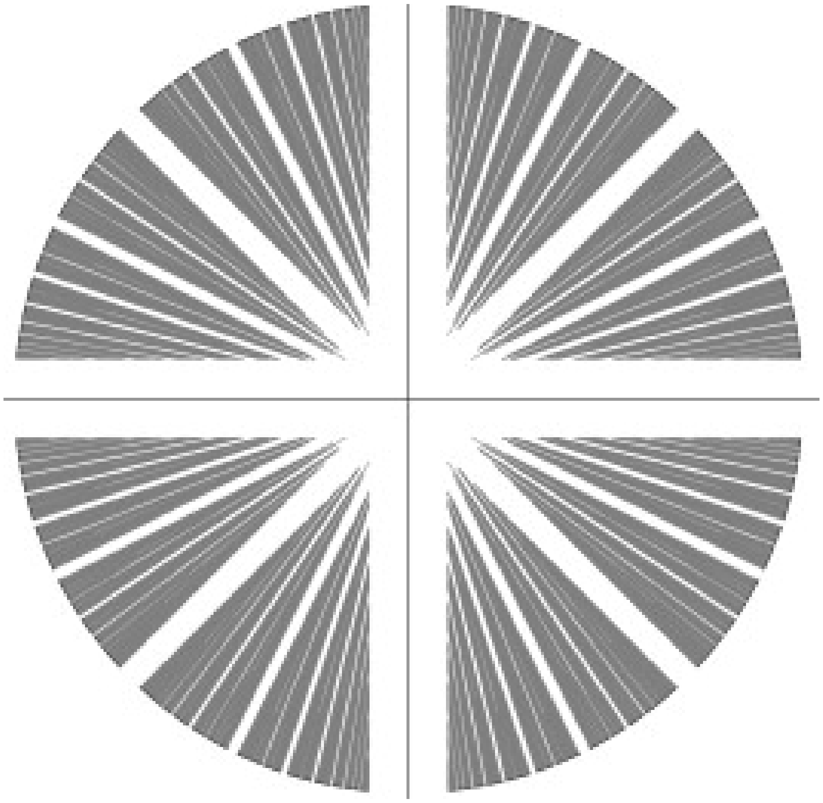
Numerical KAM theory and backward error analysis for symplectic methods applied to (quasi-)periodically perturbed Hamiltonian ODE

With applications to a tidal wave system

Francesco Carere

A thesis presented for the degree of
Master of Science



Supervisor: Prof. Jason Frank
Department of Mathematics, Universiteit Utrecht,
The Netherlands, August 29, 2022

Figure from [Han11] which illustrates Diophantine step sizes.

Abstract

Recently, a model for tidal waves in shallow areas has been reconsidered [Wal21], previously studied in [RZ92; BRZ94]. The goal was to study mixing and transport due to chaotic motion in the periodic Poincaré map. In this thesis, the study of this system is continued, shifting, however, the focus to regular instead of chaotic motion. Regular motion has been studied using Liouville (complete) integrability, action-angle coordinates and KAM theorems (we do not consider Nekhoroshev type theorems).

The Hamiltonian of the unperturbed tidal wave system is equal to $H_0(q,p) = \cos(p) + \eta\cos(q)$ for $\eta \in \mathbb{R}$ and $(q,p) \in \mathbb{R}^2$. Since this system is planar, it is Liouville integrable and a first result of this thesis is the explicit form of the action angle coordinates of the unperturbed tidal wave system in Section 3. Furthermore, from a mathematical viewpoint the tidal wave system is also interesting, as it is the natural third example in the sequence:

$$\text{Harmonic oscillator: } p^2 + \eta q^2 \qquad \text{pendulum: } \cos(p) + \eta q^2, \qquad \text{tidal wave system: } \cos(p) + \eta\cos(q)$$

and therefore presents an interesting problem to test KAM theory.

The perturbed system tidal wave system is a periodically perturbed planar Hamiltonian system with Hamiltonian $H = H_0(q,p) + \mu H_1(q,p,t)$, so H_1 is periodic in t ($1 + 1/2$ d.o.f.), where μ is small. A second result (Section 3.3) is a proof of existence of persistent invariant tori in the periodic Poincaré map (see [Wal21]) using a KAM theorem for (quasi)-periodically perturbed systems [JS96; BSG03; Sev07].

A focus in [Wal21] was on the numerical side: a splitting method was developed for "time-affine" ODE (Section 2). Also in this thesis we are much interested in the numerical side: A third result is the proof of a "numerical" KAM theory for numerically integrated periodically perturbed systems in Section 5, where the numerical integration is done using a *symplectic* integrator. This result is based on non-autonomous backward error analysis i.e. interpolation of symplectic maps by Hamiltonian flows, as considered in [Moa03; Moa05]. This numerical KAM theorem is not able to prove, for $\mu \neq 0$, the existence of invariant tori of the periodic Poincaré map of the symplectically integrated the tidal wave system. Therefore we present an 'approximate' KAM theorem, as in [HLW06] Section X.5, which proves, *up to an assumption*, the existence of 'approximate' KAM tori in the numerical, periodic Poincaré map over exponentially long times, which is the final result of this thesis discussed in Section 6.

Finally, we mention that backward error analysis (BEA) is of central importance to this thesis. Indeed in both the theoretical system as well as the numerically integrated system, the proof of the above mentioned KAM theorems was through the continuous setting (Kolmogorov/Arnolds setting i.e. flows) and not the discrete (Mosers setting i.e. symplectic maps) and to this end BEA was used to embed the discrete symplectic integrator into a Hamiltonian flow. Since BEA explains the well-behavedness of symplectic integrators applied to autonomous Hamiltonian problems a minor part of this thesis was devoted to the development of modified equation analysis (a type of BEA) for non-autonomous ODE.

Acknowledgements

I would like to thank my thesis supervisor, professor Jason Frank, for showing me this interesting subject and guiding me through the process. I would especially like to thank you for your patience: I was not the best at making the deadlines and know that this has caused some problems in your schedule at least once. I would also like to thank the second reader professor Paul Zegeling (for his patience as well if the above remark applies). Furthermore I would like to thank my family and girlfriend. My family for providing me, on several occasions, with encouragement and a place to work *à la* bed&breakfast so that I could sit productively at my desk for many hours straight, in particular my parents in the Netherlands and my aunt, uncle and cousins in Italy. My girlfriend for the encouragement and her mental support. Finally, I would like to thank (in chronological order) prof. H. Hanßmann who was so kind as to responded a question I sent by email (and for producing the picture on the front cover of this thesis), prof. A. M. Jorba who was so kind as to respond to multiple questions I sent him via email and prof. M. Somos who helped me to think about the alternative method in Section 6.5.2 via *StackExchange*.

Contents

1 Introduction

1.1 KAM theory and time-dependence in the continuous setting

The phase plane of a harmonic oscillator, described by the Hamiltonian $H_0(q,p) = \frac{1}{2}(p^2 + q^2)$ consists of invariant circles, Figure 1. More interestingly, when perturbed with a Hamiltonian H_1 i.e. considering $H_0(q,p) + H_1(q,p)$, numerical experiments seem to indicate that most of these invariant circles are not destroyed by the perturbation i.e. they persist.

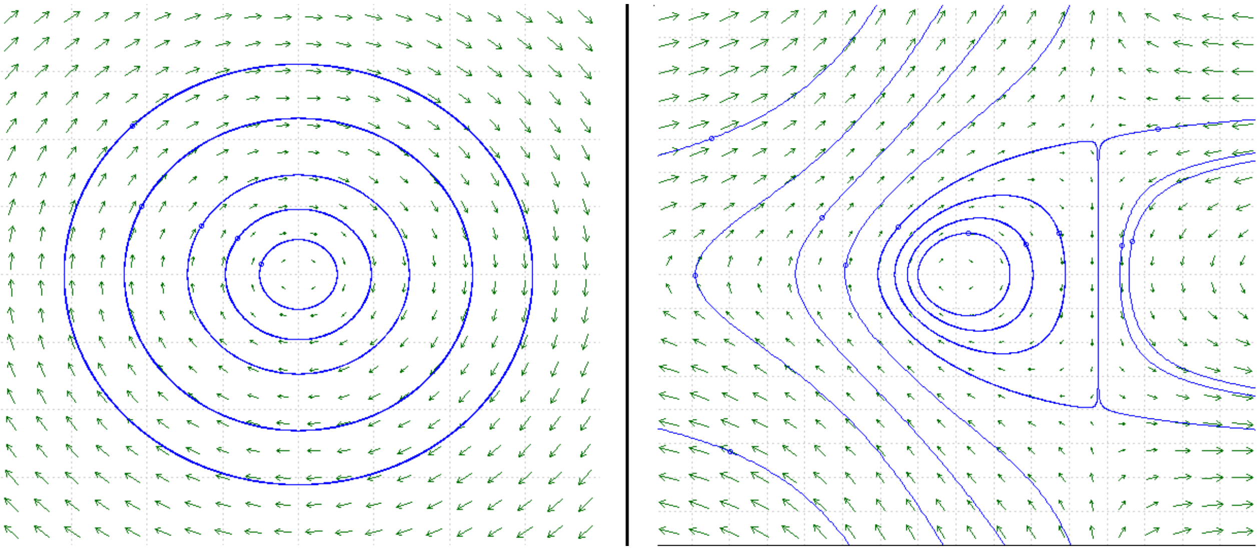


Figure 1: Typical orbits of the harmonica oscillator and the perturbed harmonic oscillator. Made in *PPlane* [Cas05].

The phase plane of a pendulum, desctribed by the Hamiltonian $H_0(q,p) = \frac{1}{2}q^2 + \cos(p)$, consists locally of invariant circles, Figure 2 (one can also see unbounded orbits), and, more interestingly, most invariant circles seem to persist when using a Hamiltonian perturbation as above.

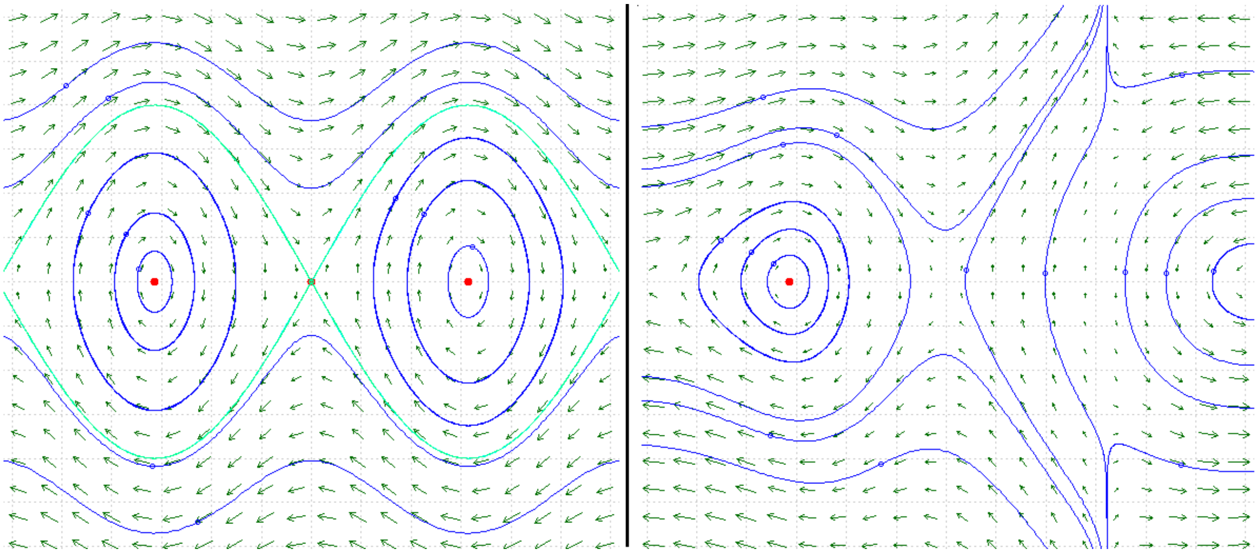


Figure 2: Typical phase plane of the pendulum ad the perturbed pendulum. Made in PPlane [Cas05].

The theory that this statement holds more generally for autonomous Hamiltonian systems in higher dimensions is one of the greater mathematical achievements of the of the 20th century in the field of perturbation theory of Hamiltonian mechanics and is attributed to Kolmogorov (in the late 50s), Arnold and Moser (in the 60s). Unsurprisingly this theory is coined *KAM theory* and the main theorem the *KAM theorem*. The statement of the KAM theorem was given by Kolmogorov[1] and two different strategies for a proof can be distinguished:

- A proof for (Hamiltonian) *flows* i.e. a continuous setting given by Arnold in 1963.
- A proof for (symplectic)[2] *maps* i.e. a discrete setting given by Moser in 1962.

and this strategy was proven equivalent by Douady [Dou82] for smooth symplectic flows defined by generating functions and maps. Douady's strategy was to use Poincaré sections and suspensions, the usual way (in dynamical systems theory) to relate continuous and discrete dynamical systems [BS02] and using this strategy he avoided entirely the consideration of any proof of KAM theory

> "Nous nous sommes efforcés de ne jamais faire appel à ces methodes et de montrer comment utiliser au mieux les énoncés proposés par la littérature." – Douady [Dou82]

The main statement of Kolmogorov is that an integrable Hamiltonian system H_0 (analytic and autonomous), which displays locally (such as the pendulum) or globally (such as the harmonic oscillator) a structure of invariant tori, has persistent invariant tori when perturbed with an autonomous Hamiltonian H_1 analytic and not too big if H_0 satisfies a regularity and strong non-periodicity/non-resonance condition (see Section 3.2). In the words of Arnold:

> "for a small perturbation, $[H = H_0(q,p) + \mu H_1(p,q)]$ $(\mu \ll 1)$, most of the tori [...] do not disappear, but are merely slightly deformed." – Arnold [Arn63]

Admittedly, this brief sketch of KAM theory and its history leaves much to be desired: First of all, there are many others who have helped, prior to and after the 1950s and '60s, to construct or extend KAM theory, for example to settings other than the Hamiltonian one. Second, exemplifying KAM theory using the harmonic oscillator and the pendulum is extremely dull (but simple). Indeed it was the three-body problem and the motion of the planets in our solar system, studied amongst others by Poincaré, which gave a big impulse to KAM theory and in this light KAM theory deserves the following, more impressive question:

> "The KAM theorem yields a troubling answer to one of the oldest questions in celestial mechanics: is the solar system stable? Will it continue eternally more or less as we see it today? Or could it be that planetary interactions, between Jupiter and Saturn e.g. will eventually lead to catastrophes, where certain planets escape from the Sun, and others collide or fall into the Sun?" – Hubbard [Hub07]

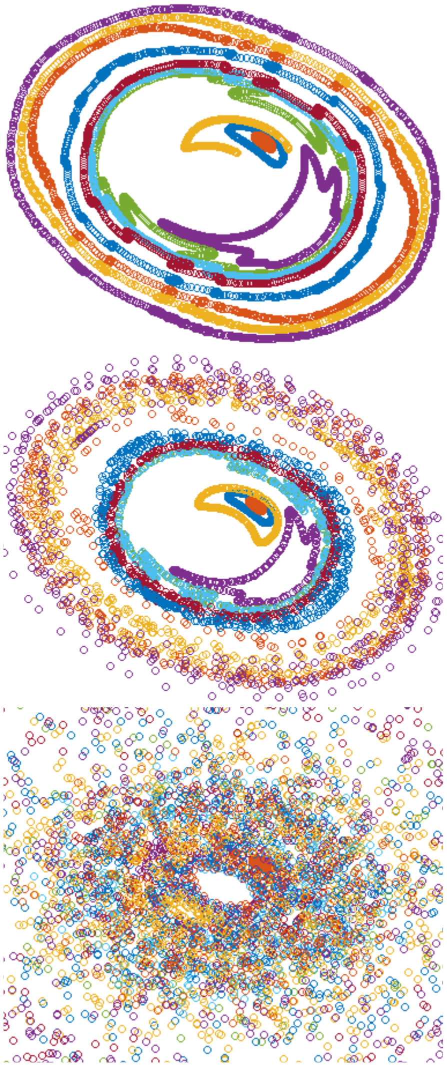


Figure 3: Periodic Poincaré map of the perturbed harmonic oscillator $H_0 + \mu H_1$ for increasing (downwards) perturbation strength.

[1] See Section 3.2 for more details and references on the question if Kolmogorov himself provided a proof.

[2] One can say that symplectic maps are the discrete version of Hamiltonian flows (Section 2.2.2).

This motivation has a long history [Mos87]. Moreover there are other fundamental motivations from physics [Ber78]. However, we continue with the simple examples and leave the more impressive history and statements of KAM theory and (Hamiltonian) perturbation theory to other sources (some references are given in Section 3.2) and emphasize that, up to this point, we have only considered autonomous Hamiltonians H_0 and autonomous Hamiltonian perturbations H_1.

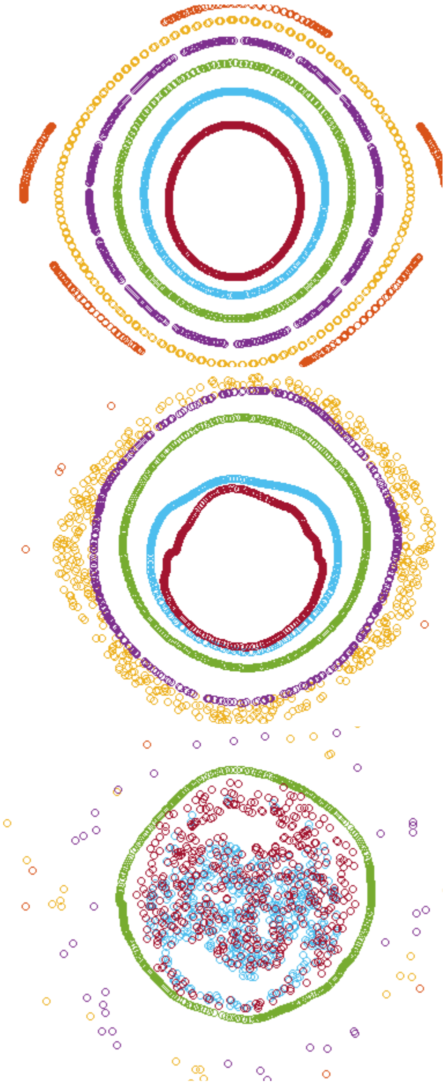


Figure 4: Periodic Poincaré map of the perturbed pendulum $H_0+\mu H_1$ for increasing (downwards) perturbation strength μ.

What if the unperturbed Hamiltonian H_0 is time-dependent?
In general, the natural strategy is to extend phase space with an extra degree of freedom, incorporating the time variable into the (symplectic) geometry of the system and making the system Hamiltonian and autonomous (Section 2.1). Subsequently, one would like to use the above, autonomous KAM theory to prove the persistence of invariant tori by providing the above mentioned conditions, but this is not always easy:
In the case of general, aperiodic[3] time-dependence the situation seems grim: A special set of coordinates, action-angle coordinates (Section 3.1), are crucial for most proofs of the KAM theorem, for which compact orbits are a condition. Time, being incorporated into the geometry (and topology), however, causes all orbits to become unbounded. The construction of action angle coordinates is still possible in some sense (although much harder) [GMS02; Fio04] and possibly one can work from there to get to a KAM theorem. Alternatively, one can avoid working with action angle coordinates all together, using more undeveloped versions of the KAM theorem [Lla+05]. In the case of (quasi)-periodic time-dependence time already lives on the circle, as an angle coordinate, thereby avoiding the grim situation in the construction of action angle coordinates in the aperiodic case so it seems that the complexities of a KAM application on the extended phase space are removed.

In view of an appliation to the tidal wave system, we are only interested in the following case: *"What if (just) the perturbation H_1 is time-dependent?*.
It seems that this question is similar to the previous one: One is again inclined to use the extended phase space and apply the usual KAM theory, as in for example [Moa03]. However, this case (that H_0 is autonomous and H_1 time-dependent) is more interesting (sometimes we may use the term "non-autonomous KAM theory" for this case), since the condition of non-degeneracy adapts accordingly:
Again, considering first the aperiodic case, there have been some movements towards "aperiodic (non-autonomous) KAM theory" [MW00; WM14; CL15; FW16]. And these interests seem to have mainly arisen from applications in mixing and transport in fluid dynamics. For example, geophysical flows, such as meandering jets, are usually aperiodic [WM14].
However, the main focus, where luckily our interest lies, has been on a KAM theorem for (quasi-periodic) perturbations:

> "Essentially all of the literature [...] concerned with time-dependent Hamiltonian systems deal

[3] "Aperiodically" indicating that the time dependence is *not* (quasi-)periodic (Defition 3.1, see also [Sev07]).

> with periodic or quasi-periodic time dependence [in the perturbation]. For such time dependence, the problems can often be cast in a form where classical results and approaches can be applied"
> – Fortunati & Wiggins [FW16]

However, the goal of this thesis is to prove the existence of persistent invariant tori (KAM tori) in the periodic Poincaré map/2π- Poincaré map (see [Wal21] where it is called the "Tidal Poincaré map" or [Wig03] chapter 10.2) of the tidal wave system, we do not "cast in a form" but use KAM theorems which are adapted to quasi-periodic perturbation as in [Jor91; JS96; BSG03; TZ09], and generalised in [JV97] (for lower-dimensional tori) and [Sev07] (for non-Hamiltonian and/or parametrised systems, for *families* of tori and under the weaker Rüsmann non-degeneracy conditions).

We check again with the elementary examples: the Harmonic oscillator $H_0(q,p) = \frac{1}{2}(p^2 + \eta q^2)$ and the pendulum $H_0(q,p) = \frac{1}{2}p^2 + \eta\cos(q)$. We perturb these systems with the non-autonomous Hamiltonian $\mu H_1(q,p,t) = \mu\sin(t)(\cos(q^2) + p^2 + \cos(p^4))$ which is 2π-periodic in t. The numerical simulations, plotting the numerically approximated 2π-Poincaré map are shown in figure 3 and 4 and we see indeed existence of persistent invariant tori in the numerically approximate 2π-Poincaré map.

1.2 Goal of the paper: theoretical part

In this thesis we are interested, as the title might suggest, in three of the topics that we have encountered so far: KAM theory for quasi-periodic non-autonomous perturbations; mixing and transport in fluid mechanics; a tidal wave system together with the elementary examples (the harmonic oscilator and the pendulum) and numerical simulations. The first three topics are connected immediately below and numerical simlutations, the fouth, is treated afterwards.

The first two topics are naturally entangled: In fluid mechanics, one sometimes encounters time-dependent Hamiltonians H_0 or perturbations H_1 [WM14] and one would like to study obstructions to mixing and transport using KAM theory, such as invariant tori. Indeed, mixing and transport in these types of systems can be due to regular and chaotic Hamiltonian dynamics [WM14; Mei15] and one would like to use Hamiltonian perturbation theory, such as non-autonomous KAM theorems, to qualitatively or quantitatively explain mixing and transport [MMP84; OO89; Mei92] (studied also in the numerical case [FS96]). As mentioned above, in geophysical flow applications H_1 is often aperiodic, but sometimes it is periodic, and we are lucky (see the previous quote).

In this thesis, we consider a recently studied model for mixing and transport in tidal waves in shallow areas [Wal21], of which a simplified model had been previously studied in [RZ92; BRZ94], which we will call the *tidal wave system.* The elementary examples and the tidal wave system can be related in the following sense: The unperturbed Hamiltonian of the tidal wave system (Section 1.5) fits right next to the elementary examples ($\eta \in \mathbb{R}$):

$$\text{Harmonic oscillator:}\quad H_0(q,p) = \frac{1}{2}(q^2{+}\eta p^2),\quad \text{Pendulum:}\quad H_0(q,p) = \frac{1}{2}q^2{+}\eta\cos(p),\quad \text{Tidal wave system:}\quad H_0(q,p) = \cos q{+}\eta\cos p$$

and the perturbations considered are periodic in time. The previous papers [RZ92; BRZ94; Wal21] focused on (mixing and transport due to) the *chaotic part* of the motion and in particular chaotic motions in the periodic Poincaré map (previously called the Tidal Poincaré map [Wal21]), the period of this map naturally equals the period of the perturbation (see e.g. [Wig03] section 10.2). Instead, in this thesis we focus on the *regular part* of the motion, in particular by applying non-autonomous KAM theory for (quasi-)periodic perturbations to prove the existence of KAM tori in the periodic Poincaré map of the tidal wave system.

To summarize, the first three topics shed two different lights on the tidal wave system and applictions of KAM theory to it: On the one hand, from a mathematical perspective, the system is relatively hard since it lies beyond the elementary examples, which are well-studied from the KAM viewpoint. On the other hand, from a physical perspective, the system is relatively easy, a 'toy problem'. Indeed it has a periodic perturbation and not aperiodic as was often the case for geophysical flows and, moreover, one is able to obtain action angle coordinates, which is not usual for fluid dynamics:

> "The KAM and Nekhoroshev theorems are stated [...] using the action-angle variables of the unperturbed integrable system. Even if one has a model that can be divided into an integrable part plus "a perturbation", it is [...] highly nontrivial to construct action-angle coordinates for the unperturbed, integrable part. For this reason there have been essentially no applications of the KAM theorem to fluid transport where the conditions for the applicability of the theorem have been verified for a model under consideration. Similarly for the Nekhoroshev theorem [...]."
> – Wiggins & Mancho [WM14]

Beside action angle coordinates, one could possibly even obtain estimates for the small parameters (e.g. μ in Arnold's quote above) in the KAM theorem, something that was done for the pendulum in the series of papers [CC87; CFP87a; CFP87b; CC88; CG88]. We will not attempt such a rigorous undertaking.

The main goal of the theoretical (non-numerical) part of this thesis, Sections 2 - 3, is to put the tidal wave system into action angle coordinates, to state a non-autonomous KAM theorem for periodic perturbations (Section 3) and to prove the existence of persistent tori *of the tidal Poincaré map* in the theoretical system (Section 3.3).

To conclude, we mention that there are *two* natural strategies to prove the existence of KAM tori in the periodic Poincaré map (as discussed in Section 3.3). The first is, as mentioned, by the use of a non-autonomous KAM theorem, which can be applied to the flow of the tidal wave system i.e. to a continuous dynamical system (defined in Appendix A). A second way is to note that the periodic Poincaré map, as a symplectic map, induces a discrete dynamical syste so that discrete KAM theorems, as the one of Moser, can be applied. We do not consider this strategy, but comment on it in the further research.

1.3 KAM theory and time-dependence in the numerical/discretised setting

We now discuss the fourth topic, numerical simulations, from the KAM perspective. In [Wal21] Figure 5 was produced, which shows two different numerical discretisations of the tidal Poincaré map. It can be seen that KAM tori persist or are destroyed depending on the numerical method used and this is a general fact. There exist methods adapted to ODE with "structure": For the Hamiltonian, with "symplectic structure" (Section 2), these are called 'symplectic methods"; For more general "structures" (see [MQ01; IQ18]), these are called "Structure preserving methods" , in particular "Geometric numerical integration" (GNI) if the structure is 'geometrical'.

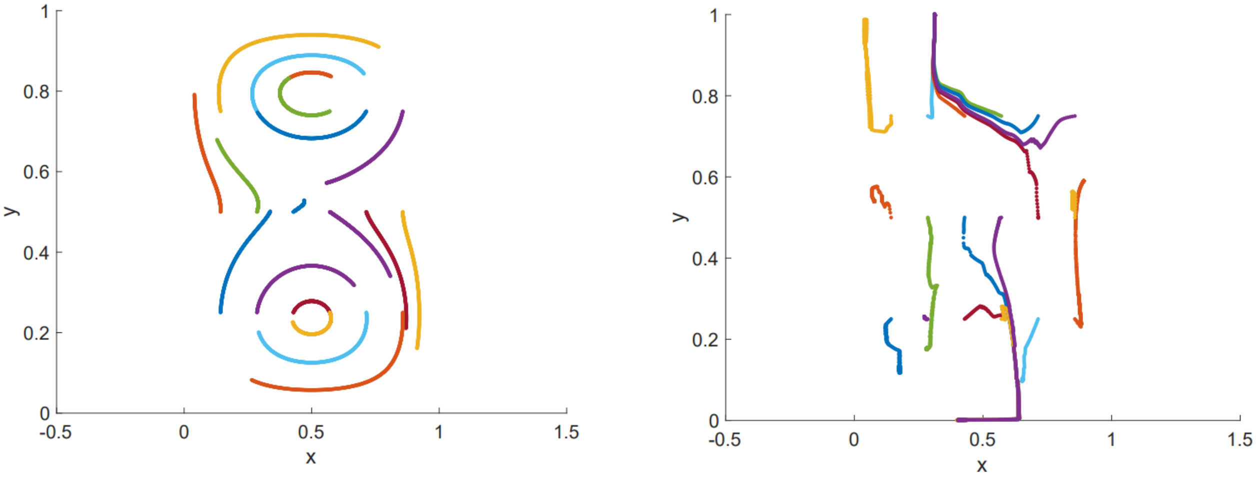


Figure 5: Not all numerical methods are good-natured towards the KAM tori. A symplectic splitting method is, while the *Matlab ODe 45* routine is not (Source: [Wal21]).

Let us first consider again the case of autonomous Hamiltonian ODE. In this case the symplectic integrators behave well: Energy is approximately conserved over exponentially long times [BG94], there is linear error growth in time when applied to integrable Hamiltonian systems over polynomially small times [HLW06]

Section X.3, and, most importantly, KAM tori persist [Sha99; HLW06]. This well-behavedness over exponentially long times has been seen in the numerical integration of the solar system and gives an answer to the above question of Hubbard about the stability of the solar system:

> "The evolution of the entire planetary system has been numerically integrated for a time span of nearly 100 million years. This calculation confirms that the evolution of the solar system as a whole is chaotic, with a time scale of exponential divergence of about 4 million years. Additional numerical experiments indicate that the Jovian planet subsystem is chaotic, although some small variations in the model can yield quasi-periodic motion. The motion of Pluto is independently and robustly chaotic." – Susmann & Wisdom [SW92]

The latter property of symplectic integrators, that KAM tori exist in the numerically integrated system, suggest that there is a "numerical KAM theory" for symplectically discretised/integrated systems. Indeed in Figures 3 and 4 symplectic integrator were indeed used so that KAM tori were seen in these Figures.

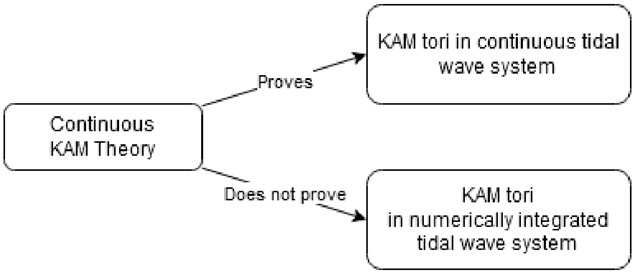

Continuous
KAM Theory
Proves
KAM tori in continuous tidal wave system
Does not prove
KAM tori
in numerically integrated
tidal wave system

Figure 6: The problem with using (continuous) KAM theory on the numerically integrated system.

We saw above that 'discrete' KAM theory, for symplectic twist maps, has already been developed: it was proven by Moser. However, similar to the theoretical setting, there is also a continuous approach to the proof and statement of discrete KAM theory (Figure 7):

1. Use a discrete version, done by Shang [Sha99] around 1999, who generalised Moser's 1962 results on twist maps on the annulus to higher dimensions. According to [HLW06], section X.6, Shang used an Arnold type construction. The authors of [HLW06] give a complementary proof based instead on a Kolmogorov type construction.

2. Use a continuous version, made possible by embedding the symplectic integrator into a Hamiltonian flow and applying continuous KAM theory. This strategy seems discussed mainly by [Moa03] and [MO10] (the latter for lower dimensional tori), inspired by the above mentioned suspension of Douady. Similarly to Douady they remark again

    > "Many KAM style results for lower dimensional invariant tori already exist. Using interpolation one can avoid redoing lengthy proofs for maps and so it is this approach we take in this paper." – McLachlan & O'Neale [MO10]

A symplectic integrator can naturally be seen as a symplectic map ψ_h with step-size $h > 0$. Thus, it seems natural to apply a discrete KAM theorem to symplectically integrated systems to prove the existence of invariant tori in the system. We are interested, moreover, in KAM tori of the 2π-Poincaré map which is also a symplectic map, on which it again seems natural to use a discrete KAM theorem. Therefore, both in the theoretical tidal wave system and in the symplectically integrated tidal wave system it seems natural to use discrete KAM theorems. However, in this thesis we will use the numerical route and the main reason for this is that backward error analysis (BEA) is a central topic in this thesis (see also Section 10).

In short, BEA (Appendix C for an introduction) constructs for a symplectic map/integrator an autonomous or non-autonomous *modified* vector field so that this symplectic map/integrator is (almost) interpolated by the flow of this vector field (Sections 5 and 6). The autonomous case is usually called *modified equation analysis* (MEA), e.g. [GS86; CMS94] and we call the other *non-autonomous flow interpolation.* The reason that we are interested in BEA is as follows: For autonomous Hamiltonian systems, the structure preserving methods are symplectic integrators, which can be proven rigorously using BEA. For example, BEA shows that the modified vector field which interpolates the symplectic map is again Hamiltonian. We therefore develop BEA for non-autonomous (Hamiltonian) ODE (Section 6.3.6 and Appendix D) in the hope to contribute to the problem of identifying what the 'structure preserving methods' of non-autonomous Hamiltonian ODE are.

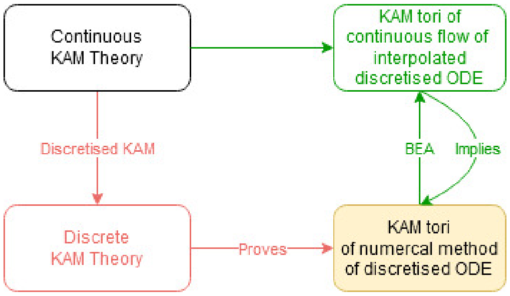


Figure 7: The two approaches (red and green) to prove KAM tori in the numerically integrated tidal wave system.

Before asking what structure preserving methods are in the non-autonomous Hamiltonian case, it is useful to first identify what this 'structure' is i.e. ask the question "*What is the 'structure' of non-autonomous Hamiltonian ODE?*". There has been some research into this structure and this is considered briefly in Section 2.4. However, coming back to the question of structure-preserving integrators, reasearch into non-autonomous Hamiltonian struture-preserving integrators is extremely scarce. For example we quote

> "Although research in geometric numerical integrators for differential equations has experienced a tremendous boost during the last decades, it is fair to say that this has been mainly restricted to autonomous problems" – Blanes, Casas & Murua [BCM12]

and more recently, for the Hamiltonian case

> "Symplectic integration of autonomous Hamiltonian systems is a well-known field of study in geometric numerical integration, but for non-autonomous systems the situation is less clear" – Marthinsen & Owren [MO14]

On top of that, an even more recent paper discussing De Vogelaere [De 56] (the first paper on symplectic integrators) shows that this research started on a bad footing and never got a hold:

> "The title of the preprint [of De Vogelaere] is a little bit misleading because De Vogelaere considers only transformations in the phase space (q, p) treating time as a parameter. Now, such transformations are rather called symplectic (or canonical). Contact transformations, usually related to the extended phase space (including the time variable), are more general.
> The problem of extending symplectic integrators on non-autonomous (time-dependent) Hamiltonian systems is much more difficult and still only particular results are available. It seems that in this aspect the way of introducing time dependence of integrators presented in the famous preprint is not so fruitful as the symplecticity of the integrators in the autonomous case" – Skeel & Cieslinski [SC20]

1.4 Goal of the paper: numerical part

In this thesis, we will not attempt to start such a theory for non-autonomous, structure preserving integrators, although we will discuss some structure of non-autonomous (Hamiltonian) ODE and flows in Section 2. As mentioned above, we will try to develop and use BEA, the *tools* in proofs concerning structure-preserving methods, as much as possible to the non-autonomous setting, so that it may possibly help in future work on the well-behavedness (including the persistence of KAM tori via application of the (periodic) non-autonomous KAM theory) of 'structure preserving methods' in the non-autonomous Hamiltonian case, whatever the structure may be. Thus, a first goal is to discuss BEA (for non-autonomous ODE).

Furthermore, we discuss a numerial method developed in [Wal21], which is due to the special form of the tidal wave system. This special form is that the considered Hamiltonian H and its perturbations are of the form $(n \in \mathbb{N})$:

$$H(q, p, t) = \sum_{i=1}^{n} g_i(t) H_i(q, p), \tag{1.1}$$

where g_i, H_i are scalar functions ($g_i = 1$ may be possible). These types of Hamiltonians are called *time-affine*. In the case $n = 1$, the Hamiltonian is of the form

$$g(t)K(q, p),$$

which we will call a *forced* Hamiltonian/vector field.

For forced ODE $\dot{y}(t) = g(t)f(y)$ there exist a natural numerical method, discussed in Section 4, which is called the *induced method.* Theoretically, in this case we will see that the system is (locally) equivalent, up to time transformation, to the autonomous system with ODE f. In turn, this will imply that, given a numerical method ψ for the ODE with vector field f, one may construct (Figure 8) a numerical method for the ODE with vector field gf by using a 'time-adaptive' step-size, adaptive only to time, not to space. Thus, for a forced ODE which is Hamiltonian, a natural choice is a symplectic method on the autonomous part, with time-adaptive step-size.

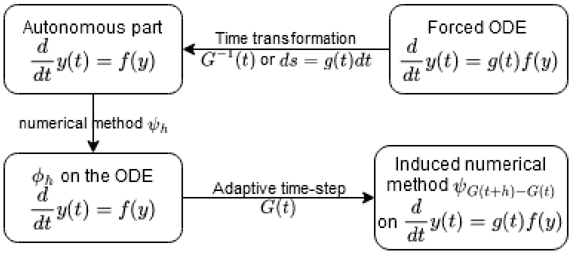

Autonomous part
$\frac{d}{dt}y(t) = f(y)$
Time transformation
$G^{-1}(t)$ or $ds = g(t)dt$
Forced ODE
$\frac{d}{dt}y(t) = g(t)f(y)$
numerical method ψ_h
ϕ_h on the ODE
$\frac{d}{dt}y(t) = f(y)$
Adaptive time-step
$G(t)$
Induced numerical method $\psi_{G(t+h)-G(t)}$
on $\frac{d}{dt}y(t) = g(t)f(y)$

Figure 8: Forced ODE are equvialent to their autonomous part. The induced method is constructed from this perspective

The tidal wave system is time-affine (Equation (1.1)). The numerical method used in [Wal21] is a splitting method (Section 4): It splits the time-affine Hamiltonian $H(q, p, t) = \sum_{i=1}^{n} g_i(t)H_i(q, p)$ into n terms $g_i(t)H_i(q, p)$ and uses the induced method to solve solve the n forced system. We argue in Section 4 that, from the perspective of structure preservation, this splitting method for time-affine Hamiltonian ODE has no special properties, except that it is symplectic if the induced methods are symplectic.

The main goal of the numerical part of this thesis (Sections 4- 6) is to show existence of KAM tori for the numerically approximated 2π- Poincaré map of the periodically perturbed tidal wave system, which is integrated using the splitting method for time-affine systems as developed in [Wal21]. In Section 5 a KAM theorem is developed for numerically integrated, periodically perturbed complete integrable systems. This theorem, however, has many assumptions and, moreover, does not prove the existence of KAM tori in the numerically approximate 2π-map of the perturbed system. Therefore, (and also due to considerations of round-off error) we will also state an 'approximate' KAM theorem in Section 6, which will prove, *up to an assumption*, the existence of 'almost invariant' tori in the numerically approximated 2π-Poincaré map of the periodically perturbed system.

Before discussing the main goals of the theoretical and numerical part of this thesis, we first discuss (in Section 2) the relation between non-autonomous (Hamiltonian) ODE and autonomous ODE is discussed. We will show that forced ODE are locally equivalent, up to a time transformation, to autonomous ODE (Figure 8) and describe the (symplectic) structure of (non-)autonomous Hamiltonian ODE.

Thus, the structure of the paper is as follows:

- In Section 2 we will show how non-autonomous and autonomous ODE are related and we show thatr forced ODE are locally equivalent the an autonomous part of the system. Afterwards we describe the symplectic structure of autonomous (and non-autonomos) Hamiltonian ODE . Finally, we given discuss what possible structures may be for non-autonomous Hamiltonian ODE.

- In Section 3 we develop action angle coordinates for completely integrable systems, the framework for KAM theory, and state a KAM theorem for periodically perturbed completely integrable Hamiltonian system. Afterwards we apply this KAM theorem to the theoretical tidal wave system, proving our first goal, the persistence of invariant tori in the 2π-Poincaré map.

- In Section 4 we discuss numerical methods, in particular symplectic methods, the induced method for forced (Hamiltonian) ODE and splitting methods for time-affine (Hamiltonian) ODE.

- In Section 5 we discuss some strategies for the proof of KAM tori in the 2π-Poincaré map of the numerically integrated tidal wave system. Furthermore, we develop a KAM result for symplectically integrated, periodically perturbed Hamiltonian ODE. However, this KAM theorem is not able to prove the existence of KAM tori in the numerically approximated 2π-Poincaré map of the periodically perturbed Hamiltonian system.

- In Section 6, we therefore develop an 'approximate' KAM theorem, which proves, *up to an assumption*, the existence of 'almost invariant' tori in the numerically approximated 2π-Poincaré map of the perturbed tidal wave system, when the discussed splitting method for time-affine Hamiltonian ODE is used.

Next, however, we will finish the introduction by giving a description of the tidal wave system.

1.5 The tidal wave system

We end the introduction by giving a mathematical formulation of the tidal wave system, as considered in [Wal21] (the formulation is similar to the one in [BRZ94], as noted in the appendix of [Wal21]). Throughout the thesis we use notation and definitions of Appendix A.

The model is based on the shallow-water equations together with some assumptions and simplifications[4]. In this thesis we will consider the Hamiltonian part of the tidal wave system, without dissipative terms. This is modelled by the Hamiltonian ODE

$$\begin{aligned}\dot{q} = \cos(t) + \ & C_\delta l\Big(k[rl\sin(kq)\cos(lp) - fk\cos(kq)\sin(lp)] \\ & + 2\gamma(r\cos t + \sin t)[fk\sin(kq)\sin(lp) + rl\cos(kq)\cos(lp)]\Big)\end{aligned} \tag{1.2}$$

$$\begin{aligned}\dot{p} = \ & C_\delta k\Big(k[-rl\cos(kq)\sin(lp) + fk\sin(kq)\cos(lp)] \\ & + 2\gamma(r\cos t + \sin t)[fk\cos(kq)\cos(lp) + rl\sin(kq)\sin(lp)]\Big),\end{aligned} \tag{1.3}$$

where $C_\delta = \delta\left(k^2 + l^2\right)^{-1}\left(k^2 + 2r^2 + 2\right)^{-1}$, $\delta, k, l, f, r \in \mathbb{R}$ and $\gamma \in \{0, 1\}$ (in the notation of [Wal21] $\delta = [h]/H$), with Hamiltonian

$$H(q, p, t) = p\cos t + H_1(q, p) + \gamma H_2(p, q, t), \tag{1.4}$$

where

$$\begin{aligned}H_1(q, p) &= C_\delta k[rl\sin(kq)\sin(lp) + fk\cos(kq)\cos(lp)] \\ H_2(p, q, t) &= 2C_\delta\left(r\cos(t) + \sin(t)\right)[rl\cos(kq)\sin(lp) - fk\sin(kq)\cos(lp)].\end{aligned}$$

In the thesis [Wal21], the linear transformation

$$(q, p) \mapsto A \cdot (q, p)^T = \begin{pmatrix} k & l \\ k & -l \end{pmatrix}(q, p)^T,$$

is considered, which is a $-2kl$ scaling of a symplectic matrix (Section 2.2.4). Thus, the transformed ODE is induced by the Hamiltonian $L(q, p, t) = \det(A)H(A^{-1}(q, p), t)$ (see Proposition 2.17 and Definition 2.16) which is of the form

$$L(q, p, t) = k(p - q)\cos(t) + L_1(q, p) + \gamma L_2(q, p, t) \tag{1.5}$$

where (correcting a factor 2 error in [Wal21])

$$L_1(q, p) = H_1(A^{-1}(q, p)) = -C_\delta k^2 l\big(\alpha\cos(p) + \beta\cos(q)\big) \tag{1.6}$$

$$L_2(q, p, t) = H_2(A^{-1}(q, p), t) = C_\delta 2kl(r\cos(t) + \sin(t))(\alpha\sin(p) + \beta\sin(q)), \tag{1.7}$$

and $\alpha := fk + rl$, $\beta := fk - rl$. Equivalently, these equations are found because $\frac{d}{dt}\big(A(p(t), q(t))^T\big) = A(\dot{p}, \dot{q})^T(t)$. The transformed system satisfies the ODE

$$\dot{q}(t) = k\cos(t) + C_\delta kl\big(k\alpha\sin(p) + 2\gamma\alpha\cos(p)(r\cos(t) + \sin(t))\big) \tag{1.8}$$

$$\dot{p}(t) = k\cos(t) - C_\delta kl\big(k\beta\sin(q) + 2\gamma\beta\cos(q)(r\cos(t) + \sin(t))\big). \tag{1.9}$$

[4] Assumptions and simplifications comprise of: Rigid-lid approach, constant water density, tidal current depending on time only, bottom height-variations being much smaller than average depth, Fourier series approximation of bottom-topography [Wal21]

The advantage of the system with Hamiltonian L is that it is easier to work with numerically, in particular the splitting method developed in [Wal21] is explicit when applied to the ODE with Hamiltonian L.

The parameter sets considered in [Wal21] are g

	r	f	γ	δ	α	β
Default	2	0	0	0.3	2l	-2l
Simple-B	2	0	0	$\frac{6(k^2+2r^2+2)}{10(kr)}$	2l	-2l
Coriolis	2	1	0	0.3	k + 2l	k - 2l
Vorticity-harmonic	2	0	1	0.3	2l	-2l
Cor-VortHarm	2	1	1	0.3	k + 2l	k - 2l

In [Wal21] $l = k$ and $\delta = 0.3$ or, in the 'Simple-B' case, $\delta = 0.6\frac{k^2+2r^2+2}{rk}$. Therefore the Hamiltonians reduce to

$$
\begin{aligned}
H(q,p,t) &= p\cos(t) + \frac{3}{20k(k^2+2r^2+2)}\bigg(k[r\sin(kq)\sin(kp) + f\cos(kq)\cos(kp)] \\
&\quad + 2\gamma(r\cos t + \sin t)[-f\sin(kq)\cos(lp) + r\cos(kq)\sin(lp)]\bigg) \\
L(q,p,t) &= k(p-q)\cos(t) + \frac{3k}{20(k^2+2r^2+2)}\bigg(-k[\cos(p)(f+r) + \cos(q)(f-r)] \\
&\quad + 2\gamma(r\cos(t) + \sin(t))[\sin(p)(f+r) + \sin(q)(f-r)]\bigg),
\end{aligned}
$$

except for the"simple-B" case, where

$$
\begin{aligned}
H(q,p,t) &= p\cos(t) + \frac{3k}{5r}\bigg(\frac{1}{2}[r\sin(kq)\sin(kp) + f\cos(kq)\cos(kp)] \\
&\quad + 2\gamma(r\cos t + \sin t)[-f\sin(kq)\cos(lp) + r\cos(kq)\sin(lp)]\bigg) \\
L(q,p,t) &= k(p-q)\cos(t) + \frac{3}{5r}\bigg(-k[\cos(p)(f+r) + \cos(q)(f-r)] \\
&\quad + 2\gamma(r\cos(t) + \sin(t))[\sin(p)(f+r) + \sin(q)(f-r)]\bigg).
\end{aligned}
$$

Remark 1.1. *This means that, varying the values of $k \neq 0$, one finds two different scaling factors $\nu, \tilde{\nu} : (0,\infty) \to \mathbb{R}$ defined by*

$$
\nu(k) = \frac{3k}{10(k^2+2r^2+2)} \qquad \tilde{\nu}(k) = \frac{3}{5r}
$$

This relates the parameter sets in [Wal21] with different δ ("default" (to ν) and "simple-B" (to $\tilde{\nu}$)).

2 Non-autonomous ODE and non-autonomous Hamiltonian ODE

The first goal of this section is to introduce forced, time-affine and Hamiltonian ODE. In particular, the perturbed tidal wave system is an example of a time-affine, Hamiltonian ODE. Afterwards we treat the symplectic structure of autonomous Hamiltonian ODE, symplectic maps and canonical maps, which is useful for the theory of structure-preserving integrators 4.2. Finally, using we give an outlook on the structure of non-autonomous Hamiltonian ODE.

Definition 2.1 (Forced ODE/vector field)**.** A non-autonomous ODE on $D \times I$ (with henceforth $D \times I \subset \mathbb{R}^n \times \mathbb{R}$ open, $(g, f) : D \to I \times \mathbb{R}^n$) (so g is a scalar function) of the form

$$
\dot{y}(t) = \hat{f}(y(t), t) = g(t)f(y(t))
$$

is called a *forced ODE* and $\hat{f}$ a *forced vector field* (*periodically forced* if g is periodic). $\varnothing$

Definition 2.2 (Time-affine ODE/vector field)**.** A vector field $\hat{f}$ on $\mathbb{R}^n$ which is a sum of forced vector fields

$$\hat{f}(y,t) = \sum_i g_i(t) f_i(y)$$

is called a *time-affine*[5] vector field. An ODE with time-affine vector field $\hat{f}$ is called a *time-affine* ODE. $\varnothing$

Definition 2.3 ((Separable) Hamiltonian ODE)**.** Given a scalar function $H \in C^1(D \times I)$. The Hamiltonian ODE (with Hamiltonian H) is defined as the ODE

$$\dot{q}(t) = D_p H(q(t), p(t), t) \quad \dot{p}(t) = -D_q H(q(t), p(t), t),$$

where, abusing notation, q, p represent points in (phase) space $D \times I$ as well as paths (i.e. $q \in C^1(\tilde{I}, \mathbb{R}^n)$, $\tilde{I} \subset I$ open, see also Appendix A). If the Hamiltonian can be written as $H(q,p,t) = V(q,t) + T(p,t)$ for two scalar function $V, T \in C^1(D \times I)$ then the Hamiltonian ODE (and the induced vector field) is called separable. $\varnothing$

Denoting $z = (q,p)$ a Hamiltonian ODE can be rewritten as

$$\dot{z}(t) = J^{-1}\nabla H(z(t),t) = \begin{pmatrix} 0 & Id_n \\ -Id_n & 0 \end{pmatrix} \begin{pmatrix} D_q H \\ D_p H \end{pmatrix} (z(t),t), \quad \text{i.e.} \quad J = \begin{pmatrix} 0 & -Id_n \\ Id_n & 0 \end{pmatrix},$$

where Id_n is the n-dimensional identity matrix and $\nabla H = (dH)^T$ denotes the gradient. Hamiltonian ODE satisfy the property that, along a solution $q(t), p(t)$ one finds

$$\frac{d}{dt} H(q(t), p(q), t) = \frac{\partial}{\partial \tau} H(q(t), p(t), \tau)|_{\tau = t} \tag{2.1}$$

so that autonomous Hamiltonians are preserved by the flow ($\frac{d}{dt} H(q(t), p(t)) = 0$). In particular, defining the Poisson bracket $\{\cdot,\cdot\}$ of two scalar functions F, G as

$$\{F, G\} = \nabla F^T J^{-1} \nabla G \tag{2.2}$$

one finds that [GPS02]

$$\frac{d}{dt} F(q(t), p(t), t) = \{F, H\} + \frac{\partial F}{\partial t}. \tag{2.3}$$

Finally, the flow of a forced Hamiltonian $g(t)H(q,p)$ preserves the function H since

$$\frac{d}{dt} H(q(t), p(t)) = g(t)\{H, H\} = 0. \tag{2.4}$$

2.1 Forced ODE, canonical autonomous extension and extended phase space

We first show how forced ODE are equivalent to their autonomous part. Afterwards we introduce a simple way to make non-autonomous ODE autonomous by *autonomous extension* of the ODE. Finally, we discuss forced Hamiltonian ODE and a special kind of autonomous extension for Hamiltonian systems.

2.1.1 Forced ODE: Equivalence to its autonomous part

Given a non-autonomous vector field $f : D \times I \to \mathbb{R}^n$, and suppose y is an integral curve of the ODE with vector field f. Then, given a function $\tau : \mathbb{R} \to \mathbb{R}$, we may consider $y \circ \tau$, which is an integral curve of the vector field $\tilde{f}(y,t) = \dot{\tau}(t) f(y(\tau(t)), \tau(t))$.

[5] The name time-affine is taken from control theory, where the functions g_i are interpreted as a control parameter and such systems are called control-affine [Jur96]. As is usual for control-affine system, one must be warned that control-affine systems are not affine in the control parameters. Similarly time-affine systems are not affine in time.

Definition 2.4. If $f : D \times I \to \mathbb{R}^n$ is a vector field and $\tau : \mathbb{R} \supset I \to \mathbb{R}$ is differentiable. Then the *time-reparametrised* (ODE with) vector field f by the the *time-raparmetrisaion* τ is the (ODE with) vector field $\tau^*(f)(y,t) := \dot{\tau}(t) f(y, \tau(t))$. $\varnothing$

Given a forced vector field $g(t)f(y)$. Suppose $G(t) = \int_0^t g(t)$ is well-defined and y is an integral curve of the ODE with vector field f i.e. $\dot{y} = f(y)$ then

$$\frac{d}{dt}(y \circ G) = g(t) f(y).$$

Thus, using the time-reparametrisation G one can, from the integral curves of the ODE with vector field f, find integral curves of the forced ODE with vector field $\dot{G}(t)f(y) = G^*(f)(y,t)$. In this sense the forced ODE and its autonomous part are equivalent up to time-reparametrisation.

Similarly, if ϕ is the flow of the autonomous ODE with vector field f and $\tilde{\phi}$ is the flow of the forced ODE with vector field gf then

$$\tilde{\phi}_{t,t_0} := \phi_{G_{t_0}(t)}, \quad \text{where } G_{t_0}(t) = \int_{t_0}^t g(s)\, ds. \tag{2.5}$$

For example, for $\lambda \in \mathbb{R}$, $t_0 = 0$ $y(t) = e^{\lambda t}$ is the solution of $\dot{y} = \lambda y$. By the above, this implies that $z(t) = e^{\lambda \sin(t)}$ is the solution of the forced ODE of $\dot{z}(t) = \cos(t)\lambda z(t)$.

This idea is used to construct numerical methods for forced ODE (Section 4).

2.1.2 Canonical autonomous extension

Given a non-autonomous vector field $f : \mathbb{R}^n \times \mathbb{R} \supset D \times I \to \mathbb{R}^n$ (with $D \subset \mathbb{R}^n$ and $I \subset \mathbb{R}$ open) then one can view f as an autonomous system on the space $\mathbb{R}^n \times \mathbb{R}$.

The canonical way to view the ODE of f as an autonomous system is to incorporate time into space: The vector field $\hat{f} = (f, 1)$ is considered, defined on a space extended by one dimension $D \times I$, which induces the ODE

$$\dot{y}(t) = f(y(t), \tau(t)) \quad \dot{\tau}(t) = 1, \quad \text{for } (y,t) \in D \tag{2.6}$$

so that the extra ODE $\dot{\tau}(t) = 1$ is introduced.

Definition 2.5. Given a non-autonomous vector field f. The autonomous ODE with vector field $\hat{f} = (f, 1)$ (Equation (2.6)) is called the *canonical autonomous extension* of the ODE with vector field f. The phase space $D \times I$ is called *(autonomously) extended space* and $\hat{f}$ the *(canonically) extended vector field.* $\varnothing$

Given initial points y_0, τ_0, the solutions of the non-autonomous ODE and its canonical autonomous extension are related up to inclusion/projection: the integral curves $\tilde{y}$ of the ODE with vector field $\hat{f}$ satisfies $(y(t), \tau(t)) = \tilde{y}(t)$, where $\tau(t) = \tau_0 + t$ and y the integral curve of the ODE with vector field f.

2.1.3 Hamiltonian autonomous extension and forced systems

Now we treat the Hamiltonian version of the canonical autonomous extension.

The canonical autonomous extension is easy to adapt to the Hamiltonian case: Adding not only a variable τ but also a variable s conjugate to 'time' τ, the canonical autonomous extension becomes, using Equation (2.1),

$$\dot{z} = J^{-1}\nabla H(z,\tau), \qquad \dot{\tau} = 1, \qquad \dot{s} = -\frac{\partial}{\partial \tau} H(z,\tau) = -\frac{d}{dt} H(z(t), \tau(t)) \text{ with Hamiltonian } \tilde{H}(q,\tau,p,s) = H(q,p,\tau) + s. \tag{2.7}$$

Thus, we find an autonomous Hamiltonian ODE with Hamiltonian $\tilde{H}(q,\tau,p,s) = H(q,p,\tau) + s$.

The Hamiltonian H is in general not preserved by the flow. The extended Hamiltonian $\tilde{H}$ is trivially preserved, since (if $\tau(t_0) = t_0$)

$$s(t) = s(t_0) + \int_{t_0}^{t} -\frac{d}{dt}H(z(t),t) = s(t_0) - H(z(t),t) + H(z(t_0),t_0).$$

Phase space together with time τ and its conjugate s is called *extended phase space.*

2.2 Non-autonomous Hamiltonian ODE and symplectic flows

Hamiltonian ODE and their surrounding theory (the Hamiltonian/canonical formalism) are very useful in the description of plethora of physical systems e.g. the tidal wave system. Heuristically, the power of the autonomous Hamiltonian formalism can be explained using the fact that it separates geometry and dynamics[6]:

> "The Hamiltonian formalism is easier to deal with [than the Lagrangian formalism] because geometry—the symplectic structure—and dynamics—the Hamiltonian 1 form dH—are independent ingredients." – Asorey, Cariñena & Ibort [ACI83]

However, considering a geometric framework for non-autonomous systems, they note "dynamics and geometry are coupled again" and the structure of non-autonomous Hamiltonian ODE is not so clear (we present an outlook of this in Section 2.4). We introduce now symplectic maps (the structure-preserving maps of the 'geometry' i.e. the symplectic structure) and canonical transformations (the structure-preserving maps of the 'dynamics' i.e. the Hamiltonian form of the ODE) and equivalence of these two concepts.

For our purposes, structure preserving numerical methods and backward error analysis, the most important results of the autonomous formalism are

1. the fact that Hamiltonian ODE are characterised (locally) by symplectically symmetric vector fields (defined below, Definition 2.7);
2. the fact that Hamiltonian ODE are characterised by symplectic flows.

In this section we show that these two statements hold as well for non-autonomous Hamiltonian ODE on $\mathbb{R}^{2n}$, luckily without the need of any other geometries, beside symplectic geometry on flat space. Thus, possibly the right 'structure' for non-autonomous systems on $\mathbb{R}^{2n} \times \mathbb{R}$ is again the symplectic structure (the matrix J) on $\mathbb{R}^{2n}$.

Besides the wide range of applications and the heuristic separation of geometry and dynamics. Another strength of Hamiltonian ODE are generating functions and the Hamilton-Jacobi equation e.g. [Arn89; GPS02] or [HLW06] chapter VI. In Figure 9 it is shown how the the mentioned objects are related.

2.2.1 Hamiltonian ODE

As mentioned, the anti-symmetric, invertible matrix (non-degenerate 2-form) J has an intrinsically geometric meaning and is usually called the *standard symplectic form.* This geometric meaning can be tied to the fact that, for $n = 2$, $v^t J w$ is the outer product of two vectors $v, w \in \mathbb{R}^2$ which is the (oriented) area of the parallelogram which v, w span. In higher dimensions there is a similar geometric interpretation, Section 2.2.2.

Definition 2.6. A vector field $f \in C^0(D \times I, \mathbb{R}^{2n})$ is *Hamiltonian* if

$$f = J^{-1}\nabla H$$

for some $H \in C^1(D \times I)$ and *locally Hamiltonian* if $\forall y \in D$, $\exists y \in U_y \subset D$ open such that $f|_{U_y}$ (i.e. the restriction of f to $U_y \times I$) is Hamiltonian. $\varnothing$

[6] In our case the symplectic geometry is flat: $\mathbb{R}^{2n}$ together with a matrix J defined below.

We may denote f_H to make the Hamiltonian explicit.

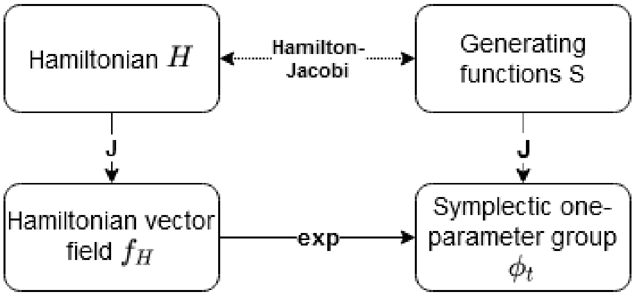

Hamiltonian H
Hamilton-Jacobi
Generating functions S
J
J
Hamiltonian vector field f_H
exp
Symplectic one-parameter group ϕ_t

Figure 9: A diagram that shows how Hamiltonian vector fields and symplectic one-parameter groups, generating functions and the Hamilton-Jacobi equations come into play. The left arrow and the bottom arrow depict respectively the important statements 1. and 2. as discussed in the text.

Definition 2.7. A vector field $f \in C^1(D \times I, \mathbb{R}^{2n})$ is *symplectically symmetric*[7] if $J\nabla_y f(y,t)$ is symmetric on its domain i.e. $JD_y f(y,t) = -D_y f(y,t))^T J$. $\varnothing$

The C^1 vector fields f which are locally Hamiltonian on Euclidean space (the left arrow in Figure 9) are characterised by symplectic symmetry.

Lemma 2.8 (Integrability Lemma [HLW06]). *A vector field $f \in C^1(D \times I, \mathbb{R}^{2n})$ is locally Hamiltonian (with Hamiltonian $H \in C^2(D \times I)$) if and only if it is symplectically symmetric (even globally Hamiltonian if D is simply connected).*

Proof. The proof is constructive and we are only interested in this construction of the (local) Hamiltonian. We refer to [HLW06] chapter VI for details. There they treat the case for autonomous f, but the proof is identical for non-autonomous f.

If f is locally Hamiltonian then $D_y f = J^{-1}\nabla^2 H$ (∇^2 denotes taking the Hessian matrix), implying $JD_y f = \nabla^2 H$, symmetric since H is C^2. Conversely, for every $y \in D$ we pick a small ball U_y (or any convex set) around y and define $H_{[y]} \in C^2(U_y \times I, D)$ by

$$H_{[y]}(z,t) = \int_0^1 Jf(\gamma(\tau),t) \cdot \gamma'(\tau) d\tau = J \int_0^1 f(\gamma(\tau),t) \cdot \gamma'(\tau) d\tau, \tag{2.8}$$

where $\gamma(\tau) = y + \tau(z-y)$ which determines a well-defined Hamiltonian for $f|_{U_y}$ ([HLW06], chapter VI). □

2.2.2 Hamiltonian flows on symplectic Euclidean space

In this section we show that Hamiltonian flows are characterised by their symplecticness (the bottom arrow of Figure 9):

> “Everybody knows that the time-1-shift of the flow of a hamiltonian vector field is a symplectic diffeomorphism” – Kuksin & Pöschel [KP94b].

If $f = f_H$ is Hamiltonian then we may write $\phi_{H,t}$ for the Hamiltonian flow (or ϕ_t if it is clear from which Hamiltonian it is induced).

Definition 2.9. If $g \in C^1(D \times I, \mathbb{R}^{2n})$ and $z \in D$ then $g_t(z) := g(\cdot, t)$ is called *extended symplectic with factor λ at y* if

$$(D_y g_t)^T(y) J^{-1} D_y g_t(y) = \lambda J^{-1}$$

for all $t \in I$ and some $\lambda \neq 0$. Furthermore, g is called *extended symplectic with factor* λ if g is symplectic at z with factor λ for all $z \in D$. $\varnothing$

Definition 2.10. A map g is called *symplectic* (at z) if it is extended symplectic with $\lambda = 1$ (at z). $\varnothing$

If $g = (Q,P)$ is symplectic (possibly depending on $t \in I$) then

$$D_{(q,[)} g(p,q) = \begin{pmatrix} D_q Q & D_p Q \\ D_q P & D_p P \end{pmatrix}$$

and the symplecticness condition reads[8]

$$Q_q^T P_q = P_q^T Q_q \qquad Q_p^T P_p = P_p^T Q_p \qquad -Q_q^T P_p + P_q^T Q_p = I.$$

[7] Usually these vector fields are called Hamiltonian, but confusion arises on flat space, where the manifold and the tangent space are hardly distinguishable.

[8] One may check that in $\mathbb{R}^2$ every invertible map $g : D \to \mathbb{R}^2$ is extended symplectic with $\lambda = \det(Dg)$.

Remark 2.11. *An equivalent definition of symplecticness is obtained by replacing $J^{-1}(= J^T = -J)$ by $-J^{-1}$ and therefore also by switching the transpose of the two Jacobian matrices: $g'J^{-1}(g')^T = J^{-1}$.*

An important property of the flow of Hamiltonian vector fields is volume preservation.

Theorem 2.12 (Liouville's Theorem)**.** *If $f \in C^1(D \times I, \mathbb{R}^{2n})$ is a vector field with zero divergence $\nabla \cdot f(z, t_0) = tr(f'(z, t_0)) = 0$, then $\det D_z\phi_t(z, t_0) = \det \phi'_{f,t}(z, t_0) = 1$ for all z, t_0, t in the domain.*

Corollary 2.13. *If $H \in C^2(D \times I)$ then the flow ϕ_t preserves volume.*

There is a more general and more important property of the flow. It is not only area preserving, but also preserves sum of the infinitesimal 2-dimensional areas of parallelograms in the subspaces spanned by q_i, p_i for $1 \leq i \leq n$ (pictures and more rigorous interpretation can be found in [Arn89; HLW06]).

Now we prove that Hamiltonian flows are characterised by their symplecticness, of which the 'if' implication is due to Poincaré. The proof for the autonomous case can be found in [HLW06] and below it is seen that it is identical for the non-autonomous case.

Theorem 2.14 (Poincaré 1899 (also [HLW06]))**.** *Suppose $\phi \in C^1(\tilde{D} \times I, D)$ is the flow induced by the vector field $f \in C^1(D \times I, \mathbb{R}^{2n})$, where $\tilde{D} \subset D \times I$ open. Then $\phi_{\cdot,t_0}$ is symplectic (for all t_0 in the domain) if and only if f is locally Hamiltonian (or globally Hamiltonian if D is simply connected).*

Proof. The variational equation $D_t\phi'_{t,t_0} := D_{p,q}\dot{\phi}_{t,t_0} = f'(\phi_{t,t_0}, t)\phi'_{t,t_0}$ gives

$$D_t((\phi'_{t,t_0})^T J\phi'_{t,t_0}) = (\phi'_{t,t_0})^T \left((f')^T J + Jf'\right) \phi'_{t,t_0}$$

$\Longleftarrow$: Now, if f is Hamiltonian with Hamiltonian function $H \in C^2(D \times I)$ then

$$D_t((\phi'_{t,t_0})^T J\phi'_{t,t_0}) = (\phi'_{t,t_0})^T \left(\nabla^2 HJ^{-T}J + JJ^{-1}\nabla^2 H\right) \phi'_{t,t_0} = 0.$$

since $\nabla^2 HJ^{-T}J + JJ^{-1}\nabla^2 HJ = -\nabla^2 H + \nabla^2 H = 0$. Therefore $(\phi'_{t,t_0})^T J\phi'_{t,t_0} = \left((\phi'_{t,t_0})^T J\phi'_{t,t_0}\right)|_{t=t_00} = J$ since $\phi_{t_0,t_0} = Id$, thus ϕ_{t,t_0} is symplectic.
$\Longrightarrow$: Conversely, if $\phi_{\cdot,t_0}$ is symplectic for all t_0 in the domain, then

$$(\phi'_{t,t_0})^T \left((f')^T J + Jf'\right) \phi'_{t,t_0} = 0$$

for all t in the domain, implying $(t = t_0)$ $(f')^T J + Jf' = 0$ (take $t = t_0$). Thus Jf' is symmetric and the Integrability Lemma 2.8 implies that f' is locally Hamiltonian (or globally in the simply connected case). □

Thus, on simply connected subsets of $\mathbb{R}^{2n}$ symplectic flows and flows from Hamiltonian ODE are the same and we see indeed that

> "Symplectic maps are the discrete-time analogue of Hamiltonian motion" – James Meiss [Mei92]

2.2.3 Symplectic maps and symplectic flows related

As quoted above, symplectic flows are the discrete analogue of Hamiltonian flows. Indeed, using the suspension and Poincaré map (e.g. [BS02] chapter 1) the discrete symplectic map and the continuous Hamiltonian flow can be related. This was for example, as mentioned in the Introduction 1, considered by Moser and Douady in the context of discrete and continuous KAM equivalence [Dou82; Mos87] (see also [Wig03] chapter 14).

More generally, the problem of finding a Hamiltonian flow, given a symplectic map, can be reformulated as finding an embedding of a symplectic map into Hamiltonian flows. This embedding problem (and similar versions in the smooth and analytic case of this problem) has received much attention see e.g. [TZ09] chapter 1.3 and references therein, and additionally [Gol95; Tre99; Gio12; ST16; GV18; Tre19]. In particular, we will use this embedding problem in the context of non-autonomous backward error analysis, which has already been done extensively by Moan [Moa03; Moa05; Moa06] and McLachlan and O'Neale [MO10].

2.2.4 Properties of symplectic maps

Symplectic maps have interesting properties

- Symplectic maps are locally C^1 invertible due to the inverse function theorem.
- The inverse is also symplectic on its domain of definition.
- Composition of symplectic maps generates symplectic maps.

2.3 Canonical coordinate transformations and symplectic maps

As in Appendix A we are interested in coordinate changes (diffeomorphisms) and how they transform the ODE. In particular, we are interested in coordinate changes which preserve the Hamiltonian form, called *canonical coordinate transformations (canonical maps).* The importance of these coordinate transformations is immense: An idea used by Jacobi to solve general Hamiltonian ODE [HLW06] was to use time-dependent or time-independent coordinate transformations so as to find coordinates such that the dynamics become trivial, which lead to the Hamilton-Jacobi equation (Appendix [HLW06] chapter VI). In this section we will see furthermore that canonical maps are equivalent to symplectic maps.

We start with a time-independent coordinate transformation $g \in C^1(D, \mathbb{R}^{2n})$. If $z \in C^1(I, \mathbb{R}^{2n})$ is an integral curve of an ODE with vector field $f \in C^1(D, \mathbb{R}^{2n})$ then $y := g \circ z$ is a solution of the ODE

$$\dot{y} = Dg(z)f(z) = Dg^{-1}(y) \cdot (f \circ g^{-1}(y)) = g_*(f)$$

$g_*(f)$ is called the *pushforward vector field*, Appendix A.3.

Definition 2.15. A coordinate transformation $g \in C^1(D, \mathbb{R}^{2n})$ with D simply connected is (a) *canonical (coordinate transformation) with factor* $\lambda \neq 0$ if, given an autonomous or non-autonomous Hamiltonian vector field f_H on D, $g_*(f_H)$ is again Hamiltonian with Hamiltonian

$$K(y,t) = \lambda^2 H(g^{-1}(y), t).$$

∅

In [HLW06] chapter VI.2, it is proven that canonical coordinate transformation with factor $\lambda \neq$ are equivalent to extended symplectic maps with factor $\lambda \neq 0$.

We now discuss time-dependent coordinate transformations g. Then the pushforward vector field is (Appendix A.3)

$$g_*(f)(y,t) = D_y(g^{-1})(y,t) \cdot \left(f(g^{-1}(y,t),t) - D_t(g^{-1})(y,t)\right). \tag{2.9}$$

Definition 2.16. If $g \in C^1(D \times I, \mathbb{R}^{2n})$ is a coordinate transformation and D is simply connected then g is called *extended canonical with factor* $\lambda \neq 0$ if

$$D_y(g^{-1})(y,t)D_t(g^{-1})(y,t) \tag{2.10}$$

is symplectically symmetric (Definition 2.7) and for any Hamiltonian vector field f_H, the pushforward vector field $\tilde{f} = g_*(f_h)$ is induced by the Hamiltonian

$$\tilde{H} = \lambda^2 K - \int_{y_0}^{y} J D_y(g^{-1})(\gamma,t) D_t g^{-1}(\gamma,t) + c\, d\gamma, \tag{2.11}$$

where γ a path connecting y and a fixed $y_0 \in D$ (see the Integrability lemma 2.8), where $K(y,t) := H(g^{-1}(y,t),t)$ and where $c > 0$ is an unidentified constant. ∅

We will see next that actually, Equations (2.10) and (2.11) are superfluous (i.e. below stated as 1. $\implies$ 2.) , so that canonical maps may also be thought of as the maps which preserve Hamiltonian form, for any Hamiltonian and that these are precisely the symplectic maps.

Proposition 2.17 ([LY68; MO17; Car71], section 1.2)**.** *For an invertible map $g \in C^2(D \times I, \mathbb{R}^{2n})$ the following are equivalent*

1. *g is such that if f_H is a Hamiltonian vector field, then $g_*(f_h)$ is Hamiltonian.*
2. *g is extended canonical with factor $\lambda \neq 0$*
3. *g is extended symplectic with factor $\lambda \neq 0$ and invertible.*

Proof. A proof can be found in [LY68], [MO17] section 2.6 or [Car71], section 1.2. □

The equivalence of proposition 2.17 is mainly useful in two cases: The first is that, if one has a symplectic, time-independent map $g \in C^1(D, \mathbb{R}^{2n})$ and a Hamiltonian $H(q, p, t)$ then we will without further notice use that the pushforward vector field is induced by the Hamiltonian $K(Q, P, t) = H(g^{-1}(Q, P), t)$. The second is more speculative and is that equivalence of time-dependent symplectic maps and canonical transformation indicates that also for non-autonomous Hamiltonian ODE the morphisms (see the next section, Section 2.4) may be given by symplectic maps and equivalently canonical transformations.

2.4 Outlook: Structure of non-autonomous Hamiltonian ODE

In Section 2.2 a quote of [ACI83] indicated that, heuristically, the strength of the autonomous Hamiltonian formalism lies in the separation of "geometry" (the symplectic form) and "dynamics" (the Hamiltonian). For non-autonomous Hamiltonian ODE, however, "dynamics and geometry are coupled again" [ACI83]. Additionally [ACI83] mention that "the geometrical framework for describing time-dependent systems is not so well established". Even though this was in 1983, this shows some of the troubles to define the framework for non-autonomous Hamiltonian ODE and consequently the respective structure-preserving numerical methods (Section 4), see also the quotes of Marthinsen & Owren and Skeel & Cieslinski in the introduction, Section 1.

From the perspective of category theory, we suppose that the structure of non-autonomous Hamiltonian ODE consists of 'objects' and 'morphisms'. We mention now some ideas about the objects (the "geometrical framework") and morphisms (symplectic/canonical maps) of non-autonomous Hamiltonian ODE. For the both the objects and the morphisms we mainly refer to other sources and speculate a bit in the Euclidean case (for $(q, p, t) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$).

2.4.1 Objects: Geometric framework

Even today the geometrical framework does not seem well established as there seem to be several geometrical frameworks: Using a presymplectic structure on a contact manifold [ACI83; Car+87; AM08; Bar+08], a cosymplectic structure [LS17; EG18; Leó+22] or a precosymplectic structure, also see [BCT17; Riv21] and references contained in these articles. Additionally, there are frameworks which *seem* more based on 'dynamics' (instead of 'geometry'): E.g. in [VV21] or a jet bundle representation [GMS97; Bar+08; Sar13a], but it is not clear to the author how they relate to the geometric frameworks (if at all). Any of these may be interesting to construct non-autonomous Hamiltonian structure-preserving methods.

We consider the case of Euclidean space. Phase space $(q, p) \in \mathbb{R}^{2n}$ can be extended either to autonomously extended space $(q, p, t) \in \mathbb{R}^{2n} \times \mathbb{R}$ through the introduction of the new time variable t using a canonical autonomous extension (Definition 2.5) or to extended phase space $(q, p, t, s)\mathbb{R}^{2n} \times \mathbb{R}^2$ where the variable s conjugate variable to t is introduced (Definition 2.16).

The space $\mathbb{R}^{2n} \times \mathbb{R}$ has a special geometry similar to the symplectic form.

Proposition 2.18 ([Arn89] chapter 44.C or [AM08] theorem 5.1.13)**.** *Given a non-autonomous Hamiltonian $H \in C^2(\mathbb{R}^{2n} \times \mathbb{R})$ ($H = H(q, p, t)$. The flow ϕ_{t,t_0} of a non-autonomous Hamiltonian ODE with Hamiltonian*

H preserves the matrix

$$\tilde{J} = \begin{pmatrix} 0_n & -Id_n & D_pH \\ Id_n & 0_n & D_qH \\ -D_pH & -D_qH & 0 \end{pmatrix}$$

where $0_n, Id_n$ are respectively the n-dimensional zero and identity matrix and $H_p = D_pH$, $H_q = D_qH$. In other words the flow ϕ_{t,t_0} satisfies

$$\nabla\phi_{t,t_0}\tilde{J}\nabla\phi_{t,t_0}^T = \tilde{J},$$

where the divergence $\nabla = \nabla_{q,p,t_0}$ is now taken also with respect to time.

In particular this implies again that the non-autonomous flow is symplectic. The above mentioned contact manifold is the space $\mathbb{R}^{2n} \times \mathbb{R}^n$ together with this matrix, the contact form (or on a manifold an associated two-form, see again [AM08]). Thus, we see in the Euclidean case that a contact form might induce *extra* geometric structure.

On extended phase space $\mathbb{R}^{2n} \times \mathbb{R}^2$, one may introduce the extended symplectic form $J_{2(n+1)}$. We extend now the non-autonomous flow $\phi_{t,t_0} : \mathbb{R}^{2n}$ from $\mathbb{R}^{2n}$ to $\mathbb{R}^{(2n+2)}$ with the extra solutions

$$\tau(t) = t + \tau_0, \quad s(t) = H(\phi_{t,t_0}(q_0, p_0, t_0) - H(q_0, p_0, t_0) + s_0$$

for initial values q_0, p_0, τ_0 and s_0. Then the resulting autonomous flow on $\mathbb{R}^{2n+2}$ with autonomous Hamiltonian $H(q, p, t) + s$ (Equation (2.7)) preserves the extended symplectic form. The non-autonomous flow $\phi_{t,t_0} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ of a non-autonomous Hamiltonian H as in the proposition above does not have the dimensions to preserve this symplectic flow. By construction, (q, p, t) (or (q, p, τ)) evolve independently of s and one may ask if there is really a different structure than the contact structure involved.

2.4.2 Morphisms: Transformations, non-autonomous systems and extended phase space

We have encountered three types of transformations: Most importantly we have canonical (coordinate) transformations and symplectic maps which are the structure preserving maps/morphisms of autonomous Hamiltonian ODE. Finally, we have time-reparametrisations, which seemed less important to consider for structure preservation. Nevertheless, with the purpose of finding the morphisms of non-autonomous Hamiltonian ODE, the relation between canonical transformations, symplectic transformations and time-transformations is investigated.

When considering regular phase space we have seen that, for invertible maps, the canonical and symplectic maps are equivalent even when they are time-dependent (not that they do not transform time[9], see also [AM08] chapter 5.2.). Time-reparametrisation are also considered [CMM22] and they may be seen as an infinitesimal canonical transformation [CIL87].
However, on the contact manifold and in extended phase space, where time may also be transformed, it seems that the concept of canonical transformations can be generalised in various ways [ACI83; Car+87; Joh89; GMS97; Tsi00; Str05; AM08] and the relation between canonical and symplectic maps is not set in stone. Thus it is not clear to the author what the correct morphisms are and, giving an outlook on structure preserving numerical methods, it is also not clear what type of transformation/morhpism a structure preserving numerical methods for a non-autonomous, Hamiltonian ODE should be.

3 Integrability, action-angle coordinates, KAM theory and KAM tori in the tidal wave system

Before proving the existence of persistent invariant tori (KAM tori) in the 2π-map (or the (tidal) Poincaré map as in [Wal21]) of the numerically integrated tidal wave system, we first prove existence of invariant and KAM tori in the 2π-Poincaré map of the theoretical tidal wave system. This is done using (continuous) KAM

[9] Of course the flow of a non-autonomous Hamiltonian ODE does transform time and one may really wonder whether symplecticness is the only demand for a structure preserving numerical method.

theory. Two strategies can be used (Figure 10).

Denote by ϕ_{t,t_0} the flow of the tidal wave system of Section 1.5. The first strategy is to view the 2π-Poincaré maps $\phi_{t_0+2n\pi,t_0}$ for $n \in \mathbb{Z}$ as a symplectic, discrete dynamical system (Appendix A). Indeed, since the time-dependence is periodic with period 2π one sees that $\phi_{t_0+2n\pi,t_0} = \phi^n_{t_0+2\pi,t_0}$, so that the 2π-Poincaré map forms, for fixed t_0 a discrete dynamical system induced by the map $\phi_{t_0+2\pi,t_0}$, which is symplectic (Proposition 2.14).

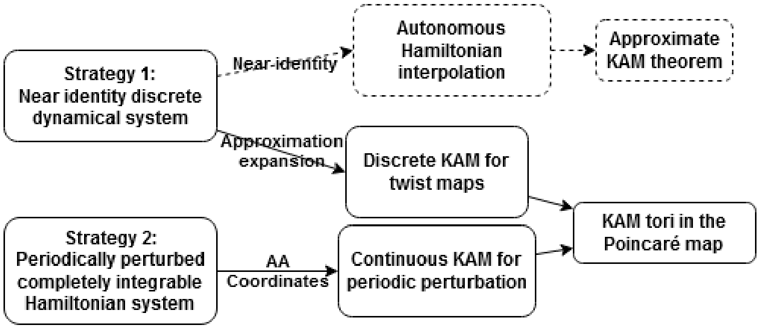

Strategy 1:
Near identity discrete dynamical system
Near-identity
Autonomous Hamiltonian interpolation
Approximate KAM theorem
Approximation expansion
Discrete KAM for twist maps
KAM tori in the Poincaré map
Strategy 2:
Periodically perturbed completely integrable Hamiltonian system
AA Coordinates
Continuous KAM for periodic perturbation

Figure 10: A diagram that shows the two strategies of proving the existence of KAM tori (and the extra strategy in the footnote denoted with dashed lines).

In this discrete 2-dimensional setting, one naturally proves the existence of invariant tori using Mosers (discrete) version of KAM theory for area-preserving twist maps [Mös62] (or see Shang [Sha99]). To prove the twist condition, one likely has to approximate the 2π-map $\phi_{t+2\pi,t}$ in some way. For example, an approximation of this map for the tidal wave system has been studied in [BRZ94] using the "orbit expansion" (defined in [RZ92]). In the case of the tidal wave system, the orbit expansion seems to reduced to a Picard iteration (as in [BBM22]). Moreover [BBM22] considers other expansions approximating the flow, which are adapted specifically to non-autonomous, non-linear ODE and therefore might be useful. Furthermore in [Wal21] section 7 it was mentioned that, for specific parameter values related to the Bessel function (the Bessel function was found in [BRZ94] using the orbit expension) the 2π-Poincaré was near the identity. Thus, it was suggested that the paper [BG94] could be used to find an autonomous Hamiltonian which interpolates the symplectic map: It has time-1 flow very close to the 2π-Poincaré map[10]. We will not consider this case in this thesis and come back to this for possible future research in Section 8.

We will adopt a second strategy: If the non-perturbed system is autonomous and phase space consists entirely of invariant tori (e.g. the unperturbed examples in the introduction, Section 1) then there already exists a KAM theory for small non-autonomous perturbations, depending quasi-periodically on time (Definition 3.1). Therefore, one finds KAM tori in extended phase space from which invariant tori can be found in the Poincaré, Section 3.3

The disadvantages of the second strategy are that one cannot say anything about the relation of the Poincaré map to the Bessel functions using the fact that it is near the identity. Furthermore, the autonomous, 'interpolating' Hamiltonian might give a different perspective of the Poincaré map. The main advantage of the second strategy is that it seems to be easier. Instead of using an approximation expansion (e.g. one in [BBM22]) of the 2π-Poincaré map and checking the twist conditions, one transforms to action-angle coordinates and checks the simple conditions of the periodically perturbed KAM theorem. Therefore we use theory which has already been developed and avoid lengthy proofs, which is in the spirit of the quotes of Douady and Mclachan & O'Neale in the introduction. Furthermore, the action angle coordinates may also give some additional information of the system.

To this end, we introduce complete integrable systems, for which action angle coordinates are available. We discuss how action angle coordinates can be constructed in the 2-dimensional case and state the KAM theorem for (quasi-)periodic perturbations. Afterwards we apply it to the tidal wave system and prove the existence of KAM tori in the theoretical system: An appetiser for the numerical case.

First let us define quasi-periodicity.

[10] At this point one may also use an approximate KAM theorem for flows which are (exponentially) close to integrable ones, see [HLW06] chapter X.5 and Section 6.

Definition 3.1. Suppose $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$, $n, m \in \mathbb{N}$, $f = f(y, t)$. Then f is called *quasi-periodic with frequencies* $\Omega \in \mathbb{R}^s$ *in the variable* t there exists a function $g : \mathbb{R}^n \times \mathbb{T}^s \to \mathbb{R}^m$ (g is periodic in its last s variables) such that

$$f(y, t) = g(y, \Omega_1 t, \dots, \Omega_s t).$$

Nota bene that a quasi-periodic function should not be confused with quasi-periodic motion (on a torus) (see Section 3.2).

3.1 Complete integrability and action-angle coordinates

In this section we discuss local Liouville and (global) Liouville/complete integrability of Hamiltonian systems and action-angle coordinates. The treatment given here is heavily based on the treatment in [HLW06] chapter X.1.

The phase space of global Liouville/complete integrable systems consists entirely of invariant tori (e.g. the examples at the start of the introduction, Section 1) and we will see that the unperturbed tidal wave system is completely integrable. Furthermore, global Liouville/complete integrability and action-angle coordinates provide the mathematical framework for the application of KAM theory and can therefore be used to prove the existence of KAM tori in the tidal-wave system.

A historical introduction for local Liouville integrability, (developed by Bour (for autonomous systems) [Bou55]) and Liouville (allowing also time-dependence)) [Lio55] and (global) Liouville (or Liouville-Arnold) integrability can be found in [HLW06] chapter X.1. In particular, the local integrability results were developed to integrate the equations of motion of Hamiltonian systems:

> "one of the great dreams of the 18th and 19th cenury analytical mechanics was to solve the equations of motion of mechanical systems by "quadrature." – Hairer, Lubich & Wanner [HLW06]

Local Liouville integrability was precisely the construction to do so, although only locally.

After the shattering of this dream by the first rigorous non-integrability results by Poincaré, focus shifted to the study of not only solutions in quadrature, but global dynamics i.e. dynamical systems theory [HLW06] chapter X.1.2. Therefore a global version was needed: Complete/Liouville integrability.

Thus, for completely (global Liouville) integrable systems phase space is very easy (only invariant tori) and the flow can be solved exactly (in quadrature). This is a very strong statement, and as a consequence covers only exceptions: very few systems are (completely) integrable [HLW06]. Nevertheless, many important systems can be seen as small perturbations of completely integrable and perturbation theory (e.g. KAM theory) can be applied to obtain results for these systems.

3.1.1 Locally Liouville integrable systems

The idea of local Liouville integrability is the following: Given a Hamiltonian H then a canonical transformation ℓ is sought which transforms the Hamiltonian ODE

$$\dot{q} = D_p H(q, p), \quad \dot{p} = -D_q H(q, p)$$

into the ODE

$$\dot{x} = 0, \quad \dot{y} = \omega(x)$$

for some function $\omega = D_x K$, where $K(x, y) = K(x)$ is a Hamiltonian depending only on x. Such a map ℓ is constructed locally and explicitly using the powerhouse of Hamiltonian dynamics: Hamilton-Jacobi theory, [HLW06] chapter VI.

Hamilton-Jacobi theory directs the search from such a symplectic map ℓ to the search of a generating function S, which then induces the transformation ℓ, where S is defined by the Hamilton-Jacobi equation

$$H(D_q S(x, q), q) = K(x).$$

Heuristically, since $dS = pdq + ydx$ (see [HLW06] chapter VI) and the x are first integrals (so $dx = 0$ along the solution), the function S is given, heuristically, by $S(x,q) = \int_{q_0}^{q} p(x,s)ds$ for some fixed q_0 (this follows heuristically from integrability lemma 2.8, which also implies that S is defined locally).

The problem is now shifted to the construction of a function $p(x,q)$ which is well-defined and has a symmetric gradient so that S is well-defined. Such a map p can be constructed for locally Liouville integrable systems.

Definition 3.2. A Hamiltonian ODE on $D \subset \mathbb{R}^{2n}$ is locally Liouville integrable at $(q_0,p_0) \in D$ if there exist n first integrals $F_i \in C^1(D)$ which are

- in involution locally around (q_0,p_0) i.e. there exists an open neighbourhood of (q_0,p_0) such that $\{F_i,F_j\} = 0$ on this neighbourhood ($\{\cdot,\cdot\}$ is the Poisson bracket, Equation (2.2))
- and have linearly independent gradients ∇F_i at (q_0,p_0). $\varnothing$

It was realised by Bour [Bou55]) (for time-dependent F and H) and Liouville [Lio55] (including also time) that this was sufficient to find such a coordinate change, which is stated in the next theorem.

Lemma 3.3 ([HLW06] chapter X.1)**.** *If an ODE on $D \subset \mathbb{R}^{2n}$ is locally Liouville integrable at $(q_0,p_0) \in D$ with first integrals $F_1 \ldots, F_n \in C^k(D)$ (analytic), then there exist $G_1,\ldots,G_n \in C^k(U)$ (analytic), such that the map*

$$\ell = (F,G) : (q,p), \mapsto (x,y) \quad \textit{where} \quad F = (F_1,\ldots,F_n),\, G = (G_1,\ldots G_n),$$

is symplectic and $C^k(U)$ (analytic), for some open $U \subset D$.
Furthermore, if $H = \sum_i v_i F_i$ for some $v_i \in \mathbb{R}$ then the flow $\phi_{H,t}$ satisfies

$$\phi_{H,t}(q,p) = \ell^{-1}(x, y + tv) \tag{3.1}$$

if $\ell(q,p) = (x,y)$.

Proof. A proof can be found in, for example, [HLW06] chapter X.1. We will sketch the idea of the construction of ℓ.
The first integrals $x_i = F_i(q,p)$, being linearly independent, can be inverted locally, so that $p = p(x,q)$. Since the first integrals are in involution, $D_q p(x,q)$ is symplectically symmetric so that the Integrability lemma can be used to construct the generating function $S(x,q) = \int_{q_0}^{q} p(x,q)dq$ *locally* (i.e. for $U \subset D$, for fixed $q_0 \in U$), which induce the map ℓ ([HLW06] chapter VI).

Furthermore, an ODE of the Hamiltonian F_i is transformed by ℓ (locally) into a constant ODE

$$\dot{x} = 0, \quad \dot{y}_i = 1, \quad \dot{y}_j = 0 \text{ for } i \neq j,$$

since the transformed Hamiltonian satisfies $H(\ell^{-1}(x,y)) = F_i(\ell^{-1}(x,y)) = x_i$ (proposition 2.17). Moreover, since $\{F_i,F_j\} = 0$, the flows commute ([HLW06]): $\phi_{F_i,t} \circ \phi_{F_j,\tilde{t}} = \phi_{F_j,\tilde{t}} \circ \phi_{F_i,t}$. Therefore, if $H_v = \sum_i v_i F_i$, then ℓ transforms the ODE of H_y locally into

$$\dot{x} = 0, \quad \dot{y} = v,$$

which proves Equation (3.1). □

3.1.2 Complete integrability and action-angle coordinates

Complete integrability is found by demanding a more global version of local integrability.

Definition 3.4. A Hamiltonian system with Hamiltonian $F_1 = H$ on $D \subset \mathbb{R}^{2n}$ is completely integrable if there exist n functions $F_i \in C^k(D)$ (analytic) such that

- All pairs F_i, F_j are in involution on D (i.e. $\{F_i,F_j\} = 0$ for all $1 \leq i,j \leq n$)
- The gradients ∇F_i are linearly independent on D

- The solutions of the Hamiltonian ODE induced by the F_i are defined globally in time (i.e. on all of $\mathbb{R}$).

The involutivity implies that the level sets $M_x = \{(q,p) \,|\, F(q,p) = x\}$ are invariant under the flows $\phi_{F_i,t}$ for all F_i, Equation (2.3). Moreover, if the gradients are independent then these sets are n-dimensional manifolds. The following theorem proves the existence of action-angle coordinates on M_x for x in an open set.

Theorem 3.5 (Arnold-Liouville; [Arn89] or [HLW06])**.** *Let $F_1, \ldots, F_d$ form a completely integrable system as in Definition 3.4. Suppose that the level sets M_x are compact and connected for all x in a neighbourhood of $x_0 \in \mathbb{R}^n$. Then, there are neighbourhoods B of x_0 and D of $0 \in \mathbb{R}^n$ such that:*

- *For every $x \in B$, the level set M_x is an n-dimensional torus that is invariant under the flow of the system with Hamiltonian F_i $(i = 1, ..., d)$.*
- *There exists a bijective symplectic transformation*

$$\psi : D \times \mathbb{T}^n \to \bigcup_{x \in B} M_x \subset \mathbb{R}^n \times \mathbb{R}^n : \quad (a, \phi) \to (q, p)$$

such that $(F_i \circ \psi)(a, \phi)$ depends only on a, i.e., $F_i(p, q) = f_i(a)$ for $(p, q) = \psi(a, \phi)$ $(i = 1, \ldots, n)$ for some functions $f_i : D \to \mathbb{R}$.

Proof. A proof can be found in [Arn89] section 49 and 50 or [HLW06] chapter X.1.3. □

The variables $(a, \phi) = (a_1, ..., a_n, \phi_1 \mod 2\pi, ..., \phi_n \mod 2\pi)$ are called action-angle variables, where a denotes the action and ϕ the angle. In particular about the regularity of the transformation [HLW06] notes.

Remark 3.6. *If the Hamiltonian is real-analytic, then the transformation ψ to action-angle variables variables is real-analytic.*

Thus, for completely integrable systems, the level sets M_x depend only on the some action coordinate a and are equal up to symplectic transformation to an (invariant) torus i.e. $M_x = \psi(a, \mathbb{T}^n)$. Therefore phase space consists, locally, of invariant tori. Furthermore, the transformed Hamiltonian $K_i = F_i \circ \psi$ depends only on a, such that the the transformed ODE of F_i satisfies

$$\dot{a} = 0 \quad \dot{\phi} = D_a K_i(a) =: \omega(a)$$

where $\omega(a) \in \mathbb{R}^n$ are the frequencies on the level set M_x. Thus, in action-angle coordinates it is a trivial task to solve the equations of motion exactly. However, the construction of ψ is in general not easy.

3.1.3 Action-angle coordinates for autonomous, planar systems

For an autonomous, planar $(n = 1)$ Hamiltonian system one only needs one invariant in involution to form a complete integrable system Definition 3.4. Therefore autonomous planar systems are very often completely integrable, in particular on parts of phase space where the orbits are compact.

For planar systems the generating function S, which induces the transformation ψ to action-angle coordinate, can be found easily using the Hamilton-Jacobi equation

$$H(D_q S(a, q), q) = K(a)$$

such that

$$p = D_q S(a, q) \quad \text{and} \quad \phi = D_a S(a, q).$$

Since a should be constant along the flow ϕ_K we find, just like in Section 3.1.1, along solutions $(q(t), p(t))$, the generating function

$$S(a, q) = \int_{q_0}^{q} p(a, q)\, ds = \int_{t_0}^{t_1} p(t) \dot{q}(t) dt.$$

The condition that ϕ is 2π periodic i.e. $2\pi = \oint_{M_x} d\phi$ implies that ([Arn89] section 50).

$$a(q,p) = \frac{1}{2\pi}\oint_{M_x} pdq = \frac{1}{2\pi}\oint_{D_x} dp \wedge dq, \quad \phi(q,p) = 2\pi\frac{y(q,p)}{T(q,p)}, \tag{3.2}$$

where M_x is the curve with Hamiltonian value $H(q,p) = x \in \mathbb{R}$, D_x is such that $\partial D_x = M_x$ i.e. M_x is the boundary and where $T(q,p)$ is the period of the orbit with initial points (q,p). Thus in this case the action-variable is the volume of D_x, the region enclosed by the orbit. Here we find that y and T are defined by

$$y(q,p) = \int_{q_0}^{q} \frac{\partial}{\partial x} p(s, H(q,p))\, ds \qquad T(q,p) = \oint_{M_x} \frac{\partial}{\partial x} p(s, H(q,p))\, ds$$

for some fixed q_0. Doing so, y is multi-valued with degeneracy addition of multiples of T but is well-defined if work modulo T i.e. set $0 \leq y \leq T$.

3.1.4 Integrability and action-angle coordinates for non-autonomous Hamiltonian systems

To be thorough we also mention briefly how integrability and action angle coordinates can be defined for non-autonomous systems. This section is not important for application to the tidal wave system.

As mentioned above, it was Liouville [Lio55] who showed that the construction of integrability could be extended to non-autonomous Hamiltonian systems. This is easily done by considering the extended Hamiltonian, Equation (2.7).

Definition 3.7. A non-autonomus Hamiltonian ODE on $D \subset \mathbb{R}^{2n}$ with Hamiltonian $H \in C^2(D \times R)$ is called completely integrable if there exists $F_1, \dots F_n \in C^k(D \times \mathbb{R})$ (for some $k \in \mathbb{N}$) possibly depending on time such that

- The F_i are first integrals of the flow of H
- The pairs F_i, F_j are in involution on $\mathbb{R}^{2n}$ i.e. $\{F_i, F_j\} = 0$
- The functions F_i have linearly independent gradients everywhere
- The flow ϕ_H, ϕ_{F_i} are exists globally in time

If the non-autonomous system is completely integrable in this sense, then one can see that the trivially extended system (with Hamiltonian $\tilde{H} = H + s$, as in Equation (2.7)) is completely integrable. The notion of complete integrability of non-automous systems was further considered in [BB98; BC10; Sar13b]. Similarly, notions of action-angle coordinates have been constructed for non-compact invariant sets [GMS02; FGS02; Fio04; FS07], which is needed since in the canonically extended system $\dot{t} = 1$ and necesarily all orbits are unbounded (non-compact).

3.2 KAM theory

Having discussed the right mathematical framework (action-angle coordinates), we now consider KAM theory. As was mentioned in the introduction, Section 1, the KAM theory proves the persistence of invariant tori when a completely integrable system i.e. with Hamiltonian $H_0(a)$ is perturbed with a small Hamiltonian perturbation $H_1(a, \phi)$. We again quote Arnold in his 1963 paper, with less omission, thus revealing the two crucial conditions of an application of KAM theory in the right framework:

- A non-degeneracy condition for the unperturbed Hamiltonian $H_0(a)$ (or for the frequency $\omega_0(a) := D_a H_0(a)$);
- a strong non-resonance condition on the frequencies ω.

> "We assume [the non-degeneracy condition det $\left|\frac{\partial \omega_0(a)}{\partial a}\right|$ = det $\left|\frac{\partial^2 H_0(a)}{\partial a^2}\right| \neq 0$] [...]. It turns out that, for a small perturbation, [$H = H_0(a) + \mu H_1(a, \phi)$] ($\mu \ll 1$), most of the tori with incommensurable frequencies ω^* satisfying [a non-resonance condition] do not disappear, but are merely slightly deformed." – V. I. Arnold [Arn63]

Often it is said that Kolmogorov stated, without proof, the first KAM result in this form in the 50s. However, Hubbard [Hub07] does state that some who were in Kolmogorov's seminar in Amsterdam in 1957 say that Kolmogorov gave a proof there. Thus, it is not entirely clear, but we refer to [Dum14] for an excellent historical treatment of KAM, in particular chapter 5.3 on this issue. Further references on (short) historical treatments of KAM theory are for example [Sev16; Val19; Fel22] and for an introduction to KAM theory one could look at
[Arn89; Sev94; Way96; Mos99; De +01; Sev03; HI03; Bro04; HLW06; AKN07; Hub07; Chi09; Pös09; TZ09; KN13; GS18; Fel22].
The non-degeneracy and strong non-resonance conditions are sufficient, but not necessary. In particular, the non-degeneracy condition in the quote is called the *Kolmogorov* degeneracy condition. More general non-degeneracy conditions are for example the isoenergetic non-degeneracy conditions and the Rüssmann degeneracy conditions, the latter of which is also necessary, see also [Han11]. The non-resonance conditions have also been relaxed, see the historical introduction in [Val19] and references therein.

Most often, the non-resonance condition is Siegel's Diophantine condition (or the *strong non-resonance condition*) [HLW06] (where $\|\cdot\|_1$ is the 1-norm)

$$|k \cdot \omega| \geq \gamma \|k\|_1^{-\sigma}, \quad k \in \mathbb{Z}^d \setminus \{0\} \text{ and some fixed } \gamma > 0, \sigma \geq 0 \tag{3.3}$$

Definition 3.8. If $\omega \in \mathbb{R}^s$ satisfies the *strong non-resonance* of Equation (3.3) condition with fixed $\gamma > 0, \sigma \geq 0$ then ω is called (σ, γ)-*Diophantine.*
If $\omega \in \mathbb{R}^n$ satisfies $|k \cdot \omega| = 0$ for some $k \in \mathbb{Z}^n \setminus \{0\}$ then we say that the frequencies ω *resonate.*

∅

Mathematically, the non-resonance condition is due to a *small divisor problem* chapter X.2.1 where the term $k \cdot \omega$ for $k \in \mathbb{Z}^d \setminus \{0\}$ arises due to a Fourier expansion [HLW06]. Intuitively, the non-resonance condition may be interpreted as follows: On the unperturbed, completely integrable Hamiltonian system with Hamiltonian $H_0(a)$ we have seen that phase space may consist (locally in a ball $B \subset \mathbb{R}^n$) of invariant tori i.e. of the form $(a, \phi) \in B \times \mathbb{T}^n$ with flow defined by the ODE

$$\dot{a} = 0, \quad \dot{\phi} = \omega(a) = D_a H_0(a).$$

On a specific torus with action variable $a_* \in B$ the flow is *conditionally periodic*: It is periodic if $|k \cdot \omega(a_*)| = 0$ for some $k \in \mathbb{Z}$ and *quasi-periodic* (e.g. [Arn89]) otherwise (quasi-periodic motion is not to be confused with a quasi-periodic function of Definition 3.1). In the quasi-periodic case it is known that motion is aperiodic and fills densely the torus ([HLW06] chapter X.1.4). Now, KAM theory tells us that, if the motion is periodic (the frequencies are resonant) or nearly periodic, then this invariant tori with action variable a_* will be destroyed under the influence of a Hamiltonian perturbation $H_1(a, \phi)$. Only when the motion is very aperiodic (strongly non-resonant frequencies i.e. (σ, γ)-Diophantine) then the invariant torus persists.

If $\sigma > n - 1$ then a large measure of ω in any fixed ball in $\mathbb{R}^n$ are (σ, γ)-Diophantine for some $\gamma > 0$, see [HLW06] chapter X.2.1. In the case $n = 2$ this is illustrated on the cover of this thesis in the figure taken from [Han11]. Thus, one may expect a large number of invariant tori to persist, but this depends on the strength of the perturbation.

3.2.1 A KAM theorem for time-dependent, periodic perturbations

Usually, KAM theorems are given for autonomous, completely integrable Hamiltonian ODE $H_0(a)$, with autonomous perturbations $H_1(a, \phi)$. If the perturbation is $(2\pi$-)periodic $H_1(a, \phi, t + 2\pi) = H_1(a, \phi, t)$ or quasi-periodic (Definition 3.1), then the KAM theorems can be adapted to this case with almost identical proof [JS96]. These adapted KAM theorems are given, for example, in [Jor91; JS96; BSG03; TZ09], and generalised in [JV97] (for lower-dimensional tori) and a very general setting in [Sev07] (for non-Hamiltonian, parameterised analytic families of lower-dimensional tori and under the weaker Rüsmann non-degeneracy conditions).

When non-autonomously perturbed, Kolmogorov's non-degeneracy condition is in general not valid on the canonically extended phase space [Moa03] and one may consider for example other non-degeneracy conditions, such as the isoenergetic one [Moa03]. However, one may notice that only the variables in the unperturbed system need to satisfy the Kolmogorov non-degeneracy condition [JS96][11].

Thus, we now consider a perturbed Hamiltonian H on $D \subset \mathbb{R}^n \times \mathbb{T}^n \times \mathbb{R}$ in action-angle coordinates (a, ϕ) of the form

$$H(a, \phi, t) = H_0(a) + H_1(a, \phi, t)$$

with H_1 quasi-periodic in t i.e. $\exists \Omega \in \mathbb{R}^s$ for some such that $H_1(a, \phi, t) = K_1(a, \phi, \Omega_1 t, \dots, \Omega_s t)$ with K_1 2π-periodic in all of its last $s \in \mathbb{N}$ variables. The KAM theorem we will use (we do not aim for the most general statement) is from [Jor91; JS96] together with an addition from [Sev07] (for which the conditions are more general so that we may state the conditions of [Jor91; JS96]). Due to quasi-periodicity, s new angle variables $\psi \in \mathbb{T}^s$ together with a conjugate variable $b \in \mathbb{R}^s$ are introduced, where $\psi = \Omega t$ so that the extended Hamiltonian K is equal to

$$K(a, b, \phi, \psi) = H_0(a) + K_1(a, \phi, \psi) + \Omega \cdot b.$$

Thus, in the case $s > 1$ the amount of variables in the perturbation increase so that the perturbation K_1 becomes periodic in its last s variables.

Theorem 3.9 ([Jor91] theorem 2.3, [JS96] theorem 3 and [Sev07] theorem 3.1)**.** *Consider the perturbed Hamiltonian*

$$K(a, b, \phi, \psi) = H_0(a) + \mu K_1(a, \phi, \psi) + \Omega \cdot b$$

defined on an open subset $D \times \tilde{D} \times \mathbb{T}^n \times \mathbb{T}^s \subset \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{T}^n \times \mathbb{T}^s$. Denote $A = (a, b)$, $\Phi = (\phi, \psi)$.

If, for some fixed $\rho > 0$, the Hamiltonian K can be extended analytically to the complex domain $F := F_\rho := \{(A, \Phi) \mid A \in G,\ \Phi \in \mathbb{T}^s \times i[-\rho, \rho]^s$ i.e. $|im(\Phi_i)| \leq \rho\}$ in a way such that K is 2π-periodic in the variables Φ_i $(i = 1 \dots n+s)$, where $G_1 \subset D, G_2 \subset \tilde{D}$ compact and $G := G_1 \times G_2 \subset \mathbb{R}^n \times \mathbb{R}^s$.
If furthermore H_0 satisfies the non-degeneracy condition on F

$$\det \frac{\partial^2 H_0(a)}{\partial a^2} = \det \frac{\partial \omega(a)}{\partial a} \neq 0,$$

where $\omega(a) = \frac{\partial H_0(a)}{\partial a}$ are the frequencies of the unperturbed system and if the frequencies Ω of the perturbation are (σ, γ)-Diophantine, Equation (3.3) i.e.

$$\langle k, \Omega \rangle \geq \frac{\gamma}{\|k\|_1^\sigma}, \quad \forall k \in \mathbb{Z}^s - \{0\}.$$

with the requirement $\sigma > s - 1, \gamma > 0$ (where $\|\cdot\|_1$ is the 1-norm).

Then $\forall \epsilon > 0$, $\exists C = C(\epsilon, \rho, G, H_0)$ such that, if

$$\mu \leq C, \text{ pointwise on } F$$

then the motion defined by

$$\dot{\Phi} = D_A K(A, \Phi) \qquad \dot{A} = -D_\Phi K(A, \Phi)$$

has the following properties:

- *Re $F = F_1 + F_2$ where F_1 is invariant and the measure of F_2 (denoted $|F_2|$) satisfies $|F_2| \leq \epsilon |F|$.*

[11] In the periodic case, another proof of the periodic KAM theorem exists. Using a trick (considering the exponential of the canonically extended Hamiltonian) to remove the non-degeneracy, an application of the KAM theorem on extended phase space *is* possible in this case, see [TZ09] chapter 2.1 and 2.2.

- *F_1 consists of a family of invariant $(n+s)$-dimensional analytic tori $I_\alpha = \{(B_\alpha, \Psi) \mid \Psi \in \mathbb{T}^{n+s}\}$, defined parametrically by*
$$A = B_\alpha + f_\alpha(\Psi), \qquad \Phi = \Psi + g_\alpha(\Psi),$$
where $(f_\alpha, g_\alpha) : \mathbb{T}^{n+s} \to \mathbb{R}^n$ are analytic and of period 2π in all its variables and the parametrisation depends on $\alpha = (\omega, \Omega)$ [JS96] (because of the non-degeneracy condition, equivalently $\alpha = (a, \Omega)$ as in [Sev07]).

- *The mapping is close to the identity*
$$|(f_\alpha, g_\alpha)| = \mathcal{O}(\epsilon) \text{ pointwise on } F$$

- *On the invariant $(n+s)$-tori I_α, the motion of the perturbed system is quasi-periodic with frequencies*[12]
$$\left(\frac{\partial H_0(B_\alpha)}{\partial a}, \Omega\right).$$

- *Suppose moreover that the perturbation H_1 can be holomorphically continued (is complex analytic) on the set*
$$\tilde{F}_\rho = \{(A, \Phi) \in \mathbb{C}^{n+s} \times (\mathbb{C}/2\pi\mathbb{Z})^{n+s} \,|\, A \in G \subset \mathbb{C}^n, \, |im(\Phi_i)| \leq \rho\}$$
for some $G \subset D \times \tilde{D}$ such that the closure $\bar{G}$ is contained in $D \times \tilde{D}$ (the only stronger condition in [Sev07]). Then, given a fixed α, the coordinate transformation mapping $(A, \Phi) \mapsto (B, \Psi)$ (which is given by $(Id + (f_\alpha, g_\alpha))^{-1}$) leaves the newly introduced angle variables ψ unchanged
$$\psi = \bar{\psi}$$
and transforms the ODE into the form
$$\begin{aligned} \dot{B} &= \mathcal{O}\left(\|B - B_\alpha\|_1^2\right) \\ \dot{\bar{\phi}} &= \omega' + \mathcal{O}\left(\|B - B_\alpha\|\right), \\ \dot{\bar{\psi}} &= \Omega. \end{aligned} \tag{3.4}$$
where B_α is the transformed coordinate parametrising the torus for the chosen α.

Thus, this theorem tells us that most invariant tori persist (all tori in F_1), if the perturbation has frequencies Ω which are strongly non-resonant ((σ, γ)-Diophantine) and satisfies the Kolmogorov non-degeneracy condition in the action variables of the unperturbed system. In particular we mentioned above that the tori with (almost) periodic motion were destroyed for autonomous perturbations, again this is the case for non-autonomous perturbations: “The tori whose frequencies are in resonance with the ones of the perturbations are destroyed” [JS96]. The surviving tori are, with increasing perturbation strength, the ones which are increasingly non-resonant, quantified by the (strong) Diophantine property, just like in the case of an autonomous perturbation. Furthermore, the mapping to the new coordinates is close to the identity.

Remark 3.10. *We consider the case of a 2π-periodic perturbation, $s = 1$. In this case $\psi = t$. Now, does the existence of KAM tori (using the KAM theorem above) in the transformed coordinates imply the existence of invariant tori in the Poincaré map?*

First of all, as stated in Sevryuk [Sev07] (p.567), the torus I_α as in Theorem 3.9, which lives in the space $D \times \tilde{D} \times \mathbb{T}^{n+s}$ and is an invariant torus of the autonomously extendeed Hamiltonian system with Hamiltonian $H_0(a) + \mu K_1(a, \phi, \psi) + s$, may be projected to the space $D \times \tilde{D} \times \mathbb{T}^n$ and so that the projected torus is an invariant torus for the original, non-autonomous Hamiltonian system with Hamiltonian $H_0(a) + \mu K_1(a, \phi, \psi)$. Furthermore, if the perturbation is periodic then we see that time $\psi = \bar{\psi} = t$ is left untransformed, which means that the projected invariant torus induces an invariant torus of the Poincaré map.

[12] The frequencies of the perturbed tori are in some cases equal [ZC10].

3.3 Persistent invariant tori in the tidal wave system

We are now able to describe the tidal wave system in a mathematical framework, action-angle coordinates, which makes rigorous the existence of invariant tori in the unperturbed system and allows for an application of a KAM theorem on the tidal wave systems to prove the existence of persistent invariant tori. This is the goal of this Section.

We use the transformed Hamiltonian of Equation (1.5)

$$L(q,p,t) = k(p-q)\cos(t) + L_1(q,p) + \gamma L_2(q,p,t)$$

where

$$\begin{aligned} L_1(q,p) &= H_1(A^{-1}(q,p)) = -C_\delta k^2 l\big(\alpha\cos(p) + \beta\cos(q)\big) \\ L_2(q,p,t) &= H_2(A^{-1}(q,p),t) = C_\delta 2kl(r\cos(t)+\sin(t))(\alpha\sin(p)+\beta\sin(q)), \end{aligned}$$

which induces the ODE

$$\begin{aligned} \dot{q}(t) &= k\cos(t) + C_\delta kl\big(k\alpha\sin(p) + 2\alpha\cos(p)(r\cos(t)+\sin(t))\big) \\ \dot{p}(t) &= k\cos(t) - C_\delta kl\big(k\beta\sin(q) + 2\beta\cos(q)(r\cos(t)+\sin(t))\big). \end{aligned}$$

3.3.1 Unperturbed system: Integrability and action-angle coordinates

For the unperturbed system we take the autonomous part K of the Hamiltonian L:

$$K(q,p) = -C_\delta k^2 l[\alpha\cos(p) + \beta\cos(q)]$$

where $k,l \geq 0$ and $f,r \in \mathbb{R}$ and $\alpha = fk+rl, \beta = fk-rl$. We rescale time by $C_\delta k^2 l \neq 0$ such that the ODE

$$\dot{q} = \alpha\sin(p), \quad \dot{p} = \beta\sin(q)$$

is obtained. One can remark that this is a special case of the ABC flow, so that symmetries (and reversible KAM theory) become very important. However, in this thesis this will not be pursued. Note here that $\alpha,\beta \in \mathbb{R}$ may take on value since, for fixed k,l, the map $(r,f) \mapsto (\alpha,\beta) = (fk+rl, fk-rl)$ is linear and surjective. The case $\alpha\beta = 0$ is not interesting, however, so we will assume that $\alpha,\beta \in \mathbb{R}\setminus\{0\}$.

Considering the further time-scaling with factor α and using the extended symplectic transformations $(q,p) \mapsto (-p,q)$ and $(q,p) \to (p,q)$ allows us to assume without loss of generality that $\eta := \alpha/\beta \in [-1,0)$. So we consider

$$\dot{q} = \sin(p), \quad \dot{p}(t) = \eta\sin(q), \quad \text{with Hamiltonian} \quad \tilde{K}(q,p) = -\cos(p) + \eta\cos(q), \quad \text{for } \eta \in [-1,0). \tag{3.5}$$

We note that the solutions of this ODE are defined on the whole time domain, since the vector field is bounded. Thus, one of the conditions of complete integrability is satisfied. Moreover, since the system is planar and autonomous, the Hamiltonian K (and the transformed Hamiltonian $\tilde{K}$) is an invariant of motion (in involution with itself. Thus, we may immediately conclude that, on sets where $\nabla K \neq 0$ (since $\tilde{K}$ is a symplectic transformation of K, equivalently $\nabla\tilde{K} \neq 0$), the **system is completely integrable**.

Before transforming to action angle coordinates, we first sketch a portrait of the phase plane. The fixed points of the ODE occur in a rectangular grid with gridpoints $q,p \in \pi\mathbb{Z}$ with energy $\tilde{K}(q,p) = \pm 1 \pm \eta$. These fixed points are, alternating in the grid/in a checkerboard pattern, saddle ($\cos(q)\cos(p) < 0$) and elliptic ($\cos(q)\cos(p) > 0$) fixed points (Figure 11):

$$J\nabla^2\tilde{K} = \begin{pmatrix} 0 & \pm 1 \\ \pm\eta & 0 \end{pmatrix} \quad \text{at the fixed points}$$

with eigenvalues $\lambda = \sqrt{\pm\eta}$. Thus, the assumption $\eta < 0$ causes the origin to be an elliptic fixed point (and the alternating stability type on the grid determines the rest). For the saddle points the eigenvectors

$(1, \sqrt{-\eta}), (1, -\sqrt{-\eta})$ can be found.

Notice that $\tilde{K} : \mathbb{R}^2 \to [-1+\eta, 1-\eta]$. The maxima and minima of the Hamiltonian are precisely the elliptic points at $(q,p) = (m\pi, n\pi)$ for $n, m \in \mathbb{Z}$, $n - m \in 2\mathbb{Z}$. The saddle points $\{(q,p) \mid q = m\pi, p = n\pi,\ m,n \in \mathbb{Z},\ m-n \in 2\mathbb{Z}+1\}$ are contained in the set of maxima/minima of $\tilde{K}$, $\{(q,p) \in \mathbb{R}^2 \mid \tilde{K}(q,p) = \pm(1+\eta)\}$, which consists of the continuous curves

$$p_{m,\pm}(q) = 2m\pi \pm \arccos(-1 - \eta(1-\cos(q))) \qquad \text{if } \tilde{K} = 1+\eta,\ \ \eta \in (-1,0) \tag{3.6}$$

$$\tilde{p}_{m,\pm}(q) = 2m\pi \pm \arccos(1 + \eta(1+\cos(q))) \qquad \text{if } \tilde{K} = -1-\eta,\ \eta \in (-1,0) \tag{3.7}$$

$$p_{m,\pm}(q) = m\pi \pm q \qquad \text{if } \eta = -1, \tag{3.8}$$

where $m \in \mathbb{Z}$ and $\arccos : [-1,1] \to [0, 2\pi)$. Using this one can show, for $\eta \in (-1,0)$, that

- Every saddle point lies on a curve $p_{m,\pm}, \tilde{p}_{m,\pm}$ with $m \in \mathbb{Z}$. Two saddle points with equal coordinate q lie on the same curve $p_{m,\pm}$ or $\tilde{p}_{m,\pm}$ for some fixed $m \in \mathbb{Z}$.
- Conversely, the curves $p_{m,\pm}, \tilde{p}_{m,\pm}$ consist completely of the saddle points together with the heteroclinic orbits of these saddle points (and form the unstable and stable manifolds).
- The curves $p_{m,\pm}, \tilde{p}_{m,\pm}$ for fixed $m \in \mathbb{Z}$ form pairs of heteroclinic orbits, which partition phase space into interior regions, containing elliptic fixed points and bounded orbits, and exterior regions.
- For $\eta \neq -1$ these unbounded orbits exist in these exterior regions.

These are illustrated in Figure 11, where the symmetry of phase space due to the invariance of the ODE under the map $(q,p) \to (q + 2\pi n, p + 2\pi m)$ is evident.

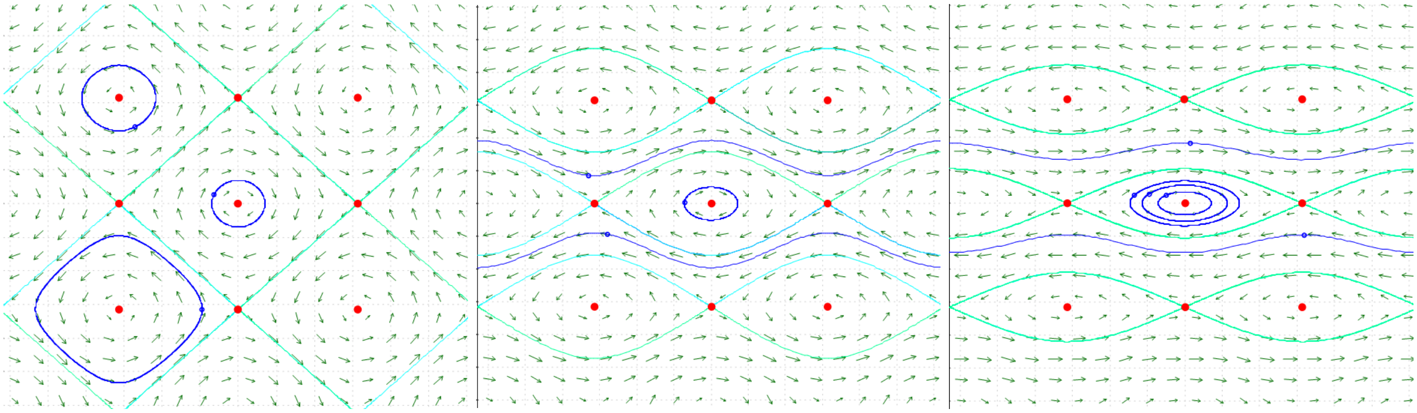


Figure 11: The phase plane (q,p) for different values of $\eta \in \{-1, -\frac{1}{2}, -\frac{1}{4}\}$ (from left to right). We see the periodicity of the phase plane in (q,p) with period π. For $\eta \neq -1$ we see also unbouned orbits in blue. Equilibrium points are shown in red and stable/unstable orbits of the saddle points in cyan. Image made in *PPlane* [Cas05].

For $\eta = -1$ similar statements hold, but this case is special in the sense that all unstable fixed points have the same energy $\tilde{K} = 0$. The heteroclinic orbits are lines and the bounded domains are rotated squares. No two heteroclinic orbits meet at the same two saddle points, however, and no unbounded orbits exist (Figure 11 left).

Understanding the phase space, we now show that locally, phase space consists of invariant circles (action-angle coordinates are found). Because of the symmetry arguments discussed above, we focus on a single 'cell' of the grid around the elliptic fixed point $(q,p) = (0,0)$, as in Figure 12. In other words we look at the set $B := \{(q,p) \in \mathbb{R}^2 \mid \tilde{K}(q,p) < -1-\eta,\ q \in (-\pi,\pi),\ p \in (-\pi,\pi)\} = \{(q,p) \in \mathbb{R}^2 \mid q \in (-\pi,\pi),\ p_{0,-}(q) < p <$

$p_{0,+}\}$. This space is bounded, open and, as shown above, contains a single elliptic fixed point at the origin. Therefore the orbits are compact and we may find action-angle coordinates on $B \setminus \{0\}$ (at 0, $\nabla\tilde{K} = 0$ so the conditions of the Arnold-Liouville theorem are not satisfied).

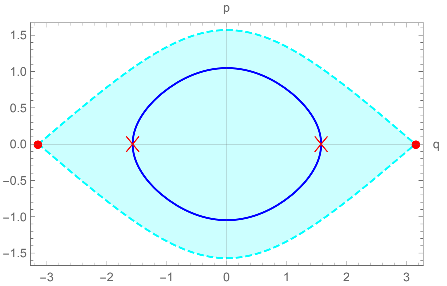


Figure 12: A single 'cell' of the system of Equations (3.5) for $\eta = -0.5$. Plotted are the heteroclinic orbits $p_{0,\pm}$ (dashed cyan lines), saddle points (red dots), the set B (the interior region bounded by $p_{0,\pm}$) (cyan shade) and an example orbit (blue line, with Hamiltonian value -1), with q_0 and q_1 (red crosses).

From Section 3.1.3 the symplectic map $(a, \phi) : B \setminus \{0\} \to \mathbb{R} \times \mathbb{T}$ defining the action-angle variables (using Equation (3.2)) is given by

$$a(q,p) = \frac{1}{2\pi}\Pi(q,p) = \frac{1}{\pi}\int_{q_0}^{q_1} \arccos\Big(-\tilde{K}(q,p) + \eta\cos(\tau)\Big)\, d\tau, \qquad \phi(q,p) = 2\pi\frac{y_1(q,p)}{T(q,p)}, \tag{3.9}$$

which is well-defined since $\tilde{K} \in (-1+\eta, 1-\eta)$. Here $q_0(q,p), q_1(q,p) \in (-\pi, \pi)$ are such that $\tilde{K}(q,p) - \eta\cos(q_i) = \pm 1$ and $q_0 < q_1 = -q_0$, see also Figure 12. For example, for $(q,p) = (\pi/2, 0)$ we find $-q_0 = q_1 = \pi/2$. The period $T(q,p) = T(\phi_t(q,p))$ ($t \in \mathbb{R}$) and the phase $y_1(q,p)$ take the form

$$y_1(q,p) = \pm\int_{q_0}^{q} \frac{1}{\sqrt{1 - (\eta\cos(\tau) - \tilde{K}(q,p))^2}}\, d\tau, \qquad T(q,p) = y_1(q_1(q,p), 0) > 0,$$

where the sign of y_1 is non-negative if and only if $p \geq 0$. Furthermore $y_1(q,p)$ depends on the path of integration and is multivalued, but well-defined if taken modulo $T(q,p)$.

3.3.2 Perturbed system: KAM theory for the tidal-wave system

We now apply periodic KAM theory (Theorem 3.9) to periodic perturbations of the system of Section 3.3.1 and in particular to the "default" and "simple-B" case (Remark 1.1). In [Wal21] these parameter sets are defined without vorticity term ($\gamma = 0$) i.e. with Hamiltonian

$$\mathcal{L}(q,p,t) = k(p-q)\cos(t) - \mu C_\delta k^2 l[\alpha\cos(p) + \beta\cos(q)] \tag{3.10}$$

where $\alpha = fk + rl$, $\beta = fk - rl$. The first step is a time-rescaling $t \mapsto \left(C_\delta \alpha k^2 l\right)^{-1} t = \chi t$ of the ODE, transforming the Hamiltonian into the form

$$\tilde{\mathcal{L}}(q,p,t) = -\cos(p) + \eta\cos(q) + \mu\chi(p-q)\cos(\chi t) \tag{3.11}$$

where we assume without loss of generality that $\eta = \eta(k,l) = \beta/\alpha \in [-1, 0)$ (see Section 3.3.1). Here $\mu \in \mathbb{R}$ is an artificially introduced perturbation parameter. The unperturbed Hamiltonian $\tilde{\mathcal{L}}$ (with $\mu = 0$) is equal to the system $\tilde{K}(q,p) = -\cos(p) + \eta\cos(q)$ (similarly $\mathcal{L}$ is equal to K for $\mu = 0$).
Three of the four conditions of the KAM Theorem 3.9 (using notation thereof) are easily verified to be true:

- The unperturbed system ($\mu = 0$) is completely integrable, which follows from the previous section (together with the fact that $\tilde{K}$ and K are related by extended symplectic transformations).

- The Hamiltonian $\tilde{\mathcal{L}}$ is complex analytic on the complex domain $F = F_\rho$ for any $\rho > 0$, since cos, sin are complex analytic on the whole of $\mathbb{C}$. Moreover $\tilde{K}$ is analytic (Remark 3.6) so it may be locally extended to a complex analytic function. One may check using a Taylor series on T, y_1 that they define analytic functions (globally) on $\mathbb{C}$.

- In any case the perturbation is periodic, therefore $\Omega = 2\pi$ is trivially (σ, γ)-Diophantine for every $\sigma > 0$ and sufficiently large $\gamma > 0$.

This only leaves the Kolmogorov non-degeneracy condition for the unperturbed Hamiltonian $\tilde{K}$ (or equivalently for K) written in action angle coordinates (denoted again by $\tilde{K}$)

$$\frac{d^2\tilde{K}(a)}{da^2} = \frac{\partial}{\partial a}\frac{2\pi}{T((q(a,\phi), p(a,\phi)))} = \frac{\partial}{\partial a}\frac{2\pi}{T(a)} = -\frac{2\pi}{T(a)^2}\frac{\partial T(a)}{\partial a} \neq 0,$$

which is checked next. It was used that the period $T((q(a,\phi),p(a,\phi)))$ only depends on the action a (since the frecuency ω only depends on a and $\omega(a)=\frac{2\pi}{T(a)}$), abusing notation we therefore write $T=T(a)$. In particular, we have $T(q,p)=T(q_0(q,p),0)$ and q_0 depends on the action a only, such that $q_0=q_0(a)$. Therefore we consider $T(q_0(a))$ (with abuse of notation) and check Kolomogorov's theorem using this function.

In the following, we make the assumption that integrals and derivatives without may be switched. This is not straightforward, as the integrals are improper. First we prove that $q_0(q,p)$ depends only on the action ($q_0=q_0(a)$), which follows from the inverse function theorem and the fact that $a'(q_0)\neq 0$:

$$a(q_0)=\frac{1}{\pi}\int_{q_0}^{q_1}\arccos\left(-\tilde{K}(q_0,0)+\eta\cos(\tau)\right)\,d\tau,\quad \text{so}\quad a'(q_0)=0+\frac{1}{\pi}\int_{q_0}^{q_1}\frac{D_q\tilde{K}(q_0,0)}{\sqrt{1-\left(-\tilde{K}(q_0,0)+\eta\cos(\tau)\right)^2}}=\frac{\eta\sin(q_0)}{2\pi}T(q_0,0)<0,$$

where it was used that $D_q\tilde{K}(q_0,0)=\eta\sin(q_0)<0$. Here we are allowed to switch (impoper) integral and derivative, because the integral on the right hand side is equal to $D_q\tilde{K}(q_0,0)\frac{1}{2\pi}T(q_0(a))$ which is well-defined (thus we may use the dominated convergence theorem with the dominant function simply the integrand itself).

Next we find that $q_0<\tilde{q}_0$ implies that $T(q_0)>T(\tilde{q}_0)$. Since ω is differentiable ($\tilde{K}$ is analytic and Remark 3.6) we find $\frac{dT(q_0)}{dq_0}>0$. Therefore

$$\frac{dT(q_0(a))}{da}=\frac{dT(q_0(a))}{dq_0}\frac{dq_0(a)}{da}>0,$$

for all actions $a(q,p)$ with $(q,p)\in B$. Intuitively one might expect that $\frac{dT(q_0(a))}{da}>0$ since increasing a means that we are approaching the heteroclinic orbits. This proves the Kolmogorov non-degeneracy condition.

3.3.3 The "default" case

In the default case, $k=l$ and $\gamma=0$ (without vortivity), the rescaled system of Equation (3.11) has Hamiltonian

$$-\cos(p)+\eta\cos(q)+\mu k\chi_1(p-q)\cos(\chi_1 t).$$

here $\mu\in\mathbb{R}$ is introduced as a perturbation parameter and $\chi_1:=\chi_1(k,r,f):=\frac{20(k^2+2r^2+2)}{3k}\to\infty$ as $k\to 0$ or $k\to\infty$. Varying k not only changes the strength of the perturbation, but also the frequency, see Figure 13.

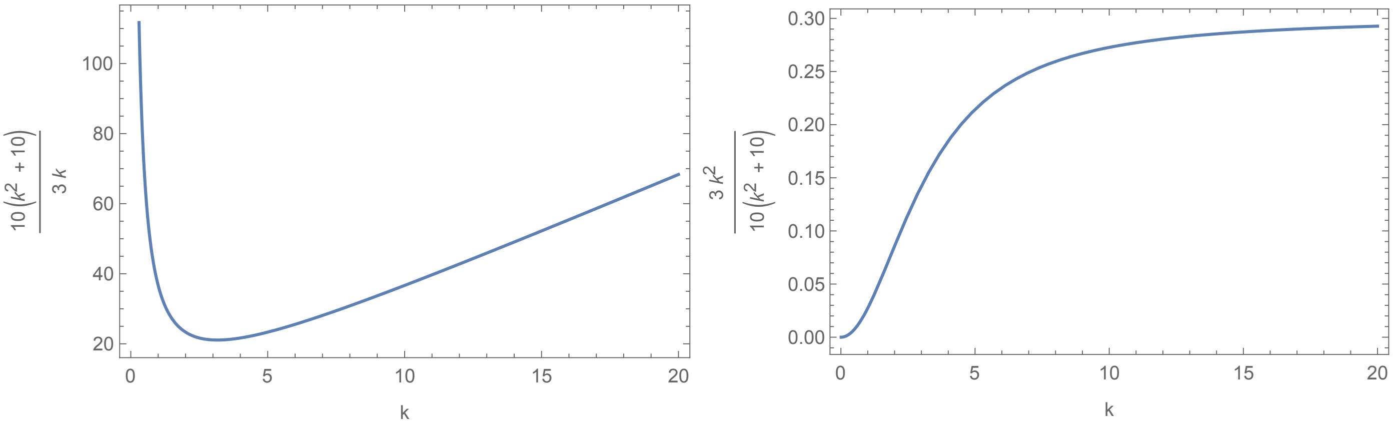


Figure 13: The amplitude $\frac{10(k^2+10)}{3k}$ and the (scaled) period $\frac{3k^2}{10(k^2+10)}$ of the perturbation, for $r=2$, $f=-1$. (Made in *Wolfram Mathematica, 10.2*).

Since the transformation into action-angle coordinates, of Equation (3.9) is symplectic and time-independent, one may simply substitute $(q,p) \to (q(a,\phi), p(a,\phi))$ such that the Hamiltonian becomes,

$$\mathcal{L}(a,\phi,t) = \underbrace{-\cos(p(a,\phi)) + \eta\cos(q(a,\phi))}_{H_0(a)} + \mu \underbrace{k\chi_1\cos(\chi_1 t)H_1(q(a,\phi),p(a,\phi))}_{H_1(a,\phi,t)}$$

defined on $D \times \mathbb{T}^2$.

Theorem 3.9 now states that for every $\mu > 0$, there exists $C = C(\mu, \rho, G, \gamma, \tau, H_0)$ such that, if $\mu \leq C$ on F_ρ (for a fixed $\rho > 0$), then there exists a symplectic mapping, near to the identity, to a coordinate system which define a family of invariant tori in a subset of $D \times \mathbb{T}^2$. From Remark 3.10 we may also conclude that there exist persistent invariant tori of the Poincaré map.

Finally, as mentioned in the Introduction 1, one could, for a particular persistent torus, try to find the value where it breaks, such as in [CFP87a], but this is not pursued in this thesis.

3.3.4 The "Simple-B" case

In the "Simple-B" case, with $k = l$ and $\gamma = 0$ (without vorticity), the time-rescaled Hamiltonian of Equation (3.11) takes the form

$$\mathcal{H}(a,\phi,t) = -\cos(p(a,\phi)) + \eta\cos(q(a,\phi))\mu k\chi_2(p(a,\phi) - q(a,\phi))\cos(\chi_2 t),$$

where $\chi_2 = \chi_2(k,r,f) := \frac{5r}{3(f+r)k}$. This time, varying k only varies the period and not the strength of the perturbation. We may use the same arguments as in the "default" setting and Remark 3.10 to prove that, for small μ, KAM tori exist in the periodic Poincaré map of the perturbed system.

We see furthermore, that the rescaled "default" and "Simple-B" systems, without vorticity term ($\gamma = 0$) and $k = l$, are easily related: The "default" system with parameters (k, μ) is equal to the 'Simple-B' system with parameter $(\frac{kr}{4(k^2+2r^2+2)}, \mu\frac{4(k^2+2r^2+2)}{kr})$. In [Wal21], $r = 2$ is used so we expect the Simple-B case to have less KAM tori (see also Section 4.4 and Figure 15).

4 Symplectic methods, induced methods for forced ODE and splitting methods for time-affine systems

In this section four topics in numerical methods are discussed. The focus is again on numerical (splitting) methods for non-autonomous Hamiltonian ODE, in particular forced and time-affine ODE (Definitions 2.1 and 2.2). We also consider briefly splitting methods for non-autonomous ODE and for near-integrable Hamiltonian systems.

The first topic is a general presentation of (partitioned) Runge-Kutta ((P)RK) methods, splitting and composition methods and the adjoint method to solve (non-autonomous) ODE numerically. We present also a particular (numerical) method introduced in [Wal21]. This method is adapted to forced ODE and we call it the *induced method*. Second, we consider structure preserving methods, in particular symplectic methods and we show that the induced method is symplectic. Third, we briefly give an outlook on structure preserving numerical methods for non-autonomous Hamiltonian ODE which is a follow-up of Section 2.4. Fourth, the splitting for the tidal wave system (time-affine) in [Wal21] is presented. This method splits the time-affine Hamiltonian of the tidal wave system into 'forced; Hamiltonians on which induced methods are used.

4.1 Induced methods and splitting methods for time-affine ODE

We will first define (partitioned) Runge-kutta and splitting methods for autonomous and non-autonomous ODE. For forced ODE the induced method is discussed, which can be combined with the splitting method to form numerical methods for time-affine systems. Finally, we discuss how splitting methods are sometimes adapted to non-autonomous systems.

4.1.1 Partitioned Runge-Kutta, adjoint, composition and splitting methods

We consider (partitioned) Runge-Kutta ((P)RK) methods, splitting methods, the adjoint method and composition methods.

For an ODE with vector field f, an *s-stage Runge-Kutta method* with step size $h > 0$ is a map $\psi_h : \mathbb{R}^{n+1} \to \mathbb{R}^n$, which can be constructed given a Butcher tableau $\mathcal{B}$ of coefficients (consisting of a matrix $A \in \mathbb{R}^{s \times s}$ (with elements $a_{i,j}$ and two vectors $b, c \in \mathbb{R}^{s}$[13]). The Butcher tableau is usually written as [HLW06]

$$\mathcal{B} \quad = \quad \begin{array}{c|ccc} c_1 & a_{1,1} & \dots & a_{1,s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{1,s} & \dots & a_{s,s} \\ \hline & b_1 & \dots & b_s \end{array} \; .$$

We may write $\mathcal{B} = (A, b, c)$ to make the notation of the coefficients explicit. The RK method is then defined as the map $\psi_h : \mathbb{R}^{n+1} \to \mathbb{R}^n$, given by

$$\psi_h(y,t) = y + h \sum_{i=1}^{s} b_i k_i, \qquad k_i = f\left(y + \sum_{i=1}^{s} a_{ij} k_j, t + h \sum_{i=1}^{s} c_i\right).$$

RK methods for autonomous ODE are a special case in the sense that the values of c_i can be omitted (and may be considered as maps $\psi_h : \mathbb{R}^n \to \mathbb{R}^n$).

Example 4.1. *The explicit and implicit Euler method have Butcher tableaus ([HLW06] chapter II.1)*

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array} \quad \textit{respectively} \quad \begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}.$$

Definition 4.2. The *local error* of a numerical method $\psi_h : \mathbb{R}^n \to \mathbb{R}^n$ applied to an ODE with flow ϕ is (for the Euclidean norm $\|\cdot\|_2$) given by $\psi_h - \phi_h$ and is of order $p \in \mathbb{N}$ if

$$\|\psi_h(y) - \phi_h(y)\|_2 = \mathcal{O}(h^{p+1}) \text{ as } h \to 0,$$

for all $y \in \mathbb{R}^n$ and all smooth f. A method is consistent if it is of order 1.
The *global error after m steps* is given by

$$\sup_{1 \le j \le m} \left\| \psi_h^j(y) - \phi_h(y) \right\|_2 . \tag{4.1}$$

$\varnothing$

A RK method can be guaranteed to be consistent and of higher order if there are restrictions to the parameters a, b (and c) in the Butcher tableau e.g. as in [HLW06] chapter II and III.

Partitioned RK methods form a bigger class of numerical methods. A partitioned RK (PRK) method splits the variable y into multiple parts, e.g. $y = (z, w)$ and use different Butcher tableaus for these variables. Suppose we have an ODE with vector field $\hat{f} = (f, g)$ where $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$ and $g : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n-m}$ then we can use two Butcher tableaus $\mathcal{B}$, $\tilde{\mathcal{B}}$ which describe s-stage (the number of stages must be the same) RK methods to form the PRK method $\phi = (\phi_{[1]}, \phi_{[2]})$ given by

$$\phi_{[1],h}(y,z,t) = y + h \sum_{i=1}^{s} b_i k_i, \qquad k_i = f\left(y + h \sum_{i=1}^{s} a_{ij} k_j,\; z + h \sum_{i=1}^{\tilde{s}} \hat{a}_{ij} \ell_j,\; t + h c_i\right),$$

$$\phi_{[2],h}(y,z,t) = z + h \sum_{i=1}^{\tilde{s}} \hat{b}_i \ell_i, \qquad \ell_i = g\left(y + h \sum_{i=1}^{s} a_{ij} k_j,\; z + h \sum_{i=1}^{\tilde{s}} \hat{a}_{ij} \ell_j,\; t + h \tilde{c}_i\right).$$

[13] Sometimes complex coefficients are considered [Sey16]

The order conditions can be found in [HLW06] chapter II and III. A simple example of a PRK method is a non-autonomous system in autonomously extended space (Definition 2.5) where one splits time $y = (z, t)$ and the vector field $\hat{f} = (f, 1)$ (where 1 is the constant function with value 1).

Remark 4.3. *If $f(y,t)$ is (real-)analytic in $y \in \mathbb{R}^n$ and $h > 0$ is sufficiently small so that the conditions of the analytic implicit function theorem hold, then a Taylor expansion of (partitioned) RK methods exists, such that a numerical method $\psi_h : \mathbb{R}^n \to \mathbb{R}^n$ method can be given by a power series in h*

$$\psi_h(y, t_0) = \sum_{i=0}^{\infty} \frac{h^i}{i!} d_i(y, t_0),$$

For consistent methods $d_0(y,t) = y$ and $d_1(y,t) = f(y,t)$. The coefficient function d_i are analytic for sufficiently small y ([HLW06] Section IX.7.1).

Next, we discuss the *adjoint (numerical) method*, useful for composition methods and to construct *symmetric integrators* which respect reversible symmetries (Section 4.2 or [HLW06] chapter V).

Definition 4.4. The adjoint method ψ^* of a method $\psi_h : \mathbb{R}^n \to \mathbb{R}^n$ is given by $\psi^*_h = (\psi_{-h})^{-1}$[14] i.e. the inverse of the map ψ_{-h}. $\varnothing$

Given numbers $\alpha_1, \dots, \alpha_s$, $\beta_1, \dots, \beta_s$ with $\sum_i (\alpha_i + \beta_i) = 1$ and a numerical method $\psi_h : \mathbb{R}^n \to \mathbb{R}^n$, the *composition method* Ψ_h with step sizes $\alpha_1 h, \beta_1 h \dots, \alpha_s h, \beta_s h$ is given by

$$\Psi_h = \psi_{\alpha_s h} \circ \psi^*_{\beta_s h} \cdots \circ \psi_{\alpha_1 h} \circ \psi^*_{\beta_1 h}.$$

One can construct composition methods with higher order than the original method, [HLW06] chapter II and III. For example, if ψ is of order 1 then $\psi^*_{h/2} \circ \psi_{h/2}$ is of order 2.

Finally we consider *splitting methods*. There are three steps to splitting methods

1. Split the vector field $\sum_i f_i$.
2. Integrate the ODE with vector field f_i in some way with a method ψ_{f_i}.
3. Use a composition method to obtain a good approximation of ϕ_f.

The main idea is that the flows of f_i are either 'simpler' or easier to integrate numerically [MQ02; BCM08].

Given a vector field f, we split $f = f_1 + f_2$. If we use two numerical methods $\psi_{[1],h}, \psi_{[2],h} : \mathbb{R}^n \to \mathbb{R}^n$ to solve these methods then we may consider the numerical method

$$\Psi_h = \psi_{[2],h} \circ \psi_{[1],h} \qquad \Psi^*_h = \psi^*_{[1],h} \circ \psi^*_{[2],h}$$

which is called the *Lie-Trotter splitting* and its adjoint [HLW06], which are methods of order 1. One may also consider

$$\Psi_h = \psi^*_{[1],h} \circ \psi^*_{[2],h/2} \circ \psi_{[2],h/2} \circ \psi_{[1],h/2}$$

which is called the *Strang* or *Marchuk splitting* [HLW06] (compare with the composition method defined above). Order conditions (using B-series, P-series or Lie algebra techniques) to produce methods of higher order can be found in [HLW06] chapter II and III. We mention in Section 4.1.3 how splitting methods can be adapted so as to become more suitable to the non-autonomous case.

4.1.2 Induced method for forced ODE

[14] For C^1 vector fields and small $h > 0$, the inverse ψ^{-1}_{-h} exists locally by the inverse function theorem.

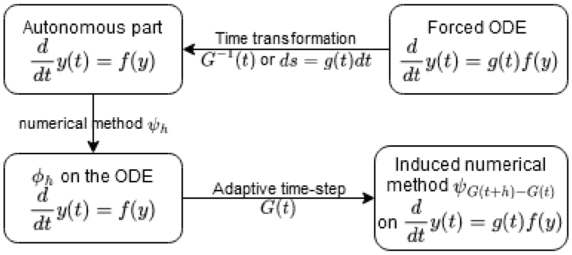

Autonomous part
$\frac{d}{dt}y(t) = f(y)$
Time transformation
$G^{-1}(t)$ or $ds = g(t)dt$
Forced ODE
$\frac{d}{dt}y(t) = g(t)f(y)$
numerical method ψ_h
ϕ_h on the ODE
$\frac{d}{dt}y(t) = f(y)$
Adaptive time-step
$G(t)$
Induced numerical method $\psi_{G(t+h)-G(t)}$
on $\frac{d}{dt}y(t) = g(t)f(y)$

Figure 14: The idea behind the induced method for forced ODE.

If ψ_h is a method for the ODE with autonomous vector field f, then for any time-dependent $g \in C^0(\mathbb{R})$ there exists a naturally induced method $\tilde{\psi}_h$ for the forced ODE with vector field $g(t)f(y)$. Indeed, from Equation (2.5), the solution $\tilde{\phi}_{t,t_0}$ of the non-autonomous ODE with vector field $g(t)f(y)$ satisfies

$$\tilde{\phi}_{t,t_0} = \phi_{G(t)-G(t_0),0} = \phi_{G(t),G(t_0)},$$

where $G(t) = G_\tau(t) = \int_\tau^t g(s)\, ds$. Therefore, heuristically (e.g. Figure 14) we expect for one timestep (at time t_0)

$$\psi_h \approx \phi_h = \phi_{t_0+h,t_0} \qquad \Longrightarrow \qquad \psi_{\eta(t_0,h)} \approx \phi_{G(t_0+h),G(t_0)} = \tilde{\phi}_{t_0+h,t_0},$$

where $\eta = \eta(t_0, h) = G(t_0 + h) - G(t_0)$. Therefore, $\psi_{\eta(t_0,h)} \approx \tilde{\phi}_{t_0+h,t_0}$ so it defines a numerical method for the forced ODE.

Definition 4.5. The *(naturally) induced (numerical) method (of gf) (from ψ)* is defined as

$$\tilde{\psi}_{h,t_0} := \psi_{\eta(t_0,h)}, \quad \text{where} \quad \eta(t_0, h) = G_\tau(t_0 + h) - G_\tau(t_0) = \int_{t_0}^{t_0+h} g(s)\, ds \tag{4.2}$$

with domain the extended phase space $\mathbb{R}^n \times \mathbb{R}$ (note that η is independent of τ). $\varnothing$

Definition 4.6. Suppose that $f \in C^\infty(D, \mathbb{R}^n)$, $g \in C^\infty(\mathbb{R})$ (with $D \subset \mathbb{R}^n$), G is well-defined, ψ is a numerical method of the ODE with vector field f and $h > 0$ a fixed stepsize. Then the numerical method $\Psi_{gf,h} = \Psi_h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}$ defined by

$$\Psi_h(y, t_0) = \begin{pmatrix} \tilde{\psi}_{h,t_0}(y) \\ t_0 + h \end{pmatrix} \tag{4.3}$$

is called the *extended induced method (of gf) (from ψ)*[15], where $\tilde{\psi}_h$ is the induced numerical method (of gf) (from ψ). $\varnothing$

If ψ_h is of order N and $\eta(t_0, h) = \mathcal{O}(h)$, then $\Psi_h : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ is a well-defined method of order N on the (canonically extended) ODE with vector field $(gf, 1)$ (and $\tilde{\psi}_{h,t_0}$ on the ODE of gf), since

$$\psi_h = \phi_h + \mathcal{O}(h^{N+1}) \qquad \Longrightarrow \qquad \Psi_h(y, t_0) = \begin{pmatrix} \psi_{\eta(t_0,h)}(y) \\ t_0 + h \end{pmatrix} = \begin{pmatrix} \phi_\eta(y) + \mathcal{O}(\eta^{N+1}) \\ t_0 + h \end{pmatrix} = \begin{pmatrix} \tilde{\phi}_{t_0+h,t_0}(y) \\ t_0 + h \end{pmatrix} + \mathcal{O}(h^{N+1}).$$

Remark 4.7. *As for the adjoint method, one can check that the adjoint of Ψ_h (well-defined locally, for sufficiently small h) satisfies*

$$(\Psi_h)^*(y, t_0) = \begin{pmatrix} \psi^*_{-\eta(t_0+h,-h)}(y) \\ t + h \end{pmatrix}.$$

Furthermore, we will need, in the numerical integration of the tidal wave system, the adjoint method of

$$\tilde{\Psi}_h(y, t_0) = \begin{pmatrix} \tilde{\psi}_{h,t_0}(y) \\ t_0 \end{pmatrix}$$

so that time t_0 is not advanced. Its adjoint is given by

$$\tilde{\Psi}^*_h(y, t_0) = \begin{pmatrix} \psi^*_{-\eta(t_0,-h)}(y) \\ t_0 \end{pmatrix}.$$

[15] cf. [HLW06] chapter VIII, equation (3.2)

Example 4.8. *In the case that $g(t) = \cos(t)$ and f on $\mathbb{R}^2$ one finds for example that $\eta(t_0, h) = \sin(t_0 + h) - \sin(t_0)$. If ψ_h is the SE method (Example 4.18) then*

$$\tilde{\psi}_{h,t_0}(y_n, z_n) = \psi_{\sin(t_0+h),\sin(t_0)}(y_n, z_n) = \begin{pmatrix} y_n + (\sin(t_0 + h) - \sin(t_0))\, f_1(y_n, z_{n+1}) \\ z_{n+1} \end{pmatrix}$$

where

$$z_{n+1} = z_n + (\sin(t_0 + h) - \sin(t_0))\, f_2(y_n, z_{n+1}),$$

if $f = (f_1, f_2)$. Since $\eta = \mathcal{O}(h)$ we have a method of order 4.

*Furthermore. there are two ways of interpreting Ψ_h: From the perspective of the autonomous part of the vector field f, we have a numerical method $\Psi_\eta(t_0, h)$ solving the ODE with vector field f **which is a PRK method but with non-constant time-step**. From the perspective of forced ODE gf, we have the method Ψ_h which has **constant time-step but is not a PRK method**.*

This example shows the following important remark.

Remark 4.9. *Given a numerical method ψ for a forced ODE $g(t)f(y)$. Then the induced method can be seen either as a PRK method with non-constant step-size solving the autonomous ODE with vector field f (Figure 14), or a method with constant step-size (but not a PRK method) on the non-autonomous ODE with vector field gf.*

We now look for a Taylor expansion of $\tilde{\psi}$, as in Remark 4.3. Suppose $\psi_h = \sum_{j=0}^{\infty} \frac{h^i}{i!} d_i$. Since $\tilde{\psi}_{h,t_0} = \psi \circ \eta(t_0, h)$ is a composition, the Taylor expansion is given by Faà di Bruno's formula (e.g. [Wik22])

$$\tilde{\psi}_{h,t_0} = \psi \circ \eta_{t_0}(h) = \sum_{j=0}^{\infty} \frac{d_j}{j!} \left(\sum_{j=1}^{\infty} h^i g_i(t_0) \right)^j = id + \sum_{i=1}^{\infty} \frac{h^i}{i!} \sum_{j=1}^{i} d_j B_{i,j}(g_0(t_0), \dots, g_{i-j}(t_0)), \tag{4.4}$$

since $\int_{t_0}^{t_0+h} g(s)\, ds = \sum_{j=1}^{\infty} h^i g_i(t_0)$ for analytic g. Here $g_j = D^j g$ and $B_{i,j}$ are the exponential, partial Bell polynomials (Appendix B). Thus, $\tilde{\psi}_{h,t_0}$ has coefficients $\tilde{d}_i$ given by

$$\tilde{d}_i(y, t_0) = \sum_{j=1}^{i} d_j(y) B_{i,j}(g_0(t_0), \dots, g_{i-j}(t_0)).$$

In Appendix D it is seen (Equation (6.6)) that the exact flow of the forced ODE with vector field $g(t)f(y)$, denoted $\tilde{\phi}_{t+h,t}$, can be written as the (Lie-)series

$$\tilde{\phi}_{t+h,t} = id + \sum_{i=1}^{\infty} \frac{h^i}{i!} \sum_{j=1}^{i} B_{i,j}(g_0(t_0), \dots, g_{i-j}(t_0)) D_f^j(id).$$

Comparing this series with the coefficient $\tilde{d}_i$ one finds the following equivalence of the order of the methods and the induced method.

Proposition 4.10. *Suppose f is a vector field on $\mathbb{R}^n$ and $g \in C^\infty(\mathbb{R})$. Then ψ_h is a method of order N on the ODE with vector field f if and only if $\tilde{\psi}_h$ is of order N on the ODE with vector field gf.*

4.1.3 Splitting methods for non-autonomous ODE

In the non-autonomous case, splitting methods have to be adapted: time has to be split as well. One can do this in a natural way by considering the canonically extended system (Section 2.1.3). However, this does not seem to be a good idea in general.

> "The simplest and most used trick for avoiding the time-dependent functions is to consider t as a new coordinate [...], whereupon one solves the [extended ODE] with a standard algorithm [...]. In many cases this transformation is not very efficient for numerically solving the problem." – Blanes & Moan [BM01]

and

> “Splitting methods are frequently used as geometric numerical integrators, and they have been designed for autonomous separable systems. A substantial number of methods tailored for different [forms] of the equations [e.g. separable or near-integrabile equations] have recently appeared, showing excellent performances in many cases. When these methods are used on non-autonomous problems, usually their performance diminishes considerably, and they can even lose the order of accuracy observed for the corresponding autonomous problems [...].” – Blanes, Diele, Marangi & Ragni [Bla+10]

The papers [BM01; BC17] mention three problems which could occur:

- If the explicit time dependency of the vector field has a relatively short time scale, then the main contribution to the error will originate from it (e.g. highly oscillatory equations as, example 2 of [Bla+10] or the example in [BC06]).
- If the time-dependent functions appearing in the vector field are expensive to evaluate.
- In the case that the ODE has a special form which is lost when extending (for example separability and applying RKN methods).

To avoid this “averaging” vs “frozen” treatment of time [Bla+10; BCM12], in particular integrators based on the Magnus series [BM01; BC06], can be used.

Separable, Hamiltonian, non-autonomous systems which are near-integrable, such as the tidal wave system, are considerd in [Bla+10], see also [BC06] for the separable, non-autonomous case. These methods might make be suitable for an application to the tidal wave system. However, in this thesis we stick to the splitting method as used in [Wal21].

We emphasise that the three disadvantages mentioned above, using splitting methods on extended phase space, are disadvantages with respect to (time-)efficiency and accuracy due to causes which are unrelated to structure preserving integration (which are treated in Section 4.2)

4.2 Structure preservation and symplectic integrators for autonomous Hamiltonian ODE

Next we discuss structure preserving methods and symplectic integrators. The idea of structure preserving numerical integration is to focus on the global and qualitative behaviour, as opposed to the local behaviour. In other words, one tries to preserve the structures of the orbits and of phase space instead of having the focus only on the minimisation of the local error.

This shift of perspective (one could even say a paradigm shift) from the local behaviour to the global behaviour of ODE happened, in the theoretical case, already around the 1900s with Poincaré’s results on non-integrability and lead to dynamical systems theory. In the numerical case, this happened around the 1980s (see e.g. [LR04; HLW06] seen again as a paradigm shift in the former). Also in the numerical case, it is important that the numerical method is a dynamical system: A first demand for structure preserving numerical methods is therefore that they form a (discrete) dynamical system. The simplest example of this are one-step methods with fixed step size:

Definition 4.11. A numerical method ψ_h ($h > 0$ the timestep) is a *one-step method* if domain and codomain are equal e.g. $\psi_h : \mathbb{R}^n \to \mathbb{R}^n$. ∅

One-step methods only depend on the previous step. For example: RK methods are one-step methods; multistep methods are not. For a forced ODE, the extended induced method, induced by a one-step method is again a one-step method.

One-step methods with fixed step size $h > 0$ form discrete dynamical systems.

Remark 4.12. *The approximations ψ_h forms a discrete dynamical system (forward-in-time so with respect to $\mathbb{N}$) generated by ψ_h. However, this is in general not a one-parameter (semi)group in h since $\psi_{h+\eta} \neq \psi_h \circ \psi_\eta$ (in particular $\psi_{-h} \neq \psi_h^{-1}$).*

Thus one-step methods open the door to (discrete) dynamical systems theory.

Beside the structure of a discrete dynamical system, other structures may be preserved, for example:

- may preserve inequalities i.e. if $t > 0$ then $\phi_t(y) \leq y$;
- may preserve conserved quantities, conserved by the flow ϕ;
- may preserve volume i.e. $\det(D_y \phi_{f,t}) = 1$ for all t in the domain;
- may preserve a (reversible) symmetry $\rho : D \to D$ i.e. for some diffeomorphism ρ satisfying $f \circ \rho = (-)f \circ \rho$ (or equivalently $\rho \circ \phi_t \circ \rho^{-1} = \phi_t^{\pm 1}$). A numerical method preserving this reversible symmetry is called *ρ-reversible*. In particular, RK methods and many partitioned methods are $\rho-$reversible, if f is $\rho-$reversible and if the method is *symmetric* ($\psi_{-h} = \psi_h^{-1}$, where the inverse exists at least locally) then the reversible symmetry is preserved by ψ [HLW06] Chapter V.1;
- or, in the Hamiltonian case, may be *symplectic*, $(D_y \phi_t)^T J D_y \phi_t = J$ (Definition 2.7),

also see [MQ01; MO14; IQ18]. Symplectic methods are closely related to methods which conserve of quadratic invariants, see [HLW06] chapter VI.4 and [Jay21].

4.2.1 Well-behavedness of symplectic integrators (and BEA)

Symplectic and reversible integrators are very important for accurate long-time integration and allow for an application of KAM theory to perturbed integrable systems [HLW06] chapter X & XI. As shown in Section 3.3, the unperturbed tidal wave system is both autonomous, Hamiltonian and most likely contains many (reversible) symmetries (it is a special case of the ABC flow). Therefore, we would like to choose symplectic methods which are preferably symmetric as well and the method considered in [Wal21] is indeed symmetric. In this thesis, we will not consider symmetric methods and reversible ODE any further, apart from a mentioning in the future research, Section 8.

We now ask the question of why it is important to choose a symplectic method. Intuitively, it is important because, by Theorem 2.14, symplecticity of the flow characterises autonomous and non-autonomous Hamiltonian flows on (a simply connected subset of) $\mathbb{R}^{2n}$. More rigorous arguments in favour of symplectic integrators have been made over the years and include linear error growth for integrable systems over long-times ([HLW06] chapter X.3), almost conservation of energy over exponentially long times ([HLW06] chapter XI.8) and KAM (and Nekhoroshev) theorems on the discretised system ([Sha99; Moa03; FQ10; MO10]).

Backward error analysis (BEA), in particular Modified Equation Analysis (MEA), is very important for these rigorous arguments[16]: Given a vector field f and a numerical method ψ for the ODE with vector field f and a stepsize $h > 0$. If ψ is of order N with respect to the ODE with vector field f then MEA searches, for given $M > N$ for a *modified vector field* $\tilde{f}_h^{[M]} = \sum_{j=1}^{n} h^{j-1} f_j$, for some vector fields f_j, such that ψ is of order M with respect to the ODE with this modified vector field i.e.

$$\psi_h = \tilde{\phi}_{\tilde{f}_h^{[M]},h} + \mathcal{O}(h^M). \tag{4.5}$$

Using BEA one finds that the structure is indeed preserved and that energy is almost conserved over exponentially long times:

Proposition 4.13 (extension of [HLW06] theorem 3.1, chapter IX.3 to non-autonomous case)**.** *Given a Hamiltonian ODE with Hamiltonian $H \in C^\infty(D \times I)$, $D \times I \subset \mathbb{R}^{2n} \times \mathbb{R}$ and a symplectic numerical method ψ of order $N \in \mathbb{N}$, then the modified vector fields f_j are locally Hamiltonian (globally if D is simply connected), such that $f_j = J^{-1} \nabla H_j$ for some $H_j : D \supset D' \to \mathbb{R}^n$.*

[16] An introduction to BEA is given in Appendix C.

Proof. The proof for the autonomous case is in [HLW06], theorem 3.1 of chapter IX.3 and is identical for the non-autonomous case, as shown next. The proof is by induction. Suppose $f_j(y,t) = J^{-1}\nabla H_j(y,t)$ for $j = 1 \dots k$. Then, by definition of the flow $\tilde{\phi}^{[k-1]} := \tilde{\phi}_{\tilde{f}_h^{[k-1]}}$, for all $h > 0$

$$D_y \psi_h = D_y \tilde{\phi}_h^{[k-1]} + h^k D_y f_k + \mathcal{O}(h^{k+1})$$

for some non-autonomous vector field f_k. This implies that

$$J = (D_y \psi_h)^T J D_y \psi_h = J + h^k \left(J D_y f_k - (J D_y f_k)^T \right) + \mathcal{O}(h^{k+1}).$$

Dividing by h^k and letting $h \to 0$ we find that $(J D_y f_k)^T = J D_y f_k$. Thus $D_y f_k$ is symplectically symmetric (Definition 2.7) and by Lemma 2.8 f_k is locally Hamiltonian (globally on simply connected domains). □

Thus, a modified Hamiltonian exist, which we denote by $\tilde{H}^{[M]} := \sum_{j=1}^{N} h^{j-1} H_j$. Using this modified Hamiltonian one may prove *in the autonomous case* almost conservation of energy up to exponentially long times.

Proposition 4.14 ([HLW06] chapter IX.7 & 8)**.** *Suppose $H : \mathbb{R}^{2n} \supset D \to \mathbb{R}$ is analytic, ψ a symplectic numerical method of order N with respect to the (autonomous) Hamiltonian ODE with vector field H and $K \subset D$ compact. Then, for $y \in D$, as long as $\{\psi_h^n(y)\}_{n\in\mathbb{N}} \subset K$, there exists $h_0 > 0$ and $M = M(h)$ such that*

$$\begin{aligned} \tilde{H}^{[M]}(\psi_h^n(y)) &= \tilde{H}^{[M]}(y) + \mathcal{O}(\exp -h_0/2h) \\ H(\psi_h^n(y)) &= H(y) + \mathcal{O}(h^{p+1}) \end{aligned}$$

over exponentially long time intervals: $nh \leq e^{h_0/2h}$.

Usually, the well-behavedness of symplectic methods is solely attributed to autonomous systems [LR04; HLW06]. Of course, there is no conservation of the Hamiltonian H for non-autonomous systems which obstructs the generalisation of Proposition 4.14. Nevertheless, one can still look for other positive advantageous of symplectic integrators applied to non-autonomous Hamiltonian systems, such as the linear error growth for integrable systems over long-times and KAM (and Nekhoroshev) theorems on the discretised system, as mentioned above. It was stated that BEA is very important to prove this and the first step towards this is Proposition 4.13 which (as opposed to Proposition 4.14) *was* a generalisation to the non-autonomous case. In this thesis, a goal is to prove an ‘approximate’ KAM theorem for the non-autonomous case (Section 6) without considering extended phase space and therefore to generalise BEA, in particular MEA, to the non-autonomous case (done in Section D and 6.3.6), contradicting the following statement.

> “Backward error analysis is a significant tool for obtaining the qualitative behaviour of the numerical solution provided by a symplectic method when integrating autonomous Hamiltonian systems. This theoretical informations is not directly available with non-autonomous Hamiltonian systems. By extending the phase space [...], it is possible to construct an autonomous Hamiltonian, the solutions of which project onto the solutions of the non-autonomous Hamiltonian[17]. Integrating the new Hamiltonian system with a symplectic method leads to a numerical solution which also has as a projection the numerical solution provided by the same integrator with the non-autonomous Hamiltonian. However, the information obtained with backward error analysis about the numerical solution in the extended phase space cannot be applied effectively to the projection.” – Cano & Lewis [CL01]

4.2.2 Symplectic PRK methods and their generating function

Only some (P)RK methods are symplectic.

[17] This “projection” of the solutions is also mentioned in Section 2.1

Proposition 4.15 ([HLW06] chapter VI.4 or [Jay21])**.** *If the Butcher tableau (A, b, c) of an s-stage RK method respectively $(A, b, c), (\hat{A}, \hat{b}, \hat{c})$ of a PRK method satisfies/satisfy (element-wise)*

$$b_i a_{ij} + b_j a_{ji} = b_i b_j, \quad \textit{respectively,} \quad \begin{matrix} b_j \hat{a}_{ij} + \hat{b}_j a_{ji} = b_i \hat{b}_j, \\ b_i = \hat{b}_i, \quad c_i = \hat{c}_i \end{matrix} \tag{4.6}$$

for $i, j = 1 \dots s$, then it is symplectic.
However, for a separable Hamiltonian $H = H(q, p, t) \in C^2$ (Definition 2.3) the conditions $b_i = \hat{b}_i$ and $c_i = \hat{c}_i$ are superfluous.

Proof. The proof in the autonomous case can be found in [HLW06] chapter IV.2 and VI.4 and the proof in the non-autonomous case is almost identical. A proof adapted to the non-autonomous case can be found (fo PRK methods) in [Jay21] theorem 2.1 using generating functions ([HLW06] chapter VI). □

As in the case of separable Hamiltonians, in the case of forced ODE the demand $c_i = \hat{c}_i$ is not necessary: If ψ is a symplectic method on $\mathbb{R}^{2n}$, then, since $D_y \tilde{\psi}_h(y, t) = D_y \psi_{\eta(t,h)}(y)$, we find immediately that the induced numerical method is symplectic:

$$D_y \tilde{\psi}_h(y, t)^T J^{-1} D_y \tilde{\psi}_h(y, t) = D_y \psi_{\eta(t,h)}(y)^T J^{-1} D_y \psi_{\eta(t,h)}(y) = J^{-1}. \tag{4.7}$$

Next we find the generating function for these symplectic maps ([HLW06] chapter VI). Consider the symplectic (P)RK method $\psi_h : (q, p) \to (Q, P)$. If q, P can be seen as a pair of independent variables then a near-identity, type-2 generating function $S(q, P)$ ([HLW06] chapter VI) i.e. such that

$$p = P + D_q S(q, P), \quad Q = q + D_P S(q, P),$$

can be found. This generating function will be useful for Hamiltonian BEA (Section 6.3.6). In the autonomous case this generating function can be found, for (P)RK methods, in [HLW06] chapter VI.5. For PRK methods applied to non-autonomous ODE we refer to [Jay21].

Proposition 4.16 ([Jay21])**.** *For a symplectic RK method respectively PRK method as in Proposition 4.15 (using notation of Section 4.1.1), mapping $(q, p) \mapsto (Q, P)$ the generating function $S(q, P, t, h) = S_h(q, P, t)$ is given by ($H_q = \nabla_q H, H_p = \nabla_p H, T_i = t + c_i h = t + \hat{c}_i h$)*

$$S_h(q, P, t) = h \sum_{i=1}^{s} b_i H(k_i, T_i) - h^2 \sum_{i,j=1}^{s} b_i a_{ij} H_q(k_i, T_i)^T H_p(k_j, T_i) \quad \textit{resp.} \quad h \sum_{i=1}^{s} b_i H(k_i, \ell_i, T_i) - h^2 \sum_{i,j=1}^{s} b_i \hat{a}_{ij} H_q(k_i, \ell_i, T_i)^T H_p(k_j, \ell_j, T_i).$$

4.2.3 Symplectic splitting methods: Symplectic Euler and Störmer-Verlet

Splitting methods for Hamiltonian ODE should be constructed by splitting into Hamiltonian vector fields $H = \sum_i H_i$. Indeed, if the splitted Hamiltonians H_i are solved exactly/with symplectic methods, then the method (e.g. Lie-Trotter/Strang splitting) are again symplectic (as was shown in Section 2.2.4). In this sense, splitting methods are a useful example of symplectic integration methods and are widely used e.g. [HLW06] chapter I.

Two important examples of symplectic PRK methods are presented as splitting methods: The symplectic Euler method and the Störmer-Verlet method (e.g. [HLW06] chapter I, II & VI and [Jay21]). We consider the autonomous case (Example 4.17) before the non-autonomous case (Example 4.18) and consider a system partitioned into two variables $q, p \in \mathbb{R}^n$ (though one could equivalently take q, p to have different dimensions) with respective vector field (f, g).

Example 4.17. *In the autonomous case, the* ***symplectic Euler (SE)*** *method (with step-size $h > 0$) $\psi_{SE,h}(q_n, p_n) = (q_{n+1}, p_{n+1})$ is a PRK method of order 1, given by*

$$q_{n+1} = q_n + h f(q_n, p_{n+1})$$
$$p_{n+1} = p_n + h g(q_n, p_{n+1})$$

with Butcher Tableaus $\mathcal{B} = (0, 1)$ and $\hat{\mathcal{B}} = (1, 1)$ used on the q variables respectively p variables. In other words, one uses (the Butcher tablueas of) explicit and implicit Euler (Example 4.1). The adjoint of the SE method ψ_{SE} (Section 4.1.1) is called the ***adjoint symplectic Euler*** *method $\psi^*_{SE,h}$ and given by*

$$
\begin{aligned}
q_{n+1} &= q_n + hf(q_{n+1}, p_n) \\
p_{n+1} &= p_n + hg(q_{n+1}, p_n)
\end{aligned}
$$

In the autonomous case, the symplectic Euler method and its adjoint can be interpreted as a splitting method: It is to the Lie-Trotter splitting for the splitting $(f, g) = (f, 0) + (0, g)$ where the forward and backward Euler are used for the splitting elements [Jay21].

The symplectic Euler methods are very useful when they are applied to an ODE with a separable Hamiltonian (Definition 2.3) such as the tidal wave system. Indeed, then it is explicit ([HLW06] chapter VI.3) so that it is a relatively cheap method (with respect to computation time).

Still in the autonomous case, the symplectic Euler methods in a composition method one can construct the Störmer-Verlet (SV) methods [HLW06; Jay21], which is a symplectic method of order 2 (explicit if the Hamiltonian is separable and split into separable pieces). They are given by

$$
\begin{aligned}
q_{n+1/2} &= q_n + \frac{h}{2} f(q_{n+1/2}, p_n) & \qquad q_{n+1/2} &= q_n + \frac{h}{2} f(q_{n+1}, p_n) \\
p_{n+1} &= p_n + \frac{h}{2}\big(g(q_{n+1/2}, p_n) + g(q_{n+1/2}, p_{n+1}\big) & \qquad q_{n+1} &= q_n + \frac{h}{2}\big(f(q_n, p_{n+1/2}) + f(q_{n+1}, p_{n+1/2})\big) \\
q_{n+1} &= q_n + \frac{h}{2} f(q_{n+1/2}, p_{n+1}) & \qquad p_{n+1} &= p_n + \frac{h}{2} f(q_{n+1}, p_{n+1/2})
\end{aligned}
$$

in other words by the Butcher tableaus [Jay21]

$$
\begin{array}{c|cc}
 & 0 & 0 \\
 & 1/2 & 1/2 \\
\hline
 & 1/2 & 1/2
\end{array}
\qquad\qquad
\begin{array}{c|cc}
 & 1/2 & 0 \\
 & 1/2 & 0 \\
\hline
 & 1/2 & 1/2
\end{array}.
$$

The Störmer-Verlet (SV) methods can be seen as a composition method of the symplectic Euler methods ([HLW06] chapter VI.3 theorem 3.4 or [Jay21]) or as a composition of the midpoint method and the trapezoidal method [Jay21].

Example 4.18. *In the non-autonomous case, if the vector fields depend also on time $(f, g) = (f(q, p, t), g(q, p, t))$ then the* ***SE methods*** *are given by*

$$
\begin{aligned}
q_{n+1} &= q_n + hf(q_n, p_{n+1}, t + c_1 h) & \qquad q_{n+1} &= q_n + hf(q_{n+1}, p_n, t + c_1 h) \\
p_{n+1} &= p_n + hg(q_n, p_{n+1}, t + \hat{c}_1 h) & \qquad p_{n+1} &= p_n + hg(q_{n+1}, p_n, t + \hat{c}_1 h)
\end{aligned}
$$

which can, for special values of $c_1, \hat{c}_1$ still be seen as a splitting method on which explicit and implicit Euler are used ([Jay21] section 3). This method is symplectic if $c_1 = \hat{c}_1$ or if the vector field is separable (Proposiiton 4.15). In particular the SE methods are symplectic if and only if this is the case, [Jay21] section 3.

In the non-autonomous case, the Störmer-Verlet method can be generalised to a method with Butcher tableaus

$$
\begin{array}{c|cc}
c_1 & 0 & 0 \\
c_2 & 1/2 & 1/2 \\
\hline
 & 1/2 & 1/2
\end{array}
\qquad\qquad
\begin{array}{c|cc}
\hat{c}_1 & 1/2 & 0 \\
\hat{c}_2 & 1/2 & 0 \\
\hline
 & 1/2 & 1/2
\end{array}.
$$

In particular, we define the ***(non-autonomous) Störmer-Verlet method*** *method to be the numerical method with $c_1 = \hat{c}_1 = 0$ and $c_2 = \hat{c}_2 = 1$, which is therefore symplectic (Proposition 4.15).*

4.2.4 Splitting methods for near-integrable systems

We briefly address splitting methods adapted to near-integrable systems i.e. of the form $H_0 + \mu H_1$ with H_0 Liouville integrable (Definition 3.4) and μ small (e.g the tidal-wave system). Splitting methods adapted to this case have been found to have good behaviour and have been developed [McL95] (and other constructions in [Sey16]). Furthermore, splitting methods for perturbed, non-autonomous ODE have been considered in e.g. [Bla+10]. In this thesis, however, we will stick to the splitting method used in [Wal21].

4.2.5 Induced method as method with time-adaptive step size and splitting time-affine systems

From Remark 4.9 it follows that the induced method can be seen either as well as a method with adaptive-step size for the *autonomous* ODE with vector field $f(y)$ (instead of a method with fixed step size for the forced ODE with vector field $g(t)f(y)$). Thus, possibly results about autonomous symplectic methods imply results about the induced methods (which are symplectic) on the respective forced ODE (using Equation (4.2))[18]. However, the fact is that symplectic methods with adaptive step size (the adaptation depending on time) are, at least in general, not as well-behaved as fixed step size symplectic integrators [RF11], see also Remark 4.12. Therefore, we do not expect this implication of results for the induced methods about well-behavedness/structure preservation.

A time-affine vector field $\sum_{i=1}^{n} g_i(t)f_i(y)$ can be split into $n > 1$ forced ODE $g_i f_i$ and induced methods can be used to integrate the split ODE numerically. Since all step sizes are different, one cannot compare in this case to an autonomous ODE (e.g. $\sum_{i=1}^{n})f_i$ and the above perspective, of a time-adaptive method on an autonomous ODE is, moreover, lost.

Remark 4.19. *Thus, the view of the induced method as a time-adaptive step-size on an autonomous Hamiltonian ODE and the form of the tidal-wave system as a time-affine system seem to be of no help to prove (approximate) KAM theorems or other similar results (based on structure preservation) for induced methods applied to non-autonomous forced/time-affine (Hamiltonian) ODE.*

The main advantage of the fact that we are dealing with forced and time-affine ODE seems to be that it is easy to find high order (induced) methods for the forced/time-affine system since we only need to find a high-order method for the autonomous part. Furthermore, the expressions for the flow using Lie-Gröbner series and expressions for MEA are relatively simple (see Section D)).

Finally, for future research we speculate that a further advantage could be related to the "form" of the equations as in Section 4.1.3 (see the quotes therein) and 4.2.4. In [BC06; Bla+10] effort has been given to extend splitting methods in a proper way to the non-autonomous case so as to preserve the advantageous properties of splitting methods adapted to the autonomous version of this (separable/near-integrable Hamiltonian) form. Perhaps for forced/time-affine ODE, the induced method is easily found to imply such a preservation of the advantageous properties. [19]

4.3 Outlook: Non-autonomous structure preserving methods

We continue the discussion of Section 2.4 (this section may be seen as an *intermezzo*) and discuss now the numerical side: non-autonomous structure preservation. We saw in that section that the structure of non-autonomous Hamiltonian ODE is not entirely clear: There are multiple geometric frameworks (types of objects) and, for example, multiple ways to define canonical transformations (types of morphisms) in extended phase space. Furthermore, little progress has been made for structure-preserving methods of non-autonomous Hamiltonian systems (see also the quotes in the introduction). We therefore will not make any

[18] Already in the theoretical case (without considering numerical methods) this is possible using Equation (2.5). For example the flow of a forced ODE preserves the autonomous part of the Hamiltonian (Equation (2.4)

[19] One might complain at this point, since we have just argued that we do *not* expect the preservation of advantageous properties related to structure preservation. However, it was emphasised in Section 4.1.3 that the papers [BC06; Bla+10] tried to preserve advantages which were *unrelated* to structure preservation.

rigorous statements about structure-preserving methods for non-autonomous Hamiltonian systems

Regarding the objects: presymplectic integrators have been constructed [Fra+21], although this was constructed for constrained Hamiltonian systems rather than non-autonomous ones. Moreover, KAM theorems for presymplectic integrators have been constructed [AL12]. Regarding the morphisms: numerical methods were constructed which were also canonical transformations [MO14], defined on extended phase space and the contact manifold mentioned in Section 2.4. Furthermore, symplectic PRK methods were examples of such a method ([MO14] Corollary 4.4.4).

Variational integrators ([HLW06] chapter VI.6) and generating function methods ([HLW06] chapter VI.5 or [FQ10]), which construct symplectic integrators, are both generalised to the non-autonomous case [CL01] respectively [Qin96].

The paper [Qin96] uses a generating function approach on the extended phase space and states moreover that this seems to be a better approach than to adapt this approach directly to the autonomous case:

> "In this way, the accuracy of the construction of symplectic schemes for nonautonomous systems in some cases is not high. Another way is through [the extended phase space], so nonautonomous systems are transformed to autonomous systems, then using [a generating function method], we can get symplectic schemes for nonautonomous systems. In this way the construction is simple and the accuracy of the construction of symplectic schemes is high." – Q. Mengzhao [Qin96]

Thus, the symplectic structure on extended phase space could also be the right structure for non-autonomous systems, in which case a symplectic PRK method is a good method to use ([MO14] corollary 4.4.4 or [Jay21] section 5).

Finally, it might be the case that the symplectic structure on the original phase space is the right structure after all. Indeed, [Jay21] section 1 and references therein mention that preservation of the symplectic structure is a "desirable property"[20]. Furthermore, we will see in Section 6 that, for these (symplectic) methods MEA can be adapted to the non-autonomous Hamiltonian case, which allows for a proof of an approximate KAM theorem.

4.4 An affine splitting method for the tidal wave system

We now consider the affine splitting method as used in [Wal21]. In this section we integrate the tidal wave system, with the Hamiltonian of Equation (1.5)

$$L(q,p,t) = k(p-q)\cos(t) - C_\delta k^2 l(\alpha\cos(p) + \beta\cos(q)) + 2\gamma C_\delta kl[(r\cos(t) + \sin(t))(\alpha\sin(p) + \beta\sin(q))]$$

where $\alpha = fk + lr$, $\beta = fk - lr$ and $C_\delta = \delta(k^2+l^2)^{-1}(k^2+2r^2+2)^{-1}$. The goal of the remaining part of this thesis is to try to prove that KAM tori persist in the numerically integrated system e.g. by applying MEA and using KAM theory on this particular numerical method as considered in [Wal21].

In [Wal21], the following parameter sets are considered, of which we only integrate the "default" and "Simple-B" parameter sets.

	r	f	γ	δ	α	β
Default	2	0	0	0.3	2l	-2l
Simple-B	2	0	0	$\frac{6(k^2+2r^2+2)}{10(kr)}$	2l	-2l
Coriolis	2	1	0	0.3	k + 2l	k - 2l
Vorticity-harmonic	2	0	1	0.3	2l	-2l
Cor-VortHarm	2	1	1	0.3	k + 2l	k - 2l

[20] However, due to the fact that symplectic PRK methods are both symplectic on the original as well as the extended phase space ([MO14] corollary 4.4.4 or [Jay21] section 5) it is not entirely clear what structure should be preserved when only using symplectic PRK methods, such as in this thesis.

As in [Wal21] we split the Hamiltonian into forced, Hamiltonian ODE $L(q,p,t)$ as

$$L_1(q,p,t) = k(p-q)\cos(t), \quad L_2(q,p,t) = C_\delta k^2 l(\alpha\cos(p)+\beta\cos(q)), \quad L_3(q,p,t) = \gamma 2C_\delta kl[(r\cos(t)+\sin(t))(\alpha\sin(p)+\beta\sin(q))]$$

These pieces are separable, so that the induced symplectic Euler methods are explicit, Remark 4.17 and Equation (4.2). Since the system is non-autonomous, we also need to 'split time'. To split time, we consider the extended Hamiltonian $L(q,p,t)+s$ (Equation (2.7)), where s is the variable conjugate to time. Time will be treated independently, i.e. we add the split Hamiltonian $L_4(s) = s$. In the "default" and "Simple-B" case, $\gamma = 0$ so that L_3 is not considered in the numerical integrations shown in this thesis. However, we will consider L_3 when describing the construction of the symplectic splitting method.

We discuss now the numerical method from the two perspectives as discussed in Remark 4.9. From the autonomous perspective we use a method similar to the Strang-splitting (Section 4.1.1) but now split into 4 parts. On these part we will use either the exact flow or the (autonomous) SE method (Example 4.17). From the time-affine perspective we split into 4 vector fields and use 4 induced methods with constant step-size. Then, denoting the induced SE methods on L_i $(i = 1,2,3)$ as $\tilde{\psi}_i$ (with time-reparametrisation G_i, Equation 2.5) and the extended induced method Ψ_4 of L_4 we find

$$\begin{aligned}
\tilde{\psi}_{1,h}(q,p,t) &= \begin{pmatrix} q \\ p \end{pmatrix} + k\left(G_1(t+h) - G_1(t)\right)\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{with} \quad G_1(t) = \sin(t), \\
\tilde{\psi}_{2,h}(q,p) &= \begin{pmatrix} q \\ p \end{pmatrix} + \left(G_2(t+h) - G_2(t)\right)C_\delta k^2 l \begin{pmatrix} \alpha\sin(p) \\ -\beta\sin\left(q + hC_\delta hk^2 l\alpha\sin(p)\right)) \end{pmatrix}, \quad \text{with} \quad G_2(t) = t, \\
\tilde{\psi}_{3,h}(q,p) &= \begin{pmatrix} q \\ p \end{pmatrix} + 2\gamma C_\delta kl\left(G_3(t+h) - G_3(t)\right)\begin{pmatrix} \alpha\cos(p) \\ -\beta\cos\left(q + 2C_\delta kl\left(G_3(t+h) - G_3(t)\right)\frac{kl}{\alpha}\cos(p)\right) \end{pmatrix}, \; G_3(t) = r\sin(t) + \cos(t), \\
\Psi_{4,h}(q,p,t) &= \begin{pmatrix} q \\ p \\ t+h \end{pmatrix}.
\end{aligned} \tag{4.8}$$

Note that the maps $\tilde{\psi}_{1,h}$ and $\Psi_{4,h}$ are equal to the exact time-h flow of the Hamiltonian L_1 respectively L_4.

Now, denoting the extended induced SE methods (without time-step) as $\tilde{\Psi}_{i,h}(y,t) = (\tilde{\psi}_{i,h}(y,t),t)$, we use the method

$$\Phi_h = \Xi^*_{h/2} \circ \Xi_{h/2}$$

where

$$\Xi_h = \Psi_{4,h} \circ \tilde{\Psi}_{3,h} \circ \tilde{\Psi}_{2,h} \circ \tilde{\Psi}_{1,h}. \tag{4.9}$$

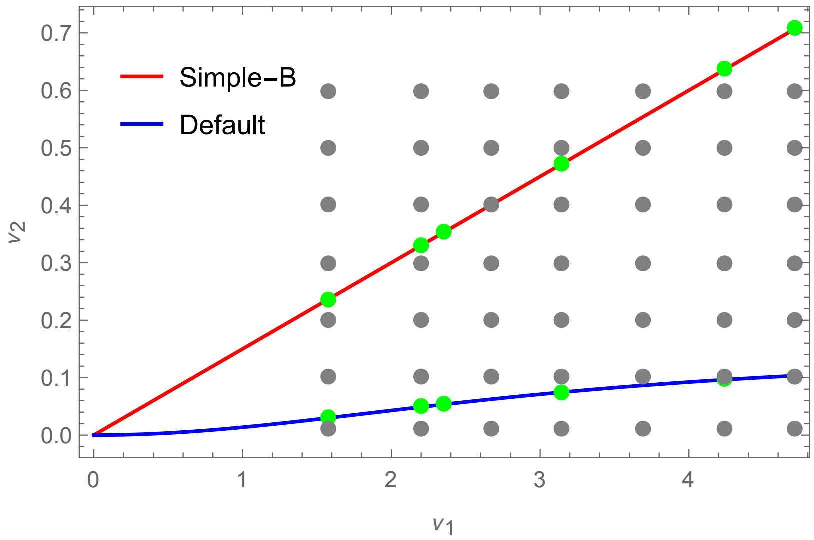


Figure 15: This figure shows (for $k = l, r = 2$): The values of ν_1 against ν_2 which agree with the "Simple-B" parameter set (red, $\nu_2 = \frac{3}{20}\nu_1$) and with the "default" parameter set (blue, $\nu_2 = \frac{3\nu_1^2}{20(\nu_1^2+10)}$). Furthermore, it shows the value of ν_1 against ν_2 for the plots in Table 1 (grey dots) and for the plots as in the paper [Wal21] (green dots).

In other words, denoting $\tilde{\psi}_h = \tilde{\psi}_{3,h} \circ \tilde{\psi}_{2,h} \circ \tilde{\psi}_{1,h}$ one finds the method

$$\begin{aligned}\Psi_h(q,p,t) &= \begin{pmatrix} \left(\tilde{\psi}\right)^*_{h/2}(q,p,t+h) \circ \tilde{\psi}_{h/2}(q,p,t) \\ t+h \end{pmatrix} \qquad (4.10)\\ &= \begin{pmatrix} \left(\psi^*_{1,\tilde{h}_1(t+h/2,h/2)} \circ \psi^*_{2,h/2} \circ \psi^*_{3,\tilde{h}_3(t+h/2,h/2)}\right) \circ \left(\psi_{3,\tilde{h}_3(t,h/2)} \circ \psi_{2,h/2} \circ \psi_{1,\tilde{h}_1(t,h/2)}\right)(q,p) \\ t+h \end{pmatrix}, \qquad (4.11)\end{aligned}$$

where Remark 4.7 was used to calculate the adjoint methods of the induced methods, such that

$$\left(\tilde{\psi}\right)^*_{h/2}(q,p,t+h) = \psi_{1,-\tilde{h}_1(t+h,-h/2)} \circ \psi_{2,h/2} \circ \psi_{3,-\tilde{h}_3(t+h,-h/2)}(q,p) = \left(\psi^*_{1,\tilde{h}_1(t+h/2,h/2)} \circ \psi^*_{2,h/2} \circ \psi^*_{3,\tilde{h}_3(t+h/2,h/2)}(q,p)\right)$$

since $-\tilde{h}_i(t+h,-h/2) = -(G_i(t+h/2) - G_i(t+h)) = \tilde{h}_i(t+h/2,h/2)$. Equation (4.10) implies that the method is a symplectic method with respect to (q,p), reversible and of order 2. Equation (4.11) is the one used in the code in Appendix E and equals the numerical integration method used in [Wal21], proving again that it is indeed of order 2, symplectic (w.r.t. (q,p)) and symmetric.

As $\gamma = 0$ we find that $\tilde{\psi}_3$ of Equation (4.8) equals the identity. Furthermore, the perturbation parameter μ is introduced as in Section 3.3.3 and 3.3.4. Since $k = l$ the Hamiltonian takes the form

$$L(q,p,t) = \mu k(p-q)\cos(t) - C_\delta k^4 (\alpha\cos(p) + \beta\cos(q)) = \nu_1(p-q)\cos(t) + \nu_2((f+r)\cos(p) + (f-l)\cos(q))$$

for $\nu_1 = \mu k$, $\nu_2 = -C_\delta k^4$. The results of the numerical integration using the induced splitting method are shown in Table 1 for various values of ν_1, ν_2 (or equivalently, k and μ, see Figure 15), cf. the figures in [Wal21].

The Table shows very interesting figures. We see that for small perturbations ν_1 the tidal Poincaré has seems to have many invariant tori but are destroyed when the perturbation μ is increased, if we may trust the numerical figures.

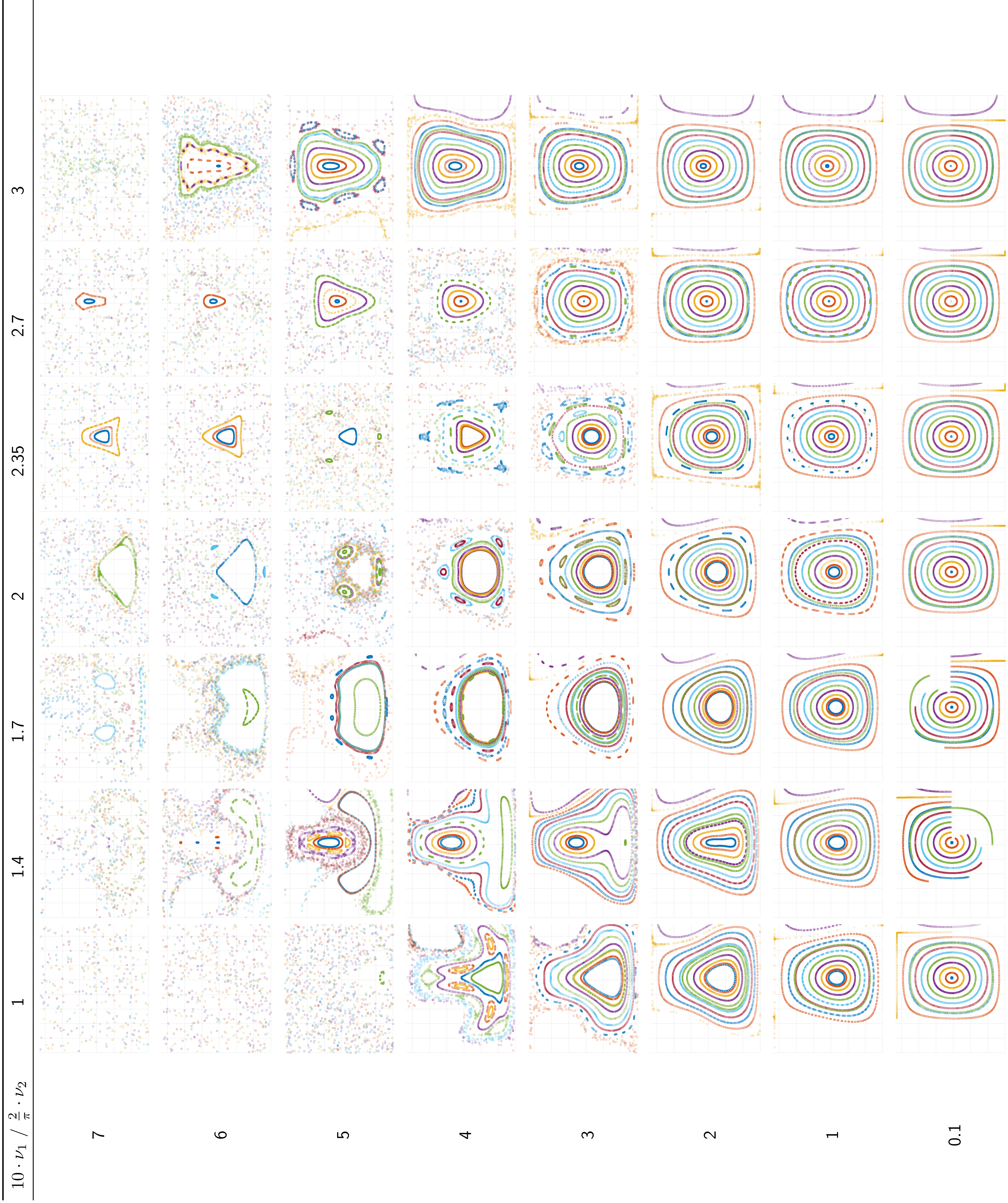


Table 1: Plots of the tidal map system without vorticity term $k = l, r = 2$ for $(q, p) \in [-1, 4] \times [0, 3]$ (coordinate system for the Hamiltonian H as in Equation (1.4)) for various scaling factors of the splittings ν_1, ν_2.

5 Non-autonomous flow interpolation and exact KAM theory applied to the tidal wave system

In the next two sections, Sections 5 and 6, the goal is to show the existence of KAM tori (persistent invariant tori) in the periodic Poincaré map (e.g. [Wal21] or [Wig03] chapter 10.2) of the numerically integrated tidal wave system using the symplectic integrator of Section 4.4. Table 1 shows that we may indeed expect KAM tori for some values of k and the added perturbation paramter μ.

To prove the existence of KAM tori in the numerically integrated, unperturbed and perturbed tidal wave system we will use KAM theorems in the numerical case i.e. *numerical KAM theorems.* As discussed in the Introduction 1, as in the theoretical case there exist a discrete [Sha99; HLW06] and continuous [KP94b; Moa03; HLW06; MO10] way to construct numerical KAM theorems. In both cases, the step-size $h > 0$ enters explicitly in the conditions of these discrete or continuous, numerical KAM theorems. Indeed, if some $h_* > 0$ resonates with the frequencies $\omega_* \in \mathbb{R}^n$ of a particular invariant tori in the unperturbed system in $\mathbb{R}^{2n}$, then this particular tori is destroyed when numerically integrating with the step-size h_* [Sha99; Moa03; HLW06] (this destruction due to resonance is also discussed in Section 3.2).

Two strategies for a proof of KAM tori in the *theoretical* tidal wave system were presented in Section 3 (see also Figure 10): One using a discrete map/dynamical system i.e. considering the 2π-Poincaré map $\phi_{t_0+2\pi,t_0}$, the other using a continuous counterpart i.e. the non-autonomous (Hamiltonian) flow ϕ_{t_0+h,t_0} for $h \in \mathbb{R}$ (see also Figure 16). In the *numerical* case we have again two such strategies.

We will first discuss both strategies, one of which uses backward error analysis BEA. Afterwards we discuss two types of BEA and arrive, using one of them, at a numerical KAM theorem which is applied to the tidal wave system.

5.1 Strategies for the application of numerical KAM theorems to the tidal wave system

Two strategies exist to prove KAM tori in the 2π-Poincaré map of the tidal wave system. The first strategy is to consider the 2π-Poincaré map $\phi_{t_0+2\pi,t_0}$ of the tidal wave system and use the *symplectic* numerical integration method $\Psi_h(q,p,t) = (\psi_h(q,p,t), t+h)$ of Section 4.4 to approximate this map. This is done by choosing $h = 2\pi/K$, $K \in \mathbb{N}$, and considering the map $\psi_h^K(\cdot, t_0)$. For large K (small step-size h) $\psi_h^K(\cdot, t_0)$ approximates (for a consistent method) the 2π-Poincaré map, $\psi_h^K(\cdot, t_0) \approx \phi_{t_0+2\pi,t_0}$, and $\psi_h^K(\cdot, t_0)$ can be seen as a *numerical* 2π*-Poincaré map.* If $\phi_{t_0+2\pi,t_0}$ is a twist map then for large K, $\psi_h^K(\cdot, t_0)$ is likely a twist map (possibly locally, on a subset of $\mathbb{R}^{2n}$). Since it is also a symplectic map, one may then apply a discrete KAM theorem [Sha99] to prove the existence of KAM tori in the numerical Poincaré map, which was the goal.

For this first strategy, it seems unavoidable that results from the theoretical case as stated in Section 3 are needed e.g. that the Poincaré map $\phi_{t_0+2\pi,t_0}$ is a twist map (locally). Since we have not done so we therefore use a different strategy: Instead of considering the 2π-Poincaré one considers the numerical method Ψ_h of Definition (4.3). The map Ψ_h is a symplectic map with respect to $(q,p) \in \mathbb{R}^2$, so that discrete KAM is naturally applied (if it is a twist map) using again [Sha99] (or [Mös62]). One may also apply a continuous KAM theorem[21] by the use of backward error analysis (BEA) [LR04; HLW06]. BEA views the (symplectic) map Ψ_h as the time-h flow of a (Hamiltonian) flow $\tilde{\Phi}_h \approx \Psi_h$, so that a continuous KAM theorem can be applied to $\tilde{\Phi}$ [GS86; CMS94; Moa05; Moa06]. The map $\tilde{\Phi}$ is seen as an (approximately) interpolating Hamiltonian flow of the symplectic map Ψ_h (as discussed in 2.2.3).

The main reason to apply a continuous KAM theorem rather than a discrete KAM theorem is that BEA is used. BEA explains, in the case of autonomous Hamiltonian ODE, why symplectic numerical integrators behave very well (e.g. Proposition 4.14).

[21] This continuous KAM theorem then proves the existence of KAM tori in the numerically integrated tidal wave system. As in Section 3 one then deduces, using Remark 3.10, the existence of KAM tori in the (numerical) 2π-Poincaré map

> "In our opinion, to the numerical analyst, the most appealing feature of symplectic integration is the possibility of *backward error interpretation*." – Sanz-Serna [San92]

As in Section 4.3, we are interested in the structure-preserving integrators of non-autonomous Hamiltonian ODE and BEA will help us to show the well-behavedness of symplectic integrators on non-autonomous Hamiltonian ODE in Section 6. Therefore we adopt the second strategy discussed above (Figure 16) and apply a continuous, numerical KAM theorem but first discuss BEA.

5.2 Backward error analysis: Modified equation analysis and non-autonomous flow interpolation

2π-Poincaré An introduction to backward error analysis is given in Appendix C. It can be summarised as follows: Forward error analysis considers the difference between the *exact solution* ϕ and the *approximate solution* ψ (called the forward error), which one wants minimised. For numerical methods on ODE we find, unsurprisingly, that ϕ is the flow and ψ the numerical method and the forward error is nothing other than the global error (Equation (4.1)) i.e. of the form $\phi - \psi$.

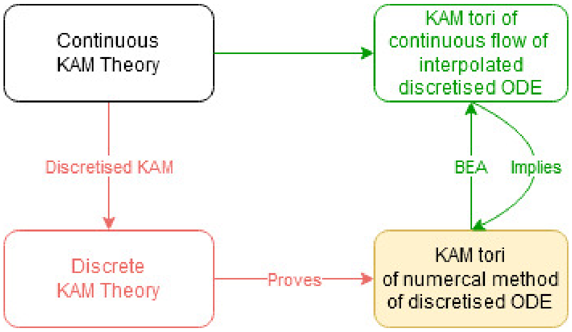

Continuous KAM Theory
KAM tori of continuous flow of interpolated discretised ODE
Discretised KAM
BEA
Implies
Discrete KAM Theory
Proves
KAM tori of numercal method of discretised ODE

Figure 16: The two strategies to prove the existence of KAM theory are via the discrete (2π twist-map) or via the continuous (using BEA)

In backward error analysis one instead considers (and sometimes minimises, Appendix C) the difference between the *exact problem* and the *approximate problem*. For ODE the problem is defined using a vector field and therefore the backward error is of the form $\tilde{f} - f$, where f is the vector field of the exact problem (thus having flow ϕ) and $\tilde{f}$ a vector field of an approximate problem (with flow $\tilde{\phi}$ which approximates ψ in some sense). For example, the modified vector field $\tilde{f} := \tilde{f}^{[M]}$ as in Section 4.2 may be used. The flow $\tilde{\phi}_{\tilde{f}^{[M]}}$ interpolates up to order M the flow of the differential equation as in Equation (4.5). More generally, BEA for ODE one constructs a *modified/interpolative vector field* $\tilde{f}$ (or equivalently a modified/interpolative flow $\tilde{\phi}_{\tilde{f}}$)

In the Hamiltonian case, one looks for a modified Hamiltonian $\tilde{H}$. Therefore, in this case BEA has much in common with the embedding of symplectic maps into Hamiltonian flows. Besides MEA, the embedding of symplectic maps into Hamiltonian flows, or *(non-autonomous) flow interpolation* is a second method for BEA, as discussed in Section 2.2.3 and the references in therein.

Thus, as mentioned in the Introduction 1, there are two types of BEA (discussed in this thesis): One using modified equations, called modified equation analysis (MEA), which constructs an autonomous modified vector field $\tilde{f}$ [GS86; CMS94; San92; SC94; LR04; HLW06], the other using *non-autonomous flow interpolation* which constructs a non-autonomous modified vector field $\tilde{f}$ [San92; Wan94; SC94; Moa05; Moa06]. In both cases, the Hamiltonian structure is preserved *when symplectic integrators are used* and the autonomous [HLW06] (or Proposition 4.13) and non-autonomous [Wan94] modified vector fields are again Hamiltonian.

The BEA methods can be described as follows.

- Modified equation analysis (e.g [GS86; LR04; HLW06]) finds a *modified vector field* $f^{[M]}$, of which (given step-size $h > 0$) the flow $\phi_{f^{[M]},h}$ approximates the numerical method ψ_h to an order $M > N$ (Equation (4.5)). We will see that $f^{[M]}$ is autonomous when applied to an ODE with autonomous vector field f.
- *Non-autonomous (Hamiltonian) flow interpolation* constructs the modified vector field $\tilde{f}$ by using the numerical method ψ_t as a casting, e.g. [Moa05] and the references in Section 2.2.3. For an autonomous vector field f, one finds, using non-autonomous flow interpolation, a *non-autonomous* modified vector field $\tilde{f}$.

Both methods, MEA and non-autonomous Hamiltonian interpolation, have (dis)advantages:
The advantage of MEA is that the construction is rather explicit (it can be found recursively). A big disadvantage is that one cannot, in general, interpolate the flow exactly ($M = \infty$) but only approximately ($M < \infty$): It is well-defined for $M = \infty$, but only as a formal power series which in general diverges numerically/as a function [HLW06]. For $M < \infty$ the flow approximates the numerical method and one can choose $M^* = M$ optimally [BG94; HL97; Rei99; LR04; HLW06]. From the point of view of asymptotic analysis, this value is often referred to as *optimal truncation* (of the diverging power series) [LR04; HLW06]. Truncating in such a way (or with $M < \infty$), one obtains not an 'exact', numerical KAM theorem for the numerical method, but an 'approximate', numerical KAM theorem e.g. [BG94] or [HLW06] chapter X.5. The approximate KAM theorem, as opposed to the exact KAM theorem, does not prove the existence perpetual invariance of orbits (invariant tori), but existence of approximate invariance ('approximately invariant', persistent tori). The approximation is very good, however: The 'approximate invariance' is exponentially close in h^{-1} to actual invariance for exponentially long times in h^{-1}, where h the time-step.
The advantage of non-autonomous Hamiltonian interpolation is that, after a time transformation, this vector field is well-defined and interpolates the flow exactly. This allows one to apply *exact* KAM theorems [Moa03; MO10], such as the KAM theorem 3.9 in Section 3.2. A disadvantage is that this construction is much more involved, e.g. [Moa05].

In this section, Section 5, we discuss such an exact (numerical) KAM theorem using non-autonomous flow interpolation. In Section 6 we discuss an approximate (numerical) KAM theorem using MEA.

Note that, although we call the KAM theorems for the numerically integrated systems (using a symplectic method) *numerical* KAM theorems, only discretisation error is considered and **rounding error is ignored**. However, when taking rounding error into account, the approximate KAM theorem of Section 6 seems to be more useful than the exact KAM theorem of Section 5.

Finally, we mention briefly that both types of BEA can be linked to classical perturbation theory [Cal04] (linking MEA to Lie-Hori theory) and [Hen96] (linking MEA to Hori transformations in the autonomous case and Deprit's transformation in the non-autonomous case).

5.3 Non-autonomous Hamiltonian flow interpolation

In Section 2.2.3 it is seen that the theory of interpolation of (symplectic) flows by smooth/analytic (Hamiltonian) flows is well established. In the setting of BEA, it was mentioned by e.g. [San92; Wan94] but these may not the first sources to mention this.

In particular, rigorous non-autonomous flow interpolation uses a construction, the suspension construction, which is used to embed discrete dynamical systems into continuous dynamical systems [BS02]. This construction was used by Douady to prove the equivalence of discrete and continuouns KAM theory and was used by Moan [Moa05; Moa06] for a rigorous treatement of this type of BEA, see also [Moa03; MO10]. In this section we will discuss this treatment of non-autonomous flow interpolation, mainly based on the papers of Moan.

In this Section, Section 5.3, only autonomous systems are considered. The treatment for non-autonomous systems is found by using the canonical autonomous extension of Section 2.1.

5.3.1 Constructing the non-autonomous interpolated vector field and flow

Given a numerical method ψ_t on $\mathbb{R}^n$ for an autonomous ODE, which is differentiable in $t \in [0, h]$. For a fixed $h > 0$ and $y_0 \in \mathbb{R}^n$ the numerical method ψ_t, for $t \in [0, h]$, defines an integral curve $y(t) = \psi_t(y_0)$ along which we can define the tangent vectors

$$\dot{y}(t) = D_t\psi_t(y_0) = D_t\psi_t \circ \psi_t^{-1}(y(t)) \tag{5.1}$$

assuming $h > 0$ is sufficiently small such that ψ_t is invertible. This vector field has ψ_t as a solution flow. This construction can be extended to any y in the domain and $t \in [0, h]$ to find the modified vector field [Moa05;

Moa06]

$$\mathfrak{f}(y,t) = D_t\psi_t(\psi^{-1}(y)).$$

This vector field is periodically extended from $t \in [0,h]$ to $t \in \mathbb{R}_+$ (and for $t \in \mathbb{R}$ one may use ψ_h^{-1}). Doing so, one finds that $\mathfrak{f}$ has the numerical method ψ_t as solution flow $\phi_{\mathfrak{f},t}$ for $t \in [0,h]$ and, moreover $\phi^n_{\mathfrak{f},h} = \phi_{\mathfrak{f},nh}$, and the backward error analysis seems done since we have found a modified vector field. However, $\mathfrak{f}$ is a discontinuous vector field at the points $t \in h\mathbb{Z}$ (so that the solution flow may be non-existent), a problem dealt with in for example [Moa06] (see also the references in Section 2.2.3) by smoothening the discontinuity in time as discussed in Section 5.3.2.

Again, the modified equation is Hamiltonian (e.g. [Wan94]) since, by the symplecticness of Ψ, $\nabla\Psi_t J \Psi_t = J$, one finds

$$\nabla\mathfrak{f}(y,t)J(\nabla\Psi_t(y,t))^T - \left(\nabla\mathfrak{f}(y,t)J(\nabla\Psi_t(y,t))^T\right)^T = 0.$$

Which means that $(\nabla\mathfrak{f}(y,t)J)^T = \nabla\mathfrak{f}(y,t)J$ if $\nabla\psi_t$ is invertible, which holds for $h > 0$ sufficiently small (and was already assumed) so that $\mathfrak{f}$ is locally Hamiltonian from the Integrability Lemma 2.8. The modified Hamiltonians can be found using generating functions. In [San92] it is seen that, if $\tilde{S}(q,p_0,t)$ is a generating function, such that

$$\psi_t(D_{p_0}\tilde{S}(q,p_0,t),p_0) = \begin{pmatrix} q \\ D_q\tilde{S}(q,p_0,t) \end{pmatrix}$$

for $t \in (0,h)$, then the modified, non-autonomous Hamiltonian $\tilde{H}$ is given by

$$\tilde{H}(q, D_q\tilde{S}(q,p_0,h),t) = -D_t\tilde{S}(q,p_0,t)$$

.

5.3.2 Smoothening the disconituity in time

The modified non-autonomous vector field $\mathfrak{f}$ and Hamiltonian K are h-periodic in time and can be found explicitly, but have discontinuities. These discontinuities can be smoothened and even be made analytic, see [Moa06].

Theorem 5.1 ([Moa06] theorem 5)**.** *If $\psi_h : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is a symplectic method ($h > 0$ fixed) applied to the ODE with autonomous vector field f, where f is analytic for $y \in \mathcal{D}_{\delta_1+\delta_2} := \{z \in \mathbb{C}^{2n} \mid \|z - y_0\|_\infty \leq \delta_1 + \delta_2\}$, for some $y_0 \in \mathbb{R}^{2n}$. Then there exists a non-autonomous modified vector field*

$$\tilde{f}_h(y,t) = f(y) + \epsilon\,(r_1(y) + r_2(y,t)) \tag{5.2}$$

analytic in $\mathcal{D}_{\delta_1}$ and analytic and h-periodic in t ($r_2(y,t+h) = r_2(y,t)$). Moreover, the flow of the ODE with modified vector field $\tilde{f}$, $\phi_{\tilde{f},t,t_0}$ exactly interpolates Ψ_h

$$\phi_{\tilde{f},h,0} = \Psi_h$$

and satisfies, for $h\|f\|_{\delta_1+\delta_2} < \frac{2\pi\delta_2}{e}$,

$$\|\epsilon r_2\| \leq C\exp\left(\frac{-2\pi\delta_2}{2h\|f\|_{\delta_1+\delta_2}}\right),$$

where $\|f\|_\delta = \sup_{y\in\mathcal{D}_\delta}\|f(y)\|_\infty$.

From [Moa05], *Note 3* it follows that $\tilde{f}$ is again Hamiltonian if f is Hamiltonian, so that the modified vector field, may be seen as a Hamiltonian perturbation. However, it seems from [Moa03; Moa05; Moa06; MO10] that there are no strong conclusions about the size of ϵr_1 and therefore it is not clear if h is a (small) perturbation parameter, although it is stated once in [Moa03] it is stated that $\epsilon(r_1 + r_2) = \mathcal{O}(h^N)$, N the order of the method. There is a smooth (non-analytic) version of Theorem 5.1, where $r_1 = \mathcal{O}(h^{N-1})$ if ψ_h is of order N [Moa05] section 1, so that the step size h *is* a perturbation parameter. Again, it is not clear to us whether, in the analytic case, there exists similar behavior of h as a perturbation parameter.

5.4 An exact, numerical KAM theorem for symplectic methods

The non-autonomous flow interpolation (Theorem 5.1) leads to an exact, numerical KAM theorem for symplectic integrators.

5.4.1 Exact numerical KAM theorem for symplectic methods of autonomously perturbed completely integrable Hamiltonian systems

Suppose we have a Hamiltonian in action angle coordinates (Section 3.1), where H_0 is real analytic on a simply-connected domain $D \subset \mathbb{R}^n$. We consider in particular an autonomously perturbed, completely integrable Hamiltonian H

$$H(a, \phi) = H_0(a) + H_1(a, \phi).$$

If a symplectic numerical method ψ_h is applied to this system and H can be extended analytically to a complex domain $\mathcal{D}_{\delta_1+\delta_2}$ as in Theorem 5.1, then one can this Theorem to find a Hamiltonian

$$\hat{H}_h(a, \phi, t) = H_0(a) + H_1(a, \phi) + H_2(a, \phi, t) \tag{5.3}$$

so that $\hat{H}_h$ is analytic on $\mathcal{D}_{\delta_1}$ and H_2 is h-periodic in t i.e. $\tilde{H}_1(a, \phi, t + h) = \tilde{H}_1(a, \phi, t)$. In particular, the time-h flow $\Phi_{\hat{H}_h}$ of the ODE with Hamiltonian $\hat{H}_h$ is equal to the numerical method ψ_h Furthermore, if $h > 0$ is small enough then we have explicit bounds on the time-dependent part of H_2 as in Theorem 5.1. One may now apply the KAM theorem 3.9 to arrive at an exact, numerical KAM theorem.

Theorem 5.2. *Consider the perturbed, completely integrable system*

$$L(q, p) = L_0(q, p) + L_1(q, p),$$

written in action-angle coordinates as

$$H(a, \phi) = H_0(a) + H_1(a, \phi).$$

where $L_i(q, p) = H_i(a(q, p), \phi(q, p))$. Suppose a symplectic method ψ_h has been used on the ODE with Hamiltonian L and that L can be extended analytically to a complex set $\mathcal{D}_{\delta_1+\delta_2}$ (defined as in Theorem 5.1) extending the set $D \subset \mathbb{R}^{2n}$ for some $\delta_1, \delta_2 > 0$ so that there exists a Hamiltonian $\hat{L}_h$, given by

$$\hat{L}_h(q, p, t) = L_0(q, p) + L_1(q, p) + L_2(q, p, t)$$

which is h-periodic in t and of which the time-h flow interpolates the numerical method ψ_h. Similarly, in action-angle variables (a, ϕ) we find a Hamiltonian $\hat{H}_h$ given by

$$\hat{H}_h(a, \phi, t) = H_0(a) + H_1(a, \phi) + H_2(a, \phi, t)$$

which is h-periodic in t.

We consider the now the variable b conjugate to time t by extending the system so as to consider the Hamiltonian

$$\hat{K}_h(q, p, t, b) = H_0(a) + H_1(a, \phi) + H_2(a, \phi, t) + \frac{2\pi}{h} b.$$

We write $A = (a, b) \in \mathbb{R}^n \times \mathbb{R}$ and $\Phi = (\phi, t) \in \mathbb{T}^n \times \mathbb{T}$ so $F = F_\rho \subset \mathbb{R}^{n+1} \times \mathbb{C}^{n+1}$ as in Theorem 3.9. can be defined for $\rho > 0$.

Suppose that $\hat{K}$ can be analytically extended to F_ρ for some $\rho > 0$ and suppose as well that H_0 satisfies the non-degeneracy condition

$$\det \frac{\partial^2 H_0(a)}{\partial a^2} = \det \frac{\partial \omega(a)}{\partial a} \neq 0,$$

on F_ρ, where $\omega(a) = \frac{\partial H_0(a)}{\partial a}$.

Then $\forall \epsilon > 0$, $\exists C = C(\mu, \rho, G, H_0)$ such that, if

$$|H_1 + H_2| \leq C, \text{ pointwise on } F$$

then the motion defined by

$$\dot{\Phi} = D_A K(A, \Phi) \qquad \dot{A} = -D_\Phi K(A, \Phi)$$

has the following properties:

- *Re $F = F_1 + F_2$ where F_1 is invariant and the measure of F_2 (denoted $|F_2|$) satisfies $|F_2| \leq \mu |F|$.*
- *F_1 consists of a family of invariant $(n+1)$-dimensional analytic tori $I_\alpha = \{(B_\alpha, \Psi) \mid \Psi \in \mathbb{T}^{n+1}\}$, defined parametrically by*

  $$A = B_\alpha + f_\alpha(\Psi), \qquad \Phi = \Psi + g_\alpha(\Psi),$$

  where $(f_\alpha, g_\alpha) : \mathbb{T}^n \to \mathbb{R}^n$ are analytic and of period 2π in all its variables and the parametrisation depends on $\alpha = (\omega, \Omega)$ [JS96] (because of the non-degeneracy condition, equivalently $\alpha = (a, \Omega)$ as in [Sev07]).
- *The mapping is close to the identity*

  $$|(f_\alpha, g_\alpha)| = \mathcal{O}(\epsilon) \text{ pointwise on } F$$
- *On the invariant $(n+s)$-tori I_α, the motion of the perturbed system is quasi-periodic with frequencies*[22]

  $$\left(\frac{\partial H_0(B_\alpha)}{\partial a}, \Omega \right).$$

Similar to Remark 3.10 one may use the results of [Sev07] to show that time is not transformed.

Proof. From Theorem 5.1 one finds the Hamiltonian

$$\hat{L}_h(q, p, t) = L_0(q, p) + L_1(q, p) + L_2(q, p, t)$$

which is h-periodic in t and of which the time-h flow interpolates the numerical method ψ_h. Written in action-angle coordinate this can be rewriten to

$$\hat{H}_h(a, \phi, t) = H_0(a) + H_1(a, \phi) + \epsilon H_2(a, \phi, t),$$

where $H_i(a, \phi, t) = L_i(q(a, \phi), p(a, \phi), t)$ if the variables (a, ϕ) can be transformed to (q, p) via the symplectic map $(q(a, \phi), p(a, \phi))$ (see also Section 3). Ths $\hat{H}_h$ is h-periodic in t.

The remaining part of this theorem can be proven by applying to $\hat{H}_h$ the KAM theorem 3.9. □

5.4.2 Approach to an exact periodic KAM theorem for symplectic methods on periodically perturbed, completely integrable Hamiltonian systems

In view of an application to the numerically integrated tidal wave system, we consider a periodically perturbed (2π-periodic in t), completely integrable Hamiltonian system in action angle coordinates,

$$H(a, \phi, t) = H_0(a) + H_1(a, \phi, t)$$

defined on $D \times \mathbb{T}^n \subset \mathbb{R}^n \times \mathbb{T}^n$. Suppose a symplectic method ψ_h (symplectic with respect to (a, ϕ) is used to integrate this system. Then we use Theorem 5.1 either on the canonically autonomised space (Definition 2.5) $D \times \mathbb{T}^{n+1}$ or on the extended phase space (Section 2.1.3) $(D \times \mathbb{R}) \times \mathbb{T}^{n+1}$. Doing the latter, we consider the variable s conjugate to time t and the extended Hamiltonian (Equation (2.7))

$$\tilde{H}(a, \phi, t, s) = H_0(a) + H_1(a, \phi, t) + s.$$

[22] The frequencies of the perturbed tori are in some cases equal [ZC10].

and find, if the conditions of Theorem 5.1 are met, a non-autonomously, periodically perturbed (in a new time variabe τ) Hamiltonian which interpolates

$$\tilde{K}(a,\phi,t,s,\tau) = H_0(a) + H_1(a,\phi,t) + s + H_2(a,\phi,t,s,\tau).$$

The advantage of this method is that it is clear from [Moa05], *Note 3* that the $\tilde{K}$ and H_2 are again Hamiltonian. The disadvantage, however, is that H_2 may depend on s, so that in general $\dot{t} \neq 1$, which impedes the use of the KAM Theorem 3.9.

If we use canonically autonomised space, without a conjugate variable s and with the vector field

$$f(a,\phi,t) = \begin{pmatrix} J\nabla_{a,\phi}(H_0(a) + H_1(a,\phi,t)) \\ 1 \end{pmatrix}$$

then, if the conditions of Theorem 5.1 are met, a non-autonomously, periodically perturbed (in a new time variabe τ) vector field

$$\tilde{f}(a,\phi,t,s,\tau) = \begin{pmatrix} J\nabla H_0(a) + J\nabla H_1(a,\phi,t) + s + f_2(a,\phi,t,\tau) \\ 1 + \tilde{f}_2(a,\phi,t,\tau) \end{pmatrix}$$

is found, which interpolates the flow of the symplectic method (on the canonically extended phase space). The advantage now is that, in the smooth case i.e. using the construction of Equation (5.1), the vector field $\tilde{f}_2 = 0$. Therefore it is likely that $\dot{t} = 1$, which we will see implies that $\tilde{f}$ can be seen as a quasi-periodically perturbed ODE. However, it is not immediately clear whether f_2 is Hamiltonian and whether f_2 is again periodic in t with the same period (2π).

As mentioned in Section 5.2, a disadvantage of non-autonomous flow interpolation (as opposed to modified equation analysis) is that the construction is very involved (e.g. [Moa05; KP94b]). Therefore we will make the following assumption on the analytic modified vector field $\tilde{f}$ (as in Theorem 5.1).

Assumption 5.3. *We will assume that $\tilde{f}_2 = 0$. We will assume furthermore that the vector field f_2 is again Hamiltonian in the sense that $f_2(a,\phi,t,\tau) = J\nabla_{a,\phi}H_2(a,\phi,t,\tau)$. We will assume additionally that $\tilde{f}_2$ are periodic in t with the same period (2π).*

It is not clear to the author if they are met in general.

With these assumptions it becomes not too difficult to apply a KAM theorem to this system.

Theorem 5.4. *Consider the 2π-periodically, non-autonomously perturbed completely integrable system*

$$\tilde{H}(a,\phi,t) = H_0(a) + H_1(a,\phi,t)$$

Suppose a symplectic method ψ_h, symplectic with respect to (a,ϕ) is used to integrate the canonically autonomised system (Definition 2.5). If the Hamiltonian $\tilde{H}$ is real analytic on $D \times \mathbb{T}^n$ and can be analytically extended to a domain $\mathcal{D}_{\delta_1+\delta_2}$ as in Theorem 5.1, then there exists a vector field $\tilde{f}$ given by

$$\tilde{f}(a,\phi,t,s,\tau) = \begin{pmatrix} J\nabla H_0(a) + J\nabla H_1(a,\phi,t) + s + f_2(a,\phi,t,\tau) \\ 1 + \tilde{f}_2(a,\phi,t,\tau) \end{pmatrix},$$

which is h-periodic in τ and the time-h flow of this vector field is equal to ψ_h.

If we assume the Assumptions 5.3, then a Hamiltonian $\hat{H} = \hat{H}(a,\phi,t,\tau)$ exists given by

$$\hat{H}(a,\phi,t,\tau) = H_0(a) + H_1(a,\phi,t) + H_2(a,phi,t,\tau),$$

which is 2π-periodic in t and h-periodic in τ such that the vector field $\tilde{f}$ can be written as a vector field $\tilde{f}(a,\phi,t,\tau)$ given by

$$\tilde{f}(a,\phi,t,\tau) = \begin{pmatrix} J\nabla_{a,\phi}\hat{H}(a,\phi,t,\tau) \\ 1 \end{pmatrix},$$

of which the time-h flow (with time $\tau = h$) is still equal to the numerical method ψ_h.

We extend now the system by adding two variables $b = (b_1, b_2) \in \mathbb{R}^2$ conjugate to t, τ so as to consider the Hamiltonian

$$\hat{K}(a, \phi, b, t, \tau) = H_0(a) + H_1(a, \phi, t) + H_2(a, phi, t, \tau) + \langle \Omega, b \rangle,$$

where $\Omega = (1, \frac{2\pi}{h})$. We write $A = (a, b) \in \mathbb{R}^n \times \mathbb{R}$ and $\Phi = (\phi, t) \in \mathbb{T}^n \times \mathbb{T}$ so $F_\rho \subset \mathbb{R}^{n+1} \times \mathbb{C}^{n+1}$ as in Theorem 3.9. can be defined for $\rho > 0$.

Suppose that $\hat{K}$ can be analytically extended to F_ρ for some $\rho > 0$ and suppose as well that H_0 satisfies the non-degeneracy condition

$$\det \frac{\partial^2 H_0(a)}{\partial a^2} = \det \frac{\partial \omega(a)}{\partial a} \neq 0,$$

on F_ρ, where $\omega(a) = \frac{\partial H_0(a)}{\partial a}$.
Suppose furthermore that H_0 satisfies the non-degeneracy condition

$$\det \frac{\partial^2 H_0(a)}{\partial a^2} = \det \frac{\partial \omega(a)}{\partial a} \neq 0,$$

on the complex domain $\mathcal{D}_{\delta_1}$, where $\omega(a) = \frac{\partial H_0(a)}{\partial a}$. And that $\Omega = (2\pi, h)$ is (σ, γ)-Diophantine i.e.

$$\langle k, (\Omega) \rangle \geq \frac{\gamma}{\|k\|_1^\sigma}, \quad \forall k \in \mathbb{Z}^2 - \{0\}.$$

for some $\sigma > 2$.

Then $\epsilon\mu > 0$, $\exists C = C(\epsilon, \delta_1, \mathcal{D}, H_0)$ such that, if

$$|H_1(a, \phi, t) + \epsilon H_2(a, phi, t, \tau)| < C, \text{ pointwise on } \mathcal{D}_{\delta_1} \times \mathbb{T}^2$$

then, denoting $A = (a, b) \in \mathbb{R}^n \times \mathbb{R}^2$ and $\Phi = (\phi, t) \in \mathbb{T}^n \times \mathbb{T}^2$, there exists a set $F \subset \mathcal{D}_{\delta_1} \times \mathbb{T}^n$ such that

- *Re $F = F_1 + F_2$ where F_1 is invariant with respect to the flow of $\hat{H}$ and $|F_2| \leq \mu|F|$.*
- *F_1 consists of a family of invariant $(n+1)$-tori $I_\epsilon = \{(B_\epsilon, \Psi), \mid \Psi \in T^{n+1}\}$, defined parametrically by*

  $$A = B_\epsilon + f_\epsilon(\Psi), \qquad \Phi = \Psi + g_\epsilon(\Psi),$$

  where f_ϵ, g_ϵ are analytic of period 2π in all its variables and the parametrisation depends on $\epsilon = (\omega, \Omega)$ [JS96] (because of the non-degeneracy condition, equivalently $\epsilon = (a, \Omega)$ as in [Sev07]).
- *The mapping (f_ϵ, g_ϵ) close to the identity*

  $$|(f_\epsilon, g_\epsilon)| = \mathcal{O}(\epsilon) \text{ on } F$$

- *On the invariant $(d+n)$-tori I_ϵ, the motion of the perturbed system is quasi-periodic with frequencies*

  $$\left(\frac{\partial H_0(B_\epsilon)}{\partial a}, \Omega \right).$$

Proof. The proof is similar to the one of Theorem 5.2. □

Remark 5.5. *Again, the result by [Sev07] could be used, as in Remark 3.10, to find that t, τ are left unchanged in the new coordinates i.e. $\Psi = (\psi, t, \tau)$ and that $\dot{t} = \dot{\tau} = 1$.*

This result together implies, if all the conditions and the Assumptions 5.3 are met, the existence of invariant tori in the extended system with variables (a, b, ϕ, t, τ).
If $\frac{2\pi}{h} = N \in \mathbb{N}$ then these conditions are *not* met, so we cannot use Theorem 5.4 to conclude that invariant tori of the numerically approximated 2π-Poincaré map ψ_h^N exist, when integrating the periodically perturbed system.

In the theoretical system described by the Hamiltonian $\tilde{H}(a, \phi, t) = H_0(a) + H_1(a, \phi, t)$ one can prove the existence of invariant tori in the Poincaré map $\phi_{\tilde{H},t_0+2\pi,t_0}$, Section 3.3. However, if one wants to 'see' these invariant tori by numerically integrating the system with a symplectic method ψ_h, then one must choose $\frac{2\pi}{h} = N \in \mathbb{N}$ so as to approximate the Poincaré map i.e. $\psi_h^N \approx \frac{2\pi}{h} = N \in \mathbb{N}$. Thus, Theorem 5.4 cannot be used to conclude the existence of invariant tori in the symplectic integrated tidal wave system (since we choose $\frac{2\pi}{h} \in \mathbb{N}$).

5.5 KAM tori in the Poincaré map of the unperturbed tidal wave system

We have discussed that we cannot use Theorem 5.4 to prove the existence of KAM tori in the numerically approximated 2π-Poincaré map, when symplectically integrating the periodically *perturbed* tidal wave system.

However, we may use 5.2 to prove the existence of KAM tori (locally) in the numerically approximated 2π-Poincaré map, when symplectically integrating the periodically *unperturbed* tidal wave system, which is done below and similar to [MO10]. Again, a problem is that we are not sure that h is a perturbation parameter.

Indeed, we have shown already in Section 3.3 that the unperturbed system with Hamiltonian $K(q, p) = \cos(p) + \eta \cos(q)$ is completely integrable and can be analytically extended to a complex set domain $\mathcal{D}_{\delta_1+\delta_2}$ for any such domain. Therefore, the perturbed Hamiltonian $\hat{H}_h$ as in Theorem 5.2 exists and, for sufficiently small h, the time-dependent part of $H_2(a, \psi, t)$ is very small (from Theorem 5.1).

Furthermore we know that $\hat{H}_h$ is analytic on the domain $\mathcal{D}_{\delta_1}$ and therefore also on F_ρ (as in Theorem 3.9) for $\rho > 0$ small enough and that H_0 satisfies the Kolmogorov non-degeneracy condition. Moreover, since the frequency of the perturbation equals $\Omega = 2\pi/h$, it is trivially (σ, γ)-Diophantine for some choice of $\gamma > 0$, $\sigma > 0$.

Therefore, we may use, as in Section 3.3, Theorem 5.2 to conclude that there exist KAM tori in the extended system $D \times \mathbb{T}^2$ (for the action angle coordinates $(a, \phi) \in D \times \mathbb{T}$) and use a remark similar to Remark 3.10 to conclude that there exists invariant tori in the numerically approximated 2π-Poincaré map **if the perturbation size is small enough**.

It is not clear if the step size can be used as a perturbation parameter, since only the time-dependent part of H_2 (as in Theorem 5.2) can be made small by decreasing the step size. This fact does give an indication that one could expect to see more KAM tori if one makes decreases the step-size but we need a bound on the time-independent of H_2 with respect to h to make rigorous statements.

6 Modified equation analysis and approximate KAM theory applied to the tidal wave system

In this section, modified equation analysis (MEA) is considered and an 'approximate' numerical KAM theorem is constructed. As mentioned in Section 5.2, MEA is a type of BEA which is an alternative to the non-autonomous flow interpolation considered in Section 5.

Furthermore, besides the fact that the exact numerical KAM theorem of Section 5 is not able to prove the existence of invariant tori in numerically approximated 2π-Poincaré map of the perturbed tidal wave system, another reason to consider such an approximate, numerical KAM theorem is round-off error. Indeed, round-off error and (exact) KAM theory do not behave well together:

> "Further difficulties [in the verification of KAM theory for symplectic integrators] are introduced by the finite precision effects of computers, guaranteeing that all steps are resonant. Furthermore, it has also been numerically verified that roundoff effects are sufficient for a trajectory to cross a torus, whereby showing that invariant tori are not strict barriers to the flows approximated by [the symplectic integrator] on computers." – Moan [Moa03]

Therefore, even if a proof was given of the existence of KAM tori in the numerically approximatesd 2π-Poincaré map of the perturbed tidal wave system, round-off error could have a negative impact on the 'invariance' of these tori. A partial solution to this is the 'approximate' KAM theorem which is presented in this Section. This approximate KAM theorem is based on [HLW06] chapter IX.7 and X.5.

Furthermore, Moan [Moa03] mentions that, similar to the exact numerical KAM Theorem 5.2, the orbits/invariant tori which are not behaving 'invariantly', when numerically integrating a (perturbed) completely integrable system, are due to resonances between the frequencies of the torus and the step-size h.

> "A more careful analysis of the close-to-integrable Hamiltonian systems [...] shows that the exponential remainder term [...] originates from the step-size resonances, i.e. when some oscillatory component of the exact solution has period equal to the step-size." – Moan [Moa03]

Therefore we consider MEA and construct an approximate KAM theorem as in [HLW06]

6.1 Lie-Gröbner series and splitting of vector field for autonomous and Hamiltonian ODE

In this section, *Lie-Gröbner series* (often *Lie Series*) will be introduced. Lie series are a characterisation of the flow $\phi_{f,t}(y)$ in the form of a formal power series, therefore generalising the matrix exponential map to the non-linear setting. For autonomous ODE, the Lie-Gröbner series are presented in this Section. For non-autonomous ODE, we refer to Appendix D. For smooth vector fields, this power series in general diverges and is therefore only a *formal power series*. For analytic vector fields they are locally convergent.

Using Lie series, expressions for the modified vector fields can be found. Additionally, the error made when splitting vector fields can be expressed in a Lie algebraic sense using the Baker-Campbell-Hausdorff formula and Lie series are therefore useful to study numerical splitting methods.

6.1.1 The exponential map as flow of linear systems

For linear systems the exponential comes in very naturally.

Scalar, linear, autonomous IVP:

For a system

$$\dot{y}(t) = \lambda y(t), \qquad y : \mathbb{R} \to \mathbb{R},\, \lambda \in \mathbb{R},$$

the flow is analytic and given by the function $\phi_t(y) = e^{\lambda t} y$, the exponent $e^{\cdot}$ being defined in a plethora of ways.

Systems of linear, autonomous ODE:

For a system

$$\dot{y}(t) = Ay(t), \qquad y : \mathbb{R} \to \mathbb{R}^n,\, A \in \mathbb{R}^{n \times n},$$

the flow is analytic and given by the function $\phi_t(y) = \exp(tA)y$, where $\exp(tA)$ is the matrix exponential, which is often defined using power series, the Laplace transform or, if the field is $\mathbb{C}$, via the Cauchy-integral.

parabolic linear PDE:

For a system

$$\dot{y} = \mathcal{A}y \qquad y \in \mathcal{F},\, \mathcal{A} \in \mathrm{Lin}(\mathcal{F},\mathcal{G}),$$

where $\mathcal{F},\mathcal{G}$ are appropriate function spaces and $\mathcal{A}$ is some sufficiently regular (e.g. sectorial), elliptic, linear operator, then $\phi_t(y) = \exp(t\mathcal{A})y$, ($y$ a function), defined using for example Laplace transform (in general not via power series) [ENB00].

Thus, for **linear** ODE, or (in the infinite-dimensional case) parabolic, **linear** PDE, the flow is denoted by the exponential function

$$\phi_{f,t}(y) = \exp(tf) \tag{6.1}$$

where f is the corresponding linear operator.

6.1.2 Lie derivative

In the case of systems on Euclidean space a nonlinear generalisation of Equation (6.1) exists, which is a power-series representation and is called the *Lie series* or *Lie-Gröbner series* [GK67; Grö67; Ste84; Ste86].

To this purpose we introduce the Lie derivative, as noted in [Ste86]: "Some of the results [...] need to be translated to the Hamiltonian [and non-linear] setting. This is done by replacing operators (e.g., A) by Lie derivatives (e.g., $[f,\cdot]$).". For a vector field $f \in C^\infty(\mathbb{R}^n,\mathbb{R}^n)$, the Lie derivative $L_f : C^\infty(\mathbb{R}^n,\mathbb{R}^m) \to \mathbb{C}^\infty(\mathbb{R}^n,\mathbb{R}^m)$ is defined by

$$L_f(g) := \lim_{t\to 0} \frac{g \circ \phi_{f,t} - g}{t} = \frac{d}{dt} g(\phi_{f,t})|_{t=0} \tag{6.2}$$

(the limit defined pointwise), which for Euclidean space reduces to

$$L_f(g)(x) = \nabla g(x) \cdot f(x) = Dg(x)(f(x))$$

One finds that

1. L_f is linear;
2. If g is a scalar function then $L_{gf} = gL_f$;
3. L_f satisfies Leibniz rule $L_f(gh) = gL_f(h) + L_f(g)h$ (where multiplication is defined component-wise in $\mathbb{R}^m$ for the vector-valued case);
4. (1) and (3) above imply that L_f is a derivation, such that we may also write $D_f := L_f$, where D_f in this case generalises the directional derivative: if f a constant function equal to $f(0) = v$, then $D_f g =: D_v g$ where D_v is the directional derivative. In particular

$$D_f = \sum_{j=1}^{n} f_j \frac{\partial}{\partial x_j}.$$

Additionally $L_f^j = D_f^j$ and for $g = Id$

$$\frac{d^j}{dt^j}\phi_t|_{t=0} = D_f^j(Id) \tag{6.3}$$

in particular, in the case of linear f we see that $D_f^j(Id) = f^n$ and $\sum_{j\geq 0} \frac{h^j}{j!} D_f^j(Id) = e^{tf}$ cf. Equation (6.1).

6.1.3 Lie-Gröbner series for autonomous nonlinear ODE

The Lie derivative is now used to generalise Equation (6.1) to nonlinear f. For analytic flows $\phi_t(y)$ at $t = 0$, one finds using the Taylor series for small t

$$\phi_t(y) = \left(\sum_{j=0}^{\infty} \frac{t^j}{j!} \left(\frac{d^j}{dt^j} \phi_0(y) \right) \right). \tag{6.4}$$

Which suggests heuristically that $\phi_s = \left(\exp\left(s \frac{d}{dt}\right) \phi_t|_{t=0} \right)$ with $\exp(s\frac{d}{dt}) = \sum_{j=1}^{\infty} \frac{1}{j!} \frac{d^j}{dt^j}|_{t=0}$. From Equation (6.3) we now find that

$$\frac{1}{j!} \frac{d^j}{dt^j} \left((e^{t\frac{d}{ds}} \phi_s|_{s=0}) \right) = D_f^j(Id), \quad \text{or more generally} \quad \frac{1}{j!} \frac{d^j}{dt^j} \left((e^{t\frac{d}{ds}} g(\phi_s)|_{s=0}) \right) = D_f^j(g)$$

for $g \in C^\infty(\mathbb{R}^n, \mathbb{R}^m)$, which implies that, formally,

$$g(\phi_{f,t}(y)) = (e^{tD_f} g)(y) \implies \phi_{f,t}(y) = (e^{tD_f} Id)(y) = (e^{tL_f} Id)(y). \tag{6.5}$$

Definition 6.1. Given a smooth vector field, f on $\mathbb{R}^n$. The formal power series

$$e^{tD_f} = \sum_{j=0}^{\infty} \frac{t^j}{j!} D_f^j$$

is called the *Lie-Gröbner series* for the (ODE with vector field) f.

In the special case that f and g are real-analytic (or holomorphic), the formal power series are also at least locally convergent (and equal the flow) [GK67; Grö67; Ste84].

6.1.4 Lie-Gröbner series for forced ODE

Like in Section 4.1.2, the Lie series of forced ODE can be found via Faà di Bruno's formula for the higher order chain rule: For $x \in C^\infty(\mathbb{R}), y \in C^\infty(\mathbb{R}, \mathbb{R}^n)$ one has (again $x_i = D^i x$)

$$\frac{d^i}{dt^i} y(x(t)) = \sum_{j=1}^{i} B_{i,j}(x_1, \ldots, x_{n-k+1})(D^j y(x(t))).$$

Combining this equation with the fact that $\phi_{gf,t,t_0} = \phi_{f,G(t)}$ where $G(t) = \int_{t_0}^t G(t)$ (Equation (2.5)), one finds

$$\frac{d^i}{dt^i} \phi_{gf,t,t_0}|_{t=t_0} = \sum_{j=1}^{i} B_{i,j}(g_0, \ldots, g_{n-j})\, D_s^j \phi_{gf,s,t_0}|_{s=G(t_0)} = \sum_{j=1}^{i} B_{i,j}(g_0, \ldots, g_{n-j})\, D_f^j(Id).$$

Then one finds the Lie Series

$$\phi_{gf,t,t_0} = Id + \sum_{i\geq 1} \frac{t^i}{i!} \sum_{j=1}^{i} B_{i,j} D_f^j(Id) = Id + \sum_{i\geq 1} \frac{t^i}{i!} \sum_{j=1}^{i} \left(\sum_{l \in p_j(i)} \frac{i!}{l! \prod_{k=1}^{i} (k!)^{l_k}} \prod_{\ell=1}^{i} g_{i-1}^{l_i} \right) D_f^j(Id), \tag{6.6}$$

where notation as in Section B.1 was used ($p_j(i)$ denoting partition of the integer i into j integers).

6.1.5 Lie-Gröbner series for Hamiltonian ODE and Poisson brackets

For Hamiltonian ODE with Hamiltonian H, the symplectic structure can be used to get a different characterisation of Lie series. The derivation in this case becomes $D_f = \sum_i f_i \frac{\partial \cdot}{\partial x_i} = (\nabla \cdot)^T (J \nabla H)$. In other words one finds, for $g : \mathbb{R}^{2n} \to \mathbb{R}$ that

$$\frac{d}{dt} g(\phi_{H,y_0}(t)) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} \right) = \{g, H\}.$$

where the Poisson bracket $\{\cdot,\cdot\}$ is used. Combining the last equation and Equation (6.2), one finds that $D_{\nabla H} := L_{\nabla H} := -\{H,\cdot\}$ is a derivation/vector field and induces the Lie derivative $L_{\nabla H} := \{\cdot, H\} : C^\infty(\mathbb{R}^{2n}) \to C^\infty(\mathbb{R}^{2n})$. One then finds the Lie-Gröbner series

$$\phi_{H,t}(q,p) = (e^{t\{\cdot,H\}}Id)(q,p). \tag{6.7}$$

6.2 Splitting of the vector field and the BCH formula

Besides numerical splitting methods, one might be interested in a theoretical splitting of the vector field: If $f = g+h$ then how are the (exact/theoretical) flows ϕ_f, ϕ_g, ϕ_h related? In particular, can we compose ϕ_g, ϕ_h so that we can get accurate approximations to ϕ_h? The Baker-Campbell-Hausdorff (BCH) formula is helpful to answer the latter question and is useful for BEA of splitting methods.

6.2.1 Baker-Campbell-Hausdorff formula for linear and non-linear ODE

We first consider linear, autonomous ODE of the form

$$\dot{y}(t) = (A+B)y(t)$$

where $A, B \in \mathbb{R}^{n\times n}$, with solution $\phi_{A+B,t}(y) = e^{t(A+B)}y$. When splitting the vector field $A+B$ into the vector fields $g_1 = A$ and $g_2 = B$, the composition of the solutions to g_1, g_2 gives

$$\phi_{B,t} \circ \phi_{A,t}(t) = e^{tB}e^{tA}.$$

As mentioned above, we are interested in the order of accuracy i.e. the difference $e^{tB}e^{tA} - e^{t(A+B)}$. Taylor expanding both sides one finds

$$e^{tB}e^{tA} - e^{t(A+B)} = \frac{t^2}{2}(A^2 + 2AB + B^2 - (A+B)^2) + \mathcal{O}(t^3)$$

so that the order is $\mathcal{O}(t^2)$ with coefficient $[A,B] := AB - BA$.

The order of accuracy can also be found using the *Baker-Campbell-Hausdorff (BCH)* formula which determines a matrix $C(t)$ such that $e^{tB}e^{tA} = e^{C(t)}$. The matrix C is determined [HLW06] by the differential equation

$$\dot{C}(t) = A + B + \frac{1}{2}[A-B, C(t)] + \sum_{k\geq 2} \frac{\mathcal{B}_k}{k!}\mathrm{ad}^k_{C(t)}(A+B), \tag{6.8}$$

where $\mathcal{B}_k$ are the Bernoulli numbers, $ad_C = [C,\cdot]$ and ad^k_C the k-th composition of ad_C. Using Equation (6.8) and the ansatz $C = \sum_i t^i C_i$ one finds the terms

$$\begin{aligned} C_1 &= A + B \\ C_2 &= \frac{1}{2}[A,B] \\ C_3 &= \frac{1}{12}([A,[A,B]] + [B,[B,A]]) \\ C_j &= \frac{1}{2}[A-B, C_{j-1}] + \sum_{k=2}^{j-1} \sum_{\ell \in p_k(j-1)} \frac{\mathcal{B}_k}{k!}\mathrm{ad}_{C_{l_1}} \dots \mathrm{ad}_{C_{l_k}}(A+B)\dots \end{aligned} \tag{6.9}$$

where the partitions $p_k(j-1)$ are used (Section B.2). Other equations exist for the matrix $C(t)$, such as the (explicit) Dynkin equation, but the expressions for the lower order C_i suffice in this thesis.

We now generalise these equations to the non-linear case, using again the Lie derivatives: The BCH formula can be (formally) constructed for any Lie algebra. This includes the Lie algebra of derivations D_f (or equivalently Lie derivatives L_f) which implies, similar to the matrix version, that

$$\phi_{g_2,s} \circ \phi_{g_1,t} = \exp(tD_1)\exp(sD_2)Id = \exp(\tilde{D})Id, \tag{6.10}$$

(note the switching of the order or "Vertauschungssatz" as Gröbner calls it [GK67; Grö67]), where

$$\tilde{D} = sD_1 + tD_2 + \frac{st}{2}[D_1, D_2] + \mathcal{O}(s^2 t + t^2 s), \tag{6.11}$$

and the higher order terms in s, t consist again of higher order Lie brackets of D_1, D_2. These series are only formal, but can be used numerically if truncated, see [HLW06] chapter III.5.1. The Lie derivative $\tilde{D}$ can be seen as an *interpolative/modified Lie derivative.*

As for numerical splitting methods, more general theoretical splittings may be considered: Splitting into more vector fields $f = \sum_{i=1}^{n} f_i$ or using different compositions. For example, the Strang-splitting

$$\phi_{g_2, \frac{t}{2}} \circ \phi_{g_1, t} \circ \phi_{g_2, \frac{t}{2}} = \exp(\tilde{D}(s,t)) Id \tag{6.12}$$

may be used. The BCH formula can be first applied $\phi_{g_2, \frac{t}{2}} \circ \phi_{g_1, \frac{t}{2}}$ and then to $\phi_{g_1, \frac{t}{2}} \circ \phi_{g_2, \frac{t}{2}}$ to find that (setting $t = s$)

$$\tilde{D}_1 = D_{g_1} + D_{g_2}, \quad \tilde{D}_2 = 0 \quad \tilde{D}_3 = -\frac{1}{24}[D_{g_1}, [D_{g_1}, D_{g_2}]] + \frac{1}{12}[D_{g_2}, [D_{g_2}, D_{g_1}]] \tag{6.13}$$

and due to to the symmetric composition, $C_i = 0$ for all even i so that the order of accuracy is $\mathcal{O}(t^3)$.

6.2.2 BCH for time-affine and Hamiltonian ODE

For a non-autonomous vector field $\hat{f}$ we may look at the canonical autonomous extension $(\hat{f}, 1)$ (Appendix D.1.1). One now splits the temporal part as well. For example, for the splitting $\hat{f} = \tilde{f}_1 + \tilde{f}_2 = (f_1, a_1) + (f_2, a_2)$ (with $a_1 + a_2 = 1$) one may use the Lie-series with $D_{\tilde{f}_i} = D_{f_i} + D_{a_i t}$ to find the error in the spatial part:

$$[D_{\tilde{f}_1}, D_{\tilde{f}_2}] == [D_{f_1}, D_{f_2}] + \frac{1}{2}\left(\frac{\partial (f)_j}{\partial t} - \frac{\partial f_j}{\partial t}\right)\frac{\partial}{\partial x_j} = [D_{f_1}, D_{f_2}] - a_2 D_{\dot{f}_1} + a_1 D_{\dot{f}_2}, \tag{6.14}$$

Remark 6.2. *We consider a time-affine ODE with vector field $\sum_{i=1}^{n} g_i(t) f_i(y)$ split into n forced ODE $g_i f_i$. We now show that the modified Lie derivative $\tilde{D}$ is a power series, in the step sizes, with time-affine vector fields/Lie derivatives as coefficients.*

We consider first the case $n = 2$, the splitting $(g_1 f_1, a_1) + (g_2 f_2, a_2)$ where $a_1 + a_2 = 1$ split time and the Lie-Trotter splitting

$$\phi_{g_1 f_1, t_1} \circ \phi_{g_2 f_2, t_2}.$$

In this case Equation (6.9) implies that higher order terms of the modified Lie derivative $\tilde{D}$ consist of higher order Lie brackets. Combining this with the fact that (6.14) implies that the Lie bracket of two 'time-affine' Lie derivatives produces a time-affine Lie derivative we may argue inductively that the modified Lie derivative $\tilde{D}$ is a power series, in t_1, t_2, with time-affine vector fields/Lie derivatives as coefficients. In other words, writing

$$\tilde{D} = \sum_{i,j \geq 0} t_1^i t_2^j D_{i,j}$$

we find that $D_{i,j}(Id)$ are time-affine vector fields.

Moreover, this holds for more general n, a_i and more general splitting methods than the Lie-Trotter splitting. The importance of this lies in modified equation analysis. This will imply that the modified equation is again time-affine, so that, in the Hamiltonian case (treated directly below), it can be seen as a periodically perturbed Hamiltonian system, on which the KAM theorem 3.9 can be used.

For Hamiltonian ODE with Hamiltonian $H = H_1 + H_2$ with vector fields f_{H_1}, f_{H_2}, the Lie-bracket of vector fields satisties $[D_{f_{H_1}}, D_{f_{H_2}}] = \{H_2, H_1\}$ such that

$$[D_{H_1}, D_{H_2}]g = D_{H_1}D_{H_2}g - D_{H_2}D_{H_1}g = \{\{g, H_2\}, H_1\} - \{\{g, H_1\}, H_2\} = \{g, \{H_2, H_1\}\}$$

where the Jacobi identity was used. So, one may replace $[D_{H_1}, D_{H_2}]$ by $\{H_2, H_1\}$ in the BCH formula. Thus, one finds the formula

$$\phi_{t,H_2} \circ \phi_{s,H_1} = e^{t\{\cdot,H_2\}}e^{s\{\cdot,H_1\}} = e^{\{\cdot,\tilde{H}(s,t)\}}$$

with $\tilde{H} = \sum_{i,j=1}^{\infty} t^i s^j \tilde{H}_{i,j}$ given by

$$\tilde{H}_{1,0} = H_1, \quad \tilde{H}_{0,1} = H_2, \quad 2\tilde{H}_{1,1} = [D_{H_2}, D_{H_1}] = \{H_1, H_2\}, \quad \ldots,$$

cf. Equation (6.11). As in the nonlinear case, for different composition methods or splitting into multiple terms, similar equations can be found by applying the BCH formula recursively, see for example [HLW06] chapter III.5.4, III.5.5.

6.3 Modified equation analysis

In much of the mathematical literature about modified equation analysis (MEA) (e.g. [CMS94; Moi10]) it is stated that MEA found its first applications in the study of linear PDE, as in [WH74], in which applications for nonlinear PDE are also discussed at the end. Here we apply it to autonomous ODE and IVP.

6.3.1 Formal modified equation analysis of (non-)autonomous ODE

As introduced in Section 4.2, MEA finds a formal expression for a vector field $\tilde{f}$ such that, for any fixed step size $h > 0$, the time-h flow $\phi_{\tilde{f},h}$ is equal to the numerical method ψ_h (up to some order $M \in \mathbb{N}$). The expression is formal: It is an asymptotic power series which generally diverges for non-linear ODE [LR04; HLW06].

Suppose a consistent numerical method $\psi_h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ of order N has been used on the ODE with vector field $f : \mathbb{R}^n \times \mathbb{R} \supset D \times I \to \mathbb{R}^n \times \mathbb{R}$ (thus $\phi_h - \psi_h = \mathcal{O}(h^{N+1})$), of the form

$$\psi_h(y, t) = y + hf(y, t) + \sum_{i=2}^{\infty} \frac{h^i}{i!} d_i(y, t), \tag{6.15}$$

see Remark 4.3. Given $M > N$ and step size $h > 0$, MEA constructs a vector field $\tilde{f}_h^{[M]} : D \times I \to \mathbb{R}^n$ such that ψ_h is a numerical method of order M with respect to the ODE with vector field $\tilde{f}_h^{[M]}$. Note, the case $M \leq N$ is not interesting as then $\tilde{f}_h^{[M]} := f$ is *the* interpolant, unique in the set of C^M interpolants.

Remark 6.3. *If $f, \tilde{f}_h^{[M]} \in C^M(D, \mathbb{R}^n)$ as above and ψ_h is consistent, then the uniqueness of the Taylor expansion of Equation* (6.16) *up to order $M + 1$ in h implies that the interpolant $\tilde{f}_h^{[M]}$ of order M is unique (in the set of C^M vector fields).*

Definition 6.4. The vector field $\tilde{f}_h^{[M]}$ (or the flow $\phi_{\tilde{f}_h^{[M]}}$) depending on the step size $h > 0$, *interpolates to order M the numerical method ψ_h* if the associated (non-)autonomous flow $\tilde{\phi}_t^{[M]} := \phi_{\tilde{f}_h^{[M]},t} : D \times I \to \mathbb{R}^n$ or integral curve v (with $v(t_0) = y_0$) satisfies

$$\tilde{\phi}_{t_0+h,t_0}^{[M]}(y_0) - \psi_h(y_0, t_0) = \mathcal{O}(h^{M+1}) \text{ (pointwise)} \quad \text{ or } \quad v(t_0 + h) - \psi_h(y_0, t_0) = \mathcal{O}(h^{M+1}) \tag{6.16}$$

for all $t_0, y_0 = v(t_0)$ in the domain of the solution. Then $\tilde{f}_h^{[M]}$ is called the *modified/interpolative vector field of order M (with step-size $h > 0$).* If a vector field $\tilde{f}_h : D \times I \to \mathbb{R}^n$ interpolates f up to order M for all $M \in \mathbb{N}$ then $\tilde{f}_h$ interpolates ψ_h *exactly.* ∅

Suppose that for fixed step size $h > 0$, the flow $\tilde{\phi}_h^{[M]} =: \tilde{\phi}^{[M]} : (\mathbb{R}^n \times \mathbb{R}) \times \mathbb{R} \to \mathbb{R}^n$ interpolates to order M the numerical method ψ_h used on the ODE with vector field f, such that

$$\sum_{j=0}^{\infty} \frac{h^j}{j!}\left(\tilde{\phi}^{[M]}_{t,\tilde{t}_0} - d_j(\cdot,t)\right) = \mathcal{O}(h^{M+1}), \quad \text{or} \quad \sum_{j=0}^{\infty} \frac{h^j}{j!}\left(\frac{d^j v}{dt^j}(\tilde{t}_0) - d_j(\tilde{y}_0,\tilde{t}_0)\right) = \mathcal{O}(h^{M+1}). \tag{6.17}$$

If we assume that $D_t^j \tilde{\phi}^{[M]}_{t,\tilde{t}_0}|_{t=t_0} = \mathcal{O}(1)$ for all $j \in \mathbb{N}$ then equivalently

$$\sum_{j=0}^{M} \frac{h^j}{j!}\left(D_t^j \tilde{\phi}^{[M]}_{t,\tilde{t}_0} - d_j(\cdot,t)\right)|_{t=\tilde{t}_0} = 0, \quad \text{or} \quad \sum_{j=0}^{M} \frac{h^j}{j!}\left(D_t^j v(\tilde{t}_0) - d_j(\tilde{y}_0,\tilde{t}_0)\right) = 0, \tag{6.18}$$

which is an ODE of order M, thus needing $M-1$ extra initial conditions $\tilde{y}_0^{[j]}$. Existence of $\tilde{\phi}^{[M]}$ is guaranteed for 'sufficiently small' $h \in \mathbb{R}$ from ODE theory (for example [DE02] section 2.6 and 2.7).

Definition 6.5. Given a numerical method ψ_h (with step-size $h > 0$ fixed). The *modified equation (of order M)* is the M-th order ODE given by Equation (6.18). ∅

The modified equation of a numerical method is the basis for MEA. In particular, it induces two different ways to solve the modified equation. First on the level of the flow $\tilde{\phi}^{[M]}$ or integral curve $v^{[M]}$ with modified vector field $\tilde{f}^{[M]} := D_t\tilde{\phi}^{[M]}$, treated immediately below and in Section 6.3.2. Second on the level of vector fields: Replacing $D_t^j\tilde{\phi}^{[M]} = D_t^{j-1}\tilde{f}^{[M]}$ and trying to find equations which determine $\tilde{f}^{[M]}$ explicitly in terms of d_j and f, treated in Sections 6.3.3 and 6.3.4.

To find the modified equation, it was assumed that $D_t^j\tilde{\phi}^{[M]} \in \mathcal{O}(1)$. However, this is not true in general as is shown in the following example, where $N = 1$ and $M = 2$ and behaviour is of order $\mathcal{O}(\frac{1}{h})$. Consequently, the local error is not of order $\mathcal{O}(h^3)$. and the error after n-steps i.e. $\phi_{\tilde{f},nh} - \psi_h^n$ might behave badly.

Example 6.6. *Consider the autonomous, scalar, linear initial value problem [GS86],*

.

$$\begin{aligned} y_0 &= 0, t_0 = 0 \\ v'(t) &= f(v(t)) = \lambda v(t), \end{aligned}$$

for $\lambda \in \mathbb{R}$, on which forward Euler (first order, $N = 1$) is applied: $\psi_h(U) = U + hf(U) = U + h\lambda U$. The modified equation, Equation (6.18), with $M = 2$, together with a perturbed initial condition $\tilde{y}_0$ and an extra initial condition $\tilde{y}^{[2]}$ is then given by

$$\begin{gathered} \tilde{y}_0 := U_0, \quad \tilde{y}_0^{[2]} := U_1, \\ v' + \frac{h}{2}v'' = \lambda v, \end{gathered}$$

which has solution

$$v(t) = c_1 e^{tr_-} + c_2 e^{tr_+}$$

where $r_\pm = (\pm\sqrt{1 + 2\lambda h} + 1))/h$. Thus it is clear that, if U_0, U_1 are not such that $c_2 = 0$ then the r_+ term causes behavior $v' = \mathcal{O}(1/h)$ (so not $\mathcal{O}(1)$). Thus $\mathcal{O}(h^3)$ behaviour of Equation (6.17) is not attained, as is shown for a similar example in [GS86].

6.3.2 Formal modified equation via differentiation and substitution

To solve the problem encountered in Example 6.6, approximations of the derivatives $D^j\tilde{\phi}^{[M]}$ which are $\mathcal{O}(h^{M+1-j})$ are given next. The start is again Equation (6.17) rewritten as

$$D_t\phi_t = d_1 + \sum_{j=2}^{M} \frac{h^{j-1}}{j!}\left(d_j - D_t^j\phi_t\right) + \mathcal{O}(h^M), \text{ or } v' = d_1(\tilde{y}_0,\cdot) + \sum_{j=2}^{M} \frac{h^{j-1}}{j!}\left(d_j(\tilde{y}_0,\cdot) - \frac{d^j v}{dt}\right) + \mathcal{O}(h^M) \tag{6.19}$$

and approximations can be found by differentiation of this equation and substitution into itself [GS86; Moi10], as shown in Example 6.7. This method of substitution actually reduces the modified equation to a first order ODE, thus removing as well the problem of finding extra initial conditions.

Example 6.7. *Continuing Example 6.6 we try to find an order $M = 3$ interpolant, such that Equation (6.19) becomes,*

$$v' + \frac{h}{2}v'' + \frac{h^2}{6}v''' = f(v) + \mathcal{O}(h^3).$$

Differentiating $1 \leq i \leq 2$ times one obtains

$$v'' = -\frac{h}{2}v''' + f'(v)v' + \mathcal{O}(h^2) \tag{6.20}$$

$$v''' = (f''(v)v'v' + f'(v)v'') + \mathcal{O}(h). \tag{6.21}$$

Substituting Equation (6.21) into Equation (6.20) then (6.20) into itself on the RHS we find an expression for v'' in terms of v' and v. This expression can be substituted into equation (6.21) thereafter substituting all into the original equation such that a first order ODE can be obtained. This method, and the order of substituting the approximations, is described as well in, for example, [KP94a; AC97] where in the latter an implementation in Maple is presented.

After substitution one finds the modified ODE

$$v' = f(v) - \frac{h}{2}f'(v)v' + \frac{h^2}{12}(f''(v)v'v' + f'(v)v'') + \mathcal{O}(h^3) \tag{6.22}$$

This examples shows that the modified equation, obtained by differentiating and repeated substitution, is of the form

$$v' = \tilde{f}^{[M]} \quad \text{with} \quad \tilde{f}^{[M]} = f + \sum_{j=2}^{M} h^{j-1} f_j \tag{6.23}$$

with terms $f_j : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$. This holds in general and can be used as an ansatz to find the modified ODE in a different way, shown in Section 6.3.3.

Thus, differentiating Equation (6.19), this type of MEA produces the terms $D^i v$ $(1 \leq i \leq M)$ up to $\mathcal{O}(h^{M-i})$ of the form (with $D_{\tilde{f}} d_i = \frac{d}{dt} d_i(v(t))$ the Lie derivative)

$$h^{i-1} D^i v = \sum_{j=i+1}^{M} \frac{-h^{j-1}}{(j-(i-2))!} D^j v + \sum_{j=i}^{M} \frac{h^{j-1}}{(j-(i-1))!} D_{\tilde{f}}^{i-1} d_{j-i+1} + \mathcal{O}(h^M) \tag{6.24}$$

for one-step methods of the form of Equation (6.15). As in Example 6.7, the equation for $2 \leq i \leq M$ can be substituted (with order increasing in i) in the original equation ($i = 1$) to find the modified equation. For $M = 2$, in other words after substituting $\frac{1}{2}hD^2 v$ into the equation we find the equation

$$v' + \sum_{j=3}^{M} h^{j-1}\left(\frac{1}{j!} - \frac{1}{2(j-1)!}\right) D^j v = \sum_{j=1}^{M} \frac{h^{j-1}}{j!} d_j - \sum_{j=2}^{M} \frac{h^{j-1}}{2!(j-1)!} D_{\tilde{f}} d_{j-1} + \mathcal{O}(h^M).$$

Denoting now by $a_{i,j}$ the coefficient of the $h^{j-1}D^j v$ in the modified equation after substituting the equation for $D^i v$ into the original Equation (6.18), then we find $a_{1,j} = \frac{1}{j!}$ and $a_{2,j} = a_{1,j} - \frac{1}{2(j-1)!}$. More generally, suppose that we have substituted for increasing $2 \leq \ell \leq i-1$ Equation (6.24) for $D^\ell v$ into the modified Equation (6.18), then the substitution of $D^i v$ gives the equation

$$v' + \sum_{j=i+1}^{M} h^{j-1} \left(a_{i-1,j} - \frac{a_{i-1,i}}{(j-(i-1))!} \right) D^j v + \mathcal{O}(h^M) = \sum_{k=1}^{i} \sum_{j=k}^{M} h^{j-1} \frac{-a_{k-1,k}}{(j-(k-1)!} D_{\tilde{f}}^{k-1} d_{j-k+1}$$

defining $a_{0,1} := -1$, $a_{0,j} = 0$ for $j > 1$. Thus, the recursive formula for the coefficients $a_{i,j}$ are given by

$$a_{0,1} = -1, \quad a_{1,1} = 1, \quad a_{0,j} = 0, \quad a_{i,j} = \begin{cases} 0 & \text{if } 1 \leq j \leq i \\ a_{i-1,j} - \frac{a_{i-1,i}}{(j-(i-1))!} & \text{if } 1 \leq i < j, \end{cases}$$

with the recursively defined solution

$$a_{i,j} = -\sum_{k=1}^{i} \frac{a_{k-1,k}}{(j-(k-1))!}$$

for $1 \leq i < j$. Defining $c_i := -a_{i-1,i}$ one finds the modified equation

$$v' + \mathcal{O}(h^M) = \sum_{i=1}^{M} \sum_{j=i}^{M} h^{j-1} \frac{c_i}{(j-(i-1))!} D_{\tilde{f}}^{i-1} d_{j-(i-1)} = \sum_{i=1}^{M} h^{i-1} c_i \sum_{j=1}^{M+1-i} \frac{h^{j-1}}{j!} D_{\tilde{f}}^{j+i-2} d_j, \tag{6.25}$$

with

$$c_1 = 1 \quad c_i = -\sum_{k=1}^{i-1} c_k \frac{1}{(i-(k-1))!},$$

with solution a rescaling of the Bernoulli numbers $\mathcal{B}_i$ (with $\mathcal{B}_1 = -\frac{1}{2}$)

$$c_i = \frac{1}{(i-1)!} \mathcal{B}_{i-1}$$

so that we arrive at

$$\begin{aligned} v' + \mathcal{O}(h^M) &= \sum_{j=1}^{M} \frac{h^{j-1}}{j!} d_j + \sum_{i=1}^{M-1} \frac{\mathcal{B}_i}{i!} \sum_{j=i+1}^{M} \frac{h^{j-1}}{(j-i)!} D_{\tilde{f}}^{i} d_{j-i} \\ &= \sum_{j=1}^{M} \frac{h^{j-1}}{j!} d_j + \sum_{i=1}^{M-1} \frac{\mathcal{B}_i}{i!} \sum_{j=i+1}^{M} \frac{h^{j-1}}{(j-i)!} \sum_{k=i}^{M+i-j} h^{k-i} \sum_{\ell \in \mathcal{P}_i(k)} D_{f_{\ell_1}} \dots D_{f_{\ell_i}} d_{j-i} \end{aligned}$$

Collecting the terms of order h^{M-1} and calling this the vector field f_M (cf. Section 6.3.3) we find

$$f_q = \frac{1}{M!} d_M + \sum_{i=1}^{M-1} \frac{\mathcal{B}_i}{i!} \sum_{j=1}^{M-i} \frac{1}{j!} \sum_{\ell \in \mathcal{P}_i(M-j)} D_{f_{\ell_1}} \dots D_{f_{\ell_i}} d_j = \frac{1}{M!} d_M + \sum_{i=1}^{M-1} \frac{\mathcal{B}_i}{i!} \sum_{\ell_1 + \dots + \ell_{i+1} = M} \frac{1}{\ell_{i+1}!} D_{f_{\ell_1}} \dots D_{f_{\ell_i}} d_{\ell_{i+1}} \tag{6.26}$$

This seems to be a new recursive expression for the modified vector fields.

Remark 6.8. *In the case that more information is known about the coefficient functions d_j then the modified equation in the form* (6.25) *can be rewritten accordingly. For example if ψ_h is a B-series method (cf. equation (9.7) in [HLW06], section IX.*9.1*), a P-series method (cf. (10.6) in Section IX.*10.2 *of [HLW06]) or a symplectic method (see Section 6.3.6, or [HLW06] chapter XI.*3*)*

6.3.3 Formal modified equation via ansatz of polynomial modified vector field

A more common way of finding the modified equation [CMS94; HLW06] is discussed next. As discussed in Section 6.3.1, the idea (e.g. [LR04; HLW06]) is to search immediately for a modified vector field $\tilde{f}^{[M]}$ and not for a solution $\tilde{\phi}^{[M]}$ or $v^{[M]}$. In Example 6.7 and Equation (6.25) it was seen that a polynomial ansatz (polynomial in h, Equation (6.28) below) is suitable.

Starting at Equation (6.18), we substitute $D_t^k \tilde{\phi}^{[M]} = D_t^{j-1} \left(\tilde{f}^{[M]} \circ \tilde{\phi}^{[M]} \right)$ or $D_t^k v = D_t^{k-1} \tilde{f}^{[M]}(v)$ such that

$$\sum_{j=1}^{M} \frac{h^j}{j!} \left(D_t^j \left(\tilde{f}^{[M]} \circ \tilde{\phi} \right) - d_j \right) = 0, \qquad \text{or} \qquad \sum_{j=1}^{M} \frac{h^j}{j!} \left(D_t^j (\tilde{f}^{[M]}(v(\tilde{t}_0)) - d_j(\tilde{y}_0) \right) = 0, \tag{6.27}$$

Now, as in Equation (6.23), the idea is to use the ansatz

$$\tilde{f}^{[M]} = \sum_{j=1}^{M} h^{j-1} f_{j,M} \qquad \tilde{f}_{0,M} = f, \tag{6.28}$$

with $f_{j,M} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ where in particular $f_{j,M} =: f_j$ is independent of M (cf. Remark 6.3 and Example 6.6).

Substituting the ansatz into Equation (6.27) and comparing equal powers of h to the modified functions we find in the autonomous case

$$\begin{aligned} f_1 &= d_1 \\ 2f_2 + f_1'(f_1) &= d_2 \\ 6f_3 + 3\left(f_2'(f_1) + f_1'(f_2)\right) + f_1''(f_1, f_1) + f_1'(f_1'(f_1)) &= d_3 \\ 4!f_4 + 12\left(f_3'(f_1) + f_2'(f_2) + f_1'(f_3)\right) + \\ 4\big(f_2''(f_1, f_1) + 2f_1''(f_1, f_2) + f_1'(f_1'(f_2)) + f_1'(f_2'(f_1)) + f_2'(f_1'(f_1))\big) + \\ D^3 f_1(f_1, f_1, f_1) + 3f_1''(f_1'(f_1), f_1) + f_1'(f_1''(f_1, f_1)) + f_1'(f_1'(f_1'(f_1))) &= d_4 \\ \ldots &= d_5. \end{aligned} \tag{6.29}$$

We use now the notation of Section 6.1.4. Gor a vector field f we write the lie derivative as $D_f(g)(y, t) = D_y g(y, t) f(y, t)$, where $g \in C^1(\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n)$ and we denote by $p_i(j)$ the partitions of a number j into i numbers, as in Section B. Then one finds the usual expression for the

Proposition 6.9 ([BG94] or [HLW06]). *The* ***autonomous*** *modified vector field f_j satisfy*

$$f_j = \frac{1}{j!} d_j - \sum_{i=2}^{j} \frac{1}{i!} \sum_{k \in \mathcal{P}_i(j)} D_{f_{k_1}} \cdots D_{f_{k_{i-1}}} f_{k_i} \tag{6.30}$$

Proof. The proof is in e.g. [HLW06], chapter IX.7.2. □

For example

$$f_3 = \frac{1}{6} d_3 - \frac{1}{2}(D_{f_2} f_1 + D_{f_1} f_2) - \frac{1}{6} D_{f_1} D_{f_1} f_1$$

which agrees with Equation (6.29).

6.3.4 Formal modified eqation using a recursive definition

Finally, we briefly mention that another approach exists. A recursive approach was considered in [Rei99]. This recursive approach has been implemented numerically at least for scalar ODE [HL00]. A similar result can be for higher order methods.

6.3.5 Modified equations of the induced method on forced ODE

We consider the induced method $\Psi_h(y, t) = (\tilde{\psi}_{h,t}(y), t + h) \in \mathbb{R}^n \times \mathbb{R}$ as in Section 4.1.2, induced from $\psi_h = Id + \sum_{i=1}^{\infty} d_i h^i$, such that (using Equation (4.4)) $\tilde{\psi}_h$ has coefficients $\tilde{d}_j$, $j \in \mathbb{N}_0$ given by

$$\tilde{d}_0(y, t) = y, \quad \tilde{d}_i(y, t) = \sum_{j=1}^{i} d_j(y) B_{i,j}(g_0(t), \ldots, g_{j-i}(t)).$$

We now fix $h > 0$. Then from Proposition 4.10 one finds, if g is analytic, that

$$\psi_{f,h} - \phi_{\tilde{f}_h^{[M]}, h} = \mathcal{O}(h^{M+1}) \qquad \Longleftrightarrow \qquad \tilde{\psi}_{t+h,t} - \tilde{\phi}_{t+h,h} = \mathcal{O}(\eta^{M+1}),$$

for all t in the domain, where $\tilde{\phi}_{t,t_0} = \phi_{\int_{t_0}^{t} g(s)\, ds}$, and $\tilde{f}_h^{[M]} = \sum_{j=1}^{M} h^{j-1} f_j$ is the modified vector field of ψ_h of order M and $\eta(t, h) = \sum_{j \geq 1} h^j D_t^{j-1} g(t)$. In other words for the step-size $\eta(t, h)$ one finds

$$\tilde{\psi}_{t+h,t} - \tilde{\phi}_{t+h,h} = \psi_{f,\eta} - \phi^{[M]}_{\tilde{f}_\eta^{[M]}, \eta}.$$

meaning that $\tilde{f}_\eta^{[M]}$ is the modified vector field of the induce method ψ_η of order M. Using again Faà di Bruno's formula and the fact that $\int_{t_0}^t g(s)\,ds = \sum_{j\geq 1} h^j g_{j-1}(t_0)$ we find that

$$\tilde{f}_\eta^{[M]} = \sum_{j=1}^{M} h^{j-1} \sum_{i=1}^{j} B_{i,j}(g_0, \dots, g_{j-i}) f_j, \tag{6.31}$$

where f_j are the modified vector fields of ψ_h applied to the non-forced vector field of f and where $B_{i,j}$ are the commutative Bell polynomials (Appendix B). Thus, Equation (6.31) expresses the modified equations of the induced method.

6.3.6 Modified Hamiltonians using generating functions

If symplectic numerical methods are used, then Proposition 4.13 implies that there exists modified Hamiltonians H_j such that the modified vector fields satisfy $f_j = J\nabla H$. Furthermore, the modified Hamiltonian H_j are constructed explicitly (in quadrature) using the integrability lemma. Using this construction, the second term H_2 of the modified Hamiltonian satisfies

$$H_2(q,p,t) = \int_{(q_0,p_0)}^{(q,p)} \frac{1}{2} J \left(f'f + \dot{f} - d_2 \right) d\gamma = \int_0^1 \frac{1}{2} J f' f(sq, sp, t)\, ds - \frac{1}{2} \int_{(q_0,p_0)}^{(q,p)} J d_2 \, d\gamma.$$

for some path γ with endpoints $(q_0,p_0),(q,p)$ in the domain. We note that

$$f'f = (J\nabla^2 HJ)\nabla H) = \begin{pmatrix} -H_{pp} & H_{pq} \\ H_{qp} & -H_{qq} \end{pmatrix} \nabla H = \begin{pmatrix} -H_{pp}H_q + H_{pq}H_q \\ H_{pq}H_q - H_{qq}H_p \end{pmatrix} = \nabla \left(H_q H_p \right)$$

such that

$$H_2(q,p,t) = \frac{1}{2} H_q H_p + \frac{1}{2}\dot{H} + \int_{(q_0,p_0)}^{(q,p)} J d_2 \, d\gamma.$$

The next step is to find H_3 in terms of d_i, H and H_2 However, it is not easily done using path integrals (the Integrability Lemma) on the formulas using Equations (6.26) or (6.30).

However, using generating functions, such expressions for the modified Hamiltonians (solving the quadrature) can be found. Instead of looking for an interpolative vector field $\tilde{f}$ one looks at an interpolative generating function $\tilde{S}$. The starting point is now not the modified equation 6.18 but the Hamilton-Jacobi equation ([HLW06] chapter VI). Still, there are many similarities with modified equation analysis of Section 6.3.3: One uses a power series expansion of $\tilde{H}$ and $\tilde{S}$ in h and the goal is to find an expression for $\tilde{S}(q,P,t)$ which interpolates $S(q,P,h)$ The following lemma is from [HLW06], where it is mentioned that it was found earlier in [BG94] and [CMS94]. Here we provide a more thorough proof and a recursive expression for the modified Hamiltonians.

Theorem 6.10 ([HLW06] chapter IX.3)**.** *If the symplectic method Φ_h (step-size h) has a generating function of the form*

$$S(q,P,h) = \sum_{i=1}^{\infty} h^i S_i(q,P)$$

(e.g. Proposition 4.16) where the S_j are smooth and defined on an open set $D \subset \mathbb{R}^{2n}$.
Then the modified vector field f_j have associated Hamiltonians H_j, smooth and defined on the whole of D, such that the modified Hamiltonian $\tilde{H}$ is has a formal expansion of the form

$$\tilde{H}(q,p) = H(q,p) + \sum_{j=1}^{\infty} h^j \tilde{H}_j(q,p)$$

given by

$$H_1 = S_1, \quad H_n = S_n - \sum_{r \in \mathcal{P}_2(n+1) - \{(1,n)\}} \tilde{S}_{r_1 r_2}, \tag{6.32}$$

where $\tilde{S}_h = \sum_{n,m\geq 1} t^n h^{m-1} \tilde{S}_{nm}$ is the interpolative generating function of S. Equation 6.32 can be solved recursively combined with the fact that

$$n\tilde{S}_{nm} = \sum_{j=1}^{n-1} \frac{1}{j!} \sum_{k\in\mathcal{P}_j(n-1)} \sum_{l\in\mathcal{P}_{j+1}(m+j)} (D_q^j H_{l_1})(D_p\tilde{S}_{k_1 l_1}, \dots, D_p\tilde{S}_{k_j l_{j+1}})$$

for $n > 1, m \geq 1$.

Proof. The exact solution $P(t), Q(t)$ of the Hamiltonian system corresponding to $\tilde{H}$ (for t in the maximal interval of existence) is given by ([HLW06] chapter VI)

$$p = P(t) + D_q\tilde{S}_h(q, P(t), t), \quad Q(t) = q + D_P\tilde{S}(q, P(t), t),$$

where $\tilde{S}_h$ is the solution to the Hamilton-Jacobi differential equation

$$D_t\tilde{S}_h(q, P, t) = \tilde{H}_h(q + D_P\tilde{S}_h(q, P, t), P), \quad \tilde{S}_h(q, P, 0) = 0. \tag{6.33}$$

Now, a Taylor expansion of Equation (6.33) in the second variable leads to

$$D_t\tilde{S}(q, P, t) = \tilde{H}(q, P) + D_q\tilde{H}(q, P)(D_P\tilde{S}(q, P, t)) + \sum_{j=2}^{\infty} \frac{1}{j!}(D_q^j H(q, P))(D_P\tilde{S}, \dots, D_P\tilde{S})(q, P, t).$$

One considers now a power series expansion of $\tilde{S}_h$ in the variables (t, h). Writing $\tilde{S}_h(q, P, t) = \sum_{i=1}^{\infty} t^i \tilde{S}_i(q, P, h)$ and substituting this into Equation (6.33), one finds the expression

$$\sum_{i=1}^{\infty} i t^{i-1} \tilde{S}_i(q, P, h) = \tilde{H}(q, P) + \sum_{j=1}^{\infty} \frac{1}{j!}(D_q^j H(q, P))(D_P \sum_{i=1}^{\infty} t^i \tilde{S}_i, \dots, D_P \sum_{i=1}^{\infty} t^i \tilde{S}_i)(q, P, h),$$

which is similar to the case in Section 6.3.3 but on the level of generating functions. Rewriting and comparing powers leads to the expression

$$\begin{aligned}
\tilde{S}_1(q, P, h) &= \tilde{H}(q, P) \\
2\tilde{S}_2(q, P, h) &= (D_q\tilde{H} \cdot D_P\tilde{S}_1)(q, P, h) \\
3\tilde{S}_3(q, P, h) &= (D_q\tilde{H} \cdot D_P\tilde{S}_2)(q, P, h) + \frac{1}{2} D_q^2 H(D_P\tilde{S}_1, D_p\tilde{S}_1)(q, P, h) \\
4\tilde{S}_4(q, P, h) &= (D_q\tilde{H} \cdot D_P\tilde{S}_3)(q, P, h) + D_q^2 H(D_P\tilde{S}_1, D_P\tilde{S}_2) + \frac{1}{6} D_q^3 H(q, P)(D_P\tilde{S}_1, D_P\tilde{S}_1, D_P\tilde{S}_1)
\end{aligned} \tag{6.34}$$

and more generally (for $2 \leq n \in \mathbb{N}$)

$$n\tilde{S}_n(q, P, h) = \sum_{j=1}^{n-1} \frac{1}{j!} \sum_{k\in\mathcal{P}_j(n-1)} (D_q^j\tilde{H})(D_p\tilde{S}_{k_1}, \dots, D_p\tilde{S}_{k_j})$$

where $\mathcal{P}_j(n-1)$ denotes the set of (ordered) partitions, as in Appendix B.1. Since $D_q^j\tilde{H}$ is a symmetric multilinear function map, one might replace $\mathcal{P}_j(n-1)$ by the ordered partition, together with a multinomial factor. Substituting now $\tilde{S}_i = \sum_{j=1}^{\infty} h^{j-1}\tilde{S}_{ij}$ and $\tilde{H} = \sum_{j=1}^{\infty} h^{j-1} H_j$ one finds

$$\begin{aligned}
\tilde{S}_{1m}(q, P) &= H_m(q, P) \\
2\tilde{S}_{2m}(q, P) &= \sum_{l\in p_2(m+1)} D_q H_{l_1}(q, P)(D_P\tilde{S}_{1l_2})(q, P) \\
3\tilde{S}_{3m}(q, P) &= \sum_{l\in p_2(m+1)} (D_q H_{l_1}(q, P)(D_P\tilde{S}_{2l_2})(q, P) + \frac{1}{2} \sum_{l\in p_3(m+2)} D_q^2 H_{l_1}(D_P\tilde{S}_{1l_2}, D_P\tilde{S}_{1l_3})(q, P)
\end{aligned}$$

or more generally (for $2 \leq n \in \mathbb{N}$, $m \in \mathbb{N}$)

$$n\tilde{S}_{nm}(q,P) = \sum_{j=1}^{n-1} \frac{1}{j!} \sum_{k \in \mathcal{P}_j(n-1)} \sum_{l \in \mathcal{P}_{j+1}(m+j)} (D_q^j H_{l_1})(D_p \tilde{S}_{k_1 l_1}, \dots, D_p \tilde{S}_{k_j l_{j+1}}). \tag{6.35}$$

Similarly, by symmetry of $D_q^j H_{l_1}$ one might substitute $\mathcal{P}_j(n-1), \mathcal{P}_{j+1}(m+j)$ by the ordered partitions together with a multinomial factor (distinguishing l_1). It is clear from this expression that $\tilde{S}_{nm}$ only depends on $D_p\tilde{S}_{ij}, D_q\tilde{S}_{ij}$ with $i < n$, $j \leq m$ i.e. (i,j) element-wise smaller or equal than (n,m). Thus, the last equation is not implicit, but can be used recursively.

Finally, the interpolation requirement (which has not been used up to this point) $S(q,P,h) = \tilde{S}_h(q,P,h)$ or $\sum_{i\geq 1} h^i S_i(q,P) = \sum_{i \geq 1} \sum_{j \geq 1} h^{i+j-1} \tilde{S}_{ij}(q,P)$ gives

$$\begin{aligned}
S_1 &= \tilde{S}_{11} = H \\
S_2 &= \tilde{S}_{12} + \tilde{S}_{21} = H_2 + \tilde{S}_{21} \\
S_3 &= \tilde{S}_{13} + \tilde{S}_{22} + \tilde{S}_{31} = H_3 + \tilde{S}_{22} + \tilde{S}_{31} \\
S_4 &= \dots
\end{aligned}$$

and more generally

$$S_n = \sum_{r \in \mathcal{P}_2(n+1)} \tilde{S}_{r_1 r_2} = \sum_{r_1=1}^{n} \tilde{S}_{r_1(n+1-r_2)}. \quad \text{or} \quad H_n = S_n - \sum_{r \in \mathcal{P}_2(n+1) - \{(1,n)\}} \tilde{S}_{r_1 r_2}. \tag{6.36}$$

This implies that the domain of the H_j is again D (i.e. they are globally defined) and they are smooth. Combining now the fact that $\tilde{S}$ is a solution to the Hamilton-Jacobi equation (leading to Equation (6.35)) and the interpolative property (leading to (6.36)), then

$$\begin{aligned}
S_1 &= \tilde{S}_{11} = H_1 \\
H_n &= S_n - \sum_{r \in \mathcal{P}_2(n+1)-\{(1,n)\}} \sum_{j=1}^{r_1-1} \frac{1}{j! r_1} \sum_{k \in \mathcal{P}_j(r_1-1)} \sum_{l \in \mathcal{P}_{j+1}(r_2+j)} (D_q^j \tilde{H}_{l_1})(D_p \tilde{S}_{k_1 l_2}, \dots, D_p \tilde{S}_{k_j l_{j+1}}),
\end{aligned} \tag{6.37}$$

which proves the theorem. □

A corollary of this lemma is that $S_1 = H$ must hold for such methods.

If $\tilde{S}(q,p,t)$ is a C^2 the solution to the (non-autonomous) Hamilton-Jacobi equation

$$D_t \tilde{S}(q,p,t) = H(q + D_P \tilde{S}(q,P,t), P, t) \quad \tilde{S}(q,P,0) = 0$$

for a non-autonomous Hamiltonian H, then $\tilde{S}$ still generates, for sufficiently small $t > 0$ and some $t_0 \in \mathbb{R}$, the solution of the Hamiltonian ODE via the equations

$$p = P(t) + D_q \tilde{S}(q, P(t), t_0 + t) \quad Q(t) = q + D_P \tilde{S}(q, P(t), t_0 + t).$$

Indeed differentiating both equations with respect to time, one finds

$$\dot{P}(I + D_P D_q \tilde{S}) = -H_q(I + D_q D_P \tilde{S}) \quad \dot{Q} = D_P D_P \tilde{S} \cdot \dot{P} + D_P H + D_P D_p \tilde{S} \cdot H_q$$

and one can use that $\tilde{S}$ is C^2 and that $I + D_p D_q \tilde{S}(q, P(t), t + t_0)$ is invertible for sufficiently small t such that $(\dot{Q}(t), \dot{P}(t)) = J \nabla H(Q(t), P(t), t_0 + t)$.

Theorem 6.11. *In the setting of Theorem 6.10 one finds the non-autonomous modified Hamiltonian*

$$\tilde{H}(q,p,t) = \sum_{i,j \geq 0}^{\infty} h^j H_j(q,p,t)$$

given by

$$S_1 = H_1, \quad S_n = \sum_{r\in p_2(n+1)} \sum_{j=0}^{r_1-1} \sum_{k\in p_{j+1}(r_1+1)} \sum_{l\in p_j(r_2+j)} \frac{1}{(k_1-1)!(r_1-k_1+1)!}(D_q^j D_t^{k_1-1}\tilde{H}_{l_1})(D_p\tilde{S}_{k_2l_2},\dots,D_p\tilde{S}_{k_{j+1}l_{j+1}}). \tag{6.38}$$

which, given S_j, recursively define the H_j.

Proof. The proof similar to the one of Theorem 6.10, we look at how Equations (6.35) and (6.36) change. Again, the start for Equation (6.35) is the Hamilton-Jacobi equation, shown above to be of the form

$$D_t\tilde{S}_h(q,P,t_0+t) = \tilde{H}_h(q + D_P\tilde{S}_h(q,P,t_0+t), P, t_0+t) \qquad \tilde{S}(q,p,t_0) = 0$$

and Taylor expanding

$$D_t\tilde{S}_h(q,p,t_0+t) = \sum_{j,k\geq 0} \frac{t^k}{j!k!} D_q^j D_t^k H(q,p,t_0)\left(D_p\tilde{S}_h \dots D_p\tilde{S}_h\right)$$

writing again $\tilde{S}_h(q,p,t_0+t) = \sum_{j\geq 1} t^j \tilde{S}_j(q,p,t_0,h)$ one finds

$$\begin{aligned} \tilde{S}_1 &= \tilde{H}_h \\ 2\tilde{S}_2 &= D_t\tilde{H}_h + (D_q\tilde{H}_h \cdot D_p\tilde{S}_1) \\ 3\tilde{S}_3 &= \frac{1}{2}D_t^2\tilde{H}_h + (D_q\tilde{H}_h \cdot D_p\tilde{S}_2 + D_qD_t\tilde{H}_h \cdot D_p\tilde{S}_1) + \frac{1}{2}D_q^2\tilde{H}_h(D_p\tilde{S}_1, D_p\tilde{S}_1) \end{aligned} \tag{6.39}$$

and more generally (for $n \in \mathbb{N}$)

$$n\tilde{S}_n(q,P,h) = \sum_{j=0}^{n-1} \sum_{k\in p_{j+1}(n+1)} \frac{1}{(k_1-1)!(n-k_1+1)!}(D_q^j D_t^{k_1-1}\tilde{H}_h)(D_p\tilde{S}_{k_2},\dots,D_p\tilde{S}_{k_{j+1}})$$

where $p_\alpha(\beta)$ denotes the set of (ordered) partitions, as in Appendix B.1. Substituting now $\tilde{S}_i(q,p,t_0,h) = \sum_{j=1}^\infty h^{j-1}\tilde{S}_{ij}(q,p,t_0)$ and $\tilde{H}_h(q,p,t_0) = \sum_{j=1}^\infty h^{j-1}H_j(q,p,t_0)$ one finds

$$\begin{aligned} \tilde{S}_{1m} &= H_m \\ 2\tilde{S}_{2m} &= D_tH_m + \sum_{l\in p_2(m+1)} D_qH_{l_1}(D_P\tilde{S}_{1l_2}) \\ 3\tilde{S}_{3m}(q,P) &= D_t^2H_m + \sum_{l\in p_2(m+1)} \Big(D_qD_tH_{l_1}(q,P)(D_P\tilde{S}_{2l_2}) + D_qD_t^2H_{l_1}(q,P)(D_P\tilde{S}_{1l_2})\Big) \\ &\quad + \frac{1}{2}\sum_{l\in p_3(m+2)} D_q^2H_{l_1}(D_P\tilde{S}_{1l_2}, D_P\tilde{S}_{1l_3})(q,P) \end{aligned}$$

or more generally

$$n\tilde{S}_n(q,P,h) = \sum_{j=0}^{n-1} \sum_{k\in p_{j+1}(n+1)} \sum_{l\in p_j(m+j)} \frac{1}{(k_1-1)!(n-k_1+1)!}(D_q^j D_t^{k_1-1}\tilde{H}_{l_1})(D_p\tilde{S}_{k_2l_2},\dots,D_p\tilde{S}_{k_{j+1}l_{j+1}}).$$

The interpolative property, Equation (6.36) does not change, such that finally

$$S_1 = H_1, \quad S_n = \sum_{r\in p_2(n+1)} \sum_{j=0}^{r_1-1} \sum_{k\in p_{j+1}(r_1+1)} \sum_{l\in p_j(r_2+j)} \frac{1}{(k_1-1)!(r_1-k_1+1)!}(D_q^j D_t^{k_1-1}\tilde{H}_{l_1})(D_p\tilde{S}_{k_2l_2},\dots,D_p\tilde{S}_{k_{j+1}l_{j+1}}).$$

□

Consistency conditions, require $H = S_1$ such that one finds, without making use of the Integrability lemma, that

$$S_2 = \tilde{S}_{12} + \tilde{S}_{21} = H_2 + \dot{H} + (D_q H)^T D_p H,$$

which agrees with the calculations at the start of this Section.

Finally, we look, given a vector field f and a numerical method ψ at the induced method $\tilde{\psi}$ applied to the forced ODE with vector field gf for some time-forcing $g \in C^\infty(\mathbb{R})$. By Equation (4.7), the induced method $\tilde{\psi}$ is symplectic. Using Equation (6.31) one finds that the modified Hamiltonians satisfy

$$\hat{H}_i(q,p,t) = \sum_{j=1}^{i} B_{i,j}(g_0(t), \ldots, g_{i-j}(t)) H_j(q,p) \tag{6.40}$$

where the H_j are determined recursively, as in Theorem 6.10 and hwere $B_{i,j}$ are the Bell polynomials (Appendix B).

6.4 BEA for splitting methods

Suppose we split an autonomous vector field into the parts $f = f_1 + f_2$ and it is possible to find the flows ϕ_{f_i} of f_i exactly. The BCH formula then allows us to find the modified vector fields: From Equation (6.10) we find formally

$$\phi_h^{[2]} \circ \phi_h^{[1]} = e^{hD_1} e^{hD_2} = e^{\tilde{D}(h)}$$

where, form Equation (6.11)

$$\tilde{D} = D_1 + D_2 + \frac{h}{2}[D_2, D_1] + \frac{h^2}{12}\left([D_2,[D_2,D_1]] + [D_1,[D_1,D_2]]\right) + \mathcal{O}(h^3),$$

where $D_i = D_{f_i}$. Thus, the modified vector field $\tilde{f}$ satisfies

$$\tilde{f} = \tilde{D} Id = f_1 + f_2 + \frac{h}{2}(f_1' f_2 - f_2' f_1) + \mathcal{O}(h^2).$$

Similarly, for the Strang-splitting one may use Equation (6.13) to find the first three modified vector fields.

In the case that the flows of f_i cannot be solved exactly, one simply substitutes, formally, D_i by $D_{\tilde{f}_i}$, where $\tilde{f}_i$ is the modified vector field.

Finally, for Hamiltonian splittings with symplectic integrators, one may substitute the Poisson bracket for the Lie bracket of vector fields, as in Section 6.2.2.

6.4.1 BEA for splitting methods on time-affine Hamiltonian ODE

We consider now splitting methods for time-affine, Hamiltonian ODE i.e. with Hamiltonian $H(q,p,t) = \sum_{i=1}^n g_i(t) H_i(q,p)$ and show, if induced methods are used, that the modified Hamiltonian is again time-affine.

Proposition 6.12. *Suppose a splitting method is applied to the Hamiltonian ODE with Hamiltonian H so that it is split into n forced Hamiltonians $g_i H_i$ on which an induced method is used. Then the modified Hamiltonian $\tilde{H} = \sum_{i=1}^\infty h^i \tilde{H}_i(q,p,t)$ has the property that the $\tilde{H}_i$ are time-affine.*

Proof. Suppose first that one integrates the n Hamiltonian systems with Hamiltonian $g_i(t) H_i(q,p)$ exactly with flow ϕ_i.

Then, from Section 6.2.2 it follows that the modified Hamiltonians $\tilde{H}_i$ consist sums of repeated extended Poisson brackets (i.e Poisson brackets in the extended phase space) which are equal to the Poisson bracket as in Equation (6.14) (replacing there the Lie bracket with the Poisson bracket).

Furthermore (cf. Remark 6.2), one sees easily from Equation (6.14) that the extended Poisson bracket of two time-affine Hamiltonian is again a time-affine function. Thus, repeated evaluation of the extended Poisson brackets produce time-affine vector fields, so that the modified Hamiltonians $\tilde{H}_i$, which are sums of these repeated extended Poisson brackets, are also time-affine.

In the case that $g_i H_i$ cannot be solved exactly i.e. one finds, using an approximate flow $\psi_{i,h} \approx \phi_{i,t}$. In the BCH formula one then replaces $g_i H_i$ with the modified Hamiltonian $\hat{H}_i$ given by Equation (6.40). Thus $\hat{H}_i$ is time-affine and the proof is done. □

The importance of this Proposition is that, using a splitting method on a time-affine Hamiltonian ODE with the induced methods on the split Hamiltonians, the modified Hamiltonian can be seen in the case of the tidal wave system as a periodically perturbed Hamiltonian ODE on which the KAM theorem 3.9 can be used.

6.5 Approximate KAM theorem for symplectic PRK integrators on non-autonomous periodically perturbed completely integrable systems

We follow [HLW06] chapter IX.7 and X.5 in this Section. The goal is to state an approximate KAM theorem for a symplectic integrator ψ_h applied to a periodically perturbed, completely integrable Hamiltonian systems. This is done in two parts: First, Sections 6.5.3-6.5.1 find a local error $\left\|\psi_h - \tilde{\phi}_h^{[M_*]}\right\|$ between a PRK method ψ_h applied to an ODE with vector field f and the solution $\tilde{\phi}^{[M_*]}$ of the modified equation $\tilde{f}^{M_*}$, for some optimally chosen M_*. Here f is assumed to be complex analytic and bounded on a complex neighbourhood of $B_{2R}(y_0) := \{y \in \mathbb{C}^n \mid \|y - y_0\| \leq 2R\}$ of some $y_0 \in \mathbb{R}^n$ (for $R > 0$, $t \in I \subset \mathbb{R}$ and $\|\cdot\|$ a norm on $\mathbb{C}^n$):

$$\|f(y,t)\| \leq M \qquad \text{for} \qquad y \in B_{2R}(y_0). \tag{6.41}$$

To this end we follow [HLW06], who clearly state the strategy of Section 6.5.1-6.5.3:

> "Our strategy is the following: using [(6.41)] and Cauchy's estimates we derive bounds for the coefficient functions $d_j(y)$ [as in Remark 4.3] on $B_R(y_0)$ [...], then we estimate the vector fields $f_j(y)$ of the modified differential equation on $B_{R/2}(y_0)$ [...], and finally we search for a suitable truncation for the formal series $[\tilde{f}^{M_*}]$ and we prove the closeness of the numerical solution to the exact solution of the truncated modified equation [...]." – Hairer, Lubich & Wanner [HLW06].

The second part, Section 6.5.4, uses the bound on the local error $\left\|\psi_h - \tilde{\phi}_h^{[M_*]}\right\|$ together with the KAM Theorem 3.9 to show that a symplectic PRK method, applied to a periodically perturbed, completely integrable Hamiltonian system, has 'almost invariant' tori.
At two points we differ from [HLW06]. First, when estimating of the coefficients f_j, not only the recursive expressions of Equation (6.30)

$$f_j(y) = d_j(y) - \sum_{i=2}^{j} \frac{1}{i!} \sum_{k_1+\cdots+k_i=j} \left(D_{k_1} \ldots D_{k_{i-1}} f_{k_i}\right)(y).$$

is used (following [HLW06]) but also estimates using the recursive expressions of Equation (6.26)

$$f_j(y) = d_j(y) - \sum_{i=1}^{j-1} \frac{\mathcal{B}_i}{i!} \sum_{k_1+\cdots+k_{i+1}=j} \left(D_{k_1} \ldots D_{k_i} d_{k_{i+1}}\right)(y).$$

are explored in Section 6.5.2) ($\mathcal{B}_i$ the Bernoulli numbers). The latter (new) approach is not fully developed but numerical simulations suggest that this approach may lead to better estimates in some cases.

The second difference is that we consider periodically perturbed systems, whereas [HLW06] only treats autonomously perturbed systems.

6.5.1 Estimation of the derivatives of a PRK method

We consider a consistent, s-stage PRK method $\psi_h : \mathbb{R}^n \to \mathbb{R}^n$ (Section 4.1.1) with butcher tableaus $B = (A, b, c)$, $\tilde{B} = (\tilde{A}, \tilde{b}, \tilde{c}) \in \mathbb{R}^{s\times s} \times \mathbb{R}^s$ with (Remark 4.3)

$$\psi_h(y) = y + hf(y) + \sum_{j\geq 2} h^j d_j(y). \tag{6.42}$$

Suppose the PRK method ψ_h splits the variable y as $y = (z, w) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ and similarly the vector field $f = (\zeta, \chi)$ and assume that $f(z, w, t)$ is analytic for $(z, w) \in B_{2R_1}(z_0) \times B_{2R_2}(w_0)$ (and $t \in I \subset \mathbb{R}$) on which

$$\begin{aligned} \|\zeta(z, w, t)\| &\leq M_1 \\ \|\chi(z, w, t)\| &\leq M_2. \end{aligned} \tag{6.43}$$

Then the d_j are analytic and can be bounded for sufficiently small $h > 0$.

Proposition 6.13 ([HLW06] theorem 7.2)**.** *Given an s-stage PRK method ψ_h as above, with Butcher tableaus $B = (A, b)$, $\tilde{B} = (\tilde{A}, \tilde{b})$ on the splitting $y = (z, w)$. Suppose that*

$$\begin{aligned} \mu &= \sum_{i=1}^s |b_i| \qquad & \kappa &= \max_{i=1\ldots s} \sum_{j=1}^s |a_{ij}| \\ \tilde{\mu} &= \sum_{i=1}^s |\tilde{b}_i| & \tilde{\kappa} &= \max_{i=1\ldots s} \sum_{j=1}^s |\tilde{a}_{ij}|. \end{aligned},$$

If $f(z, w, t)$ is analytic on the complex domain $(z, w) \in B_{2R_1}(z_0) \times B_{2R_2}(w_0)$ $(t \in I)$ and satisfies Equation (6.43)*. Then the functions $d_j(y, t)$ of Equation* (6.42) *are analytic on $B_{2R_1}(z_0) \times B_{2R_2}(w_0)$ (and $t \in I$) and satisfy*

$$\|d_j(z, w, t)\| \leq (\mu M_1 + \tilde{\mu} M_2)\left(2\left[\frac{\kappa M_1}{R_1} + \frac{\tilde{\kappa} M_2}{R_2}\right]\right)^{j-1} \qquad for (z, w) \in B_{2R_1}(z_0) \times B_{2R_2}(w_0). \tag{6.44}$$

Furthermore suppose that $\|\cdot\|$ is a p-norm $(1 \leq p \leq \infty)$. If f is analytic on $B_{2R}(y_0)$ and satisfies Equation (6.41) *then*

$$\|d_j(y, t)\| \leq (\mu_0 M)\left(\frac{2\kappa_0 M}{R}\right)^{j-1} \qquad for \qquad \|y - y_0\| \leq R. \tag{6.45}$$

where

$$\mu_0 = \sum_{i=1}^s \max\left(|b_i|, |\tilde{b}_i|\right), \qquad \kappa_0 = \max_{i=1\ldots s} \sum_{j=1}^s \max\left(|a_{ij}|, |\tilde{a_{ij}}|\right)$$

Proof. The first estimate, for PRK methods is almost identical to the one for RK methods, for which we refer to [HLW06] chapter IX theorem 7.2. In the case that $\|\cdot\|$ is a p-norm, then fact that

$$\left\|\begin{pmatrix} a_{ij} z \\ \tilde{a}_{ij} w \end{pmatrix}\right\| \leq \max\left(|a_{ij}|, |\tilde{a}_{ij}|\right) \left\|\begin{pmatrix} z \\ w \end{pmatrix}\right\|$$

will lead to the second estimate. □

6.5.2 Estimation of the coefficients of the modified vector field

As mentioned, we not only state the approach of [HLW06] (using in its proof Equation (6.30)) but also use the Equation (6.26). This latter approach is not made completely rigorous, since estimates are missing. But a start is given and numerically the estimates may predict better results.

In this Section we also assume that the vector field is *autonomous.* For non-autonomous vector fields one can either extend the system (Sections 2.1 and 2.1.3) and use the estimates for the non-autonomous vector field. Alternatively, one can use the non-autonomous versions of the modified vector fields (Appendix D), but on first sight this seems to be much harder.

First we state a proposition from [HLW06], where the proof can be found.

Proposition 6.14 ([HLW06] chapter IX theorem 7.5)**.** *Suppose $\|\cdot\|$ is a p-norm. Let $f(y)$ be analytic on $y \in B_{2R}(y_0)$, let the Taylor series coefficients d_j of the numerical method ψ_h be analytic on $B_{2R}(y_0)$ and assume that Equation* (6.41) *and* (6.45) *are satisfied. Then, we have for the coefficients of the modified differential equation*

$$\|f_j(y)\| \leq \ln 2\eta M \left(\frac{\eta M j}{R}\right)^{j-1} \qquad \textit{for} \qquad \|y - y_0\| \leq R/2, \tag{6.46}$$

where $\eta = 2\max\left(\kappa_0, \mu_0/(2\ln 2 - 1)\right) \geq 5.1773989$.

The proof of Proposition 6.14 (see [HLW06] chapter IX theorem 7.5), Equation (6.30) is used. Next we consider a similar idea as in the proof Equation, using instead (6.26).

We fix $J \in \mathbb{N}$ and try to estimate

$$\|f_J\|_{R/2}$$

where $\delta = R/(2(J-1))$ and for $r > 0$ we denote $\|f_j\|_r := \max\{\|f_j(y)\| \mid y \in B_r(y_0)\}$. Following the proof in [HLW06] chapter IX theorem 7.5 identically, up until equation (7.9), we then find

$$f_j(y) = d_j(y) - \sum_{i=1}^{j-1} \frac{\mathcal{B}_i}{i!} \sum_{k_1+\cdots+k_{i+1}=j} \left(D_{k_1} \ldots D_{k_i} d_{k_{i+1}}\right)(y),$$

so that $\|f_j\|_j \leq \delta\chi_j$ for $1 \leq j \leq J$ with

$$\chi_j = \frac{\gamma}{\delta}\left(q^{j-1} + \sum_{i=1}^{j-1} \frac{|\mathcal{B}_i|}{i!} \sum_{r=1}^{j-i} q^r \sum_{k \in p_i(j-r)} \chi_{k_1} \cdots \chi_{k_i}\right), \tag{6.47}$$

where $q = \frac{2\kappa_0 M}{R}$, $\gamma = \mu_0 M$ and $p_i(j-r)$ denotes the set partitions of $j-r$ into i integers (as in Section B).

Defining the generating function $c(\zeta) = \sum_{j\geq 1} \chi_j \zeta^{j-1}$ we find formally

$$\begin{aligned}
c(\zeta) &= \gamma/\delta\left(\frac{1}{1-\zeta q} + \sum_{j\geq 0}\sum_{i=1}^{j-1} \frac{|\mathcal{B}_i|}{i!} \sum_{r=1}^{j-i} \zeta^{r-1} q^{r-1} \sum_{k \in p_i(j-r)} \zeta^{k_1}\chi_{k_1} \cdots \zeta^{k_i}\chi_{k_i}\right) \\
&= \gamma/\delta\left(\frac{1}{1-\zeta q} + \sum_{i\geq 1} \frac{|\mathcal{B}_i|}{i!} \sum_{r\geq 0} \zeta^r q^r \sum_{k_1+\cdots+k_i \geq i} \zeta^{k_1}\chi_{k_1} \cdots \zeta^{k_i}\chi_{k_i}\right) \\
&= \frac{\gamma/\delta}{1-\zeta q}\left(1 + \sum_{i\geq 1} \frac{|\mathcal{B}_i|}{i!} c(\zeta)^i\right).
\end{aligned}$$

Now, for $x \in \mathbb{R}$ with $|x|$ sufficiently small we find that (e.g. [DK04] in the preview exercises)

$$\pi x \cot(\pi x) = \sum_{n \in \mathbb{N}_0} (-1)^n (2\pi)^{2n} \frac{\mathcal{B}_{2n}}{(2n)!} x^{2n} = 1 - \frac{\pi^2}{3}x^2 - \frac{\pi^4}{45}x^4 - \mathcal{O}(x^6)$$

where cot is the cotangent. Combining this with the fact that the Bernoulli numbers $\mathcal{B}_i$ have changing signs e.g. have the first couple of values (with the convention $\mathcal{B}_1 = -1/2$)

$$\mathcal{B}_0 = 1, \quad \mathcal{B}_1 = -\frac{1}{2}, \quad \mathcal{B}_2 = \frac{1}{6}, \quad \mathcal{B}_{2n+1} = 0, \quad \mathcal{B}_4 = -\frac{1}{30}, \quad \mathcal{B}_6 = \frac{1}{42}, \ldots$$

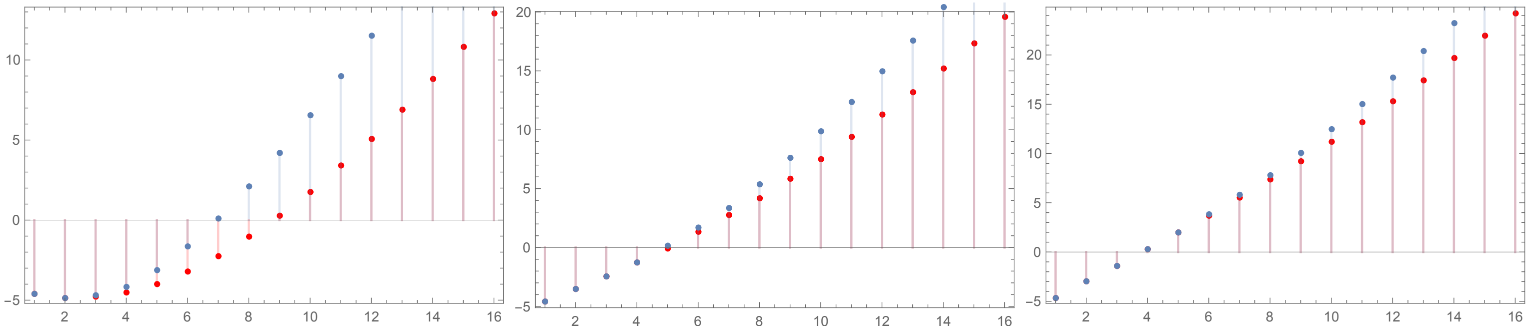
(so that $|\mathcal{B}_{2n}| = (-1)^{n+1}\mathcal{B}_{2n}$ for $n \geq 1$) we find then for at least $|c(\zeta)|$ small enough that

$$1 + \sum_{n\geq 1} \frac{|\mathcal{B}_n|}{n!} c(\zeta)^n = 2 + \frac{1}{2}c(\zeta) + \sum_{n\geq 0} (-1)^{n+1} (2\pi)^{2n} \frac{|\mathcal{B}_{2n}|}{(2n)!} \left(\frac{c(\zeta)}{2\pi}\right)^{2n} = 2 + \frac{1}{2}c(\zeta)\left(1 - \cot\left(\frac{1}{2}c(\zeta)\right)\right)$$

Therefore, formally, the generating function $c(\zeta)$ satisfies

$$c(\zeta) = \frac{\gamma/\delta}{1-\zeta q}\left(2 + \frac{1}{2}c(\zeta)\left[1 - \cot\left(\frac{1}{2}c(\zeta)\right)\right]\right)$$

and, depending on where this holds not only formally but also numerically, this is the function which we may need. Comparing with the proof of [HLW06] chapter IX theorem 7.5 at this point the implicit function theorem is used and, after some estimates, Cauchy's theorem but this approach does not seems to be harder in this case. Numerically, we see that for $\mu_0 = 1$[23] the estimates of χ_J seem to be better than the equivalent of χ_J in the proof of [HLW06] chapter IX theorem 7.5, denoted β_J, Figure 17. In particular, the quotient χ_J/β_J seems to be independent of M, R, Figure 18. For $\mu_0 > 1$ one could see that β_J started to become smaller than χ_J, so estimates from this approach are probably better only for $\mu_0 \approx 1$.

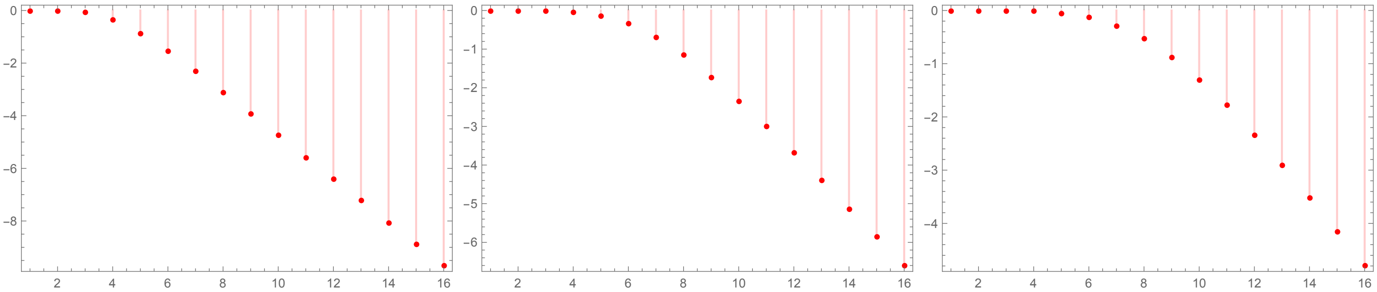


(a) $\kappa_0 = 3.5$

(b) $\kappa_0 = 13.5$

(c) $\kappa_0 = 23.5$

Figure 17: Here we see $\ln(\delta\chi_J)$ (red) and $\ln(\delta\beta_J)$ (blue) without green (for $J = 1 \dots 16$). One can see more clearly that, for different κ_0 the $\chi_J < \beta_J$. We set $\mu_0 = 1, R = 0.01, M = 0.1$ in all three figures.



(a) $\kappa_0 = 3.5$

(b) $\kappa_0 = 13.5$

(c) $\kappa_0 = 23.5$

Figure 18: Here we see $\ln(\chi_J/\beta_J)$ (for $J = 1 \dots 16$). In particular, this value has been numerically seen to be independent of M, R. We set $\mu_0 = 1$ and different values of κ_0

6.5.3 Optimal truncation and closeness of the flow

Using the bounds of the previous Sections on the coefficints d_j and the modified vector fields f_j we state (similar to [BG94; HL97; Rei99]) chapter IX theorem 7.6 from [HLW06] (to which we refer for the proof) about the local error between the numerical method and the flow of an optimally truncated modified vector field.

Proposition 6.15 ([HLW06] chapter IX theorem 7.6)**.** *Let $f(y)$ be analytic in $B_{2R}(y_0)$, let the coefficients $d_j(y)$ of the numerical method (of order N) be analytic in $B_R(y_0)$, and assume that Equation* (6.41) *and*

[23] One has $\mu_0 = 1$ for PRK methods with Butcher tableau (A, b), $(\tilde{A}, \tilde{b})$ such that $\tilde{b}_j = b_j > 0$, for example symplectic PRK methods on non-separable Hamiltonian ODE e.g. Section 4.2.2

(6.46) *hold. If $h \leq h_0/4$ with $h_0 = R/(e\eta M)$, then there exists $\tilde{N} = \tilde{N}(h)$ (namely $\tilde{N}$ equal to the largest integer satisfying $h\tilde{N} \leq h_0$) such that the difference between the numerical solution $y_1 = (y_0)$ and the exact solution $\tilde{\phi}^{[\tilde{N}]}t(y_0)$ of the truncated modified equation $f + \sum_{j=N+1}^{\tilde{N}} h^{j-1} f_j$ satisfies*

$$\left\| \psi_h(y_0) - \phi^{[\tilde{N}]}h(y_0) \right\| \leq h\gamma M e^{-h_0/h}, \tag{6.48}$$

where $\gamma = e(2 + 1.65\eta + \mu_0) > 31.37$ depends only on the method.

6.5.4 An approximate KAM theorem for (quasi)-periodically perturbed completely integrable systems

We now consider a quasi-periodically perturbed Hamiltonian system $H : B \times \mathbb{T}^n \times \mathbb{R} \to \mathbb{R}$, where $B \subset \mathbb{R}^n$, given by

$$H(a, \theta, t) + H_0(a) + \mu H_1(a, \theta, t), \tag{6.49}$$

where H_1 is quasi-periodic with frequencies $\Omega = (\Omega_1, \ldots, \Omega_s) \in \mathbb{R}^s$ in the variable t (Definition 3.1) on which a symplectic method ψ_h with step size $h > 0$ is used. The following theorem, which is almost identical to [HLW06] chapter X.5 theorem 5.4, shows that there are are $(n+s)$-dimensional 'almost invariant' tori of the map ψ_h when extended to the phase space $\mathbb{R}^n \times \mathbb{T}^n \times \mathbb{T}^s$. The difference with [HLW06] chapter X.5 theorem 5.4 is that we now consider quasi-periodic perturbations, instead of autonomous perturbations.

Theorem 6.16. *Suppose a symplectic method ψ_h of order N is used on a periodically perturbed completely integrable Hamiltonian $K(q,p,t) = K_0(q,p) + \mu K(q,p,t)$ defined on $D \times \mathbb{R} \subset \mathbb{R}^{2n} \times \mathbb{R}$ (in action-angle coordinates (a, θ) given by Equation (6.49)) satisfying the bound*

$$\left\| \psi_h(q,p,t) - \phi^{[\tilde{N}]}_{t+h,t}(q,p) \right\|_1 \leq hCe^{-h_0/h} \quad \text{for} \quad (q,p) \in \mathcal{D} \subset D \tag{6.50}$$

for some $C, h_0 > 0$ and for all $t \in \mathbb{R}$, where $\phi^{[\tilde{N}]}$ is the flow of the modified Hamiltonian $\tilde{K}^{[\tilde{N}]}(q,p,t)$ (as in Theorem 6.11) truncated at order $h^{\tilde{N}}$.

Suppose furthermore that the unperturbed Hamiltonian $H_0(a) := K(q,p)$ (in action-angle coordinates (a,θ)) and the frequencies Ω satisfy the (non-degeneracy, complex analyticity, non-resonance) conditions of the KAM Theorem 3.9 on some domain $B \times \mathbb{T}^n$.

Then for 'most' (see Theorem 3.9) invariant tori $T_\omega = T_{\omega(a)}$ of the unperturbed system with Hamiltonian H_0 with $a \in B$, there exists an $n+s$-dimensional torus $\tilde{T}_{(\tilde{\omega},\Omega)}$, $\mathcal{O}(h^N + \mu)$ close to T_ω, carrying a quasi-periodic flow with frequencies $(\tilde{\omega}, \Omega)$ and a Hamiltonian $\tilde{H}$ (related to the modified Hamiltonian $\tilde{K}^{[\tilde{N}]}$ defined on $D \times \mathbb{R}$), such that $\tilde{T}_{(\tilde{\omega},\Omega)}$ is an invariant torus for the flow of $\tilde{H}$. Furthermore, the difference between any numerical solution (p_n, q_n) starting on the torus $\tilde{T}_{(\tilde{\omega},\Omega)}$ and the solution $(p(t), q(t))$ of the modified Hamiltonian system with the same starting values remains exponentially small in $1/h$ over exponentially long times:

$$\|(p_n, q_n) - (p(t), q(t))\|_1 \leq Ce^{-\kappa/h}, \quad \text{for} \quad t = nh \leq e^{\kappa/h}. \tag{6.51}$$

The constants C and κ are independent of n, h, μ (for h, μ sufficiently small) and of any initial value $(p_0, q_0, s_0) \in \tilde{T}_{(\omega,\Omega)}$.

Proof. The proof is similar to the proof of [HLW06] chapter X theorem 5.4, but is adapted to the case of a quasi-periodic perturbations. As in [HLW06] chapter X theorem 5.4 we divide the proof into parts (a),(b) and (c) (part (c) is unchanged).

(a). By the BEA of Section 6.3.6, the symplectic method induces the modified Hamiltonian

$$\tilde{K}^{[\tilde{N}]}(q,p,t) = H_0(a) + \mu H_1(a,\theta,t) + \sum_{j=N+1}^{\tilde{N}} h^{j-1} H_j(a,\theta,t)$$

where the $H_j(a,\theta,t)$ consist of terms involving only sums, scalings and producs of the $H_{\tilde{j}}$, H_0, H_1 and their derivatives for $N \leq \tilde{j} \leq j$. This implies that $H_j(a,\theta,t)$ are again quasi-periodic with frequency $\Omega \in \mathbb{R}^s$. We denote by $\tilde{H}_j$ their quasi-periodic extensions, with variable $\Theta \in \mathbb{R}^s$, such that

$$\mathcal{K}^{[\tilde{N}]}(a,\theta,\Theta) = H_0(a) + \mu\tilde{H}_1(a,\theta,\Theta) + \sum_{j=N+1}^{\tilde{N}} \tilde{H}_j(a,\theta,\Theta),$$

where $H_j(a,\theta,t) = \tilde{H}(q,\theta,t\Omega)$ and

$$\tilde{K}^{[\tilde{N}]}(q,p,t) = \mathcal{K}^{[\tilde{N}]}(a,\theta,t\Omega).$$

We introduce next the conjugate action variables $b = (b_1, \cdot, b_s) \in \mathbb{R}^s$ of $\Theta \in \mathbb{R}^s$ and consider in extended phase space $\mathbb{R}^{n+s} \times \mathbb{T}^{n+s}$ the Hamiltonian (as in Theorem 3.9)

$$\tilde{\mathcal{K}}^{[\tilde{N}]}(a,b,\theta,\Theta) = \mathcal{K}^{[\tilde{N}]}(a,\theta,\Theta) + \Omega\cdot b = H_0(a) + \mu\tilde{H}_1(a,\theta,\Theta) + \sum_{j=N+1}^{\tilde{N}} \tilde{H}_j(a,\theta,\Theta) + \Omega\cdot b.$$

By applying the KAM theorem 3.9 to $\tilde{\mathcal{K}}^{[\tilde{N}]}$ we find that there exists for $\epsilon, \mu > 0$ small enough a transformations from extended phase space (with s variables $b_1 \ldots b_s$ conjugate to the angle variables $\psi = (\psi_1, \ldots, \psi_s)$ $(q,b,p,\psi) \to (c,\bar{\psi}) \in \mathbb{R}^{n+s} \times \mathbb{T}^{n+s}$ which is $\mathcal{O}(h^p + \epsilon)$ close to the identity. Furthermore, using the last part of the KAM theorem 3.9, Equation (3.4) find that we are in the setting of the proof of [HLW06] part (a) and may use from this point their arguments untill part (b).

(b). We use again identically the proof of [HLW06], with the observation that

$$\left\|\psi_h(y_0,t_0) - \phi^{[\tilde{N}]}_{t_0+h,t_0}(y_0)\right\|_1 \leq \|y_j - \tilde{y}(t_j,t_{j-1},y_{j-1})\|_1\,,$$

(where the second term is in their notation). This inequality holds since we may choose $y_j = (\psi_h^n(q_0,p_0,t_0), t_0 + jh, \tilde{s}(jh))$, where $s(t) \in \mathbb{R}^s$ is the exact solution of the newly introduced action variable of the modified Hamiltonian (transformed to the old coordinates), such that in fact

$$\left\|\psi_h(y_0,t_0) - \phi^{[\tilde{N}]}_{t_0+h,t_0}(y_0)\right\|_1 = \|y_j - \tilde{y}(t_j,t_{j-1},y_{j-1})\|_1\,.$$

We may now use the rest of the proof in [HLW06] to conclude the bound of Equation (6.51).

To conclude the proof we notice the following: At this point we have $(n+s)$-dimensional invariant tori $\mathcal{T}$ in the extended phase space $\mathbb{R}^{n+s} \times \mathbb{T}^{n+s}$, which are invariant with respect to the flow of the Hamiltonian $\tilde{\mathcal{K}}(a,b,\theta,\Theta)$. However, we want tori in the space $(a,\theta,\Theta) \in \mathbb{R}^n \times \mathbb{T}^{n+s}$ which are invariant with respect to the flow of the Hamiltonian $\tilde{K}$. This conclusion comes from using Remark 3.10. Thus, in the notation of this Theorem we have $\tilde{H} = \mathcal{K}$ and the tori are simply a projection of the tori $\mathcal{T}$ in $\mathbb{R}^{n+s} \times \mathbb{T}^{n+s}$ on the space $\mathcal{T}$ in $\mathbb{R}^n \times \{0\}^s \times \mathbb{T}^{n+s}$. □

One may notice that in the periodic case, $s = 1$ we find that $\tilde{H}$ (as in the previous Theorem) equals the modified Hamiltonian i.e. $\tilde{H} = \mathcal{K}^{[\tilde{N}]} = \tilde{K}$.

Remark 6.17. *In the periodic case $s = 1$ with $\Omega = 2\pi$, an easy consequence of this Theorem is that symplectic methods satisfying the bound* (6.50) *have 'approximately invariant' tori in the numerically approximated 2π Poincaré map (see also Remark 3.10).*

In particular, using Proposition 6.14 one sees that symplectic RK methods used on a 2π-periodically perturbed, completely integrable Hamiltonian system have 'approximately invariant' tori in the numerically approximated 2π Poincaré map.

6.6 Approximate KAM theory applied to the the induced splitting method on the tidal wave system: “default” and “Simple-B” case

In this Section we use Theorem 6.16 on the tidal wave system. In particular, we consider the symplectic splitting method of Section 4.4 applied to the parameter sets “Default” and “Simple-B”.

The splitting method applied to the tidal wave system, Ψ_h of Equation (4.10) reduces in this case ($k = l, f = 0, r = 2$) to

$$\Psi_h(q, p, t) = \begin{pmatrix} \psi^*_{1,\tilde{h}_1(t+h/2,h/2)} \circ \psi^*_{2,h/2} \circ \psi_{2,h/2} \circ \psi_{1,\tilde{h}_1(t,h/2)}(q, p) \\ t + h \end{pmatrix} = \begin{pmatrix} \tilde{\psi}^*_{1,h/2} \circ \tilde{\psi}^*_{2,h/2} \circ \tilde{\psi}_{2,h/2} \circ \tilde{\psi}_{1,h/2}(q, p) \\ t + h \end{pmatrix}.$$

As mentioned, the SE method ψ_1 is equal to the exact flow. The modified vector field of the method $\psi_{2,h/2} = \tilde{\psi}_{2,h/2}$ is calculated using Equations (6.30) and (6.32) (or the non-autonomous version) using *Mathematica* code as shown in Appendix E, where the modified vector fields are also presented.

As shown in Figure 19 it seems that the scaled Hamiltonians $(-2C_\delta k^4)^{-j} H_j$ first reduce in size, but afterwards increase, so that the expansion is seemingly only formal and non-convergent (see also [HLW06] chapter IX.7). This form of the figure very much resembles the estimates of Proposition 6.14 of the form $(cj)^{j-1}$. Using the BCH formula and the theory in Section 6.2.2 (and the *Mathematica* code in Appendix E) one finds the first few terms of the modified Hamiltonian of the method Ψ of which the first four, in the case of the “Default”

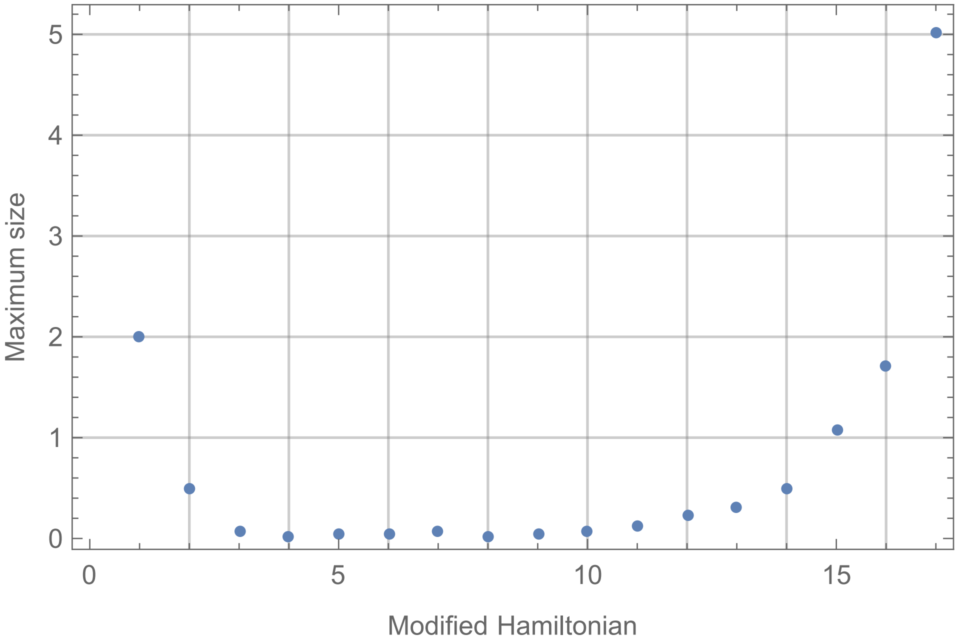


Figure 19: The maximum (see also the values of the parameter sets in Section 1.5) of the absolute value of the scaled Hamiltonians $(-2C_\delta k^4)^{-j} H_j$ for $1 \leq j \leq 17$. One sees that the size increases significantly towards the end, so that the non-scaled Hamiltonians are not suspected to converge at least for large values of $-2C_\delta k^4$.

and "Simple-B" parameter sets terms, are of the form

$$
\begin{aligned}
H_1 &= 2C_\delta k^4(\cos(p)-\cos(q)) + k(p-q)\cos(t) + s \\
H_2 &= \frac{1}{2}C_\delta k^5\left(2C_\delta k^3 \sin(p)\sin(q) + \cos(t)(\sin(q)-\sin(p))\right) \\
H_3 &= \frac{1}{24}\big(4C_\delta^3 k^{12}\left(\sin^2(p)\cos(q) - \cos(p)\sin^2(q)\right) + 2k\cos(t)\left(2C_\delta^2 k^8(2\sin(p)\cos(q)+\cos(p)\sin(q)) - p + q\right) \\
&\quad + C_\delta k^6\cos^2(t)(\cos(q)-\cos(p)) + 2C_\delta k^5 \sin(t)(\sin(q)-\sin(p))\big) \\
H_4 &= -\frac{1}{24}C_\delta k^5\big(\cos(t)\left(\sin(p)\left(2C_\delta^2 k^8\sin(p)\sin(q)-1\right)+\sin(q)\right) + C_\delta k^4\left(C_\delta^2 k^7 \sin(2p)\sin(2q) - 2\sin(p)\cos(q)\sin(t)\right) \\
&\quad + C_\delta k^5 \sin(p)\sin(q)\cos^2(t) - C_\delta k^5\cos(p)\cos(q)\cos(t)\left(2C_\delta k^3(\sin(p)-\sin(q))+\cos(t)\right)\big)
\end{aligned}
$$

We next try to find an exponential estimate of the form of Equation (6.50) (We will not consider the domain in the next section i.e. for which $(q,p) \in \mathbb{R}^2$ this holds). We start with the sympletic RK method $\psi_{2,h}$ applied to the Hamiltonian $L_2(q,p) = 2C_\delta k^4(\cos(p)-\cos(q))$. We immediately find from Proposition (6.14) the bound (ommiting the entries (q,p) of the function)

$$\left\|\psi_{2,h} - \tilde{\phi}_{2,h}^{[\tilde{N}]}\right\|_1 \leq hCe^{-h_0/h}$$

for some $C, h_0 > 0$. Where $\tilde{\phi}_{2,h}^{[\tilde{N}]}$ is the flow of the ODE induced by the modified Hamiltonian $L_2^{[\tilde{N}]}$ of the Hamiltonian L_2 truncated after terms of order $h^{\tilde{N}}$, for some $\tilde{N} \in \mathbb{N}$.

Finally, we split the error

$$\left\|\begin{pmatrix}\psi_{2,h/2}\circ\psi_{1,\tilde{h}_1(t,h/2)} \\ t+h\end{pmatrix} - \tilde{\Phi}_{h/2}^{[\tilde{N}]}\right\|_1 \leq \left\|\psi_{2,h/2}\circ\psi_{1,\tilde{h}_1(t,h/2)} - \tilde{\phi}_{2,h/2}^{[\tilde{N}]}\circ\psi_{1,\tilde{h}_1(t,h/2)}\right\|_1 + \left\|\begin{pmatrix}\tilde{\phi}_{2,h/2}^{[\tilde{N}]}\circ\psi_{1,\tilde{h}_1(t,h/2)} \\ t+h\end{pmatrix} - \tilde{\Phi}_{h/2}^{[\tilde{N}]}\right\|_1,$$

where $\tilde{\Phi}_h^{[\tilde{N}]}$ is the flow of the modified vector field of the symplectic method $\psi_{2,h/2}\circ\psi_{1,\tilde{h}_1(t,h/2)}$ applied to the Hamiltonian of the tidal wave system. The second term may be estimated using the BCH formula but in this thesis we end on the assumption that this term may be estimated by an exponential estimate of the form of Equation (6.50), and that this also implies that the conditions of the approximate KAM theorem 6.16 are satisfied.

Assumption 6.18. *We assume that the following bound*

$$\left\|\psi_{2,h/2}\circ\psi_{1,\tilde{h}_1(t,h/2)} - \tilde{\Phi}^{[\tilde{N}]}\right\|_1 \leq h\tilde{C}e^{-\tilde{h}_0/h}$$

exists for some $\tilde{C}, \tilde{h}_0 > 0$, which implies the bound (for some $\hat{C}, \hat{h}_0 > 0$)

$$\left\|\psi_{2,h/2}\circ\psi_{1,\tilde{h}_1(t,h/2)} - \tilde{\Phi}^{[\tilde{N}]}\right\|_1 \leq h\hat{C}e^{-\hat{h}_0/h}.$$

We assume furthermore that this bound implies an exponential bound of the form

$$\left\|\Psi_h - \hat{\Phi}_h\right\|_1 \leq h\mathcal{C}e^{-k_0/h}$$

for some $\mathcal{C}, k_0 > 0$.

If this assumption holds where the domains are the 'cell' B as define in Section 3.3, then we may use the KAM theorem 6.16 to conclude the existence of 'approximately invariant' tori in the tidal wave system (using also Remark 6.17) . Indeed then we have used a symplectic method, with an exponential bound of the form of Equation (6.50) on a periodially perturbed completely integrable system which satisfies the conditions of Theorem 3.9 (shown in Section 3.3).

7 Conclusion

In this thesis we have shown that, for the parameter sets "Simple-B" and "default", the theoretical, unperturbed tidal wave system (Section 1.5) is completely integrable. Furthermore we have found explicitly (in quadrature) the action-angle coordinates and have applied a KAM theory for periodic perturbations which showed that there exist KAM tori in the 2π-Poincaré map of the perturbed tidal wave system (Section 3). The KAM theorem was from [Jor91; JS96; Sev07].

Next, a special splitting method for time-affine (Definition 2.2) Hamiltonian ODE was considered, which was developed in Section [Wal21]. This method splits the time-affine Hamiltonian into forced vector fields (Definition 2.1), on which an induced method (Definition 4.5) is used which can easily be made symplectic (Section 4). We applied this method to the tidal wave system with the parameter "Simple-B" and "default" and numerically approximated the 2π-Poincaré map using this symplectic splitting method for the time-affine tidal wave Hamiltonian. There we saw, as in [Wal21] that KAM tori seemed to be present in the numerically approximated 2π-Poincaré map, for specific values of the (perturbation) parameter.

Afterwards, we used BEA to interpolate/embed the symplectic splitting method with/into a Hamiltonian flow. Two types of BEA were considered: A non-autonomous version, based on [Moa03; Moa05] and an autonomous version, based on [GS86; CMS94; LR04; HLW06], (also see Section 5.2) .

In Section 5 we used the non-autonomous version of BEA, *non-autonomous flow interpolation* (see Sections 2.2.3 and 5.2) to embed the symplectic integrator into a non-autonomous, periodically (in time) perturbed Hamiltonian flow. We applied the KAM theorem for periodically perturbed Hamiltonian systems. Doing so one could prove the existence of KAM tori in the numerically approximated 2π-Poincaré map of the *unperturbed* tidal wave system, although it was not clear if the step size was a perturbation parameter.

Therefore, and due to the consideration of rounding errors, an 'approximate' KAM theorem was stated in Section 6. This was constructed as in [HLW06] chapter X.5 but now using the non-autonomous version of BEA: modified equation analysis. This approximate KAM theorem proved, *up to an assumption* the existence of approximately invariant tori in the numerically approximated 2π-Poincaré map of the unperturbed and perturbed tidal wave ODE, where in this case the step size h *was* seen as a perturbation parameter.

Furthermore, in Section 2 and 4. The structure of non-autonomous Hamiltonian ODE and structure-preserving methods were considered. This consisted mostly of a short review of some developments. In particular, in this thesis we saw that a symplectic method applied to a non-autonomous Hamiltonian ODE has a modified vector field a non-autonomous Hamiltonian vector field. Using this modified, non-autonomous Hamiltonian, we were able to prove an approximate KAM theorem for (quasi-)periodically perturbed Hamiltonian systems. Therefore, symplecticity (Definition 2.7) of the numerical method already seems to be a good 'structure' to preserve for 'structure-preserving integrators'.

8 Further research

We give a list of topics which could be interesting for the study of the tidal wave system and/or structure preserving methods for non-autonomous Hamiltonian ODE.

- First of all, one could show the truth or negate the Assumptions 6.6
- First of all, one could develop the theory for non-autonomous Hamiltonian geometry. As mentioned this could be contact geometry, presymplectic geometric, (per)cosymplectic or via dynamical systems constructions as discussed in Section 2.2.
- Additionally, one could develop a theory for structure preserving integrators which preserve this structure for non-autonomous Hamiltonian. Possibly one can use the BEA of Section 6.3.6, for a non-autonomous modified Hamiltonian (or of Appendix D) to study symplectic integrators on non-autonomous

ODE. To this end, one could also consider other numerical integration schemes, such as the methods in [MO14] or the methods based on the Magnus series [BM01], splitting schemes for non-autonomous, perturbed ODE [Bla+10] or the Magnus expansion for non-autonomous, separable ODE as in [BM01; BC06] to integrate the tidal wave system.

- One could also take symmetries/reversibility into account, see e.g. [HLW06; Sev06] In particular, the unperturbed flow is a special case of the ABC flow, which is studied due to all its many symmetries.
- Instead of using continuous KAM theory, one could use discrete KAM theory (for twist maps) on the 2π-Poincaré map in the theoretial tidal wave system, see Figure 10. In particular, the application of KAM theory via twist maps to $1 + 1/2$ degree of freedoms has been previously studied [Con16; CF]. The approach then typically consists of finding an approximation (in the form of an expansion), to approximate the 2π map for example the "orbit expansion" in [BRZ94] or other expansions in [BBM22].
- Similarly for the numerically approximated 2π-Poincaré map, one could use a discrete strategy, possibly using again expansions as in [BBM22] as was discussed in 5.1.
- One can do better than 'approximate KAM theorems'. The stronger version is the Nekhoroshevs theorem. This has already been done by Moan [Moa03] on numerical systems. See also [WM14; FW16] for the fluid dynamics perspective.
- As mentioned, one could use the tidal wave system to test if one could find rigorous bounds for the destruction of particular invariant tori, as in [CFP87a] for the pendulum.

A Notation and some preliminaries in ODE and dynamical systems theory

First we will set notation for derivatives, then we will give some short preliminaries on ODE and dynamical systems theory. If $f : \times_{i=1}^{k} \mathbb{R}^{n_i} \to \mathbb{R}^m$ is differentiable then we will denote by $Df : \times_{i=1}^{k} \mathbb{R}^{n_i} \to \mathbb{R}^{m \times n}$ the (total) derivative (in the standard basis), where $n := \sum_{i=1}^{k} n_i$. If points in $\mathbb{R}^{n_i}$ are denoted by x_i we will denote the partial derivative with respect to the variables in $\mathbb{R}^{n_i}$ by

$$D_{x_i} f := D_i f := \partial_{x_i} f := \frac{\partial f}{\partial x_i} \in \mathbb{R}^{m \times n_i}$$

and sometimes as f_{x_i} if no confusion arises. Most often f is a vector field such that $k = 1$ and $n = m$ (autonomous) or $n = m + 1$ (non-autonomous). Then we will additionally use the notation $(x_1 =: x, x_2 := t)$

$$\dot{f} := D_t f \qquad f' := D_x f.$$

The notation t hints that it is often thought of as time. Although this could provide some intuition for the geometry of the integral curves, one should sometimes rather see t as any variable describing an integral curve (defined below), dropping the physical intuition.

One could say that the study of ODE consists of two theories: Geometry and dynamical systems. Since we are working in Euclidean space, the (differential) geometry is very implicit and we do not need a lot of tools from differential geometry, moreover, this thesis has little emphasis on the dynamical systems part so we need little tools from this part as well. In the remainder of this section, which is therefore short, we will define ODE from an (analytic) geometrical and geometrical perspective.

A.1 ODE vs IVP and flows vs integrable curves

A first distinction in ODE theory on Euclidean space $D \subset \mathbb{R}^n$ is between ordinary differential equations (ODE) and initial value problems (IVP). An ODE problem is defined by a vector field f only. An IVP problem additionally needs an initial point $y_0 \in D$ in space (and time $t_0 \in \mathbb{R}$, if non-autonomous).

Suppose f is a vector field on $\mathbb{R}^n$, i.e. $f : \mathbb{R}^n \times \mathbb{R} \supset D \times I \to \mathbb{R}^n$, with $D \times I$ open.

Definition A.1. *The IVP of* (f, y_0, t_0) is the (set of) equation(s)

$$\dot{y}(t) = f(y(t), t) \text{ where } y \in C^1(J, \tilde{D}), \quad \text{such that } y(t_0) = y_0.$$

In other words y is in the set a differentiable paths with specific 'initial value' y_0 at t_0, where $J \subset I$, $\tilde{D} \subset D$ open. ∅

A solution of this IVP is a specific differentiable path z for which the IVP is true and is called an *integrable curve*. Existence and uniqueness results of these solutions can be found in for example [Hal80; DE02].

An ODE does not need an initial value.

Definition A.2. *The ODE of* f is the (set of) equation(s)

$$\dot{y}(t) = f(y(t), t), \text{ where } y \in C^1(J, \tilde{D}). \qquad \varnothing$$

A solution is defined similarly and again called an *integral curve*. Existence and uniqueness results follow by adding an initial value (and time), transforming the ODE into an IVP and subsequently considering existence. In general, there is no more uniqueness since there is a set $B \subset D$ such that every $y_0 \in B$ produces an integral curves of the ODE. For non-autonomous ODE one includes also a set $K \subset I$ of initial times t_0.

We now encounter a second distinction, to solve this non-uniqueness. There exists another, more global version, of a solution of an ODE or initial value problem called a *flow* ϕ which naturally includes such B

(and in the non-autonomous case also K) in the domain $\phi : B \times K \times \tilde{J} \to \mathbb{R}^n$ (one actually needs to be more careful with the domain $\tilde{J}$, [DE02] chapter 2). One usually denotes $\phi(y_0, t_0, t) =: \phi_{t,t_0}(y_0)$. Solution-paths and flows are related by $y(t) = \phi((y_0, t_0), t)$ if $y(t_0) = y_0$

ODE can be defined as well using flows, now with a unique solution.

Definition A.3. *The ODE of f* is the equation

$$\dot{\phi}_{t,t_0}(y) = f(\phi_{t,t_0}(y), t), \qquad \text{or} \qquad \dot{\phi} = f \circ (\phi, \pi_t),$$

where $\phi_{t,t_0}(y) := \phi((y, t_0), t) \in C^1(\tilde{D} \times J, \tilde{D})$ for some $\tilde{D} \subset D \times I$ open and π_t is the projection on the t coordinate.

If $\tilde{D} \times J = \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ then the flow is called *entire*. We denote $\phi_t := \phi(\cdot, t)$ and $\phi_{t,t_0} := \phi((\cdot, t_0), t)$. If f is autonomous then the flow ϕ is independent of t_0 and we denote additionally $\phi_t := \phi_{t,0}$.

A flow of f which is entire defines, if f is autonomous, a *one-parameter group* $\phi \mapsto \phi_t$ ([DE02], chapter 2) or, if f is non-autonomous, a *groupoid* ϕ_{t,t_0} (such that $\phi_{t,t_1} \circ \phi_{t_1,t_0} = \phi_{t,t_0}$). In non-autonomous ODE/dynamical systems theory, this groupoid is sometimes called a *two-parameter group* or a *process* [KR11] and the groupoid property of ϕ is sometimes called the co-cycle property [Wig03].

A.2 Dynamical systems theory

Thus, an ODE induces in the autonomous one-parameter group of diffeomorphisms, which are studied from another viewpoint by dynamical systems theory. The theory of dynamical systems will is used both for theoretical systems, including Hamiltonian systems (c.f. [Ver06; MO17]), as well as for numerical methods (c.f. Section 4.2 or [SH98]), where one could say that, in particular numerical methods for Hamiltonian systems (c.f. [LR04; HLW06]).

Definition A.4. A discrete respectively continuous dynamical system (DS) is a tuple (ϕ, X), X a set and ϕ respectively a one-parameter group $\phi : \mathbb{R} \times X \to X$ or a map $\phi : X \to X$. $\varnothing$

Usually more structure is present (topological, (symplectic) geometrical) and the DS adapts accordingly. The usual notions of *phase space, orbits, limit sets* and *invariant sets* can be introduced as in e.g. [BS02].

A.3 Coordinate transformation and pushforward vector fields

The isomorphisms of geometry are coordinate transformations.

Definition A.5. A *C^k-coordinate transformation g* is a C^k-diffeomorphism. For a coordinate transformation g and vector field f, the vector field $g^*(f)$ on $\mathbb{R}^n$ defined by

$$g_*(f)(y,t) = (Dg(g^{-1}(y))) \cdot f(g^{-1}(y), t) \qquad \text{or} \qquad g_*(f) = D(g^{-1}) \circ (f \circ g^{-1}) \tag{A.1}$$

is called the *pushforward* ODE of f/vector field f by g is the vector field $g_*(f)$. $\varnothing$

The pushforward vector field $g_*(f)$ of f is natural if one notices that $g \circ \phi$ and $g \circ y$ (and even $g \circ \phi_{t,t_0} \circ g^{-1}$) are (path)solutions to the ODE of $g_*(f)$. If $\phi_{t,t_0} \in C^1$ is to the ODE of f.

The term "coordinate transformation" (as opposed to "C^k- diffeomorphism") serves to emphasize the setting of ODE (as opposed to the setting of the space/manifold). This is useful because in the setting of ODE one has available an extra variable, coming from the dynamics, the 'time t' dimension/variable, which does not exist on the underlying manifold. We see next how the time-dependent versions of the coordinate transformation and pushforward vector field arise.

Remark A.6. *If $g(y,t) =: g_t(y)$ is a time-dependent coordinate transformation (depending differentiably on time t), then $g \circ \phi$ (but not $g \circ \phi \circ g^{-1}$) is the solution to the ODE of $g_*(f)$, defined by*

$$g_*(f)(y,t) = D_y(g_t^{-1})(y) \cdot \big(f(g_t^{-1}(y), t) - D_t(g_t^{-1})(y)\big) \qquad \textit{or} \qquad (g_t)_*(f) = D(g_t^{-1}) \circ \big(f \circ g_t^{-1} - D_t(g_t^{-1})\big)$$

where the identity $D_t(g(g^{-1}(y,t),t) = 0 = D_t g(z,t) + D_y g(z,t) D_t g^{-1}(y,t)$ with $z = g^{-1}(y,t)$ was used. A special example of such a time-dependent coordinate transformation is a one-parameter group of transformations/flow ϕ_t.

If f is non-autonomous, such that initial time t_0 is relevant, then we may equivalently take $g_t : \mathbb{R}^n \to \mathbb{R}^{n+1}$, changing t_0. In this case the statements and equations above are still valid (but changed accordingly e.g. $f(g_t^{-1}(y),t) \mapsto f(g_t^{-1}(y))$.

Furthermore, one may also consider C^k diffeomorphism g_t with domain $\mathbb{R}^{n+1}$, depending as well on initial time t_0. Then, for $(g_t)_*(f)$ to make sense we need not a vector field f on $\mathbb{R}^n$ but an extension to some vector field g on $\mathbb{R}^{n+1}$, this rises the demand for an extension of phase space, incorporating initial time into the geometry.

A.4 Conjugacy of dynamical systems versus coordinate transformations

The epi-/isomorphisms of DS are respectively called *(semi)conjugacy's*.

Definition A.7. Given two discrete (continuous) DS, $D_1 = (\phi_1, X_1), D_2 = (\phi_2, X_2)$, where ϕ_i are endomorphisms (flows) then a *(semi)conjugacy* from D_1 to D_2 is a (surjective) bijective map $g : X_1 \to X_2$ such that

$$g \circ \phi_1 = \phi_2 \circ g.$$

Usually g preserves 'more structure' than only the 'dynamical systems structure', for example phase space (X_i) is a smooth manifold and g is a diffeomorphism.

Time-independent isomorphisms of ODE (coordinate transformations) and the isomorphisms of dynamical systems (conjugacy's) behave well together: Suppose that two continuous DS $D_i = (\phi_i, \mathbb{R}^n)$ are induced by ODE of f_i i.e. $\dot{\phi}_{i,t} = f_i \circ (\phi_t, \pi_t)$ for some vector fields f_i. Then by differentiation of $g \circ \phi_1 \circ g^{-1}$ we find that $f_2 = g_*(f_1)$, if g is a differentiable conjugacy from D_1 to D_2 Conversely, if g is a coordinate transformation on the ODE of f_1, and ϕ_2 is the flow of $g_*(f_1)$ then, from Section A.5, $\phi_2 := g \circ \phi_1 \circ g^{-1}$ so that g is a conjugacy.

B Bell polynomials

Commutative Bell polynomials are useful to study cumulants [Wik21], compositions of analytic formulas (Faà di Bruno's formula) and come up as well in the Lie series and BEA of forced ODE. Non-commutative Bell polynomials are used for non-autonomous Lie series, see e.g. [ELM14], [LM11].

B.1 Commutative Bell polynomials

In the study of cumulants, partitions of sets (in particular the sizes of the subsets) are an object of study and helpful to understand Bell polynomials. We define for $j \in \mathbb{N}$ the commutative/unordered partitions of j as the set of commutative tuples $p(j) := \{(n_1, \dots, n_j) \in \mathbb{N}^j \mid \sum_{i=1}^{j} i n_i = j\}\}$. For example $4 = 0 + 4 = 1 + 3 = 2 + 2 = 1 + 1 + 1 + 1$ such that $p(4) = \{(0,0,0,1), (1,0,1,0), (0,2,0,0), (4,0,0,0)\}$. We define as well $p_i(j) = \{((n_1, \dots, n_j) \in p(j) \mid \sum_{i=1}^{j} n_i\}$. For example $p_2(4) = \{(1,0,1,0), (0,2,0,0)\}$.

The commutative, exponential Bell polynomials B_i, $i \in \mathbb{N}$ are defined as polynomials in the commutative free algebra $\mathcal{D}$ over $\mathbb{R}$ on a countable set of letters (or variables) x_i, so $\mathcal{D} = \mathbb{R}\langle (x_i)_{i \in \mathbb{N}} \rangle$. An element $w \in \mathcal{D}$ is a polynomial (or *sentence*) and is called a *word* if it consists of a monomial.

The commutative algebra $\mathcal{D}$ is graded on the variables by $|x_i| = i$ and on a monomial $w \in \mathcal{D}$ of degree k (i.e. $cw = x_{j_1} x_{j_2} \dots x_{j_k}$ for $j_i \in \mathbb{N}$, $i \geq 1$, $c \in \mathbb{R}$) as $|w| = \sum_{i=1}^{k} |x_{j_i}|$ and linearly extended to polynomials. We denote by $\#(w)$ the degree of a monomial w (e.g. $\#(x_{j_1} x_{j_2} \dots x_{j_k}) = k$). For example $|x_1 x_2^3 x_4| = 1{+}3{\cdot}2{+}4 = 11$ and $\#(x_1 x_2^3 x_4) = 5$. We define $\mathcal{D}_i := \{w \in \mathcal{D} \,|\, |w| = i\}$ and $\mathcal{D}_{i,j} := \{w \in \mathcal{D}_i \,|\, \#(w) = j\}$.

A word $w \in \mathcal{D}_{i,j}$ with coefficient $c = 1$ is 1-1 related to an unordered partition $l \in p_j(i)$ by the relation $w_l = x_1^{l_1} \dots x_j^{l_j}$ with inverse $l(w)$. For example the partition $l = (2,0,1,0,0) \in p_3(5)$ is related to the monomial $w(l) = x_1^2 x_3 \in \mathcal{D}_{5,3}$ and $l(x_1^2 x_3) = (2,0,1,0,0)$. Henceforth we skip the notation of partitions (in the commutative case) and denote $w_i := l(w)_i$ (thus $w_i = (x_{j_1} \dots x_{j_k})_i \neq x_{j_i}$) and $\#(l) := \#(w_l)$ for a monomial w and $w! := l(w)! = w_1! \dots w_i!$.

Definition B.1. The *commutative, exponential Bell polynomial* $B_i = B_i(x_1, \dots, x_i)$ is defined as the sum over monomials $w \in \mathcal{D}_i$ (or partitions $l \in p(i)$) of grade i, such that

$$B_i = \sum_{w \in \mathcal{D}_i} \alpha(w) w = \sum_{l \in p(i)} \alpha(w_l) w_l,$$

where the coefficient $\alpha(w)$ can be interpreted in a combinatorial way using set-partitions (e.g. [Wik21]) and equals [ELM14]

$$\alpha(w) = \frac{i!}{\prod_{k=1}^{i} w_k! \, (k!)^{w_k}} = \frac{i!}{w! 1!^{w_1} \dots i!^{w_i}}. \tag{B.1}$$

For example

$$B_3(x_1, x_2, x_3) = 3! \left(\frac{1}{3!} x_3 + \frac{1}{2!} x_1 x_2 + \frac{1}{3!} x_1^3 \right) = x_3 + 3x_1 x_2 + x_1^3.$$

Thus

$$B_i = \sum_{w \in \mathcal{D}_i} \frac{i!}{w! \prod_{k=1}^{i} (k!)^{w_k}} w = \sum_{l \in p(i)} \frac{i!}{l! \prod_{k=1}^{i} (k!)^{l_k}} \prod_{\ell=1}^{i} x_{l_\ell} \tag{B.2}$$

Furthermore, the *partial, commutative, exponential* Bell polynomial $B_{i,j}$ is defined as the part of B_i of degree k, e.g. $B_{3,1} = x_3$, $B_{3,2} = 3x_1 x_2$, $B_{3,3} = x_1^3$ (c.f. Table 4) such that $B_i = \sum_{j=1}^{i} B_{i,j}$ and

$$B_{i,j} = \sum_{w \in \mathcal{D}_{i,j}} \frac{i!}{w! \prod_{k=1}^{i} (k!)^{w_k}} w = \sum_{l \in p_j(i)} \frac{i!}{l! \prod_{k=1}^{i} (k!)^{l(w)_k}} x_{l_1} \dots x_{l_i}. \tag{B.3}$$

B.2 Non-commutative Bell polynomials

The non-commutative, exponential Bell polynomials $\hat{B}_i$ and its partial version $\hat{B}_{i,j}$ are non-commutative versions of the commutative Bell polynomials. They are polynomials in the algebra $\hat{\mathcal{D}}$, defined as the non-commutative version of $\mathcal{D}$. We defined similarly $\hat{\mathcal{D}}_i$ and $\hat{\mathcal{D}}_{i,j}$ the subsets of grade i and degree j. A word $w \in \mathcal{D}_{i,j}$ is not in 1-1 relation with a partition in the non-commutative case, however it is in 1-1 relation with the set of tuples $l := (l_1, \dots, l_j)$ ($l_i \in \mathbb{N}$) such that $\sum_k l_k = i$, we denote the $\mathcal{P}_j(i) := \{l \in \mathbb{N}^j \mid \sum_k l_k = i\}$ (which can be seen as non-commutative partitions). For example $l = (2,3,1,2) \in \mathcal{P}_4(8)$ is related 1-1 to the word $w = x_2 x_3 x_1 x_2$. Similar, to the commutative version, we define $w! = l(w)! = \prod_{k=1}^{j} l_k!$ for $l \in \mathcal{P}_j(i)$. One has $\hat{B}_i = \sum_{j=1}^{i} \hat{B}_{i,j}$ and the non-commutative, partial, exponential Bell-polynomials $\hat{B}_{i,j}$ can be written as [LM11; ELM14]

$$\hat{B}_{i,j} = \sum_{w \in \hat{\mathcal{D}}_{i,j}} \frac{i!}{w!} \kappa(w) w = \sum_{l \in \mathcal{P}_j(i)} \frac{i!}{l!} \kappa(l) \prod_{k=1}^{j} x_{l_j} \tag{B.4}$$

where κ is defined as

$$\kappa(w) = \kappa(w_l) := \frac{\prod_{k=1}^{j} l_k}{\prod_{k=1}^{j} \left(\sum_{n=1}^{k} l_n \right)}. \tag{B.5}$$

For example (c.f. Equation (D.5))

$$\hat{B}_{3,2} = 3 \left(\frac{2}{6} x_2 x_1 + \frac{2}{3} x_1 x_2 \right) = x_2 x_1 + 2x_1 x_2 \quad \text{or} \quad \hat{B}_{4,3} = 12 \left(\frac{2}{8} x_1^2 x_2 + \frac{2}{12} x_1 x_2 x_1 + \frac{2}{24} x_2 x_1^2 \right) = 3x_1^2 x_2 + 2x_1 x_2 x_1 + + x_2 x_1^2.$$

C Introduction to backward error analysis

Backward error analysis (BEA) is a type of error analysis for numerical/approximation methods. It is an alternative to forward error analysis (FEA). The general idea of BEA is to find an approximate problem, to which a numerical approximation is the exact solution, as seen in Figure 20. There are multiple reasons to take BEA into consideration:

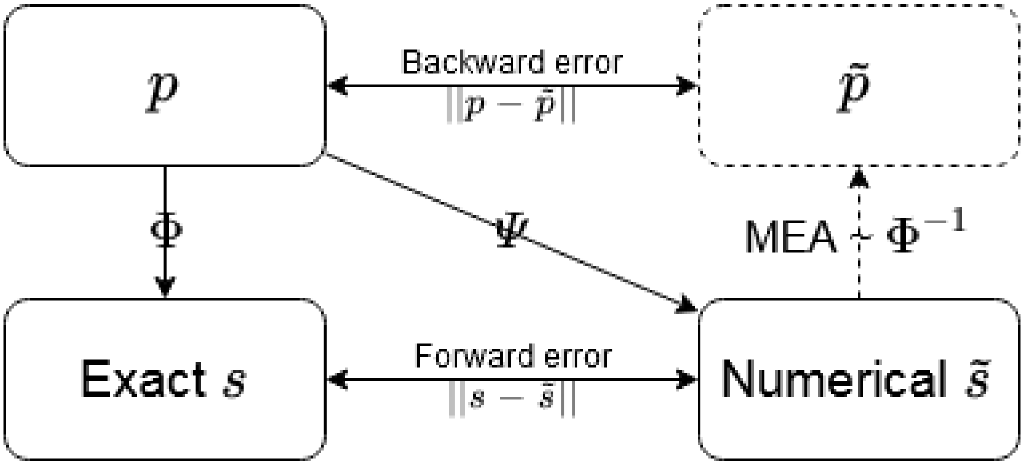


Figure 20: The approximate problem $\tilde{p} \in \mathcal{P}$ together with the BE $\|p - \tilde{p}\|$ and FE $\|s - \tilde{s}\|$ are shown in this figure. A solution operator Φ and numerical method Ψ are shown, which produce $s, \tilde{s}$ from a problem $p \in \mathcal{P}$. A way to find an approximate problem $\tilde{p}$ (in other words to find some kind of map Φ^{-1}) for ODE is by modified equation analysis (MEA), see Section 6.3.1

- **Explicitness:** The backward error, as opposed to the forward error, is sometimes **explicitly known**, or easier estimated, [CF13; Moi10].
- **Constructiveness:** It can be used to **construct approximation algorithms**, e.g. for a numerical linear algebra problem (Section C.1) or for ODE/IVP as in [Moi10; CHV07].
- **Allows forward error estimates:** It can be used as well to **analyse error** the forward error in numerical algorithms (for example to estimate the global error in IVP)[Moi10; CF13].
- **Use of perturbation theory:** Backward error relates a 'nearby problem' to the numerical method, in particular the use of perturbation theory to analyse numerical methods .
- **Justifies use of numerical methods for physical problems:** From a more philosophical point of view, the physicality of problems validates the use of BEA. In particular, the fact that physical problems are often perturbed in some sense (e.g. cars passing near a harmonic oscillator) combined with the fact that numerical methods solve a perturbed problem (shown using BEA) implies that, if both perturbations are in some sense similar, a numerical method which controls the backward error must give a useful solution, as is discussed in [Moi10], chapter 1, 2, & 5 and [Cor94].

We suppose that we have two normed spaces: A problem space $\mathcal{P}$ and solution space $\mathcal{S}$ and that we have some problem $p \in \mathcal{P}$, with solution operator $\Phi : \mathcal{P} \to \mathcal{S}$ and numerical method $\Psi : \mathcal{P} \to \mathcal{S}$. The goal is to construct Ψ such that $\|\Psi - \Phi\|$ is small. We assume for simplicity that Φ is injective.

Definition C.1. Given a problems $p \in \mathcal{P}$. The forward error is given by $\|s - \tilde{s}\|$, where $s = \Psi(p)$, $\tilde{s} = \Phi(p)$. and *backward error* is by $\|p - \tilde{p}\|$, where $\tilde{p} = \phi^{-1}(\Psi(p))$.

BEA is useful if there is some kind of continuity with respect to the problems: $\|s - \tilde{s}\| \leq k \|p - \tilde{p}\|$ such that a small backward error guarantees a useful solution (with small forward error) if $k \in \mathbb{R}$ is not too large. The number k is often called the *condition number*, and a problem is well-conditioned if k is 'not too large'[24].

Sometimes a well-conditioned problem is called "backward stable" [Cor94], but stability is usually a property which refers to numerical methods as opposed to the conditioning of a problem (and it can be argued that the perspective of BEA helps to distinguish between the two [Enr89]). However, this does indicate that there is also some kind of "forward stability" of the form $\|p - \tilde{p}\| \leq \tilde{k} \|s - \tilde{s}\|$.

As mentioned, the backward error can be used to construct algorithms or to analyse the forward error. It can be expected that an algorithm which has small backward error, applied to a well-conditioned problem, finds an accurate solution. One is inclined to think, however, that that such an algorithm is useless for ill-conditioned problems, but this is not the case, as is sometimes seen for chaotic IVP [Moi10].

C.1 Backward error analysis for numerical linear algebra

The origins of BEA lie in the theory of Numerical Linear Algebra (NLA) and are usually attributed to Wilkinson [Moi10; LR04; HLW06]. Therefore, this setting is used to introduce the method.

Consider the following NLA problem $p = (A, b)$:

$$\text{For matrix-vector pair } p = (A, b) \in \mathbb{R}^{n \times n} \times \mathbb{R}^n, \text{ calculate } x \in \mathbb{R}^n \text{ s.t. } Ax = b \tag{C.1}$$

and suppose an iterative method (BEA is also useful for direct methods [LR04]) is used with approximation x_i of x (i an iteration index).
The *forward error* of the NLA problem is $\|x_i - x\|$. The forward error is unknown since x is unknown. As mentioned, the backward error constructs an approximate problem (A_i, b_i) such that $A_i x_i = b_i$ (we assume very shortly uniqueness, which is certainly not the case). The *backward error* of the NLA problem is $\|(A_i, b_i) - (A, b)\|$. The backward error is known if (A_i, b_i) are explicitly constructed. It is sometimes called the *residual* or *defect*[25].

On well-conditioning: one finds

$$\frac{1}{\|x\|} \|x - x_i\| \leq \kappa(A) \frac{1}{\|b\|} \|b - b_i\|,$$

such that the *condition number of the matrix* A is given by $\kappa(A) = \|A\| \left\|A^{-1}\right\|$ (assuming $\det A \neq 0$). Well-conditioned problems have 'small' condition number of the associated matrix, and ill-conditioned matrices can be regularized (e.g. by preconditioning).

In general, either $A_i = A$ or $b_i = b$, as varying both is not useful, thus multiple approaches to BEA are possible. If $A_i = A$, numerical algorithms based on minimising the backward error can be constructed. These include Krylov subspace methods and in particular the celebrated GMRES.

Example C.2. *Consider the problem*

$$A = \begin{pmatrix} 0.01 & 0 \\ 0 & 1 \end{pmatrix} x = b$$

has solution $x_1 = (100, 1)$ *for* $b_1 = (1, 1)$ *and* $x_2 = (101, 1)$ *for* $b_2 = (1.01, 1)$*, such that* $\|x_1 - x_2\| / \|b_1 - b_2\| = 10000$*. The condition number is* $K(A) = 100$ *for the operator norm so the relative error can be multiplied by* 100*. Due to the ill-conditioning, this means that any algorithm which wants a forward error as small as* 10^{-n} *digits needs possibly* $n + 2$ *accurate digits in the backward solution.*

[24] Ill-conditioned problems can be regularised, which is the subject of the mathematical field of Inverse Problems

[25] Sometimes in NLA the residual is the term $\|b - b_i\|$ instead of $\|(A_i, b_i) - (A, b)\|$. For IVP the redisual, the defect and the backward error sometimes have different meaning [CF13], but not in this thesis

	Defect analysis	Shadowing	MEA
Vector field	Variable	Fixed	Variable
Initial conditions	Fixed	Variable	Variable
Analytic utility	Estimating global error using Gröbner-Alekseev formula [Moi10; CF13]	Identifying dynamics of ODE such as structural stability.	Studying qualitatively the numerical dynamical system, connect to perturbation theory. Studying ODE with geometric structures
Example of algorithm	Defect control	Many types of shadowing	Automatic MEA solvers [AC97] Structure preserving methods [CHV07] Adaptive grid-size [DM21].
Algorithmic utility	Bounding global error. Choosing step-size. Studying chaotic systems.	Calculating with small forward error in chaotic systems.	Understanding perturbed system shown in numerical figures

Table 2: Table showing properties of three types of BAE. The first two rows should not be taken to rigidly, as modified

C.2 Backward error analysis for initial value problems

The IVP $(f, (y_0, t_0))$ is considered (Definition A.2). The forward error is given by the global error, Equation (4.1). The backward error can be constructed by finding an approximate vector field $\tilde{f}$ and/or a modified initial condition/time $\tilde{y}_0$ or $\tilde{t}_0$. Moir [Moi10] suggests that this distinguishes three different types of BEA: *Defect analysis, shadowing and modified equation analysis MEA*, as shown in Table 2.

Since ODE, as opposed to IVP, have no initial value in the problem, Table 2 might suggest that defect analysis is the correct setting for BEA of ODE. However, in our opinion, the distinction is not rigid: modified equation analysis is useful for ODE as well. The distinction between BEA and defect analysis seems to be, mainly, the way in which it is constructed, see Table 3.
Thus, defect analysis is about finding an path $\tilde{u} : \mathbb{R} \to \mathbb{R}^n$ and using it to estimates the backward error $\|\tilde{u}_t(t) - f(\tilde{u}(t), t)\|$. In *defect control* algorithms, $\tilde{u} : \mathbb{R} \to \mathbb{R}^n$ is obtained by any kind of (differentiable) interpolant and quantitative results are then obtained, which can be related to the global error (using e.g. the Gröbner-Alekseev formula, [Enr89; KP94a; Moi10; CF13]) MEA, on the other hand, is about finding a one-parameter flow $\tilde{\phi}_t : \mathbb{R}^n \to \mathbb{R}^n$ such that $\left\|D_t\phi_{t,t_0}(y) - f(\tilde{\phi}_{t,t_0}(y), t)\right\| = \mathcal{O}(h^{q+1})$ for $q \in \mathbb{N}$ the order of approximation, for all y in the domain. The one-parameter flow, $\tilde{\phi}_t : \mathbb{R}^n \to \mathbb{R}^n$, is obtained by asymptotic expansion techniques to find a modified vector field $\tilde{f}_p$. Qualitative and quantitative results, such as those of perturbation type, can then be found.

	Defect analysis	Modified equation analysis
Focus	Pathwise solutions	One-parameter flow
Results	Quantitative (numerical)	Qualitative (analytic)
Interpolation type	Numerical Orbit flow interpolated using any kind of differentiable interpolation	Numerical dynamical system interpolated using modified vector field

Table 3: More (important) differences between defect analysis and MEA.

Finally, about conditioning. For IVP, ill-conditioning is somewhat equivalent to chaotic systems (in the sense of initial value perturbations). For vector field perturbations, this is harder to define, but may include Integrable systems, or systems close to bifurcation. Sometimes stiff equations are seen as a "forward stably" equivalent of ill-conditioned problems: there exist solutions with small forward error but large backward error [Cor94].

On top of the reasons given in the introduction of this chapter, there are more reasons to consider defect control/MEA for IVP:

- For Hamiltonian systems, perturbation theory is extremely well-developed, which leads to interesting results about symplectic algorithms.
- It is in general easier to relate the backward error to the global error than to relate the local error to the global error, [Moi10; KP94a], and a relation between the local error and the backward error is found. Sometimes BEA can also be used to discuss stability of numerical algorithms [WH74], but [GS86] argue that this approach must be exception rather than rule.
- This makes it interesting to construct algorithms which bound the backward error and useful for error analysis of general methods [Enr89; Moi10] i.e. to use BEA to explain the good behaviour of usual local error-minimising algorithms using BEA. In particular, sometimes the global error can be estimated without knowing the exact solution.
- BEA or backward error-minimising algorithms are sometimes useful to apply to chaotic systems [Cor92; Moi10].

In this thesis, only MEA is considered. For a quick overview of the other two, together with implementations and references, consult for example [Moi10].

D Lie-Gröbner series and modified equation analysis for non-autonomous ODE

In this Section we find expressions for the Lie-Gröbner series and modified vector fields of non-autonomous systems.

D.1 Non-autonomous Lie-Gröbner series

D.1.1 Equivalence of Lie-Gröbner series for non-autonomous ODE and canonically (autonomous) extended ODE

For non-autonomous vector fields f one finds, using the integral curve solution, that

$$y(t+h) = y(t) + hf(y(t),t) + \frac{h^2}{2}(f'(y(t),t)f(y(t),t) + f_t(y(t),t)) + \cdots .$$

In other words, Equation (6.5) becomes

$$\phi_{t_0+h,t_0}(y) = (e^{h(D_f+\partial/\partial_t)}Id)(y,t_0). \tag{D.1}$$

Equivalently one considers the canonical autonomous extension (Section 2.1) of $\bar{f} = (f,1)$, such that $D_f + \partial/\partial t = D_{\bar{f}}$ Indeed, one finds the Lie series

$$\begin{pmatrix} \phi_h(y,\tau(t)) \\ \tau(t+h) \end{pmatrix} = \begin{pmatrix} (e^{h(D_{\bar{f}})}Id)(y,\tau(t)) \\ \tau(t)+h \end{pmatrix}.$$

Thus, the Lie series of a non-autonomous system is just the projected Lie series of the canonical autonomous extension.

D.1.2 Different way to find the Lie Gröbner series of forced vector fields

This Section represents an alternative way to arrive at the Lie Series for a forced vector field i.e. Equation (6.6).

We consider a forced vector field $g(t)f(y)$, where $g \in C^\infty(\mathbb{R})$. Equation (D.1) becomes

$$\phi_{gf,t,t_0} = e^{tD_{gf}+\partial/\partial t}(Id) = \sum_{i\geq 0} \frac{t^i}{i!}(D_{gf} + D_t)^i(Id) = Id + \sum_{i\geq 1} \frac{t^i}{i!}(D_{gf} + D_t)^{i-1}(gf) \tag{D.2}$$

It is now shown that the coefficients $(D_{gf}+D_t)^i$ can be rewritten in terms of D_f i.e. of the form $\frac{t^i}{i!}\sum_{j=1}^i a_{i,j} D_f^j(Id)$ for coefficients $a_{i,j}$ depending on g and its derivatives: For the next two terms, if $\bar{f} = (gf, 1)$, this can be seen by writing

$$\begin{aligned} D_{\bar{f}}^2 &= (D_t + gD_f)^2 = g' D_f + g^2 D_f^2 \\ D_{\bar{f}}^3 &= (D_t + gD_f)^3 = g'' D_f + 3g' g D_{\bar{f}}^2 + g^2 D_f^3 \end{aligned}$$

where it was used that $D_{gf} = gD_f$. Since $(D_t + gD_f)^i = (D_t + gD_f)\left((D_t + gD_f)^{i-1}\right)$, one finds

$$a_{i,j} = \frac{d}{dt}(a_{i-1,j}) + g\, a_{i-1,j-1} \tag{D.3}$$

with initial values $a_{i,1} = \frac{d^i}{dt^i} g$ and $a_{1,j} = 0$ for $i \geq 1, j \geq 2$, which together with the recursive relations gives $a_{i,j} = 0$ for $j > i$ and $a_{i,i} = g^{i+1}$.

Defining $b_{i,j} := a_{i+j,j}$ (since $a_{i,j} = 0$ for $j > i$) we find

$$b_{i,j} = \frac{d}{dt} b_{i-1,j} + g\, b_{i,j-1}, \quad b_{i,0} = g_i \quad b_{0,j} = g^{j+1},$$

where $g_i := \frac{d^i}{dt^i} g$, of which the first terms are shown in Table 4. The recursive relation of Equation (D.3) has as a solution the commutative, partial, exponential Bell-polynomials $B_{i,j}$ as defined in Section B.1. Thus, setting $x_i := g_{i-1}$ one finds (c.f. Table 4) $a_{i,j}(g) = B_{i,j}(x_1, \ldots, x_i)$ such that (denoting $B_{i,j} := B_{i,j}(g, \dot{g}, \ldots)$)

g	g^2	g^3	g^4
g_1	$3gg_1$	$6g^2g_1$	$10g^3g_1$
g_2	$4gg_2 + 3g_1^2$	$10g^2g_2 + 15gg_1^2$	$20g^3g_2 + 45g^2g_1^2$
g_3	$5gg_3 + 10g_1g_2$	$15g^2g_3 + 15g_1^3 + 60gg_1g_2$	$35g_3g^3 + 105gg_1^3 + 210g^2g_1g_2$
g_4	$6gg_4 + 10g_2^2 + 15g_3g_1$	$21g_4g^2 + 70gg_2^2 + 105g_3gg_1 + 105g_1^2g_2$	$56g^3g_4 + 280g^2g_2^2 + 105g_1^4 + 420g^2g_3g_1 + 840gg_1^2g_2$

Table 4: Terms of the coefficients $b_{i,j}$, for $1 \leq i \leq 5, 1 \leq j \leq 4$.

$$\phi_{gf,t,t_0} = Id + \sum_{i\geq 1} \frac{t^i}{i!} \sum_{j=1}^{i} B_{i,j} D_f^j(Id) = Id + \sum_{i\geq 1} \frac{t^i}{i!} \sum_{j=1}^{i} \left(\sum_{l\in p_j(i)} \frac{i!}{l! \prod_{k=1}^{i} (k!)^{l_k}} \prod_{\ell=1}^{i} g_{i-1}^{l_i} \right) D_f^j(Id), \tag{D.4}$$

where notation as in Section B.1 was used ($p_j(i)$ thus notation partition of the integer i into j integers).

D.1.3 Lie series for non-autonomous ODE revisited

We return to the case of a non-autonomous ODE of f, such that the Lie-derivative becomes

$$\frac{d^i}{dt^i}\phi_{t,t_0}(y) = (D_t + D_f)^i(Id) = (D_t + D_f)^{i-1}(f)$$

for example

$$(D_t + D_f)^2(f) = D^2_{f^{(0)}}(f^{(0)}) + 2D_{f^{(0)}}(f^{(1)}) + D_{f^{(1)}}(f^{(0)}) + f^{(2)} \tag{D.5}$$

where $f^{(i)} = D_t^i f$. One can see that $(D_f + D_t)^i$ is a non-commutative version of the forced case $(gD_f + D_t)^i$. Therefore, the non-commutative versions $\hat{B}_i$ (see Section B.1) of the Bell polynomials are needed to define the Lie series, such that ([ELM14], chapter 3.1.2 or [LM11] chapter 4.2)

$$\phi_{t,t_0} = \sum_{i=0}^{\infty} \frac{h^i}{i!} \hat{B}_i(D_{f^{(0)}}, \ldots, D_{f^{(i-1)}})(Id) = \sum_{i=0}^{\infty} \frac{h^i}{i!} \sum_{j=1}^{i} \sum_{l\in P_j(i)} \frac{i!}{l!} \kappa(l) D_{f^{(l_1)}}, \ldots, D_{f^{(l_j)}}.(Id) \tag{D.6}$$

D.2 Formal modified equation analysis for non-autonomous systems

For non-autonomous ODE the MEA can be approached directly or via the (canonical) autonomous extension and we first show these approaches are equivalent c.f. Section D.1.1.

Afterwards, we change the MEA of the sections (6.25) and 6.3.3.

D.3 Equivalence of non-autonomous MEA and MEA for the canonical autonomous extension

We check the equivalence of non-autonomous modified equation analysis and modified equation analysis of the canonical autonomous extension.

For non-autonomous IVP $(f, (y_0, t_0)$, the non-autonomous equivalent of the modified ODE (Equation (6.17)) is

$$\begin{aligned} y_0 - \tilde{y}_0 = \mathcal{O}(h^{p+1}), \quad t_0 - \tilde{t}_0 = \mathcal{O}(h^{p+1}) \\ v - \psi_h(id, v) = \sum_{j=1}^{\infty} \frac{h^j}{j!} \left(D_t^j v - d_j(id, v) \right) = \mathcal{O}(h^{p+1}), \end{aligned} \tag{D.7}$$

which is the same as the regular modified equation. Therefore, the methods of Sections 6.3.2 - 6.3.4 are unchanged: Equation (6.25) is equal and the ansatz and recursive method just require instead non-autonomous modified vector fields $f_j : \mathbb{R} \times D \to \mathbb{R}^n$.

The second is by using a (canonical) autonomous extension. The corresponding (autonomous) modified vector fields $\tilde{g}^{[q]} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}$ are equal to the non-autonomous modified vector field $\tilde{f}^{[q]} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ in the following sense.

Proposition D.1. *Given a non-autonomous vector field f and a consistent numerical method $\psi_h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ then the modified vector fields $\tilde{g}$*

$$\tilde{g}^{[q]}(y, t) = (\tilde{f}^{[q]}, 1)$$

if the extended method $\tilde{\psi}_h(y, t) = (\psi_h(y, t), t + h)$ is used on the canonical extension.

Proof. For a consistent numerical method we find the coefficient functions of $\tilde{\psi}_h$ being $\tilde{d}_1 = (d_1, 1)$ and $\tilde{d}_j = (d_j, 0)$, $j > 1$. Therefore we find (if $\tilde{g}^{[q]} = \sum_{j=1}^{q} h^{j-1} g_j$) that $g_1 = (f_1, 1)$. Moreover since $D_{f_i}(g)(y, t) = D_y g(y, t) f(y, t) + D_t g(y, t)$

Therefore from the formulas in Equation (6.29), the fact that

□

D.3.1 Non-autonomous MEA using polynomial ansatz

Repeating the setup of Section 6.3.3 one now finds the non-autonomous equivalent of Equation (6.29) (using Lie derivatives, as in Proposition 6.9)

$$\begin{aligned} f_1 &= d_1 \\ 2f_2 + D_{f_1} f_1 + \dot{f}_1 &= d_2 \\ 6f_3 + 3\left(D_{f_1} f_2 + D_{f_2} f_1 + \dot{f}_2 \right) + D_{f_1} D_{f_1} f_1 + 2D_{f_1} \dot{f}_1 + D_{\dot{f}_1} f_1 + \ddot{f}_1 &= d_3. \\ \ldots &= d_4 \end{aligned} \tag{D.8}$$

where again the ansatz $\tilde{f}^{[M]} = \sum_{j=1}^{M} h^{j-1} f_j$ is made where $f_j : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ are non-autonomous vector fields (notation f_j is overloaded, in this case $f_j \neq D_t^j f$). We use the non-commutative, partial, exponential Bell polynomials $\hat{B}_{i,j}$ and the corresponding notation of Section B.2 to find an explicit formula for the non-autonomous modified vector fields.

Now the letters/variables $x_n = D_{\tilde{f}^{(n)}}$, where $n \in \mathbb{N}$, $\tilde{f}^{(n)} = \partial_t^n \tilde{f}$, can be seen a power series in $h \in \mathbb{R}$ i.e. $x_l = \sum_{\ell \geq 1} h^{\ell-1} x_{l,\ell}$ (since we want to apply the ansatz). Substituting this power series expansion into $\hat{B}_{i,j}$ one finds that $\hat{B}_{i,j}$ is a polynomial in the free, non-commutative algebra $\hat{\mathcal{E}} = \mathbb{R}\langle \{x_{l,\ell}\}_{l,\ell \in \mathbb{N}} \rangle$ which is a sum of polynomials $\hat{B}_{i,j,k}$ of order h^k (for $k \geq 0$) For example

$$\hat{B}_{3,2}(x_1, x_2, x_3) = 2x_1x_2 + x_2x_1 = \sum_{k \geq 0} h^k \left(\sum_{l_1 + l_2 = k+2} 2x_{1,l_1} x_{2,l_2} + x_{2,l_1} x_{1,l_2} \right) = \sum_{k \geq 0} h^k \hat{B}_{3,2,k}(x_1, x_2, x_3)$$

such that $\hat{B}_{3,2,k} = \sum_{l_1+l_2=k+2} 2x_{1,l_1} x_{2,l_2} + x_{2,k_1} x_{1,k_2}$. More generally, we can find for $k \geq 0$ (using again notation of Section B.2) that the substitution $\tilde{f} = \sum_{j \geq 1} h^{j-1} f_j$ gives

$$\hat{B}_{i,j}(D_{\tilde{f}}, \ldots, D_{\tilde{f}^{(i-j)}}) = \sum_{l \in \tilde{P}_j(i)} \frac{i!}{l!} \kappa(l) D_{\tilde{f}^{(l_1)}} \ldots D_{\tilde{f}^{(l_j)}} = \sum_{l \in \tilde{P}_j(i)} \sum_{\ell \in \mathbb{N}^j} \frac{i!}{l!} \kappa(l) h^{\ell_1 + \cdots + \ell_j - j} \prod_{n=1}^{j} D_{f_{\ell_1}^{(l_1)}} \ldots D_{f_{\ell_j}^{(l_j)}}$$

such that

$$\hat{B}_{i,j,k}(D_{\tilde{f}}) := \sum_{l \in \mathcal{P}_j(i)} \sum_{\ell \in \mathcal{P}_j(k+j)} \frac{i!}{l!} \kappa(w_l) \prod_{n=1}^{j} D_{f_{\ell_n}^{(l_n)}},$$

where $f_{\ell_n}^{(l_n)} = \partial_t^{l_n} f_{\ell_n}$.

Proposition D.2. *The non-autonomous modified vector fields f_j satisfy*

$$f_m - \frac{1}{m!} d_m = \sum_{i=2}^{m} \frac{1}{i!} \sum_{j=1}^{i} \hat{B}_{i,j,m-i}(D_{\tilde{f}})(Id) = -\sum_{i=2}^{m} \frac{1}{i!} \sum_{j=1}^{i} \sum_{l \in \mathcal{P}_j(i)} \sum_{\ell \in \mathcal{P}_j(m-i+j)} \frac{i!}{l!} \kappa(w_l) \left(\prod_{n=1}^{j} D_{f_{\ell_n}^{(l_n - 1)}} \right) (Id). \tag{D.9}$$

Proof. The proof is similar to the one of Proposition 6.9. Equation (D.6) gives

$$\phi_h = \sum_{i=0}^{\infty} \frac{h^i}{i!} \sum_{j=1}^{i} \sum_{l \in P_j(i)} \frac{i!}{l!} \kappa(l) D_{\tilde{f}^{(l_1)}}, \ldots, D_{\tilde{f}^{(l_j)}}(Id)$$

Substituting $\tilde{f} = \sum_{i=1}^{\infty} h^{i-1} f_i$ one finds for a numerical method $\psi_h = \sum_{i \geq 0} \frac{h^i}{i!} d_i$ that $\phi_h - \psi_h = 0$ implies

$$\sum_{i \geq 1} \frac{h^{i-1}}{i!} d_i = \sum_{i \geq 1} \frac{h^{i-1}}{i!} \sum_{j=1}^{i} \sum_{l \in \mathcal{P}_j(i)} \sum_{k \geq 0} h^k \sum_{\ell \in \mathcal{P}_j(k+j)} \frac{i!}{l!} \kappa(w_l) \prod_{n=1}^{j} D_{f_{\ell_n}^{(l_n - 1)}}.$$

using again notation of Section B.2. Comparing equal powers of h one finds that

$$\frac{1}{m!} d_m = \sum_{i=1}^{m} \frac{1}{i!} \sum_{j=1}^{i} \sum_{l \in \mathcal{P}_j(i)} \sum_{\ell \in \mathcal{P}_j(m-i+j)} \frac{i!}{l!} \kappa(w_l) \left(\prod_{n=1}^{j} D_{f_{\ell_n}^{(l_n - 1)}} \right) (Id) = \sum_{i=1}^{n} \frac{1}{i!} \sum_{j=1}^{i} \hat{B}_{i,j,m-i}(D_{\tilde{f}})(Id).$$

□

D.3.2 Non-autonomous MEA using substitution

We start at equation (6.25) and use that $D^i d_j = (D_{\tilde{f}} + D_t) D^{i-1} d_j$. This resembles extremely the case of Section D.1.3 and

$$D^2 d_j = D_{\tilde{f}} D_{\tilde{f}} d_j + D_{\dot{\tilde{f}}} d_j + 2 D_{\tilde{f}} \dot{d}_j + \ddot{d}_j$$

where $\dot{d}_j(y, t) = \frac{\partial}{\partial t} d_j(y, t)$. The only change is to replace in the non-commutative, partial Bell polynomials

$$\hat{B}_{i,k}(D_{f^{(1)}}, \ldots, D_{f^{(n)}}) = \sum_{l \in \mathcal{P}_k(i)} \frac{i!}{l!} \kappa(l) D_{f^{(l_1)}}, \ldots, D_{f^{(l_k)}}(Id)$$

the term $D_{\tilde{f}^{(l_k)}}(Id)$ by $d_j^{(l_k)}$, where $d_j^{(l_k)} = \frac{\partial^{l_k}}{\partial t^{l_k}} d_j(y,t)$, such that

$$\begin{aligned} v' + \mathcal{O}(h^M) &= \sum_{j=1}^{M} \frac{h^{j-1}}{j!} d_j + \sum_{i=1}^{q-1} \frac{\mathcal{B}_i}{i!} \sum_{j=i+1}^{M} \frac{h^{j-1}}{(j-i)!} \sum_{l \in \mathcal{P}(i)} \frac{i!}{l!} \kappa(l) \left(\prod_{n=1}^{k-1} D_{\tilde{f}^{(l_n)}} \right) (d_{j-i}^{(l_k)}) \\ &= \sum_{j=1}^{M} \frac{h^{j-1}}{j!} d_j + \sum_{i=1}^{q-1} \frac{\mathcal{B}_i}{i!} \sum_{j=i+1}^{M} \frac{h^{j-1}}{(j-i)!} \sum_{k=1}^{i} \sum_{l \in \mathcal{P}(i)} \sum_{\tilde{k}=k-1}^{q-j+k} \sum_{\ell \in \mathcal{P}_{k-1}(\tilde{k})} h^{\tilde{k}-k+1} \frac{i!}{l!} \kappa(l) \left(\prod_{n=1}^{k-1} D_{f_{\ell_n}^{(l_n)}} \right) (d_{j-i}^{(l_k)}) \end{aligned}$$

where $k = \#(l)$ is the degree of l. Substituting $\tilde{f} = \sum_{j \geq 1} h^{j-1} f_j$ one finds, collecting terms of h^{q-1}, that

$$f_q = \frac{1}{q!} d_j + \sum_{i=1}^{q-1} \frac{\mathcal{B}_i}{i!} \sum_{j=i+1}^{M} \frac{1}{(j-i)!} \sum_{l \in \mathcal{P}(i)} \sum_{\ell \in \mathcal{P}_{k-1}(q-j+k-1)} \frac{i!}{l!} \kappa(l) \left(\prod_{n=1}^{k-1} D_{f_{\ell_n}^{(l_n)}} \right) (d_{j-i}^{(l_k)}). \tag{D.10}$$

D.3.3 Modified equation analysis for forced ODE

We look finally at forced ODE of $g(t)f(y)$. For a consistent method, $d_1 = gf$ so that the first few modified vector field satisfy

$$\begin{aligned} f_1 &= gf \\ 2f_2 + g^2 f' f + \dot{g} f = 2f_2 + B_{2,2}(g) f' f + B_{1,2}(g) f &= d_2 \\ 6f_3 + 3\left(D_{f_1} f_2 + D_{f_2} f_1 + \dot{f}_2 \right) + D_{f_1} D_{f_1} f_1 + 2 D_{f_1} \dot{f}_1 + D_{\dot{f}_1} f_1 + \ddot{f}_1 &= d_3. \\ \ldots &= d_4 \end{aligned} \tag{D.11}$$

One can then use Equation (D.9). In the case that all d_j and all its time derivatives satisfy that the Lie derivative D_{d_j} commutes with D_{gf} then all the $D_{f_\ell^{(l)}}$ commute and the modified vector fields satisfy

$$\frac{1}{m!} d_m - f_m = \sum_{i=2}^{n} \frac{1}{i!} \sum_{j=1}^{i} B_{i,j,m-i}(D_f)(Id) = \sum_{i=2}^{m} \frac{1}{i!} \sum_{j=1}^{i} \sum_{l \in p_j(i)} \frac{i!}{l! \prod_{n=1}^{i} (n!)^{l_n}} \sum_{\ell \in p_j(m-i+j)} \left(\prod_{n=1}^{j} D_{f_{\ell_n}^{(l_n - 1)}} \right) (Id), \tag{D.12}$$

such that $B_{i,j,k}$ is the commutative version of $\hat{B}_{i,j,k}$. When using the induced method, equations simplify greatly (Equation (6.31)).

E *Mathematica* programs

E.1 Calculating the modified vector fields of the tidal wave system

Using *Mathematica Version 10.2.0.0* we use the following code to calculate the modified vector fields of autonomous, one degree of freedom vector fields.

Listing 1: Modified vector fields

```
LieDerivN[f0_, g0_] := Module[{fmod = f0, gmod = g0, A},
  D[fmod, {{q, p}}].gmod]

ModVFN[vecf_, psi_, n_, ord_] :=
(*Calculates the Modified vector field of order n, with vector field
"vecf" which is Order "ord" accurate, and to which "psi" is the numerical method applied*)
  Module[{x, modF, fTermPart, listP, lieTermL, lieList , m, time},
  modF = Table[0, n, 2];

  For[m = 1, m <= ord, m++,
   modF[[m]] = SeriesCoefficient[vecf[q, p], {h, 0, m - 1}];
   ];

```

```
 tay [q_, p_, h_] := Series[psi[q, p, h], {h, 0, n}];

 For[j = ord + 1, j <= n, j++,
     modF[[j]] = SeriesCoefficient [tay[q, p, h], {h, 0, j}] ;
     pp = IntegerPartitions [j];

   For[k = 2, k <= PartitionsP[j], k++,
     x = pp[[k]];   m = Length[x]!;
     listP = Permutations[x];

     For[ l = 1, l <= Length[listP], l++,
       lieList = listP [[ l ]];
      lieTermL = modF[[ lieList [[1]] ]];
      For[ r = 2, r <= Length[lieList ], r++,
       lieTermL = LieDerivN[lieTermL, modF[[ lieList [[ r ]] ]]];
        ];
      modF[[j]] = (modF[[j]] − 1/m lieTermL);
      ]
     ];
    modF[[j]] = Simplify[modF[[j ]]];
   ];
 modF[[ord + 1 ;; −1]]
 ]
```

The vector field for the method $\psi_{2,h}$ applied to the Hamiltonian L_2 as in Section 4 has the following modified vector fields

$$\begin{pmatrix} \sin(p) & \sin(q) \\ \frac{1}{2}\cos(p)\sin(q) & -\frac{1}{2}\sin(p)\cos(q) \\ -\frac{1}{12}\sin(p)\sin^2(q) - \frac{1}{6}\sin(p)\cos(p)\cos(q) & -\frac{1}{12}\sin^2(p)\sin(q) - \frac{1}{6}\cos(p)\sin(q)\cos(q) \\ -\frac{1}{24}\cos(2p)\sin(2q) & \frac{1}{24}\sin(2p)\cos(2q) \end{pmatrix} \tag{E.1}$$

Listing 2: Modified Hamiltonians

```
ModHam[H_, S_, nrH_, dim_] := Module[{x, sTab, tayS, i, j, time, ord},
 ord = 1/2 (−1 + Sqrt[1 + 8 Length[H]]);
 tayS = Series[S, {h, 0, nrH}];
 sTab = Table[0, ((nrH + 1) (nrH))/2];
 For[j = 1, j <= Length[H], j++,
  Print [" j␣=␣", j ];
  (*Print[" j = ",j, " invindex = ",invIndexM[ord,j ]]; *)
  sTab[[ indexM[nrH, #[[1]], #[[2]] &@invIndexM[ord, j] ]] = H[[j ]];
  ];
 (*Print [sTab];*)

 time = Timing[
   For[j = 1, j <= ord, j++,
    For[ i = ord + 2 − j, i <= nrH + 1 − j, i++,
     Print [" i␣=␣", i "␣j␣=␣", j ];
     sTab[[ indexM[nrH, i, j] ]] =
      Simplify [Scalc[ i , j, sTab, nrH, dim ]];
     (*sTab[[ i]]=Scalc[ i ,1, sTab,nrH,dim];*)
     (*Print[" i , j = ",i," ", j, " " ,sTab];*)
     ]
    ]
   ];
 Print ["time␣=␣", time [[1]]];
 (*Print [sTab];*)

 For[j = ord + 1, j <= nrH, j++,
  Print [" j␣=␣", j ];
  sTab[[indexM[nrH, 1, j ]]] =
   Simplify[ Hcalc[j , sTab, SeriesCoefficient [tayS, {h, 0, j }], nrH]];
  (*sTab[[indexM[nrH,1,j ]]] = Hcalc[j ,sTab, SeriesCoefficient [
  tayS,{ h,0, j }], nrH];*)

  time = Timing[For[i = 2, i <= nrH + 1 − j, i++,
      Print["i␣=␣", i];
      sTab[[indexM[nrH, i, j ]]] =
       Simplify[Scalc[i, j, sTab, nrH, dim]];
      (*sTab[[indexM[nrH,i,j ]]] = Scalc[i,j,sTab,nrH,dim];*)
      ]
     ];
   Print["time␣=␣", time [[1]]];
   ];
  (*sTab[[indexM[nrH,1,#]&/@Table[s,{s,1,nrH}]]]*)
  sTab
  ]

Hcalc[j_, sTab_, Sj_, nrH_] := Module[{outT, i},
  outT = Sj;
  (*Print[outT];*)
  For[i = 1, i < j, i++,
   outT = outT − sTab[[indexM[nrH, 1 + i, j − i ]]];
   (*Print[outT];*)
   ];
  (* FullSimplify [outT]*)
  outT
  ]

Scalc[n_, m_, sTab_, nrH_, dim_] :=
 Module[{termS, termSj, termSk, termSl, DsTab, r, i, ii , k, l , j,
   partK, partL, facL, facK, dHr, dHL},
  (*dHL=Table[0,m,2];*)
  (*Print[" Calculating S_{", n,"   ", m, "}"];*)
  termS = 0;
  DsTab = dStabCalc[n, m, nrH, dim, sTab];


  For[j = 1, j < n, j++,
   termSj = 0;
   partK = IntegerPartitions [n − 1, {j }];
   (*Print["j = ", j , " Partitions of K are ", partK];*)
   dHL = DqJHcalc[j, sTab, n, m, nrH, dim];

   For[r = 1, r <= m, r++,
    partL = IntegerPartitions [m + j − r, {j }];
    (*Print["r = ", r, " Partitions of L are ", partL]; *)
    dHr = dHL[[r]];

    For[i = 1, i <= Length[partK], i++,
     termSk = 0;
     k = partK[[i ]];


     For[ ii = 1, ii <= Length[partL], ii++,
      l = partL[[ ii ]];
      facL = 1/facCoeff[l, m − r + 1];
      (*Print[" Error here?"];
      Print["{k,l} = ", Transpose[{k,l }]]; *)
      (*termSk=termSk + (facL dHr Times[DsTab[[ indexM[
      nrH,#[[1]],#[[2]]]&/ @Transpose[{k,l }]]]]) [[1]]; *)
      termSk =
       termSk +
        facL dHr Times @@ (Extract[DsTab, Transpose[{k, l}] ]) ;
      (*Print[" Term = S",k,l," = ",facL dHr Times@@(Extract[DsTab,
      Transpose[{k,l}] ])]*)
      (*Print[" Error here?"];*)
      (*Print[Times[DsTab[[ indexM[2,#[[1]],#[[2]]]&/@Transpose[{k,
      l }]]]]]; *)
      ];
     facK = 1/facCoeff[k, n − j];
     termSj = termSj + facK termSk;
     ]
    ];
   termS = termS + (j!) termSj;
   ];
  1/n termS
  ]

DqJHcalc[ord_, sTab_, n_, m_, nrH_, dim_] := Module[{ouT, l, j },
  l = If[n == 1, m − 1, m];
  ouT = Table[0, l, dim/2];
  For[j = 1, j <= l, j++,
   ouT[[j ]] = D[sTab[[indexM[nrH, 1, j ]]], {q, ord}]
   ];
  ouT
  ]
dStabCalc[n_, m_, nrH_, dim_, sTab_] := Module[{dSTab, j, i},
  dSTab = Table[0, n, m , dim/2];
  For[j = 1, j < m, j++,
   For[i = 1, i <= n, i++,
     dSTab[[i , j ]] = D[sTab[[indexM[nrH, i, j ]]], p];
     ];
   ];
  For[i = 1, i < n, i++,
   dSTab[[i , m]] = D[sTab[[indexM[nrH, i, m ]]], p];
   ];
  dSTab
  (*dSTab = Table[0,((nrH+1)nrH)/2−indexM[nrH,n,m],dim/2];
  For[j=1,j\[LessEqual]Length[dSTab],j++,
  dSTab[[j ]] = D[sTab[[j ]], p];
  ];
  dSTab*)
  ]

indexM[nrH_, n_, m_] := n + (m − 1) (nrH + 1 − m/2);
invIndexM[nrH_, j_] := Module[{q, j0 = j},
  Assert [j <= ((nrH + 1) nrH)/2, ”index␣of␣Table␣out␣of␣bounds”];
  q = 0;
  While[j0 − nrH + q > 0, j0 = j0 − nrH + q; q++];
  {j0 , q + 1}]

facCoeff [ part_ , x_] := Module[{m, n},
  m = 1;
  For[n = 1, n <= x, n++,
   m = m*((Length[Select[part, # == n &]])!);
   ];
  m
  ]
```

References

[AC97] Mohammad O. Ahmed and Robert M. Corless. “The method of modified equations in Maple”. In: *Electronic Proc. 3rd Int. IMACS Conf. on Applications of Computer Algebra*. 1997.

[ACI83] Manuel Asorey, Jose F Carinena, and Luis A Ibort. “Generalized canonical transformations for time-dependent systems”. In: *Journal of mathematical physics* 24.12 (1983), pp. 2745–2750.

[AKN07] Vladimir I Arnold, Valery V Kozlov, and Anatoly I Neishtadt. *Mathematical aspects of classical and celestial mechanics*. Vol. 3. Springer Science & Business Media, 2007.

[AL12] Hassan Najafi Alishah and Rafael de la Llave. “Tracing KAM tori in presymplectic dynamical systems”. In: *Journal of Dynamics and Differential Equations* 24.4 (2012), pp. 685–711.

[AM08] Ralph Abraham and Jerrold E Marsden. *Foundations of mechanics*. 364. American Mathematical Soc., 2008.

[Arn63] Vladimir Igorevich Arnold. “Smal denominators and problems of stability of motion in classical and celestial mechanics”. In: (1963), pp. 85–191.

[Arn89] Vladimir Igorevich Arnold. *Mathematical methods of classical mechanics.* 2nd ed. Vol. 60. Springer Science & Business Media, 1989.

[Bar+08] María Barbero-Liñán et al. "Unified formalism for nonautonomous mechanical systems". In: *Journal of mathematical physics* 49.6 (2008), p. 062902.

[BB98] S Bouquet and A Bourdier. "Notion of integrability for time-dependent Hamiltonian systems: Illustrations from the relativistic motion of a charged particle". In: *Physical Review E* 57.2 (1998), p. 1273.

[BBM22] Karine Beauchard, Jérémy Le Borgne, and Frédéric Marbach. "On expansions for nonlinear systems, error estimates and convergence issues". In: (2022).

[BC06] S Blanes and F Casas. "Splitting methods for non-autonomous separable dynamical systems". In: *Journal of Physics A: Mathematical and General* 39.19 (2006), pp. 5405–5423. DOI: 10.1088/0305-4470/39/19/s05. URL: https://doi.org/10.1088/0305-4470/39/19/s05.

[BC10] David Blázquez-Sanz and Sergio A Carrillo Torres. "Group analysis of non-autonomous linear Hamiltonians through differential Galois theory". In: *Lobachevskii Journal of Mathematics* 31.2 (2010), pp. 157–173.

[BC17] Sergio Blanes and Fernando Casas. *A concise introduction to geometric numerical integration.* CRC press, 2017.

[BCM08] Sergio Blanes, Fernando Casas, and Ander Murua. "Splitting and composition methods in the numerical integration of differential equations". In: *arXiv preprint arXiv:0812.0377* (2008).

[BCM12] Sergio Blanes, Fernando Casas, and Ander Murua. "Splitting methods in the numerical integration of non-autonomous dynamical systems". In: *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas* 106.1 (2012), pp. 49–66.

[BCT17] Alessandro Bravetti, Hans Cruz, and Diego Tapias. "Contact hamiltonian mechanics". In: *Annals of Physics* 376 (2017), pp. 17–39.

[Ber78] Michael Victor Berry. "Regular and irregular motion". In: *AIP Conference proceedings.* Vol. 46. 1. American Institute of Physics. 1978, pp. 16–120.

[BG94] Giancarlo Benettin and Antonio Giorgilli. "On the Hamiltonian interpolation of near-to-the identity symplectic mappings with application to symplectic integration algorithms". In: *Journal of Statistical Physics* 74.5 (1994), pp. 1117–1143.

[Bla+10] Sergio Blanes et al. "Splitting and composition methods for explicit time dependence in separable dynamical systems". In: *Journal of computational and applied mathematics* 235.3 (2010), pp. 646–659.

[BM01] S. Blanes and P.C. Moan. "Splitting Methods for Non-autonomous Hamiltonian Equations". In: *Journal of Computational Physics* 170.1 (2001), pp. 205–230. DOI: https://doi.org/10.1006/jcph.2001.6733. URL: https://www.sciencedirect.com/science/article/pii/S0021999101967336.

[Bou55] Edmond Bour. "Sur l'intégration des équations différentielles de la Mécanique analytique." In: *Journal de mathématiques pures et appliquées* (1855), pp. 185–200.

[Bro04] Henk Broer. "KAM theory: the legacy of Kolmogorov's 1954 paper". In: *Bulletin of the American Mathematical Society* 41.4 (2004), pp. 507–521.

[BRZ94] SP Beerens, H Ridderinkhof, and JTF Zimmerman. "An analytical study of chaotic stirring in tidal areas". In: *Chaos, Solitons & Fractals* 4.6 (1994), pp. 1011–1029.

[BS02] Michael Brin and Garrett Stuck. *Introduction to dynamical systems.* Cambridge university press, 2002.

[BSG03] Liu Baifeng, Shi Shaoyun, and Wang Guoming. "KAM-type theorem for nearly integrable Hamiltonian with a quasiperiodic perturbation". In: *Northeast. Math. J* 19.3 (2003), pp. 273–282.

[Cal04] M Calvo. "Backward error analysis of numerical methods for ODEs and Lie–Hori perturbation theory". In: (2004).

[Car+87] José F Cariñena et al. “Applications of the canonical-transformation theory for presymplectic systems”. In: *Il Nuovo Cimento B (1971-1996)* 98.2 (1987), pp. 172–196.

[Car71] Jürgen K. Moser Carl L. Siegel. *Lectures on celestial mechanics*. Springer, 1971.

[Cas05] John Polking & Joel Castellanos. *PPlane*. Version 2005.10. 2005.

[CC87] Alessandra Celletti and Luigi Chierchia. “Rigorous estimates for a computer-assisted KAM theory”. In: *Journal of mathematical physics* 28.9 (1987), pp. 2078–2086.

[CC88] Alessandra Celletti and Luigi Chierchia. “Construction of analytic KAM surfaces and effective stability bounds”. In: *Communications in mathematical physics* 118.1 (1988), pp. 119–161.

[CF] Dana Constantinescu and Marie-Christine Firpo. “Integrability versus chaos in non-autonomous Hamiltonian systems. Applications to the study of some transport phenomena”. In: (). Seems unpublished. URL: https://inspirehep.net/files/8c8293cc8e28a0672b0864f5abb288a1.

[CF13] Robert M. Corless and Nicolas Fillion. “Numerical Solution of ODEs”. In: *A Graduate Introduction to Numerical Methods: From the Viewpoint of Backward Error Analysis*. New York, NY: Springer New York, 2013, pp. 509–583. ISBN: 978-1-4614-8453-0. DOI: 10.1007/978-1-4614-8453-0_12. URL: https://doi.org/10.1007/978-1-4614-8453-0_12.

[CFP87a] A Celletti, C Falcolini, and A Porzio. “Rigorous KAM stability statements for non-autonomous one-dimensional Hamiltonian systems”. In: (1987).

[CFP87b] Alessandra Celletti, Corrado Falcolini, and Anna Porzio. “Rigorous numerical stability estimates for the existence of KAM tori in a forced pendulum”. In: *Annales de l’IHP Physique théorique*. Vol. 47. 1. 1987, pp. 85–111.

[CG88] Alessandra Celletti and Antonio Giorgilli. “On the numerical optimization of KAM estimates by classical perturbation theory”. In: *Zeitschrift für angewandte Mathematik und Physik ZAMP* 39.5 (1988), pp. 743–747.

[Chi09] Luigi Chierchia. *Kolmogorov-Arnold-Moser (KAM) Theory.* 2009.

[CHV07] Philippe Chartier, Ernst Hairer, and Gilles Vilmart. “Numerical integrators based on modified differential equations”. In: *Mathematics of computation* 76.260 (2007), pp. 1941–1953.

[CIL87] José F Cariñena, Luis A Ibort, and Ernesto A Lacomba. “Time scaling as an infinitesimal canonical transformation”. In: *Celestial mechanics* 42.1 (1987), pp. 201–213.

[CL01] Begoña Cano and H Ralph Lewis. “A comparison of symplectic and Hamilton’s principle algorithms for autonomous and non-autonomous systems of ordinary differential equations”. In: *Applied numerical mathematics* 39.3-4 (2001), pp. 289–306.

[CL15] Marta Canadell and Rafael de la Llave. “KAM tori and whiskered invariant tori for non-autonomous systems”. In: *Physica D: Nonlinear Phenomena* 310 (2015), pp. 104–113.

[CMM22] José F. Carinena, Eduardo Martínez, and Miguel C. Muñoz-Lecanda. “Infinitesimal time reparametrisation and its applications”. In: *Journal of Nonlinear Mathematical Physics* (2022), pp. 1–33. DOI: https://doi.org/10.1007/s44198-022-00037-w.

[CMS94] M. P. Calvo, Ander Murua, and J. Sanz-Serna. “Modified equations for ODEs”. In: 172 (Jan. 1994). DOI: 10.1090/conm/172/01798.

[Con16] Dana Constantinescu. “REGULAR VERSUS CHAOTIC DYNAMICS IN SYSTEMS GENERATED BY AREA-PRESERVING MAPS. APPLICATIONS TO THE STUDY OF SOME TRANSPORT PHENOMENA”. In: *ROMANIAN JOURNAL OF PHYSICS* 61.1-2 (2016), pp. 52–66.

[Cor92] Robert M Corless. “Defect-controlled numerical methods and shadowing for chaotic differential equations”. In: *Physica D: Nonlinear Phenomena* 60.1-4 (1992), pp. 323–334.

[Cor94] Robert M Corless. “Error backward”. In: *Contemporary Mathematics* 172 (1994), pp. 31–31.

[De +01] Rafael De la Llave et al. “A tutorial on KAM theory”. In: *Proceedings of Symposia in Pure Mathematics*. Vol. 69. Providence, RI; American Mathematical Society; 1998. 2001, pp. 175–296.

[De 56] Rene De Vogelaere. "Methods of integration which preserve the contact transformation property of the Hamilton equations". In: *Technical report (University of Notre Dame. Dept. of Mathematics)* (1956).

[DE02] Johannes Jisse Duistermaat and Wiktor Eckhaus. *Analyse van gewone differentiaalvergelijkingen.* Epsilon Uitgaven, 2002.

[DK04] Johannes Jisse Duistermaat and Johan AC Kolk. *Multidimensional real analysis I: differentiation.* Vol. 86. Cambridge University Press, 2004.

[DM21] Ritesh Kumar Dubey and Prabhat Mishra. "Modified equation based mesh adaptation algorithm for evolutionary scalar partial differential equations". In: *Numerical Methods for Partial Differential Equations* (2021).

[Dou82] Raphaël Douady. "Une démonstration directe de l'équivalence des théorèmes de tores invariants pour difféomorphismes et champs de vecteurs". In: (1982).

[Dum14] H Scott Dumas. *Kam Story, The: A Friendly Introduction To The Content, History, And Significance Of Classical Kolmogorov-arnold-moser Theory.* World Scientific Publishing Company, 2014.

[EG18] Oğul Esen and Partha Guha. "On time-dependent Hamiltonian realizations of planar and nonplanar systems". In: *Journal of Geometry and Physics* 127 (2018), pp. 32–45. DOI: https://doi.org/10.1016/j.geomphys.2018.01.024. URL: https://www.sciencedirect.com/science/article/pii/S0393044018300408.

[ELM14] Kurusch Ebrahimi-Fard, Alexander Lundervold, and Dominique Manchon. "Noncommutative Bell polynomials, quasideterminants and incidence Hopf algebras". In: *International Journal of Algebra and Computation* 24.05 (2014), pp. 671–705.

[ENB00] Klaus-Jochen Engel, Rainer Nagel, and Simon Brendle. *One-parameter semigroups for linear evolution equations.* Vol. 194. Springer, 2000.

[Enr89] Wayne H Enright. "A new error-control for initial value solvers". In: *Applied Mathematics and Computation* 31 (1989), pp. 288–301.

[Fel22] Achim Feldmeier. *Introduction to Arnold's Proof of the Kolmogorov–Arnold–Moser Theorem.* CRC Press, 2022.

[FGS02] Emanuele Fiorani, Giovanni Giachetta, and Gennadi Sardanashvily. "The Liouville-Arnold-Nekhoroshev theorem for non-compact invariant manifolds". In: *arXiv preprint math/0210346* (2002).

[Fio04] Emanuele Fiorani. "Completely and partially integrable hamiltonian systems in the noncompact case". In: *International Journal of Geometric Methods in Modern Physics* 1.03 (2004), pp. 167–183.

[FQ10] Kang Feng and Mengzhao Qin. *Symplectic geometric algorithms for Hamiltonian systems.* Vol. 449. Springer, 2010.

[Fra+21] Guilherme França et al. "Optimization on manifolds: A symplectic approach". In: *arXiv preprint arXiv:2107.11231* (2021).

[FS07] E Fiorani and G Sardanashvily. "Global action-angle coordinates for completely integrable systems with noncompact invariant submanifolds". In: *Journal of mathematical physics* 48.3 (2007), p. 032901.

[FS96] Bernold Fiedler and Jürgen Scheurle. *Discretization of Homoclinic Orbits, Rapid Forcing and "Invisible" Chaos.* Vol. 570. American Mathematical Soc., 1996.

[FW16] Alessandro Fortunati and Stephen Wiggins. "The Kolmogorov-Arnold-Moser (KAM) and Nekhoroshev Theorems with Arbitrary Time Dependence". In: *Essays in Mathematics and its Applications.* Springer, 2016, pp. 89–99.

[Gio12] Antonio Giorgilli. "On the representation of maps by Lie transforms". In: *arXiv preprint arXiv:1211.5674* (2012).

[GK67] Wolfgang Gröbner and Herbert Knapp. *Contributions to the method of Lie series.* Vol. 802. Bibliographisches Institut Mannheim, 1967.

[GMS02] Giovanni Giachetta, Luigi Mangiarotti, and Gennadi Sardanashvily. "Action–angle coordinates for time-dependent completely integrable Hamiltonian systems". In: *Journal of Physics A: Mathematical and General* 35.29 (2002), p. L439.

[GMS97] G Giachetta, L Mangiarotti, and G Sardanashvily. "Differential geometry of time-dependent mechanics". In: *arXiv preprint dg-ga/9702020* (1997).

[Gol95] Christophe Golé. "Suspension of symplectic twist maps by Hamiltonians". In: *Hamiltonian Dynamical Systems*. Springer, 1995, pp. 163–169.

[GPS02] Herbert Goldstein, Charles Poole, and John Safko. *Classical mechanics*. American Association of Physics Teachers, 2002.

[Grö67] Wolfgang Gröbner. *Die Lie-reihen und ihre Anwendungen*. Vol. 3. Deutscher Verlag der Wissenschaften, 1967.

[GS18] Mauricio Garay and Duco van Straten. "KAM Theory. Part I. Group actions and the KAM problem". In: *arXiv preprint arXiv:1805.11859* (2018).

[GS86] DF Griffiths and JM Sanz-Serna. "On the scope of the method of modified equations". In: *SIAM Journal on Scientific and Statistical Computing* 7.3 (1986), pp. 994–1008.

[GV18] Vassili Gelfreich and Arturo Vieiro. "Interpolating vector fields for near identity maps and averaging". In: *Nonlinearity* 31.9 (2018), p. 4263.

[Hal80] J.K. Hale. *Ordinary Differential Equations*. Dover Books on Mathematics Series. Krieger Pub Co, 1980. ISBN: 9780898740110.

[Han11] Heinz Hanßmann. "Non-degeneracy conditions in kam theory". In: *Indagationes Mathematicae* 22.3 (2011). Devoted to: Floris Takens (1940–2010), pp. 241–256. ISSN: 0019-3577. DOI: `https://doi.org/10.1016/j.indag.2011.09.005`. URL: `https://www.sciencedirect.com/science/article/pii/S0019357711000474`.

[Hen96] Jacques Henrard. "Symplectic integrators". In: *Analysis and Modelling of Discrete Dynamical Systems*. Gordon and Breach, 1996.

[HI03] John Hubbard and Yulij Ilyashenko. "A proof of Kolmogorov's theorem". In: *Discrete and Continuous Dynamical systems* 4 (2003), pp. 1–20.

[HL00] Ernst Hairer and Christian Lubich. "Asymptotic expansions and backward analysis for numerical integrators". In: *Dynamics of algorithms*. Springer, 2000, pp. 91–106.

[HL97] Ernst Hairer and Christian Lubich. "The life-span of backward error analysis for numerical integrators". In: *Numerische Mathematik* 76.4 (1997), pp. 441–462.

[HLW06] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical integration: structure-preserving algorithms for ordinary differential equations*. Springer, 2006.

[Hub07] John H. Hubbard. "The KAM Theorem". In: *Kolmogorov's Heritage in Mathematics*. Ed. by Éric Charpentier, Annick Lesne, and Nikolaï K. Nikolski. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 215–238. ISBN: 978-3-540-36351-4. DOI: `10.1007/978-3-540-36351-4_11`. URL: `https://doi.org/10.1007/978-3-540-36351-4_11`.

[IQ18] Arieh Iserles and G. R. W. Quispel. "Why Geometric Numerical Integration?" In: *Discrete Mechanics, Geometric Integration and Lie–Butcher Series*. Ed. by Kurusch Ebrahimi-Fard and María Barbero Liñán. Cham: Springer International Publishing, 2018, pp. 1–28. ISBN: 978-3-030-01397-4.

[Jay21] Laurent O Jay. "Symplecticness conditions of some low order partitioned methods for non-autonomous Hamiltonian systems". In: *Numerical Algorithms* 86.2 (2021), pp. 495–514.

[Joh89] Oliver Davis Johns. "Canonical transformations with time as a coordinate". In: *American Journal of Physics* 57.3 (1989), pp. 204–215.

[Jor91] Àngel Jorba i Monte. "On quasiperiodic perturbations of ordinary differential equations". PhD thesis. Universitat de Barcelona, 1991.

[JS96] Àngel Jorba and Carles Simó. "On quasi-periodic perturbations of elliptic equilibrium points". In: *SIAM Journal on Mathematical Analysis* 27.6 (1996), pp. 1704–1737.

[Jur96] Velimir Jurdjevic. "Control affine systems". In: *Geometric Control Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1996, pp. 95–124. DOI: `10.1017/CBO9780511530036.006`.

[JV97] Angel Jorba and Jordi Villanueva. "On the persistence of lower dimensional invariant tori under quasi-periodic perturbations". In: *Journal of Nonlinear Science* 7.5 (1997), pp. 427–473.

[KN13] Victor V Kozlov and AI Neishtadt. *Dynamical Systems III*. Vol. 3. Springer Science & Business Media, 2013.

[KP94a] Peter E Kloeden and KT Palmer. *Chaotic Numerics: An International Workshop on the Approximation and Computation of Complicated Dynamical Behavior, Deakin University, Geelong, Australia, July 12-16, 1993*. Vol. 172. American Mathematical Soc., 1994.

[KP94b] Sergei Kuksin and Jürgen Pöschel. "On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications". In: *Seminar on Dynamical systems*. Springer. 1994, pp. 96–116.

[KR11] Peter E Kloeden and Martin Rasmussen. *Nonautonomous dynamical systems*. 176. American Mathematical Soc., 2011.

[Leó+22] Manuel de León et al. "Time-dependent contact mechanics". In: *arXiv preprint arXiv:2205.09454* (2022).

[Lio55] Joseph Liouville. "Note à l'occasion du Mémoire précédent de M. Edmond Bour." In: *Journal de Mathématiques Pures et Appliquées* (1855), pp. 201–202.

[Lla+05] Rafael de la Llave et al. "KAM theory without action-angle variables". In: *Nonlinearity* 18.2 (2005), p. 855.

[LM11] Alexander Lundervold and Hans Munthe-Kaas. "Hopf algebras of formal diffeomorphisms and numerical integration on manifolds". In: *Contemp. Math* 539 (2011), pp. 295–324.

[LR04] Benedict Leimkuhler and Sebastian Reich. *Simulating hamiltonian dynamics*. 14. Cambridge university press, 2004.

[LS17] Manuel de León and C Sardón. "Cosymplectic and contact structures for time-dependent and dissipative Hamiltonian systems". In: *Journal of Physics A: Mathematical and Theoretical* 50.25 (2017), p. 255205.

[LY68] GSS Ludford and DW Yannitell. "Canonical transformations without Hamilton's principle". In: *American Journal of Physics* 36.3 (1968), pp. 231–233.

[McL95] Robert I McLachlan. "Composition methods in the presence of small parameters". In: *BIT numerical mathematics* 35.2 (1995), pp. 258–268.

[Mei15] JD Meiss. "Thirty years of turnstiles and transport". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.9 (2015), p. 097602.

[Mei92] JD Meiss. "Symplectic maps, variational principles, and transport". In: *Reviews of Modern Physics* 64.3 (1992), p. 795.

[MMP84] RS MacKay, JD Meiss, and IC Percival. "Transport in Hamiltonian systems". In: *Physica D: Nonlinear Phenomena* 13.1 (1984), pp. 55–81. ISSN: 0167-2789. DOI: `https://doi.org/10.1016/0167-2789(84)90270-7`. URL: `https://www.sciencedirect.com/science/article/pii/0167278984902707`.

[MO10] Robert I McLachlan and Dion RJ O'Neale. "Preservation and destruction of periodic orbits by symplectic integrators". In: *Numerical Algorithms* 53.2 (2010), pp. 343–362.

[MO14] Håkon Marthinsen and Brynjulf Owren. "Geometric integration of non-autonomous Hamiltonian problems". In: *arXiv preprint arXiv:1409.5058* (2014).

[MO17] Kenneth R. Meyer and Daniel C. Offin. *Introduction to Hamiltonian Dynamical*. Springer International Publishing, 2017. ISBN: 978-3-319-53691-0.

[Moa03] Per Christian Moan. “On the KAM and Nekhoroshev theorems for symplectic integrators and implications for error growth”. In: *Nonlinearity* 17.1 (2003), p. 67.

[Moa05] PC Moan. *On rigorous modified equations for discretizations of ODEs*. Tech. rep. Technical Report 2005-3, Geometric Integration Preprint Server, 2005 ..., 2005.

[Moa06] Per Christian Moan. “On modified equations for discretizations of ODEs”. In: *Journal of Physics A: Mathematical and General* 39.19 (2006), p. 5545.

[Moi10] Robert HC Moir. “Reconsidering backward error analysis for ordinary differential equations”. PhD thesis. School of Graduate and Postdoctoral Studies, University of Western Ontario, 2010.

[Mös62] J Möser. “On invariant curves of area-preserving mappings of an annulus”. In: *Nachr. Akad. Wiss. Göttingen, II* (1962), pp. 1–20.

[Mos87] Jürgen Moser. “Is the solar system stable?” In: *Hamiltonian dynamical systems: a reprint selection* 1 (1987), p. 20.

[Mos99] J Moser. “Old and new applications of KAM theory”. In: *Hamiltonian Systems with Three or More Degrees of Freedom*. Springer, 1999, pp. 184–192.

[MQ01] Robert I McLachlan and GRW Quispel. “What kinds of dynamics are there? Lie pseudogroups, dynamical systems and geometric integration”. In: *Nonlinearity* 14.6 (2001), p. 1689.

[MQ02] Robert I McLachlan and G Reinout W Quispel. “Splitting methods”. In: *Acta Numerica* 11 (2002), pp. 341–434.

[MW00] Alfredo Martinez and Stephen Wiggins. “Time Aperiodic Perturbations of Integrable Hamiltonian Systems”. In: *arXiv preprint nlin/0007010* (2000).

[OO89] Julio M Ottino and JM Ottino. *The kinematics of mixing: stretching, chaos, and transport*. Vol. 3. Cambridge university press, 1989.

[Pös09] Jürgen Pöschel. “A lecture on the classical KAM theorem”. In: *arXiv preprint arXiv:0908.2234* (2009).

[Qin96] Mengzhao Qin. “Symplectic schemes for nonautonomous Hamiltonian system”. In: *Acta Mathematicae Applicatae Sinica* 12.3 (1996), pp. 284–288.

[Rei99] Sebastian Reich. “Backward error analysis for numerical integrators”. In: *SIAM Journal on Numerical Analysis* 36.5 (1999), pp. 1549–1570.

[RF11] AS Richardson and JM Finn. “Symplectic integrators with adaptive time steps”. In: *Plasma Physics and Controlled Fusion* 54.1 (2011), p. 014004.

[Riv21] Xavier Rivas Guijarro. “Geometrical aspects of contact mechanical systems and field theories”. PhD thesis. UPC, Facultat de Matemàtiques i Estadística, 2021. URL: http://hdl.handle.net/2117/361635.

[RZ92] H Ridderinkhof and JTF Zimmerman. “Chaotic stirring in a tidal system”. In: *Science* 258.5085 (1992), pp. 1107–1111.

[San92] Jesus M Sanz-Serna. “Symplectic integrators for Hamiltonian problems: an overview”. In: *Acta numerica* 1 (1992), pp. 243–286.

[Sar13a] G Sardanashvily. “Fibre bundle formulation of time-dependent mechanics”. In: *arXiv preprint arXiv:1303.1735* (2013).

[Sar13b] G Sardanashvily. “Lectures on integrable Hamiltonian systems”. In: *arXiv preprint arXiv:1303.5363* (2013).

[SC20] Robert D Skeel and Jan L Cieslinski. “On the famous unpublished preprint” Methods of integration which preserve the contact transformation property of the Hamilton equations” by René De Vogelaere”. In: *arXiv preprint arXiv:2003.12268* (2020).

[SC94] J. M. Sanz-Serna and M. P. Calvo. *Numerical Hamiltonian problems*. Chapman & Hall, 1994. ISBN: 0412542900.

[Sev03] Mikhail B Sevryuk. “The classical KAM theory at the dawn of the twenty-first century”. In: *Moscow Math. J* 3.3 (2003), pp. 1113–1144.

[Sev06] Mikhail B Sevryuk. *Reversible systems.* Vol. 1211. Springer, 2006.

[Sev07] Mikhail B Sevryuk. "Invariant tori in quasi-periodic non-autonomous dynamical systems via Herman's method". In: *Discrete & Continuous Dynamical Systems* 18.2&3 (2007), p. 569.

[Sev16] Mikhail Borisovich Sevryuk. "On the history of KAM theory". In: *Russian Journal of Nonlinear Dynamics* 12.2 (2016), pp. 289–293.

[Sev94] Mikhail B Sevryuk. "New results in the reversible KAM theory". In: *Seminar on Dynamical Systems.* Springer. 1994, pp. 184–199.

[Sey16] Muaz Seydaoglu. "Splitting methods for autonomous and non-autonomous perturbed equations". PhD thesis. Universitat Politècnica de València, 2016.

[SH98] Andrew Stuart and Anthony R Humphries. *Dynamical systems and numerical analysis.* Vol. 2. Cambridge University Press, 1998.

[Sha99] Zai-jiu Shang. "KAM theorem of symplectic algorithms for Hamiltonian systems". In: *Numerische Mathematik* 83.3 (1999), pp. 477–496.

[ST16] Sergey M Saulin and Dmitry V Treschev. "On the inclusion of a map into a flow". In: *Regular and Chaotic Dynamics* 21.5 (2016), pp. 538–547.

[Ste84] Stanly Steinberg. "Lie series and nonlinear ordinary differential equations". In: *Journal of mathematical analysis and applications* 101.1 (1984), pp. 39–63.

[Ste86] Stanly Steinberg. "Lie series, Lie transformations, and their applications". In: *Lie methods in optics: proceedings of the CIFMO-CIO workshop, held at León, México, January 7-10, 1985.* Springer, 1986, pp. 45–103.

[Str05] Jürgen Struckmeier. "Hamiltonian dynamics on the symplectic extended phase space for autonomous and non-autonomous systems". In: *Journal of Physics A: Mathematical and General* 38.6 (2005), p. 1257.

[SW92] Gerald Jay Sussman and Jack Wisdom. "Chaotic evolution of the solar system". In: *Science* 257.5066 (1992), pp. 56–62.

[Tre19] Dmitry Treschev. "Volume preserving diffeomorphisms as Poincaré maps for volume preserving flows". In: *arXiv preprint arXiv:1911.09420* (2019).

[Tre99] Dmitry V Treschev. "Continuous Averaging in Hamiltonian Systems". In: *Hamiltonian Systems with Three or More Degrees of Freedom.* Springer, 1999, pp. 244–253.

[Tsi00] AV Tsiganov. "Canonical transformations of the extended phase space, Toda lattices and the Stäckel family of integrable systems". In: *Journal of Physics A: Mathematical and General* 33.22 (2000), p. 4169.

[TZ09] Dmitry Treschev and Oleg Zubelevich. *Introduction to the perturbation theory of Hamiltonian systems.* Springer Science & Business Media, 2009.

[Val19] Lorenzo Valvo. "Hamiltonian Perturbation Theory on a Poisson Algebra. Application to a Throbbing Top and to Magnetically Confined Particles". PhDTheses. Aix Marseille University, 2019. URL: `https://tel.archives-ouvertes.fr/tel-02442078`.

[Ver06] Ferdinand Verhulst. *Nonlinear differential equations and dynamical systems.* Springer Science & Business Media, 2006.

[VV21] Lorenzo Valvo and Michel Vittot. "Hamiltonian Perturbation Theory on a Lie Algebra. Application to a non-autonomous Symmetric Top". In: *arXiv preprint arXiv:2101.01432* (2021).

[Wal21] ST van der Wal. "Particle Trajectories in Topography-Induced Currents: Looking for Chaos". B.S. thesis. 2021.

[Wan94] Daoliu Wang. "Some aspects of Hamiltonian systems and symplectic algorithms". In: *Physica D: Nonlinear Phenomena* 73.1-2 (1994), pp. 1–16.

[Way96] C Eugene Wayne. "An introduction to KAM theory". In: *Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994)* 31 (1996), pp. 3–29.

[WH74] Robert F Warming and BJ Hyett. "The modified equation approach to the stability and accuracy analysis of finite-difference methods". In: *Journal of computational physics* 14.2 (1974), pp. 159–179.

[Wig03] Stephen Wiggins. *Introduction To Applied Nonlinear Dynamical Systems And Chaos*. Vol. 4. Jan. 2003. ISBN: 0-387-00177-8. DOI: `10.1007/b97481`.

[Wik21] Wikipedia contributors. *Bell polynomials — Wikipedia, The Free Encyclopedia*. `https://en.wikipedia.org/w/index.php?title=Bell_polynomials&oldid=1062670624`. [Online; accessed 24-March-2022]. 2021.

[Wik22] Wikipedia contributors. *Faà di Bruno's formula — Wikipedia, The Free Encyclopedia*. `https://en.wikipedia.org/w/index.php?title=Fa%C3%A0_di_Bruno%27s_formula&oldid=1071944003`. [Online; accessed 2-May-2022]. 2022.

[WM14] S Wiggins and AM Mancho. "Barriers to transport in aperiodically time-dependent two-dimensional velocity fields: Nekhoroshev's theorem and" Nearly Invariant" tori". In: *Nonlinear Processes in Geophysics* 21.1 (2014), pp. 165–185.

[ZC10] Dongfeng Zhang and Rong Cheng. "On invariant tori of nearly integrable Hamiltonian systems with quasiperiodic perturbation". In: *Fixed Point Theory and Applications* 2010 (2010), pp. 1–17.