Contents

1 Introduction

In the past decades the progress of laser technology has been immense. This has led to the development of many new methods for measurement or control of processes in Science and Engineering, especially in Chemistry and Solid–State Physics, using various forms of interaction between laser fields and matter. As one of these methods, the *laser–induced excitation of molecules* is playing an increasingly important role in modern Chemistry. Particularly the usage of *ultrashort laser pulses* is a promising new technique for controlling or accelerating chemical reactions. One key problem in this new field is the simulation of the dynamical reaction of a molecule subjected to infrared laser pulses. This key problem is treated herein: The aim is the construction of an efficient simulation algorithm. In order to characterize the specific challenge of this problem class we begin with an illustration of the basic ideas of "laser–assisted" Chemistry. After this we shortly deal with the numerical difficulties arising in this context. Finally, we describe the basic pattern of the new algorithm to be presented herein.

Illustration of Chemical Problem Class

We consider a typical chemical reaction process: a well–balanced mixture of substances is heated so that the thermally activated substances *mainly* react towards the desired product. The expression "mainly" is inserted in order to underline the fact that thermal activation is a deeply *statistical* effect. The activated substances react towards *all* possible products due to the situation dependent statistical probabilities, not only towards the desired one.

Let us choose a particularly simple example: the dissociation of HOD. The two different dissociation processes

$$\mathrm{HOD} \rightleftharpoons \mathrm{H}^+ + \mathrm{OD}^- \qquad \text{and} \qquad \mathrm{HOD} \rightleftharpoons \mathrm{D}^+ + \mathrm{OH}^- \tag{1.1}$$

take place with nearly equal probabilities. Now, assume we are only interested in OD^- anions. Unfortunately, chemical standard methods do not allow us to suppress the second dissociation. Is there any alternative way to *control* the dissociation *selectively*? A promising idea is to control the behaviour of the HOD–molecules by external, well–designed laser–radiation: First, use infrared radiation to selectively excite a well–chosen vibration of the OH–bond in HOD. Second, induce the desired dissociation by use of UV–radiation in order to shift the vibrationally excited $(\mathrm{HO})^{*}\,\mathrm{D}$ into an instable, delocalized electronic state:

$$\mathrm{HOD} \xrightarrow{IR} (\mathrm{HO})^{*}\,\mathrm{D} \xrightarrow{UV} \mathrm{H}^{+} + \mathrm{OD}^{-} \tag{1.2}$$

In this situation the previously excited OH–bond is the first to break up so that, indeed, the dissociation $\mathrm{D}^{+} + \mathrm{OH}^{-}$ is suppressed. Thus, (1.2) is a *selective* dissociation process which, indeed, was experimentally realized and is well–understood theoretically (cf. [23][27][47]).

The basic idea of this "2–photon–process" (1.2) can be transferred to many similar situations (e.g. selective isomerization [32]). Its second step, the UV–photoexcitation process, is well established whereas its first step, the *selective vibrational excitation*, is a subject of much actual interest. This first step can only be realized on a very short — femto- or picosecond — time scale because otherwise the local energy concentration in the OH–bond is blurred out by the processes of intramolecular energy redistribution [3].

The possibility of a selective control of molecules by ultrashort laser pulses was first documented by Zewail in 1980 [49]. Since then, important experimental and theoretical work has been done in this field which, due to the short time scale, is usually called "Femto–Chemistry" or, more descriptive, "Laser–Assisted Molecular Control". It is part of one key problem in chemistry ("Understanding Chemical Reactivity") which, according to [39], is topic 'A' in the list of chemical research frontiers with outstanding priority.

Obviously, the task of finding a vibrational excitation of highest selectivity can be seen as an *optimization problem* which, for the time being, may be formulated as: "Find the form of laser radiation which causes an optimally selective excitation of the molecule". An *experimental* optimization is nearly impossible or at least far too expensive. This shows to be an exemplary case in which simulation calculations are absolutely necessary. In this paper, as a contribution to the theoretical investigation (cf. [10][28][30][36][35][32]) of this simulation task, an algorithm for the fast simulation of laser–molecule interaction is presented.

Numerical Challenge of Problem Class

The mathematical model for the description and simulation of selective vibrational excitation is based on Quantum Theory, i.e. on the time–dependent Schrödinger equation. In Section 2.1 the details of the quantum theoretical derivation of this model are presented. It is shown that the Schrödinger equation can be reduced to an ODE of the form

$$i\,\partial_t c \;=\; (\Omega \,+\, f(t)\;\cos(\omega t)\;V_\infty)\;c, \tag{1.3}$$

with $c \in \mathbb{C}^n$, a diagonal matrix $\Omega = \mathrm{diag}(\Omega_{11}, \ldots, \Omega_{nn}) \in \mathbb{R}^{n\times n}$ and a symmetric one $V_\infty \in \mathbb{R}^{n\times n}$. The function $f : [0,\tau] \subset \mathbb{R}^+ \to \mathbb{R}^+$ models the laser pulse shape and is slowly varying in comparison with the light oscillation itself (represented by $\cos(\omega t)$).

Only the knowledge of the values $|c_k(t)|^2$ of the solution $c(t) = (c_k(t))_{k=1\ldots n}$ of (1.3) is of quantum chemical relevance. These values are called "populations of the vibrational states $k = 1\ldots n$". Their computational determination is the *inner problem* of the optimization problem named above and, numerically, its most difficult part. In this thesis we are only interested in a highly efficient algorithm for solving this inner problem. Thus, to us, the optimization problem is of minor importance. Details concerning the quantum chemical framework of the optimization problem can be found in Section 5, along with a detailed motivation of some chemical real life applications.

But which difficulties are to be faced when solving the inner problem? Can't we simply use a standard integrator for solving (1.3), simultaneously computing the populations? The solutions of (1.3) show *multifrequency, highly oscillatory behaviour* including fast oscillations with large amplitudes. These fast oscillations confine the stepsizes of any numerical integrator. Hence, our first simple idea causes tremendous computational effort. Nevertheless, exactly this method is used in most previous investigations concerning our inner problem.

If we compute the populations $|c_k(t)|^2$ of the highly oscillatory solutions we also observe fast oscillations but with small amplitudes only. Particularly in our kind of applications, it is sufficient to know the running average of the populations

$$\left(\mathcal{A}_T|c_k|^2\right)(t) \;=\; \frac{1}{T}\int\limits_{t-T/2}^{t+T/2} |c_k|^2(s)\,ds \tag{1.4}$$

because the fast oscillations are only small deviations from the "general evolution" $(\mathcal{A}_T|c_k|^2)(t)$ of the populations. Obviously, the choice of the average length T determines which oscillations are "fast" and which not.

The problem of highly oscillatory behaviour of ODEs has received a considerable amount of attention in Numerical Mathematics: there are several different filtering and averaging techniques which directly compute smooth approximations of a considered highly oscillatory solution. As shown in Section 2.2, all these *smoothing techniques* unfortunately are *inappropriate* for our inner problem for two main reasons: First, in all these methods it is basically assumed that fast oscillations of the solution of an ODE are caused by large imaginary eigenvalues of the Jacobian of the right hand side of this ODE. This assumption arises from the study of *autonomous* ODEs. Here, however, we deal with the specific *nonautonomous* ODE (1.3). Its oscillatory behaviour is deeply transient and cannot be characterized sufficiently by the eigenvalues (see Section 2.2 for details). Second, because of the general non–identity $|\mathcal{A}_T c_k|^2 \neq \mathcal{A}_T|c_k|^2$, we cannot compute the averaged populations $\mathcal{A}_T|c_k|^2$ from the smoothed solutions $\mathcal{A}_T c_k$. Thus, the direct smoothing of the solutions is no way to solve our "inner problem".

In addition, there are different "non–smoothing" techniques for solving highly oscillatory ODEs (see Section 2.2 again). But they also are either inappropriate or not efficient enough for our kind of problem (e.g. Floquet methods [7], the Fourier–Galerkin method [45], or the quasi–envelope method [19][38]).

Basic Pattern of Quasiresonant Smoothing Algorithms

The algorithm presented herein determines the averaged populations $\mathcal{A}_T|c_k|^2$ by computing a *smooth* solution of an ODE which results from (1.3) by "deliberate sparsing" of V_∞. Therefore it is called a "smoothing algorithm". Its basic idea is the *quasiresonant approximation* which motivates the named variation of V_∞ and can shortly be explained as follows:

The observation of so–called "transition conditions" constitutes a basic physical insight into molecule–light interaction processes. With respect to the notations used in (1.3) this insight can be formulated in the following way: "The interaction between the populations $|c_k|^2$ and $|c_l|^2$ can only be important for the evolution of the populations if the frequency ω fulfills the resonant *transition condition* $|\Omega_{kk} - \Omega_{ll}| \approx \omega$." In (1.3) the interaction between $|c_k|^2$ and $|c_l|^2$ is represented by the elements V_{kl} and V_{lk} of V_∞. Now, let (k, l) be an index pair for which the transition condition is broken. Can we then set $V_{kl} = V_{lk} = 0$ without changing the evolution of the populations? The answer is: yes, if we do it carefully. This is the basis of the quasiresonance–idea allowing the construction of a quasiresonant smoothing (QRS) algorithm which is explained in more detail in Section 3. Therein, it is demonstrated that the quasiresonance idea allows a deliberate sparsing of V_∞ controlled by a natural splitting parameter δ. It is further shown that, indeed, a well–chosen δ leads to smooth solutions $d(t)$ and smooth populations $|d_k(t)|^2 \approx (\mathcal{A}_T|c_k|^2)(t)$ simultaneously. Thus, this method is useful to obtain an extreme speedup in the solution of the inner problem.

Yet, in the pure QRS–method the crucial point is the choice of the right sparsing parameter δ. If δ is too small, wrong results are obtained. If it is too large, the efficiency can be drastically reduced. In order to get a reliable algorithm, one must be prepared to adapt the parameter δ according to the local variation of the solution. Thus, the construction of an *adaptive* QRS–version is necessary which determines δ step by step according to an accuracy requirement. This is successfully done in Section 4: an error estimation scheme is constructed which allows us to study the effect of different δ directly on the smoothed populations in the next step. Since δ controls the sparsing of V_∞, the background of this error estimator is perturbation analysis of the effect of the elements of V_∞ on the smoothed populations. In Section 4 all these investigations lead to the construction of an adaptive QRS–algorithm (AQRS).

Finally, in Section 5, the performance of AQRS is demonstrated in several real life applications from "Laser–Assisted Molecular Control". We observe that, applied to these test problems, AQRS works reliably, smoothes the populations, and leads to speedup factors of the order of 10^2 in most test problems. The limits of its smoothing properties are demonstrated, as well. In cases of extremely high laser field strengths AQRS does not smooth the populations but still works reliably — and efficiently. The general advantage of AQRS, independent of any dynamic smoothing effect, is the reduction of the number of operations needed in each evaluation of the right hand side of (1.3): AQRS reduces the usual $\mathcal{O}(n^2)$ complexity to just $\mathcal{O}(n)$.

Acknowledgement

I would like to express my deep gratitude to P. Deuflhard. He has enabled me to find my way towards interdisciplinary work in Scientific Computing and supported my research all the time. Coming from purely theoretical Physics I was an absolute newcomer in Numerics. He, who himself has turned from Physics to Numerics, has always accepted my lack of numerical knowledge and, in doing so, helped me to accept that a well–grounded understanding of Numerics is absolutely necessary for doing interdisciplinary work of sufficient quality. It is also to his merit that my attention was turned to the challenging class of applications presented herein.

I also thank J. Manz and G. Paramonov for very informative discussions about the quantum chemical modelling and the simulation techniques used in the field "selective excitation with picosecond laser pulses" and for supplying me with the concrete model parameters for the "real life applications". B. Fiedler should not be forgotten who has resolved the question of whether there are tools from asymptotic analysis which could have been used.

Many thanks to the staff at the ZIB, in particular F. Bornemann, for encouraging me to take my algorithmic ideas seriously and for many helpful discussions about the characterization of oscillatory behaviour; J. Lang for taking the part of the mathematical critic in many conceptional problems; and, last but surely not least, A. Hohmann, for enduring my impatience and especially for his wonderful C++ extrapolation codes which allowed me a fast and reliable implementation of all my algorithmic ideas.

2 Preliminary Considerations

This section contains contributions to two different questions which should shortly be discussed before we start a detailed presentation of the promised AQRS–algorithm. First, we give a short description of the quantum theoretical framework of the considered kind of laser–molecule interaction (Section 2.1). In particular, some quantum theoretical aspects of model preparation and the corresponding interpretation context are explained, and a derivation of the basic dynamical equation is given.

Second, we deal with the question of whether a new simulation technique for solving the derived model equations is really necessary? Indeed, there are several well–known numerical methods which we could take for promising simulation techniques in our context. In Section 2.2 these methods are discussed in order to analyse their applicability and concrete efficiency with respect to our kind of problem.

2.1 Description of Problem Class

2.1.1 Basic Quantum Chemical Equations

The physical systems we consider are samples of dilute gas molecules subjected to monochromatic infrared laser radiation. We are only interested in the population dynamics of a set of discrete states during the interaction of a single molecule with the external laser field.

In this situation the standard method used is the *electric dipole approximation*: writing the molecular Hamiltonian in the form

$$H = H_0 + H_1 \tag{2.1}$$

$$H_1 = -\vec{\mu} \cdot \vec{E} \tag{2.2}$$

where H_0 is the pure molecular Hamiltonian, neglecting the field $\vec{E}(t)$. $\vec{\mu}$ is the molecular dipole moment operator. All operators are Hermitian and act on a separable Hilbert space $\mathcal{H}$ with scalar product $< \cdot | \cdot >$. For monochromatic radiation we may write

$$E(t) = f(t) \cos \omega t \tag{2.3}$$

with the pulse form function f, which in comparison with the light oscillations $\cos \omega t$ is a slowly varying function. To describe a constant light source we may choose $f \equiv \text{const}$; in order to model a short laser pulse, f may denote the *laser shape function.*

We consider the spectrum σ of H_0 to be given. In most real life applications σ includes a discrete part $D = \{\epsilon_k |\ k \in J\}$ with $J \subseteq \mathbb{N}$ and a nonempty subinterval C of $\mathbb{R}$. Let φ_k be eigenfunction to $\epsilon_k \in D$:

$$H_0 \, \varphi_k \ = \ \epsilon_k \, \varphi_k, \qquad \forall k \in J.$$

The key for the description of the population dynamics is the time dependent Schrödinger equation:

$$i\hbar \, \partial_t \psi \ = \ H\psi. \tag{2.4}$$

Let us use a Galerkin approximation of $\psi(t) \in \mathcal{H}$ in the linear subspace

$$\mathcal{H}_n \ := \ \text{span}\{\varphi_k, \ k = 1 \ldots n\}, \qquad n \in \mathbb{N} \text{ fixed} \tag{2.5}$$

of $\mathcal{H}$, i.e. assume

$$\psi(t) \ = \ \sum_{k=1}^{n} c_k(t) \ \varphi_k. \tag{2.6}$$

With this eigenfunction expansion of ψ Schrödinger's equation is projected into $\mathcal{H}_n$ and can be transferred into an ODE–form:

By use of the assumption $\vec{E}(t) \| \vec{\mu}$ with $\vec{\mu} \cdot \vec{E} = f(t) \cos(\omega t) \ \mu$ we deduce from (2.4):

$$\begin{aligned} i\,\hbar\,\partial_t < \varphi_k|\psi > \ &= \ < \varphi_k|\, i\,\hbar\,\partial_t\psi > \\ &= \ < \varphi_k|H_0\psi > \ - \ f(t)\cos(\omega t) \ < \varphi_k|\mu\psi > . \end{aligned}$$

Inserting (2.6) gives

$$i\,\hbar\,\partial_t c \ = \ (\Omega \ + \ f(t)\cos(\omega t)\, V_\infty) \ c \tag{2.7}$$

In (2.7) $c = (c_k) \in \mathbb{C}^n$ is the n-dimensional (complex) vector of the expansion coefficients, Ω the real and diagonal matrix of the molecular eigenvalues $\Omega \ = \ \text{diag}(\epsilon_1, \ldots, \epsilon_n)$ and $V_\infty = (V_{kl})$ the dipole matrix defined by:

$$V_{kl} \ = \ - \ < \varphi_k|\mu \ \varphi_l > . \tag{2.8}$$

Thus, according to (2.6), a solution $c(t)$ of (2.7) gives us an approximation $\psi(t)$ of the solution of Schrödinger's equation (2.4) up to a certain *projection error*. Certainly one has to analyse the quality of the approximation, i.e. its projection error, but this is not the task of this paper (see Remark 1). It rather is a part of the task of *model preparation* in each single application. Herein the concrete model is always assumed to be given, including a choice of a $n \in \mathbb{N}$ which leads to a sufficiently small projection error and including the specification of the data $(f, \omega, \Omega, V_\infty)$. Thus, once and for all, let us replace the Schrödinger equation (2.4) as the basic equation of evolution by the finite ODE system (2.7) and the solution $\psi(t)$ of (2.4) by the eigenfunction expansion in (2.6), i.e. by the solution $c(t)$ of (2.7).

Normally (2.7) is transformed to *dimensionless units* by the substitutions:

$$\begin{aligned} \tilde{t} &= \frac{\epsilon_2 - \epsilon_1}{\hbar} t \quad \text{and} \quad \tilde{\Omega} = \frac{1}{\epsilon_2 - \epsilon_1} \Omega \\ \tilde{V}_\infty &= \frac{1}{\epsilon_2 - \epsilon_1} V_\infty \quad \text{and} \quad \tilde{\omega} = \frac{\hbar\omega}{\epsilon_2 - \epsilon_1} \end{aligned} \tag{2.9}$$

from which results

$$i\,\partial_{\tilde{t}}\, c = (\tilde{\Omega} + f(\tilde{t}) \cos(\tilde{\omega}\tilde{t})\, \tilde{V}_\infty)\, c. \tag{2.10}$$

For simplicity the old symbols without tildes are held, so that the final *standard form of the basic equation of motion* is written

$$i\,\partial_t c = (\Omega + V(t)\, \cos(\omega t)\,)\, c, \tag{2.11}$$

with

$$\begin{aligned} V(t) &= V_\infty\, f(t) \\ \Omega &= \operatorname{diag}(\omega_1, \ldots, \omega_n) \\ \omega_k &= \frac{\epsilon_k}{\epsilon_2 - \epsilon_1}. \end{aligned}$$

Equation (2.11) can be transferred into the so–called "interaction picture" (which plays a central role in Section 4.1) by the *unitary* transformation

$$b(t) = \exp(i\,\Omega\, t)\, c(t) \tag{2.12}$$

The result is the *interaction picture version of* (2.11)

$$i\,\partial_t b \;=\; e^{i\Omega t}\;V(t)\;\cos(\omega t)\;e^{-i\Omega t}\;b. \tag{2.13}$$

Physically the expansion coefficients are not observable. Only the populations p_{kk} of the molecular eigenstates are measurable. They are given by the probabilities

$$p_{kk}(t) \;=\; |<\varphi_k|\,\psi>|^2 \;=\; |c_k|^2 \tag{2.14}$$

and are the diagonal elements of the *probability matrix*

$$p(t) \;:=\; c(t)\,\otimes\,\bar{c}(t). \tag{2.15}$$

In (2.15) $\bar{c}$ denotes the complex conjugate of c and $\otimes$ the common tensor multiplication which, using vectors a and b, is defined as

$$(a\,\otimes\,b)_{kl} \;:=\; a_k\,b_l. \tag{2.16}$$

Hence it is more than sufficient for our task to know the expansion coefficients $c(t) = (c_k(t))$ or any *unitary* transformed set of coefficients $\tilde{c} = U\,c$ (unitary U with $U^*U = 1$), for example the interaction picture coefficients b (the solution of (2.13)).

Equation (2.11) should be completed by introducing an initial condition. From the assumption that at $t = 0$ all systems occupy the *initial state* s we get

$$|c_j|^2(0) \;=\; \delta_{sj} \tag{2.17}$$

which (by *phase choice*) leads to the initial condition

$$c_j(0) \;=\; e^{i\pi/4}\,\delta_{sj} \;=\; \frac{1}{\sqrt{2}}(1+i)\,\delta_{sj}, \qquad \forall j. \tag{2.18}$$

In most cases the initial state is the ground state $s = 1$ of the quantum system.

2.1.2 Physical Interpretation

In Quantum Theory a *system* is a statistical sample of a lot of identically prepared subsystems (in our case each subsystem is a single molecule in interaction with the external field). The solution ψ of the Schrödinger equation (2.4) is called the *state* of the considered system while an eigenfunction φ_k of H_0 is called a *molecular eigenstate* with energy ϵ_k. If a system remains in a molecular eigenstate, i.e.

$$\psi(t) \;=\; \varphi_k \;\exp\left(i\frac{\epsilon_k}{\hbar}t\right) \tag{2.19}$$

the system is said to allow a *sharp energy measurement* (with result ϵ_k at each subsystem). In this case only one expansion coefficient c_k is different from zero ($c_l(t) \;=\; \exp(i\,\epsilon_k/\hbar\,t)\,\delta_{lk}$). If there are some nonvanishing c_k in (2.6) for the system's state ψ, the system is interpreted to *occupy a mixture of eigenstates simultaneously.* Hence on the basis of the statistical interpretation of Quantum Theory the value

$$p_{kk}(t) \;=\; |<\psi(t), \varphi_k>|^2 \tag{2.20}$$

is interpreted as the *probability* of the fact that a system which is in state ψ occupies the molecular eigenstate φ_k. This probability is equal to the *rate* of subsystems at which the energy ϵ_k can be measured at this moment. Therefore a p_{kk} is often called the *occupation rate* or simply *population of (eigen)state k*. In this manner of speaking our basic equation (2.11) describes the dynamics of the interactions between the (occupations of the) (eigen)states.

Let us use the following overloaded notation:
For a matrix $A \in \mathbb{C}^{n\times n}$, $A = (A_{kl})$

$$\mathrm{diag}\,A \;:=\; (A_{11}\ldots A_{nn}) \tag{2.21}$$

denotes the n-tupel of the diagonal elements of A, while for a n-tupel $(x_1 \ldots x_n)$

$$\mathrm{diag}\,(x_1\ldots x_n) \;=\; A \;\in\; \mathbb{C}^{n\times n} \tag{2.22}$$

denotes the diagonal matrix A with diagonal elements x_k. In the following this notation is only used in situations in which "confusion is impossible".

Thus in the herein considered systems the physical observables which we are interested in are

$$\operatorname{diag} p(t) \;=\; \Big(|c_1|^2(t)\ldots|c_n|^2(t)\Big) \tag{2.23}$$

These values are called *probabilities*. Therefore they all should fulfil the property

$$p_{kk}(t) \;=\; |c_k|^2(t) \;\in\; [0,1] \tag{2.24}$$

for all $t > 0$. Indeed this is the fact. It follows from the *conservation* property of the Schrödinger equation

$$\partial_t < \psi(t), \psi(t) > \;=\; 0 \tag{2.25}$$

and its corresponding formulation

$$\sum_k |c_k|^2(t) \;=\; \sum_k |c_k|^2(0) \;=\; 1. \tag{2.26}$$

REMARK 1. Above, the dynamical evolution of our quantum system has been projected into the linear subspace

$$\mathcal{H}_n \;:=\; \operatorname{span}\{\varphi_k,\; k = 1\ldots n\} \tag{2.27}$$

of the considered Hilbert space $\mathcal{H}$. We have already noted that the projection error of this Galerkin method has to be analysed before we can take the evolution of the populations $|c_k(t)|^2$ in $\mathcal{H}_n$ as sufficient approximations of the exact populations $| < \psi(t)|\varphi_k > |^2$ computed using the exact solution $\psi(t)$ of Schrödinger's equation. The quality of the Galerkin approximations is not discussed in this paper because it is another aspect of work and belongs to the task of *model–preparation* in each single application. But we can give a short experience–based statement which shows the physical background of the usefulness of the considered Galerkin approximations. For field strengths which are not too large an essential occupation of "states" $x \in \sigma/D = C$ is "nearly impossible" if the initial state s lies "deep enough":

$$\text{for most } l \in J: \quad |\epsilon_s - \epsilon_l \pm \omega| \;\ll\; \min_{x\in C} |\epsilon_s - x \pm \omega|. \tag{2.28}$$

Therefore we can hope to find a subspace $\mathcal{H}_n \subseteq \operatorname{span}\{\varphi_k,\; k \in J\}$, i.e. a $n \in \mathbb{N}$, with sufficient approximation properties.

2.1.3 Application Problems

The application problems considered herein may be separated into two classes:

1. selective excitation of vibrational states and
2. selective configuration changes (selective isomerizations).

For both problems the modelling can be done as explained above with (2.11) as equation of motion for the populations. A detailed description of the chemical background of the single examples and of the respective specification of the data V_∞, Ω, ω and f is given in Section 5. In general, the form of the laser shape function f must not be fixed. But in this work the model for the laser pulses always are $\sin^2$–shapes (adopting the choice in the chemical literature), written in the dimensionless time coordinate of (2.11) as

$$f(t) \; = \; E_r \, \sin^2(\eta t) \tag{2.29}$$

with $\eta = \frac{\pi\hbar}{\tau(\epsilon_2 - \epsilon_1)}$ describing a pulse length τ of e.g. one picosecond. As briefly mentioned above the essential property of f is its slow variation with respect to the light frequency ω which can be expressed as

$$\eta \ll \omega. \tag{2.30}$$

Now, assume $c^{(f,\omega)}(t)$ to be the solution of

$$i\,\partial_t c \; = \; (\Omega \, + \, f(t) \, \cos(\omega t)\, V_\infty \,) \; c, \qquad c_j(0) \; = \; \frac{1}{\sqrt{2}}(1+i) \; \delta_{sj} \, \forall j \tag{2.31}$$

for given (f, ω) and fixed (Ω, V_∞) and let

$$p^{(f,\omega)}(\tau) \; = \; \mathrm{diag}(c^{(f,\omega)} \otimes \bar{c}^{(f,\omega)})(\tau) \tag{2.32}$$

be the corresponding final populations. In Section 5 we will see that some important chemical applications require a solution of the following *optimization problem*: determine a laser pulse (f, ω) which leads to an optimal population $|c_l(\tau)|^2$ of a given target state l. The parameters of this variational problem are clearly fixed if we assume that e.g. the pulse form function f is given by

(2.29):
"*Determine* (E_r, ω) *so that* $|c_l^{(E_r,\omega)}(\tau)|^2 = \max!$".
Obviously this optimization problem includes the evaluation of the populations $|c_k^{(f,\omega)}(t)|^2$ of (2.31) as an *inner problem.* In this text we deal with the inner problem mainly: an adaptive algorithm is presented which is constructed for use as a highly efficient integrator for the solution of the inner problem. But we have to remember that there is sometimes a demand for solving the optimization problem, too. In particular this is important for the discussion in Subsection 2.2.6.

REMARK 2. Our molecule–light interaction model does not include the essential process of *energy redistribution.* For the two classes of application problems named this fact restricts the range of validity of our model to laser pulse lengths τ in the range of a few picoseconds. The period $T = 2\pi/\omega$ of the used infrared light is about 10^0–10^2 femtoseconds. Hence the light–periods per pulse length τ/T is mostly of the order of 10^2.

2.1.4 Illustrative Example

In Section 2.2 we will analyse the difficulties which arise from a numerical solution of (2.11). Therefore, an illustrative example is needed, i.e. a set of concrete data for $(f, \omega, \Omega, V_\infty)$. For this purpose a relatively simple example is used, named "the OH test problem": concrete data are computed for a sample of one–dimensional anharmonic oscillators modelling the vibrational degrees of freedom of single OH–bonds. Remember: the vibrational excitation of an OH–bond was already considered as part of a selective dissociation process of HOD in the introduction.

More details and a description of the chemical scope in which this problem appears can be found in Section 5. For the time being, we are content with Figure 1 which exemplifies the structure of the solution $c(t)$ of (2.11) and (2.13) for our test problem 'OH'.

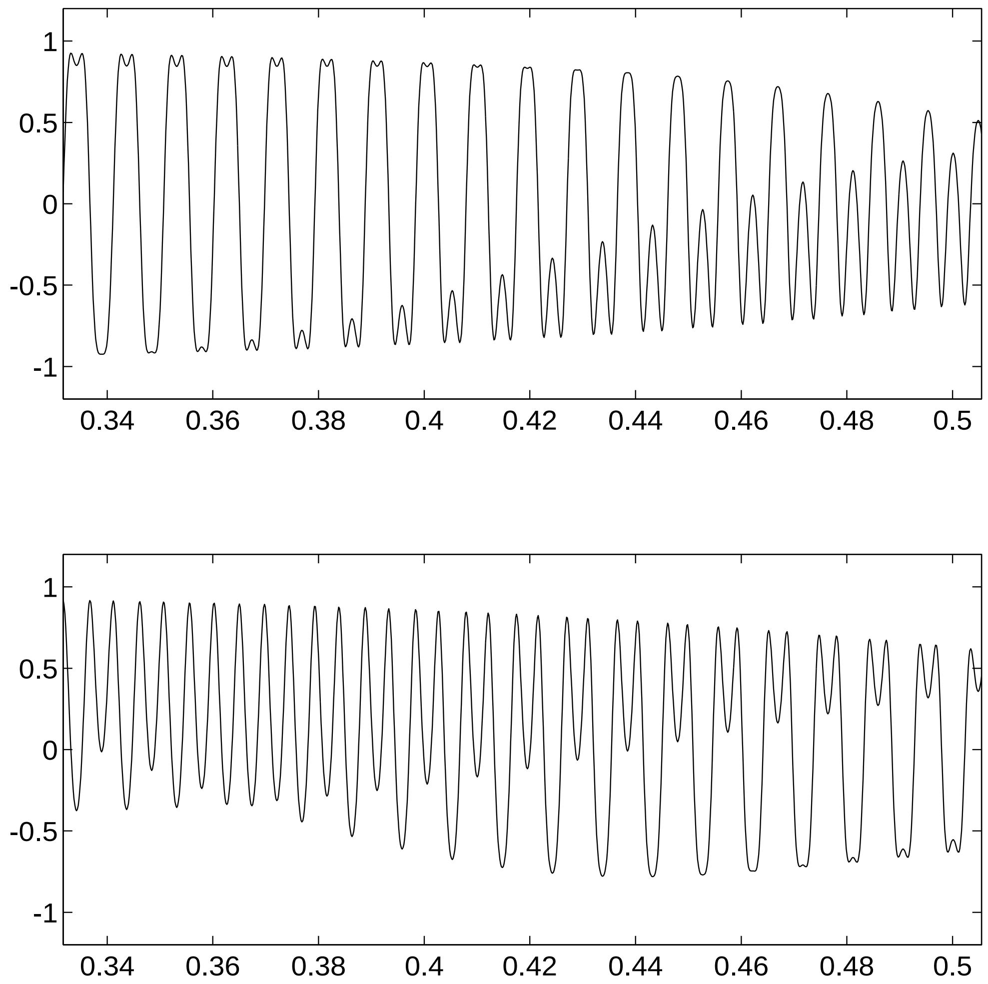


FIG. 1. Real- and imaginary part of the coefficient c_1 of a solution c of (2.11) versus time in picoseconds. The data $(f, \omega, \Omega, V_\infty)$ are taken from a real life test problem 'OH', see Table II. In this case we have $\omega_1 = 0$ for the first component of Ω. Thus, the graphs simultaneously show the coefficient $b_1 = \exp(-i\omega_1 t)c_1$ of the solution b of (2.13). Note the highly oscillatory character of the solutions.

2.2 Collection of Competing Approaches

Figure 1 demonstrates that the solutions of (2.11) can be characterized as *highly oscillatory.* There are several methods constructed by mathematicians and physicists which deal with the task of efficiently solving ODEs with highly oscillatory solutions. In this section these methods are collected and evaluated with regard to our specific class of problems (solving (2.11): the inner problem; the optimization problem). For this purpose we will go through some details of these methods. We will observe that each of them is based on certain construction principles, each depending on basic properties of a specific class of ODEs. In the following our task is to analyse whether the properties of our kind of problem meet the requirements of these construction principles or not, thus deciding whether the corresponding methods are appropriate or inappropriate for solving our kind of problem. The performance (and efficiency) of appropriate methods is discussed and compared with that of the algorithm AQRS to be presented later on.

2.2.1 Inapplicability of Linear Smoothing Techniques

First, the basic idea of a common class of filtering and averaging techniques, herein called "linear smoothing techniques", should shortly be explained, i.e. the idea of spectral analysis of the right hand side of the considered ODE. For this purpose, let us first consider the linear, autonomous ODE (our "model problem")

$$i\,\partial_t\,u \;=\; A\,u, \qquad u(t_0) = u_0 \tag{2.33}$$

with a constant coefficient matrix $A \in \mathbb{R}^{n\times n}$, which is assumed to be symmetric. This is often written as a second order ODE

$$x'' \;+\; A^2\,x \;=\; 0, \qquad x(0) = \mathrm{Re}(u_0), \quad x'(0) = A\,\mathrm{Im}(u_0). \tag{2.34}$$

We can find an orthogonal $C \in \mathbb{R}^{n\times n}$ with

$$\Lambda \;=\; C\,A\,C^T \;=\; \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \quad \lambda_i \text{ eigenvalue of } A \tag{2.35}$$

and the solution of (2.33) is

$$u(t) \;=\; C^T\;\exp\left(-i\,\Lambda\,(t-t_0)\right)\;C\,u_0 \tag{2.36}$$

Thus u can simply be characterized as highly oscillatory (or not) by checking the eigenvalues λ_i of the linear right hand side of (2.33):

$$\exists i: \ |\lambda_i| > \mu \quad \rightsquigarrow \quad u \text{ highly oscillatory.} \tag{2.37}$$

There are different *linear* transformations $\bar{u} = \mathcal{P}u$ which allow us to reduce (2.33) to an ODE

$$i\,\partial_t\,\bar{u} \;=\; C^T\,\mathcal{P}\,C\;\bar{u} \tag{2.38}$$

with smooth, i.e. slowly varying solutions. For example take $\mathcal{P} = P_C$ or $\mathcal{P} = P_A$ with

$$\begin{aligned} P_C \;&:=\; \mathrm{diag}(p_1,\ldots,p_n), \qquad &\text{with} \quad p_i \;&=\; \left\{ \begin{array}{ccc} 0 & : & |\lambda_i| > \mu \\ 1 & : & |\lambda_i| \le \mu \end{array} \right\} \\ P_A \;&:=\; \mathrm{diag}(q_1,\ldots,q_n), \qquad &\text{with} \quad q_i \;&=\; \tfrac{\sin(\lambda_i T/2)}{\lambda_i T/2}. \end{aligned} \tag{2.39}$$

The solution of (2.38) with $\mathcal{P} = P_A$ is the running average $(\mathcal{A}_T u)(t)$ of the original solution u of (2.33)

$$(\mathcal{A}_T u)(t) = \frac{1}{T} \int\limits_{t-T/2}^{t+T/2} u(s)\,ds. \tag{2.40}$$

The corresponding method is called “averaging”. The solution corresponding to P_C is the slowly oscillating part of the sum (2.36); this method is called “filtering”. These solutions are smooth (i.e. not *highly* oscillatory) and can therefore be numerically integrated more easily than u itself. We may draw the following conclusion:

In (2.33) the knowledge of the spectrum of A contains the whole structure of oscillating behaviour of its solution u. This total information can be used to essentially reduce the complexity of numerical evaluations of properties of u (e.g. its running average).

This leads us to a general basic idea: *use spectral information to facilitate numerical computations concerning ODEs with oscillating solutions.* We can take the usage of averaging (P_A) or filtering (P_C) as examples for techniques based on this idea; let us call them “linear smoothing techniques”.

Up to now, we have only analysed (2.33). But we are not really interested in linear, autonomous ODEs. Thus, are linear smoothing techniques applicable to nonlinear, nonautonomous ODEs? Linear smoothers are restricted to the significance of the eigenvalues λ_i, i.e. they rely on the possibility of a (local) "linearization" of the considered ODE.

$$
\boxed{
\begin{array}{rcll}
i\partial_t\, y & = & f(t,y) & (a) \\
 & \downarrow & \text{Linearize} & \\
i\partial_t\, u & = & A(t)\, u, \quad A(t)\ \text{sym.} & (b) \\
 & \downarrow & \text{Freeze Coefficients} & \\
i\partial_t\, u & = & A\, u, \quad A\ \text{sym.} & (c) \\
 & \downarrow & \text{Diagonalize} & \\
i\partial_t\, u & = & \lambda\, u & (d)
\end{array}
}
$$

FIG. 2. The standard paradigm of characterization of fast oscillations: reduction to a collection of scalar model problems.

Figure 2 summarizes the standard paradigm for the reduction of a nonlinear, nonautonomous ODE to linear model problems like (2.33). We begin with the ODE (a) of dimension n with differentiable f. The first step is to linearize (a) by assuming that we can find a substitution $y(t) = y_0(t) + u(t)$ with small u and a particular solution y_0 of (a) (e.g. its slow oscillating part). Now we pass from (a) to (b) by neglecting terms in u^2 and denoting the Jacobian of f in t as

$$A(t) \ := \ \partial_y f(t, y_0(t)). \tag{2.41}$$

The second step is to freeze the coefficients in A by setting $A = A(t_0)$ for some t_0. The idea here is that the problem of characterization of oscillation in solutions of ODEs is the *imaginary–axis–analogy* to the problem of characterization of stability and stiffness. Assuming A is diagonalizable, the third step is diagonalization as shown above. By this standard paradigm the characterization of oscillatory behaviour of nonlinear ODEs is transformed to the above discussed case of linear model problems, analogous to the characterization of stiffness and stability.

In [24] some difficulties of the stiffness–standard paradigm are collected which can be identically transferred into a criticism of the presented oscillation–standard–paradigm. There is one key argument: In general, oscillatory behaviour is a deeply transient phenomenon and the significance of the eigenvalues cannot be extended to a sufficiently large finite time scale. But even this extension to a sufficiently large time scale is essential because smoothing or averaging are deeply nonlocal operations (cf. (2.40)). This mainly criticizes the second step "Freeze Coefficients" in Figure 2: The spectrum of $A = A(t_0)$ *cannot* characterize the oscillatory behaviour of (b) in $[t_0, t_0 + T]$ if $A(t)$ itself oscillates (or: is of fast variation) in $[t_0, t_0 + T]$. Take

$$i\,\partial_t u \;=\; \Omega\, u \;-\; f(t)\,\cos(\omega t)\, u, \qquad \alpha \in \mathbb{R}, \Omega \in \mathbb{R} \tag{2.42}$$

as a simple example. Its solution

$$u(t) \;=\; u(0)\;\exp\left(i\int_0^t f(s)\cos(\omega s)\,ds \;-\; i\Omega t\right) \tag{2.43}$$

$$\;=\; u(0)\;\exp\left(i\left(\frac{\alpha}{\omega}\,\sin(\omega t)\;-\;\Omega t\right)\right), \quad \text{if } f(t) \equiv \alpha \tag{2.44}$$

contains double–trigonometric functions, e.g.

$$\cos\left(\alpha\,\frac{\sin(\omega t)}{\omega}\right) \;=\; \cos\left(\alpha\;\mathrm{sinc}(\omega t)\;t\right) \tag{2.45}$$

with time–dependent frequencies, e.g. $\eta(t) = \alpha\;\mathrm{sinc}(\omega t)$. The spectral information (eigenvalue of the Jacobian: $\lambda(t_0) = \Omega - \alpha\cos(\omega t_0)$) is oscillating and characterizes the oscillating solution only in a period $[t_0, t_0 + \tilde{T}]$ with $\tilde{T} \ll 2\pi/\omega$. If these T are *not sufficiently large*, i.e. if ω is comparatively large, we cannot construct a smoothing technique via use of spectral information only (cf. Figure 3).

But the class of herein considered ODEs (2.11)

$$i\,\partial_t c \;=\; (\Omega \;+\; f(t)\cos(\omega t)\,V_\infty)\;c$$

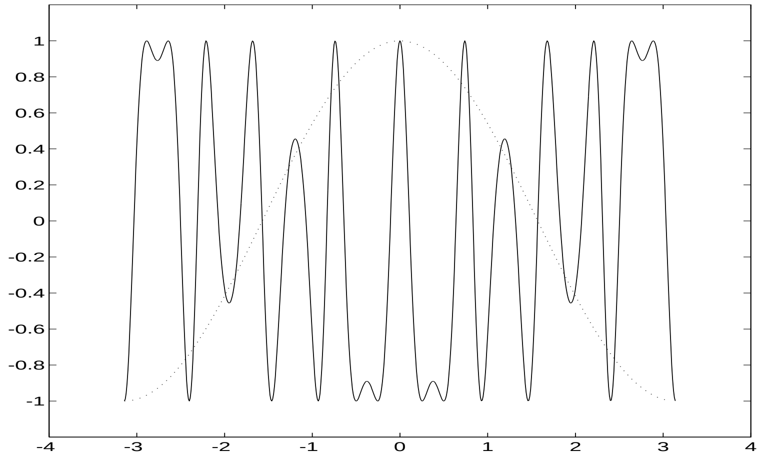


FIG. 3. The dense line represents a typical highly oscillatory solution (2.43) of the ODE (2.42). The dotted line shows the corresponding (real part $\cos((\Omega-\alpha)t)$ of the) solution of $i\partial_t u = (\Omega - \alpha)u$ which represents the characterization of the oscillatory behaviour which would have been made using the eigenvalue $\lambda(0) = \Omega - \alpha$ of the Jacobian in $t_0 = 0$. Note that this dotted line only approximates the dense line (= real solution) within an extremely small neighborhood of $t_0 = 0$ (which is not sufficiently large for means of any nonlocal averaging operation). Compare the dense line with Figure 1 and note the structural similarities of the oscillatory behaviour.

definitely belongs to the case of model problem (2.42). Indeed, (2.42) is just the one dimensional version of (2.11) and ω is comparatively large in the above sense. Thus, (2.11) leaves the domain of the applicability of the oscillation–standard–paradigm — its solutions show *multifrequency highly oscillatory behaviour.* Furthermore we see no hope in the use of linear transformations like (2.39) to efficiently solve our problem (2.11):
Linear smoothing techniques are bound to fail in our kind of problem.

This conclusion is also supported by the fact that it would simply be *useless* to compute the smoothed solution of (2.11) or (2.13) in a considered case: Let $c(t)$ be the solution of (2.11) or (2.13) and assume that we know $(\mathcal{A}_T c_k)(t)$. Unfortunately, we have to recognize that in general

$$\left(\mathcal{A}_T |c_k|^2\right)(t) \;\neq\; |(\mathcal{A}_T c_k)(t)|^2\,, \tag{2.46}$$

i.e. that smoothing the solution $c(t)$ leads to totally wrong smoothed populations $(\mathcal{A}_T|c_k|^2)(t)$ (cf. Figure 6 on page 45). But, lately, we are only interested in these smoothed populations! Because of this, $(\mathcal{A}_T c_k)(t)$ (or similarly smoothed data) is of no interest in our scope.

Surely, we can try to avoid this problem by using an ODE directly for the populations. For the probability matrix $p = c \otimes \overline{c}$ we easily derive

$$i\,\partial_t\, p \;=\; [\,\Omega \,+\, f(t)\cos\omega t\, V_\infty\ ,\ p\,]\,, \qquad p(0) = c_0 \otimes \overline{c_0} \tag{2.47}$$

with the matrix–commutator $[A,B] := AB - BA$. The averaged trajectories $\mathcal{A}_T p$ of (2.47) would directly give us the *averaged populations* $\mathcal{A}_T|c_k|^2 = \mathcal{A}_T p_{kk}$ of (2.11). Despite of this, we cannot get an increase in efficiency if we integrate (2.47) instead of (2.11), because (2.47) is an ODE of dimension n^2 if (2.11) is of dimension n. But perhaps, we can learn how to construct (nonlinear) smoothing techniques for (2.11) if we *formally* analyse the effects of linear smoothing techniques on (2.47)! Unfortunately, the right hand side of (2.47) is highly oscillating, just as the right hand side of (2.11) is. This leads to the same difficulties as mentioned above. Thus, linear smoothers are inappropriate also for studying the effect of averaging on (2.47).

Thus, our final conclusion is:

Linear smoothing techniques are inappropriate for our kind of problem.

And we add: the whole smoothing idea (linear as well as nonlinear techniques) must be transformed into an alternative concept: "Compute smoothed populations instead of smoothed solutions". In Section 3 QRS is introduced as being, firstly, based on this new concept and, secondly, as being an example for a *nonlinear* smoothing technique.

In addition to the initially described linear smoothing technique there are several other approaches which use (a-priori given) spectral information. Altogether they are also inappropriate to our kind of problem, again for one or both of the above listed reasons:

First, let us continue discussing alternative *smoothing methods*. Several authors have proposed algorithms to solve oscillating problems by computing running averages of the solutions $y(t)$ of specific ODEs. On one hand, there are alternative *linear* smoothing techniques like that documented in [34]. Therein a family of stable, λ–dependent discretizations is worked out for computing $\mathcal{A}_T(y)(t)$ for nearly linear problems

$$\partial_t y \;+\; \lambda^2 y \;=\; f(y,t) \tag{2.48}$$

with a small perturbation f.

On the other hand, indeed, there are *nonlinear* averaging techniques: In Russian (Celestial) Mechanics the *time averaging* idea (cf. $\mathcal{A}_T$) has been worked out for multifrequency systems using perturbation theory and/or resonance relations between the frequencies (cf. [1], [4], [37], and the work cited therein). But the Russian authors only pay attention to the running average *of the solutions* because they are dealing with problems from (Celestial) Mechanics. Thus, these techniques also are inappropriate for our kind of problem, again because of our smoothing problem (2.46).

Second, there are lots of investigations which propose *different discretization schemes* intended to efficiently deal with highly oscillatory behaviour. Altogether they are *linear* methods in the sense that they need a single, a-priori known frequency which dominates the oscillatory behaviour. Thus, these methods are only appropriate for nearly linear problems of type (2.48).

For these methods the restriction to the existence of a dominating frequency is already *necessary* for conceptual reasons. Two examples may demonstrate this:

One class of methods is based on construction principles like "principle of minimal phase lag" (cf. [8][9][44]) or "principle of coherence" (cf. [12][13]). These principles demand discretization schemes which "conserve the constant phase of the oscillation" or which are of the sort for which "successive approximations do not contradict each other". Both demands can only be formulated if there is something like a "constant phase" caused by the dominating frequency.

Since Implicit Runge–Kutta (IRK) methods are in some sense equivalent to Collocation Methods (cf. [21], p. 206) they exactly integrate algebraic polynomials of given degree. In analogy one can construct discretization schemes which exactly integrate appropriate *trigonometric* polynomials as it is done in [18], for example. But even in this case a main frequency λ must be known in advance and the method is bounded to problems corresponding to (2.48), too.

2.2.2 Damping Properties of Implicit Runge–Kutta Methods

It is known, from the analysis of stiffness and stability of ODEs, that some (Semi-)Implicit Runge–Kutta (IRK) schemes show damping properties for fast oscillations (cf. [22], p. 43). Can we use these damping properties for our purpose?

In Mechanical Engineering it is well known that some mechanical systems with highly oscillatory solutions and strong potential forces $F = -\frac{1}{\epsilon^2}\nabla U$ remain close to a manifold of "smooth" solutions. In [31], Ch. Lubich has shown that this manifold is associated with the limit $\epsilon \to 0$ and that specific IRK methods, using stepsizes larger than an ϵ_0, approximate such smooth solutions. Thus lower stepsize bounds in Implicit Runge–Kutta (IRK) methods can be used for efficiently solving these mechanical systems.

In some sense there is an analogy between this investigation and the AQRS–algorithm which will be presented later on: We can rewrite (2.13) as

$$\begin{aligned} i\partial_t b_k &= \frac{f(t)}{2}\sum_l V_{kl}\left(x^+_{kl} + x^-_{kl}\right) b_l, & b_k(0) &= c_k(0) \\ i\partial_t x^\pm_{kl} &= -\left(\omega_k - \omega_l \pm \omega\right) x^\pm_{kl}, & x^\pm_{kl}(0) &= 1 \end{aligned} \tag{2.49}$$

If you integrate (2.49) using e.g. the semi–implicit IRK method EULSIM with lower stepsize bound and exploiting its damping properties for large frequencies you obtain the "die out" of all frequencies $|\omega_k - \omega_l \pm \omega| > \mu$ and smooth solutions for b. Equivalently we can cancel the corresponding dipole elements:

$$V_{kl} := 0 \qquad \leftrightarrow \qquad |\omega_k - \omega_l \pm \omega| > \mu. \tag{2.50}$$

Indeed, this is very similar to the fundamental idea of AQRS (cf. Section 3). But we will neither construct nor discuss AQRS on the basis of this concept, because

- the blow–up of (2.13) to (2.49) is not necessary,
- in the class of problems (2.13) there is no analogous parameter ϵ which allows an asymptotical analysis like that given for mechanical systems,

- in the previous subsection we have seen that we may not ask for the smoothed coefficients b_k but for the smoothed populations p_{kk}.

2.2.3 Multirevolution Methods

For astronomers one key problem is the accurate computation of the long time behaviour of periodic orbits of celestial bodies including small perturbations of their exact periodicity by others. This problem has become more and more important since the long term behaviour of artificial satellites had to be computed. Originally, multirevolution methods [20][48] were introduced for use as fast integrators concerning the problem. But for these methods first, little theoretical justification has been attempted, and second, some physical a-priori knowledge about the orbits is needed. R. Gear and L. Petzold took some central ideas of multirevolution methods, constructed a more general *quasi–envelope* method, and gave a theoretical justification of it [19][38]. We will observe that this method is no competitor of AQRS — not in the class of problems (2.11) considered herein. But it should be explained shortly because it can be taken as a model for all methods which use information from one oscillation to jump over a few oscillations in one single integration step:

The central idea of the quasi–envelope method is shown in Figure 4: Assume the solution $y(t)$ of

$$y' \;=\; f(y,t), \qquad y(0) = y_0 \tag{2.51}$$

to be highly oscillatory but nearly periodic. In this situation one disclaims following each oscillation of y but integrates a quasi–envelope only. If the solution is nearly periodic, this quasi–envelope is of slow variation in time t and can therefore be integrated using very large time steps ("large" in comparison to those which must be chosen for an integration of y itself). The following inexact but short and formal introduction schematically explains how Gear and Petzold propose to compute the quasi–envelope of (2.51):

Assume that we know the evolution $T(t)$ of the "quasi–period" of y. Then we can use a substitution, which is implicitly given by

$$t(s+1) \;=\; t(s) \;+\; T(t(s)), \tag{2.52}$$

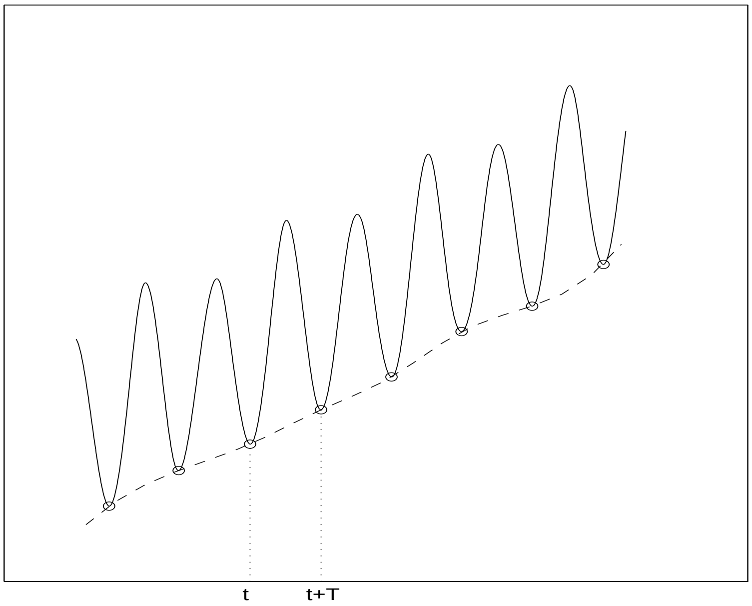


FIG. 4. For explaining the idea of the quasi–envelope: On an oscillating curve some points are marked which all have the same "phase of oscillation": they mark the beginning and end of each approximate period of the oscillation (period $\approx T$). The smooth line connecting them is called the quasi–envelope.

for transforming y to a nearly 1-periodic function $\tilde{y}$:

$$\tilde{y}(s) \;:=\; y(t(s)) \tag{2.53}$$

We find

$$\begin{aligned} \tilde{y}(s+1) - \tilde{y}(s) &= y\,(t(s) + T(t(s))) - y(t(s)) \;=:\; g(s, \tilde{y}) \\ t(s+1) - t(s) &= T(t(s)) \end{aligned}$$

which we can understand as equations of the explicit Euler discretization scheme for the ODE

$$\begin{aligned} \partial_s \tilde{y} &= g(s, \tilde{y}) \\ \partial_s t &= T(t(s)) \end{aligned} \tag{2.54}$$

if $\tilde{y}$ is slowly varying with s (slow variation in each interval $[s, s+1]$). The solution $\tilde{y}$ of (2.54) is called the quasi–envelope of y and only (2.54) is integrated. Each evaluation of $g(s,\cdot)$ requires the integration of (2.51) with initial condition $y(t(s)) = \tilde{y}(s)$ over one period $T(t(s))$. But if $\tilde{y}$ varies very slowly in s we can expect stepsizes $\Delta s \gg 1$. If we start at $t = 0$ with $y(0) = y_0$ and need $y_1 = y(t_1)$ so that $[0, t_1]$ includes many ($m \gg 1$) periods of y we should

1. Compute $\tilde{y}(m)$; this requires $k \ll m$ g–evaluations, i.e. integrations of (2.51) over one period.
2. Integrate (2.51) with initial condition $y(T(t(m))) = \tilde{y}(m)$ in $[T(t(m)), t_1]$.

Thus, the quasi–envelope method will be more efficient (i.e. k will be increasingly smaller than m) the more slowly $\tilde{y}$ varies with s. Surely one can observe that the origin of this method is somewhat like computation of satellite orbits: in these problems, $\tilde{y}$ is of very slow variation.

Inapplicability of the method to our kind of problem. It is easy to observe that any trajectory which is *oscillating but not nearly periodic* will cause difficulties for the method. These difficulties can be removable or not — this depends on the character of the oscillations. In the best case, there is a quasi–period $T(t)$ which is not nearly constant (not $T(t) \approx T$) but the corresponding quasi–envelope $\tilde{y}$ is slowly varying. Then the method will work at a high level of efficiency *only if* the determination of $T(t)$ causes only small computational effort. In the worst case, y is highly oscillatory but there is no quasi–period $T(t)$ which makes $\tilde{y}$ a slowly varying function. Then the method looses its gain in efficiency because there are too many evaluations of g.

Figure 1 (see page 21) demonstrates that the solution of our test problem 'OH' is an example for the worst case: there is no quasi–period $T(t)$ which would lead to a reasonable, slowly varying quasi–envelope for the solutions shown in Figure 1; it would rather make sense to work with two different quasi–envelopes for this kind of solution. Indeed, the character of the oscillations of the solutions is similar to that shown in Figure 1 in all real life problems (2.11) which have come to my knowledge. Thus, unfortunately, also the quasi–envelope method is inappropriate for our kind of problem.

2.2.4 Fourier–Galerkin Method

If there is multifrequency oscillatory behaviour, the use of the Fourier transformation (FT) often is a good choice. Indeed, there is some work applying FT to nonlinear and nonautonomous ODEs with periodic solutions [45][46]. For explaining this idea it should directly be applied to our basic equation (2.11). Therefore we first must symmetrically extend the solution c of (2.11) from $[0,\tau]$ to $[0,2\tau]$ for making it a periodic function. For simplicity assume (2.29) and $\omega = m\eta = m\,\pi/\tau$ with $m \in \mathbb{N}$. Now insert the Fourier representation of c

$$c(t) \;=\; \sum_{k\in\mathbb{Z}} \gamma_k \,\exp(ik\eta t) \tag{2.55}$$

into (2.11) and you get an infinite but linear system of equations determining the coefficients $\gamma_k \in \mathbb{C}^n$. The γ_k are the expansion coefficients of c in the orthonormal basis $\exp(ik\eta\cdot)$ of the space $\mathcal{P}_{2\tau}$ of 2τ–periodic L^1–functions. Hence we speak about a *Galerkin Method* if we drop all γ_l with $|l| \geq N$ in the computed linear equation, thus approximating c in the subspace

$$\mathcal{P}_{2\tau}^N \;:=\; \mathrm{span}\{\exp(ik\eta\cdot), \quad k \in \mathbb{N}, |k| < N\}.$$

The result is a *finite* system of linear equations

$$\mathcal{F}\,\gamma \;=\; \beta \tag{2.56}$$

with $\gamma = (\gamma_{1,1},\gamma_{1,2},\ldots,\gamma_{1,n},\gamma_{2,1},\ldots,\gamma_{N,n})^T \;\in\; \mathbb{C}^{Nn}$ and a complex $Nn \times Nn$ matrix $\mathcal{F}$. In [45] it is shown that (2.56) has an unique solution $\bar{\gamma}^{(N)}$ and it is discussed how the Galerkin approximation

$$\bar{c}^{(N)}(t) \;=\; \sum_{|k|\leq N} \gamma_k \,\exp(ik\eta t) \tag{2.57}$$

approximates the exact solution c.

If c always is close to a “smooth” curve $\tilde{c}$ you can expect that $\bar{c}^{(N)}$ already approximates $\tilde{c}$ for small N. Unfortunately, in our case we must not smooth the solution c of (2.11) but only the corresponding populations. Test calculations have shown that we definitely have to consider the frequencies up to 2ω to get a sufficiently good approximation of the populations. Because of Remark 2 it is $\omega/\eta \doteq 10^2$ (order of magnitude) and we have to take at least $N \doteq 10^2$.

However, this FT–Galerkin Method obviously needs $\mathcal{O}\left((nN)^3\right)$ operations in comparison to AQRS which only needs $\mathcal{O}(kn^\alpha)$ operations with an $\alpha \in [1,2]$ and with the number of integration steps $k \doteq 10^3$.

Thus our conclusion is: this Fourier–Galerkin method is appropriate in principle, but inefficient with respect to our problems.

2.2.5 Quasiresonant Approximations

The quasiresonant approximation (often called *rotating wave approximation*) has received a considerable amount of attention in physics and chemistry. Herein only the work of M. Quack is mentioned (cf. [41] and [42] and the references cited therein). Quack has applied the more exact term *weak field quasiresonant approximation* (WFQRA) to this method. WFQRA was discussed as an efficient method for solving equation (2.11) in the case of pure coherent, monochromatic light, i.e. for constant $f(t) \equiv E_r$. The following short introduction to WFQRA demonstrates that it should be seen as a purely heuristical method with less theoretical justification:

With $D_{kl} := \omega_k - \omega_l - \operatorname{sgn}(\omega_k - \omega_l)\omega$ we can rewrite the interaction picture version (2.13) of our basic equation as

$$i\,\partial_t b_m \;=\; \frac{1}{2} f(t) \sum_{k=1}^{n} V_{mk}\, b_k \;\exp(i\, D_{mk} t)\;(1 \,+\, \exp(2\,i\operatorname{sgn}(D_{mk})\,\omega t)), \quad (2.58)$$

From this we come to

$$i\,\partial_t b_m \;=\; \frac{1}{2} f(t) \sum_{k=1}^{n} V_{mk}\, b_k \;\exp(i\, D_{mk} t), \qquad (2.59)$$

by neglecting the high frequency term $\exp(2\,i\operatorname{sgn}(D_{mk})\omega t)$. According to [42] this is valid (in the sense of getting a “good approximation”) if the two conditions

$$f(t)\, V_{mk} \;\ll\; \omega, \qquad (2.60)$$
$$D_{mk} \;\ll\; \omega \qquad (2.61)$$

are fulfilled.

We can introduce the *level scheme integer* N_k by defining it as that integer which fulfills

$$\omega_k \;=\; N_k\omega \,+\, x_k \quad \text{with} \quad -\frac{\omega}{2} < x_k \leq \frac{\omega}{2}. \tag{2.62}$$

Hence equation (2.59) gets the form

$$\begin{aligned} i\,\partial_t b_m \;=\; & \frac{1}{2} f(t)\,\Big(\sum_{\substack{k \\ |N_m - N_k| = 1}} V_{mk}\, b_k\; e^{i\,(x_m - x_k)t} \\ & + \sum_{\substack{k \\ |N_m - N_k| = 0,2}} V_{mk}\, b_k\; e^{i\,(x_m - x_k)t}\; e^{\pm i\,\omega t} \\ & + \text{ terms with } e^{\pm m\,i\,\omega t}, m = 2, 3, \ldots \Big). \end{aligned} \tag{2.63}$$

Now a second approximation is made by removing all terms with $|N_m - N_k| \neq 1$. Following [42] again, this shall approximately be valid if the conditions (2.60) and (2.61) are fulfilled. Let us do this approximation and write

$$i\,\partial_t b_m \;=\; \frac{1}{2} f(t) \sum_{k=1}^{n} B_{mk}\, b_k\; \exp(i\,(x_m - x_k)t) \tag{2.64}$$

with the sparse dipole matrix

$$B_{mk} \;=\; \begin{cases} V_{mk} & : \; |N_m - N_k| = 1 \\ 0 & : \; \text{otherwise} \end{cases} \;. \tag{2.65}$$

Finally this can be brought into an interesting form by use of the unitary transformation

$$X \;:=\; \operatorname{diag}(x_1, \ldots, x_n) \tag{2.66}$$

$$a \;:=\; \exp(-i\,X\,t)\; b. \tag{2.67}$$

We obtain a system

$$i\,\partial_t a \;=\; (\,X \;+\; \frac{1}{2} f(t)\, B\,)\; a \tag{2.68}$$

which in the case $f(t) \equiv E_r$ is a system with constant coefficients.

For some cases this is a very useful and reasonable method. But condition (2.61) is nearly *never* fulfilled. This reminds us that we should use WFQRA carefully. Indeed, WFQRA is successful in some of the herein considered applications but obtains totally wrong results in most cases. Moreover, if successful it is less efficient in comparison to AQRS (for details see [43]).

2.2.6 Floquet–Analysis

As an introduction let us state a specialized version of the general Floquet–Theorem:

THEOREM 1. *Let be $T \in \mathbb{R}^+$ and let $A(t) \in \mathbb{R}^{n\times n}$ be symmetric for all $t \geq 0$ and T–periodic. In addition let $U(t, t_0)$ be the time propagation semi–group of*

$$i\,\partial_t c \;=\; A(t)\; c \tag{2.69}$$

(i.e. $c(t) = U(t, t_0)c_0$ is a solution of (2.69) for each $c(t_0) \in c_0 \in \mathbb{C}^{n\times n}$). Then there exists a T–periodic function $P : \mathbb{R}^+ \to \mathbb{C}^{n\times n}$ and a symmetric matrix $G \in \mathbb{R}^{n\times n}$ with

$$U(t, 0) \;=\; P(t)\; \exp(-i\,G\,t), \qquad \forall t \geq 0 \tag{2.70}$$

As a consequence of the periodicity of P we find

$$\exp(-i\,G\,T) \;=\; U(T, 0) \tag{2.71}$$

so that the spectrum of G can easily be evaluated by diagonalization of $U(T, 0)$, but only modulo $2\pi/T$! Assume that ε_k are the eigenvalues of G computed from (2.71). The evaluation can be made definite by the condition $\varepsilon_k \in [-\pi/T, \pi/T]$. These definite eigenvalues $\varepsilon_k \in [-\pi/T, \pi/T]$ are called *Floquet quasienergies* of (2.69).

This is the starting point of the ideas of Holthaus et al. [5][7][6][29] for an analysis of the structure of the above mentioned optimization task (cf. page 20). It is useful to follow their thoughts through a few details:

Let us discuss (2.11) with the specific form function $f(t) \equiv F$

$$i\,\partial_t c \;=\; (\Omega \,+\, F\;\cos(\omega t)\, V_\infty\,)\; c, \tag{2.72}$$

remember $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_n)$ and let us denote the quasienergies of (2.72) as $\varepsilon_k(F)$. For $F = 0$ we simply find

$$(\varepsilon_k(0) \;=\; \omega_k)\;\bmod\;\omega. \tag{2.73}$$

One can find an interval $\mathcal{F} = [0, F_0]$ in which each state k is uniquely related to a curve $\varepsilon_k : \mathcal{F} \to \mathbb{R}$, $F \mapsto \varepsilon_k(F)$ with $(\varepsilon_k(0) = \omega_k) \bmod \omega$.

Let us now fix two states k, l as initial state (k) and final state (l) and let us choose the laser frequency

$$\tilde{\omega} \ := \ \frac{\omega_k \ - \ \omega_l}{m} \qquad \text{for a certain } m \in \mathbb{Z}. \tag{2.74}$$

We can determine m in a way that makes $\varepsilon_k(0) = \varepsilon_l(0)$ degenerate quasienergies. For this case and under some specific assumptions Holthaus et al. have found the following result for the solution $c^f(t)$ of

$$i\,\partial_t c \ = \ (\Omega \ + \ f(t) \ \cos(\tilde{\omega} t)\, V_\infty \,) \ c, \qquad c_j(0) \ = \ \frac{1}{\sqrt{2}}(1+i) \ \delta_{kj} \ \ \forall j. \tag{2.75}$$

They showed that

$$|c_l^f(\tau)|^2 \ = \ \sin^2\left(\frac{1}{2}\int\limits_0^\tau \ (\varepsilon_k(f(t)) \ - \ \varepsilon_l(f(t))) \ dt\right). \tag{2.76}$$

The most central of their specific assumptions is

$$\max_{t\in[0,\tau]} |f(t)| \ \le \ F_0 \tag{2.77}$$

which guarantees the knowledge for computing (2.76). Thus, equation (2.76) directly gives the population of the target state l, if the curves $\varepsilon_k(\cdot)$ and $\varepsilon_l(\cdot)$ and the value F_0 are already evaluated. In this case also the optimization problem is simply solved by determining f due to

$$\int\limits_0^\tau \ (\varepsilon_k(f(t)) \ - \ \varepsilon_l(f(t))) \ dt \ = \ \pm\,\pi. \tag{2.78}$$

Surely all this is interesting for analytical reasons (interpretation of the optimization problem as a search for constructive resonances between Floquet states), but: What can we learn from this previous investigation about the comparative efficiencies of Floquet Analysis and AQRS? To answer this question sufficiently we have to name some properties of the AQRS–algorithm in anticipation of its concrete construction. The answer is split into four items because we have to distinguish between efficiency with respect to the optimization task and with respect to an inner problem:

- For the determination of the curves $\varepsilon_j(\cdot)$ an evaluation of $U(T,0)$ is necessary in some points F_m. Each computation of $U(T,0)$ for a given F requires the solution of the ODE

$$i\,\partial_t U(t,0) \;=\; (\Omega \,+\, F\;\cos(\tilde{\omega}t)\;V_\infty)\;U(t,0), \quad U(0,0) \;=\; 1 \tag{2.79}$$

  in $[0,T]$, an ODE with dimension n^2 ($n =$ number of states). In contrast to this AQRS needs the integration of a sparsed version of (2.11) in the interval $[0,\tau]$ to determine $|c_l(\tau)|^2$ for given (f,ω). For picosecond laser pulses τ/T is of the order of 10^2 (see Remark 2).

  Thus, for large n a determination of $U(T,0)$ is tremendously inefficient compared with an AQRS–integration on $[0,\tau]$.

  For small n the fact should be stated that an AQRS–integration on $[0,\tau]$ results in about the same computational effort as an integration of (2.11) on $[0,T]$ with an efficient but usual integrator. For this reason the evaluation of $U(T,0)$ ($\hat{=}\, n$ integrations of (2.11) on $[0,T]$) again is more expansive than any AQRS–integration.

  This means that AQRS is much more efficient than Floquet Analysis with respect to the solution of a single inner problem.

- Can we, from this statement regarding the higher efficiency of AQRS in solving a single inner problem, make the further step in showing that AQRS is also more efficient in the solution of the optimization task? If the form of the laser shape f is fixed in some way, e.g.

$$f(t) \;=\; E_r\,\sin^2(\eta t), \qquad \text{with given } \eta = \frac{\pi}{\tau}. \tag{2.80}$$

  the optimization problem is based on a low–dimensional space (here one–dimensional: variation only in E_r, $\tilde{\omega}$ is fixed) and only a few inner problems must be solved, i.e. only a few AQRS–integrations are needed for determining the optimum. Indeed, this is the fact in all investigations concerning the optimization problem in real life applications (see Section 5) which have come to my knowledge. But, particularly for large n, these few AQRS–integrations surely are cheaper than the number of $U(T,0)$–computations which are at least necessary in the Floquet method for construction of the curves $\varepsilon(\cdot)$.

Thus, AQRS also is more efficient than Floquet Analysis with respect to the solution of optimization problems with fixed laser shapes.

- If you want to extend the optimization problem to the study of the effect of various different form functions, Floquet Analysis would be a wonderful tool: after the evaluations of the curves $\varepsilon_j(\cdot)$ the effect of the variations can be studied without any additional integrations. Thus, this extended task can be solved more efficiently by use of Floquet Analysis than by use of AQRS.
- Later on we will see that the scope in which AQRS is of high efficiency is limited by the condition that $\max_{[0,\tau]}|f(t)|$ may not be too large. Exactly the same limitation holds for the use of Floquet Analysis in the form of condition (2.77). But since the Floquet Analysis presented herein simply doesn't work at all if this condition is broken, AQRS may be regarded, in comparison, to still work well, although with a smaller gain in efficiency.

2.2.7 Concluding Remarks

Many of the herein collected competing methods are inappropriate for application to our kind of problem, in particular all linear smoothing techniques, all proposed specific discretization schemes, the WFQRA method and the quasi–envelope method. Others are appropriate in principle, but not efficient enough: the Fourier–Galerkin method, the damping method from Subsection 2.2.2 and the Floquet method (at least for the inner problem and for the "customary" optimization problem).

So, there is a real need to develop a new approach in view of both, efficiency and reliability with respect to the problems considered herein.

3 Basic Ideas of Quasiresonant Smoothing

The following Section introduces the basic concept of "quasiresonant smoothing". The subsections follow the way this idea was initially found. This way joins the effects of two basic ideas

1. The solutions of the considered class of problems can be characterized as highly oscillatory. We ask for an integration tool which solves (2.11) or (2.13) with a remarkable gain in computational efficiency. Thus, we must look for an algorithm which computes only *smooth* solutions, i.e. solutions without "fast" oscillations. This smoothness would allow an integrator to make much larger time steps.

2. A basic physical interpretation of the dynamics described by (2.11) leads to the expectation that a certain *sparsing process* of the considered dipole matrix V_∞ will not change the *main characteristics* of the considered solution. These main characteristics appear to be the *averaged populations* (the running average of $\operatorname{diag} p$).

It is demonstrated that these basic ideas are linked by the fact that, indeed, the named sparsing process also leads to the required smoothness of the solutions: A well–controlled sparsing process causes a damping of the high frequency oscillations and leads to smooth solutions from which the averaged populations can be computed. The introduction of a *sparsing* or *damping* parameter δ allows the construction of a class of basic algorithms QRS(δ). For this Section the pursued way ends in a study of the properties and problems of QRS(δ).

3.1 Reduction of Computational Complexity by Smoothing

In this section we consider the task of computing the populations of the dynamical equation (2.11) (or (2.13)) for given data V_∞, Ω, ω, and f (i.e. the "inner problem" from page 19). Let us assume, that we want to solve this equation by use of a numerical integrator. So our first question is: *How can we find a computationally cheap solution of our inner problem using integrators.*

To answer this we may take a closer look to Fig. 5 on page 44, which shows data from an accurate solution of (2.11) for our real life test problem 'OH' (cf. Section 2.1.4). All over Section 3 test problem 'OH' is used to explain the basic pattern of QRS.

On the left top of Fig. 5 we see the population $p_{11} = |c_1|^2$ of the ground state ($k = 1$) dying out and the evolution of the desired full occupation $p_{55}(\tau) = 1$ of the 5th state. On the left bottom of Fig. 5, as an example for the trajectories of the solution of the considered problem, the imaginary part of the first expansion coefficient $c_1(t)$ is shown. Most of the high computational complexity of numerically computing this solution is caused by the oscillations with high frequencies and large amplitudes: they force the integrator to choose very small time steps.

Physically, namely for the class of real life applications considered herein, the knowledge of the *running average* $\mathcal{A}_T p_{kk}$ of the populations $p_{kk} = |c_k|^2$ would be sufficient. This average may be defined by introducing the *average–operator* $\mathcal{A}_T$:

DEFINITION 1. *For integrable functions $f : [t_0, t_1] \subseteq \mathbb{R} \to \mathbb{C}$ and a $T > 0$ the (time) average operator is defined by*

$$(\mathcal{A}_T f)(t) \;=\; \frac{1}{T} \int\limits_{t-T}^{t} f(s)\, ds, \qquad \forall t \in [t_0 + T, t_1] \tag{3.1}$$

if $t_0 + T \leq t_1$. For matrix–valued, componentwisely integrable maps $M : [t_0, t_1] \to \mathbb{C}^{m \times n}$ the corresponding average operator is defined simply componentwise.
Let $c(t)$ be the solution of (2.11) with given data V_∞, Ω, ω, and f and $p(t)$ the corresponding probability matrix according to (2.15). In addition let c and p be extended to $t < 0$ by

$$c_k\big|_{t<0} \;:=\; \frac{1}{\sqrt{2}}(1+i)\,\delta_{ks} \qquad \forall k \in \{1 \ldots n\} \tag{3.2}$$

(s: initial state) and the corresponding extension of p. The data from the solution of (2.11), in which we are physically interested only, is given by

$\operatorname{diag}\pi(t)$ *where* π *is the smoothed probability matrix*

$$\pi(t) \;:=\; (\mathcal{A}_T\, p)(t) \;=\; \frac{1}{T}\int\limits_{t-T}^{t} (c \otimes \bar{c})(s)\, ds \tag{3.3}$$

with T *big enough in comparison with the fast oscillation's periods.*

The effect of this averaging can again be seen in Fig. 5: At the right top we see a section from the left top picture showing $p_{55}(t)$ and its local time average $\mathcal{A}_T p_{55}(t)$; at the right bottom the smoothed probabilities π_{11} and π_{55} corresponding to p_{11} and p_{55} are shown.

But let us return to our question of how to find an appropriate integrator for solving the inner problem efficiently. We want to construct an algorithm which computes the averaged populations $\pi_{kk} = \mathcal{A}_T p_{kk}$ instead of the oscillatory populations p_{kk}. We will do this in a way which can be described by the following reformulation of the inner problem:
Replace equation (2.11) by a similar one for which the populations $q_{kk}(t)$ *(computed from its solution) are good approximations of the averaged populations* $\pi_{kk}(t)$ *of (2.11).*
Later on this 'similar one' will be called *smoothed equation*. We should call q_{kk} a good approximation of $\pi_{kk} = \mathcal{A}_T p_{kk}$ if at least for some norm:

$$||\mathcal{A}_T q_{kk} \;-\; \mathcal{A}_T p_{kk}|| \;<\; \text{tol}, \tag{3.4}$$

for a given tolerance $\text{tol} > 0$.

But there is a *central smoothing difficulty*: Using an integrator for solving the smoothed equation we can only realize the expected gain in efficiency if one condition is fulfilled: the solution of the smoothed equation must be smooth, i.e. it must allow large stepsizes of the integrator. Maybe, one firstly assumes that these smooth solutions should be the averaged solutions $\mathcal{A}_T c$ of (2.11). But unfortunately, the solutions of the smoothed equation must be different from $\mathcal{A}_T c$ *if we want to approximate* $\mathcal{A}_T |c_k|^2$, because in general we have

$$\mathcal{A}_T\, |c_k|^2 \;\neq\; |\mathcal{A}_T\, c_k|^2. \tag{3.5}$$

An impressive example for this inequality is given in Figure 6 (see page 45). For a simple *formal* example set $c_k(t) = \exp(i\eta t)$. Then, we observe

$$(\mathcal{A}_T|c_k|^2)(t) \;=\; 1 \;\neq\; \operatorname{sinc}^2(\frac{\eta T}{2}) \;=\; |\mathcal{A}_T\, c_k|^2(t).$$

Thus, we can put our task in a more concrete form:
Construct a smoothed equation with a smooth solution $d(t)$ so that the corresponding populations $q_{kk} = |d_k|^2$ are good approximations of the populations p_{kk} of (2.11) in the sense of (3.4).

In addition to all this we may remember the fact that the solution b of the interaction picture version (2.13) of equation (2.11) can also be used to compute the required probabilities:

$$\operatorname{diag} p(t) \;=\; \operatorname{diag}\Big((b \otimes \bar{b})(t)\Big) \tag{3.6}$$

As a consequence of this we can also try to find a smoothed equation to (2.13) instead to (2.11). Again this leads to the same central smoothing difficulty, but, in fact, it has got some advantages.

For simplicity let us make the following agreement: The averaged or *smoothed* outlines of the basic magnitudes are denoted using the Greek letter instead of the Latin letter which denotes the original magnitude. Hence we write $\gamma = \mathcal{A}_T c$, $\beta = \mathcal{A}_T b$, or $\pi = \mathcal{A}_T p$.

In the next section a *heuristical* solution of the reformulated problem will be presented.

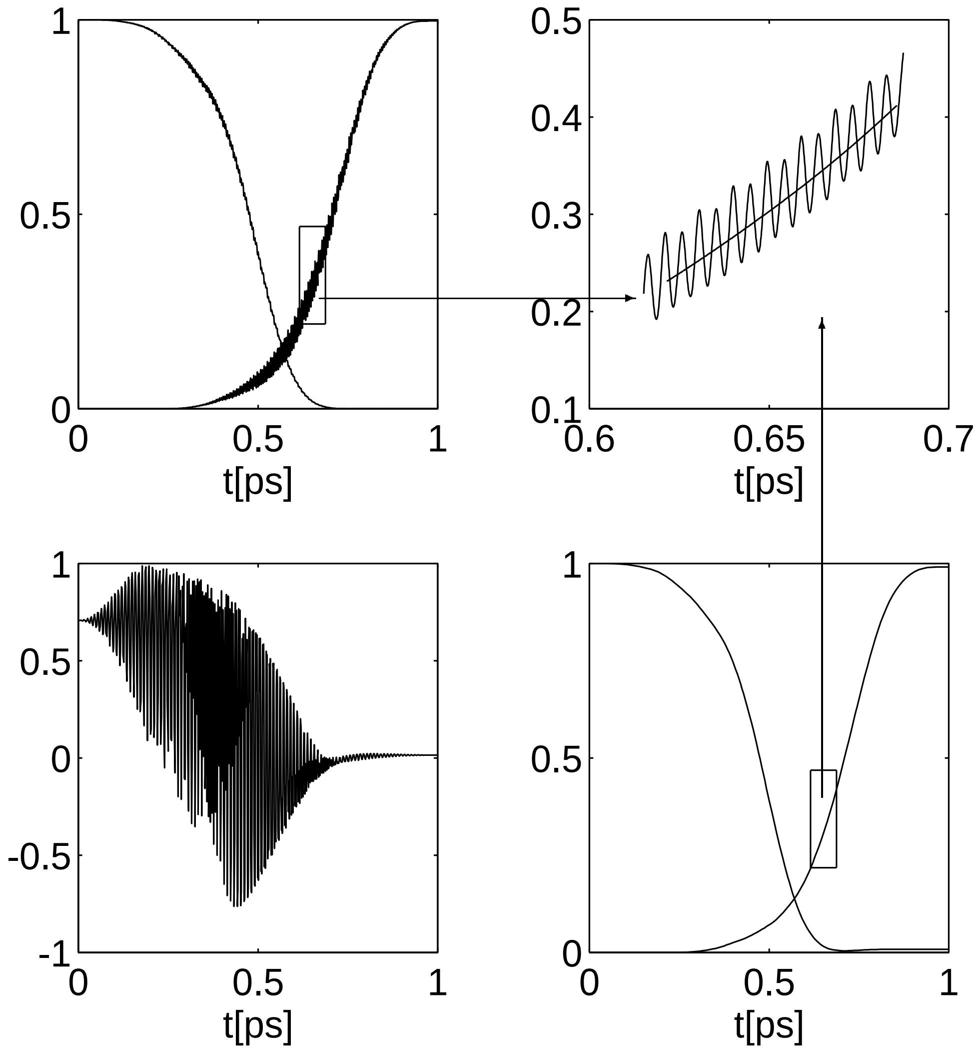


FIG. 5. For explaining the smoothing idea, data from the solution of our test problem 'OH' are shown here: At the left top two populations corresponding to the exact solution of the considered problems are shown (drawn against time in picoseconds). The right top picture shows a section of one of these populations and its local time average. For showing one of the really computed trajectories of the considered problem at the left bottom the highly oscillatory imaginary part of the corresponding expansion coefficients c_1 is drawn. The right bottom picture shows the averaged outlines of the populations from the left top one.

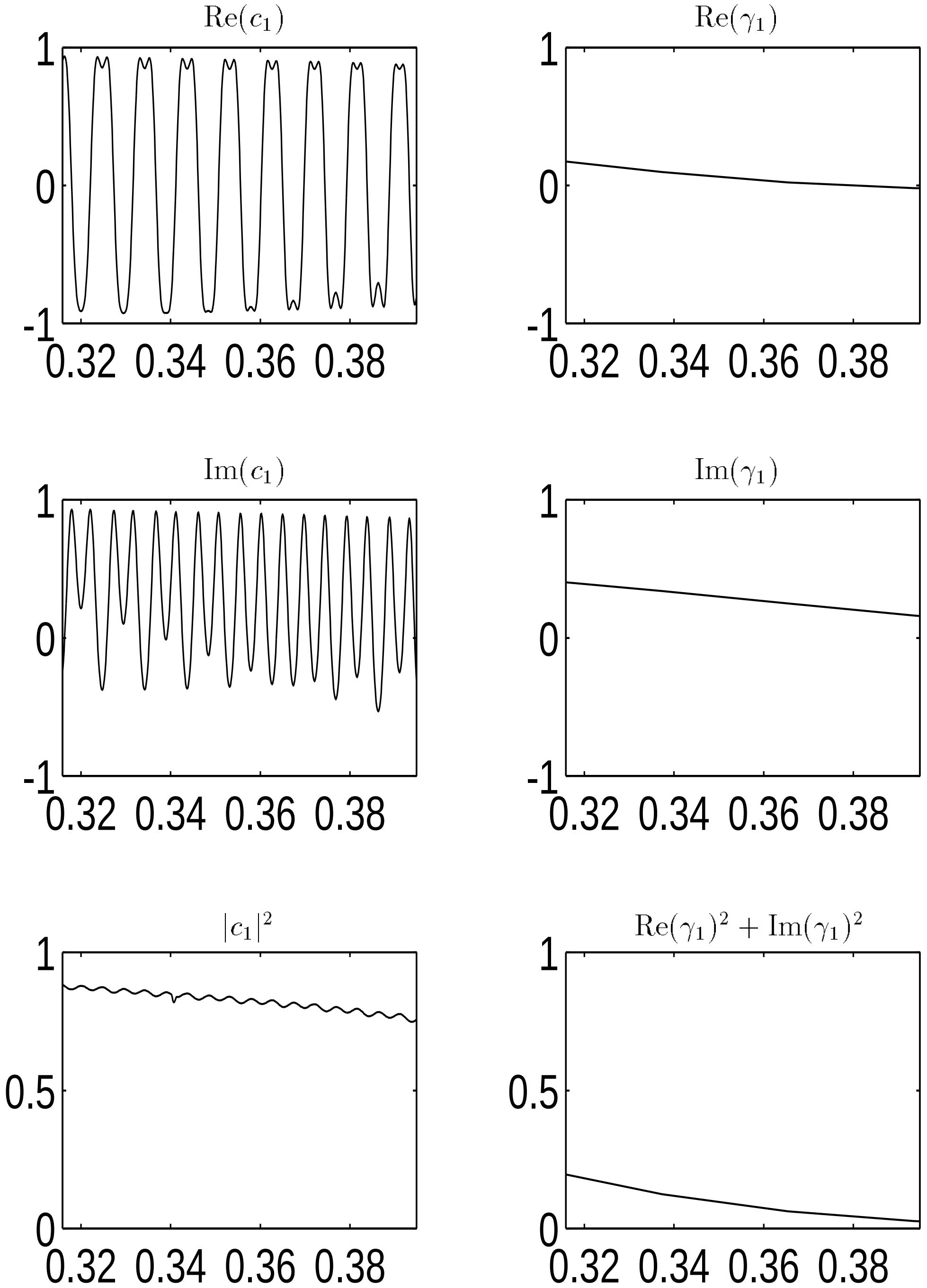


FIG. 6. This Figure demonstrates that it cannot be sufficient to smooth the coefficients directly. It shows trajectories from the test problem 'OH': On the left side from top to bottom the coefficients $\mathrm{Re}(c_1)$, $\mathrm{Im}(c_1)$ and the corresponding population $p_{11} = |c_1|^2$ of the exact solution of the problem. On the right side the corresponding smoothed coefficients $\mathrm{Re}(\gamma_1)$ and $\mathrm{Im}(\gamma_1)$ and the smoothed population $\mathrm{Re}(\gamma_1)^2 + \mathrm{Im}(\gamma_1)^2$ which we would compute from them. Note the essential difference between the both populations.

3.2 Physical Motivation of QRS

There are several ways to give a *formal* heuristic which leads to the basic idea of QRS. Herein the shortest one is used. It is presented only for the reason of introducing the method and will not appear in the later sections. In the first part the reason for the name "quasiresonant" cannot be seen. It results from the physical interpretation which closes this section.

Cutting smoothers We consider functions from

$$L \ := \ \{f \,:\, R \to R \mid f(t) = \sum_k \alpha_k \, \exp(i\eta_k t) \text{ with } \eta_k \in R, \alpha_k \in C\}. \quad (3.7)$$

For simplification of the heuristical reasoning let us assume that in the following the considered functions are such sums of $\exp(i\cdot)$–terms from L. For $\mu > 0$ the operation

$$\begin{aligned} G_\mu \,:\, L \,\to\, L \qquad & \qquad \text{with} \\ f(t) = \sum_k \alpha_k \, \exp(i\eta_k t) \qquad & \mapsto \\ (G_\mu\, f)(t) \quad = & \quad \sum_{\substack{k \\ |\eta_k| \le \mu}} \alpha_k \, \exp(i\eta_k t) \end{aligned} \qquad (3.8)$$

can be called a 'cutting smoother'.

Heuristical smoothing. Using $\cos \omega t = \frac{1}{2}(e^{i\omega t} + e^{-i\omega t})$ and the definition

$$\Delta_{mk} \ := \ \omega_m - \omega_k$$

one can rewrite our dynamical equation (2.13) (interaction picture):

$$i\,\partial_t b_m \ = \ \frac{1}{2} f(t) \ \sum_{k=1}^{n} V_{mk}\, b_k \ \left(e^{i\,(\Delta_{mk}+\omega)t} \ + \ e^{i\,(\Delta_{mk}-\omega)t}\right). \qquad (3.9)$$

We are interested in the effect of changes in the dipole matrix. Therefore, we consider the following class of equations:

$$i\,\partial_t d_m \ = \ \frac{1}{2} f(t) \ \sum_{k=1}^{n} \left(A^+_{mk}\, e^{i\,(\Delta_{mk}+\omega)t} \ + \ A^-_{mk}\, e^{i\,(\Delta_{mk}-\omega)t}\right) d_k. \qquad (3.10)$$

If we assume $d_k \in L$, $d_k(t) = \sum_l \alpha_l^{(k)} \exp(i\eta_l t)$, one obtains from (3.10) by integration:

$$d_m(t) \;=\; d_m(t_0) + \frac{f(t)}{2} \sum_{k=1}^{n} \sum_l \alpha_l^{(k)} \left(A^+_{mk}\, I^+_{mk,l} \,+\, A^-_{mk}\, I^-_{mk,l} \right) \tag{3.11}$$

with

$$I^{\pm}_{mk,l}(t,t_0) \;=\; -\frac{i}{\eta_l + \Delta_{mk} \pm \omega} \, \exp(i(\eta_l + \Delta_{mk} \pm \omega)t)\big|_{t_0}^{t} \,. \tag{3.12}$$

We want the d_m to be smooth trajectories, i.e. $d_m = G_\mu d_m$ shall hold for a small μ. Then, at least, all $I^{\pm}_{mk,l}$ with $\Delta_{mk} \pm \omega \gg \mu$ must vanish. The simplest choice in order to realize this is

$$A^{\pm}_{mk}(\delta) \;=\; \begin{cases} V_{mk} & : \;\; |\Delta_{mk} \pm \omega| < \delta\,\omega \\ 0 & : \;\; \text{otherwise} \end{cases} \tag{3.13}$$

where $\delta \geq 0$ is a free parameter.

This observation already gives the rather simple idea of the QRS(δ) algorithm (Quasi–Resonant Smoothing):

- Choose δ heuristically by physical insight (see below).
- For this δ compute the solution d of (3.10) with data (3.13).
- take $|d_m|^2$ as an approximation for the smoothed populations $|b_m|^2 = |c_m|^2$ to be computed from (2.11) or (2.13).

Because of later advantages let us introduce another notation of the smoothed equations (3.10):

- write $V_\delta = A^+(\delta)$ to denote the origin of V_δ in V_∞ (at this point the notation V_∞ itself can be understood) and
- use the variables b_δ to denote the solution of (3.10) for given δ; this reflects that the solutions of (3.10) are discussed as approximations of the solutions of the basic equation (2.11) (respectively (3.9)).

Because of $V_\delta^T = A^-(\delta)$ we may write as another version of (3.10):

$$i\,\partial_t b_{\delta,m} \;=\; \frac{1}{2}f(t)\;\sum_{k=1}^{n}\left(V_{\delta,mk}\,e^{i\,(\Delta_{mk}+\omega)t}\;+\;V^T_{\delta,mk}\,e^{i\,(\Delta_{mk}-\omega)t}\right)\,b_{\delta,k}. \tag{3.14}$$

This is the "smoothed equation" of QRS(δ).

Interpretation. A basic physical insight into molecule–light interaction processes is the observation of so–called *transition conditions*: The interaction between states k and l can only be important for the evolution of the populations if the frequency $\tilde{\omega}$ fulfills the "resonant" *first order transition condition*

$$|\Delta_{mk}| \;\simeq\; \tilde{\omega}. \tag{3.15}$$

Each present frequency ω_j effects transitions according to the corresponding condition $|\Delta_{mk}| \simeq \omega_j$.

Let us now consider our case of a laser pulse with a form function f which is slowly varying in comparison to the "light–oscillation" $\cos\omega t$. The effect of f can be understood as a *splitting* of the inducing light frequency ω. Take e.g. $f(t) = E_r \sin^2(\eta t)$ with $\eta \ll \omega$. In this case ω is split into ($\omega \pm 2\eta$, $\omega \pm \eta$, ω). This splitting softens the hard first order transition condition

$$|\Delta_{mk}| \;\simeq\; \omega \quad\Leftrightarrow\quad |\Delta_{mk} \pm \omega| \simeq 0.$$

to a "quasiresonant" condition like

$$|\Delta_{mk} \pm \omega| < \delta\,\omega \tag{3.16}$$

with a small $\delta > 0$. δ has to be chosen in a way which reflects these connections between frequency–splitting and importance of single state-state–interactions.

Hence only those "interaction–elements" V_{mk} have to be considered, which belong to interactions fulfilling (3.16). Exactly this concept is realized in QRS(δ): the matrices $V_\delta = A^+(\delta)$ and $V_\delta^T = A^-(\delta)$ only consist of these *quasi–resonant* elements in V_∞.

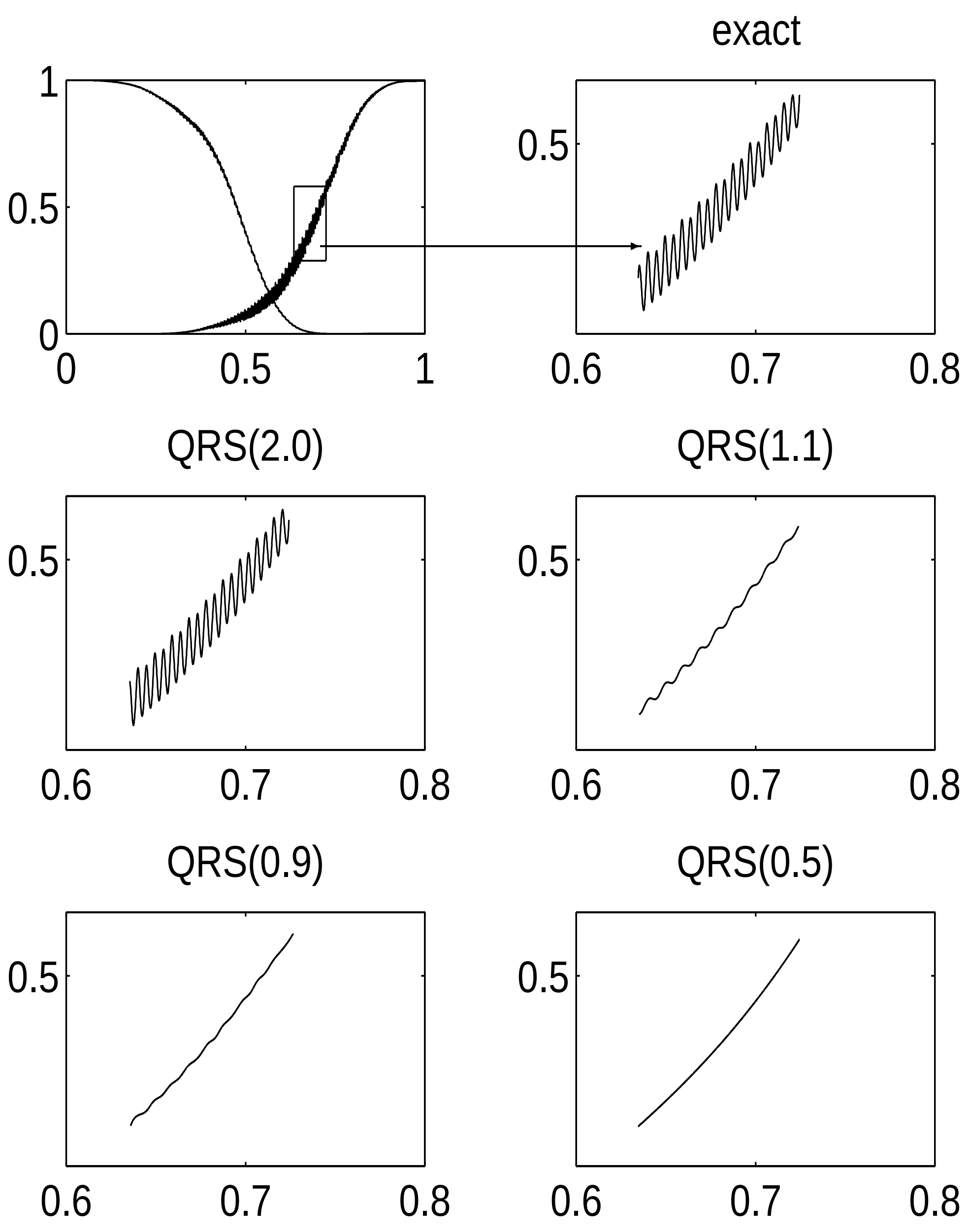


Fig. 7. QRS(δ)–solutions for different δ compared with the exact solution of the considered test problem. All subfigures are showing the same section of the p_{55}–trajectory introduced in Figure 1. Note the increasing damping of the high frequency oscillations with decreasing δ.

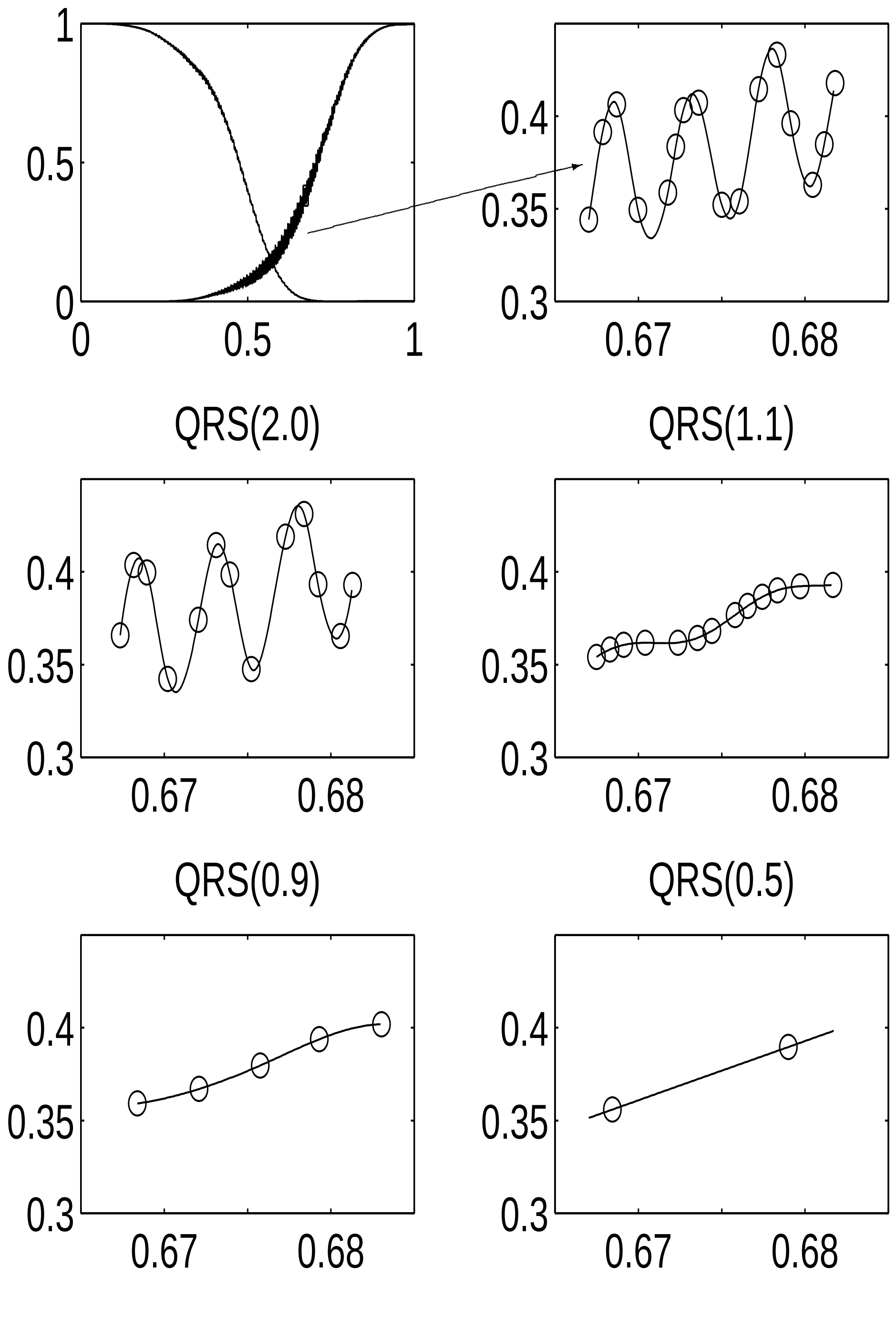


FIG. 8. Additional zoom of the p_{55}–sections shown in the previous figure for different QRS(δ). Each single time step of the algorithms is marked by an 'o'. Observe the increase in the stepsize with decreasing δ.

3.3 QRS Algorithm with fixed Sparsing Parameter δ

A few lines above the class of QRS(δ)–algorithms has been introduced. In this section some of its basic properties are presented:

Figure 7 above shows QRS(δ)–solutions for different δ compared with the exact solution of the considered test problem 'OH'. All six subfigures of Figure 7 are showing the same section of the p_{55}–trajectory introduced in Figure 1. The QRS($\delta = 2.0$)–population is very similar to the exact solution for p_{55} whereas the QRS($\delta = 0.5$)–population approximates its smoothed outline $\mathcal{A}_T\, p_{55}$. In general we observe an increasing damping of the high frequency oscillations with decreasing δ.

In addition it must be said that all QRS($\delta < 0.2$) obtain wrong results while QRS($\delta = 0.2$) still works well. Obviously, the QRS(δ)–algorithms reproduce a (more or less) smooth approximation of the exact averaged populations until δ undergoes a critical value δ_0. According to the above given physical interpretation one may assume that for $\delta < \delta_0$ the sparsed dipole matrices $A_\delta^\pm$ (respectively V_δ) from equation (3.13) do not include the most important state–state–interactions furthermore.

For the implementation of all QRS(δ) algorithms a highly efficient numerical integration tool with an adaptive order and stepsize control should be used. It automatically adapts the stepsizes to the obtained smoothness of the solution of the current δ-smoothed equation (3.14). In the scope of this work this "background–integrator" is always basing on *extrapolation methods* due to [14], [15]: all presented QRS–realizations are using the C++–implementation of the extrapolation code EULEX described in detail in [25] (for a C–version of EULEX see [26]).

From the definition (3.13) of the sparsed dipole matrices $V_\delta = A^+(\delta)$ we easily can see that QRS($\delta = \infty$) is nothing more than the used standard background–integrator itself. Hence, QRS($\delta = \infty$) produces an approximation of the *exact* solution of (2.13) up to the accuracy tol which is chosen in the used EULEX–tool (tol $= 10^{-5}$ in all following cases). For simplification we call the QRS($\delta = \infty$)–solution "exact" in the following.

Definitely, we have started the discussion of the QRS–idea for reducing the *computational complexity* of finding sufficient information about the population dynamics of the considered problem. For this purpose we need a complexity measure, which allows to compare the computational effort produced by QRS(δ) with that produced by QRS($\delta = \infty$). To prepare this note that the computational complexity of QRS(δ) mainly depends on two effects:

- the number of evaluations of the right hand side of equation (3.14) made in the whole integration process,
- the number of flops each single evaluation of the right side of (3.14) costs.

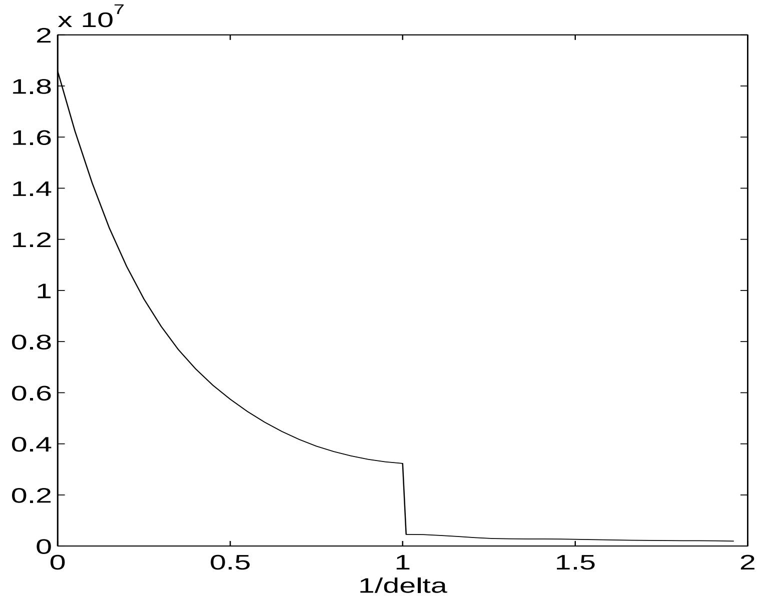


FIG. 9. Effort $C(\delta)$ versus $1/\delta$ for QRS(δ) on the considered test problem. The effort measure $C(\delta)$ is defined on page 53.

Hence let us assume that the different QRS(δ)–versions are basing on the same background–integrator and are controlled by the same given accuracy tol.

DEFINITION 2. *Let $P = (V_\infty, \Omega, \omega, f)$ denote the specific data of the problem ("solve the basic equation (2.11)") and I the considered integration interval. Let $N_P(\delta, I)$ be the number of evaluations of the right hand side of (3.14) which QRS(δ) needs for the considered computation. In addition let $\#(A)$ be the number of nonvanishing elements in a matrix $A \in \mathbb{C}^{n\times n}$. Now $\#(V_\delta)$ can be taken as a measure of the number of flops each evaluation of the right side of (3.14) costs. Hence the complexity of a QRS(δ)–integration on P, I is measured by*

$$C_{P,I}(\delta) \ := \ \#(V_\delta) \ \cdot \ N_P(\delta, I) \tag{3.17}$$

In the following we should simply write $C(\delta)$ and $N(\delta)$ if the choice of problem P and interval I is clearly fixed.

As a representative example Figure 9 shows the graph of $C(\delta) = C_{P,I}(\delta)$ versus $1/\delta$ fixing the problem P as our test–problem 'OH' and fixing I as the integration interval which corresponds to the pulse length of the considered laser pulse. We observe:

- $C(\delta)$ increases with $\delta \to \infty$ ($1/\delta \to 0$); for well–chosen δ a speedup factor of about 100 is observed:

$$\frac{C(\delta = \infty)}{C(\delta = 0.5)} \ \approx \ 100 \tag{3.18}$$

- $C(\delta)$ exhibits a remarkable jump at $\delta = 1$.

The first fact is not surprising whereas the second one should be explained: At $\delta = 1$ the diagonal elements of the dipole matrix V_∞ are set to 0 in V_δ. This has got a tremendous effect on the computed trajectories of the expansion coefficients b_δ. This is exemplarily demonstrated in Figure 10: At $\delta = 1$ the character of the shown QRS(δ)–solution changes from highly oscillatory to smooth. For this smooth trajectories big integration stepsizes are possible. (It should be stated that this effect is particularly large in the coefficient $\mathrm{Im}(b_{\delta,1})$ of the test problem 'OH' which is shown in Figure 10 for demonstration. In other coefficients the effect is merely similar to the slow damping process which is demonstrated in Figure 7.)

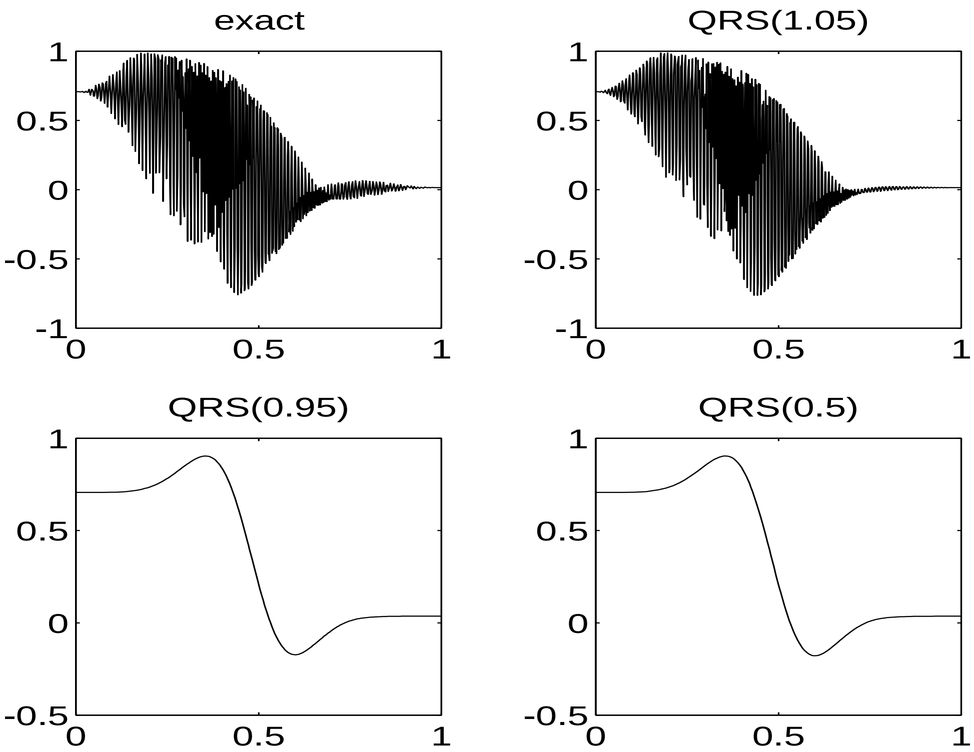


FIG. 10. Different QRS(δ)–trajectories for the imaginary part of the first expansion coefficient ($b_{\delta,1}$) for the considered test problem. While the QRS(1.05)–trajectory is still similar to the exact solution the one of QRS(0.95) demonstrates the tremendous smoothing effect which happens at $\delta = 1$.

Figure 10 demonstrates that our above *central smoothing difficulty* is removed in the case of 'OH': Both, the probabilities $|b_k|^2$ and the corresponding coefficients b_k are *simultaneously* smoothed by QRS(0.5). But 'OH' is only one example and therefore we ask now and answer later: *Is it always possible to smooth coefficients and probabilities simultaneously?* Only a positive answer to this question can lead to a reliable QRS algorithm.

4 Adaptive QRS Algorithms

In the early beginnings of numerical integration there were discretization schemes which were used with a fixed integration stepsize h in simple, uncontrolled integrators. Indeed, theoretical propositions about the convergence of these methods with $h \to 0$ have been known. But mostly there was no way to *guarantee* that a solution computed with a certain fixed h met the required accuracy. In addition it mostly was impossible to a-priori choose a stepsize h which would lead to a sufficiently accurate solution but which at the same time would not produce more computational effort than necessary.

In this situation the saving idea has been the idea of stepsize control in connection with a valid estimation of the integration error: After the kth step the stepsize h_{k+1} for the next step is automatically determined. Therefore an *error model* is constructed which is used to compute an estimation $\langle\epsilon(h)\rangle$ of the integration error $\epsilon(h)$ made in the $(k+1)$th step if h is used as stepsize. Then h_{k+1} is chosen due to the required accuracy ($\langle\epsilon(h_{k+1})\rangle < \mathrm{tol}$).

The above constructed QRS(δ) methods have got a structurally analogical problem:

- In QRS(δ) the (sparsing) parameter δ is fixed during the whole integration process; theoretically it is known that for $\delta \to \infty$ the result will converge to the exact solution,
- the smaller δ can be chosen the bigger is the gain in efficiency,
- if you use a small δ there is no way to *guarantee* that a QRS(δ)–solution meets the accuracy requirement.

We use this analogy $\delta \leftrightarrow h$ for constructing an *adaptive* δ–control in connection with a valid error estimation.

Inspired by the preceding, my proposal for this control scheme starts from the following items

- divide the whole integration interval into several subintervals $I_k = [t_k, t_k + H_k]$, on which the smooth laser shape function f from equation (2.11) is nearly constant,

- use QRS(δ_k) for the concrete integration of (2.13) on each subinterval I_k (the "kth step"),
- determine the sparsing parameter δ_{k+1} (used for the $(k+1)$th step) with respect to a given accuracy which the solution should meet. For this purpose construct an error model which allows us
  - to reliably estimate the integration error made on I_{k+1} using QRS(δ_{k+1}) and
  - to compute this error estimation cheaply,
- and care about the effect of the initial choice of δ_1 on the control mechanism — if necessary construct an automatic method to make a good initial choice.

This should be the program of the next sections. Executing it, the central task is getting sufficient but cheap a-priori information about the solution on the interval I_{k+1} after having done the kth step. Herein, this key problem is solved by studying a perturbation theoretical analysis of equation (2.11): perturbation theory of the effect of the dipole matrix on the interaction between the states. The basic idea is to use the first order correction of perturbation theory as an approximative solution on I_{k+1}. This approximation shows to be sufficient for studying the effect of δ–variations on the future integration error.

Thus, the proposed error model depends on analytical realizations of properties of the considered problems. It consists of analytical formulae which use information from the last integration step only.

But these formulae are not used in the error model only. They are also used for answering some open questions which occur in the construction process of δ–control:

- how should the length of the subintervals I_k be chosen?,
- what should be the basic length T in our average operator $\mathcal{A}_T$?,
- can we always expect to get smooth trajectories c_k and smooth probabilities $|c_k|^2$ simultaneously?,

- how can we make a good a-priori δ–choice for the QRS(δ)-algorithms constructed above?

As in all adaptive algorithms (especially in stepsize-controlled adaptive integrators) the presented error estimation for QRS is *exact* in a more or less small subclass of problems only. In solving problems not belonging to this subclass, the guaranteed accuracy of the solution is reduced to a *heuristical accuracy*. Therefore it should be discussed how this "exact" subclass of problems can be described in our case and what can make us believe in only *heuristical accuracy*?

Thus, the segmentation of the following explanations is determined: First, we extensively discuss the perturbation theory of the effect of the dipole matrix on the interaction between the states. Therein, we have to watch out not only for the effect on the trajectories, but also for the effect on the original and smoothed probability matrices. This causes many details in notation and some unsightly analytical expressions will unfortunately appear, but it must be done... Based on this analysis we are able to work out the required error estimator, to answer the above collected questions, and to construct the whole δ–control. All this results in a brief presentation of a basic *adaptive* QRS–version.

4.1 Time Dependent Perturbation Theory

The aim of this section is the introduction of the basic concept and notation of perturbation theory for the equation

$$i\,\partial_t c \;=\; (\Omega \;+\; \epsilon\, V(t)\,)\;c, \qquad c(t_0) = c_0 \tag{4.1}$$

with $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_n)$ and a perturbation V with the special structure known from above:

$$V(t) \;=\; V^+ \exp(i\omega t) + V^- \exp(-i\omega t). \tag{4.2}$$

If we compare (4.2) with equations (2.11) or (3.14) we observe that we have to set $f(t) \equiv \mathrm{const}$, $V^+ = V_\delta$, and $V^- = V_\delta^T$. In the following it is very helpful to use the following *summation convention*:

A sum like $A_+(x^+) + A_-(x^-)$ consists of two analogically built summands, which only differ from each other by the replacement of an index "+" by "−". If such a sum appears in an equation it will be shortened by $A_\pm(x^\pm)$. For example, the perturbation $V(t)$ from (4.2) can be written as

$$V(t) \;=\; V^\pm \exp(\pm i\omega t). \tag{4.3}$$

In addition to equation (4.1) we consider its *non–disturbed* form

$$i\,\partial_t c \;=\; \Omega\; c, \quad c(t_0) = c_0. \tag{4.4}$$

4.1.1 Perturbation Theory for the Time Propagator

Let us denote the time propagation half group of (4.1) by $U^\epsilon(t,s)$. We will discuss the expansion of U^ϵ in terms of ϵ. For the case $\epsilon = 0$, $U^\epsilon = U_0$ is the time propagator of (4.4) and well–known:

$$U_0(t,s) \;=\; \exp\left(-i\Omega\,(t-s)\right). \tag{4.5}$$

Both, (4.1) and (4.4) are Schrödinger equations and therefore the time propagators U_0 and U^ϵ are unitary. According to [33] (pp. 207–209) the following theorem holds:

THEOREM 2. *The asymptotic expansion of U^ϵ is given by*

$$U^\epsilon(t,t_0) \;=\; U_0(t,t_0) \;+\; \sum_{m=1}^{\infty} \epsilon^m\; U_m(t,t_0) \tag{4.6}$$

with

$$\begin{aligned} U_m(t,t_0) \;=\; (-i)^m \int_{t_0}^{t}\int_{t_0}^{\tau_1}\cdots\int_{t_0}^{\tau_{m-1}} & U_0(t,\tau_1)V(\tau_1)U_0(\tau_1,\tau_2)\,\cdot \\ & \cdot V(\tau_2)\cdots U_0(\tau_{m-1},\tau_m)V(\tau_m)U_0(\tau_m,t_0)\,d\tau_m\cdots d\tau_1. \end{aligned} \tag{4.7}$$

(4.6) converges for any $\epsilon \in \mathbb{R}$ in arbitrarily chosen operator norm.

REMARK 3. Note that this is nothing else than a special form of Picard's iteration. More clearly this can be seen in section 4.1.2 using the "interaction picture" transformation.

In the following we assume a specific norm $||\cdot||$ for the solutions $c^\epsilon \in \mathbb{C}^n$ to be chosen, e.g.

$$||c|| \;:=\; \sum_k \phi_k \, |c_k|^2, \qquad \forall k: \;\; \Phi_k \neq 0. \tag{4.8}$$

As a consequence the operator norm is fixed (as that one which is induced by $||\cdot||$).

In the following there will often appear sums $A + A^*$ of a matrix $A \in \mathbb{C}^{n\times n}$ and its adjoint. Hence

$$\Re : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n} \quad \text{with} \quad \Re(A) \;=\; A \;+\; A^* \tag{4.9}$$

will be used for keeping the notation short. The symbol $\Re$ is used because symmetric matrices ($A = A^T$) lead to $\Re(A) \;=\; 2\,\mathrm{Re}(A)$ where $\mathrm{Re}(A)$ denotes the real part of an $A \in \mathbb{C}^{n\times n}$.

Now, $c^\epsilon(t) \;=\; U^\epsilon(t,t_0)\,c_0$ is the solution of (4.1) for a given ϵ. We are interested in its probability matrix p^ϵ for which Theorem 2 gives us

$$\begin{aligned} p^\epsilon(t) &= (c^\epsilon \otimes \overline{c^\epsilon})(t) \\ &= U^\epsilon(t,t_0)\; p(t_0)\; U^{\epsilon *}(t,t_0) \\ &= U_0(t,t_0)\; p(t_0)\; U_0^*(t,t_0) \;+ \\ &\quad\; 2\epsilon\; \Re\left(U_1(t,t_0)\; p(t_0)\; U_0^*(t,t_0)\right) \;+\; \mathcal{O}(\epsilon^2) \end{aligned} \tag{4.10}$$

where $p(t_0) = c_0 \otimes \overline{c_0}$ and its property $p(t_0)^* = p(t_0)$ is used in addition to the fact that U_1 is unitary because of $V = V^*$. Hence, if we want to compute the first order correction of p^ϵ in an elementary, i.e. integral free, form we have to evaluate the integral

$$U_1(t,t_0) \;=\; -\,i \int_{t_0}^{t} U_0(t,s)\,V(s)\,U_0(s,t_0)\;ds. \tag{4.11}$$

This can be done in a global notation which is prepared in the next section and can be used in further sections, too.

4.1.2 Interaction Picture

In the following let us consider equation (4.1) to be fixed, i.e. assume $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n\times n}$, $V^\pm$ and $\omega \in \mathbb{R}$ as fixed given. Hence U_0 (and with it all U_n) are explicitly computable. For simplification of those computations we need the definition of two helpful operations:

DEFINITION 3. *Let Γ be the set of maps $[t_0, \infty) \to \mathbb{C}^{n\times n}$ and γ the set of maps $[t_0,\infty) \to \mathbb{C}^n$. The actions of the bijective, linear operators*

$$\begin{aligned} T : \Gamma &\to \Gamma, & T(B)(t) &:= U_0^*(t,t_0)\, B(t)\, U_0(t,t_0) \\ \tau : \gamma &\to \gamma, & \tau(c)(t) &:= U_0^*(t,t_0)\, c(t) \end{aligned}$$

are called transformation into interaction picture. For simplification it will be written

$$B_I(t) \;:=\; T(B)(t) \quad \text{and} \quad c_I(t) \;:=\; \tau(c)(t) \tag{4.12}$$

The maps $Q_\pm : \Gamma \to \Gamma$ are called resonance denominators and are componentwisely defined as

$$(Q_\pm B)_{kl} \;=\; \frac{B_{kl}}{\Delta^\pm_{kl}} \quad \text{with} \quad \Delta^\pm_{kl} \;:=\; \omega_k - \omega_l \pm \omega. \tag{4.13}$$

Basic Calculus In addition to the obvious linearity of $Q_\pm$, we need two other important properties of the resonance denominators: First, let $A, D \in \Gamma$ and let D be diagonal. We observe

$$Q_\pm(AD) \;=\; Q_\pm(A)\, D \qquad \text{and} \qquad Q_\pm(DA) \;=\; D\, Q_\pm A. \tag{4.14}$$

Second, by a simple componentwise evaluation we see that the $Q_\pm$ could also be defined as solution operators of a class of matrix equations:

$$\Omega\, v \;-\; v\,(\Omega \mp \omega 1) \;=\; B \quad \Rightarrow \quad v \;=\; Q_\pm B \tag{4.15}$$

The transformations T and τ essentially reduce the effort of notation. This starts with the behaviour of equation (4.1) under transformation to interaction picture: The transformation

$$b^\epsilon(t) \;:=\; c^\epsilon_I(t) \;=\; e^{i\Omega(t-t_0)} c^\epsilon(t) \tag{4.16}$$

of the solution c^ϵ of equation (4.1) fulfills

$$i\,\partial_t b^\epsilon \;=\; \epsilon\, V_I(t)\; b^\epsilon(t), \qquad b(t_0) = c(t_0) \tag{4.17}$$

and the probability matrix $(b^\epsilon \otimes \overline{b^\epsilon})(t)$ of equation (4.17) is nothing else than p_I^ϵ:

$$p_I^\epsilon(t) \;=\; U_0^*(t,t_0)\,(c^\epsilon \otimes \overline{c^\epsilon})(t)\,U_0(t,t_0) \;=\; (b^\epsilon \otimes \overline{b^\epsilon})(t) \tag{4.18}$$

In addition, Theorem 2 gives us an integral form of U_1 which due to transformation of V into the interaction picture can be written as (cf. (4.11))

$$U_1(t,t_0) \;=\; -\,i\;U_0(t,t_0)\int_{t_0}^{t} V_I(s)\,ds. \tag{4.19}$$

With (4.19) we get another version of the first order formula (4.10):

$$\begin{aligned} p_I^\epsilon(t) \;&=\; p(t_0) \;+\; 2\,\epsilon\,\Re\,(U_0^*(t,t_0)\;U_1(t,t_0)\;p(t_0)) \;+\; \mathcal{O}(\epsilon^2) \qquad (4.20)\\ &=\; p(t_0) \;+\; 2\,\epsilon\,\Re\left(-i\int_{t_0}^{t} V_I(s)\,ds \;\; p(t_0)\right) \;+\; \mathcal{O}(\epsilon^2) \end{aligned}$$

Hence the following propositions are of some interest

LEMMA 1. *Let $B \in \Gamma$ be of the form $B(t) = B^\pm \exp(\pm i\omega t)$. Then the following formula for $B_I = T(B)$ holds:*

$$\int_{t_0}^{t} B_I(s)ds \;=\; -i\,\left((Q_\pm B^\pm)_I(t)\,e^{\pm i\omega t} \;-\; (Q_\pm B^\pm)_I(t_0)\,e^{\pm i\omega t_0}\right) \tag{4.21}$$

Proof. Simple partial integration of $v(t) \;:=\; \int_{t_0}^{t} B_I(s)ds$ gives

$$\begin{aligned} v(t) \;&=\; (i\Omega)^{-1}\;[\,B_I(s)\,]_{s=t_0}^{s=t} \\ &\quad -\,(i\Omega)^{-1}\;v(t)\;\;((-i\,(\Omega \mp \omega 1))\,. \end{aligned}$$

From multiplication with Ω on both sides results

$$\Omega v \;-\; v\,(\Omega \mp \omega 1) \;=\; -i\;[\,B_I(s)\,]_{s=t_0}^{s=t} \tag{4.22}$$

which together with (4.15) gives the proposition of the lemma. □

Perturbation Theory and Picard–Iteration. Now, we are able to fulfill the promise from Remark 3, i.e. to discuss our perturbation theory Theorem 2 as a special case of Picard's iteration.

Let us start with the time propagator $\mathcal{U}^\epsilon(t,t_0)$ of $b^\epsilon = c_I^\epsilon$:

$$\mathcal{U}^\epsilon(t,t_0) \;=\; U_0^*(t,t_0)\; U^\epsilon(t,t_0). \tag{4.23}$$

Because of equation (4.17) this fulfills the following differential equation:

$$i\,\partial_t \mathcal{U}^\epsilon(t,t_0) \;=\; \epsilon\, V_I(t)\, \mathcal{U}^\epsilon(t,t_0), \qquad \mathcal{U}^\epsilon(t_0,t_0) \;=\; 1 \tag{4.24}$$

Hence, applied to this equation, Picard's theorem gives us

$$\begin{aligned} \mathcal{U}_{k+1}^\epsilon(t,t_0) &= 1 - i\,\epsilon \int_{t_0}^{t} V_I(s)\,\mathcal{U}_k^\epsilon(s,t_0)\; ds \qquad (4.25)\\ \mathcal{U}_0^\epsilon(t,t_0) &= 1 \\ \mathcal{U}^\epsilon(t,t_0) &= \lim_{k\to\infty} \mathcal{U}_k^\epsilon(t,t_0) \end{aligned}$$

Obviously, the limes exists in any operator norm. By induction it is simple to prove that

$$\begin{aligned} \mathcal{U}_{k+1}^\epsilon(t,t_0) &= 1 + \sum_{m=1}^{k} \epsilon^m\, u_m(t,t_0) \\ u_m(t,t_0) &= (-i)^m \int_{t_0}^{t}\int_{t_0}^{\tau_1}\cdots\int_{t_0}^{\tau_{m-1}} V_I(\tau_1)\cdots V_I(\tau_m)\; d\tau_m\cdots d\tau_1 . \end{aligned}$$

But now we remember the U_m from Theorem 2 and we find

$$U_m(t,t_0) \;=\; U_0(t,t_0)\; u_m(t,t_0) \tag{4.26}$$

so that Theorem 2 is proved by using

$$\begin{aligned} U^\epsilon(t,t_0) &= U_0(t,t_0)\,\mathcal{U}^\epsilon(t,t_0) \\ &= U_0(t,t_0) + \sum_{m=1}^{\infty} U_0(t,t_0)\,u_m(t,t_0) \end{aligned}$$

Hence, a perturbation correction of order m in ϵ is nothing else than the mth Picard–iterate (cf. (4.25)) in the interaction picture.

4.1.3 First Order Perturbation Theory

With Lemma 1 we are able to rewrite formula (4.20) with the first order perturbation correction to p_I^ϵ. If this is done the first order form of our smoothed probability matrix $\Pi^\epsilon(t) = \frac{1}{T}\int_{t-T}^t p_I^\epsilon(s)\,ds$ can be computed due to integration. The results are given in the next theorems:

THEOREM 3. *The solution $b^\epsilon = c_I^\epsilon$ of equation (4.17) and its probability matrix $p_I^\epsilon = b^\epsilon \otimes \overline{b^\epsilon}$ fulfill*

$$\begin{aligned} b^\epsilon(t) &= c(t_0) - \epsilon\left(e^{i\Omega(t-t_0)}\,(Q_\pm V^\pm)\,e^{-i\Omega(t-t_0)}\,e^{\pm i\omega t} - \right. \\ &\qquad \left. (Q_\pm V^\pm)\,e^{\pm i\omega t_0}\right)\,c(t_0) + \mathcal{O}(\epsilon^2) \qquad (4.27) \\ p_I^\epsilon(t) &= p(t_0) + 2\,\epsilon\Re\left(Q_\pm V^\pm\; e^{\pm i\omega t_0}\; p(t_0)\right) \\ &\quad - 2\epsilon\,\Re\left(e^{i\Omega(t-t_0)}\;(Q_\pm V^\pm)\; e^{-i\Omega(t-t_0)}\; e^{\pm i\omega t}\; p(t_0)\right) \\ &\quad + \mathcal{O}(\epsilon^2) \qquad (4.28) \end{aligned}$$

Proof. To proof the first formula use Theorem 2 and equation (4.19) to get

$$\begin{aligned} b^\epsilon(t) &= U_0^*(t,t_0)\,(U_0(t,t_0) + \epsilon\,U_1(t,t_0))\,c(t_0) + \mathcal{O}(\epsilon^2) \\ &= \left(1 - \epsilon\int_{t_0}^{t} V_I(s)\,ds\right)\,c(t_0) + \mathcal{O}(\epsilon^2) \end{aligned}$$

and insert the equation from Lemma 1. The second formula analogically follows from Lemma 1, formula (4.10), and the definition of the interaction picture. □

Again using Lemma 1 direct evaluations of $\Pi^\epsilon(t) = \frac{1}{T}\int_{t-T}^t p_I^\epsilon(s)\,ds$ lead to

THEOREM 4. *The smoothed probability matrix*

$$\Pi^\epsilon(t_0+T) = \frac{1}{T}\int_{t_0}^{t_0+T} b^\epsilon \otimes \overline{b^\epsilon}(t)\,dt \qquad (4.29)$$

of the solution b^ϵ of our basic interaction picture equation (4.17) fulfills

$$\begin{aligned}\Pi^\epsilon(t_0+T) &= p(t_0) \;+\; 2\,\epsilon\Re\left(Q_\pm V^\pm\; e^{\pm i\omega t_0}\; p(t_0)\right)\\ &\quad -2\epsilon\,\Re\left(-\frac{i}{T}\left[e^{i\Omega T}\,(Q_\pm^2 V^\pm)\,e^{-i\Omega T}\,e^{\pm i\omega T} \;-\; (Q_\pm^2 V^\pm)\right]\,p(t_0)\,e^{\pm i\omega t_0}\right)\\ &\quad +\;\mathcal{O}(\epsilon^2). \end{aligned} \tag{4.30}$$

This looks really terrible, but it contains a lot of usable and interesting structure. The best way for discovering this structure is to rewrite componentwisely the formulae in Theorems 3 and 4 for the elements on the diagonals.

4.1.4 Componentwise First Order Corrections

THEOREM 5. *The diagonal elements of the probability matrices p_I^ϵ and Π^ϵ of the solution b^ϵ of equation (4.17), the transformation of equation (4.1) into the interaction picture, are fulfilling*

$$\begin{aligned}|b_k^\epsilon(t)|^2 &= |c_k(t_0)|^2 \\ &\quad -2\epsilon\sum_l \mathrm{Re}\left(\frac{V_{kl}^\pm}{\Delta_{kl}^\pm}\,(c_l\overline{c_k})(t_0)\;\left(e^{i\Delta_{kl}^\pm(t-t_0)}\;-\;1\right)\;e^{\pm i\omega t_0}\right)\\ &\quad +\;\mathcal{O}(\epsilon^2)\end{aligned} \tag{4.31}$$

$$\begin{aligned}\Pi_{kk}^\epsilon(t_0+T) &= |c_k(t_0)|^2 \\ &\quad -2\epsilon\sum_l \mathrm{Re}\left(\frac{V_{kl}^\pm}{\Delta_{kl}^\pm}\,(c_l\overline{c_k})(t_0)\;\left(\frac{e^{i\Delta_{kl}^\pm T}\;-\;1}{i\,\Delta_{kl}^\pm\,T}\;-\;1\right)\;e^{\pm i\omega t_0}\right)\\ &\quad +\;\mathcal{O}(\epsilon^2)\end{aligned} \tag{4.32}$$

whereas the components of b^ϵ itself are given by

$$b_k^\epsilon(t) \;=\; c_k(t_0)\;-\;\epsilon\sum_l\frac{V_{kl}^\pm}{\Delta_{kl}^\pm}\,c_l(t_0)\;\left(e^{i\Delta_{kl}^\pm(t-t_0)}\;-\;1\right)\;e^{\pm i\omega t_0}+\mathcal{O}(\epsilon^2) \tag{4.33}$$

Proof. The three formulae are direct results from the corresponding theorems. In the evaluation of $p_{kk}^\epsilon(t)=|c_k^\epsilon(t)|^2$ the being diagonal of all $\exp(i\Omega s)$

terms eases the calculations of the matrix products. Exactly the same property simplifies the expression for Π^ϵ_{kk}. □

It should be remarked that the diagonal elements of Π^ϵ are equal to those of $\pi^\epsilon(t) = \frac{1}{T}\int_{t-T}^{t}(c^\epsilon \otimes \overline{c^\epsilon})(s)\,ds$.

4.1.5 Normwise Concepts for Error Estimation

We have made a lot of effort in order to evaluate an elementary formula for the first order correction of perturbation theory. One will guess that this first order correction could be a good approximation of the real properties of equation (4.1) if ϵ is sufficiently small. In this section that guess will be investigated. Hence, first of all, we have to define what is exactly meant by "mth order correction":

DEFINITION 4. *Let c^ϵ denote the family of solutions of equation (4.1) for $\epsilon \in \mathbb{R}$ and $p^\epsilon = c^\epsilon \otimes \overline{c^\epsilon}$ the corresponding probability matrices. The perturbation theory correction of the orders $m \in M \subset \mathbb{N}$ is defined as*

$$\begin{aligned} c^\epsilon_M(t) &= \sum_{k\in M}\left(\left.\frac{\partial^k c^\epsilon}{\partial \epsilon^k}\right|_{\epsilon=0}\right)(t)\;\epsilon^k \\ p^\epsilon_M(t) &= \sum_{k\in M}\left(\left.\frac{\partial^k p^\epsilon}{\partial \epsilon^k}\right|_{\epsilon=0}\right)(t)\;\epsilon^k \end{aligned}$$

For simplicity let us write the summands of order $m \in \mathbb{N}$ with single indices c^ϵ_m and the sum of the summands of order $\leq m$ with $c^\epsilon_{\leq m}$:

$$c^\epsilon_m \;:=\; c^\epsilon_{\{m\}} \quad \text{and} \quad c^\epsilon_{\leq m} \;:=\; c^\epsilon_{\{1\ldots m\}} \tag{4.34}$$

This is analogically defined for p^ϵ.

Furthermore the exactness of an approximation needs a measure in which it can be formulated. Naturally, this measure will be a normwise one. And if we are just involved in definitions of particular norms, we can define them in a generality which will be sufficient for the description of the algorithm following later:

DEFINITION 5. *A $\phi \in \mathbb{R}^n$ with $\phi_k \neq 0\ \forall k = 1 \ldots n$ is called weight vector and a $\Phi \in \mathbb{R}^{n\times n}$ with $\Phi_{kl} \neq 0\ \forall k,l$ weight matrix. They define weighted 2-norms $||\cdot||_\phi$ and $||\cdot||_\Phi$ in the linear spaces $\mathbb{C}^n$ and $\mathbb{C}^{n\times n}$ by*

$$||c||_\phi := \sum_{k=1}^{n} \phi_k\ |c_k|^2 \qquad \text{for} \quad c = (c_k) \in \mathbb{C}^n$$

$$||A||_\Phi := \sum_{k,l=1}^{n} \Phi_{kl}\ |A_{kl}|^2 \qquad \text{for} \quad A = (A_{kl}) \in \mathbb{C}^{n\times n}.$$

For simplicity the matrix norm induced by $||\cdots||_\phi$ is written

$$||A||^\phi := \sup_{||c||_\phi = 1} ||Ac||_\phi \tag{4.35}$$

For a weight vector ϕ the notation $|\phi| = \sum_k \phi_k$ is defined. A weight matrix Φ is called multiplicative iff there are weight vectors ϕ, ψ with $\Phi = \phi \otimes \psi$.

LEMMA 2. *Let U_0 be the time propagator of equation (4.4) and U_m the time propagators which occur in the asymptotic expansion in Theorem 2. In addition let $b, c \in \mathbb{C}^n$ and ϕ, ψ be weight vectors. Then the following propositions hold:*

a) $||\, b \otimes \bar{c}\, ||_{\phi\otimes\psi} = ||\, b\, ||_\phi\ ||\, c\, ||_\psi$

b) $||\, U_0(t, t_0)\, ||^\phi = 1$

c) For our $V(t) = V^\pm e^{\pm i\omega t}$ from above we define $v_\phi := \max_t ||V(t)||^\phi$. For all $m \in \mathbb{N}$ it is

$$||\, U_m(t, t_0)\, ||^\phi \le \frac{1}{m!}\, v_\phi^m\ (t - t_0)^m \tag{4.36}$$

d) The perturbation theory correction of orders $m \in M \subset \mathbb{N}$ to the solution c^ϵ of equation (4.1) can be estimated as

$$||\, c_M^\epsilon\, ||_\phi \le \sum_{m\in M} \frac{(t - t_0)^m}{m!} (\,\epsilon\, v_\phi\,)^m\ ||c(t_0)||_\phi. \tag{4.37}$$

Proof. Written in components a) is trivial just as b) if we remember $U_0(t,t_0) = \exp(-i\Omega(t-t_0))$ with real Ω. For c) we take a look at Theorem 2, exchange integration and norm, use the submultiplicativity of $||\cdot||^\phi$, and the fact that

$$\int_{t_0}^{t}\int_{t_0}^{\tau_1}\cdots\int_{t_0}^{\tau_{m-1}} d\tau_n\cdots d\tau_1 \;=\; \frac{1}{m!}\,(t-t_0)^m. \tag{4.38}$$

With c), proposition d) follows directly from

$$c^\epsilon_M(t) \;=\; \sum_{m\in M} \epsilon^m\; U_m(t,t_0)\,c(t_0). \tag{4.39}$$

□

Now, it is easy to give an estimation for the exactness of the perturbation theory corrections. We observed above that our asymptotic expansion (Theorem 2) is a special case of Picard's iteration. Thus, we cannot be surprised about the particular form of our results.

THEOREM 6. *c^ϵ denotes the solution of equation (4.1) and $p^\epsilon = c^\epsilon\otimes\overline{c^\epsilon}$ its probability matrix. Let ϕ be a weight vector, v_ϕ the time independent maximal norm of the perturbation matrix $V(t)$ and $g(x)=\exp(x)-x-1$. Then the following formulae hold:*

a) $||\,c^\epsilon(t)\;-\;c^\epsilon_{\leq 1}(t)\,||_\phi \;\leq\; g\,(\,\epsilon\, v_\phi\,(t-t_0))\;\;||c(t_0)||_\phi$

b) $||\,p^\epsilon(t)\;-\;p^\epsilon_{\leq 1}(t)\,||_{\phi\otimes\phi} \;\leq\; g\,(\,2\,\epsilon\, v_\phi\,(t-t_0))\;\;||c(t_0)||_\phi$

Proof. Because of $c^\epsilon - c^\epsilon_{\leq 1} = c^\epsilon_{\{2,\ldots\}}$, a) follows directly from Lemma 2, d). For b) let us start with

$$p^\epsilon - p^\epsilon_{\leq 1} \;=\; \sum_{\substack{m,k\in\mathbb{N}_0\\ m+k>1}} c^\epsilon_k\otimes\overline{c^\epsilon_m} \tag{4.40}$$

Inserting Lemma 2, d) gives for $\rho(t) := ||\, p^\epsilon(t) \,-\, p^\epsilon_{\leq 1}(t)\, ||_{\phi\otimes\phi}$:

$$\begin{aligned}\rho(t) \;&\leq\; \sum_{M=2}^{\infty} \sum_{\substack{m,k\in\mathbb{N}_0\\ m+k=M}} \frac{(t-t_0)^M}{k!\,m!}\,(\epsilon\, v_\phi)^M \;||c(t_0)||_\phi \\ &\leq\; \sum_{M=2}^{\infty} (t-t_0)^M\,(\epsilon\, v_\phi)^M \;||c(t_0)||_\phi \;\sum_{m=0}^{M} \frac{1}{(M-m)!\,m!}\end{aligned}$$

The last inequality gives b), because of

$$\sum_{m=0}^{M} \frac{1}{(M-m)!\,m!} \;=\; \frac{2^M}{M!} \tag{4.41}$$

□

REMARK 4. Obviously, we can give another version of Theorem 6 d) with a *nonmultiplicative* weight matrix Φ. Therefore, we may start at

$$\left(p^\epsilon \,-\, p^\epsilon_{\leq 1}\right)(t) \;=\; \sum_{\substack{k,l\in\mathbb{N}_0\\ k+l>1}} \epsilon^{k+l}\, U_k(t,t_0)\, p(t_0)\, U_l^*(t,t_0) \tag{4.42}$$

with the U_k from Theorem 2. Let $||\cdot||^\Phi$ be the operator norm which is induced by $||\cdot||_\Phi$. We again find $||U_0(t,t_0)||^\Phi = 1$ for all $t \geq t_0$ and again define $v_\Phi := \max_t ||V(t)||^\Phi$. All this results in

$$||U_k(t,t_0)||^\Phi \;\leq\; \frac{1}{k!}\; v_\Phi^k\;(t-t_0)^k. \tag{4.43}$$

By using (4.43) for estimating (4.42) we finally find

$$\begin{aligned}||\left(p^\epsilon \,-\, p^\epsilon_{\leq 1}\right)(t)\,||_\Phi \;&\leq\; \sum_{M=2}^{\infty} (t-t_0)^M\,(\epsilon\, v_\Phi)^M \;||p(t_0)||_\Phi \;\sum_{m=0}^{M} \frac{1}{(M-m)!\,m!} \\ &\leq\; g\left(\,2\,\epsilon\, v_\Phi\,(t-t_0)\right)\;||p(t_0)||_\Phi\end{aligned} \tag{4.44}$$

4.2 A-Priori Error Estimation for QRS

In Section 4.1 we had worked out that we need *cheap* approximate knowledge about the solution of our basic equation

$$i\,\partial_t c \;=\; (\Omega \;+\; f(t)\,V(t)\,)\;\;c, \qquad c(t_0) = c_0 \tag{4.45}$$

in the interval $[t_0,\, t_0 + H]$ with $V(t)$ according to (4.2); more exactly about the diagonal elements of its smoothed probability matrix

$$\pi(t) \;=\; \frac{1}{T}\int\limits_{t-T}^{t} (c \otimes \bar{c})(s)\,ds. \tag{4.46}$$

Instead of π we can also ask for the interaction picture probability matrix

$$\Pi(t) \;=\; \frac{1}{T}\int\limits_{t-T}^{t} (c_I \otimes \bar{c}_I)(s)\,ds \tag{4.47}$$

because we have got $\operatorname{diag}\pi \;=\; \operatorname{diag}\Pi$.

Let the laser shape function $f(t)$ again be approximately constant in $[t_0,\, t_0 + H]$:

$$f\big|_{[t_0,\,t_0+H]} \;\cong\; E \tag{4.48}$$

For getting a cheap approximation we consider the first order correction $c^{\epsilon}_{\leq 1}$ of the disturbed equation

$$i\,\partial_t c \;=\; (\Omega \;+\; \epsilon\,V(t)\,)\;\;c, \qquad c(t_0) = c_0. \tag{4.49}$$

More exactly we define

$$\tilde{c}(t) \;:=\; c^{\epsilon}_{\leq 1}(t)\Big|_{\epsilon=E} \qquad \text{and} \qquad \tilde{b}(t) \;:=\; \left(c^{\epsilon}_{\leq 1}\Big|_{\epsilon=E}\right)_I(t) \tag{4.50}$$

respectively their smoothed magnitudes

$$\begin{aligned}
\tilde{\beta}(t) \;&=\; \frac{1}{T}\int\limits_{t-T}^{t} \tilde{b}(s)\,ds \\
\tilde{\Pi}(t) \;&=\; \frac{1}{T}\int\limits_{t-T}^{t} (\tilde{b} \otimes \bar{\tilde{b}}_I)(s)\,ds.
\end{aligned} \tag{4.51}$$

For $\tilde{c}$, $\tilde{b}$, and $\tilde{\Pi}$, Theorems 4 and 5 give analytical formulae. Hence both can be computed cheaply (see Subsection 4.3.5). The next subsection demonstrates that $\tilde{c}$, $\tilde{b}$, and $\tilde{\Pi}$ are worth to be considered.

4.2.1 An Exactly Approximated Subclass of Problems

We can easily prove the following approximation proposition:

THEOREM 7. *Let H, T, E, Π, $\tilde{\Pi}$, c, and $\tilde{c}$ be as defined on top of this section and assume $H \geq T$. Beside this, let ϕ be a weight vector, $v_\phi := max_t \, ||V(t)||^\phi$ the maximal norm of $V(t)$, and let $N(a)$ be the unique zero of the equation $\exp(x) - x - 1 = a$ for $a > 0$. For any $\eta > 0$ it holds:*

$$v_\phi \;\leq\; \frac{1}{2EH}\, N\left(\frac{\eta}{||c(t_0)||_\phi}\right) \qquad \Rightarrow \qquad ||\,(\Pi \,-\, \tilde{\Pi})(t_0+H)\,||_{\phi\otimes\phi} \;\leq\; \eta \tag{4.52}$$

Proof. With $I = [t_0 + H - T, t_0 + H]$ it is

$$||\,(\Pi \,-\, \tilde{\Pi})(t_0+H)\,||_{\phi\otimes\phi} \;\leq\; \max_{s\in I} ||(c \otimes \bar{c} \,-\, \tilde{c}\otimes\bar{\tilde{c}})(s)||_{\phi\otimes\phi} \tag{4.53}$$

from which the whole proposition easily follows with Theorem 6 and $\epsilon = E$. □

Thus, we have found a non–empty subclass of problems (for which the dipole matrix fulfills the inequality $v_\phi \;\leq\; \frac{1}{2EH}\, N\left(\frac{\eta}{||c(t_0)||_\phi}\right)$). For this subclass the first order correction $\tilde{\Pi}$ can serve as an “exact” (up to a small error η) approximation of Π. But let us generally call $\tilde{c}$, $\tilde{b}$, $\tilde{\Pi}$ *approximative solutions* of (2.11) in $[t_0, t_0 + H]$ even if the above inequality is hurt.

4.2.2 Construction of the Error Estimation Scheme

But now, let us step forward to our aim (the construction of an error estimation for QRS) and remember: The approximative solution $\tilde{c}$ ($\tilde{\Pi}$) should serve for an a-priori study of the effect of changes in the dipole matrix V on the solution $c\,(\Pi)$ of (4.45). Therefore we must think of $V(t)$ as a *free parameter* in (4.45) and (4.49). In the original ("exact") problem it has got the form

$$V(t) \;=\; V_\infty e^{i\omega t} \;+\; V_\infty e^{-i\omega t} \tag{4.54}$$

with the full molecular dipole matrix V_∞. In each QRS(δ)–interval we want to use

$$V(t) \;=\; V_\delta e^{i\omega t} \;+\; V_\delta^T e^{-i\omega t} \tag{4.55}$$

with V_δ computed according to (3.13) depending on the used sparsing parameter δ.

To denote the dependence of the solutions on the used dipole matrix let us write

$$c\{\mathcal{V}\}(t), \quad \tilde{c}\{\mathcal{V}\}(t), \quad \Pi\{\mathcal{V}\}(t), \quad \tilde{\Pi}\{\mathcal{V}\}(t), \quad \text{and so on} \tag{4.56}$$

if the solutions for

$$V(t) \;=\; \mathcal{V} e^{i\omega t} \;+\; \mathcal{V}^T e^{-i\omega t} \tag{4.57}$$

are meant.

Collecting the notations of the last few paragraphs we can finally define our error estimation scheme:

DEFINITION 6. *Let $c\{\mathcal{V}\}$ be the solution of our basic equation (4.45) with dipole matrix*

$$V(t) \;=\; \mathcal{V} e^{i\omega t} \;+\; \mathcal{V}^T e^{-i\omega t} \tag{4.58}$$

and let $b\{\mathcal{V}\}$ be its transformation into interaction picture and

$$\Pi\{\mathcal{V}\}(t) \;=\; \frac{1}{T}\int_{t-T}^{t} (b\{\mathcal{V}\} \otimes \overline{b\{\mathcal{V}\}})(s)\, ds \tag{4.59}$$

the corresponding smoothed interaction picture probability matrix. We are interested in the difference between $\Pi\{V_\infty\}$ (smoothed probabilities for the exact problem) and $\Pi\{V_\delta\}$ (those for a δ-sparsed dipole matrix). Let Φ be a weight matrix and $||\cdot||_\Phi$ the corresponding norm. The considered difference ("sparsing error") can be measured by

$$\epsilon_\Phi(t,\delta) \;:=\; ||\,(\Pi\{V_\infty\} \,-\, \Pi\{V_\delta\})\,(t)\,||_\Phi. \tag{4.60}$$

As an estimator of ϵ_Φ we define

$$\langle\epsilon_\Phi(t,\delta)\rangle \;:=\; ||\left(\tilde{\Pi}\{V_\infty\} - \tilde{\Pi}\{V_\delta\}\right)(t)\,||_\Phi \tag{4.61}$$

using the approximative smoothed probability matrix $\tilde{\Pi}$ constructed above (cf. equation (4.51)).

Now, let us consider one of the problems which do not belong to the "exactly" approximated subclass described in section 4.2.1. Theoretically we do not know anything about the question of whether $\tilde{c}$, $\tilde{\Pi}$ are still approximations of c, Π, or how good approximations they are. Why shall we trust the error estimator $\langle\epsilon_\Phi\rangle$ in this case?

First, we should remember the fact that this question *often* appears in the context of adaptive algorithms! Even the error models of adaptive ODE–integrators with stepsize control are heuristical ones.

In our case the error estimator is based on the effect of changes in the used dipole matrix V on the first order perturbation theory correction $\tilde{c}$. From equation (4.50) and Theorem 7 it can easily be seen that the estimate is an "exact" approximation (for *every* considered problem) if the external field strength E is small enough. Thus the considered error estimate is basing on a "blow-up technique": it is reliable under the assumption that *the effect, which changes in V cause in the considered smoothed probabilities, can linearly be extrapolated from sufficiently small to bigger values of the external field strength E.*

There are some heuristical reasons which can make this assumption reliable:

- The physical interpretation of QRS given above (see Section 2.1) and the connected success indicates that the herein considered optical excitation processes are *pure first order effects.* This statement cannot surprise: it appears in the discussion of nearly all direct light–molecule interactions with moderately large field strengths. This fact is in deep connection to the next item.
- The assumption named is a wise restriction of the considered class of problems: In Section 2.1 we had deduced the model equations using the

electric dipole approximations for the construction of the basic Hamiltonian. This approximation is only valid for external field strengths which are not too large. The electric dipole approximation itself is only a first order term of a multipole expansion. Hence we may trust our assumptions above as long as the physical model itself is reliable (i.e. as long as the field strength E is small enough to remain in the framework of the electric dipole approximation).

- Another important reason for the acceptance of a heuristical assumption is the success of its consequences. In Section 5 the performance of the QRS–error estimator is demonstrated. The success of the error model documented therein speaks: "it works well".

Now, we can start to design the required δ-control for QRS by studying the properties of the approximations $\langle\epsilon_\Phi\rangle$, $\tilde{\Pi}$, and $\tilde{c}$ (instead of those of ϵ_ϕ, Π, and c: the assumption, on which the error model depends, should even be the base of the details of the δ-control).

4.3 Control Mechanisms for Adaptive QRS

In the introduction to Section 4 the basic mechanism of an adaptive QRS–version was introduced: the division of the integration interval in subintervals $I_k = [t_k, t_k + H_k]$ and the choice of δ for the action of QRS(δ) on I_k using an a-priori error estimator $\langle\epsilon_\Phi\rangle(\delta)$. After we have already constructed $\langle\epsilon_\Phi\rangle$ we can answer the four questions collected above:

- How should the length of the subintervals I_k be chosen?
- What should be the basic length T in our average operator $\mathcal{A}_T$?
- Can we always expect to get smooth trajectories b_k and smooth probabilities $|b_k|^2$ simultaneously?,
- How can we make a good a-priori δ–choice for QRS(δ) in the first step?

For this purpose let us consider any subinterval $I = [t_0, t_0 + H]$ and a given initial value $c(t_0)$. But before turning to each single question in detail let us collect some notations and some basic properties of our approximations

$\tilde{\Pi}$ and $\tilde{b}$. In all subsections 4.3.x we should remember the notation from Subsection 4.1: $V^+ = V_\delta$ and $V^- = V_\delta^T$ for a given $\delta \in [0, \infty]$.

4.3.1 Interaction–Contribution Measures

If we go back to Theorem 5, we find

$$\begin{aligned}\tilde{\Pi}_{kk}(t_0 + T) \;&=\; |c_k(t_0)|^2 \\ &\quad -2 \sum_l \mathrm{Re}\left(\frac{E\, V_{kl}^{\pm}}{\Delta_{kl}^{\pm}}\, (c_l \overline{c_k})(t_0)\, \left(\frac{e^{i\Delta_{kl}^{\pm} T} - 1}{i\, \Delta_{kl}^{\pm}\, T} - 1 \right) e^{\pm i\omega t_0} \right)\end{aligned}$$

and

$$\tilde{b}_k(t) \;=\; c_k(t_0) - \sum_l \frac{E\, V_{kl}^{\pm}}{\Delta_{kl}^{\pm}}\, c_l(t_0)\, \left(e^{i\Delta_{kl}^{\pm}(t - t_0)} - 1 \right) e^{\pm i\omega t_0} \tag{4.62}$$

Hence we can call the expressions

$$R_{kl}(t_0, T) \;:=\; -2\, \frac{E\, V_{kl}^{\pm}}{\Delta_{kl}^{\pm}}\, \mathrm{Re}\left((c_l \overline{c_k})(t_0)\, \left(\frac{e^{i\Delta_{kl}^{\pm} T} - 1}{i\, \Delta_{kl}^{\pm}\, T} - 1 \right) e^{\pm i\omega t_0} \right) \tag{4.63}$$

the *contribution of the state k–state l–interaction to the smoothed populations of state k.* The decision to write R_{kl} as a function of (only) t_0 and T reflects the scope of the following discussion: We should have a closer look on the effect of changes of T on $\tilde{\Pi}$ while fixing the problem–dependent data E, $V^{\pm}$, $\Delta_{kl}^{\pm}$, ω. For studying the structure of these expressions let us define

DEFINITION 7. *The two functions $\xi_R, \xi_I : \mathbb{R} \to \mathbb{R}$*

$$\xi_R \;:\; x \;\longmapsto\; \begin{cases} \frac{1}{x}\left(\frac{\sin x}{x} - 1\right) & \text{for} \quad x \neq 0 \\ 0 & \text{for} \quad x = 0 \end{cases}$$

$$\xi_I \;:\; x \;\longmapsto\; \begin{cases} \frac{1 - \cos x}{x^2} & \text{for} \quad x \neq 0 \\ 0.5 & \text{for} \quad x = 0 \end{cases}$$

are called interaction–contribution measures. The matrix $R(t_0, T) \in \mathbb{R}^{n \times n}$, $R(t_0, T) = (R_{kl}(t_0, T))$ with

$$R_{kl}(t_0, T) \;:=\; -2\, E\, V_{kl}^{\pm}\, T\, \mathrm{Re}\left((c_l \overline{c_k})(t_0)\, e^{\pm i\omega t_0} \left(\xi_R(\Delta_{kl}^{\pm} T) \,+\, i\, \xi_I(\Delta_{kl}^{\pm} T) \right) \right) \tag{4.64}$$

is called interaction contribution matrix.

The two interaction contribution measures can be seen as *weight functions* in $R(t_0, T)$ if we want to discuss the dependence of $R(t_0, T)$ from the choice of the average length T in $\mathcal{A}_T$.

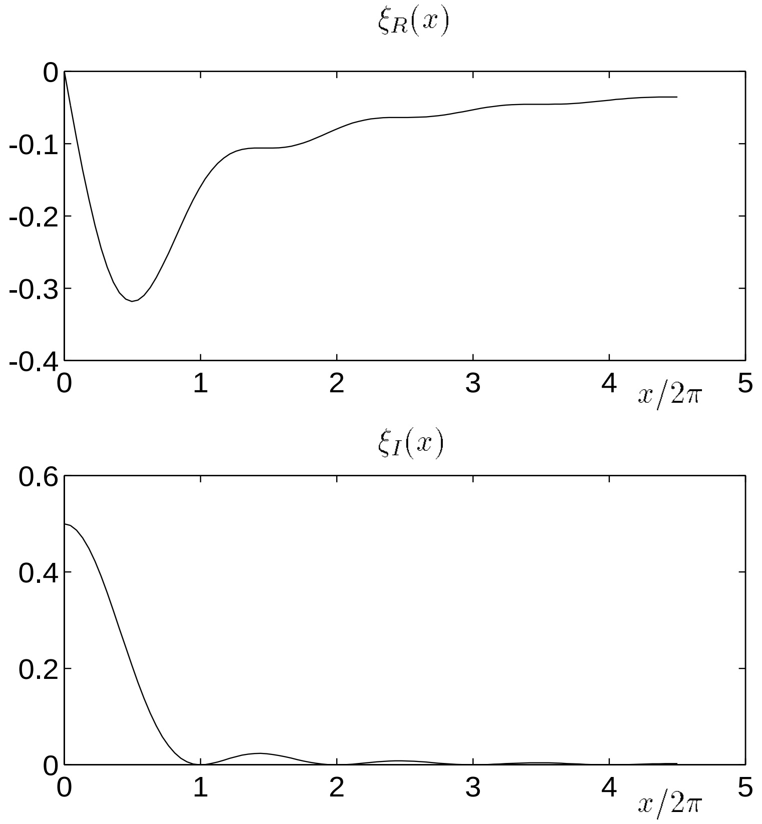


FIG. 11. Outlines of the above defined interaction contribution measures ξ_R and ξ_I.

Figure 11 shows the graphs of ξ_R and ξ_I. In both measures the damping of large values $x := \Delta^{\pm}_{kl} T$ is relatively strong. This reflects the damping of oscillations with large $\Delta^{\pm}_{kl}$–frequencies: with increasing T only the slow oscillations survive.

In equation (4.64) both, ξ_R and ξ_I, appear with two different arguments: $\xi_R(\Delta^{\pm}_{kl} T)$ and $\xi_I(\Delta^{\pm}_{kl} T)$. For studying the structure of the corresponding sums let us assume that $t_0 = 2m\pi/\omega$ and $T = 2j\pi/\omega$ with $m, j \in \mathbb{N}$. These specific values are preferred because they are significant and simplifying and because they will play a central role later on. Before evaluating the considered sums we should note that

$$\Delta^-_{kl}T \;=\; \Delta^+_{kl}T \;-\; 2\omega T \;=\; \Delta^+_{kl}T \;-\; 4\pi j. \tag{4.65}$$

With this we get

$$\begin{aligned} V^\pm_{kl}\;\; \mathrm{Re}\left(e^{\pm i\omega t_0}\left(\xi_R(\Delta^\pm_{kl}T) \;+\; i\,\xi_I(\Delta^\pm_{kl}T)\right)\right) & \\ =\; V^+_{kl}\;\; \xi_R(\Delta^+_{kl}T) \;+\; V^-_{kl}\;\; \xi_R(\Delta^+_{kl}T - 4\pi j) & \\ =:\; \Theta_R(V^+_{kl},\, V^-_{kl},\, \Delta^+_{kl}T,\, j) & \end{aligned} \tag{4.66}$$

and

$$\begin{aligned} V^\pm_{kl}\;\; \mathrm{Im}\left(e^{\pm i\omega t_0}\left(\xi_R(\Delta^\pm_{kl}T) \;+\; i\,\xi_I(\Delta^\pm_{kl}T)\right)\right) & \\ =\; V^+_{kl}\;\; \xi_I(\Delta^+_{kl}T) \;+\; V^-_{kl}\;\; \xi_I(\Delta^+_{kl}T - 4\pi j) & \\ =:\; \Theta_I(V^+_{kl},\, V^-_{kl},\, \Delta^+_{kl}T,\, j) & \end{aligned} \tag{4.67}$$

We are talking about the QRS–idea ("cutting the dipole element $V^\pm_{kl}$ if $|\Delta^\pm_{kl}| \notin [0, \omega)$"). Hence, we should give priority to the following task:

Let be $\Delta^+_{kl} \in (-\omega, \omega)$ (the case $\Delta^-_{kl} \in (-\omega, \omega)$ is analogical) and compare the two cases

1. $V^+_{kl} \;=\; V^-_{kl} \;=\; V_{kl}$ (QRS(∞)–calculation)
2. $V^+_{kl} \;=\; V_{kl}, \quad V^-_{kl} \;=\; 0$ (QRS($\delta < 1$)–calculation)

In these cases (4.66) and (4.67) get the form

1. $\Theta_{R,I}(V_{kl}, V_{kl}, x, j)/V_{kl} \;=\; \xi_{R,I}(x) \;+\; \xi_{R,I}(x - 4\pi j)$
2. $\Theta_{R,I}(V_{kl}, 0, x, j)/V_{kl} \;=\; \xi_{R,I}(x)$

For comparing these two cases we should take a look at Figure 12 which shows the two functions

$$\rho(x) \;=\; |\xi_R(x)|^2 \;+\; |\xi_I(x)|^2 \tag{4.68}$$

$$\Delta\rho_j(x) \;=\; |\xi_R(x - 4\pi j)|^2 \;+\; |\xi_I(x - 4\pi j)|^2 \tag{4.69}$$

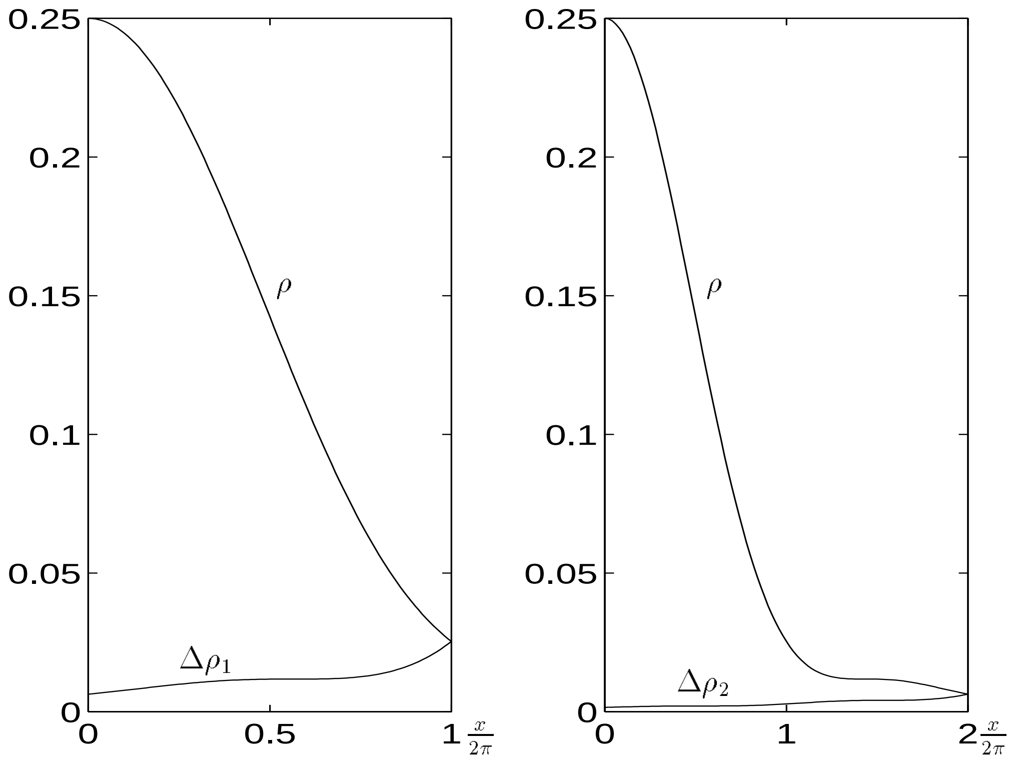


FIG. 12. Outlines of the above defined function ρ and $\Delta\rho_j$ for $j = 1, 2$. They demonstrate that the contribution (to π_{kk}) of the smaller of the both $\Delta_{kl}^{\pm}$–frequencies strongly dominates that of the larger one. The value of j corresponds to the chosen average length $T = 2\pi j/\omega$.

in the most interesting interval

$$x \in [0, 2\pi j] \quad \leftrightarrow \quad \Delta_{kl}^{+} \in [0, \omega] \tag{4.70}$$

for $j = 1, 2$.

It gets clear that the smaller frequency (Δ_{kl}^{+} herein) is strongly dominant in this $\pm$-sums. As an example this demonstrates the rightness of the QRS–idea: the contribution of the small frequency $\Delta_{kl}^{+} \approx 0$ to the smoothed probabilities is much more important than that of the larger frequency $\Delta_{kl}^{-} > \omega$.

For short let us go back to equation (4.62). Therein, the V–dependency of $\tilde{b}$ is omitted. But investigating the effect of its diagonal $\operatorname{diag} V$ we find an interesting result (notation from Section 4.2):

$$\left(\tilde{b}\{V\}(t_0+T) \; - \; \tilde{b}\{V - \operatorname{diag} V\}(t_0+T)\right)\Big|_{T=2\pi j/\omega} \; = \; 0, \qquad \forall j \in \mathbb{N} \tag{4.71}$$

using $\Delta^{\pm}_{kk} = \pm\omega$.

Conclusion. The observations, which we have made when studying $\xi_{R,I}$, ρ, $\Delta\rho_j$ and equation (4.62), tell us the story of the effect of the choice of the average length T. The main results of this subsection are:

- Using an average operator $\mathcal{A}_T$ we never reach the effect of cutting smoothers G_μ (see Section 3.2) which simply kill all fast oscillations. Only the $\mathcal{A}_T$ with very large T act like cutting smoothers somehow. Moderately large T ($\approx 2\pi/\omega$) lead to a pronounced damping of high frequencies without killing them.
- The contribution of the state k–state l–interaction to the smoothed populations is strongly dominated by the smaller of the $\Delta^{\pm}_{kl}$–frequencies. This effect grows with increasing T for which only frequencies $\Delta^{\pm}_{kl} \approx 0$ are dominating.
- Choosing the average length $T = 2\pi/\omega$ we do not get any effect of the diagonal of the dipole matrix V on the coefficients $\tilde{b}$.

4.3.2 *Choice of Average Length T and Interval Length H*

Let us now turn to the questions asked above:
How should the length of the subintervals I_k be chosen? and *What should be the basic length T in our average operator $\mathcal{A}_T$?*
There are several requirements to which should be paid attention. We collect them and then we look for a compromise:

- The adaptive control scheme needs an error estimate at the end of each subinterval, i.e. the evaluation of $\langle\epsilon_\Phi\rangle(\delta, t_0 + H)$. In that evaluation the knowledge of $\left(\tilde{\Pi}\{V_\infty\} - \tilde{\Pi}\{V_\delta\}\right)(t_0 + H)$ is demanded, i.e. the knowledge of $\left(\tilde{b} \otimes \bar{\tilde{b}}\right)\{V'\}|_{[t_0+H-T,t_0+H]}$ with $V' = V_\infty, V_\delta$. The $\langle\epsilon_\Phi\rangle$–computation has to be cheap and fast. Thus we get problems if $T > H$, because then we must use values $\tilde{\Pi}\{V'\}$ from $[t_0 + H - T, t_0)$. In that case we would have to evaluate and store many data from the previous QRS–subinterval. But if

(R1) $\quad T \le H,$

the needed approximative populations $\tilde{\Pi}\{V'\}$ are known analytically.

- The error estimator $\langle\epsilon_\Phi\rangle(\cdot, t)$ depends on results of a first order perturbation theory. Theorem 7 shows that this is only exact for small $(t - t_0)$. If we want to compute $\langle\epsilon_\Phi\rangle(\cdot, t_0 + H)$ for large H this fact demands us for care. We should require

  (R2) $\quad H$ should not be too large.

  This condition meets with the demand on the laser shape function f to be approximately constant in $[t_0, t_0 + H]$. (But remember: this last demand on f is only needed in the error estimator not yet in the QRS–integrations.)

- The previous subsection has told us that

  (R3) $\quad T$ should be large enough

  to give $\mathcal{A}_T$ sufficient damping properties for high frequencies ($> \omega$).

- In Section 3.3 the important influence of the diagonal of the dipole matrix V_∞ on the coefficients b was demonstrated. According to equation (4.71) $\operatorname{diag} V$ effects oscillations with frequencies $\pm\omega$ whose amplitudes are relatively large because of the relatively large V_{kk}–values. But according to Theorem 5, just this V_{kk}–values have *no* effect on the $\tilde{\Pi}_{kk}$–values! Planing the construction of the "right" QRS–integration interval $[t_0, t_0 + H]$ we have to consider the influence of the H–choice on the initial value $b(t_0 + H)$ for the next QRS–interval. If we are able to suppress the effect of the large $\pm\omega$–oscillations on $b(t_0 + H)$ we should certainly do this. From equation (4.62) and the last item in the previous subsection it gets clear that therefore we should require

(R4) $\quad H = \frac{2\pi j}{\omega}, \qquad j \in \mathbb{N}.$

Let us draw some conclusions from the requirements (R1)–(R4). (R1)–(R3) lead to

$$T = H, \qquad \text{but } H \text{ not too large.} \tag{4.72}$$

But the average length T must be fixed over all subintervals. Thus, we can link (R4) and (4.72) together to a final compromise:

Choose $T = 2\pi j/\omega$ with a small $j \in \mathbb{N}$ as average length and fix all subinterval–lengths to $H = T$.

For all QRS–intervals $I_k = [t_0^{(k)}, t_0^{(k)} + H]$ this choice implies that the values of their left hand sides always are

$$t_0^{(k)} = \frac{2\pi kj}{\omega}, \qquad kj \in \mathbb{N} \tag{4.73}$$

But how to choose j? One possibility for the proceeding is the following: The *user* defines T by choosing a small $j \in \mathbb{N}$ and all subintervals are fixed automatically. On the one hand giving limited choice of T to the user seems to be the right idea. But on the other hand it must be checked whether his choice does not hurt condition (R2). This can be realized by the following *checking procedure*:

REMARK 5. Let us stop the QRS–integration on a subinterval I_k iff the change in the populations intermediately gets too large. This interrupt–concept bases on the idea that the first order perturbation theory used in the error estimator can only be valid if the change in the populations is a "small correction". Taking (R4) into account this can be realized in the following way:

Assume that it is planned to take $I_k = [t_0^{(k)}, t_0^{(k)} + 2\pi j/\omega]$ for $j \in \mathbb{N}$ as kth subinterval. In addition assume $t_m^{(k)} \in I_k$, $m = 1, \ldots, M$ to be the inner integration steps of QRS(δ) on I_k and $p(t_m^{(k)})$ the populations computed therein. For each $m = 1, \ldots, M$ define $l(m) \in \{1 \ldots j\}$ by the condition

$$t_m^{(k)} \in \left(t_0^{(k)} + \frac{2\pi(l(m)-1)}{\omega},\, t_0^{(k)} + \frac{2\pi l(m)}{\omega} \right] \tag{4.74}$$

Inside QRS(δ_k) check after each step:

If $\left(||p(t_m^{(k)}) - p(t_0^{(k)})|| > \text{ctol} \right)$

Then $I_k \;:=\; \left[t_0^{(k)},\, t_0^{(k)} + 2\pi l(m)/\omega \right]$

i.e. let the subinterval I_k, i.e. the actual QRS–integration, end at the next possible time $t_0^{(k)} + 2\pi l(m)/\omega$ iff the change in the populations exceeds a given limitation. The upper bound ctol and the norm $|| \cdot ||$ must be given (in the implementation the maximum norm $|| \cdot ||_0$ is taken and it is set e.g. ctol $= 0.1$).

REMARK 6. Surely, we should try to construct an adaptive j–control, i.e. an automatic control scheme for choosing j according to the concrete behaviour of the solution computed actually. Using the error estimator $\langle \epsilon_\Sigma \rangle$ this is no serious problem: Planing the QRS–interval I_k we may determine the j–choice j_k on I_k by computing the maximal value $j \in \mathbb{N}$ which leads to the same δ–choice as the initial proposal $j_k^{(0)}$, e.g. $j_k^{(0)} = 1$. But the computational effort of this additional control mechanism cannot be neglected. Indeed, when implemented it led to an essential decrease in efficiency in all considered real life applications. Thus, the idea was dropped. Nevertheless, it can immediately be picked up if necessary.

We should shortly come back to the question: *Can we always expect to get smooth trajectories b_k and smooth probabilities $|b_k|^2$ simultaneously?* For the following reason we (again heuristically) can answer this question with "*Yes!*": Let us go back to Theorem 5, to equations (4.31) and (4.33). From these equations we can observe the δ–dependency of our approximations $\tilde{b}_k\{V_\delta\}(t)$ and $|\tilde{b}_k|^2\{V_\delta\}(t)$: For both, approximated solutions and populations, all fast oscillations (those with frequencies $\geq \omega$) *simultaneously* vanish iff $\delta < 1$. Therefore $\delta < 1$ is called the *smoothing domain* of the QRS–method. Thus, according to the first order approximation scheme, we expect smooth coefficients by smoothing the populations and vice versa.

4.3.3 Interaction Skeletons and Initial Choice of δ

In Section 4.3.1 we have seen that the interaction contribution of the state k–state l interaction ("(k,l)–interaction") depends on three parts: $\xi^{\pm}_{R,I}(\Delta^{\pm}_{kl}T)$, $V^{\pm}_{kl}$, and $(c_k\overline{c_l})(t_0)$. Let us assume that we want to analyse these interaction contributions *a–priori*, i.e. before any QRS(δ)–subintegration. A-priori, we have not got any information about the solution, i.e. no information about $(c_k\overline{c_l})(t_0)$. In this situation (a–priori for the whole integration process) the most useful contribution measures are the expressions

$$V^{\pm}_{kl}\ \xi^{\pm}_{R,I}(\Delta^{\pm}_{kl}T). \tag{4.75}$$

Obviously, we should not consider (4.75) as the sum it is according to our summation convention. Each of the $\pm$–summands should be considered separately. Therefore, we cancel our summation convention for this section.

The central question of the following paragraphs is this: If we fix the problem (i.e. fix $(V_\infty,\ \Omega,\ \omega,\ f)$ and T; thus $\Delta^{\pm}$, $V^{\pm} = V_\infty$) *which (k,l)–interactions are of most importance for the considered process?* Can we construct a kind of *interaction skeleton* representing these main (k,l)–interactions and their interconnections?

In Section 4.3.1 we have clearly seen that the smaller of the both frequencies $\Delta^{\pm}_{kl}$ dominates the interaction contribution. Thus one can make a distinction between the (k,l)–interactions whether they are "+" (for $|\Delta^{+}_{kl}| < |\Delta^{-}_{kl}|$) or "-" interactions ($|\Delta^{-}_{kl}| < |\Delta^{+}_{kl}|$). Thus, we separate + from − contributions, in addition to the differentiation between real(R)– and imaginary part(I) contributions.

Before we start, some assumptions should precede which can be made without loss of generality. First, assume the eigenmatrix Ω to be ordered: $l < k \ \Rightarrow\ \omega_l \le \omega_k$. Second, let ω be positive, and third, let the set of all state(numbers) be $J = \{1 \ldots n\}$.

Now we can fix our above a–priori interaction measures and some neighboring notations in a definition:

DEFINITION 8. *Let $V^{\pm}$, $\Delta^{\pm}$ be as above, $\xi_{R,I}$ as defined in Definition 7 and for $k \in J$ let us define*

$$J_{\pm}(k) \ := \ \left\{ l \in J, \quad |\Delta^{\pm}_{kl}| < |\Delta^{\mp}_{kl}| \right\}. \tag{4.76}$$

The following interaction measures are defined for all $k, l \in J$:

$$\begin{aligned} \mu^{\pm}_{R,I}(k,l) \ &:= \ \left| V^{\pm}_{kl} \right| \ \left| \xi^{\pm}_{R,I}(\Delta^{\pm}_{kl} T) \right| \\ \mu_{R,I}(k,l) \ &:= \ \max_{\pm} \ \mu^{\pm}_{R,I}(k,l) \\ \lambda^{\pm}_{R,I}(k) \ &:= \ \max_{l \in J_{\pm}(k)} \ \mu^{\pm}_{R,I}(k,l) \end{aligned}$$

For any $\eta \in [0,1]$ the functions

$$\begin{aligned} r^{\pm}_{R,I}(k,\eta) \ &:= \ \{(k,l) \in J^2, \quad l \in J_{\pm}(k), \quad \mu^{\pm}_{R,I}(k,l) > \eta \, \lambda^{\pm}_{R,I}(k)\} \\ r_{R,I}(k,\eta) \ &:= \ r^{+}_{R,I}(k,\eta) \ \cup \ r^{-}_{R,I}(k,\eta), \end{aligned}$$

$r_{R,I}(\cdot,\eta)$, $r^{\pm}_{R,I}(\cdot,\eta) : \ J \to 2^{J \times J}$ *are called sets of resonant indices.*

Thus, the $r^{\pm}_{R,I}(k,\eta)$ are the sets of all (k,l)–interactions which possibly give important (up to a safety factor η maximal) contributions to the smoothed population π_{kk}.

To compose the wanted *interaction–skeleton* we still have to ask whether $(c_k\overline{c_l})(t_0)$ can ever be big enough: If $(c_k\overline{c_l})$ is small all over the process the interaction (k,l) will not play a central role although it perhaps seems to be important according to $(k,l) \in r_{R,I}(k,\eta)$.

If we assume that $|c_k|^2$ and $|c_l|^2$ are small all over the process this also is the fact for $(c_k\overline{c_l})(t_0)$ for all t_0. Because we have got no a–priori information about the $|c_k|^2$ we should go on heuristically:

First of all it is clear that the initial state $s \in J$ (with $|c_s|^2(0) = 1$) and all main (s,l)–interactions play a central role: Starting with s all states l with $(s,l) \in r_{R,I}(s,\eta)$ are initially populated ("2nd generation"). After this the population can eventually concentrate in all states connected to these 2nd generation states l via $r_{R,I}(l,\eta)$, and so on. Thus the construction of our *interaction–skeleton* should be a *recursive* procedure representing this "grow of generations".

DEFINITION 9. *Adopting all notations from Definition 8 the following recursive sets are defined for* $\eta \in [0,1]$ *and an initial state* s*:*

$$\begin{aligned}
S^{(1)}_{R,I}(\eta) &:= r_{R,I}(s,\eta) \\
S^{(n)}_{R,I}(\eta) &:= \Big\{(k,l) \in J^2 \setminus \mathcal{R}^{(n-1)}_{R,I}(\eta), \quad (k,l) \in r_{R,I}(k,\eta) \text{ and} \\
&\qquad\qquad \text{there exists a } p \in J \text{ with } (p,k) \in S^{(n-1)}_{R,I}(\eta)\Big\} \\
\mathcal{R}^{(n)}_{R,I}(\eta) &:= \bigcup_{k=1}^{n} S^{(k)}_{R,I}(\eta).
\end{aligned}$$

The sets $\mathcal{R}^{(n)}_{R,I}(\eta)$ *are called sets of resonant indices of* n *generations.*

Obviously the probability of high population in states belonging to the nth generation drastically decreases with increasing n. In the definition of the interaction skeleton this effect can be considered by a *limitation of the allowed recursion level*: one may *choose* an $m \in \mathbb{N}$ and call $\mathcal{R}^{(m)}_{R,I}(\eta)$ skeletons. Here the choice of m is left to physical insight (or experience).

In this heuristical scope the limitation of the skeleton by *choice of maximal state* is just as good as the limitation by choice of maximal recursion level: choosing a maximal state(number) m the index pairs in the skeleton must not include state(number)s $l \geq m$. (Physicists say "starting with s the population will not come higher than m"). Let us fix these two different skeletons in a definition:

DEFINITION 10. *Adopting all notations from above and letting* $\eta \in [0,1]$ *the sets*

$$\mathcal{S}_{R,I}(\eta) := \mathcal{R}^{(\infty)}_{R,I}(\eta) := \bigcup_{k=1}^{\infty} S^{(k)}_{R,I}(\eta) \tag{4.77}$$

are called full interaction skeletons. Moreover for any $m \in \mathbb{N}$

$$\mathcal{S}^{(m)}_{R,I}(\eta) := \mathcal{R}^{(m)}_{R,I}(\eta) \tag{4.78}$$

are called recursion limited skeletons while for $m \in J, m > s$ *the sets*

$$\mathcal{S}_{R,I}(\eta, m) := \{(k,l) \in \mathcal{S}_{R,I}(\eta), \qquad k,l \leq m\} \tag{4.79}$$

are called interaction skeletons limited by maximal state. The choice of $\mathcal{S}_{R,I}$ in all cases may be forgiven; it simplifies the notation and causes no confusion because of the added indices and arguments.

REMARK 7. Obviously the construction of the interaction skeleton can be formulated in a graph theoretical pattern: The two graphs $G_{R,I} = (J, J \times J)$ include connections between all states (=nodes). In boths graphs the connections $(k,l) \in J \times J$ are double–weighted by the two measures $\mu_x^{\pm}(k,l)$ (with $x = R, I$). The skeletons are connected trees with father–node s. Their connections fulfill some conditions (they must be locally maximal); together they include the "paths of main population exchange".

Now, let us construct the initial δ–choice for QRS(δ) answering the initial question *How can we make a good a-priori δ–choice in the first step of a δ–controlled QRS–version or for a QRS(δ)–solution in a single step?*

If we have constructed interaction skeletons $\mathcal{S}_{R,I}$ we can simply choose the initial $\delta_{R,I}$ so that the sparsed dipole matrix V_δ includes all interaction coefficients V_{kl} with $(k,l) \in \mathcal{S}_{R,I}$. We find

$$(V_{\delta,lk} \;=\; V_{kl} \;\vee\; V_{\delta,kl} \;=\; V_{kl}) \quad \Leftrightarrow \quad \min_{\pm} \frac{\Delta_{kl}^{\pm}}{\omega} \;<\; \delta. \tag{4.80}$$

Hence, this idea leads to

$$\delta_{R,I} \;=\; \max_{(k,l)\in\mathcal{S}_{R,I}} \; \min_{\pm} \left| \frac{\Delta_{kl}^{\pm}}{\omega} \right|. \tag{4.81}$$

To denote this for the three cases introduced in the previous definition we have to write

DEFINITION 11. *According to any $\eta \in [0,1]$ and to the interaction skeletons defined in Definition 10 the real positive numbers*

$$\delta_{R,I}(\eta) \;:=\; \max_{(k,l)\in\mathcal{S}_{R,I}(\eta)} \; \min_{\pm} \left| \frac{\Delta_{kl}^{\pm}}{\omega} \right|$$

$$\begin{aligned} \delta_{R,I}^{(m)}(\eta) &:= \max_{(k,l)\in\mathcal{S}_{R,I}^{(m)}(\eta)} \min_{\pm} \left|\frac{\Delta_{kl}^{\pm}}{\omega}\right| \\ \delta_{R,I}(\eta,m) &:= \max_{(k,l)\in\mathcal{S}_{R,I}(\eta,m)} \min_{\pm} \left|\frac{\Delta_{kl}^{\pm}}{\omega}\right| \end{aligned}$$

are used to define the initial δ–choices up to a threshold value $\kappa > 0$:

$$\delta(\eta,\kappa) \;:=\; \min\{\kappa,\, \max\{\delta_R(\eta), \delta_I(\eta)\}\} \tag{4.82}$$

alternatively for $\delta_{R,I}$, $\delta_{R,I}^{(m)}$ or $\delta_{R,I}(\cdot, m)$.

REMARK 8. Using our approximate solution $\tilde{c}\,(\tilde{\Pi})$ we are able to measure the contribution of each single (k,l)–interaction to the smoothed populations π_{kk}. From this a question results: Why do we remain at the idea of an adaptive δ–control taking a whole δ–section of the dipole matrix? Why don't we try to consider the main (k,l)–interactions only: this would lead to a much sparser dipole matrix? All attempts in order to construct such an adaptive control mechanism have failed. The reason for this failure is the loss of important interconnections between the states in this kind of control procedures: These interconnections may be *measured* as less important for some subintervals (e.g. for the reason of small $(c_k\overline{c_l})(t_0)$). But they belong to the interaction skeletons connecting each (important) state in $+$ and $-$ direction with its main interaction states. The effect of the proposed control procedure ("main (k,l)–interactions only") is the decomposition of the above described interaction graphs in several non–connected subgraphs. This destroys the "paths of main population exchange" between the states. This destruction happens if some of these paths of main population exchange are not sufficiently represented in a *normwise* concept (sign structure of the $\tilde{\Pi}_{kk}$–expressions). These experiences have finally led to the idea of interaction skeletons which are used to "guarantee" or "conserve" these paths of main population exchange. Thus, $\delta(\eta,\kappa)$ from Definition 11 should be taken as a lower bound for all δ–choices ($\delta < \delta(\eta,\kappa)$ would destroy some important interconnections of the interaction skeleton). Indeed, this is done in the final adaptive QRS–algorithm (see Subsection 4.4).

4.3.4 Automatic Scaling

In our description of the error estimator $\langle \epsilon_\Phi \rangle$ we used an abstract norm $||\cdot||_\Phi$. The choice of a suitable norm plays an important role for the performance of the algorithm. We require the algorithm to be *scaling invariant*. This demands for the elements of Φ to be solution dependent *scaling weights*. To construct them we should use the above introduced analogy of δ–controlled QRS and integrators with adaptive stepsize control (see introducing remarks of Section 4). In the case of integrators the current scaling of a vector y in a subinterval $[t_0, t_0 + H]$ with initial value $y(t_0)$ is e.g. given by

$$||y||_s^2 \;=\; \frac{1}{n}\sum_k \left(\frac{y_k^2}{s_k}\right) \qquad \text{with} \qquad s_k \;=\; \max\{|y_k(t_0)|^2, s\} \tag{4.83}$$

using a threshold value $s > 0$ (for more details see [26]). First of all let us transfer this invariant scaling idea to the smoothed populations $\Pi = \mathcal{A}_T(b \otimes \overline{b})$ of a QRS–integration on $[t_0, t_0 + T]$ or more exactly to the approximations $\tilde{\Pi}(t_0 + T) = \mathcal{A}_T(\tilde{b} \otimes \overline{\tilde{b}})(t_0 + T)$ starting at t_0 with $b(t_0)$. Taking the notation from Definition 5 and a threshold value $\sigma > 0$ we introduce the *scaling weight matrix* $\Sigma(t_0) = (\Sigma_{kl}(t_0))$ with

$$\Sigma_{kl}(t_0) \;:=\; \frac{1}{n^2}\,\left(\max\{\sigma,\; \left|(b_k\overline{b_l})(t_0)\right|^2\}\right)^{-1}. \tag{4.84}$$

We get the scaling invariant norm $||\cdot||_{\Sigma(t_0)}$ on the $\tilde{\Pi}$–level (which is needed in our error estimator $\langle \epsilon_\Phi \rangle$). Herein $n = |J|$ is the total number of states.

With $|b_k|^2(t) \le 1$ for all k and t we find for all $k, l \in J$:

$$|b_k|^2(t_0) \le \sigma \;\vee\; |b_l|^2(t_0) \le \sigma \quad \Rightarrow \quad \Sigma_{kl}(t_0) = \Sigma_{lk}(t_0) = \sigma \tag{4.85}$$

Hence, with respect to $||\cdot||_{\Sigma(t_0)}$, the states from

$$\mathcal{P}_{\sqrt{\sigma}}(t_0) \;:=\; \{k \in J, \quad |b_k|(t_0) \le \sqrt{\sigma}\} \tag{4.86}$$

are less important. The $\mathcal{P}_\vartheta(t_0)$ with small $\vartheta \ll 1$ are called *sets of less populated states*.

4.3.5 Evaluation of the Error Estimate

Definitely we must try to minimize the effort which is caused by the computation of the error estimator

$$\langle \epsilon_{\Sigma(t_0)} \rangle(\delta, t_0) \;=\; \sum_{kl} \Sigma_{kl}(t_0) \; \left| \left(\tilde{\Pi}_{kl}\{V_\infty\} - \tilde{\Pi}_{kl}\{V_\delta\} \right)(t_0 + T) \right|^2 . \tag{4.87}$$

Let us define the single (k,l)–*contributions*

$$\varrho_{kl} \;:=\; \Sigma_{kl}(t_0) \; \left| \left(\tilde{\Pi}_{kl}\{V_\infty\} - \tilde{\Pi}_{kl}\{V_\delta\} \right)(t_0 + T) \right|^2 \tag{4.88}$$

and the diagonal–contribution

$$\varrho \;:=\; \sum_{k} \varrho_{kk} . \tag{4.89}$$

Surely we can neglect all (k,l)–contributions which fulfill

$$\varrho_{kl} \;\le\; \frac{s}{n^2}\, \varrho \tag{4.90}$$

with a small safety factor s (e.g. $s < 10^{-1}$). Can we save the computations of those ϱ_{kl}? The following result is of some help for answering this question.

LEMMA 3. *Assume $k,l \in \mathcal{P}_\vartheta(t_0)$ with $\vartheta < \sqrt{\sigma}$ and let the vector $v = (v_k)$ be given by the line–sums*

$$v_k \;:=\; \sum_{j} |V_{kj}|^2 \tag{4.91}$$

of the dipole matrix. Then the (k,l)–contribution ϱ_{kl} fulfills the following inequality:

$$\varrho_{kl} \;\le\; \frac{1}{2\, n^2}\; \frac{1}{\sigma}\; (E\, T\, \vartheta)^2 \; (v_k + v_l) \tag{4.92}$$

Let us displace the proof of this Lemma to the end of this subsection. We should firstly use it to get another result which is implied directly:

THEOREM 8. *Adopting the assumptions from Lemma 3, particularly $k, l \in \mathcal{P}_\vartheta(t_0)$, the diagonal–contribution ϱ from equation (4.89), and letting σ be the threshold value of the scaling matrix $\Sigma(t_0)$ we find:*

$$\vartheta \;\leq\; \frac{1}{|E|\,T}\,\sqrt{\frac{2s\sigma}{v_k + v_l}\,\varrho} \qquad \Rightarrow \qquad \varrho_{kl} \;\leq\; \frac{s}{n^2}\,\varrho \tag{4.93}$$

Now we can collect our findings in a conclusion which describes the way of how to compute the error estimate $\langle\epsilon_{\Sigma(t_0)}\rangle(\delta, t_0)$. We should drop certain $\tilde{\Pi}_{kl}$–computations in the evaluation of ϱ — the computations of the $\tilde{\Pi}_{kl}$–values which are indicated as unimportant by (4.93):

ALGORITHM 1.

Compute $v = (v_k)$

$\varrho \;=\; 0$

For $k = 1$ **to** n **do**

  $\vartheta \;=\; \frac{1}{|E|T}\,\sqrt{\frac{s\sigma}{v_k}\,\varrho}$

  If $(k \notin \mathcal{P}_\vartheta(t_0))$ **Then**

    Compute ϱ_{kk}

    $\varrho \;=\; \varrho + \varrho_{kk}$

  End of If

End of For

For $k = 1$ **to** n **do**

  For $l = 1$ **to** n **do**

    $\vartheta \;=\; \frac{1}{|E|T}\,\sqrt{\frac{2s\sigma}{v_k+v_l}\,\varrho}$

    If $((k \neq l)$ *and* $(k \notin \mathcal{P}_\vartheta(t_0)$ *or* $l \notin \mathcal{P}_\vartheta(t_0)))$ **Then**

      $\varrho \;=\; \varrho + \varrho_{kl}$

  End of For

End of For

$$\langle \epsilon_{\Sigma(t_0)} \rangle(\delta, t_0) \;=\; \varrho$$

An analogical algorithm is used to reduce the effort of the computation of the finally needed $\tilde{\Pi}_{kl}(t_0 + T)$ which have to be computed in the evaluation of the considered ϱ_{kl}.

REMARK 9. The effort of each evaluation of $\langle \epsilon_\Sigma \rangle$ is the more reduced the more states are "less populated". In real life applications, fortunately, *nearly all* states are less populated (cf. Section 5.3).

But this is only one way which leads to a cheaply computable error estimate. Another possibility is the choice of a simplified error estimator. For example, we may choose

$$\langle \epsilon_{\phi(t_0)} \rangle(\delta, t_0) \;:=\; \sum_{k=1}^{n} \phi_k(t_0) \left| \left(\tilde{\Pi}_{kk}\{V_\infty\} - \tilde{\Pi}_{kk}\{V_\delta\} \right)(t_0 + T) \right|^2 \qquad (4.94)$$

with a weight vector $\phi(t_0)$ defined in analogy to $\Sigma(t_0)$. It can be shown that $\langle \epsilon_{\phi(t_0)} \rangle$ contains information about the $\tilde{\Pi}_{kl}$–errors ($k \neq l$), too, and that Theorem 8 can again be used for saving "unnecessary computations" . In concrete applications $\langle \epsilon_{\phi(t_0)} \rangle$ mostly causes the same δ–choices as $\langle \epsilon_{\Sigma(t_0)} \rangle$ does. Herein, $\langle \epsilon_{\Sigma(t_0)} \rangle$ is preferred for mainly conceptual reasons (it contains the clearer concept of error control with respect to $\tilde{\Pi}_{kl}$, $k \neq l$).

Let us return to the proof of Lemma 3:

Proof. According to Theorem 4 and its definition, $\tilde{\Pi}_{kl}$ has got the form

$$\begin{aligned}\tilde{\Pi}_{kl}(t_0 + T) \;=\; & (c_k \bar{c}_l)(t_0) - \sum_j \frac{EV^{\pm}_{kj}}{\Delta^{\pm}_{kj}} (c_j \overline{c_l})(t_0) \left(\frac{e^{i\Delta^{\pm}_{kj} T} - 1}{i\,\Delta^{\pm}_{kj}\, T} - 1 \right) e^{\pm i\omega t_0} \\ & - \sum_j \frac{EV^{\pm}_{lj}}{\Delta^{\pm}_{lj}} (c_k \overline{c_j})(t_0) \left(\frac{1 - e^{-i\Delta^{\pm}_{lj} T}}{i\,\Delta^{\pm}_{lj}\, T} - 1 \right) e^{\mp i\omega t_0}.\end{aligned}$$

This implies for the (k, l)–contribution

$$\varrho_{kl} \;\leq\; \Sigma_{kl}(t_0) \left| \sum_{j \in J_k(\delta)} \frac{EV^{\pm}_{kj}}{\Delta^{\pm}_{kj}} (c_j \overline{c_l})(t_0) \left(\frac{e^{i\Delta^{\pm}_{kj} T} - 1}{i\,\Delta^{\pm}_{kj}\, T} - 1 \right) e^{\pm i\omega t_0} \right.$$

$$- \sum_{j \in J_l(\delta)} \frac{EV_{lj}^{\pm}}{\Delta_{lj}^{\pm}} (c_k \overline{c_j})(t_0) \left(\frac{1 - e^{-i\Delta_{lj}^{\pm} T}}{i\, \Delta_{lj}^{\pm}\, T} - 1 \right) e^{\mp i\omega t_0} \Bigg|^2 .$$

with

$$J_k(\delta) \ := \ \{ j \in J, \quad V_{\delta,kj} = 0 \} \tag{4.95}$$

Because of $k, l \in \mathcal{P}_{\vartheta}(t_0)$ we have

$$\Sigma_{kl}(t_0) \ = \ \frac{1}{n^2\, \sigma} \tag{4.96}$$

and looking back to Section 4.3.1 we find a simple estimate for the $\pm$–sum:

$$\left| \frac{1}{\Delta_{kj}^{\pm}} \left(\frac{e^{i\Delta_{kj}^{\pm} T} - 1}{i\, \Delta_{kj}^{\pm}\, T} - 1 \right) \right|^2 \ \leq \ \frac{1}{2} T^2, \qquad \forall k, j, T. \tag{4.97}$$

In addition we have $|c_j| \leq 1$ for all $j \in J$ so that we can finally estimate

$$\begin{aligned} |\varrho_{kl}| \ &\leq \ \frac{1}{2\, n^2} \frac{1}{\sigma} (E\, T\, \vartheta)^2 \left(\sum_{j \in J_k(\delta)} |V_{kj}|^2 + \sum_{j \in J_l(\delta)} |V_{lj}|^2 \right) \\ &\leq \ \frac{1}{2\, n^2} \, \frac{1}{\sigma} \, (E\, T\, \vartheta)^2 \, (v_k + v_l) \end{aligned}$$

□

4.4 The AQRS–Algorithm

After we have prepared all details, the adaptive, δ–controlled QRS–algorithm (AQRS) should finally be formulated. Let us start with some concluding remarks on the matter of notation: Let the problem data $(V_{\infty}, \Omega, \omega, f)$ and the initial state s be given, let the pulse length of the laser shape f be τ (so that the integration interval is $[0, \tau]$) and define $\tau_0 = 2\pi/\omega$. Moreover assume that a small $j \in \mathbb{N}$ is chosen so that

$$N \ := \ \lfloor \frac{\tau}{j\, \tau_0} \rfloor \ \in \ \mathbb{N}. \tag{4.98}$$

With this j–choice the average and subinterval length is chosen to be $T = j\tau_0$ implying that there are $N + 1$ subintervals

$$\begin{aligned} I_k \ &:= \ [t_k,\, t_k + T) \quad \text{ with } t_k \ := \ kT \text{ for } k \in \{0, \ldots, N-1\} \\ I_N \ &:= \ [NT,\, \tau) \end{aligned}$$

Finally let us fix a formal notation for the statement "Use QRS(δ) to compute $b(t_2)$ as solution of the integration of (2.13) with initial value $b(t_1)$ and data $(V_\infty, \Omega, \omega, f)$ on interval $[t_1, t_2]$":

$$b(t_2) \;=\; \mathrm{QRS}(\delta)\,[b(t_1), t_1, t_2] \tag{4.99}$$

Remembering the interaction skeleton $\mathcal{S}_{R,I}(\eta, m)$ and the corresponding initial δ–choice $\delta^{(m)}(\eta,\kappa)$ from Definitions 10 and 11, the automatic scaling matrix $\Sigma(t)$, the corresponding error estimator $\langle\epsilon_{\Sigma(t)}\rangle$ from Section 4.3.4, we can denote the AQRS algorithm:

ALGORITHM 2.

AQRS–Algorithm

Compute $\mathcal{S}^{(m)}_{R,I}(\eta)$ *and* $\delta_0 = \delta^{(m)}(\eta,\kappa)$

Initialize $b_k(0) \;=\; \frac{1}{\sqrt{2}}(1+i)\delta_{ks}$

For *k=0* **to** *N-1* **do**

$\delta \;=\; \delta_0$

Compute $\varepsilon_0 \;=\; \langle\epsilon_{\Sigma(t_k)}\rangle(0, t_k)$

Repeat

Compute $\varepsilon_\delta \;=\; \langle\epsilon_{\Sigma(t_k)}\rangle(\delta, t_k)$

If $\left(\left|\frac{\varepsilon_\delta}{\varepsilon_0}\right| > \theta\right)$ **Then** $\delta \;=\; \delta + \Delta\delta$

Until $\left(\left|\frac{\varepsilon_\delta}{\varepsilon_0}\right| \;\leq\; \theta\right)$

$b(t_{k+1}) \;=\; QRS(\delta)\,[b(t_k), t_k, t_{k+1}]$

End of For

$\delta \;=\; \delta_0$

Compute $\varepsilon_0 \;=\; \langle\epsilon_{\Sigma(t_N)}\rangle(0, t_N)$

Repeat

$$\textbf{Compute} \quad \varepsilon_\delta \;=\; \langle \epsilon_{\Sigma(t_N)} \rangle(\delta, t_N)$$
$$\textbf{If} \;\left(\left| \tfrac{\varepsilon_\delta}{\varepsilon_0} \right| > \theta \right) \;\textbf{Then} \quad \delta \;=\; \delta + \Delta\delta$$

$$\textbf{Until} \;\left(\left| \tfrac{\varepsilon_\delta}{\varepsilon_0} \right| \leq \theta \right)$$

$$b(\tau) \;=\; QRS(\delta)\,[b(t_N), t_N, \tau]$$

REMARK 10. Some comments should be given on the used *local relative accuracy control* $\varepsilon_\delta/\varepsilon_0 < \theta$. In advanced integrators the stepsize control scheme depends on a local absolute accuracy control $\langle\epsilon\rangle <$ tol (see e.g. [14] or [21]): the estimated local error must remain under the required accuracy. In AQRS the δ–control scheme depends on controlling the local error, too. But instead of an absolute control scheme a relative one is used: ε_0 represents the effect of all (k,l)–interaction contributions on the populations in $(t_0{+}T)$ while ε_δ represents the effect of those (k,l)–contributions which are neglected only taking V_δ as dipole matrix. Hence $\varepsilon_\delta/\varepsilon_0$ represents the *percentage* of effects of (k,l)–contributions which are *missed* taking V_δ instead of V_∞. $\varepsilon_\delta/\varepsilon_0 < \theta$ means that this *missed percentage* has to undergo a required value θ which should be called the *allowed error percentage.*

REMARK 11. For simplification of notation we have not included additional control procedures (concerning exceptional cases) in this first AQRS–algorithm. There are the following two additional procedures:

1. the checking procedure from Remark 5 (see page 80) which saves the interval lengths from being too large and
2. the switch to symmetric mode which will be described in the next subsection.

In both cases it should be clear how to implement these algorithmic additions.

In the presented form AQRS is a *family* of algorithms which contains some control parameters. For each specific AQRS–algorithm these parameters have

Symbol	Range	Place of Use	Recommended Value
m	$\mathbb{N}$	recursion limit (interaction skeleton)	10
η	$[0,1]$	safety factor (interaction skeleton)	0.9
κ	> 0	threshold value (interaction skeleton)	0.9
σ	$[0,1]$	threshold value (automatic scaling)	10^{-4}
j	$\mathbb{N}$	factor for average and subinterval length	c
ctol	> 0	upper bound for change in population	0.1
$\Delta\delta$	> 0	increment for δ	0.1
θ	$[0,1]$	allowed error percentage	c
s	$[0,1]$	computation of error estimate	0.1

TABLE I. *Hidden parameters in AQRS and their place of use. The last column of the table contains some values on which the corresponding parameters are fixed all over the AQRS–tests presented herein. An entry "c" in this column means "to be chosen". It indicates that these parameters may be chosen problem dependent. Methods of choice are given in this section.*

to be chosen. Let us collect them and let us add a short description of each. You find this collection in Table I.

Before censuring "lots of parameters" you should notice the explanation of Table I and pay attention to the following items:

- Only two parameters may be chosen problem–dependent. The others are only hidden ones which are fixed. They only are remains of our theoretical investigations but listed up frankly. In addition, j and θ can be chosen automatically: j should be adapted to the property of the laser shape f (according to its slow or fast variation), its default value is $j = 6$; θ leaves scope to the user and his accuracy requirements (my usual choice is $\theta = 10^{-1}$). For m there is a dangerless choice ($m = n$ which is the implemented default value) but some physical insight might imply a better choice which can gain a speedup factor up to 2.

- In the above sections we have always compared the δ–controlled AQRS with advanced integrators with stepsize–control. Even in the implementation of these advanced integrators “lots of parameters” are included. Mostly they are hidden (scaling thresholds, different factors for stepsize reduction, ...). But some are elements of the user interface (tolerance, initial stepsize).

4.4.1 Switch to Symmetric Mode

We have observed that AQRS works as a smoothing algorithm *only if* it remains in its “smoothing domain” $\delta < 1$ (see Figure 8). We must expect a decrease in its efficiency if the choice $\delta \geq 1$ is necessary (see Figure 9). Thus, is there any possibility to damp this decrease? The answer is “Yes!”. The basic idea starts at the fact that (k,l)–interactions with $|\Delta^+_{kl}| \approx \omega$ *and* $|\Delta^-_{kl}| \approx \omega$ become important if $\delta < 1$ but $\delta \approx 1$. For them the contributions of the “+” and “−”–terms to $\tilde{\Pi}_{kk}$ are nearly equal (see Section 4.3). But for considering both, “+” and “−”, AQRS is forced to choose $\delta > 1$. Maybe other (k,l)–interactions which would lead to $\delta > 1$, are of no importance in comparison to those ones. In this case we can remain at a $\delta < 1$ by using the symmetrized dipole matrix

$$\mathcal{V}_\delta \;:=\; \frac{1}{2}\left(V_\delta \,+\, V_\delta^T\right) \tag{4.100}$$

instead of V_δ *and* V_δ^T in the QRS(δ)–integration in the considered subinterval. Let us call this “*Switch to symmetric mode*”. In our AQRS–algorithm 2 this can be inserted by changing the line

If $\left(\left|\frac{\varepsilon_\delta}{\varepsilon_0}\right| > \theta\right)$ **Then** $\delta \;=\; \delta + \Delta\delta$

to

If $\left(\left|\frac{\varepsilon_\delta}{\varepsilon_0}\right| > \theta\right)$ **Then**

  If $(\delta + \Delta\delta < 1)$

    Then $\delta \;=\; \delta + \Delta\delta$

Else $\quad \delta = \delta_0, \quad \text{symMode} = \text{true}$

starting with symMode = false and using $\mathcal{V}_\delta$ from (4.100) instead of V_δ,V_δ^T in the error estimations and QRS(δ)–integration iff symMode = true.

4.4.2 Adjustment of the background integrator

Assume we have done the QRS(δ_k)–integration on subinterval I_k and all preparation for the integration on I_{k+1}. How must the accuracy requirement $\text{tol}^{(k+1)}$ and the initial stepsize $h_0^{(k+1)}$ of the background integrator in the next QRS–integration be adjusted?

For $h_0^{(k+1)}$ the answer is simple: Let $\tilde{h}$ be the last stepsize of the QRS(δ_k)–integration on $I_k = [t_0^{(k)}, t_0^{(k)} + T]$ and let $h_M^{(k)}$ be the stepsize choice the integrator made in $t_M^{(k)} = t_0^{(k)} + T - \tilde{h}$ before restricting its stepsize to $\tilde{h}$ to exactly reach $t_0^{(k)} + T$. Then, choose $h_0^{(k+1)} \; := \; h_M^{(k)}$.

Now, let the AQRS–subintervals be $I_k = [t_0^{(k)}, t_0^{(k+1)})$, $k = 1, \ldots, N$ with $t_0^{(N+1)} = \tau$ and for given δ let

$$\hat{b}(t_0^{(k+1)}) \; := \; \text{QRS}(\delta) \left[b(t_0^{(k)}), t_0^{(k)}, t_0^{(k+1)} \right] \tag{4.101}$$

be the numerical and $b(t_0^{(k+1)})$ the exact solution of (3.14) for this δ. In addition let $\hat{B}(\tau)$ be the final numerical solution of the whole AQRS–integration and $B(\tau)$ the corresponding exact solution of (2.13).

If (in the background integrator) $\text{tol}^{(k+1)}$ is chosen as tolerance, we may assume

$$\left| \hat{b}_j(t_0^{(k+1)}) \; - \; b_j(t_0^{(k+1)}) \right| \; < \; C \; \text{tol}^{(k+1)}, \quad \forall j \tag{4.102}$$

with a constant $C \geq 1$. From this follows via simple estimations

$$\left| \left| \hat{b}_j \, \overline{\hat{b}_l} \right| (t_0^{(k+1)}) \; - \; \left| b_j \, \overline{b_l} \right| (t_0^{(k+1)}) \right| \; < \; (3 - \delta_{jl}) \, C \; \text{tol}^{(k+1)}, \quad \forall j, l. \tag{4.103}$$

The choice of $\text{tol}^{(k+1)}$ is mainly limited by two demands: First, there is a *global* accuracy requirement for the numerically computed final populations $|\hat{B}_k|^2(\tau)$. For them at least the first two leading digits should be right.

This forces us to note a stronger *basic local* accuracy requirement for the intermediate populations $|b_j|^2(t_0^{(k+1)}) = |c_j|^2(t_0^{(k+1)})$, e.g.:

$$\Big|\, |b_j|^2(t_0^{(k+1)}) \;-\; |\hat{b}_j|^2(t_0^{(k+1)}) \Big| \;<\; 10^{-3}, \quad \forall k. \tag{4.104}$$

Second, those $\hat{b}_j\,\overline{\hat{b}_l}(t_0^{(k+1)})$ which are used in the error estimator must be exact enough. They *are* used in the error estimator iff their index pair is in

$$\mathcal{J} \;:=\; \Big\{ (j,l), \;\; \Big|\hat{b}_j\,\overline{\hat{b}_l}(t_0^{(k+1)})\Big| \;>\; \sqrt{\sigma} \Big\} \tag{4.105}$$

and they must be exact enough *not* to disturb the evaluation of $\langle \epsilon_{\Sigma(t_0^{(k+1)})} \rangle$. We can realize this by a condition like

$$\frac{\Big|\, \Big|\hat{b}_j\,\overline{\hat{b}_l}\Big|\,(t_0^{(k+1)}) \;-\; \Big|b_j\,\overline{b_l}\Big|\,(t_0^{(k+1)}) \Big|}{\Big|\hat{b}_j\,\overline{\hat{b}_l}\Big|\,(t_0^{(k+1)})} \;<\; 10^{-1}, \quad \forall (j,l) \in \mathcal{J}. \tag{4.106}$$

Taking (4.103) and (4.105) into account we can replace this by

$$\frac{1}{\sqrt{\sigma}}\,3C\;\mathrm{tol}^{(k+1)} \;<\; 10^{-1} \qquad \text{or} \qquad \mathrm{tol}^{(k+1)} \;<\; \frac{\sqrt{\sigma}}{3C} \cdot 10^{-1}. \tag{4.107}$$

Making a "normal choice" $\sigma = 10^{-4}$ we can "fulfill" (4.104) and (4.107) by setting

$$\mathrm{tol}^{(k+1)} \;:=\; 10^{-5}, \quad \forall k \tag{4.108}$$

4.4.3 Complexity measures for AQRS

Let us first ask: How can we measure the computational effort produced by the sequence of QRS(δ)–integrations. Remember the construction of the QRS(δ) effort measure $C_{P,I}(\delta)$ (see Definition 2 in Section 3.3) depending on the integration interval I and the problem data $P = (V, \Omega, \omega, f)$. It leads to the following definition:

DEFINITION 12. *Let $P = (V, \Omega, \omega, f)$ be the data of the considered problem. Let I be the whole integration interval* AQRS *acts on. Assume the* AQRS*–subintervals to be I_k, $k = 0 \ldots N$ and $(\delta_0, \ldots, \delta_N)$ to be the sequence of δ–choices made by* AQRS *on these subintervals. The integration complexity measure of* AQRS *is defined as*

$$\mathcal{I}_{P,I} \;:=\; \sum_{k=0}^{N} C_{P,I_k}(\delta_k) \tag{4.109}$$

which (using the notation from Definition 2) can also be written

$$\mathcal{I}_{P,I} \;:=\; \sum_{k=0}^{N} \#(V_{\delta_k}) \,\cdot\, N_P(\delta_k, I_k). \tag{4.110}$$

According to the algorithm given in Section 4.3.5 the costs of the computation of the error estimate $\langle \epsilon_{\Sigma(t_0)} \rangle(\delta, t_0)$ is dominated by

- the number of elements $\tilde{\Pi}_{kl}$ which really are computed,
- the number of flops the computation of each considered $\tilde{\Pi}_{kl}$ costs.

Therefore we define

DEFINITION 13. *Let P, I_k, δ_k be as in Definition 12. In addition let χ^j_{kl} be the number of dipole matrix elements V_{km} or V_{lm} which are really needed in the computation of $\tilde{\Pi}_{kl}(t_j + T)$ by the error estimation algorithm 1 from page 89 preparing interval I_j. Then we have $\chi^j_{kl} = 0$ iff $\tilde{\Pi}_{kl}(t_j + T)$ is not needed. The error estimation complexity measure for* AQRS *is defined by*

$$\mathcal{E}_{P,I} \;:=\; \sum_{j=0}^{N} \sum_{kl} \chi^j_{kl} \tag{4.111}$$

Finally the sum

$$\mathcal{C}_{P,I} \;:=\; \mathcal{E}_{P,I} \,+\, \mathcal{I}_{P,I} \tag{4.112}$$

is called total complexity measure for AQRS.

5 Application to Laser–Molecule Interaction

In the following first few paragraphs the introductory explanations from Section 1 and Section 2.1 are again summarized and complemented in some points. This is made in order to prepare for the concrete description of the real life applications which serve as test problems for the demonstration of the performance of AQRS.

The field of *Laser–Assisted Molecular Control* has received a considerable amount of attention recently (e.g. [2] and the lot of work cited in [32]). It is of great importance in several areas of Physics and Chemistry, e.g. in the attempt to control chemical reactions or isotope separation by laser pulses in Chemical Physics. However, until the 1980's the majority of previous investigations remained restricted to a constant amplitude of the driving laser force. But in a number of important cases, the physical situation forces us to allow rapid changes of this amplitude: For instance, it is indispensable to work with very short laser pulses if it is necessary to beat fast molecular processes such as the redistribution of vibrational energy [3]. Because of this fact a new working field has emerged which may be called "Selective Excitation of Molecules using Ultrashort Laser Pulses". Let us consider applications from this field.

The key to the mathematical modelling of these laser–molecule–interactions is the Schrödinger equation $i\partial_t\psi = H\psi$ (see (2.4) in Section 2.1). The starting point for a fully quantum mechanical treatment of the dynamics of a molecule exposed to a laser field would be a Hamiltonian of the form $H = H_0 + H_1 + H_f$ where H_0 is the Hamiltonian of the unperturbed molecule, H_f that of the electromagnetic field, and H_1 describes the laser–molecule interaction. H_0 represents electronic, vibrational and rotational degrees of freedom. But in most cases it is sufficient to restrict H_0 to only one kind of these molecular degrees of freedom: one is allowed to assume that the reaction path is decoupled from all other degrees of freedom. In addition, the laser field needs mostly not be treated as a quantized electromagnetic field (thus $H_f = 0$), but rather, as a classical external force which acts on H_0. Thus, in the framework of Born–Oppenheimer and semiclassical electric dipole approximation one restricts the model to exactly the situation described in Section 2.1 with the ODE (2.11) as the basic equation describing the dynamical process.

We pay attention to the field of "Laser–Assisted Molecular Control". Therein, one is interested in designing an external laser pulse which allows a *controlled, selective excitation* of the molecule. For example, let all molecules of a sample occupy the same initial (eigen)state of H_0, and assume that the requirement is a strong occupation of another (eigen)state k, the stronger the better. In the language of Quantum Theory and with respect to our basic dynamical equation (2.11) (see Section 2.1) this task can be formulated as follows:

Determine the optimal laser data (laser shape function $f_o : [0,\tau] \to \mathbb{R}$ and light frequency ω_o) with respect to the conditions

$$p(0) = p_0 \quad \text{and} \quad p_{kk}^{(f_o,\omega_o)}(\tau) = |c_k^{(f_o,\omega_o)}|^2(\tau) = \max_{(f,\omega)} |c_k^{(f,\omega)}|^2(\tau), \tag{5.1}$$

where $c^{(f,\omega)}(t)$ is the solution of (2.11) according to the laser data (f,ω) and fixed molecular data (Ω, V_∞). The first condition represents the initial occupation of the molecule's eigenstates and the second the requirement of maximal occupation of the target state k after pulse length τ.

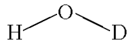
But there are additional conditions: f must be a technically producible laser shape (changing smoothly in time) with a pulse length τ short enough to exclude the redistribution of energy in $[0,\tau]$ (see Remark 2). Thus, mathematically this is a formidable optimization problem with constraints. In the chemical literature (e.g. [10][28][30][36][35][32]) the latter additional conditions are considered by fixing the form of the laser shape: for example, one restricts the optimization task to picosecond $\sin^2$–laser shapes (or analogical pulse shapes [17][28]), e.g.

$$f(t) \ := \ E_r \ \sin^2(\eta t), \qquad \eta \ = \ \frac{\pi}{\tau} \text{ with } \tau = 1\,\text{ps}. \tag{5.2}$$

Now the problem is simpler:
Determine (ω, E_r) *so that* $|c_k^{(\omega,E_r)}|^2(\tau) \ = \ \max$,
whereas the basic dynamical equation (2.11), its initial condition $c(0) = c_0$, its molecular data (Ω, V_∞), and the $\sin^2$–shape and length τ of f are fixed given.

Whatever we should use as an algorithm for solving this optimization task we definitely need a fast algorithm for solving the inner problem ("compute the populations of (2.11)") for each given pair (ω, E_r) (and all other fixed data). This is the field AQRS was designed for. Some real life applications should be taken exactly from this field as test problems to demonstrate the performance of AQRS.

5.1 Selective Vibrational Excitations

One way of controlling chemical processes of molecules is the selective excitation of a certain single bond of a multiatomic molecule, e.g. the selective excitation of the OD–bond in a  molecule. In this case the distinguished degrees of freedom are the vibrations of the molecule's bonds, which can be modelled as anharmonic oscillations. Thus, in this case, the molecular Hamiltonian H_0 is given by a sum of two coupled Hamiltonians of anharmonic oscillators, both based on Morse potentials (see Figure 13: Morse potential as a model for a single OH–bond and see [28] or [30] for details).

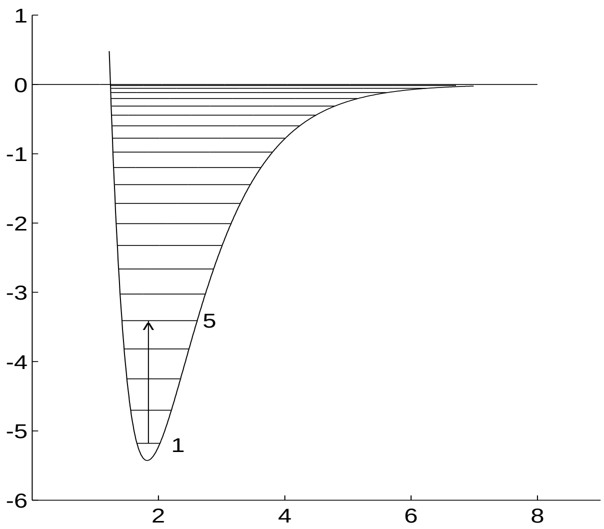


FIG. 13. Morse potential v of a single OH bond versus space coordinate x/a_0. The horizontal lines represent the energy ϵ_k of the vibrational eigenstates k of OH. Both, v and ϵ_k are given in eV. The vertical vector symbolizes the excitation from ground state 1 to the 5th state.

Some information about the vibrational eigenstates ϵ_k of this HOD–Hamiltonian H_0 is given in Figure 14.

Therein, the main location of the eigenfunctions φ_k is marked schematically. We detect states which are located "in the OD–bond" only (pseudo pure OD–vibrations), e.g. state $k = 24$. Thus, we can take state $k = 24$ as target state of our optimization task "Excite vibrations of the OD–bond only". Assuming the molecule's initial occupation to be $|c_k|^2(0) = \delta_{k1}$ (pure ground state occupation) we have got the following optimization task:
Determine (ω, E_r) *so that* $|c_{24}^{(\omega,E_r)}|^2(\tau) \approx 1$.

This problem can be solved. Figure 15 shows a section of the parameter space $(\omega, E_r) \in \mathbb{R} \times \mathbb{R}$ where $|c_{24}|^2(\tau)$ has got a maximum with $|c_{24}|^2(\tau) \approx 1$ (for more details see [35] or Figure 21 which shows the decay of $|c_1|^2(t)$ and the birth of $|c_{24}|^2(t)$ in the optimally controlled process).

Two problems from the controlled HOD–excitation task will serve as test problems for a detailed description of the performance of AQRS (see Table II).

5.2 Selective Isomerizations

For many molecules there are different stable configurations (isomers), e.g. $Be_2H_3D^-$ has got the isomers $Be_2H_3D^-(C_{2\nu})$ (with symmetry $C_{2\nu}$) and the slightly less stable one $Be_2H_3D^-(C_{3\nu})$. The structures of these two configurations are shown in Figure 16.

For controlling chemical reactions or for separating isotopes [11] it can be very helpful to *selectively switch* to the configuration of the molecule which e.g. is needed in the reaction [10]. In the case of $Be_2H_3D^-$ the ground state configuration is the $C_{2\nu}$ one. We are interested in a reaction path for a *selective isomerization* $C_{2\nu} \to C_{3\nu}$.

The original study of the $Be_2H_3D^-$ system [40] has shown, that both, $C_{2\nu}$ and $C_{3\nu}$, isomers can be converted into each other by rotations of the Be–H bonds. We restrict the investigation to these degrees of freedom (see [32] for details). From these assumptions a first simple, one–dimensional model Hamiltonian H_0 for the $Be_2H_3D^-$ isomerization can be deduced (see [32] again). Its investigation can be seen as a stimulating example for a favourable strategy for the isomerization of other anions. Moreover, we can

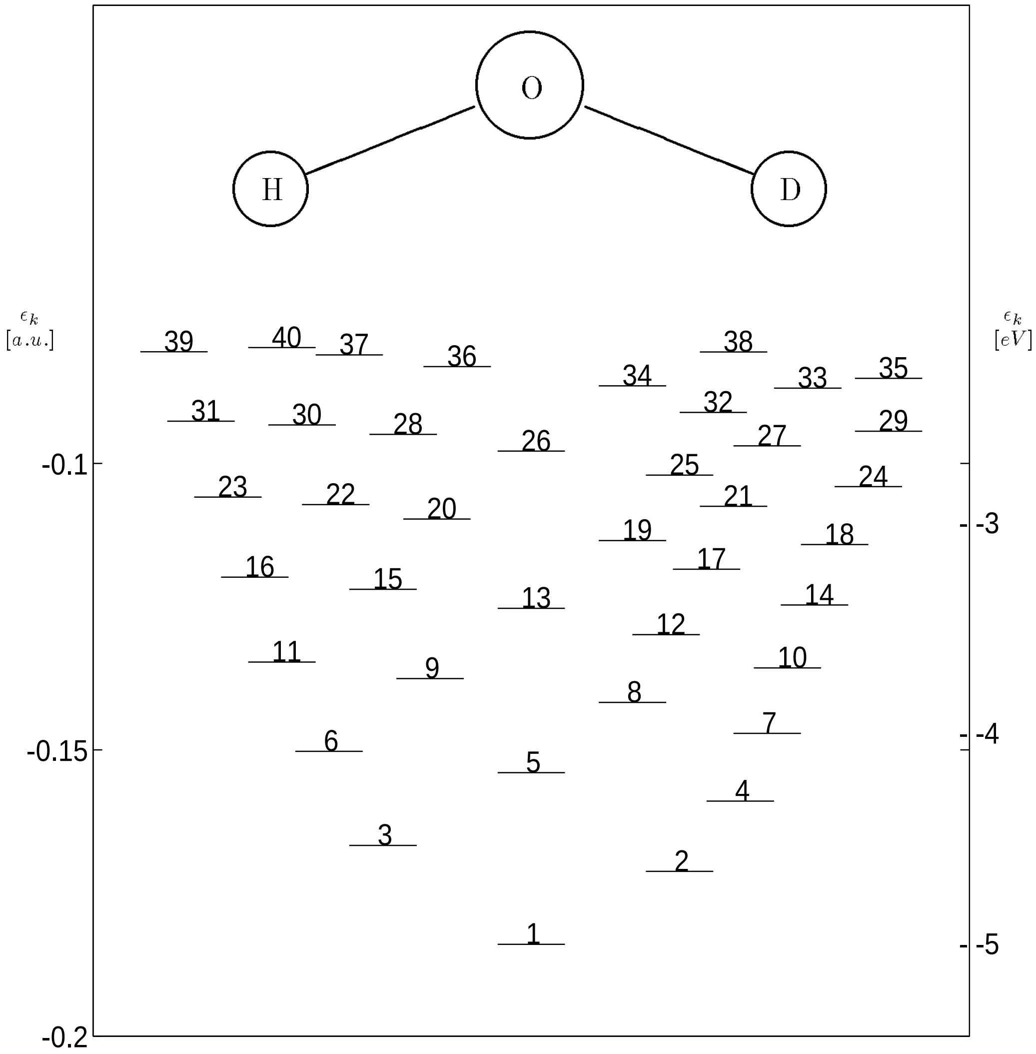


FIG. 14. Representation of the first 40 vibrational states of HOD. Each state is represented by its number k and by a line which marks its energy ϵ_k (which is given on left axis in atomic units, on the right in eV). The location of the lines with respect to the abscissa represent the center of the corresponding wavepacket, i.e. if state k appears on the left, then the OH bond is vibrating mainly, if it appears on the right hand side, then OD is vibrating.

again use (2.11) as the basic dynamical equation.

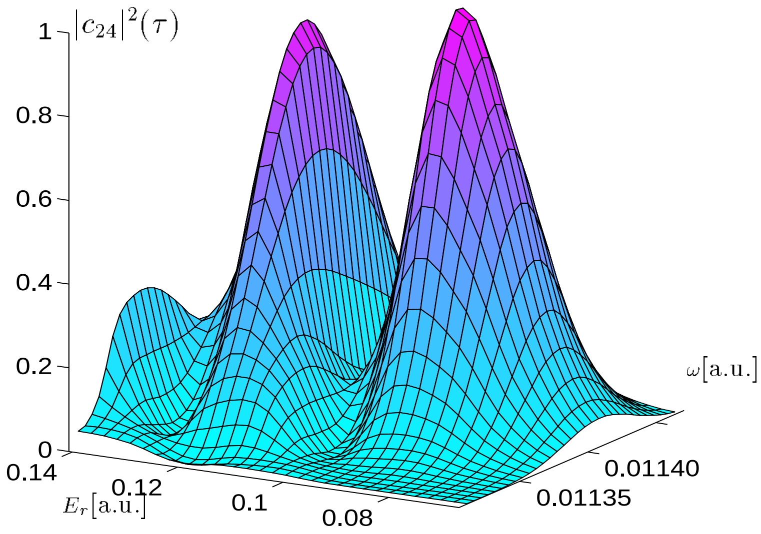


FIG. 15. Final population $|c_{24}|^2(\omega, E_r)$ of the 24th vibrational state of HOD versus ω and E_r in atomic units. Initial state was the ground state $k = 1$. Obviously several pronounced maxima lay close to each other.

The computations of the dipole matrix V_∞ and the molecular eigenenergies Ω are based on ab initio calculations of the potential energy surface and the dipole function for the electronic ground state of $Be_2H_3D^-$ at the M P4/6 – $31^{++}G^{**}$ level. In these calculations we observe that the potential energy surface is symmetric. This implies that the molecular eigenstates φ_k of H_0 have gerade or ungerade parity. Since the ground state $k = 1$ has gerade parity all states $2k + 1$ have so also, whereas all states $2k$ have ungerade parity. By usage of x-polarized light all matrix elements

$$V_{2k+1,2l} \;=\; - < \varphi_{2k+1}|\mu\ \varphi_{2l} > = \; V_{2l,2k+1} \tag{5.3}$$

between gerade and ungerade states $(2k+1)$ and $(2l)$ $(k, l \in N)$ vanish due to the gerade parity of μ. This implies: If we use x-polarized light and start with an initial occupation with gerade parity (i.e. $|c_1|^2(0) = 1$) we have to consider the gerade states and their dipole matrix V'_∞ only.

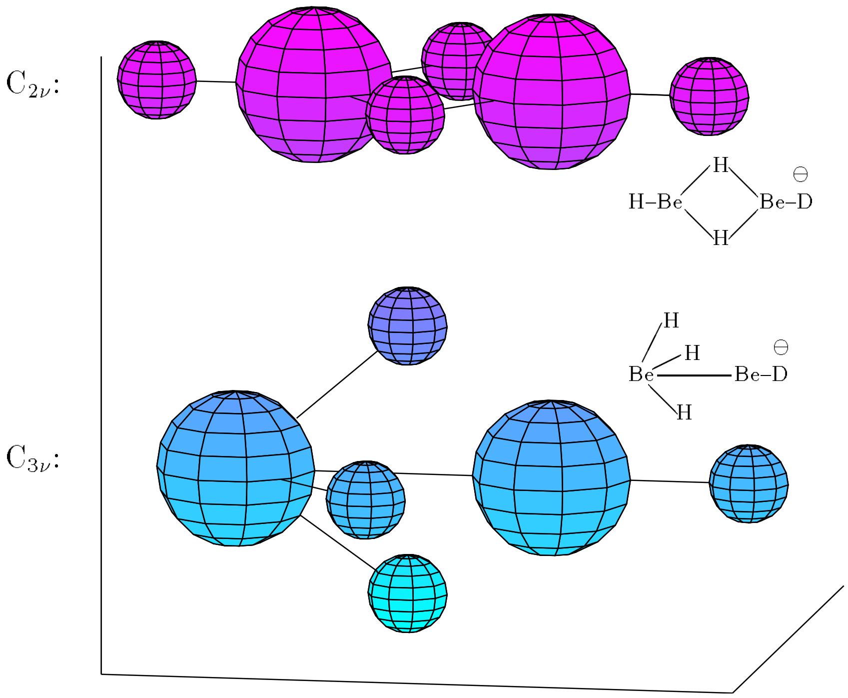


FIG. 16. Two isomers of $Be_2H_3D^-$: the stable configuration with $C_{2\nu}$ symmetry (top) and the slightly less stable one with $C_{3\nu}$ symmetry (bottom).

For demonstration of the internal structure of the gerade dipole matrix Figure 17 shows a level plot of V'_∞. The molecular eigenspectrum Ω is characterized in Figure 18 restricted to the states with gerade parity. Apparently, the states $k = 1, 3, 5, \ldots$ belong to the $C_{2\nu}$ isomer while the states $k = 7, 11, \ldots$ belong to the target $C_{3\nu}$ configuration. Others ($k = 21, 23, \ldots$) are delocalized.

Fixing the usage of x-polarized light the task of finding a selective isomerization reaction path can be formulated as: Start with $|c_1|^2(0) = 1$ and optimize $|c_{11}|^2(\tau)$. Unfortunately we cannot find a single optimal laser pulse with $|c_{11}|^2(\tau) \approx 1$. But we can succeed if we search for *series of picosecond laser pulses*. For example, Figure 19 shows such an optimized process: Four single pulses are iterated. Each of them corresponds to a selective excitation

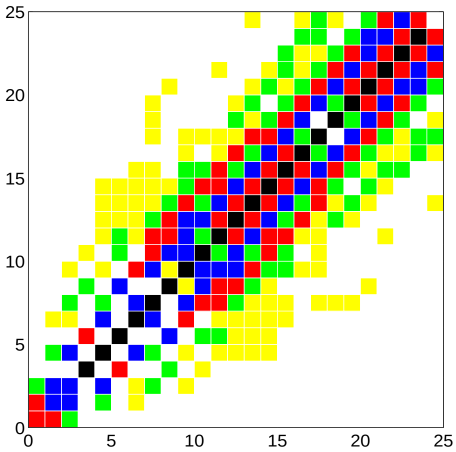


FIG. 17. Level plot of the dipole matrix V'_∞ of the Be1 test problem. The darker the square at (k,l) the larger is $|V'_{kl}|$: the large values of V'_∞ are concentrated near its diagonal.

(Pulse I: 1→ 3, Pulse II: 3→ 15, Pulse III: 15→ 23, and Pulse IV: 23→ 11). Their succession effects the required selective isomerization $C_{2\nu} \to C_{3\nu}$.

Three problems from this "controlled $Be_2H_3D^-$ isomerization task" will serve as test problems for a detailed description of the performance of AQRS (see Table II). Two of them ('Be1' and 'Be3') represents the case of x-polarized light and a pure interaction of the states with gerade parity. 'Be1' corresponds to above Pulse II. The other ('Be2') represents the "full" problem including all states and the excitation of an ungerade state starting from an gerade initial occupation.

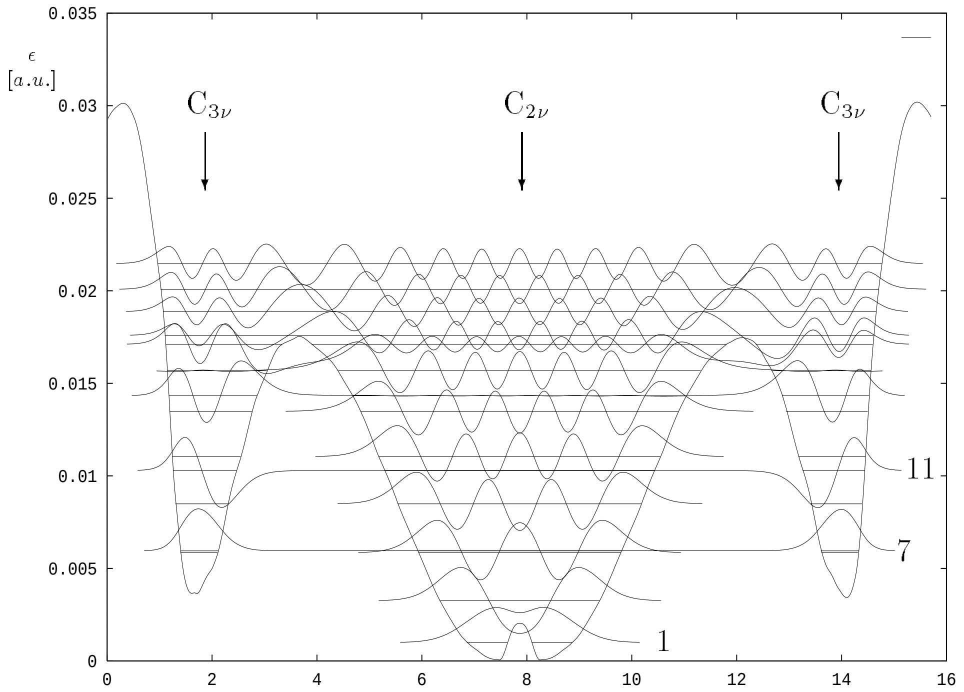


FIG. 18. Vibrational potential of $Be_2H_3D^-$ together with the energies ϵ_{2k+1} and the eigenvalues φ_{2k+1} of its vibrational eigenstates with gerade parity. The potential minima on the left and the right represent the configuration with $C_{3\nu}$ symmetry while the main minimum corresponds to the stable $C_{2\nu}$ isomer. Note that the eigenstates $\varphi_{2k+1}(x)$ of some states (e.g. 7 or 11) are mainly localized in the $C_{3\nu}$–minima while others (e.g. 1) are localized in the main minimum only. Thus, the states 7 or 11, for example, represent the $C_{3\nu}$–isomer and the ground state 1, for example, the stable $C_{2\nu}$–isomer.

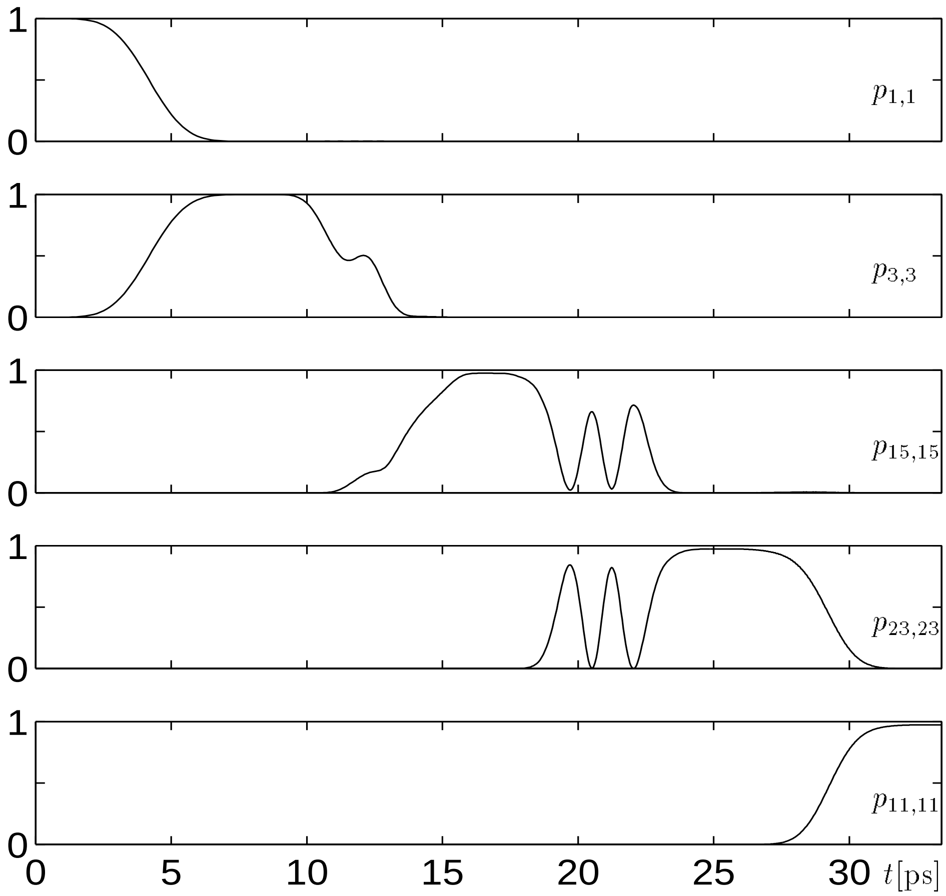


Fig. 19. Isomerization process for $Be_2H_3D^-$ consisting of four laser pulses each 8.35 ps long. Each subfigure belongs to the evolution of the population of a single state. You see the states $1 \to 3 \to 15 \to 23 \to 11$ being totally occupied one after the other. Anyone, for whom the 33.4 picoseconds needed for this isomerization appear too long, may look for a shorter process in [32].

5.3 Performance of AQRS

Selective vibrational excitation and selective isomerization are two fields in which Quantum Chemical modelling leads to the basic equation (2.11). AQRS is designed for efficiently computing the populations of (2.11), i.e. for solving the corresponding inner problem. Hence, we demonstrate the performance of AQRS in real life problems from the fields "vibrational excitation of HOD" and "selective isomerization of $Be_2H_3D^-$". Table II presents the chosen test problems. Note that they represent all typical situations which have been evoked in the previous subsections. The last test problem ('Be3') is something of a rare case: the field amplitude E_r is extraordinarily high. 'Be3' is separated from the other ones and discussed at the end of this subsection.

The performance of AQRS in these test problems is presented in Figures 20-23 on pages 115-118 and Figure 25 on page 120. Altogether, these figures are described in Table IV on page 114. Each figure is divided into 6 subfigures, fills one page and shows data which is significant for the AQRS–performance on the current test problem. These significant data contain answers to the most important questions concerning AQRS:

- Are the populations computed by AQRS sufficient approximations of the smoothed exact populations? (Subfigures 1&2)
- How do the δ–control and the local relative accuracy control work? (Subfigures 3&4)
- Does AQRS essentially reduce the computational complexity of solving the original problem (2.11) (which is represented by the QRS(∞)–solution)? (Subfigure 5)
- How much of the computational complexity of AQRS is caused by the error estimator, how much by the pure QRS–subintegrations? (Subfigure 6)

The control parameters of AQRS are equally fixed over all presented computations: the allowed error percentage was set to $\theta = 0.1$, the recursion depth limit to $m = 10$. For all other control parameters the entries in Table I are used. In all cases the subinterval length $T = 2j\pi/\omega$ was fixed by the choice $j = 6$ (in no case this T–choice was corrected by the checking procedure).

In addition to the graphical answer to the four questions named, Table III

problem name	description of problem	parameters ω [d.u.]	ω [a.u.]	E_r [a.u.]	s	τ [ps]	dim. (ODE)
OH	single OH bound excitation $1 \to 5$	0.8866	0.0160	0.0658	1	1.0	44
HOD	HOD–molecule excitation $1 \to 24$	0.89752	0.0114	0.0960	1	1.0	120
Be1	$Be_2H_3D^-$ gerade, pulse II of isomerization	1.13064	0.00255	0.0152	2	8.35	50
Be2	$Be_2H_3D^-$ full, nonoptimal excitation $5 \to 10$	2.0628	0.00128	0.00712	5	8.35	100
Be3	$Be_2H_3D^-$ gerade, isomerization, pulse II in [32]	3.383	0.00763	0.2443	3	1	50

TABLE II. Specification of the considered test problems. The laser frequency ω is given in dimensionless units ([d.u.]) described in Section 2.1 and in atomic units ([a.u.]) normally chosen in Quantum Chemistry (ω[a.u.]= $1 \to \omega$[Hz]=$4.152 \cdot 10^{16}$). In all test problems the laser shape f is a $\sin^2$–function; its field amplitude E_r is given in [a.u.] (E_r[a.u.]= $1 \to E_r$[V/m]=$5.142 \cdot 10^{11}$), its pulse length τ in picoseconds ([ps]). s denotes the number of the initial state of the considered test problem. The last column gives the dimension of the problem's basic ODE (2.11). In all cases V_∞ and Ω are computed as explained above. Note the extremely high field amplitude E_r in the last problem 'Be3'.

contains conclusive comparisons of the different total computational complexities.

problem	$\left(\frac{\mathcal{C}_{P,I}}{C_{P,I}(\infty)}\right)^{-1}$	$\frac{\mathcal{E}_{P,I}}{\mathcal{C}_{P,I}} \times 100\%$	speedup factor
OH	120	14 %	90
HOD	30	0.4 %	40
Be1	180	26 %	150
Be2	330	23 %	285
Be3	12	29 %	10

TABLE III. Computational effort caused by AQRS on the whole integration interval I for the test problems from Table II. The first column contains the reduction factor by which AQRS reduces the computational complexity of the "exact" QRS(∞)–solution. The second column gives the share of computational effort produced by error estimation in AQRS in per cent. For a description of the complexity measures $C_{P,I}$, $\mathcal{C}_{P,I}$ and $\mathcal{E}_{P,I}$ see pages 53 and 98. In the last column the real–time speedup factor for AQRS in comparison with QRS(∞) is listed. 'Be3' represents the rare case "switching to symmetric mode" because of extremely high laser field strengths (see page 113).

Altogether, these tables and figures allow the following conclusions:

1. AQRS is a smoothing integrator. Its results give sufficient approximations of $\operatorname{diag} \Pi(t)$, the smoothed evolution of the populations of (2.11).

2. In comparison to the usual (QRS(∞))–integration of (2.11) AQRS obtains a speedup factor of the order of 10^2 (for field amplitudes E_r which are not *extremely* high).

3. The computational complexity of the QRS–subintegrations in AQRS strongly dominate that complexity caused by the computations in error estimation scheme.

The presentation of the performance of AQRS in our first four test problems still leaves some important questions unanswered:

First of all, let us discuss the case of *nonresonant* excitations. AQRS is intended to be used as the internal integrator in an optimization procedure used for determining the absolute maximum of the population $p_k(\omega, E_r, \tau) = |c_k|^2(\omega, E_r, \tau)$ in the parameter space $(\omega, E_r) \in \mathbb{R} \times \mathbb{R}$. Altogether, our test problems correspond to resonant cases (with $|c_k|^2(\omega', E_r', \tau) \gg 0$). Naturally, *nonresonant* parameter sets (ω', E_r') will also appear in such optimization procedures ($|c_k|^2(\omega', E_r', \tau) \ll 1$). Is AQRS still *reliable* and *efficient* in such nonresonant cases? The reliability is already demonstrated by Figure 15 which is computed using AQRS and which can identically be reproduced using QRS(∞), i.e. an "exact" integrator. In all computations needed for Figure 15 (for all used parameter sets (ω, E_r)) the speedup factor of AQRS in comparison to QRS(∞) was about 30, in resonant and in nonresonant cases. Thus, AQRS is in nonresonant cases as highly efficient as in resonant ones. A similar result is found in all considered test problems.

Second, we should still ask for a direct comparison of the *estimated* relative error $\langle\varepsilon\rangle = \langle\epsilon_\Sigma\rangle(t,\delta)/\langle\epsilon_\Sigma\rangle(t,0)$ and the *exact* relative error $\varepsilon = \epsilon_\Sigma(t,\delta)/\epsilon_\Sigma(t,0)$: Can the estimate be taken as a good approximation of the original? Figure 24 shows a comparison of $\langle\varepsilon\rangle(t,\delta)$ and $\varepsilon(t,\delta)$ produced for the test problem 'HOD': the estimate $\langle\varepsilon\rangle$ reproduces the original ε beautifully. Again, a similar result is found in all considered test problems.

Last, but not least, an example should be given in which the accuracy requirement forces AQRS to leave its "smoothing domain" $\delta < 1$. This is the case in our last test problem 'Be3':

'Be3' is taken from a series of pulses leading to an isomerization of $Be_2H_3D^-$ ($C_{2\nu} \rightarrow C_{3\nu}$) (cf. Pulse II in [32]). Figure 25 shows the performance of AQRS on the test problem 'Be3'. We observe that AQRS is forced to switch to its *symmetric mode* in some subintervals. In these subintervals AQRS cannot work as a smoothing integrator: in the symmetric mode elements V_{kl} of the dipole matrix are considered which correspond to e.g. $|\Delta_{kl}^{-}| > \omega$. But although AQRS doesn't work as a smoothing integrator it obtains a speedup factor of about 10 (in comparison with QRS(∞)). This is caused by the usage of sparsed dipole matrices only; it is not due to stepsize increases. The right bottom subfigure of Figure 25 shows the relative error estimate computed by AQRS *before* the switch to the symmetric mode is done. The estimated relative errors reach up to 40%, an effect we can explain looking back to Figure 12: Dipole elements V_{kl} with $|\Delta_{kl}^{\pm}/\omega| \approx 1$ (corresponding to $\delta \approx 1$) play a central role in the process. For them, both the $\pm$–terms give similar contributions to Π_{kk} as Figure 12 shows. Thus, we cause an error of up to 50%, if we only consider the term $\min_{\pm} |\Delta_{kl}^{\pm}|$.

Some selected components of the "exact" QRS(∞)–populations $p(t)$ of the considered problem as functions of time t in picoseconds.	AQRS–populations $p(t)$ of the considered problem as functions of time t; same components as in the left subfigure.
Sequence $(\delta_0, \ldots, \delta_N)$ of the δ–values $\delta_k = \delta_{I_k}$ which are chosen by AQRS on the subintervals $I_0 \ldots I_N$, $I_k = [kT, (k+1)T]$.	Estimated relative local error $(\varepsilon_{\delta_k}/\varepsilon_0)(t)$ computed in the local relative accuracy control of AQRS as a function of t. Its upper bound θ (allowed error percentage) is shown as a dotted line. See page 92 for an explanation of the relative local error scheme.
Evolution of the cumulative computational effort produced by AQRS for problem P until time t in comparison with that of QRS(∞): $\mathcal{I}_{P,[0,t]}/C_{P,[0,t]}(\infty)$ and $\mathcal{C}_{P,[0,t]}/C_{P,[0,t]}(\infty)$ as functions of t. See pages 53 and 98 for a definition of the complexity measures.	Evolution of the computational effort produced by the AQRS–error estimator for problem P on subinterval I_k in comparison with the pure AQRS–integration effort on I_k: $\mathcal{E}_{P,I_k}/\mathcal{I}_{P,I_k}$ as a function of t. See page 98 for a definition of the complexity measures.

TABLE IV. Explanation of the figures on the following pages: The 6 subfigures of each figure contain essential information about the AQRS–performance on the 4 considered test problems.

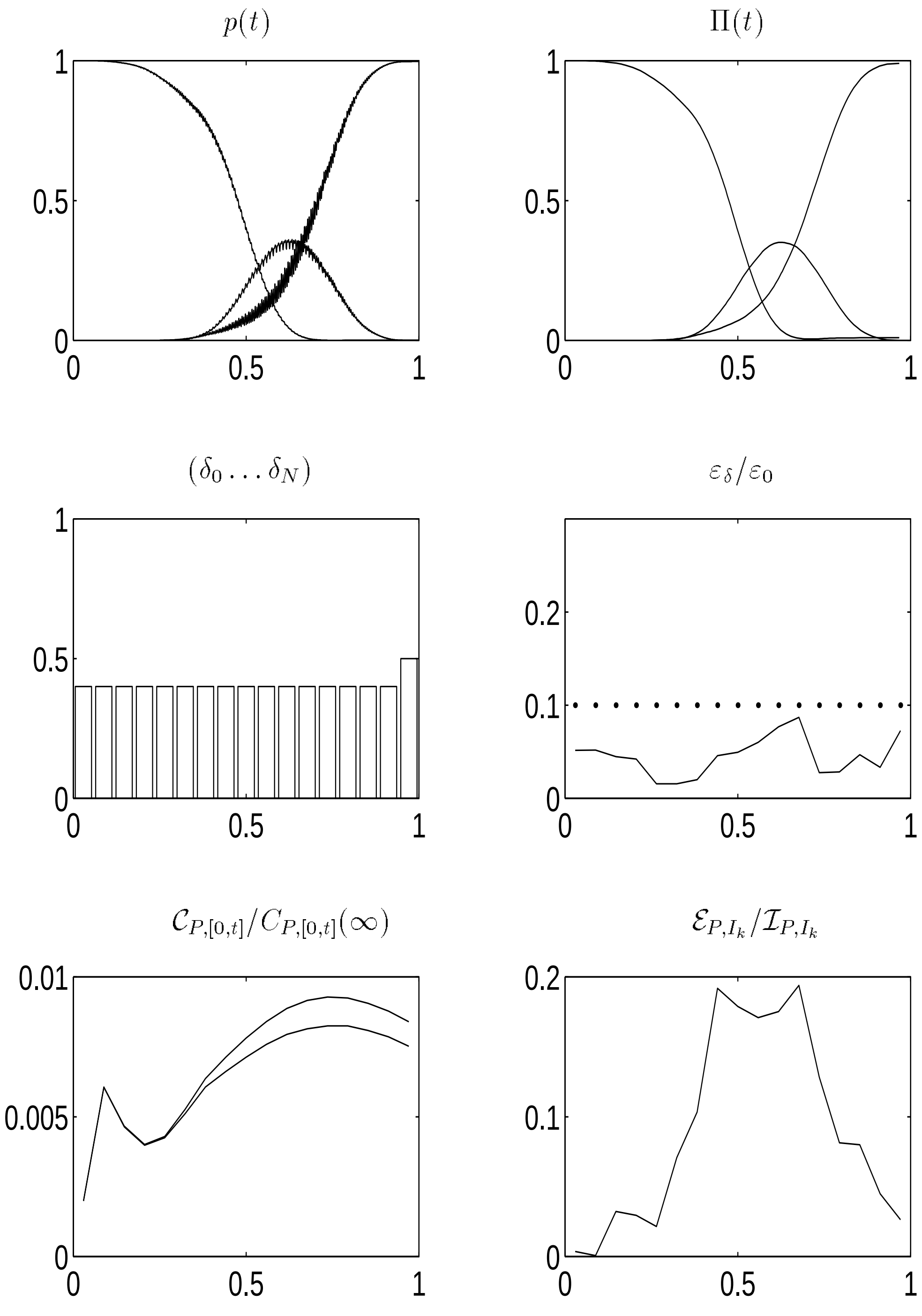


FIG. 20. Documentation of AQRS performance for problem 'OH'. See Table IV on page 114 for description of the single subfigures.

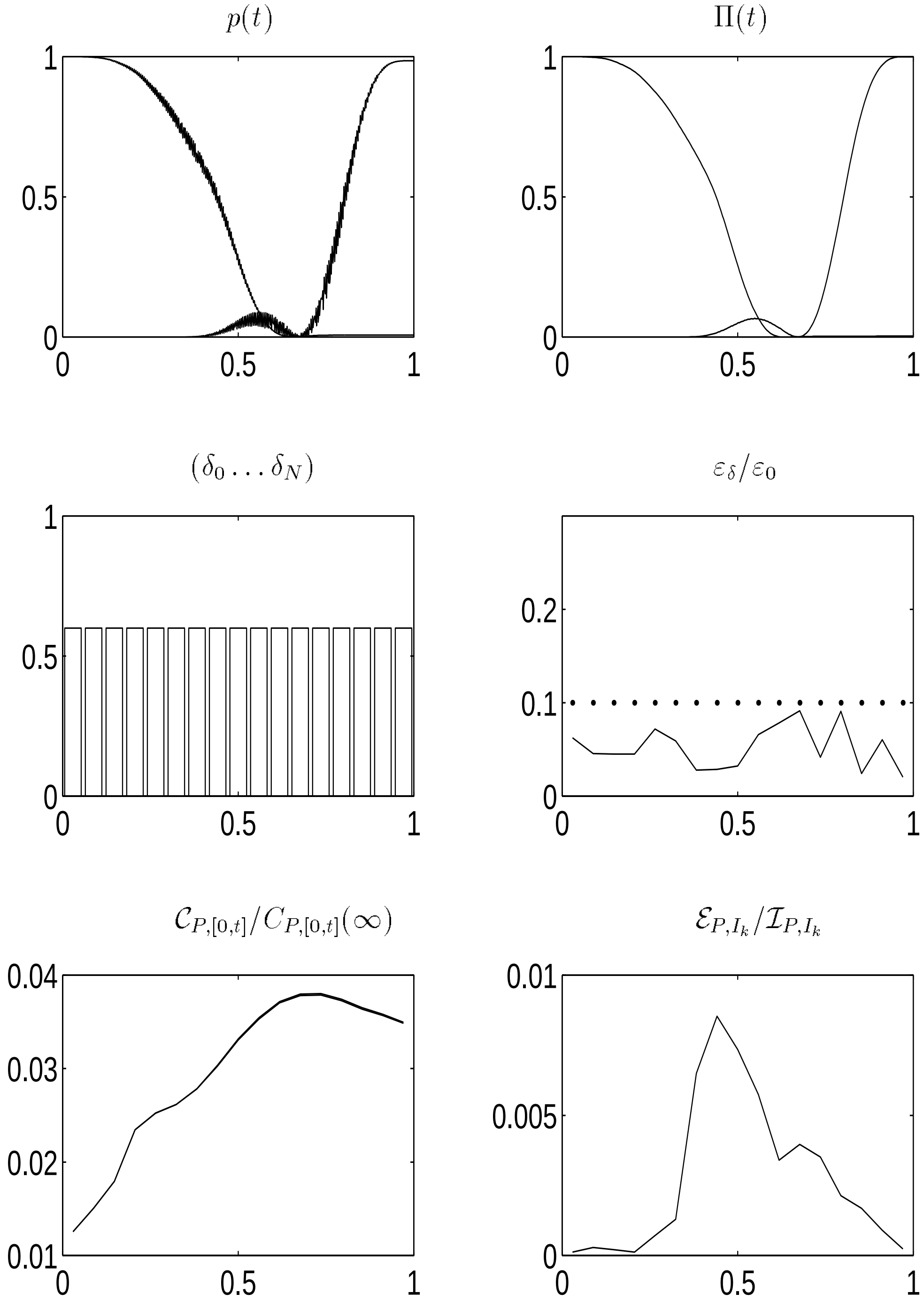


FIG. 21. Documentation of AQRS performance for problem 'HOD'. See Table IV on page 114 for description of the single subfigures.

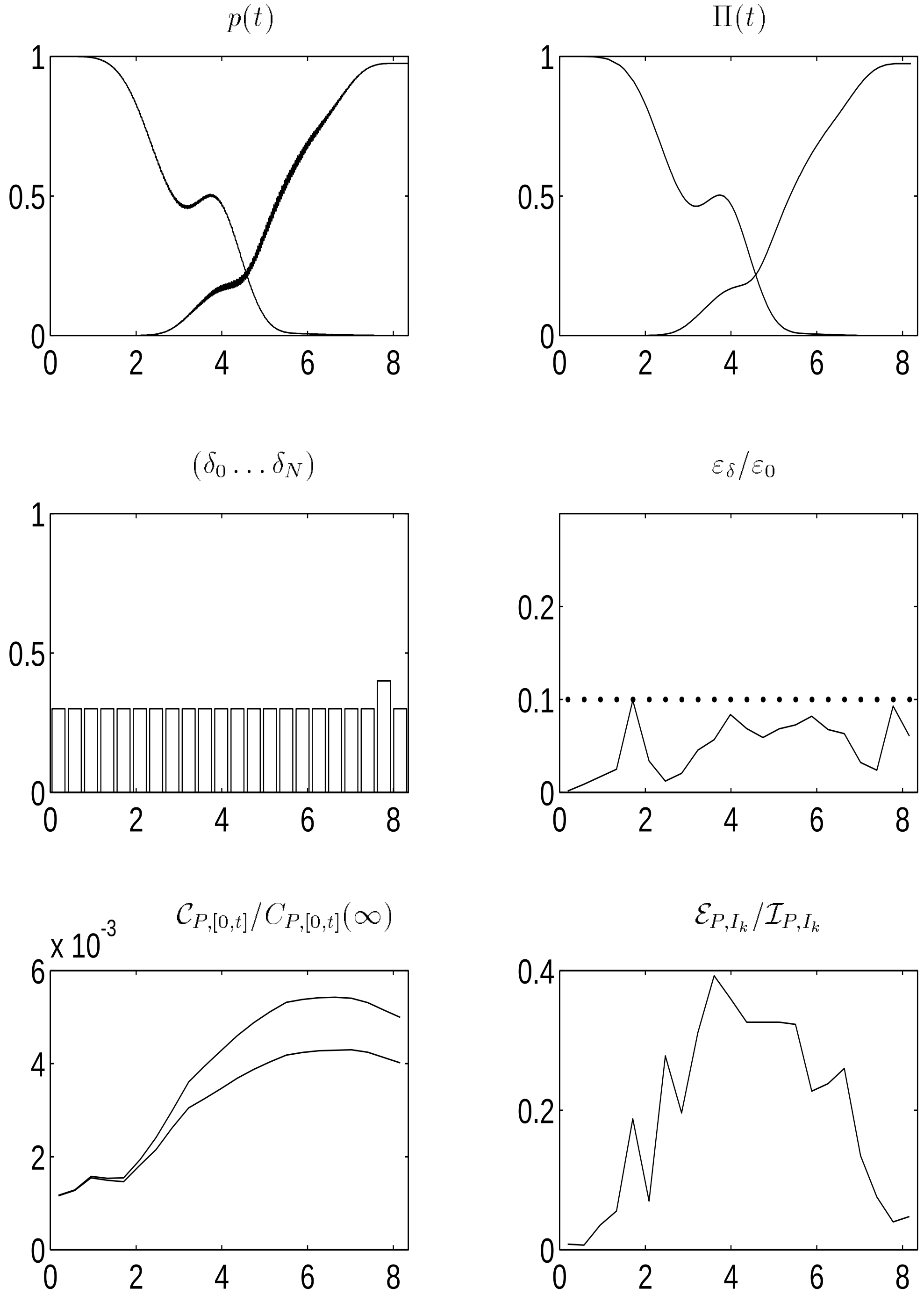


Fig. 22. Documentation of AQRS performance for problem 'Be1'. See Table IV on page 114 for description of the single subfigures.

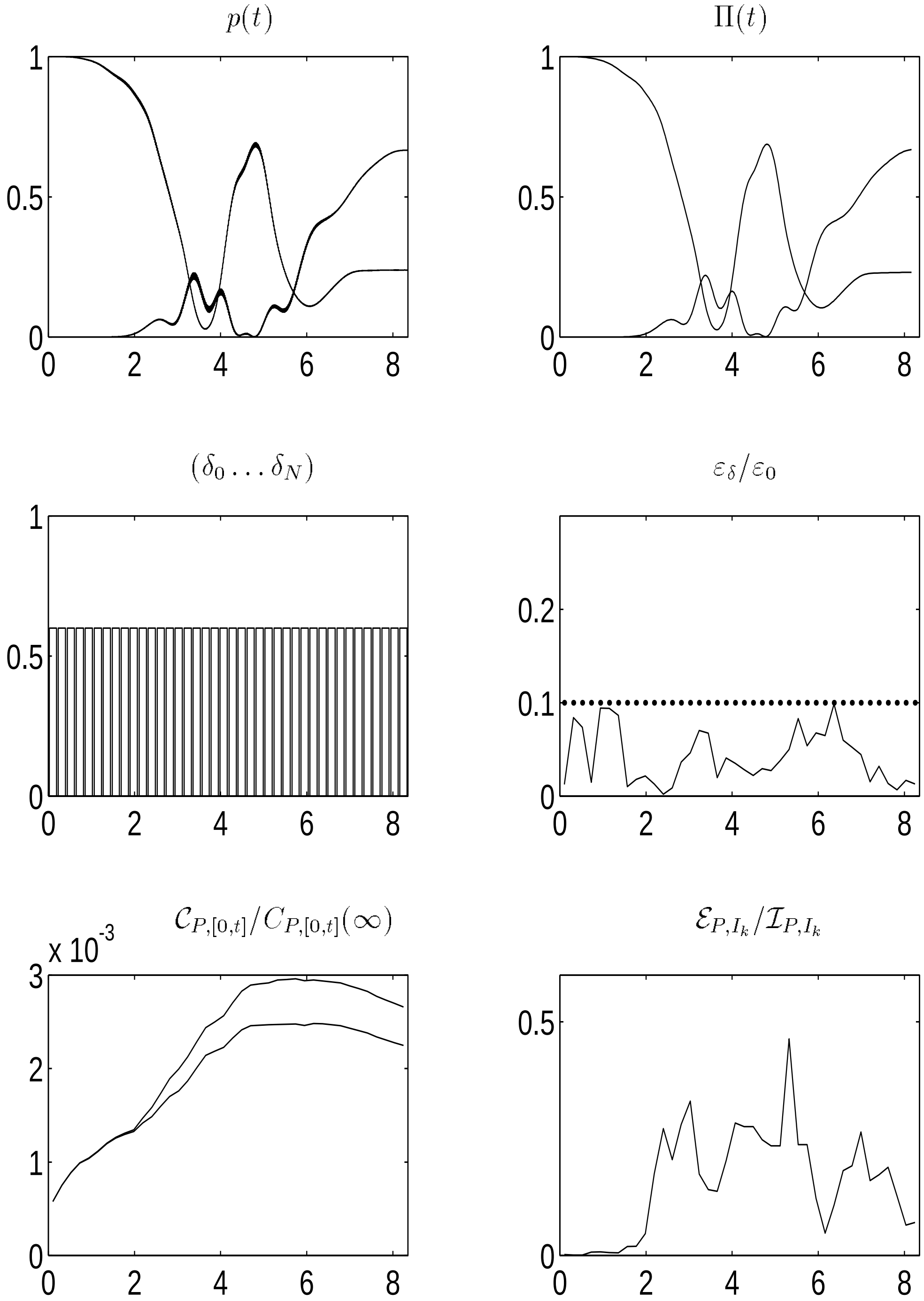


FIG. 23. Documentation of AQRS performance for problem 'Be2'. See Table IV on page 114 for description of the single subfigures.

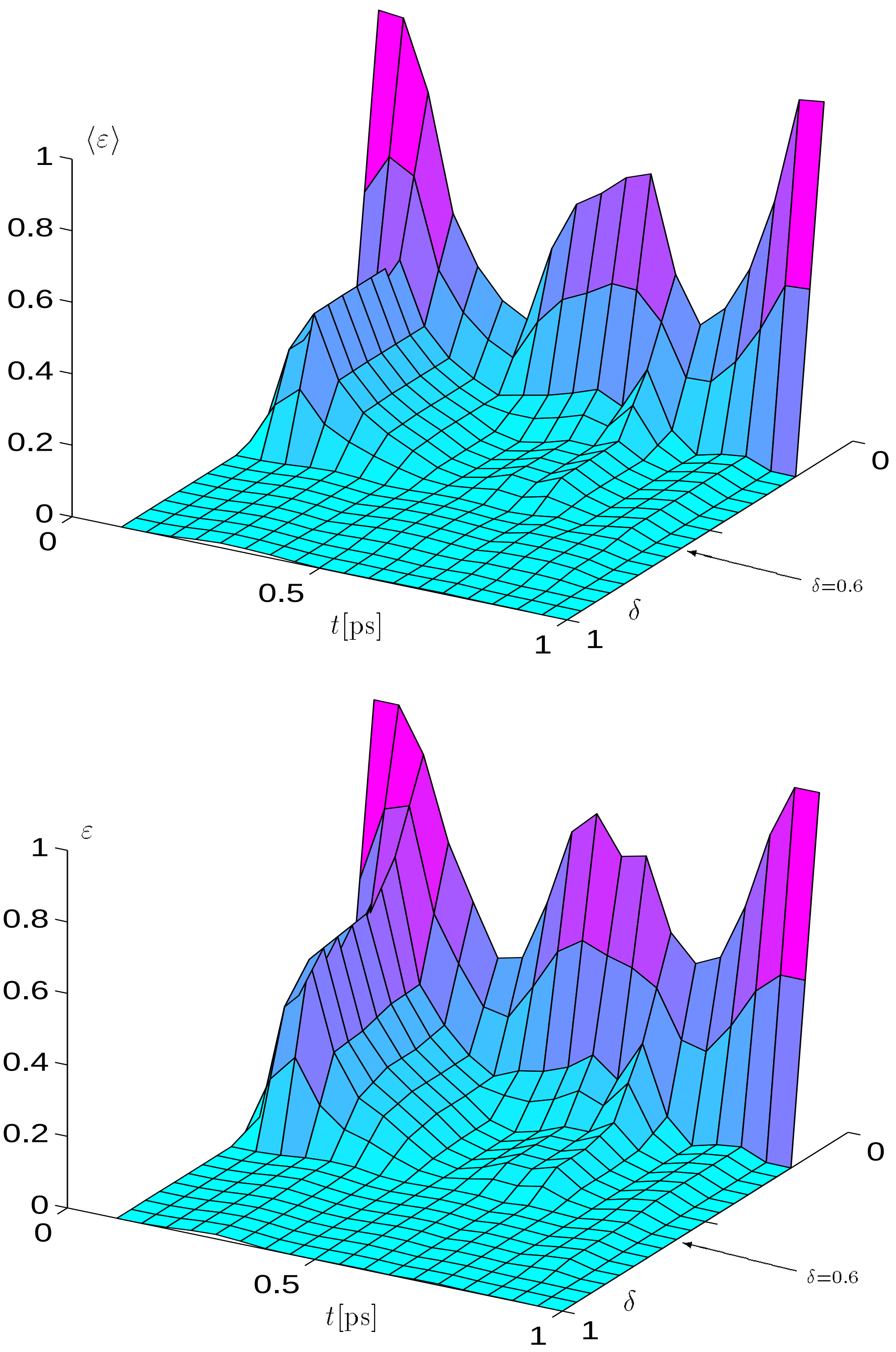


FIG. 24. Comparison of estimated relative error $\langle \varepsilon \rangle = \langle \epsilon_\Sigma \rangle(t,\delta)/\langle \epsilon_\Sigma \rangle(t,0)$ and exact relative error $\varepsilon = \epsilon_\Sigma(t,\delta)/\epsilon_\Sigma(t,0)$ for HOD.

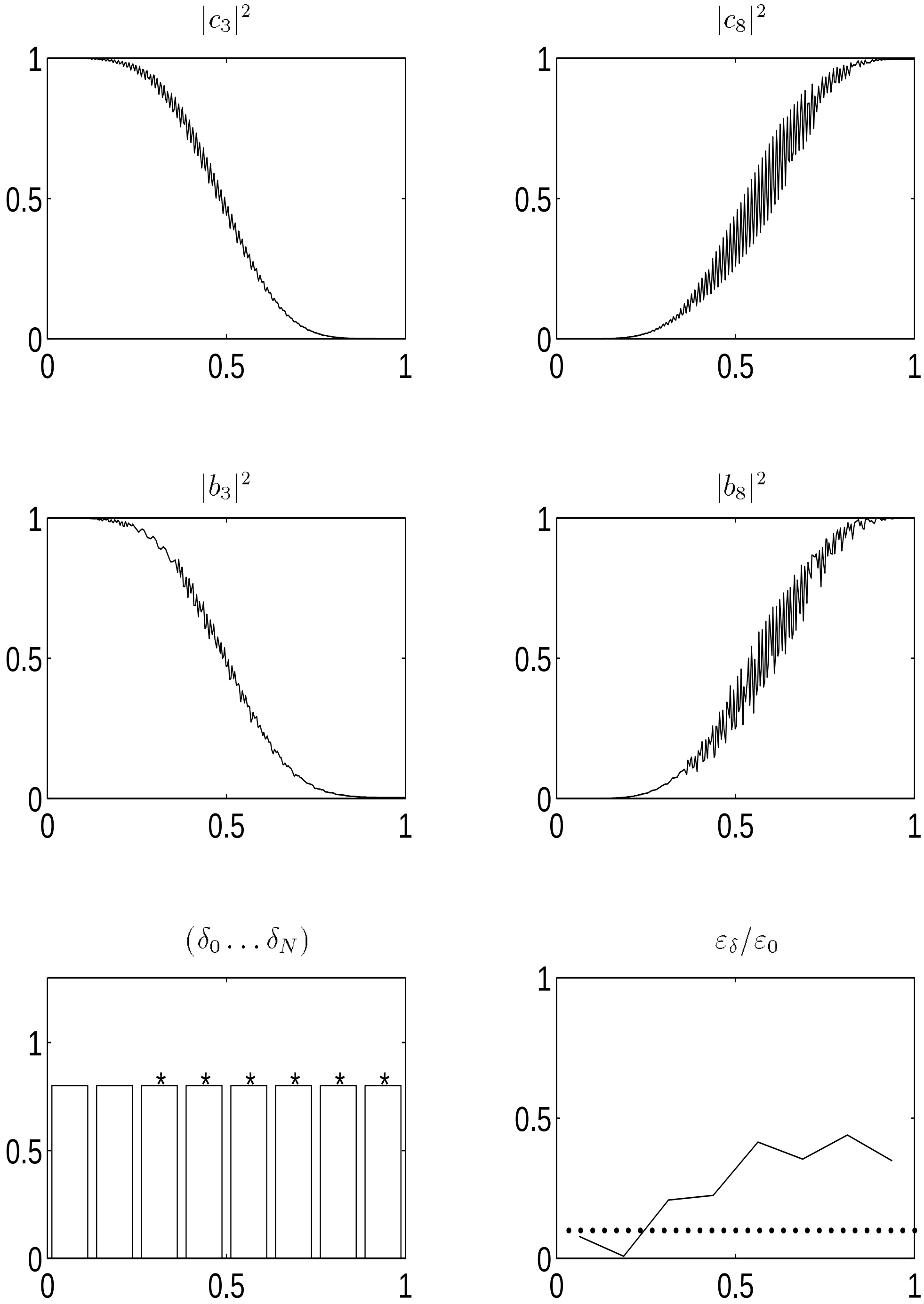


FIG. 25. Documentation of the performance of AQRS on test problem 'Be3'. The top 4 subfigures show the evolution of the exact ($|c_k|^2$) and AQRS–populations ($|b_k|^2$) of initial state 3 and target state 8. Bottom left the δ–sequence chosen by AQRS is shown (with separation of the single subintervals). In subintervals carrying a '$*$' AQRS was forced to switch to its symmetric mode. Bottom right the relative error estimate before switching to symmetric mode is shown.

6 Conclusions

We presented a new adaptive algorithm for an efficient simulation of the influence of short laser pulses on the dynamical behaviour of single molecules. We observed that the solution of this simulation task is the indispensable key problem with respect to the realization of a laser–assisted, optimally selective vibrational excitation of a molecule.

The quantum chemical modelling of these laser–molecule interactions led us to large ODEs $i\dot{c} = (\Omega + f(t)\cos\omega t\, V)c$. We saw that their solutions $c = (c_k)$ show multifrequency highly oscillatory behaviour, and that the corresponding populations $|c_k|^2$, or their running averages $\mathcal{A}_T\,|c_k|^2$ respectively, contain all information of chemical relevance. A discussion of the applicability of numerical methods which succeed in other highly oscillatory contexts resulted in the conclusion that all these methods are inappropriate for solving this kind of problem efficiently. Thus, the efficient evaluation of these averaged populations $\mathcal{A}_T\,|c_k|^2$ was discovered to be the numerical challenge.

Starting from an inspection in the physical interpretation of resonant transition conditions in light–matter interaction processes, we developed the basic quasiresonant smoothing idea which claims that dropping specific elements of the ODE's interaction matrix V does not disturb the evolution of its populations. Indeed, this idea allowed the construction of <u>q</u>uasi<u>r</u>esonant <u>s</u>moothing algorithms QRS(δ). In QRS(δ), the proposed sparsing of V is controlled by a splitting parameter δ and leads to smooth solutions and smooth populations. The smoothness of the solutions allows comparatively large stepsizes while the smooth populations approximate the running average $\mathcal{A}_T\,|c_k|^2$ of the original ones. We observed that an integration using QRS(δ) with a well–chosen δ obtains a large gain in efficiency in comparison with standard integration of the original ODE.

We used perturbation analysis of the effect of V on the running average of the populations for developing an error estimator which allows an adaptive choice of the splitting parameter δ with respect to a given accuracy requirement. This led us to an adaptive QRS–version (AQRS) which can be used as an efficient and reliable black box integrator for our kind of problem. Its reliability was proved for a subclass of problems for which the error estimate is exact. The performance of AQRS in some test problems demonstrated its efficiency and reliability in real life applications: in comparison with standard integrators it obtained speedup factors of the order of 10^2.

Of course this is not the end of the story. The presented results draw attention to further perspectives:

- So far, AQRS was not applied to *very large* real life problems, because today extended chemical models are not available. This is because the quantum chemical preparation of such models itself is an ambitious problem. In the framework of [32] a first step in this direction will be done.

- We observed that the simulation of laser–molecule interaction is only the inner problem of a chemically fundamental optimization task ("determine the most selective laser pulse"). Thus, the most important extension of our method is its incorporation into an adapted optimization framework. This is already prepared by the object–oriented implementation of AQRS and QRS(δ) and will be realized next.

Symbols

$\mathbb{R}$, $\mathbb{C}$	real and complex numbers
$\mathbb{C}^{n\times n}$	complex $n \times n$ matrices
$v \otimes w$	usual tensor product of vectors v, w
n	number of states of the considered physical model
Ω	diagonal matrix of the eigenfrequencies of the molecule
ω	light–frequency of the external laser pulse
$f(t)$	shape function of the external laser pulse
V_∞	matrix representation of the full electric dipole operator of the molecule
$\mathcal{A}_T$	local time average operator, see page 41
$c(t)$	vector of expansion coefficients (see page 14), solution of the basic equation of motion (2.11)
$\gamma(t)$	local time average of $c(t)$: $\gamma = \mathcal{A}_T c$
$b(t)$	interaction picture expansion coefficients: $b(t) = c_I(t)$, see page 60
$\beta(t)$	local time average of $b(t)$: $\beta = \mathcal{A}_T b$
$p(t)$	probability matrix of the problem: $p = c \otimes \bar{c}$, see page 16
$\pi(t)$	local time average of $p(t)$: $\pi = \mathcal{A}_T p$
$\Pi(t)$	local time average of $p_I(t)$: $\Pi = \mathcal{A}_T p_I$
$\tilde{\Pi}(t)$	approximative local time average of $p_I(t)$, see page 69.
$c\{\mathcal{V}\}(t)$	solution of the basic equation (2.11) using dipole matrix $\mathcal{V}$, see 71.
$\Pi\{\mathcal{V}\}(t)$	probability matrix using dipole matrix $\mathcal{V}$
$\epsilon_\Phi(t,\delta)$	sparsing error measure for smoothed probabilities
$\langle \epsilon_\Phi(t,\delta) \rangle$	estimate for $\epsilon_\Phi(t,\delta)$
$U_0(t,s)$	time propagator of the non-disturbed problem, see page 58
$\mathcal{S}_{R,I}$	real– or imaginary part interaction skeleton, see page 84.

References

[1] V.I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations*. Springer–Verlag, Berlin, Heidelberg, New York, Tokyo, 2nd edition, 1988.

[2] A.D. Bandrauk. *Atomic and molecular processes with short intense laser pulses*. Plenum Press, New York, first edition, 1988.

[3] N. Bloembergen and A.H. Zewail. Energy redistribution in isolated molecules and the question of mode–selective laser chemistry revised. *J. Phys. Chem.*, 88:5459–5465, 1984.

[4] N.N. Bogolyubov and V.S. Korolyuk. The research of Y.A. Mitropolski in the realm of the theory of nonlinear oscillations. *Ukr. Math. J.*, 29:3–14, 1977.

[5] H.P. Breuer, K. Dietz, and M. Holthaus. Selective excitations of molecular vibrations by interference of floquet states. Preprint SC 93-10, Universität Bonn, Physikalisches Institut, 1990.

[6] H.P. Breuer, K. Dietz, and M. Holthaus. *Radiation Effects and Defects in Solids*, 122-123:91–106, 1991.

[7] H.P. Breuer, K. Dietz, and M. Holthaus. Selective excitations of molecular vibrations by interference of floquet states. *J. Phys. B*, 24:1343–1357, 1991.

[8] M.M. Chawla and P.S. Rao. Noumerov–type methods with minimal phase–lag for the integration of second order periodic initial–value problems. *J. Comput. Appl. Math.*, 11:277–281, 1984.

[9] M.M. Chawla and P.S. Rao. Noumerov–type methods with minimal phase–lag for the integration of second order periodic initial–value problems ii: explicit methods. *J. Comput. Appl. Math.*, 15:329–337, 1986.

[10] J.E. Combariza, B. Just, J. Manz, and G.K. Paramonov. Isomerization controlled by ultrashort infrared laser pulses: Model simulations. *J. Phys. Chem.*, 95:10351, 1991.

[11] J.E. Combariza and other, editors. *Isotope Effects in Gas–Phase Chemistry*, volume 502 of *ASC Symposium Series*. American Chemical Society, Washington, 1992.

[12] G. Denk. A new numerical method for the integration of highly oscillatory second–order ordinary differential equations. *Applied Numerical Mathematics*, 13:57–67, 1993.

[13] P. Deuflhard. A study of extrapolation methods based on multistep schemes without parasitic solutions. *Angewandte Mathematische Physik*, 30:177–189, 1979.

[14] P. Deuflhard. Order and stepsize control in extrapolation methods. *Numer. Math.*, 41:399–422, 1983.

[15] P. Deuflhard. Recent progress in extrapolation methods for ordinary differential equations. *SIAM Rev.*, 27:505–535, 1985.

[16] P. Deuflhard and A. Hohmann. *Numerische Mathematik I — Eine algorithmisch orientierte Einführung*. Walter de Gruyter, Berlin, New York, 2nd edition, 1993.

[17] Z.E. Dolya, N.B. Nazarova, G. Paramonov, and V. Sacca. Localization of population at specific vibrational levels of molecules pumped by ultrashort IR laser pulses. *Chem. Phys. Lett.*, 145(6):499–504, 1988.

[18] W. Gautschi. Numerical integration of ordinary differential equations based on trigonometric polynomials. *Numer. Math.*, 3:381–397, 1961.

[19] C.W. Gear and K.A. Gallivan. Automatic methods for highly oscillatory ordinary differential equations. Scientific report, University Urbana, 1982.

[20] O.F. Graff and D.G. Bettis. Modified multirevolution integration methods for satellite orbit computation. *Celestial Mechanics*, 11:443–448, 1975.

[21] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I, Nonstiff Problems*. Springer Verlag, Berlin, Heidelberg, New York, Tokyo, 1987.

[22] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems.* Springer Verlag, Berlin, Heidelberg, New York, Tokyo, 1991.

[23] B. Hartke, J. Manz, and J. Mathies. Mode–selective control of unimolecular dissociation:... *Chem. Phys.*, 139:123, 1989.

[24] D.J. Higham and L.N. Trefethen. Stiffness of ODEs. *BIT*, 33:285–303, 1993.

[25] A. Hohmann. An implementation of extrapolation codes in C++. Technical Report TR 93–8, Konrad-Zuse-Zentrum, Berlin, 1993.

[26] A. Hohmann and C. Wulff. Modular design of extrapolation codes. Technical Report TR 92-5, Konrad-Zuse-Zentrum, Berlin, 1992.

[27] D.G. Imre and J. Zhang. Dynamics and selective bond–breaking in photodissociation. *Chem. Phys.*, 139:89, 1989.

[28] W. Jakubetz, B. Just, and J. Manz. Mechanism of state–selective vibrational excitation by a picosecond laser pulse studied by two techniques. *J. Phys. Chem.*, 94:2294–2300, 1990.

[29] B. Just. PhD thesis, Freie Universität Berlin, 1993.

[30] B. Just, J. Manz, and G. Paramonov. Series of ultrashort infrared laser pulses with analytical shapes for selective vibrational excitations. *Chem. Phys. Letters*, 193:429–433, 1992.

[31] Ch. Lubich. Integration of stiff mechanical systems by Runge–Kutta methods. SAM-Report 92-04, ETH Zürich, 1992.

[32] J. Manz, G. Paramonov, M. Polasek, and Ch. Schütte. Overtone state–selective isomerization by series of infrared picosecond laser pulses: Model simulations for $Be_2H_3D^-$. *Israel Journal of Chemistry*, 34, 1994.

[33] A. Messiah. *Quantenmechanik*, volume 2. de Gruyter, Berlin, New York, second edition, 1985.

[34] W.L. Miranker and G. Wahba. An averaging method for the stiff highly oscillatory problem. *Math. Comp.*, 30:383–399, 1976.

[35] G. Paramonov. Blue shift and ultrafast selective excitation of high–lying molecular vibrational states. *Phys. Letters A*, 152:191–198, 1991.

[36] G. Paramonov and V. Sacca. Resonance effects in molecule vibrational excitation by picosecond laser pulses. *Phys. Letters A*, 97:340–342, 1983.

[37] R. Petrishin. Averaging with regard to resonance relations between frequencies in oscillatory systems. *Ukr. Math. J.*, 33:204–208, 1981.

[38] L. Petzold. An efficient numerical method for highly oscillatory ordinary differential equations. *SIAM J. Numer. Anal.*, 18:455–479, 1981.

[39] G.C. Pimentel, editor. *Opportunities in Chemistry*. National Accademy Press, Washington D.C., 1985.

[40] M. Polasek and R. Zahradnik. *International Journal of Quantum Chemistry*, 1993. in press.

[41] M. Quack. Theory of unimolecular reactions induced by monochromatic infrared radiation. *J. Chem. Phys.*, 69:1282–1307, 1978.

[42] M. Quack. Reaction dynamics and statistical mechanics of the preparation of highly excited states by intense infrared radiation. *Adv. Chem. Phys.*, 50:397–473, 1982.

[43] Ch. Schütte. A quasiresonant smoothing algorithm for the fast analysis of selective vibrational excitation. *Impact of Computing in Science and Engineering*, 5:176–200, 1993.

[44] T.E. Simons and A.D. Raptis. Numerov–type methods with minimal phase–lag for the numerical integration of the one–dimensional Schrödinger equation. *Computing*, 45:175–181, 1990.

[45] M. Urabe. Galerkin's procedure for nonlinear periodic systems. *Arch. Rat. Math. Anal.*, 20:120–152, 1965.

[46] M. Urabe and A. Reiter. Numerical computation of nonlinear forced oscillations by Galerkin's procedure. *J. Math. Anal. Appl.*, 14:107–140, 1966.

[47] R.L. Vander Wal, J.L. Scott, and F.F. Crim. Selectively breaking the OH–bond in HOD. *J. Chem. Phys.*, 92:803, 1990.

[48] C.E. Velez. Numerical integration of orbits in multirevolution steps. Technical Note D-5915, NASA, Goddard Space Flight Center, Greenbelt, Maryland.

[49] A.H. Zewail. Laser selective chemistry — is it possible? *Physics Today*, 33:27–33, Nov. 1980.