at time t . Then, the following figure can help us to write equations for the dynamics of the occupation process, which will be caused by the transitions between the states induced by an external light beam consisting the "right" frequencies ω_n .



In spectroscopic experiments the electromagnetic field is that of the external light beam. Therefore the numbers M_n of ω_n -photons is a constant of the process given by the spectral intensities of the external beam. This implies that only the induced transitions appear while the spontaneous ones don't take place. Under all these assumptions the evolution of the occupation can be given by

$$(2.14) \quad \partial_t u_0 = \beta (-M_0 u_0 + (M_0 + 1) u_1)$$

$$(2.15) \quad \partial_t u_n = \beta \left(-(n+1)^2 M_n u_n + n^2 M_{n-1} u_{n-1} + (n+1)^2 (M_n + 1) u_{n+1} - n^2 (M_{n-1} + 1) u_n \right)$$

Planing a simulation of the occupation dynamic we still have to make realistic assumptions for the "intensities" M_n . For reasons of simplification we choose

$M_n = M$ for all n , that means we suppose a constant infrared spectrum. This choice is made only because no other intrudes itself. There are no necessities in the numerical treatment which would force us to do so. Thus our final description of the occupation dynamic of Infrared spectroscopy of rotation spectra sounds

$$(2.16) \quad \partial_t u_0 = \beta (-M u_0 + (M+1) u_1)$$

$$(2.17) \quad \partial_t u_n = \beta \left(n^2 M u_{n-1} - ((n+1)^2 M + n^2 (M+1)) u_n + (n+1)^2 (M+1) u_{n+1} \right)$$

It is very important to see (2.16) fulfilling the conservation law

$$\partial_t \left(\sum_{n=0}^{\infty} u_n(t) \right) = 0$$

which in this attempt have to be interpreted as the conservation of the number of systems.

Later in Section 4 we want to test the solution of (2.16) comparing it with experimental results like those in figure (2.1). Therefore we have to compute the evolution of the intensities $I(\omega_n)$ in the emission spectrum of the process. As the frequency $\omega_n = \frac{1}{\hbar} \Delta E_n$ is connected with the emission $(n+1) \rightarrow n$ its intensity at time t is given by the number of such transitions at that time. Hence

$$(2.18) \quad I(\omega_n)(t) = \gamma (n+1)^2 (M+1) u_{n+1}, \quad n \in N_0$$

with a constant factor γ .

3 COUNTABLE SYSTEMS OF ORDINARY DIFFERENTIAL EQUATIONS

The general form of a (scalar, nonlinear) initial value problem for a countable system can be written as

$$(3.1) \quad u'_n(t) = f_n(u_1(t), u_2(t), \dots) \quad , \quad n = 1, 2, \dots, n_{\max} .$$

The whole vector (sequence, distribution) of the $u_n(t)$ will also be called $u(t)$. The right-hand side functions $f_n(u) = f_n(u_1, u_2, \dots)$ (time dependency omitted) can arise from problems in quite different fields (physics, astrophysics, chemistry, biology). We will call a system *hierarchical*, if for infinite many n the functional f_n depends on components $u_{\bar{n}}$ with $\bar{n} > n$. Whereas in the present context the variables $u_n(t)$ denote the occupation of states, e.g. in polymer chemistry $u_n(t)$ describes the concentrations of macromolecules with chain-length n at time t . In contrary to an ordinary differential equation, the upper index n_{\max} (i.e. the dimension of the system) is usually very large ($10^3 - 10^6$ in practical examples) or even *infinite*. If for computational reasons an infinite system has to be truncated at chain-length n_{\max} , a suitable value is rarely known a priori and may vary with time t . Thus a system (3.1) cannot be treated by standard numerical methods for ordinary differential equations in general. The theory of countable systems shows, that a CODE is quite different from a large scale system of ordinary differential equations (ODE). The techniques for proving existence and uniqueness of solutions - in particular the choice of appropriate spaces - resembles more to the theoretical treatment of partial differential equations PDEs. A survey on this field is given in [9] and [24]. In our special case, we will not discuss the theory, but concentrate on the numerical treatment of countable systems. We only want to mention, that the solution of (2.16) exists in certain weighted l^2 -spaces.

Numerical treatment of countable systems. In the following we discuss advantages and disadvantages of certain numerical methods for CODE'S.

- *Direct integration of a truncated system.* If a countable system is truncated at dimension n_{\max} , the solution of the resulting finite dimensional system depends on this index in general ($u_n(t) \rightarrow u_n^{n_{\max}}(t)$). Under certain conditions (which are fulfilled in the present example, see [9]) for fixed index \bar{n} the convergence

$$\lim_{n_{\max} \rightarrow \infty} u_{\bar{n}}^{n_{\max}}(t) = u_{\bar{n}}(t)$$

is ensured. This means, that in principle for a given final time t_{end} indices \bar{n} and n_{max} can be given, such that $|u_n(t) - u_n^{n_{max}}(t)|$ is small enough for $n \leq \bar{n} < n_{max}$ and an interesting range of indices is considered (e.g. in Section 3, where for $M = 10$ the index $n_{max} = 150$ is chosen in view of the process and the given initial distribution). The problem is, that n_{max} is not known a priori in general cases, and many attempts with different choices have to be made. If the chosen n_{max} is too small, essential properties of the infinite system can be destroyed (e.g. conservation variables, stationary states, see Figure 4.3). In any case, the resulting n_{max} -dimensional system may be very large and stiff, such that the numerical integration can be very time-consuming or even impossible in practice.

- *Moments, lumping, statistical methods, model approximations.* There are several numerical methods for CODEs, which are developed for special applications.
 - *Computation of moments.* Whenever one is not interesting in the distribution u_n itself, but in characterizing variables as e.g. mean values, one can try to compute the statistical moments of u_n . For time-dependent problems the derivation of differential equations of the moments is necessary, which can only be done, if the systems are not hierarchical. It is in general not possible to reveal u_n from a finite number of moments.
 - *Lumping.* A number of components u_n is summarized for a super-species in this approach. A lumping is only possible for certain linear problems and if the behavior of the system is known in principle.
 - *Statistical methods.* In contrary to our example above, one can avoid the coarsening step to the quasi-macroscopic level of description on which CODEs emerge. Simulations dealing with the real microscopic (and therefore often statistical) description of the processes are called *statistical*. They are extremely time consuming and in most cases really impossible to carry out.
 - *Model approximation.* Such approximations like the *quasi-steady-state-approximation* in chemistry are only valid with good insight into the problem to be studied. If a system leaves the region of such a model assumption, this behavior cannot be detected.

We must say that none of the above methods can serve as a comprehensive approach to countable systems.

– *Approximation in weighted sequence spaces (Discrete Galerkin method).*

A more promising approach is based on a theory of CODEs formulated in certain weighted sequence spaces, which are motivated by many applications. The general idea of a discrete Galerkin method, first presented in [12], later specified in several versions in [1] and [24], is to expand the solution $u_n(t)$ of a CODE into a series of certain special functions of a discrete variable. This expansion is truncated after $k + 1$ coefficients leading to a *Galerkin approximation* $u_n^k(t)$

$$(3.2) \quad u_n(t) = \Psi(n) \sum_{i=0}^{\infty} a_i(t) l_i(n) \longrightarrow u_n^k(t) = \Psi(n) \sum_{i=0}^k a_i(t) l_i(n) ,$$

with a weight function Ψ and polynomials $l_i(n)$. The polynomials $l_i(n)$ are connected with the weight function by the *orthogonality* relation

$$(3.3) \quad \sum_{n=1}^{\infty} l_i(n) l_j(n) \Psi(n) = \begin{cases} \gamma_i , & \text{for } i = j \\ 0 , & \text{for } i \neq j \end{cases}$$

(for details see [12] or [24]). The above approach is very efficient, whenever there are certain similarities between the u_n and the weight function Ψ (for a discussion see [1]). Therefore weight functions are chosen, which are known to be solutions of well-known processes in chemistry (geometric, Poisson, etc.).

In order to obtain the expansion coefficients $a_i(t)$, there are two principle ways. In a *method of lines* the expansion (3.2) is inserted into the CODE. Then analytical properties of the polynomials and their orthogonality with respect to Ψ are employed to derive differential equations for the $a_i(t)$. Such an approach is implemented in the program package MACRON, which is written for problems from polymer chemistry.

An improvement compared to the method-of-lines approach is a time discretization by means of a *Rothe's method*. This technique was introduced in [5] for parabolic differential equations. We write a countable system in a closed form

$$u'(t) = A u(t) , \quad u(0) \text{ given .}$$

The idea is to discretize this equation as if it was an ODE. In order to obtain an approximation $u^1 = u_{t+\tau}$ of $u(t + \tau)$ in a time step of size τ , we apply the *implicit Euler* scheme :

$$(3.4) \quad (I - \tau A) u^1 = \varphi ,$$

with $\varphi = u(t)$. This *stationary* subproblem is also very large or infinite. It must be solved by the discrete Galerkin method within a certain accuracy.

Then the task is to get an estimation η^1 of the time error $\|u^1 - u(t + \tau)\|$ ($\|\cdot\|$ an appropriate norm) for predicting a new reasonable step-size $\bar{\tau}$. This is done by solving a *correction equation* with the same left-hand side as in (3.4). Finally we mention, that when a time step with size τ has been performed, a new step size $\bar{\tau}$ can be computed by

$$\bar{\tau} = \tau \sqrt{\frac{\text{tol}}{\|\eta^1\|}},$$

where tol is the required tolerance for the solution $u(t)$. Details of this kind of a Rothe's method can be found in [6], a (complete adaptive) algorithm for CODEs based on this theory is implemented in the program CODEX [24].

Discrete h-p-method. In this paper, we suggest a new development, which extends and generalizes the idea of a discrete Galerkin method. The methods described above have got the drawback, that due to the *global* character of the expansion the approach may become inefficient whenever the solution u_n cannot be approximated well by the chosen class of weight functions. In particular this can be observed for distributions u_n with jumps, discontinuities or some local extrema, which require a *local adaptation*. On the other hand, a purely *local* approach as a discrete analogue to *finite elements* is not the appropriate choice. A careful analysis of interesting CODE-problems shows, that the amount of work for a Galerkin approximation is described in terms of the degrees of freedom used. This number would be too large, if the standard linear or quadratic finite elements were applied. Consequently we suggest a mixture of local and global approaches in the literature on the numerical solution of PDEs known as *h-p-method* [3]. We adapt the principle idea and add the typical details for a discrete treatment in the context of CODEs. By a multi-level algorithm we

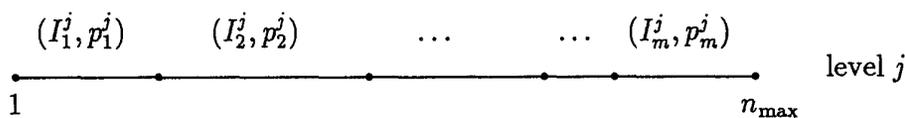


Figure 3.1: General *h-p*-pattern.

try to construct a sub-division of the n -axis as shown in Figure 3.1 and use

on each interval I_l a (local) expansion with special functions up to an order p_l . This number of expansion coefficients may differ from interval to interval. The resulting node-order-distribution

$$\Delta = \{(I_1, p_1), \dots, (I_m, p_m)\}$$

has to be chosen such that the amount of work to compute the whole approximation is as small as possible. The construction is usually started with one initial interval $I_1 = \{1, \dots, n_{\max}\}$, n_{\max} finite, but possibly very large, and proceeds from level to level, using information from the previous levels. On each sub-interval I_l^j we have an approximation u_n^j of u_n by

$$(3.5) \quad u_n^j|_{I_l^j} = \sum_{i=0}^{p_l^j} a_{i,l}^j t_{i,l}^j(n)$$

where the polynomials $t_{i,l}^j$ are the well-known *discrete Chebyshev polynomials* [8], [21], which are orthogonal on I_l^j with respect to the weight function $\Psi \equiv 1$. An example for two successive levels is shown in Figure 3.2. Whereas on the

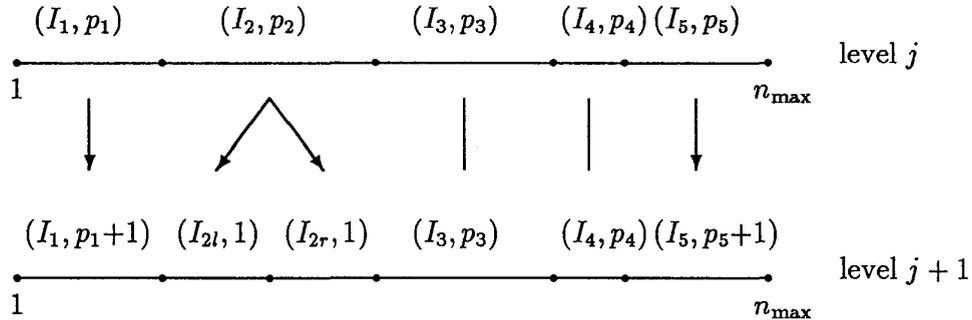


Figure 3.2: Example for a transition in a h - p -multilevel method.

intervals I_1 and I_5 the order of the (local) approximation is increased by one, I_2 is bisected and the orders are set to one. The remaining intervals are not changed. The above situation arises from cases as in Figure 4.4 in Section 4, where the solution isn't "smooth" for all n – smooth with respect to the possibility of an efficient polynomial expansion.

If worked out this idea requires the treatment of a lot of details, which cannot

be explained here. Thus we restrict to the main modules of the h - p -algorithm for the solution of a problem (3.4). Assume that a situation as in Figure 3.2, level j , is given and that there are also error estimates for the local expansions on the I_l . These information (possibly also from level $n - 1$) must be used, to predict a new grid and perform the Galerkin method on it.

Sketchy algorithm.

1. *Decide type of improvement:* Decide, whether an interval I_l has to be bisected or the present order of the truncation index p_l has to be increased. For that the behavior of the error of the local expansion is compared to a somehow ideal behavior in terms of a type parameter ([3]) or an error prediction model.
2. *Cut-value:* Compute a threshold value, to decide which intervals have to be changed.
3. *Assemble and solve:* Built up the so-called Galerkin equations and solve them.
4. *Estimate:* Estimate the error of the obtained solution.
5. *Loop control:* If the error is smaller that the required tolerance stop, otherwise go back to 1.

4 NUMERICAL RESULTS

In this section we discuss some numerical results for the countable system (2.16) under physical as well as numerical aspects.

The computations have been performed on a SPARC-station 1+ (direct integrations) and a Mac IIsi (Galerkin method). All computing times (CPU) for the Galerkin method are in the scale of seconds or one minute. Most of the computations were performed with the program CODEX-HP, which uses a Rothe method with multiplicative error correction as time discretization and a discrete h-p-method as stationary solver. In cases where direct integrations of the respective systems of ordinary differential equations were possible, they were performed with the extrapolation code EULSIM [10].

Physical Aspects Simulations of the process (2.16) for different M and different initial occupation distributions turned out, that in each case the system tends to a stationary state of the form

$$u_n^M = C \rho^n, \quad n = 1, 2, \dots,$$

with

$$\rho = \frac{M}{M+1}.$$

Indeed this can be verified by inserting u_n^M into the countable system (2.16). An important consequence of this observation is, that the stationary state only depends on the photon density M and *is independent of the initial distribution* (see also the numerical discussion below).

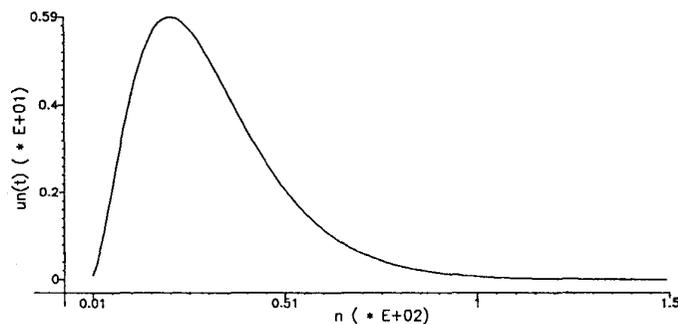


Figure 4.1: Steady-state of the intensities for $M = 10$.

Figure 4.1 shows the associated distribution of the intensities (2.18) for $M = 10$, which looks quite similar to each of the both branches of the measured spectrum in figure (2.1). Figure 4.2 presents the time evolution of the spectra starting from the initial occupation distribution $u_1(0) > 0, u_s(0) = 0, s = 2, \dots$

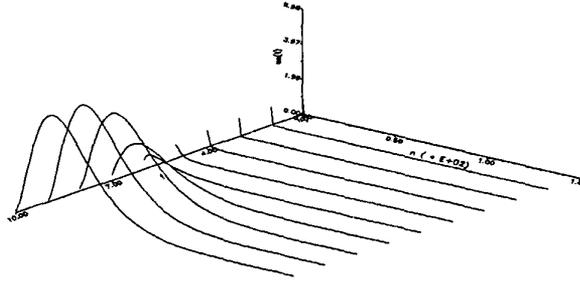


Figure 4.2: Time evolution of intensities.

These are interesting results justifying the simplifications of our model. Let us have a look to other current explanations of the infrared emission spectra (e.g. [17], p.122f): They all use terms from statistical equilibrium theory like the "Boltzmann factor" $\exp(-\frac{E_n}{kT})$ as basic occupation distributions. So they all are very heuristic explanations, more or less based on traditions in physics as on a real knowledge of the running process. It would be very interesting to answer the question, whether the predicted influence of external photon density in the spectra can be found experimentally.

Surely our discussion leaves many physically interesting questions open:

- What happens if we enlarge our model considering the underlying vibration-process of the molecules? In this case the basic physical theory of Section 2 leads to a system of *coupled* CODEs. This requires an extended implementation of the suggested numerical algorithm, which is in progress. Because the single CODEs of the system are similar to (2.16), we can expect a successful numerical computation by the extended method. Perhaps an explanation of the little differences of the outlines of the two branches in figure (2.1) will be possible in this way.
- How do other spectral distributions M_n affect on the stationary states?
- Which are the effects of short light pulses on the occupation distribution? Can we get new informations about the molecules observing their spec-

tral reaction to short light pulses? Note, that the numerical solution of CODEs like (2.16) with time-dependent $M(t)$ isn't a harder task than the given problem. At least we can say, that numerical tests with a suddenly switched-off photon field has lead to the expected results in form of an exponential decrease of the intensities.

We leave such examinations for later work, because appropriate models are still missing and our main aim was to show the usefulness of a numerical treatment of CODEs. Tools for the solution of time-dependent problems are available, models fitting real experiments are still missing.

Numerical Aspects Dealing with the countable system (2.16) as a large ordinary differential equation leads to some difficulties. Truncating this hierarchical system at an index s_{\max} leads to a very stiff system and it turned out, that the results strongly depend on s_{\max} . Figure 4.3 illustrates on the obtained spectra for $M = 10$ and $t_{\text{end}} = 5$ sec. for $s_{\max} = 50, 75, 100, 150$, that for $s_{\max} > 10$ the results begin to be satisfying.

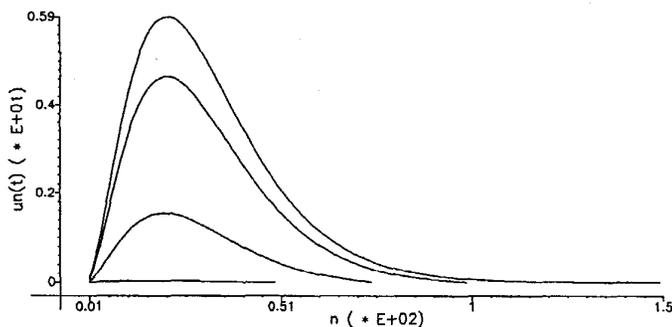


Figure 4.3: The effect of truncating a CODE at indices $s_{\max} = 50, 75, 100, 150$ (from below).

In particular for $s_{\max} = 50$ and $s_{\max} = 75$ the results are hopelessly wrong. As a rule of thumb we can say, that in this example $s_{\max} \approx 10 M$ must be chosen for a reasonable direct integration. Thus the simulation of a system with $M = 100$ as shown in the next figure computed by a h - p Galerkin method would require the treatment of about 1000 components of u_n . The h - p -method adapts not only the index n_{\max} with time automatically (see Figure 4.5), but computes an adaptive approximation for the occupation distribution even when the initial distribution seems to look quite difficult for numerical approximations.

In order to demonstrate this, we started a simulation with

$$(4.1) \quad u_n(0) = \begin{cases} c & , \text{ for } n \leq 20 , \\ c(1-c)^{n-1-20} & , \text{ for } n > 20 , \end{cases}$$

where $c = 0.1$ was set.

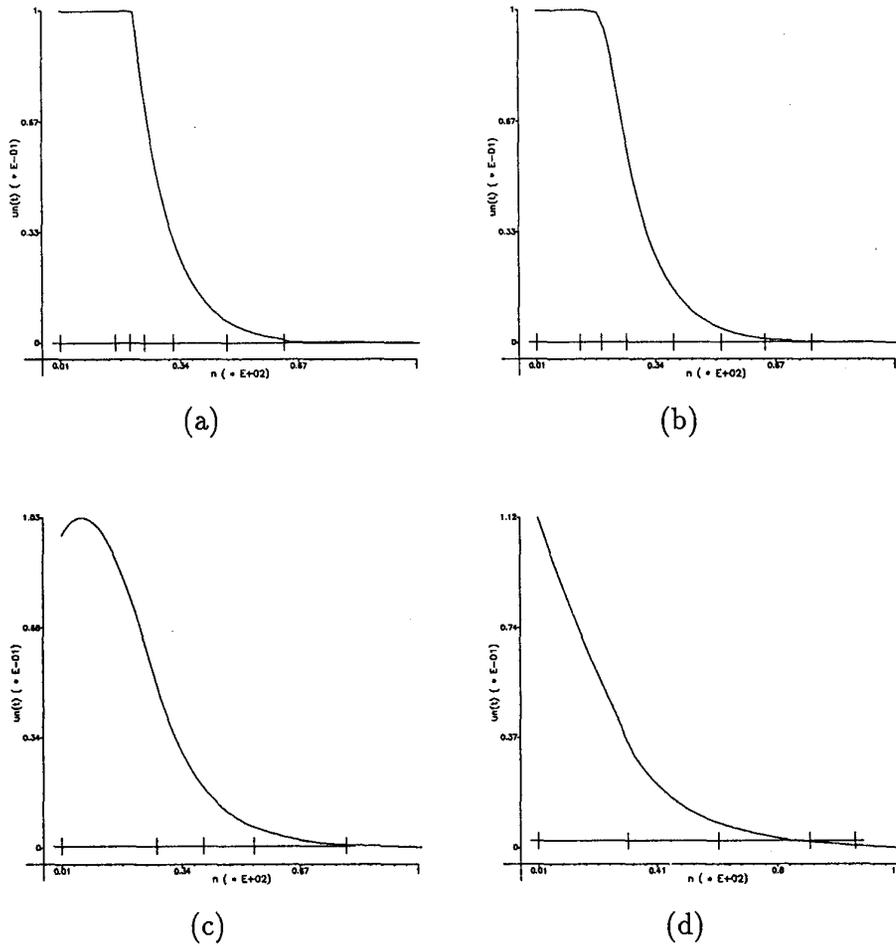


Figure 4.4: Initial phase of a computation with $M = 100$ and initial distribution (4.1)

In Figure 4.4 four time layers from the initial phase are presented in this case, where the effect of smoothing can nicely be seen. The grids and orders are also seen documenting the working of the h-p-method. Obviously this simulation

also leads to a stationary state u_n^M . Figure 4.5 shows the grids chosen by the h-p-method for the whole process up to $t = 0.5$ sec., reflecting how the stationary state is reached.

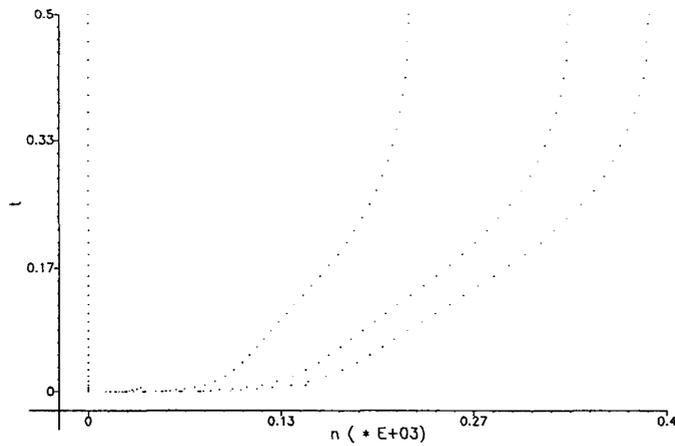


Figure 4.5: Evolution of grids in the computation with named initial data.

The number of degrees of freedom is 27 in the beginning and essentially decreases to 14 for the stationary state. The stepsizes chosen by the algorithm increases until the stationary state is reached, then it remains constant. Other tests with very large M and computations for the *harmonic oscillator* showed, that the h-p-method works reliable and independent of s_{\max} . This opens the possibility of very extended models in this context.

ACKNOWLEDGEMENT

We are pleased to thank Prof. Dr. P. Deuffhard. He put us in the way by calling our attention to the structural similarities between CODEs and physical problems of discrete nature.

REFERENCES

- [1] J. Ackermann, M. Wulkow: MACRON – A Program Package for Makromolekular Reaction Kinetics. Konrad-Zuse-Zentrum Preprint SC-90-14 (1990).
- [2] M. Wulkow, J. Ackermann: *The Treatment of Macromolecular Processes with Chain-Length-Dependent Reaction Coefficients – An Example from Soot Formation*. Konrad-Zuse-Zentrum Preprint SC-91-18 (1991).
- [3] W. Gui, I. Babuška: *The h , p and h - p Versions of the finite Element Method in 1 Dimension, Part I-III*, Numer. Math. 49 (1986).
- [4] Banwell: *Fundamentals of Molecular Spectroscopy*, McGraw Hill, Second Edition
- [5] F. A. Bornemann: *An Adaptive Multilevel Approach to Parabolic Equations I. General Theory and 1D Implementation*. IMPACT Comput. Sci. Engrg. 2, 279–317 (1990).
- [6] F. A. Bornemann: *An Adaptive Multilevel Approach to Parabolic Equations II. Variable-Order Time Discretization Based on a Multiplicative Error Correction*. IMPACT Comput. Sci. Engrg. 3, 93–122 (1991).
- [7] U. Budde, M. Wulkow: *Computation of Molecular Weight Distributions for Free Radical Polymerization Systems*. Chem. Eng. Sci., Vol. 46, No. 2 (1991).
- [8] P. L. Chebyshev: *Sur les fractions continues*. J. Math. Ser. 2, 3 (1855).
- [9] K. Deimling: *Ordinary Differential Equations in Banach Spaces*. Lecture Notes in Mathematics, Vol. 596, Springer (1977).
- [10] P. Deuffhard: *Recent Progress in Extrapolation Methods for Ordinary Differential Equations*. SIAM Rev. 27 (1985), pp. 505-535.
- [11] P. Deuffhard, A. Hohmann: *Numerische Mathematik. Eine algorithmisch orientierte Einföhrung*, de Gruyter (1991).
- [12] P. Deuffhard, M. Wulkow: *Computational Treatment of Polyreaction Kinetics by Orthogonal Polynomials of a Discrete Variable*. IMPACT Comput. Sci. Engrg., 1 (1989) 269–301.
- [13] W. Ebeling: *Applications of Evolutionary Strategies.*, Syst. Anal. Model. Simul. 7 (1990).
- [14] *Encyclopedia of Physics*, Vol. XXVI, *Light and Matter II*, Springer (1958).
- [15] H. Haken: *Synergetik*. Springer, 2. Auflage (1983).
- [16] H. Haken: *Quantenfeldtheorie des Festk6rpers*, Teubner 1. Auflage (1973).
- [17] Herzberg: *Spectra of diatomic molecules*, Van Nostrand, Second Edition
- [18] E. Hille: *Pathology of infinite systems of linear first order differential equations with constant coefficients*. Ann. Mat. Pura. Appl. 55 (1961) 133 – 148.
- [19] G.H. Mateescu, M. Avram *Infrared Spectroscopy*, R.E. Krieger (1978).

- [20] A. Messiah: *Quantum Mechanics II*, North-Holland (1970).
- [21] A. F. Nikiforov, V. B. Uvarov: *Special Functions of Mathematical Physics*. Birkhäuser (1988).
- [22] W. H. Ray: *On the Mathematical Modeling of Polymerization Reactors*. J. Macromol. Sci.-Revs. Macromol. Chem. C8 (1) (1972), pp. 1-56.
- [23] P. Canu, W.H. Ray: *Discrete Weighted Residual Methods Applied to Polymerization Reactors*, to appear in Comp. Chem. Engrg.
- [24] M. Wulkow: *Adaptive Treatment of Polyreactions in Weighted Sequence Spaces*. Konrad-Zuse-Zentrum Berlin, Preprint 91-17 (1991), to appear in IMPACT Comput. Sci. Engrg.