Continuum Limits of Variational Systems

vorgelegt von

M.Sc.
Mats Vermeeren

geb. in Leuven

von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
Dr.rer.nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Fredi Tröltzsch
Gutachter: Prof. Dr. Yuri B. Suris
Gutachter: Prof. Dr. Christian Lubich
Gutachter: Prof. Dr. Frank W. Nijhoff

Tag der wissenschaftlichen Aussprache: 19.10.2018

Berlin, 2018

Contents

Abstract

In this thesis we examine how to recover continuous systems from discrete systems, i.e. differential equations from difference equations. In particular, we are interested in equations with a variational (Lagrangian) structure and the transferal of this structure from the discrete to the continuous.

In the context of numerical integration, the differential equation corresponding to a given difference equation is known as its modified equation. Studying the modified equation to learn about a numerical integrator is a form of backward error analysis. It is well known that for a symplectic integrator applied to a Hamiltonian system, the modified equation is again a Hamiltonian equation. We will prove the corresponding result on the Lagrangian side: the modified equation for a variational integrator applied to a Lagrangian system is Lagrangian.

In the context of integrable systems, discrete models are often better understood than their continuous counterparts, so continuum limits are a useful tool to construct and study integrable hierarchies of differential equations. Over the last decade, a variational perspective on integrable systems has been developed, known as pluri-Lagrangian or Lagrangian multiform theory. It has analogous continuous and discrete versions. We will discuss how to take the continuum limit of a pluri-Lagrangian lattice equation to obtain a hierarchy of differential equations, together with its pluri-Lagrangian structure. We will apply this to most of the lattice equations of the ABS list and to some members of the lattice GD hierarchy. This way, we obtain many previously unknown examples of continuous pluri-Lagrangian systems, including a multi-component system.

Zusammenfassung

Diese Arbeit behandelt die Frage, wie man kontinuierliche Systeme aus diskreten Systemen herleiten kann, d.h. Differentialgleichungen aus Differenzengleichungen. Insbesondere sind wir an Gleichungen mit einer variationellen (Lagrangeschen) Struktur interessiert, und an der Frage wie diese Struktur vom Diskreten ins Kontinuierliche übertragen werden kann.

Im Bereich der numerischen Integration nennt man die Differentialgleichung, die mit einer gegebenen Differenzengleichung übereinstimmt, die modifizierte Gleichung. Einen numerischen Integrator mittels seiner modifizierten Gleichung zu untersuchen, ist eine Form der Rückwärtsanalyse. Es ist bekannt, dass die modifizierte Gleichung für einen symplektischen Integrator, angewandt auf eine Hamiltonsche Gleichung, auch eine Hamiltonsche Gleichung ist. In dieser Arbeit leiten wir die entsprechende Aussage auf der Lagrangeschen Seite her: die modifizierte Gleichung für einen variationellen Integrator, angewandt auf eine Lagrangesche Gleichung, ist wieder eine Lagrangesche Gleichung.

Im Bereich der integrablen Systemen sind diskrete Gleichungen oft leichter zu verstehen als ihre kontinuierliche Ebenbilder. Deswegen sind stetige Limes ein hilfreiches Werkzeug um Hierarchien integrabler Gleichungen herzuleiten und ihre Eigenschaften zu erforschen. Im letzten Jahrzehnt ist eine variationelle Theorie für integrablen Systeme entwickelt worden: die Theorie der Pluri-Lagrangeschen Systeme, oder auch der Lagrangeschen Multiformen. Diese Theorie hat analoge Versionen auf der diskreten und der kontinuierlichen Ebene. In dieser Arbeit zeigen wir, wie man den stetigen Limes eines Pluri-Lagrangeschen Systems bildet. Das Ergebnis ist eine Hierarchie integrabler Differentialgleichungen, inklusive der Pluri-Lagrangeschen Struktur. Dieses Verfahren wenden wir auf die meisten diskreten Gleichungen der ABS-Liste und auf einigen Gleichungen der diskreten GD-Hierarchie an. Auf diese Weise finden wir verschiedene neue Beispiele kontinuierlicher Pluri-Lagrangescher Systeme, insbesondere auch ein Mehrkomponentensystem.

Acknowledgments

Science fiction writer Isaac Asimov is reported to have said that

> 'The most exciting phrase to hear in science, the one that heralds new discoveries, is not "Eureka!" (I found it!) but "That's funny…"'

The first part of this quote is certainly true in mathematics as well. I've never heard a colleague exclaim "Eureka!". But in my experience the most exciting phrase is not "That's funny…" either. It is "I'm an idiot!". Mathematical progress almost always comes at the moment one finds a retrospectively obvious mistake in the previous attempt.

These last few years I have felt like an idiot often, but only a few times did I get to proclaim that with joy. I would like to thank everyone who helped me accept the feeling of idiocy, everyone who was an idiot with me, and everyone who nudged me in the direction towards those poorly named Eureka-moments when I realized why exactly I was an idiot.

> First and foremost I would like to thank my Yuri Suris, my supervisor. I am sure his influence can be found on every page of this thesis: in the topics covered, the mathematical taste, or even the style of writing. As much as for his mathematical teachings, I am grateful for his unwavering confidence in me and my work, especially in those moments when my own confidence was shaken.
>
> Additionally I am thankful to all the people form the Geometry and Mathematical Physics group on the 8th floor of the math building at TU Berlin, with whom I sat through countless research seminars, and with whom I got cooked each summer as the heat rises through the floors of the building. Thank you for the camaraderie and for the inspiring discussions.

Of course I am not only grateful for support of people, but also of institutions. In particular:

> The Berlin Mathematical School, which gave me the opportunity to come to Berlin for my master's degree and has remained a central part of my mathematical experience afterwards.
>
> The DFG collaborative research center *SFB 109 Discretization in Geometry and Dynamics*, for providing me with funding and scientific opportunities.

Many more people away from the university have supported me in various ways. In particular I would like to thank:

> My parents, who always encouraged my siblings and me to find our own paths in life. Surely they did not expect that those paths would lead all of their children abroad, but they never once asked us to come back.
>
> The rest my family and friends in Belgium. I'm sorry for all the missed gatherings.
>
> My classmates from my undergraduate studies in Leuven, many of whom have remained my friends (even after I cited their lack of scientific ambition as a reason to come to Berlin). Your goodbye gift still hangs on my wall.
>
> My friends in Berlin, who occasionally have had to compete for my attention with the mathematical thoughts running through my head after hours. You have made these last few years good ones on a social level too.

Preface

Why variational systems?

Physicists have long ago realized the danger of assuming that nature works in the way we like to think about the world. The earth is not the center of the universe, even though we see everything else rotate around us. A feather will fall just as fast as a hammer, provided we eliminate wind resistance. There is no clear distinction between particles and waves, even though they seem completely different. Applying your experience in daily life to the fundamentals of physics is a fallacy.

It is tempting to think that mathematicians are immune to this fallacy. After all, mathematical truth should not depend on any physical reality, should it? But even if we are fully convinced that it shouldn't, the way we view the world certainly influences which mathematical tools we consider most natural. Take the notion of time for example. In our real-world experience, time flows steadily forward. This makes it seem most natural to describe dynamical processes by differential equations. Understanding nature in this mathematical language was Isaac Newton's great achievement.

Time progressed, and Newton's era made way to that of minds like Euler, Lagrange, and later Hamilton. They realized that there is more structure to the world than can be understood through the differential equations themselves. Out of their insights (and those of many others) two mostly equivalent concepts arose, which we now refer to as Lagrangian and Hamiltonian mechanics.

The Hamiltonian formalism constructs differential equations with powerful geometric properties. Even though this geometry is not obvious in our experience of the world, the fact that differential equations lie at the heart of the Hamiltonian picture fits very well with our intuition of forward flowing time. In contrast, the Lagrangian formalism looks at the system at all moments in time at once by requiring that the action, which is an integral over time, takes a critical value. Of course differential equations can be derived from this principle as well, but the fundamental idea is at odds with our human experience of time. It seems to involve looking both at the past and at the future to determine how the system evolves at any given time. Perhaps this is why the Hamiltonian theory is more developed in many areas of mathematical physics than the Lagrangian one.

The Lagrangian point of view, often referred to as the principle of least action, can seem like magic. It sounds like a teleological principle, where the universe is consciously trying to extremize its action. The philosophical aspects of the Lagrangian formalism have puzzled researchers and laymen alike. The principle of least action has not just inspired scientific research, but pretty much everything from philosophical treatises to science fiction stories.

It is important to realize that the objection about the apparent teleological nature of the least action principle might have more to do with our experience of the world than with nature itself. Ted Chian's short story "Story of Your Life" imagines an intelligent alien species – called heptapod because of their seven-fold rotational symmetry – that does not observe the arrow of time. For a heptapod, time isn't flowing steadily forward. For them, it's just an additional dimension. Because of this, heptapods consider the principle of least action completely natural. On the other hand, they might see the Hamiltonian formalism as a mathematical trick invented by heptapod physicists.

Though fiction, "Story of Your Life" is a wonderful illustration of how our understanding of nature is shaped by our human experience. Our physics, and even our mathematics, might not be as universal as we like to think.

Aside from philosophical considerations, there are several solid arguments why the Lagrangian point of view should not be neglected. In some contexts, for example in fully discrete models, there is no way to consider a continuously flowing time. Here the Lagrangian perspective is more natural. By extension, if one wants to study connections between discrete systems and their continuous counterparts, it is probably a good idea to have a Lagrangian point of view on both sides.

Another area where the Lagrangian point of view is extremely valuable are Lorentz-invariant theories, which are much more naturally expressed in Lagrangian terms than in Hamiltonian terms. This is why Lagrangian formulations are commonplace in quantum field theory, even though quantum mechanics is traditionally a very Hamiltonian business. Back in 1933, Dirac himself suggested that the Lagrangian formulation of quantum mechanics is more natural [24].

I do not claim that the Lagrangian point of view is superior in general. It very much depends on the problem under investigation – and arguably on the taste of the researcher – which formalism is more suitable. What I am claiming is that the Lagrangian point of view is neglected in some areas. There is a gap in the scientific literature when it comes to Lagrangian theories. This thesis provides a little bit of material to help fill that gap.

Aims of this thesis

Variational principles are a powerful tool in physics and mathematics. One of their many appealing features is that they can be applied with equal ease in both continuous and discrete contexts. One obvious area where this is of use is numerical integration. Numerical methods for solving Lagrangian differential equations that discretize the variational principle rather than the equations themselves are known as *variational integrators*. For mechanical systems, where the Legendre transform allows us to switch between the Lagrangian and the Hamiltonian points of view, variational integrators are equivalent to symplectic integrators.

Most of the literature on geometric numerical integration seems to prefer the symplectic perspective over the variational one, even when this is somewhat unnatural. Consider for example the famous long-term near conservation of energy for symplectic integrators. This is proved using the fact that symplectic maps are interpolated by Hamiltonian systems. Looking at interpolating systems is a form of backward error analysis, known by the term *modified equations*. Powerful as the results are, comparing symplectic maps to continuous Hamiltonian flows is somewhat unsatisfactory. On the discrete side the symplecticity is the key notion, whereas on the continuous side the symplecticity usually appears as a consequence of the Hamiltonian form of the equations. If instead we look on the Lagrangian side, the fundamental property is the action principle, which is the same in both the continuous and the discrete world. Therefore it is worth investigating the story of modified equations from the variational perspective. Can we construct a Lagrangian for the modified equation for a given variational difference equation? That is the central question of the first part of this thesis.

A numerical integrator is *consistent* if its error converges to zero for decreasing step size, that is, if its continuum limit is again the original equation. This notion of continuum limit essentially means setting the step size equal to zero and looking at the surviving leading order term of the modified equation. It destroys all information about the integrator, except for whether or not it is consistent. What if we look beyond this leading order interpretation of continuum limit?

For generic difference equations, a higher order term in the modified equation does not have an interpretation independent of the other terms. However, that changes if we start with a difference equation that is integrable in the sense of *multidimensional consistency*. This means that we can consistently impose copies of the difference equation in a higher-dimensional lattice. If this is the case, and if we choose a suitable parameterization of the difference equation, then the terms in the power series defining the modified equation become compatible differential equations. This way a single integrable difference equation can produce a hierarchy of commuting differential equations. We consider this whole hierarchy to be the continuum limit of the difference equation.

This idea of continuum limits – embedding the discrete system in a clever way in the continuous hierarchy – goes back to Miwa [55] and we will refer to the suitable embedding by his name. The main theme of part II of this thesis is to combine this kind of continuum limit with a variational notion of integrability, that of *pluri-Lagrangian* (or *Lagrangian multiform*) systems. Just like in the case of modified equations, we will show that the variational nature of the equations is preserved when moving from discrete to continuous.

The concept of pluri-Lagrangian systems is quite new. Looking through the literature on integrable systems, there is a staggering absence of variational principles. Even though many integrable systems do posses a Lagrangian, this is not a sufficient requirement for integrability, just as having a Hamiltonian does not suffice for integrability. But while an important way of characterizing integrability is the existence of many Poisson-commuting Hamiltonians, there is no well-established definition of integrability in terms of Lagrangians.

The pluri-Lagrangian notion of integrability has its roots in the theory of multidimensionally consistent lattice equations. A classification of such equations on quadrilateral lattices was given by Adler, Bobenko and Suris [4]. All of these *ABS equations* have a variational formulation. The pluri-Lagrangian formulation combines the variational nature of an equation with its multidimensional consistency. It was first proposed by Lobb and Nijhoff [47].

An appealing feature of the pluri-Lagrangian principle is that it applies in a perfectly analogous way to both difference equations and differential equations. This raises the question how the two worlds are related. How to take the continuum limit of a discrete pluri-Lagrangian system? Answering this is the main goal of the second part of this thesis. In addition, we aim to give an accessible introduction to the topic of pluri-Lagrangian systems.

Part I.

Variational principles in numerical integration

1. Variational integrators and modified equations

Chapters 1 and 2 are an adaptation of [90]

An important technique to study the long-time behavior of numerical integrators is backward error analysis. This consists in finding a *modified equation*, a perturbation of the original differential equation whose solutions exactly interpolate the numerical solutions. When a modified equation has been found, one can study the behavior of the numerical solutions by comparing two differential equations, rather than comparing a differential equation with a difference equation.

It is a well-known and essential fact that if a symplectic integrator is applied to a Hamiltonian equation, then the resulting modified equation is Hamiltonian as well. This strongly suggests that when a variational integrator is applied to a Lagrangian system, the resulting modified equation is Lagrangian as well. The aim of Part I of this thesis is to confirm this. In this introductory chapter we review the essentials of variational integrators on the one hand, and modified equations on the other.

1.1. Variational integrators

In this section we give a concise introduction to variational integrators, inspired on Hairer, Lubich, and Wanner [35, Section VI.6]. For a detailed overview of the concept and an extensive bibliography, we refer to Marsden and West [52].

A continuous *Lagrangian* or *variational* system on the Euclidean space $\mathbb{R}^N$ is described by a smooth function $\mathcal{L} : T\mathbb{R}^N \cong \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ and the corresponding *action integral*

$$S(x) = \int_a^b \mathcal{L}(x(t), \dot{x}(t)) \, \mathrm{d}t. \tag{1.1}$$

A smooth curve $x : [a, b] \to \mathbb{R}^N : t \mapsto (x_1(t), \ldots, x_N(t))$ is a solution of the system if and only if it is a critical point of the action S in the set of all smooth curves with the same endpoints $x(a)$ and $x(b)$. Formally, this condition can be written as

$$\begin{aligned} 0 = \delta S(x) = \int_a^b \delta \mathcal{L}(x(t), \dot{x}(t)) \,\mathrm{d}t &= \int_a^b \sum_{i=1}^N \left(\frac{\partial \mathcal{L}}{\partial x_i} \delta x_i + \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \delta \dot{x}_i \right) \mathrm{d}t \\ &= \int_a^b \sum_{i=1}^N \left(\frac{\partial \mathcal{L}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) \delta x_i \,\mathrm{d}t. \end{aligned} \tag{1.2}$$

When integrating by parts to obtain the last equality we could ignore the boundary term because the boundary values of the curve are fixed, hence $\delta x(a) = \delta x(b) = 0$. Since Equation (1.2) holds for any such variation δx, the criticality of the action is characterized by the conditions

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = 0 \qquad \text{for } i = 1, \ldots, N,$$

which are known as the *Euler-Lagrange equations*. We will usually write them as a single vector-valued equation,

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{x}} = 0. \tag{1.3}$$

In general this is a second order differential equation. We will assume that the Lagrangian is *regular*, i.e. $\det \frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2} \neq 0$. Then the Euler-Lagrange equation can always be solved for $\ddot{x}$.

One approach to discretizing the Euler-Lagrange equation (1.3) is to discretize the action integral (1.1) and to consider discrete curves that are a critical points of this discrete action. Usually, one looks for a discrete Lagrangian, which is a smooth function $L_{\text{disc}} : \mathbb{R}^N \times \mathbb{R}^N \times (0, \infty) \to \mathbb{R}$, and defines the discrete action as

$$S_{\text{disc},h}\big((x_j)_{j \in \{0,\ldots,n\}}\big) = \sum_{j=1}^n h L_{\text{disc}}(x_{j-1}, x_j, h).$$

A discrete curve $x = (x_0, \ldots, x_n)$ is a critical point of $S_{\text{disc},h}$ in the set of all discrete curves with the same endpoints x_0 and x_n if and only if it satisfies the *discrete Euler-Lagrange equation*

$$\mathrm{D}_2 L_{\text{disc}}(x_{j-1}, x_j, h) + \mathrm{D}_1 L_{\text{disc}}(x_j, x_{j+1}, h) = 0 \qquad \text{for } j = 1, \ldots, n-1, \tag{1.4}$$

where $\mathrm{D}_1 L_{\text{disc}}$ and $\mathrm{D}_2 L_{\text{disc}}$ denote the vectors of partial derivatives of L_{disc} with respect to the first and second entry, respectively.

The discrete Lagrangian can be seen as a generating function for a symplectic map $(x_j, p_j) \mapsto (x_{j+1}, p_{j+1})$, determined by

$$p_j = -\mathrm{D}_1 L_{\text{disc}}(x_j, x_{j+1}, h) \qquad \text{and} \qquad p_{j+1} = \mathrm{D}_2 L_{\text{disc}}(x_j, x_{j+1}, h). \tag{1.5}$$

In this way, a variational integrator for $\mathcal{L}$ leads to a symplectic integrator for the corresponding Hamiltonian system

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial p}, \qquad \dot{p} = -\frac{\partial \mathcal{H}}{\partial x}, \tag{1.6}$$

where $p = \frac{\partial \mathcal{L}}{\partial \dot{x}}$ and the Hamilton function is given by $\mathcal{H} = \langle p\,, \dot{x}\rangle - \mathcal{L}$, considered as a function of x and p. The brackets $\langle \cdot\,, \cdot \rangle$ denote the standard scalar product on $\mathbb{R}^N$.

Example 1.1. There are many ways to obtain a discrete Lagrangian L_{disc} from a given continuous Lagrangian $\mathcal{L}$. Some examples are:

(*a*) $L_{\text{disc}}(x_j, x_{j+1}, h) = \mathcal{L}\left(\frac{x_j + x_{j+1}}{2}, \frac{x_{j+1} - x_j}{h} \right),$

in which case the symplectic map (1.5) is the one obtained by applying the implicit midpoint rule to (1.6).

(*b*) $L_{\text{disc}}(x_j, x_{j+1}, h) = \frac{1}{2}\mathcal{L}\left(x_j, \frac{x_{j+1} - x_j}{h} \right) + \frac{1}{2}\mathcal{L}\left(x_{j+1}, \frac{x_{j+1} - x_j}{h} \right),$

in which case the symplectic map (1.5) is the one obtained by applying the Störmer-Verlet method to (1.6), assuming $\mathcal{L}$ is separable. The Störmer-Verlet is a prime example of a geometric numerical integrator, as it can be used to illustrate many different concepts of geometric integration [34].

(*c*) $L_{\text{disc}}(x_j, x_{j+1}, h) = \mathcal{L}\left(x_j, \frac{x_{j+1} - x_j}{h} \right)$

or

(*d*) $L_{\text{disc}}(x_j, x_{j+1}, h) = \mathcal{L}\left(x_{j+1}, \frac{x_{j+1} - x_j}{h} \right),$

for which the symplectic maps (1.5) are the ones obtained by applying the two variants of the symplectic Euler method to (1.6).

The very least one can expect from a good numerical integrator is that the local error converges to zero as the step size gets smaller. If we could apply it with a step size of zero, we would get an exact solution. This is called consistency, but the actual definition we use for it is slightly more technical:

Definition 1.2. (*a*) A smooth function $\Phi : \left(\mathbb{R}^N\right)^2 \times (0,\infty) \to \mathbb{R}$ is a *consistent discretization* of a smooth function $g : T\mathbb{R}^N \to \mathbb{R}$ if there exist smooth functions $g_i : \left(\mathbb{R}^N\right)^{n_i} \to \mathbb{R}$ such that for any smooth curve x and for sufficiently small h there holds

$$\Phi(x(t), x(t+h), h) = g(x(t), \dot{x}(t)) + \sum_{i=1}^{\infty} h^i g_i[x(t)],$$

where the square brackets denote dependence on x and an arbitrary number of its derivatives, $g_i[x(t)] = g_i(x(t), \dot{x}(t), \ldots, x^{(n_i)}(t))$. If x or Φ is not analytic this should be interpreted as an asymptotic expansion.

(*b*) Consider a smooth function $g : T^{(2)}\mathbb{R}^N \to \mathbb{R}$, where $T^{(2)}\mathbb{R}^N$ is the second order tangent bundle of $\mathbb{R}^N$. A smooth function $\Phi : \left(\mathbb{R}^N\right)^3 \times (0,\infty) \to \mathbb{R}$ is a *consistent discretization* of g if there exist smooth functions $g_i : \left(\mathbb{R}^N\right)^{n_i} \to \mathbb{R}$ such that for any smooth curve x and for sufficiently small h there holds

$$\Phi(x(t-h), x(t), x(t+h), h) = g(x(t), \dot{x}(t), \ddot{x}(t)) + \sum_{i=1}^{\infty} h^i g_i[x(t)].$$

If x or Φ is not analytic this should be interpreted as an asymptotic expansion.

In both cases Definition 1.2 implies that a consistent discretization Φ of g, evaluated as in the definition, satisfies $\Phi = g + \mathcal{O}(h)$ as $h \to 0$.

Proposition 1.3. *If $L_{\text{disc}} : \left(\mathbb{R}^N\right)^2 \times (0,\infty) \to \mathbb{R}$ is a consistent discretization of $\mathcal{L} : T\mathbb{R}^N \to \mathbb{R}$, then the left hand side of the discrete Euler-Lagrange equation (1.4) is a consistent discretization of the left hand side of the continuous Euler-Lagrange equation (1.3).*

Proof. Fix a point t in time. From the definition of consistency it follows that there exist functions g_i such that

$$L_{\text{disc}}(x(t), x(t+h), h) = \mathcal{L}(x(t), \dot{x}(t)) + \sum_{i=1}^{\infty} h^i g_i[x(t)]$$

for any smooth curve x. Taking a constant variation of the curve, $\delta x(s) = 1$, we find

$$\begin{aligned} &\mathrm{D}_1 L_{\text{disc}}(x(t), x(t+h), h) + \mathrm{D}_2 L_{\text{disc}}(x(t), x(t+h), h) \\ &\qquad = \frac{\partial \mathcal{L}}{\partial x}(x(t), \dot{x}(t)) + \sum_{i=1}^{\infty} h^i \frac{\partial g_i[x(t)]}{\partial x(t)}. \end{aligned} \tag{1.7}$$

Taking a linear variation $\delta x(s) = s - t$ we find

$$h\mathrm{D}_2 L_{\text{disc}}(x(t), x(t+h), h) = \frac{\partial \mathcal{L}}{\partial \dot{x}}(x(t), \dot{x}(t)) + \sum_{i=1}^{\infty} h^i \frac{\partial g_i[x(t)]}{\partial \dot{x}(t)}, \tag{1.8}$$

hence there exist $\bar{g}_i[x(t)]$ such that

$$\sum_{k=2}^{\infty} \frac{(-h)^k}{k!} \frac{\mathrm{d}^k}{\mathrm{d}t^k}\big(\mathrm{D}_2 L_{\text{disc}}(x(t), x(t+h), h)\big) = \sum_{i=1}^{\infty} h^i \bar{g}_i[x(t)]. \tag{1.9}$$

Combining the last three equations, (1.7) $-\frac{\mathrm{d}}{\mathrm{d}t}$(1.8) + (1.9), we find that

$$\begin{aligned}
\mathrm{D}_1 L_{\text{disc}}(x(t), x(t+h), h) &+ \mathrm{D}_2 L_{\text{disc}}(x(t-h), x(t), h) \\
&= \frac{\partial \mathcal{L}}{\partial x}(x(t), \dot{x}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}}(x(t), \dot{x}(t)) \\
&\quad + \sum_{i=1}^{\infty} h^i \left(\frac{\partial g_i[x(t)]}{\partial x(t)} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial g_i[x(t)]}{\partial \dot{x}(t)} + \bar{g}_i[x(t)] \right). \qquad \square
\end{aligned}$$

We close this section by noting that many authors choose the discrete Lagrangian L_{disc} to be a consistent discretization of $h\mathcal{L}$, rather than of $\mathcal{L}$.

1.2. Modified equations

An important tool for studying the long-term behavior of numerical integrators is *backward error analysis*. Instead of comparing a discrete solution $(x_j)_{j \in \{0,\ldots,n\}}$ to a solution $x : [a,b] \to \mathbb{R}^N$ of the continuous system, backward error analysis compares the original differential equation to another differential equation satisfied by a curve $\widetilde{x} : [a,b] \to \mathbb{R}^N$ that interpolates the discrete solution. The latter differential equation is known as the *modified equation*.

1.2.1. First order equations

For first order equations the notion of modified equations is well-known, see for example [14, 32, 56, 74], [35, Chapter IX], and the references therein. Nevertheless, defining a modified equation is a subtle matter. Let $\Psi(x_j, x_{j+1}, h) = 0$ be a discretization of the differential equation. We would like to define a modified equation along the following lines.

Pseudodefinition. A parameter-dependent differential equation $\dot{x} = f(x,h)$ is a *modified equation* for the difference equation $\Psi(x_j, x_{j+1}, h) = 0$ if for any solution $(x_j)_{j \in \{0,\ldots,n\}}$ of the difference equation, the differential equation has a solution x that satisfies $x(jh) = x_j$ for all j.

However, we need to be more careful because the right hand side of the modified equation will generally be a power series in h that does not converge. We write

$$f(x,h) = f_0(x) + h f_1(x) + h^2 f_2(x) + \cdots,$$

and denote by $\mathcal{T}_k$ the operator which truncates a power series in h after order k,

$$\mathcal{T}_k\left(\sum_{i=0}^{\infty} f_i h^i \right) = \sum_{i=0}^{k} f_i h^i.$$

We call this the *k-th truncation* of the power series. We say that two power series f and g are equal up to order k if $\mathcal{T}_k(f) = \mathcal{T}_k(g)$, hence "up to" is to be understood as "up to and including."

Furthermore, we will need to consider families of curves parameterized by the step-size h, rather than just individual curves. Admissible families are those which are bounded and whose derivatives do not blow up as $h \to 0$.

Definition 1.4. A family $(x_h)_{h\in(0,\infty)}$ of smooth curves $x_h : [a_h, b_h] \to \mathbb{R}^N$ is called *admissible* if there exists a $h_{\max} > 0$ such that for each $k \geq 0$, $\big\| x_h^{(k)} \big\|_\infty$ is bounded as a function of $h \in (0, h_{\max}]$, where $\|\cdot\|_\infty$ denotes the supremum norm.

Admissibility of a family of curves $(x_h)_{h\in(0,\infty)}$ guarantees that in power series expansions like $x_h(t+h) = x_h(t) + h\dot{x}_h(t) + \frac{h^2}{2}\ddot{x}_h(t) + \ldots$ the asymptotic behavior of each term is determined by the exponent of h in that term. This is essential in much of what follows and would not be the case for general families of curves. Now we are in a position to define a modified equation.

Definition 1.5. Let $\Psi : \left(\mathbb{R}^N\right)^2 \times (0,\infty) \to \mathbb{R}^N$ be a consistent discretization of some $g : T\mathbb{R}^N \to \mathbb{R}$, with $\det \frac{\partial g}{\partial \dot{x}} \neq 0$. The formal differential equation $\dot{x} = f(x,h)$, where

$$f(x,h) = f_0(x) + hf_1(x) + h^2 f_2(x) + \cdots,$$

is a *modified equation* for the difference equation $\Psi(x_j, x_{j+1}, h) = 0$ if, for every k, every admissible family of solutions $(x_h)_{h\in(0,\infty)}$ of the truncated differential equation

$$\dot{x}_h = \mathcal{T}_k\left(f(x_h,h)\right), \qquad h \in (0,\infty),$$

satisfies

$$\Psi(x_h(t), x_h(t+h), h) = \mathcal{O}(h^{k+1})$$

as $h \to 0$, for all t.

Note that the discrete dynamics is invariant under scaling of Ψ by a nonzero h-dependent factor, but the condition that $\Psi(x(t), x(t+h), h) = \mathcal{O}(h^{k+1})$ is not. This is not a problem because the scaling is constrained by the fact that Ψ is a consistent discretization of some function g.

Proposition 1.6. *Let* $\Psi : \left(\mathbb{R}^N\right)^2 \times (0,\infty) \to \mathbb{R}^N$ *be a consistent discretization of some smooth* $g : T\mathbb{R}^N \to \mathbb{R}^N$*, with* $\det \frac{\partial g}{\partial \dot{x}} \neq 0$*. Then the difference equation* $\Psi(x_j, x_{j+1}, h) = 0$ *has a unique modified equation.*

Proof. Because of the consistency, the Taylor expansion of $\Psi(x(t), x(t+h), h)$ takes the form

$$\Psi(x(t), x(t+h), h) = g(x,\dot{x}) + hg_1(x,\dot{x},\ddot{x},\ldots) + h^2 g_2(x,\dot{x},\ddot{x},\ldots) + \cdots. \tag{1.10}$$

We look for a modified equation of the form

$$\dot{x} = f(x,h) = f_0(x) + hf_1(x) + h^2 f_2(x) + \cdots.$$

This ansatz allows us to write the higher derivatives of x as linear combinations of *elementary differentials* [35, Chapter III.1],

$$\begin{aligned} \dot{x} &= f, \\ \ddot{x} &= f'f, \\ x^{(3)} &= f''(f,f) + f'f'f, \\ &\vdots \end{aligned}$$

where a prime $'$ denotes differentiation with respect to x, and the arguments x and h of f and its derivatives are omitted. Plugging these expressions into Equation (1.10) we get

$$\Psi(x(t), x(t+h), h) = g(x,f) + hg_1(x, f, f'f, f''(f,f) + f'f'f, \ldots) + \cdots,$$

where again the arguments of f and its derivatives were omitted. By definition of modified equation this should be zero up to any order,

$$g(x,f) + hg_1(x, f, f'f, f''(f,f) + f'f'f, \ldots) + \cdots = 0.$$

The h^k-term of this expression is of the form

$$\frac{\partial g}{\partial \dot{x}} f_k + \text{terms depending only on } x, f_0, \ldots, f_{k-1}, g, g_1, \ldots, g_k.$$

Since $g, g_1, g_2, \ldots$ are determined by Ψ, this gives us a recurrence relation for the f_k. $\square$

Some authors (e.g. Calvo, Murua, and Sanz-Serna [14], Hairer [32]) use the following property as their definition of a modified equation.

Proposition 1.7. *Consider a difference equation of the form*

$$x_{j+1} = x_j + h\Phi(x_j, x_{j+1})$$

and let $(x_h)_{h\in(0,\infty)}$ *be an admissible family of solutions of the truncated modified equation* $\dot{x}_h = \mathcal{T}_k(f(x_h,h))$. *Then*

$$x_h(t+h) = x_h(t) + h\Phi(x_h(t), x_h(t+h)) + \mathcal{O}(h^{k+2}).$$

Proof. The difference equation can be written as $\Psi(x_j, x_{j+1}, h) = 0$, where

$$\Psi(x_j, x_{j+1}, h) = \frac{x_{j+1} - x_j}{h} - \Phi(x_j, x_{j+1}),$$

which is a consistent discretization of $\dot{x} - \Phi(x,x)$. Hence any admissible family of solutions $(x_h)_{h\in(0,\infty)}$ of the modified equation, truncated after order k, satisfies

$$\begin{aligned} x_h(t+h) - x_h(t) - h\Phi(x_h(t), x_h(t+h)) &= h\Psi(x_h(t), x_h(t+h), h) \\ &= \mathcal{O}(h^{k+2}). \end{aligned}$$

$\square$

1.2.2. Second order equations

For the purposes of this thesis we need to generalize Definition 1.5. Since we want to consider variational integrators, we need to introduce a notion of modified equations for second order difference equations.

Definition 1.8. Let $\Psi : \left(\mathbb{R}^N\right)^3 \times (0,\infty) \to \mathbb{R}^N$ be a consistent discretization of $g : T^{(2)}\mathbb{R}^N \to \mathbb{R}^N$, with $\det \frac{\partial g}{\partial \ddot{x}} \neq 0$. The formal differential equation $\ddot{x} = f(x,\dot{x},h)$, where

$$f(x,\dot{x},h) = f_0(x,\dot{x}) + hf_1(x,\dot{x}) + h^2 f_2(x,\dot{x}) + \cdots,$$

is a *modified equation* for the difference equation $\Psi(x_{j-1},x_j,x_{j+1},h) = 0$ if, for every k, every admissible family $(x_h)_{h\in(0,\infty)}$ of solutions of the truncated differential equation

$$\ddot{x}_h = \mathcal{T}_k\left(f(x_h,\dot{x}_h,h)\right)$$

satisfies

$$\Psi(x_h(t-h),x_h(t),x_h(t+h),h) = \mathcal{O}(h^{k+1})$$

as $h \to 0$, for all t.

As in the first order case, we have existence and uniqueness.

Proposition 1.9. *Let $\Psi : \left(\mathbb{R}^N\right)^3 \times (0,\infty) \to \mathbb{R}^N$ be a consistent discretization of some smooth function $g : T^{(2)}\mathbb{R}^N \to \mathbb{R}^N$, with $\det \frac{\partial g}{\partial \ddot{x}} \neq 0$. Then the difference equation $\Psi(x_{j-1},x_j,x_{j+1},h) = 0$ has a unique modified equation.*

Proof. The Taylor expansion of Ψ takes the form

$$\Psi(x(t-h),x(t),x(t+h),h) = g(x,\dot{x},\ddot{x}) + hg_1(x,\dot{x},\ddot{x},\ldots) + h^2 g_2(x,\dot{x},\ddot{x},\ldots) + \cdots. \tag{1.11}$$

We look for a modified equation of the form

$$\begin{cases} \dot{x} = v \\ \dot{v} = f(x,v,h) = f_0(x,v) + hf_1(x,v) + h^2 f_2(x,v) + \cdots. \end{cases}$$

This first order formulation of the modified equation allows us to write the higher derivatives of x as linear combinations of elementary differentials [35, Chapter III.2],

$$\begin{aligned} \ddot{x} &= f, \\ x^{(3)} &= f_x v + f_v f, \\ x^{(4)} &= f_{xx}(v,v) + 2f_{xv}(f,v) + f_x f + f_{vv}(f,f) + f_v f_x v + f_v f_v f, \\ &\vdots \end{aligned} \tag{1.12}$$

where the arguments x, v, and h of f and its derivatives were omitted, and the subscripts denote partial derivatives. Plugging these expressions into Equation (1.11) we get

$$0 = \Psi(x(t-h), x(t), x(t+h), h) = g(x, \dot{x}, f) + hg_1(x, \dot{x}, f, f_x\dot{x} + f_v f, \ldots) + \cdots,$$

where the arguments x, $\dot{x}$, and h of f and its derivatives were omitted. The h^k-term of this expression is of the form

$$\frac{\partial g}{\partial \ddot{x}} f_k + \text{terms depending only on } x, \dot{x}, f_0, \ldots, f_{k-1}, g, g_1, \ldots, g_k.$$

Since $g, g_1, g_2, \ldots$ are determined by Ψ_h, this gives us a recurrence relation for the f_k. $\square$

Example 1.10. Consider the differential equation $\ddot{x} = -U'(x)$, where $U : \mathbb{R}^N \to \mathbb{R}$ is some smooth potential, and its Störmer-Verlet discretization

$$x_{j+1} - 2x_j + x_{j-1} = -h^2 U'(x_j).$$

The modified equation is of the form

$$\ddot{x} = f(x, h) = f_0(x, \dot{x}) + h^2 f_2(x, \dot{x}) + \mathcal{O}(h^4).$$

In general we should also include odd order terms, but in this example they all vanish because of the symmetry of the difference equation. We evaluate a smooth curve x on a mesh of size h. In particular we consider $x_j = x(t)$ and

$$x_{j\pm1} = x(t \pm h) = x \pm h\dot{x} + \frac{h^2}{2}\ddot{x} \pm \frac{h^3}{6}x^{(3)} + \frac{h^4}{24}x^{(4)} \pm \frac{h^5}{120}x^{(5)} + \mathcal{O}(h^6).$$

We write $v = \dot{x}$, plug the above expansion into the difference equation, and replace derivatives using Equation (1.12). This gives us

$$\begin{aligned} -h^2 U'(x) &= h^2\ddot{x} + \frac{h^4}{12}x^{(4)} + \mathcal{O}(h^6) \\ &= h^2(f_0 + h^2 f_2) + \frac{h^4}{12}\big(f_{0,xx}(v,v) + 2f_{0,xv}(f_0, v) + f_{0,x}f_0 \\ &\qquad + f_{0,vv}(f_0, f_0) + f_{0,v}f_{0,x}v + f_{0,v}f_{0,v}f_0\big) + \mathcal{O}(h^6), \end{aligned} \tag{1.13}$$

where the arguments x and v of the f_i were omitted and the subscripts x and v denote partial derivatives, for example

$$f_{0,xv}(f_0, v) = \sum_{i,j=1}^N \frac{\partial^2 f_0}{\partial x^i \partial v^j} f_0^i v^j \qquad \text{and} \qquad f_{0,v}f_{0,v}f_0 = \sum_{i,j=1}^N \frac{\partial f_0}{\partial v^i}\frac{\partial f_0^i}{\partial v^j} f_0^j$$

where the upper indices denote the components in $\mathbb{R}^N$.

The h^2-term of Equation (1.13) gives us $f_0(x,v) = -U'(x)$. In particular, partial derivatives of f_0 with respect to v are zero. The h^4-term then reduces to $f_2 = \frac{1}{12}(U^{(3)}(x)(v,v) - U''(x)U(x))$. We find the modified equation

$$\ddot{x} = -U' + \frac{h^2}{12}\left(U^{(3)}(\dot{x},\dot{x}) - U''U\right) + \mathcal{O}(h^4),$$

where the argument x of U and of its derivatives has been omitted.

Observe that the truncation after the second order term of this modified equation is *not* an Euler-Lagrange equation because the second order term $\frac{h^2}{12}\left(U^{(3)}(\dot{x},\dot{x}) - U''U'\right)$ contains first derivatives of x but no second derivative of x. However, we will see that it can be obtained from an Euler-Lagrange equation by solving it for $\ddot{x}$ and truncating the resulting power series.

Example 1.11 (Harmonic oscillator)**.** The simplest instance of the last example is the case that x is real-valued and $U(x) = \frac{1}{2}x^2$, which gives us the difference equation

$$\frac{x_{j+1} - 2x_j + x_{j-1}}{h^2} = -x_j.$$

The modified equation for this difference equation is of the form

$$\ddot{x} = f(x,h) = f_0(x) + h^2 f_2(x) + h^4 f_4(x) + \mathcal{O}(h^6). \tag{1.14}$$

The fact that the f_i do not depend on $v = \dot{x}$ in this example vastly simplifies the calculations. It should be noted that this is very atypical behavior. In almost all other examples at least some f_i do depend on $v = \dot{x}$. From Equation (1.14) we obtain the following simplified form of the expressions in Equation (1.12)

$$\begin{aligned}
x^{(3)} &= f'\dot{x},\\
x^{(4)} &= f''\dot{x}^2 + f'f,\\
x^{(5)} &= f^{(3)}\dot{x}^3 + 3f''f\dot{x} + (f')^2\dot{x},\\
x^{(6)} &= f^{(4)}\dot{x}^4 + 6f^{(3)}f\dot{x}^2 + 5f''f'\dot{x}^2 + 3f''f^2 + (f')^2 f,\\
&\ \vdots
\end{aligned}$$

where the arguments x and h of f and its derivatives were omitted. If $x(t) = x_j$, then

$$\begin{aligned}
x_{j\pm1} = x(t \pm h) = x \pm h\dot{x} + \frac{h^2}{2}\ddot{x} \pm \frac{h^3}{6}x^{(3)} + \frac{h^4}{24}x^{(4)}\\
\pm \frac{h^5}{120}x^{(5)} + \frac{h^6}{720}x^{(6)} \pm \frac{h^7}{5040}x^{(7)} + \mathcal{O}(h^8).
\end{aligned}$$

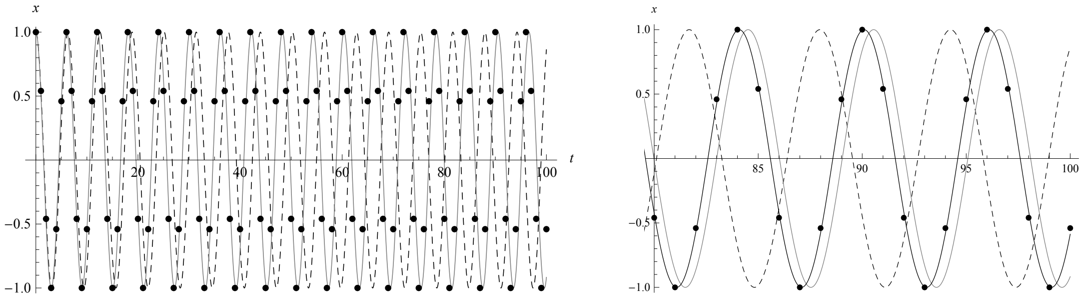


Figure 1.1. *The harmonic oscillator. The right hand image is a magnification of the time interval* $[80, 100]$. *The initial values for the differential equations are* $x(0) = 1$ *and* $\dot{x}(0) = 0$. *For the difference equation we take* $x_0 = 1$ *and* $x_1 = x(1)$, *evaluated on the exact solution.*

Dashed line: exact solution.

Bullets: solution of the Störmer-Verlet discretization with step size $h = 1$.

Gray solid line: solution of the modified equation truncated after order two.

Black solid line: solution of the modified equation truncated after order four.

Plugging this into the difference equation we find

$$\begin{aligned}-h^2 x &= h^2 \ddot{x} + \frac{h^4}{12} x^{(4)} + \frac{h^6}{360} x^{(6)} + \mathcal{O}(h^8) \\ &= h^2 \left(f_0 + h^2 f_2 + h^4 f_4 \right) \\ &\quad + \frac{h^4}{12} \left(f_0'' \dot{x}^2 + h^2 f_2'' \dot{x}^2 + f_0' f_0 + h^2 f_0' f_2 + h^2 f_2' f_0 \right) \\ &\quad + \frac{h^6}{360} \left(f_0^{(4)} \dot{x}^4 + 6 f_0^{(3)} f_0 \dot{x}^2 + 5 f_0'' f_0' \dot{x}^2 + 3 f_0'' f_0^2 + (f_0')^2 f_0 \right) + \mathcal{O}(h^8).\end{aligned}$$

The h^2-term of this equation gives us $f_0(x) = -x$, and hence $f_0'(x) = -1$ and $f_0''(x) = 0$. The h^4-term then reduces to $f_2(x) = \frac{-x}{12}$, hence $f_2'(x) = -\frac{1}{12}$ and $f_2''(x) = 0$. Finally, the h^6-term gives

$$f_4(x) = -\frac{1}{12} \left(\frac{x}{12} + \frac{x}{12} \right) + \frac{x}{360} = -\frac{x}{90}.$$

Therefore, the modified equation is

$$\ddot{x} = -x - \frac{h^2}{12} x - \frac{h^4}{90} x + \mathcal{O}(h^6).$$

In Figure 1.1 we see that the solution of the fourth truncation of the modified equation agrees very well with the discrete flow, even with a large step-size.

Also apparent from Figure 1.1 is that the discrete system is 6-periodic for $h = 1$. This can be easily verified from the difference equation as well. Of course, for generic h no periodicity is observed, so the periodic behavior is a coincidence. Or is it? The periodicity implies that the series $1 + \frac{1}{12} + \frac{1}{90} + \cdots$ converges to $\frac{\pi^2}{9}$, which in turn is related to the famous series $1 + \frac{1}{4} + \frac{1}{9} + \cdots = \frac{\pi^2}{6}$. Going into detail would lead us far out on a tangent. We refer the reader who wants to go down this rabbit hole to [92].

2. Modified Lagrangians

Chapters 1 and 2 are an adaptation of [90]

Given a variational integrator, we would like to find a Lagrangian that produces the modified equation as its Euler-Lagrange equation. The idea is to look for a modified Lagrangian $\mathcal{L}_{\mathrm{mod}}(x, \dot{x}, h)$ such that the discrete Lagrangian is its *exact discrete Lagrangian*, i.e.

$$\int_0^h \mathcal{L}_{\mathrm{mod}}(x(t), \dot{x}(t), h) \,\mathrm{d}t = h L_{\mathrm{disc}}(x_0, x_1, h), \tag{2.1}$$

where $x(t)$ is a critical curve for $\mathcal{L}_{\mathrm{mod}}$ with $x(0) = x_0$ and $x(h) = x_1$. However, unless we already know the modified equation, we have no idea which curves are critical for $\mathcal{L}_{\mathrm{mod}}$, so we will try to realize Equation (2.1) without the assumption that $x(t)$ is a critical curve. Since modified equations are generally non-convergent power series in h, the best we can hope for is to find such a modified Lagrangian up to an error of arbitrarily high order in h. Its Euler-Lagrange equation will then agree with the modified equation up to an error of the same order.

This chapter presents the construction of such a modified Lagrangian. The first incarnation of this Lagrangian, which we will construct in Section 2.2, depends on higher derivatives of the curve instead of just on x and $\dot{x}$. Furthermore, the variational principle that this Lagrangian represents is unconventional: one looks for critical curves in a set of curves that need not be differentiable everywhere. In Section 2.1 we study this *meshed variational principle* by itself and argue that it is natural to consider it.

In Section 2.3 we study some properties of admissible families of curves, which we will need in Section 2.4 to make sense of variational principles involving non-convergent power series. Once we have this understanding, we revisit the first incarnation of our modified Lagrangian and study it in more detail in Section 2.5. This paves the way for a simplified calculation of the modified equation in Section 2.6 and, finally, the derivation of a modified Lagrangian depending only on x and $\dot{x}$ in Section 2.7. In Section 2.8, we clarify our approach with some examples.

The procedure presented in Section 2.2 is similar to an approach taken by Oliver and Vasylkevych [67], who discuss the analogous problem for a variational semi-discretization of the semi-linear wave equation. The rest of this chapter has no close relation to that work. For the time being, we will assume that the Lagrangians involved are nondegenerate. In Chapter 3 we will consider a particular class of degenerate Lagrangians.

2.1. Natural boundary conditions and meshed variational problems

A variation of the curve x this is supported within the interval $(0,h)$ does not affect the right hand side of Equation (2.1), so it must also leave the action integral in the left hand side unchanged. More generally, the modified Lagrangian will be such that variations within any mesh interval do not affect the action, even if these variations lead to nonsmoothness at mesh points. Before we make this claim precise, we must make sense of variations that leave the set of smooth curves.

Definition 2.1. A *classical variational problem* consists in finding critical curves of some action $\int_a^b \mathcal{L}[x(t)]\,\mathrm{d}t$, where $[\cdot]$ denotes dependence on any number of derivatives of x, in the set of smooth curves $\mathcal{C}^\infty([a,b])$ with fixed boundary values.

A *meshed variational problem* with mesh size h consists in finding smooth curves that for every $t_0 \in \mathbb{R}$ are critical for the action $\int_a^b \mathcal{L}[x(t)]\,\mathrm{d}t$ in the set of piecewise smooth curves $\mathcal{M}^{t_0,h}$ whose nonsmooth points lie in the mesh $t_0 + h\mathbb{Z}$,

$$\mathcal{M}^{t_0,h} = \left\{ x \in \mathcal{C}^0([a,b]) \mid x \text{ is smooth on } [a,b] \setminus (t_0 + h\mathbb{Z}) \right\},$$

again with fixed boundary values.

In other words, a smooth curve x solves the meshed variational problem if it satisfies $\delta \int_a^b \mathcal{L}[x(t)]\,\mathrm{d}t = 0$ for all variations

$$\begin{aligned} \delta x &\in \bigcup_{t_0\in\mathbb{R}} \left\{ v \in \mathcal{C}^0([a,b]) \;\middle|\; v \text{ is smooth on } [a,b]\setminus(t_0+h\mathbb{Z}) \text{ and } v(a)=v(b)=0 \right\} \\ &= \bigcup_{t_0\in\mathbb{R}} \left(\mathcal{C}_0^\infty([a,b]) \;+ \sum_{t\in t_0+h\mathbb{Z}} \mathcal{C}_0^\infty([t,t+h]\cap[a,b]) \right), \end{aligned} \tag{2.2}$$

where C_0^∞ denotes the space of smooth functions that are zero on the boundary. Furthermore, it is understood that each element of $\mathcal{C}_0^\infty([t,t+h]\cap[a,b])$ is zero on $[a,b]\setminus[t,t+h]$ and thus an element of $\mathcal{C}^0([a,b])$. A few of these variations are illustrated in Figure 2.1.

Consider a classical variational problem on the interval $[a,b]$ with a Lagrange function $\mathcal{L}[x]$. The condition for criticality reads

$$\int_a^b \frac{\delta \mathcal{L}}{\delta x}\delta x\,\mathrm{d}t + \sum_{j=0}^\infty \frac{\delta \mathcal{L}}{\delta x^{(j+1)}}\delta x^{(j)}\bigg|_a^b = 0, \tag{2.3}$$

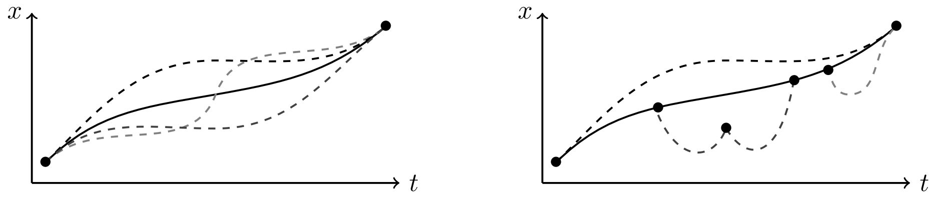


Figure 2.1. *A smooth curve and a few of its variations for classical variational problem (left) and a meshed variational problem (right).*

where

$$\frac{\delta \mathcal{L}}{\delta x^{(j)}} = \sum_{i=0}^{\infty} (-1)^i \frac{\mathrm{d}^i}{\mathrm{d}t^i} \frac{\partial \mathcal{L}}{\partial x^{(j+i)}}$$

are *variational derivatives* of $\mathcal{L}$. Equation (2.3) is easily verified using the identity

$$\frac{\delta \mathcal{L}}{\delta x^{(j)}} + \frac{\mathrm{d}}{\mathrm{d}t} \frac{\delta \mathcal{L}}{\delta x^{(j+1)}} = \frac{\partial \mathcal{L}}{\partial x^{(j)}}.$$

We assume that each of the quantities $x(a), x(b), \dot{x}(a), \dot{x}(b), \ddot{x}(a), \ddot{x}(b), \ldots$ is either fixed independently of the others or left completely free. Depending on which of those are fixed and which are left free, the following necessary and sufficient conditions follow from (2.3):

(a) $\dfrac{\delta \mathcal{L}}{\delta x} = 0.$

(b) $\forall j \geq 0:$ if $x^{(j)}(a)$ is free, $\left. \dfrac{\delta \mathcal{L}}{\delta x^{(j+1)}} \right|_{t=a} = 0,$

$\forall j \geq 0:$ if $x^{(j)}(b)$ is free, $\left. \dfrac{\delta \mathcal{L}}{\delta x^{(j+1)}} \right|_{t=b} = 0.$

Condition (a) is the Euler-Lagrange equation. Conditions (b) are known as the *natural boundary conditions* [30, Sec. 6].

Now consider a meshed variational problem on the interval $[a, b]$ with Lagrange function $\mathcal{L}[x]$. A necessary condition for criticality is that on each interval $[t, t+h] \subset [a, b]$ the corresponding classical variational problem, with boundary conditions on x but not on the derivatives, is solved. This gives the following conditions on the whole time interval $[a, b]$:

(a) $\dfrac{\delta \mathcal{L}}{\delta x} = 0.$

(b) $\forall j \geq 2: \dfrac{\delta \mathcal{L}}{\delta x^{(j)}} = 0,$ (2.4)

or equivalently: $\forall j \geq 2: \dfrac{\partial \mathcal{L}}{\partial x^{(j)}} = \dfrac{\delta \mathcal{L}}{\delta x^{(j)}} + \dfrac{\mathrm{d}}{\mathrm{d}t} \dfrac{\delta \mathcal{L}}{\delta x^{(j+1)}} = 0.$

These conditions are also sufficient, because any variation consistent with the meshed structure can be written as the sum of a smooth variation on $[a, b]$ and variations on intervals $[t, t+h]$ that vanish at the endpoints, as shown in Equation (2.2). In analogy with the classical case we call (2.4)(b) the *natural interior conditions*. They can also be seen as a higher-order version of the Weierstrass-Erdmann corner conditions [30, Sec. 15], where the time of a corner is not allowed to be varied, but every point is a corner.

Note that the conditions (2.4) are independent of h. In (2.4)(a) there is no reason to expect h-dependence because it comes from the classical part of the variational problem. Equation (2.4)(b) does depend on the mesh, in the sense that it applies at time t if we can find a mesh containing t as a mesh point, but this can be achieved for any h by setting $t_0 = t$. In other words, since we are allowed to shift the mesh, (2.4)(b) applies everywhere regardless of the mesh size h.

Since the Euler-Lagrange equation (2.4)(a) together with suitable boundary conditions already determine a unique solution, meshed variational problems are overdetermined. This should not be surprising. After all we are looking for critical curves in sets $\mathcal{M}^{t_0,h}$ of piecewise smooth curves, but at the same time require the critical curve to be in the subset $\mathcal{C}^\infty \subset \mathcal{M}^{t_0,h}$ of smooth curves.

For general Lagrangians there is no reason to hope for the existence of a nontrivial meshed critical curve, but for the modified Lagrangian that we are about to construct, any solution of the classical variational problem will solve the meshed variational problem. With this observation, to be made precise in Lemma 2.13 of Section 2.5, Equation (2.4)(b) changes from restrictive condition into useful information.

2.2. A meshed modified Lagrangian

Here we begin the construction of a modified Lagrangian from a given discrete Lagrangian L_{disc} that is a consistent discretization of some continuous Lagrangian. Using a Taylor expansion we can write the discrete Lagrangian $L_{\text{disc}}\big(x\big(t-\frac{h}{2}\big), x\big(t+\frac{h}{2}\big), h\big)$ as a function of a smooth curve x and its derivatives, all evaluated at time t,

$$\begin{aligned}\mathcal{L}_{\text{disc}}([x(t)], h) = L_{\text{disc}}\Bigg(& x(t) - \frac{h}{2}\dot{x}(t) + \frac{1}{2}\left(\frac{h}{2}\right)^2 \ddot{x}(t) - \dots, \\ & x(t) + \frac{h}{2}\dot{x}(t) + \frac{1}{2}\left(\frac{h}{2}\right)^2 \ddot{x}(t) + \dots,\ h \Bigg).\end{aligned} \tag{2.5}$$

From Equation (2.5) we proceed by expanding $L_{\text{disc}}(\cdot, \cdot, h)$ around the point $(x(t), x(t))$ to write $\mathcal{L}_{\text{disc}}([x], h)$ explicitly as a power series in h.

We could also have chosen $t-\frac{h}{2}$, $t+\frac{h}{2}$, or any other point in the interval $\left[t-\frac{h}{2}, t+\frac{h}{2}\right]$ to expand around. Choosing the midpoint has the computational advantage that the expansions of some common terms like $\frac{1}{h}\left(x\big(t+\frac{h}{2}\big)-x\big(t-\frac{h}{2}\big)\right)$ and $\frac{1}{2}\left(x\big(t-\frac{h}{2}\big)+x\big(t+\frac{h}{2}\big)\right)$ only contain even powers of h.

Proposition 2.2. *If the discrete Lagrangian L_{disc} is a consistent discretization of some $\mathcal{L}(x,\dot{x})$, then the h^k-term of $\mathcal{L}_{\mathrm{disc}}$ can depend on $x,\dot{x},\ldots,x^{(k+1)}$, but not on higher derivatives of x.*

Proof. Let $y = x + \frac{h^2}{8}\ddot{x} + \ldots$ and $z = \dot{x} + \frac{h^2}{24}x^{(3)} + \ldots$. Fix a time t_0 and consider the curve $v(t) = y(t_0) + \big(t-\frac{h}{2}\big)\, z(t_0)$. Then for some functions g_i ($i \in \mathbb{N}$) there holds

$$\begin{aligned}
\mathcal{L}_{\mathrm{disc}}([x(t_0)],h) &= L_{\mathrm{disc}}\left(y(t_0)-\frac{h}{2}z(t_0), y(t_0)+\frac{h}{2}z(t_0), h\right)\\
&= L_{\mathrm{disc}}(v(0),v(h),h)\\
&= \mathcal{L}(v(0),\dot{v}(0)) + hg_1[v(0)] + h^2 g_2[v(0)] + \ldots.
\end{aligned}$$

Since v is linear, we can assume the g_i to depend on v and $\dot{v}$ only. This means we can also write the g_i as a function of $y(t_0)$ and $z(t_0)$, say $g_i[v(0)] = \bar{g}_i(y(t_0),z(t_0))$. Then

$$\begin{aligned}
\mathcal{L}_{\mathrm{disc}}([x(t_0)],h) &= \mathcal{L}(v(0),\dot{v}(0)) + \sum_{i=1}^{\infty} h^i \bar{g}_i(y(t_0),z(t_0))\\
&= \mathcal{L}\left(y(t_0)-\frac{h}{2}z(t_0), z(t_0)\right) + \sum_{i=1}^{\infty} h^i \bar{g}_i(y(t_0),z(t_0))\\
&= \mathcal{L}(y(t_0),z(t_0)) + \sum_{i=1}^{\infty} h^i \hat{g}_i(y(t_0),z(t_0)),
\end{aligned}$$

where the Taylor series of $\mathcal{L}$ around $(y(t_0),z(t_0))$ has been absorbed into the $\hat{g}_i$. When we substitute $y = x + \frac{h^2}{8}\ddot{x} + \ldots$ and $z = \dot{x} + \frac{h^2}{24}x^{(3)} + \ldots$ in this equation, the claim follows immediately. □

We have written the discrete Lagrangian as a function of a continuous curve and found $\mathcal{L}_{\mathrm{disc}}([x(t)],h)$, but as the subscript indicates it is still very much a discrete quantity from the variational point of view. The action is still a sum,

$$S_{\mathrm{disc},h} = \sum_{j=1}^{n} hL_{\mathrm{disc}}(x(jh-h),x(jh),h) = \sum_{j=1}^{n} h\mathcal{L}_{\mathrm{disc}}\left(\left[x\big(jh-\tfrac{h}{2}\big)\right],h\right).$$

We want to write this action as an integral. To do this we require a lemma, based on the Euler-Maclaurin Formula.

Lemma 2.3. *For any smooth function $f : \mathbb{R} \to \mathbb{R}^N$ we have*

$$\sum_{j=1}^{n} hf\left(jh - \frac{h}{2}\right) \simeq \int_0^{nh} \left(\sum_{i=0}^{\infty} h^{2i} \left(2^{1-2i} - 1\right) \frac{B_{2i}}{(2i)!} f^{(2i)}(t) \right) \mathrm{d}t,$$

where B_i are the Bernoulli numbers. The symbol $\simeq$ denotes an asymptotic expansion for $h \to 0$. In general, the power series in the right hand side does not converge.

Remark. The first few terms can easily be obtained by Taylor expansion. We have

$$\begin{aligned}
\int_0^h f(t)\,\mathrm{d}t &= \int_0^h f\left(\tfrac{h}{2}\right) + \left(t - \tfrac{h}{2}\right) f'\left(\tfrac{h}{2}\right) + \frac{1}{2}\left(t - \tfrac{h}{2}\right)^2 f''\left(\tfrac{h}{2}\right) + \mathcal{O}(t^3)\,\mathrm{d}t \\
&= hf\left(\tfrac{h}{2}\right) + \frac{h^3}{24} f''\left(\tfrac{h}{2}\right) + \mathcal{O}(h^4) \\
&= hf\left(\tfrac{h}{2}\right) + \int_0^h \frac{h^2}{24} f''\left(\tfrac{h}{2}\right) \mathrm{d}t + \mathcal{O}(h^4) \\
&= hf\left(\tfrac{h}{2}\right) + \int_0^h \frac{h^2}{24} f''(t)\,\mathrm{d}t + \mathcal{O}(h^4),
\end{aligned}$$

which gives the result up to order 2 after summation:

$$\sum_{j=1}^{n} hf\left(jh - \frac{h}{2}\right) = \int_0^{nh} f(t) - \frac{h^2}{24} f''(t)\,\mathrm{d}t + \mathcal{O}(nh^4).$$

We could prove the general statement by an iteration of this procedure, but that would reinventing the wheel; a wheel known as the Euler-Maclaurin formula.

Proof of Lemma 2.3. The Euler-Maclaurin formula [1, Sec. 23.1] gives the asymptotic expansion

$$\sum_{j=1}^{n-1} g(j) \simeq \int_0^n g(t)\,\mathrm{d}t - \frac{1}{2}(g(0) + g(n)) + \sum_{i=1}^{\infty} \frac{B_{2i}}{(2i)!} \left(g^{(2i-1)}(n) - g^{(2i-1)}(0) \right)$$

for any smooth function $g : \mathbb{R} \to \mathbb{R}^N$. If we double n in this formula, we get

$$\sum_{j=1}^{2n-1} g(j) \simeq \int_0^{2n} g(t)\,\mathrm{d}t - \frac{1}{2}(g(0) + g(2n)) + \sum_{i=1}^{\infty} \frac{B_{2i}}{(2i)!} \left(g^{(2i-1)}(2n) - g^{(2i-1)}(0) \right).$$

If we double the the argument of g instead, we get

$$\sum_{j=1}^{n-1} g(2j) \simeq \int_0^n g(2t)\,\mathrm{d}t - \frac{1}{2}(g(0) + g(2n)) + \sum_{i=1}^{\infty} 2^{2i-1} \frac{B_{2i}}{(2i)!} \left(g^{(2i-1)}(2n) - g^{(2i-1)}(0) \right).$$

Taking the difference yields

$$\sum_{j=1}^{n} g(2j-1) \simeq \int_0^n g(2t)\,\mathrm{d}t + \sum_{i=1}^{\infty} \left(1-2^{2i-1}\right) \frac{B_{2i}}{(2i)!} \left(g^{(2i-1)}(2n) - g^{(2i-1)}(0) \right),$$

hence

$$\sum_{j=1}^{n} g(2j-1) \simeq \int_0^n \left(g(2t) + \sum_{i=1}^{\infty} \left(2-2^{2i}\right) \frac{B_{2i}}{(2i)!} g^{(2i)}(2t) \right) \mathrm{d}t.$$

Now set $f(t) = g\left(\frac{2}{h}t\right)$. Then

$$\sum_{j=1}^{n} f\left(hj - \frac{h}{2}\right) \simeq \int_0^n \left(f(ht) + \sum_{i=1}^{\infty} \left(2-2^{2i}\right) \frac{B_{2i}}{(2i)!} \left(\frac{h}{2}\right)^{2i} f^{(2i)}(ht) \right) \mathrm{d}t,$$

which is equivalent to the claimed result. □

Definition 2.4. We call the formal power series

$$\begin{aligned}
\mathcal{L}_{\text{mesh}}([x(t)],h) &= \sum_{i=0}^{\infty} \left(2^{1-2i}-1\right) \frac{h^{2i} B_{2i}}{(2i)!} \frac{\mathrm{d}^{2i}}{\mathrm{d}t^{2i}} \mathcal{L}_{\text{disc}}([x(t)],h) \\
&= \mathcal{L}_{\text{disc}}([x(t)],h) - \frac{h^2}{24} \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathcal{L}_{\text{disc}}([x(t)],h) + \frac{7h^4}{5760} \frac{\mathrm{d}^4}{\mathrm{d}t^4} \mathcal{L}_{\text{disc}}([x(t)],h) + \cdots
\end{aligned}$$

the *meshed modified Lagrangian* of L_{disc}.

Note that the higher order terms of the meshed modified Lagrangian do not contribute to the Euler-Lagrange equations because they are time derivatives. However, they do contribute to the natural interior conditions. Furthermore, they are needed to have (formal) equality between the discrete and the meshed modified action,

$$\begin{aligned}
S_{\text{disc},h}\big((x(t_0+jh))_{j\in\{0,\ldots,n\}}\big) &= \sum_{j=1}^{n} h \mathcal{L}_{\text{disc}}\big(\left[x\left(t_0 + jh - \tfrac{h}{2}\right)\right], h\big) \\
&\simeq \int_{t_0}^{t_0+nh} \mathcal{L}_{\text{mesh}}([x(t)],h)\,\mathrm{d}t
\end{aligned}$$

for any piecewise smooth curve $x \in \mathcal{M}^{t_0,h}$. This implies that if x is a curve such that $(x(t_0+jh))_j$ is critical for the discrete action, then x formally solves the meshed variational problem for $\mathcal{L}_{\text{mesh}}$. To state properties like this in an analytically meaningful way, we need to introduce a few new concepts. This will be done in the next two sections. Once we have those tools at our disposal, we will work towards constructing a classical, first-order Lagrangian $\mathcal{L}_{\text{mod}} : T\mathbb{R}^N \to \mathbb{R}$ for the modified equation.

2.3. Properties of admissible families of curves

Recall from Definition 1.4 that a family of curves in $\mathbb{R}^N$ is called admissible if the curves and their derivatives of any order are bounded as $h \to 0$. An admissible family of real valued curves (i.e. with $N = 1$) is called an admissible family of functions. In particular, for a family of Lagrangians $(\mathcal{L}_h)_{h\in(0,\infty)}$ that is given by a power series in h and an admissible family of curves $(x_h)_{h\in(0,\infty)}$, the compositions $\mathcal{L}_h[x_h]$ form an admissible family of functions.

Let us list a few useful lemmas. The first two follow immediately from the definition of admissibility.

Lemma 2.5. *If $(x_h)_{h\in(0,\infty)}$ is an admissible family of curves, then for every $k \in \mathbb{N}$ the family of derivatives $\big(x_h^{(k)}\big)_{h\in(0,\infty)}$ is admissible as well.*

Lemma 2.6. *If $(x_h)_{h\in(0,\infty)}$ is an admissible family of curves, then for a suitable $h_{\max} > 0$, the family $(x_h)_{h\in(0,h_{\max}]}$ is equicontinuous.*

The next lemma states that if an admissible family of functions tends to zero, then so does the family of their derivatives. The proof is an ε-δ-argument, but the idea behind it is very simple: if the functions get smaller but their derivatives do not, then they must be oscillating faster and faster, which leads to a blow up of the second derivative.

Lemma 2.7. *Let $(f_h)_{h\in(0,\infty)}$ be an admissible family of functions on the same domain $[a, b]$ and let $(h_k)_{k\in\mathbb{N}}$ be a sequence with $h_k \to 0$. If $\lim_{k\to\infty} f_{h_k} = 0$, then $\lim_{k\to\infty} f'_{h_k} = 0$. (And hence $\lim_{k\to\infty} f^{(n)}_{h_k} = 0$ for all n.)*

Because we are considering equicontinuous functions on a compact domain, the pointwise limit $\lim_{k\to\infty} f_{h_k}(t) = 0$ is equivalent to the uniform limit $\lim_{k\to\infty} \big\| f_{h_k} \big\|_\infty = 0$, so the potential ambiguity in the notation above is irrelevant.

Proof of Lemma 2.7. Suppose towards a contradiction that there exists an $\varepsilon > 0$ and a $t \in [a, b]$ such that $|f'_{h_\ell}(t)| > \varepsilon$ for all members of a subsequence of $(h_k)_{k\in\mathbb{N}}$. Without loss of generality we can assume $\varepsilon < 1$. Since $\lim_{\ell\to\infty} f_{h_\ell} = 0$, for every $j \in \mathbb{N}$ we can find a $k \in \mathbb{N}$ such that for all $\ell \geq k$ there holds $\|f_{h_\ell}\|_\infty < \frac{1}{8}\varepsilon^{j+1}$.

We claim that for every $\ell \geq k$ there exists an $s_\ell \in \big[t - \frac{1}{2}\varepsilon^j, t\big]$ such that $\big|f'_{h_\ell}(t) - f'_{h_\ell}(s_\ell)\big| \geq \frac{\varepsilon}{2}$. Indeed, if this were not the case there would hold that

$$\begin{aligned}\left|f_{h_\ell}(t) - f_{h_\ell}\left(t - \tfrac{1}{2}\varepsilon^j\right)\right| &\geq \frac{\varepsilon^j}{2} \inf_{\tau\in[t-\frac{1}{2}\varepsilon^j, t_\ell]} \left|f'_{h_\ell}(\tau)\right| \\ &> \frac{\varepsilon^j}{2}\left(\left|f'_{h_\ell}(t)\right| - \frac{\varepsilon}{2}\right) > \frac{\varepsilon^{j+1}}{4},\end{aligned}$$

which contradicts the fact that $\|f_{h_\ell}\|_\infty < \frac{1}{8}\varepsilon^{j+1}$.

Since such an s_ℓ exists, we can find an $r_\ell \in [s_\ell, t] \subset \left[t - \frac{1}{2}\varepsilon^j, t\right]$ such that $|f''_{h_\ell}(r_\ell)| > \frac{\varepsilon}{2} \big/ \frac{\varepsilon^j}{2} = \varepsilon^{1-j}$. It follows that $\limsup_{\ell\to\infty} \|f''_{h_\ell}\|_\infty \geq \lim_{j\to\infty} \varepsilon^{1-j} = \infty$, which contradicts the assumption that $(f_h)_{h\in(0,\infty)}$ is admissible. □

Lemma 2.7 can be extended to include the rate of convergence. The following lemma states this both for families on a constant domain and for families on a shrinking domain.

Lemma 2.8. (*a*) *Let $(f_h)_{h\in(0,\infty)}$ be an admissible family of functions on the same domain $[a,b]$. If $f_h = \mathcal{O}(h^\ell)$, then for all $k \in \mathbb{N}$ there holds $f_h^{(k)} = \mathcal{O}(h^\ell)$.*

(*b*) *Let $(f_h)_{h\in(0,\infty)}$ be an admissible family of functions on a shrinking domain $[a_h, b_h] = [a_h, a_h + h]$. If $f_h = \mathcal{O}(h^\ell)$, then for all $k \in \mathbb{N}$ there holds $f_h^{(k)} = \mathcal{O}(h^{\ell-k})$.*

Proof. (*a*) Since the derivatives of an admissible family of functions form an admissible family, it is sufficient to show this for $k = 1$.

Assume towards a contradiction that f'_h is not $\mathcal{O}(h^\ell)$. Then there exists a sequence $(h_j)_{j\in\mathbb{N}}$ with $h_j \to 0$ such that $\|f'_{h_j}\|_\infty > j h^\ell$. Hence

$$\lim_{j\to\infty} \frac{f_{h_j}}{\|f'_{h_j}\|_\infty} = 0.$$

Lemma 2.7, applied to the family $\left(f_h / \|f'_h\|_\infty\right)_{h\in(0,\infty)}$, implies that

$$\lim_{j\to\infty} \frac{f'_{h_j}}{\|f'_{h_j}\|_\infty} = 0,$$

but since pointwise and uniform convergence are equivalent (see Lemma 2.6), this leads to a contradiction:

$$1 = \lim_{j\to\infty} \left\| \frac{f'_{h_j}}{\|f'_{h_j}\|_\infty} \right\|_\infty = 0,$$

(*b*) Consider the functions $g_h : [0,1] \to \mathbb{R}^N$ defined by rescaling f_h:

$$g_h(t) = f_h(a_h + ht).$$

Then $\|g_h^{(k)}\|_\infty = h^k \|f_h^{(k)}\|_\infty$, so the g_h form an admissible family. Hence from part (*a*) it follows that $h^k f_h^{(k)} = \mathcal{O}(h^\ell)$. □

The final lemma of this section makes precise the fact that a function must be small if the average of any two nearby evaluations is small.

Lemma 2.9. *Let $(f_h)_{h\in(0,\infty)}$ be an admissible family of functions on the same domain $[a,b]$. If for every $t \in [a, b-h]$ there holds $f_h(t) + f_h(t+h) = \mathcal{O}(h^\ell)$, then $f_h = \mathcal{O}(h^\ell)$.*

Proof. We proceed by induction on ℓ. If $\ell = 0$ the claim follows from the definition of admissibility. Assume the statement holds for $\ell-1$. Observe that $f_h(t+h)-f_h(t) = \mathcal{O}(h)$, so $f_h(t) + f_h(t+h) = \mathcal{O}(h^\ell)$ implies $f_h(t) = \mathcal{O}(h)$. By Lemma 2.8(a) this implies that $f_h^{(k)} = \mathcal{O}(h)$ for every k. Therefore $\left(\frac{1}{h} f_h\right)_{h\in(0,\infty)}$ is an admissible family of functions. Since $\frac{1}{h} f_h(t) + \frac{1}{h} f_h(t+h) = \mathcal{O}(h^{\ell-1})$, the induction hypothesis implies that $\frac{1}{h} f_h = \mathcal{O}(h^{\ell-1})$. $\square$

Studying meshed variational problems for admissible families of curves instead of individual piecewise smooth curves is much more subtle. The reason for this is that the higher derivatives of variations on a mesh interval $[t, t+h]$ tend to increase without bound as $h \to 0$. Such variations take us outside the set of admissible families and are therefore not allowed. The next section provides us with the tools to avoid this.

2.4. k-critical families of curves

Modified equations generally are nonconvergent power series and so are modified Lagrangians. To make sense of these analytically we need to truncate the power series. It will be useful to allow an unspecified truncation error in the notion of a critical curve.

Definition 2.10. In all three cases below we assume that a full set of boundary conditions is provided and that the variations respect these boundary conditions.

(a) An admissible family $(x_h)_{h\in(0,\infty)}$ of curves $x_h : [a,b] \to \mathbb{R}$ is *k-critical* for a family of actions $S_h = \int_a^b \mathcal{L}_h \,\mathrm{d}t$ if for every admissible family of smooth variations δx_h there holds
$$\delta S_h = \mathcal{O}\big(h^{k+1} \left\| \delta x_h \right\|_1\big).$$
The set of k-critical families of curves is denoted by $\mathcal{C}_k(\mathcal{L}_h)$.

(b) An admissible family $(x_h)_{h\in(0,\infty)}$ of curves $x_h : [a,b] \to \mathbb{R}$ is *meshed k-critical* for a family of actions $S_h = \int_a^b \mathcal{L}_h \,\mathrm{d}t$ if for every admissible family of variations $\delta x_h \in \mathcal{M}^{t_h,h}$ (i.e. δx_h piecewise smooth with nonsmooth points in a mesh of size h) there holds
$$\delta S_h = \mathcal{O}\big(h^{k+1} \left\| \delta x_h \right\|_1\big).$$
The set of meshed k-critical families of curves is denoted by $\mathcal{C}_k^{\mathcal{M}}(\mathcal{L}_h)$.

(c) A family $(x_h)_{h\in(0,\infty)}$ of discrete curves $x_h = (x_{h,j})_{j\in\{0,\ldots,n_h\}}$, where $n_h \sim h^{-1}$, is *k-critical* for a family of actions $S_{\mathrm{disc},h} = \sum_j h L_{\mathrm{disc}}(\cdot,\cdot,h)$ if for every family of variations $\delta x_h = (\delta x_{h,j})_{j \in \{0,\ldots,n_h\}}$ there holds

$$\delta S_{\mathrm{disc},h} = \mathcal{O}\big(h^{k+1} \, \|\delta x_h\|_{\mathrm{disc}}\big),$$

where $\|\delta x_h\|_{\mathrm{disc}} = \sum_j h|\delta x_{h,j}|$.

Note that the scaling of the norm in the discrete case is such that for any smooth variation δx there holds

$$\|\delta x\|_1 = \big(1 + \mathcal{O}(h)\big) \|(\delta x(jh))_{j \in \{0,\ldots,n_h\}}\|_{\mathrm{disc}}.$$

The following Lemma characterizes k-critical families of curves by a natural relaxation of the usual criticality conditions.

Lemma 2.11. (a) *An admissible family $(x_h)_{h \in (0,\infty)}$ of curves $x_h : [a,b] \to \mathbb{R}$ is k-critical for the family of actions $S_h = \int_a^b \mathcal{L}_h \, \mathrm{d}t$ if and only if it satisfies the corresponding Euler-Lagrange equations with a defect of order $\mathcal{O}(h^{k+1})$:*

$$\frac{\delta \mathcal{L}_h}{\delta x} = \mathcal{O}(h^{k+1}). \tag{2.6}$$

(b) *An admissible family $(x_h)_{h \in (0,\infty)}$ of curves $x_h : [a,b] \to \mathbb{R}$ is meshed k-critical for the family of actions $S_h = \int_a^b \mathcal{L}_h \, \mathrm{d}t$ if and only if it satisfies*

$$\frac{\delta \mathcal{L}_h}{\delta x} = \mathcal{O}(h^{k+1}) \quad \text{and} \quad \frac{\partial \mathcal{L}_h}{\partial x^{(\ell)}} = \mathcal{O}(h^{k+\ell+1}) \quad \text{for all } \ell \geq 2. \tag{2.7}$$

(c) *A family $(x_h)_{h \in (0,\infty)}$ of discrete curves $x_h = (x_{h,j})_{j \in \{0,\ldots,n_h\}}$ is k-critical for the family of actions $S_{\mathrm{disc},h} = \sum_j h L_{\mathrm{disc}}(x_{h,j}, x_{h,j+1}, h)$ if and only if it satisfies the corresponding discrete Euler-Lagrange equations with a defect of order $\mathcal{O}(h^{k+1})$:*

$$\mathrm{D}_2 L_{\mathrm{disc}}(x_{h,j-1}, x_{h,j}, h) + \mathrm{D}_1 L_{\mathrm{disc}}(x_{h,j}, x_{h,j+1}, h) = \mathcal{O}(h^{k+1}).$$

Proof. (a) Consider a family of Lagrangians $\mathcal{L}_h$ and a smooth curve x. It is sufficient to consider variations that have a fixed 1-norm, say $\|\delta x_h\|_1 = 1$. For any such family of variations $(\delta x_h)_{h \in (0,\infty)}$ we have

$$\delta S_h = \int_a^b \frac{\delta \mathcal{L}_h}{\delta x} \delta x_h \, \mathrm{d}t.$$

It follows that Equation (2.6) holds if and only if $\delta S_h = \mathcal{O}(h^{k+1}) = \mathcal{O}\big(h^{k+1} \|\delta x_h\|_1\big)$ for all families of variations with $\|\delta x_h\|_1 = 1$.

(*b*) Any variation can be written as the sum of variations supported on single mesh intervals and a smooth variation, as in Equation (2.2). It is sufficient to look at these types of variations separately. Smooth variations are treated as in (*a*). For a variation supported on a mesh interval $[t_0, t_0 + h]$ we find

$$\delta S_h = \int_{t_0}^{t_0+h} \frac{\delta \mathcal{L}_h}{\delta x} \delta x_h \,\mathrm{d}t + \sum_{i=2}^{\infty} \sum_{j=0}^{i-2} (-1)^j \frac{\mathrm{d}^j}{\mathrm{d}t^j} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h^{(i-j-1)} \bigg|_{t_0}^{t_0+h}.$$

Note that we did not include $j = i-1$ in the summation range, because the variation δx_h must vanish at the endpoints t_0 and t_0+h. If $\|\delta x_h\|_1 = \mathcal{O}(h^{\ell+1})$ for some ℓ, then $\|\delta x_h\|_\infty = \mathcal{O}(h^\ell)$, hence by Lemma 2.8(*b*) we have $\big\|\delta x_h^{(i-j-1)}\big\|_\infty = \mathcal{O}(h^{\ell-i+j+1})$. Then the conditions (2.7) imply that $\delta S_h = \mathcal{O}(h^{k+\ell+2})$. Since $\|\delta x_h\|_1 = \mathcal{O}(h^{\ell+1})$, this is sufficient for meshed k-criticality.

By considering smooth variations as in part (*a*), we can conclude that also in this case the Euler-Lagrange equations up to order k are necessary conditions. More subtle to show is the necessity of the natural interior conditions. The difficulty is that the higher derivatives of variations supported on a mesh interval $[t, t+h]$ are usually unbounded as $h \to 0$, so the set of admissible families of such variations is rather small.

We will use induction on m to show that

$$\forall m \geq 0 : \ \forall k \geq 0 : \ \forall \ell \geq 2 : \quad \frac{\partial \mathcal{L}_h}{\partial x^{(\ell)}} = \mathcal{O}(h^{k+\ell+1} + h^m) \quad \text{on } k\text{-critical families.}$$

For $m = 0$ this follows from the admissibility of the family of curves.

Now fix some M and suppose the claim holds for $m = M - 1$. Take any $k \geq 0$ and $\ell \geq 2$. To construct admissible variations we consider the family of polynomials $p_{\ell,h}(t)$ of degree ℓ in t, parameterized by $h \in (0, \infty)$, that satisfies

$$\begin{aligned}
&p_{\ell,h}(0) = p_{\ell,h}(h) = 0, \\
&p'_{\ell,h}(0) = p'_{\ell,h}(h), \quad p''_{\ell,h}(0) = p''_{\ell,h}(h), \quad \ldots \quad p_{\ell,h}^{(\ell-2)}(0) = p_{\ell,h}^{(\ell-2)}(h), \\
&p_{\ell,h}^{(\ell)} \equiv 1.
\end{aligned}$$

For each ℓ and h these conditions uniquely define a polynomial because they are equivalent to $\ell + 1$ independent linear equations in the coefficients of $p_{\ell,h}$. Note that these polynomials satisfy the scaling relation $p_{\ell,h}(ht) = h^\ell p_{\ell,1}(t)$, from which it follows that $\max_{[0,h]} \big|p_{\ell,h}^{(j)}\big| = \mathcal{O}(h^{\ell-j})$. In particular, we have that $p_{\ell,h}^{(\ell-1)}(t) = t - \frac{h}{2}$.

Fix a family of real numbers $(t_h)_{h\in(0,\infty)}$. Consider the family of variations

$$\delta x_h(t) = p_{\ell,h}(t - t_h)\, \mathbb{1}_{[t_h, t_h+h]}(t)\, v,$$

where $\mathbb{1}_A$ denotes the indicator function of A and $v \in \mathbb{R}^N$ is a constant vector. Note that $\delta x_h \in \mathcal{C}_0^\infty([t_h, t_h + h])$, so it is compatible with the meshed variational problem. Since $\max_{[0,h]}|p_{\ell,h}| = \mathcal{O}(h^\ell)$ we have $\|\delta x_h\|_1 = \mathcal{O}(h^{\ell+1})$, hence for meshed k-critical families of curves there holds that

$$\begin{aligned}
\delta S_h &= \int_{t_h}^{t_h+h} \sum_{i=0}^{\ell} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h^{(i)} \,\mathrm{d}t \\
&= \int_{t_h}^{t_h+h} \sum_{i=0}^{\ell} (-1)^i \frac{\mathrm{d}^i}{\mathrm{d}t^i} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h \,\mathrm{d}t + \sum_{i=2}^{\ell} \sum_{j=0}^{i-2} (-1)^j \frac{\mathrm{d}^j}{\mathrm{d}t^j} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h^{(i-j-1)} \bigg|_{t_h}^{t_h+h} \\
&= \mathcal{O}(h^{k+\ell+2}),
\end{aligned} \tag{2.8}$$

where we could limit the summation range to $i \leq \ell$ because higher derivatives of the variation δx_h vanish. We have already established that $\frac{\delta \mathcal{L}_h}{\delta x} = \mathcal{O}(h^{k+1})$ on meshed k-critical families and the induction hypothesis implies that $\frac{\partial \mathcal{L}_h}{\partial x^{(i)}} = \mathcal{O}(h^{k+1} + h^{M-1})$ for $i \geq 2$ on meshed k-critical families. It follows that

$$\int_{t_h}^{t_h+h} \sum_{i=0}^{\ell} (-1)^i \frac{\mathrm{d}^i}{\mathrm{d}t^i} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h \,\mathrm{d}t = \mathcal{O}(h^{k+\ell+2} + h^{M+\ell}),$$

hence Equation (2.8) implies

$$\sum_{i=2}^{\ell} \sum_{j=0}^{i-2} (-1)^j \frac{\mathrm{d}^j}{\mathrm{d}t^j} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h^{(i-j-1)} \bigg|_{t_h}^{t_h+h} = \mathcal{O}(h^{k+\ell+2} + h^{M+\ell}). \tag{2.9}$$

Using the fact that $\left\| \delta x_h^{(i-j-1)} \right\|_\infty = \mathcal{O}(h^{\ell-i+j+1})$ and the induction hypothesis, we find that for all i and j in the range of this sum, except $(i, j) = (\ell, 0)$,

$$\frac{\mathrm{d}^j}{\mathrm{d}t^j} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h^{(i-j-1)} = \mathcal{O}(h^{k+\ell+2} + h^{M+1}),$$

hence

$$(-1)^j \frac{\mathrm{d}^j}{\mathrm{d}t^j} \frac{\partial \mathcal{L}_h}{\partial x^{(i)}} \delta x_h^{(i-j-1)} \bigg|_{t_h}^{t_h+h} = \mathcal{O}(h^{k+\ell+2} + h^{M+1}).$$

Equation (2.9) now implies that the term with $(i, j) = (\ell, 0)$ satisfies the same order condition:

$$\frac{h}{2} \left(\frac{\partial \mathcal{L}_h}{\partial x^{(\ell)}} \bigg|_{t_h} + \frac{\partial \mathcal{L}_h}{\partial x^{(\ell)}} \bigg|_{t_h+h} \right) = \mathcal{O}(h^{k+\ell+2} + h^{M+1}).$$

By Lemma 2.9 it follows that

$$\frac{\partial \mathcal{L}_h}{\partial x^{(\ell)}} = \mathcal{O}(h^{k+\ell+1} + h^M).$$

This concludes the induction step and thus the proof that the interior conditions, up to the appropriate order, are necessary for k-criticality.

(c) If the family $(x_h)_{h\in(0,\infty)}$ of discrete curves is k-critical, then

$$\begin{aligned}\sum_j h\left(\mathrm{D}_2 L_{\mathrm{disc}}(x_{h,j-1},x_{h,j},h)+\mathrm{D}_1 L_{\mathrm{disc}}(x_{h,j},x_{h,j+1},h)\right)\delta x_{h,j} &= \delta S_{\mathrm{disc},h} \\ &= \mathcal{O}(h^{k+1}\|\delta x_h\|_{\mathrm{disc}}).\end{aligned}$$

For some index ℓ, set $\delta x_{h,\ell}=\frac{1}{h}$ and $\delta x_{h,j}=0$ for $j\neq\ell$, then $\|\delta x_h\|_{\mathrm{disc}}=1$. It follows that

$$\mathrm{D}_2 L_{\mathrm{disc}}(x_{h,\ell-1},x_{h,\ell})+\mathrm{D}_1 L_{\mathrm{disc}}(x_{h,\ell},x_{h,\ell+1})=\mathcal{O}(h^{k+1}).$$

On the other hand, for any family of variations $(\delta x_h)_{h\in(0,\infty)}$ with norm $\|\delta x_h\|_{\mathrm{disc}}=1$ we have

$$\begin{aligned}|\delta S_{\mathrm{disc},h}| &\leq \sum_j h\left|\left(\mathrm{D}_2 L_{\mathrm{disc}}(x_{h,j-1},x_{h,j})+\mathrm{D}_1 L_{\mathrm{disc}}(x_{h,j},x_{h,j+1})\right)\delta x_{h,j}\right| \\ &\leq \left(\sum_j h|\delta x_{h,j}|\right)\max_j\left(\left|\mathrm{D}_2 L_{\mathrm{disc}}(x_{h,j-1},x_{h,j})+\mathrm{D}_1 L_{\mathrm{disc}}(x_{h,j},x_{h,j+1})\right|\right) \\ &= \max_j\left(\left|\mathrm{D}_2 L_{\mathrm{disc}}(x_{h,j-1},x_{h,j})+\mathrm{D}_1 L_{\mathrm{disc}}(x_{h,j},x_{h,j+1})\right|\right).\end{aligned}$$

Hence $\left((x_{h,j})_{j\in\{0,\dots,n_h\}}\right)_{h\in(0,\infty)}$ is k-critical if the discrete Euler-Lagrange equations are satisfied up to order k. □

2.5. Properties of the meshed modified Lagrangian

Now that we have established the analytic framework, it is time to list some important properties of the meshed modified Lagrangian.

Lemma 2.12. *Let L_{disc} be a consistent discretization of a regular Lagrangian $\mathcal{L}(x,\dot{x})$. Then the zeroth order term of the modified Lagrangian is the original continuous Lagrangian, i.e. $\mathcal{L}_{\mathrm{mesh}}([x],h)=\mathcal{L}(x,\dot{x})+\mathcal{O}(h)$.*

Proof. We have

$$\begin{aligned}\mathcal{L}_{\mathrm{mesh}}([x(t)],h) &= \mathcal{L}_{\mathrm{disc}}([x(t)],h)+\mathcal{O}(h^2) \\ &= L_{\mathrm{disc}}\left(x\left(t-\tfrac{h}{2}\right),x\left(t+\tfrac{h}{2}\right),h\right)+\mathcal{O}(h^2) \\ &= \mathcal{L}(x(t),\dot{x}(t))+\mathcal{O}(h).\end{aligned}$$

□

An essential property of the meshed modified Lagrangian is that any curve that solves the Euler-Lagrange equations automatically satisfies the natural interior conditions. Hence for this particular class of Lagrangians, the natural interior conditions are not a additional restrictions compared to the classical variational problem, but instead they provide us with useful information about critical curves.

Lemma 2.13. *If a family of curves $(x_h)_{h\in(0,\infty)}$ satisfies the Euler-Lagrange equations of $\mathcal{L}_{\text{mesh}}$ up to order k then it satisfies the natural interior conditions*

$$\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x^{(\ell)}} = \mathcal{O}(h^{k+\ell+1}) \qquad \text{for all } \ell \geq 2.$$

In other words, for $\mathcal{L}_{\text{mesh}}$ every k-critical family of curves is also meshed k-critical, $\mathcal{C}_k^{\mathcal{M}}(\mathcal{L}_{\text{mesh}}) = \mathcal{C}_k(\mathcal{L}_{\text{mesh}})$.

Proof. Consider the same family of polynomials $p_{\ell,h}(t)$ as in the proof of Lemma 2.11(b) and the corresponding family of variations $\delta x_h(t) = p_{\ell,h}(t-t_0)\mathbb{1}_{[t_h,t_h+h]}v$. Since these variations do not affect the discrete action

$$S_{\text{disc},h} = \sum_j L_{\text{disc}}(x_h(t_h+(j-1)h), x_h(t_h+jh), h),$$

there holds for every curve that

$$\int_a^b \sum_{i=0}^{\ell} \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x^{(i)}} \delta x_h^{(i)} \,\mathrm{d}t \simeq \delta S_{\text{disc},h} = 0.$$

In particular this implies Equation (2.8). Since the Euler-Lagrange equations are satisfied up to order k, we can proceed exactly as in the proof of Lemma 2.11(b). □

The modified Lagrangian depends on fewer derivatives of x than $\mathcal{L}_{\text{disc}}$ (cf. Proposition 2.2):

Proposition 2.14. *For $\ell \geq 1$ the h^ℓ-term of $\mathcal{L}_{\text{mesh}}$ (as a power series in h) can depend on $x, \dot{x}, \ldots, x^{(\ell)}$, but not on higher derivatives of x.*

Proof. Any admissible family of curves satisfies the Euler-Lagrange equations up to order -1:

$$\frac{\delta \mathcal{L}_{\text{mesh}}}{\delta x} = \mathcal{O}(1).$$

Hence it follows from Lemma 2.13 that for all $\ell \geq 2$:

$$\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x^{(\ell)}} = \mathcal{O}(h^\ell),$$

which implies that $x^{(\ell)}$ can only occur in $\mathcal{L}_{\text{mesh}}$ in terms of order at least h^ℓ. □

2.6. The modified equation

From Lemma 2.13 it follows that k-critical families of curves for $\mathcal{L}_{\mathrm{mesh}}$ satisfy the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial \dot{x}} = \mathcal{O}(h^{k+1})$$

even though $\mathcal{L}_{\mathrm{mesh}}$ depends on higher derivatives of x. By Proposition 2.14, this equation takes the form

$$\begin{aligned} &\mathcal{E}_0(x,\dot{x},\ddot{x}) + h\mathcal{E}_1(x,\dot{x},\ddot{x}) \\ &\quad + h^2\mathcal{E}_2(x,\dot{x},\ddot{x},x^{(4)}) + \ldots + h^k\mathcal{E}_k(x,\dot{x},\ldots,x^{(2k)}) = \mathcal{O}(h^{k+1}). \end{aligned} \tag{2.10}$$

If we replace the error term by an exact zero, this is a singularly perturbed equation, whose solutions in general have increasingly steep boundary layers as $h \to 0$. However, the condition that $(x_h)_{h\in(0,\infty)}$ is an admissible family of curves excludes this behavior and allows us to write Equation (2.10) as a second order differential equation with an $\mathcal{O}(h^{k+1})$ defect. This is done by a simple recursion.

If L_{disc} is a consistent discretization of some regular continuous Lagrangian, then for sufficiently small h we can solve $\mathcal{E}_0(x,\dot{x},\ddot{x}) + h\mathcal{E}_1(x,\dot{x},\ddot{x}) = \mathcal{O}(h^2)$ for $\ddot{x}$, say

$$\ddot{x} = F_1(x,\dot{x},h) + \mathcal{O}(h^2).$$

Now suppose we know that Equation (2.10) implies $\ddot{x} = F_k(x,\dot{x},h) + \mathcal{O}(h^{k+1})$. Then there exist functions $F_k^2, F_k^3, \ldots : T\mathbb{R} \times (0,\infty) \to \mathbb{R}$ such that

$$\begin{aligned} \ddot{x} &= F_k^2(x,\dot{x},h) + \mathcal{O}(h^{k+1}) = F_k(x,\dot{x},h) + \mathcal{O}(h^{k+1}), \\ x^{(3)} &= F_k^3(x,\dot{x},h) + \mathcal{O}(h^{k+1}), \\ x^{(4)} &= F_k^4(x,\dot{x},h) + \mathcal{O}(h^{k+1}), \\ &\;\vdots \end{aligned}$$

Then Equation (2.10) (with k replaced by $k+2$) implies

$$\begin{aligned} &\mathcal{E}_0(x,\dot{x},\ddot{x}) + h\mathcal{E}_1(x,\dot{x},\ddot{x}) \\ &\quad + \left(h^2\mathcal{E}_2(x,\dot{x},\ddot{x},x^{(4)}) + \ldots + h^{k+2}\mathcal{E}_k(x,\dot{x},\ldots,x^{(2k+2)}) \right)\Big|_{x^{(j)}=F_k^j(x,\dot{x},h)} = \mathcal{O}(h^{k+3}). \end{aligned}$$

After making the replacements, the terms between the parentheses only depend on x and its first derivative. Hence we can solve this equation for $\ddot{x}$ to find an expression of the form $\ddot{x} = F_{k+2}(x,\dot{x},h) + \mathcal{O}(h^{k+3})$.

Note that each step of this recursion increases the order of accuracy by two. This is the case because we only replace derivatives in terms of second and higher order.

Alternatively, we could start from $\mathcal{L}_{\mathrm{disc}}$, because it differs from $\mathcal{L}_{\mathrm{mesh}}$ by a time derivative. In that case we do not have the property that the Euler-Lagrange equations imply the natural interior conditions, so we cannot truncate the variational derivative. Hence we should start from the equation

$$\frac{\delta \mathcal{L}_{\text{disc}}}{\delta x} = \mathcal{O}(h^{k+1}).$$

2.7. A classical modified Lagrangian

Definition 2.15. The *modified Lagrangian* is the formal power series

$$\mathcal{L}_{\text{mod}}(x, \dot{x}, h) = \mathcal{L}_{\text{mesh}}([x], h)\Big|_{\ddot{x}=f(x,\dot{x},h),\ x^{(3)}=\frac{\mathrm{d}}{\mathrm{d}t}f(x,\dot{x},h),\ \ldots},$$

where $\ddot{x} = f(x, \dot{x}, h)$ is the modified equation. The k-th truncation of the modified Lagrangian is denoted by $\mathcal{L}_{\text{mod},k}$,

$$\mathcal{L}_{\text{mod},k}(x, \dot{x}, h) = \mathcal{T}_k(\mathcal{L}_{\text{mod}}(x, \dot{x}, h)) = \mathcal{T}_k\left(\mathcal{L}_{\text{mesh}}([x], h)\Big|_{x^{(j)}=F^j_{k-2}(x,\dot{x},h)}\right),$$

where $\mathcal{T}_k$ denotes truncation after the h^k-term.

From the definition it follows that $\mathcal{L}_{\text{mod},k}(x_h, \dot{x}_h, h) = \mathcal{L}_{\text{mesh}}([x_h], h) + \mathcal{O}(h^{k+1})$ for families of curves $(x_h)_{h\in(0,\infty)}$ that are k-critical for $\mathcal{L}_{\text{mesh}}$. Since this does not hold for general curves, it does not immediately imply that the Euler-Lagrange equations of both Lagrangians agree up to order k. Indeed, to get the Euler-Lagrange equations we need to take arbitrary variations, which take us away from critical curves. Nevertheless, this property holds true.

Lemma 2.16. *The meshed modified Lagrangian $\mathcal{L}_{\text{mesh}}([x], h)$ and the first-order modified Lagrangian $\mathcal{L}_{\text{mod},k}(x, \dot{x}, h)$ have the same k-critical families of curves.*

Proof. We proceed by induction. Since $\mathcal{L}_{\text{mesh}}([x], h) = \mathcal{L}_{\text{mod},0}(x, \dot{x}, h) + \mathcal{O}(h)$ they have the same 0-critical families of curves, $C_0(\mathcal{L}_{\text{mesh}}) = C_0(\mathcal{L}_{\text{mod},0})$. Now suppose that $C_{k-1}(\mathcal{L}_{\text{mesh}}) = C_{k-1}(\mathcal{L}_{\text{mod},k-1})$. Since k-critical families of curves are also $(k-1)$-critical, this set contains all k-critical families of curves of both $\mathcal{L}_{\text{mesh}}$ and $\mathcal{L}_{\text{mod},k}$.

For every family of curves in $C_{k-1}(\mathcal{L}_{\text{mesh}})$, or even in $C_{k-2}(\mathcal{L}_{\text{mesh}})$, Lemma 2.13 implies that $\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x^{(\ell)}} = \mathcal{O}(h^{k+1})$ for $\ell \geq 2$, and there holds that $x^{(j)} = F^j_{k-2}(x, \dot{x}, h) + \mathcal{O}(h^{k-1})$. Therefore,

$$\begin{aligned}
\frac{\partial \mathcal{L}_{\text{mod},k}}{\partial x} &= \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x} + \sum_{\ell=2}^{\infty} \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x^{(\ell)}}\Big|_{x^{(j)}=F^j_{k-2}(x,\dot{x},h)} \frac{\partial F^\ell_{k-2}(x, \dot{x}, h)}{\partial x} + \mathcal{O}(h^{k+1}) \\
&= \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x} + \mathcal{O}(h^{k+1})
\end{aligned}$$

and

$$\begin{aligned}\frac{\partial \mathcal{L}_{\mathrm{mod},k}}{\partial \dot{x}} &= \frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial \dot{x}} + \sum_{\ell=2}^{\infty} \left.\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial x^{(\ell)}}\right|_{x^{(j)}=F^j_{k-2}(x,\dot{x},h)} \frac{\partial F^{\ell}_{k-2}(x,\dot{x},h)}{\partial \dot{x}} + \mathcal{O}(h^{k+1}) \\ &= \frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial \dot{x}} + \mathcal{O}(h^{k+1}),\end{aligned}$$

hence

$$\frac{\partial \mathcal{L}_{\mathrm{mod},k}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\mathrm{mod},k}}{\partial \dot{x}} = \sum_{\ell=0}^{\infty} (-1)^{\ell} \frac{\mathrm{d}^{\ell}}{\mathrm{d}t^{\ell}} \frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial x^{(\ell)}} + \mathcal{O}(h^{k+1}),$$

which shows that $C_k(\mathcal{L}_{\mathrm{mesh}}) = C_k(\mathcal{L}_{\mathrm{mod},k})$. □

We now arrive at our main result: up to truncations, the modified equation is Lagrangian in the classical sense.

Theorem 2.17. *For a discrete Lagrangian L_{disc} that is a consistent discretization of a regular Lagrangian $\mathcal{L}$, the k-th truncation of the Euler-Lagrange equation of the modified Lagrangian $\mathcal{L}_{\mathrm{mod},k}(x,\dot{x},h)$, solved for $\ddot{x}$, is the k-th truncation of the modified equation.*

Proof. Let $(x_h)_{h\in(0,\infty)}$ be an admissible family of solutions of the k-th truncation of the Euler-Lagrange equation for $\mathcal{L}_{\mathrm{mod},k}$. Then $(x_h)_{h\in(0,\infty)}$ is k-critical for the family of actions $\int_a^b \mathcal{L}_{\mathrm{mod},k}(x,\dot{x},h)\,\mathrm{d}t$. Consider the family of discrete curves $(z_h)_{h\in(0,\infty)}$ with $z_{h,j} = x_h(jh)$, an admissible family of variations δx_h of x_h, and the corresponding family of variations δz_h with $\delta z_{h,j} = \delta x_h(jh)$. Then $\|\delta x_h\|_1 = \big(1+\mathcal{O}(h)\big)\|\delta z_h\|_{\mathrm{disc}}$.

By Lemma 2.16, the family $(x_h)_{h\in(0,\infty)}$ is k-critical for $\mathcal{L}_{\mathrm{mesh}}$. By construction, the actions $\sum_j hL_{\mathrm{disc}}(y(jh), y((j+1)h), h)$ and $\int_a^b \mathcal{L}_{\mathrm{mesh}}([y(t)],h)\,\mathrm{d}t$ are (formally) equal for any smooth curve y. Therefore

$$\begin{aligned}\delta S_{\mathrm{disc},h}(z_h) &= \delta \sum_j hL_{\mathrm{disc}}(x_h(jh), x_h((j+1)h), h) \\ &\simeq \delta \int_a^b \mathcal{L}_{\mathrm{mesh}}([x_h(t)],h)\mathrm{d}t = \mathcal{O}\big(h^{k+1}\|\delta x_h\|_1\big) = \mathcal{O}\big(h^{k+1}\|\delta z\|_{\mathrm{disc}}\big),\end{aligned}$$

so the family of discrete curves z_h is k-critical for the family of discrete actions $S_{\mathrm{disc},h}$. Hence the $z_{h,j} = x_h(jh)$ satisfy the discrete Euler-Lagrange equation up to order h^k, i.e.

$$\mathrm{D}_2 L_{\mathrm{disc}}(x_h(t-h), x_h(t), h) + \mathrm{D}_1 L_{\mathrm{disc}}(x_h(t), x_h(t+h), h) = \mathcal{O}(h^{k+1}).$$

By Proposition 1.3 the left hand side of this expression is a consistent discretization of the continuous Euler-Lagrange equation, so this order condition is the one defining a modified equation as in Definition 1.8. □

Theorem 2.17 provides an alternative proof of the following well-known result [35, Chapter IX.3].

Corollary 2.18. *If a symplectic method is applied to a Hamiltonian system with a regular Hamiltonian $H(p,q)$, then any truncation of the resulting modified equation, written in its first order form $\dot{p} = \ldots$, $\dot{q} = \ldots$, is Hamiltonian.*

Proof. Fix any truncation order k. By applying the Legendre transformation we obtain a Lagrangian system with regular Lagrangian. We can apply Theorem 2.17 to this Lagrangian to find a modified Lagrangian $\mathcal{L}_{\mathrm{mod},k}(x,\dot{x},h)$. For sufficiently small h, the modified Lagrangian is regular as well. Therefore we can take the Legendre transformation again and obtain a modified Hamiltonian $\mathcal{H}_{\mathrm{mod},k}$. Its equations of motion

$$\dot{p} = -\frac{\partial \mathcal{H}_{\mathrm{mod},k}}{\partial q}, \qquad \dot{q} = \frac{\partial \mathcal{H}_{\mathrm{mod},k}}{\partial p}$$

agree with the modified equation up to order k. Hence its k-th truncation $\mathcal{T}_k(\mathcal{H}_{\mathrm{mod},k})$ is a Hamiltonian for the k-th truncation of the modified equation

$$\dot{p} = \mathcal{T}_k\left(-\frac{\partial \mathcal{H}_{\mathrm{mod},k}}{\partial q} \right), \qquad \dot{q} = \mathcal{T}_k\left(\frac{\partial \mathcal{H}_{\mathrm{mod},k}}{\partial p} \right). \qquad \square$$

2.8. Examples

2.8.1. Störmer-Verlet discretization of mechanical Lagrangians

A Lagrangian $\mathcal{L} : T\mathbb{R}^N \to \mathbb{R}$ is called *separable* if there exists functions K and U such that $\mathcal{L}(x,\dot{x}) = K(\dot{x}) - U(x)$. The Euler-Lagrange equation of such a Lagrangian is

$$\frac{\partial^2 K(\dot{x})}{\partial \dot{x}^2}\, \ddot{x} = -\frac{\partial U(x)}{\partial x}.$$

If $\mathcal{L}$ is separable, then the discrete Lagrangians (*b*), (*c*), and (*d*) from Example 1.1 are equivalent (but their discrete Legendre transforms are different). A separable Lagrangian with $K(\dot{x}) = \frac{1}{2}|\dot{x}|^2$ is called a *mechanical Lagrangian*.

Second order

We consider some mechanical Lagrangian $\mathcal{L}(x,\dot{x}) = \frac{1}{2}|\dot{x}|^2 - U(x)$ and use the Störmer-Verlet discretization, whose discrete Lagrangian is given in Example 1.1(*b*),

$$L_{\mathrm{disc}}(x_j, x_{j+1}, h) = \frac{1}{2}\left| \frac{x_{j+1} - x_j}{h} \right|^2 - \frac{1}{2} U\left(x_j\right) - \frac{1}{2} U\left(x_{j+1}\right).$$

Its Euler-Lagrange equation is

$$\frac{x_{j+1} - 2x_j + x_{j-1}}{h^2} = -U'(x_j).$$

We have

$$\begin{aligned}
\mathcal{L}_{\text{disc}}([x], h) &= \frac{1}{2}\left|\dot{x} + \frac{h^2}{24}x^{(3)} + \dots\right|^2 - \frac{1}{2}U\left(x - \frac{h}{2}\dot{x} + \frac{h^2}{8}\ddot{x} - \dots\right) \\
&\qquad\qquad\qquad\qquad\qquad\quad - \frac{1}{2}U\left(x + \frac{h}{2}\dot{x} + \frac{h^2}{8}\ddot{x} + \dots\right) \\
&= \frac{1}{2}|\dot{x}|^2 - U + \frac{h^2}{24}\left(\left\langle \dot{x}\,, x^{(3)}\right\rangle - 3U'\ddot{x} - 3U''(\dot{x},\dot{x})\right) + \mathcal{O}(h^4),
\end{aligned}$$

where the argument x of U and of its derivatives has been omitted. Note that $U^{(k)}$ is a symmetric k-tensor, hence the notations $U'\ddot{x}$ and $U''(\dot{x},\dot{x})$.

From $\mathcal{L}_{\text{disc}}([x],h)$ we calculate the meshed modified Lagrangian as follows:

$$\begin{aligned}
\mathcal{L}_{\text{mesh}}([x], h) &= \mathcal{L}_{\text{disc}}([x], h) - \frac{h^2}{24}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{L}_{\text{disc}}([x], h) + \mathcal{O}(h^4) \\
&= \frac{1}{2}|\dot{x}|^2 - U + \frac{h^2}{24}\left(\left\langle \dot{x}\,, x^{(3)}\right\rangle - 3U'\ddot{x} - 3U''(\dot{x},\dot{x})\right) \\
&\qquad\qquad - \frac{h^2}{24}\left(|\ddot{x}|^2 + \left\langle \dot{x}\,, x^{(3)}\right\rangle - U'\ddot{x} - U''(\dot{x},\dot{x})\right) + \mathcal{O}(h^4) \\
&= \frac{1}{2}|\dot{x}|^2 - U + \frac{h^2}{24}\left(-|\ddot{x}|^2 - 2U'\ddot{x} - 2U''(\dot{x},\dot{x})\right) + \mathcal{O}(h^4).
\end{aligned} \tag{2.11}$$

Only one of the natural interior conditions for $\mathcal{L}_{\text{mesh}}$ is nontrivial at this truncation order, it reads

$$0 = \frac{\partial \mathcal{L}}{\partial \ddot{x}} = \frac{h^2}{12}(-\ddot{x} - U') + \mathcal{O}(h^4).$$

As predicted by Lemma 2.13, this is a consequence of the Euler-Lagrange equation.

The modified equation up to second order is then obtained from

$$\mathcal{O}(h^4) = \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial \dot{x}} = -\ddot{x} - U' + \frac{h^2}{12}\left(U''\ddot{x} + U^{(3)}(\dot{x},\dot{x})\right).$$

We solve this recursively for $\ddot{x}$. In the leading order we have $\ddot{x} = -U'$, the original equation, so in the second order term we can substitute $\ddot{x} = -U'$. Hence the modified equation is

$$\ddot{x} = -U' + \frac{h^2}{12}\left(U^{(3)}(\dot{x},\dot{x}) - U''U'\right) + \mathcal{O}(h^4), \tag{2.12}$$

as we already found in Example 1.10.

To obtain the classical modified Lagrangian we need to replace higher derivatives in the meshed modified Lagrangian (2.11) using the modified Equation (2.12). In fact, to get the modified Lagrangian up two order two, we only need the leading order term $\ddot{x} = -U'$ of the modified equation. We find

$$\mathcal{L}_{\mathrm{mod},3}(x,\dot{x},h) = \frac{1}{2}|\dot{x}|^2 - U + \frac{h^2}{24}\Big(\big|U'\big|^2 - 2U''(\dot{x},\dot{x})\Big). \tag{2.13}$$

Observe that the modified Lagrangian $\mathcal{L}_{\mathrm{mod},3}(x,\dot{x},h)$ is not separable for general U because the term $U''(\dot{x},\dot{x})$ depends on both x and $\dot{x}$. The Euler-Lagrange equation of $\mathcal{L}_{\mathrm{mod},3}$ is

$$-\ddot{x} - U' + \frac{h^2}{12}\Big(U''U' + U^{(3)}(\dot{x},\dot{x}) + 2U''\ddot{x}\Big) = 0.$$

Note that this equation does not contain an error term. However when we solve it for $\ddot{x}$ we again get (2.12), including the $\mathcal{O}(h^4)$ error term. In other words, $\ddot{x} = -U' + \frac{h^2}{12}\big(U^{(3)}(\dot{x},\dot{x}) - U''U'\big)$ is *not* the Euler-Lagrange equation for $\mathcal{L}_{\mathrm{mod},3}$, but it is $\mathcal{O}(h^4)$-close to it.

Fourth order

We extend the calculations of Section 2.8.1 to include the h^4-terms. We find

$$\begin{aligned}\mathcal{L}_{\mathrm{disc}}([x],h) = \frac{1}{2}|\dot{x}|^2 - U &+ \frac{h^2}{24}\Big(-3U''(\dot{x},\dot{x}) - 3U'\ddot{x} + \big\langle \dot{x}\,, x^{(3)}\big\rangle\Big) \\ &+ \frac{h^4}{5760}\Big(-45U''(\ddot{x},\ddot{x}) - 90U^{(3)}(\ddot{x},\dot{x},\dot{x}) - 60U''(x^{(3)},\dot{x}) \\ &\qquad + 5\big|x^{(3)}\big|^2 - 15U^{(4)}(\dot{x},\dot{x},\dot{x},\dot{x}) - 15U'x^{(4)} + 3\big\langle \dot{x}\,, x^{(5)}\big\rangle\Big) + \mathcal{O}(h^6)\end{aligned}$$

and

$$\begin{aligned}\mathcal{L}_{\mathrm{mesh}}([x],h) = \frac{1}{2}|\dot{x}|^2 - U &+ \frac{h^2}{24}\left(-2U''(\dot{x},\dot{x}) - 2U'\ddot{x} - |\ddot{x}|^2\right) \\ &+ \frac{h^4}{720}\Big(3U''(\ddot{x},\ddot{x}) + 6U^{(3)}(\ddot{x},\dot{x},\dot{x}) + 4U''(x^{(3)},\dot{x}) + 2\big|x^{(3)}\big|^2 \\ &\qquad + U^{(4)}(\dot{x},\dot{x},\dot{x},\dot{x}) + U'x^{(4)} + \big\langle \ddot{x}\,, x^{(4)}\big\rangle\Big) + \mathcal{O}(h^6).\end{aligned}$$

To eliminate higher derivatives of x in the h^4-term we can use $\ddot{x} = -U' + \mathcal{O}(h^2)$ as before. To do this in the h^2-term, the second order term of the modified equation (2.12) is also necessary. We apply it repeatedly until all higher derivatives are eliminated. We find

$$\begin{aligned}\mathcal{L}_{\mathrm{mod},5}(x,\dot{x},h) = \frac{1}{2}|\dot{x}|^2 - U &+ \frac{h^2}{24}\Big(\big|U'\big|^2 - 2U''(\dot{x},\dot{x})\Big) \\ &+ \frac{h^4}{720}\big(3U''(U',U') - 6U^{(3)}(U',\dot{x},\dot{x}) - 2U''(U''\dot{x},\dot{x}) + U^{(4)}(\dot{x},\dot{x},\dot{x},\dot{x})\big).\end{aligned}$$

Remark. The derivatives of U should be considered as covariant, contravariant, or mixed tensors depending on the context. For example:

$$U^{(3)}(U', \dot{x}, \dot{x}) = \sum_{i,j,k} \frac{\partial^3 U}{\partial x_i \partial x_j \partial x_k} \frac{\partial U}{\partial x_i} \dot{x}_j \dot{x}_k,$$

$$U''(U''\dot{x}, \dot{x}) = \sum_{i,j} \frac{\partial^2 U}{\partial x_i \partial x_j} \left(\sum_k \frac{\partial^2 U}{\partial x_i \partial x_k} \dot{x}_k \right) \dot{x}_j.$$

The fourth truncation of the modified equation is most easily found from the meshed modified Lagrangian. We have

$$\begin{aligned}\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial \dot{x}} = -\ddot{x} - U' &+ \frac{h^2}{12}\Big(U''\ddot{x} + U^{(3)}(\dot{x},\dot{x})\Big) \\ &+ \frac{h^4}{240}\Big(-6U^{(4)}(\ddot{x},\dot{x},\dot{x}) - 3U^{(3)}(\ddot{x},\ddot{x}) \\ &\qquad - 4U^{(3)}(x^{(3)},\dot{x}) - U''x^{(4)} - U^{(5)}(\dot{x},\dot{x},\dot{x},\dot{x})\Big) + \mathcal{O}(h^6).\end{aligned}$$

Equating this to zero and solving for $\ddot{x}$ we obtain the modified equation

$$\begin{aligned}\ddot{x} = -U' &+ \frac{h^2}{12}\Big(U^{(3)}(\dot{x},\dot{x}) - U''U'\Big) \\ &+ \frac{h^4}{720}\Big(20U^{(3)}(U''\dot{x},\dot{x}) - 8U''(U''U') + 18U^{(4)}(U',\dot{x},\dot{x}) \\ &\qquad - 9U^{(3)}(U',U') - 3U^{(5)}(\dot{x},\dot{x},\dot{x},\dot{x})\Big) + \mathcal{O}(h^6).\end{aligned}$$

2.8.2. Comparison with the modified Hamiltonian

We consider the symplectic Euler discretization of a mechanical Lagrangian. Its discrete Lagrangian is given in Example 1.1(c),

$$L_{\mathrm{disc}}(x_j, x_{j+1}, h) = \frac{1}{2}\left|\frac{x_{j+1} - x_j}{h}\right|^2 - U(x_j).$$

The discrete Euler-Lagrange equation is

$$\frac{x_{j+1} - 2x_j + x_{j-1}}{h^2} = -U'(x_j).$$

Since we are dealing with a separable continuous Lagrangian, this is the same difference equation as the one obtained by the Störmer-Verlet method.

For this discretization we have

$$\mathcal{L}_{\mathrm{disc}}([x], h) = \frac{1}{2}|\dot{x}|^2 - U + \frac{h}{2}U'\dot{x} + \frac{h^2}{24}\Big(\big\langle \dot{x}\,, x^{(3)}\big\rangle - 3U'\ddot{x} - 3U''(\dot{x},\dot{x})\Big) + \mathcal{O}(h^3)$$

and

$$\mathcal{L}_{\text{mesh}}([x],h) = \frac{1}{2}|\dot{x}|^2 - U + \frac{h}{2}U'\dot{x} - \frac{h^2}{24}\left(|\ddot{x}|^2 + 2U'\ddot{x} + 2U''(\dot{x},\dot{x})\right) + \mathcal{O}(h^3).$$

The second truncation of the modified Lagrangian is

$$\mathcal{L}_{\text{mod},2}(x,\dot{x},h) = \frac{1}{2}|\dot{x}|^2 - U + \frac{h}{2}U'\dot{x} + \frac{h^2}{24}\left(|U'|^2 - 2U''(\dot{x},\dot{x})\right). \tag{2.14}$$

Note that the first order term $\frac{h}{2}U'\dot{x} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{h}{2}U\right)$ does not contribute to the Euler-Lagrange equations, hence this Lagrangian is equivalent to the corresponding modified Lagrangian (2.13) of the Störmer-Verlet method.

We compare this to the symplectic Euler discretization of the Hamiltonian system with Hamiltonian $\mathcal{H}(x,p) = \frac{1}{2}|p|^2 + U(x)$. The modified Hamiltonian for this system, truncated after order 2, is

$$\begin{aligned}\mathcal{H}_{\text{mod},2}(x,p,h) &= \mathcal{H} - \frac{h}{2}\mathcal{H}_x\mathcal{H}_p + \frac{h^2}{12}\big(\mathcal{H}_{pp}(\mathcal{H}_x,\mathcal{H}_x) + \mathcal{H}_{xx}(\mathcal{H}_p,\mathcal{H}_p) + 4\mathcal{H}_{px}(\mathcal{H}_x,\mathcal{H}_p)\big) \\ &= \frac{1}{2}|p|^2 + U - \frac{h}{2}U'p + \frac{h^2}{12}\big(|U'|^2 + U''(p,p)\big). \end{aligned} \tag{2.15}$$

Its derivation can be found for example in [35, Example IX.3.4].

Now we take the Legendre transformation of the modified Lagrangian (2.14). We have

$$p = \frac{\partial \mathcal{L}_{\text{mod},2}}{\partial \dot{x}} = \dot{x} + \frac{h}{2}U' - \frac{h^2}{6}U''\dot{x}$$

and hence

$$\dot{x} = p - \frac{h}{2}U' + \frac{h^2}{6}U''p + \mathcal{O}(h^3).$$

The Hamiltonian corresponding to $\mathcal{L}_{\text{mod},2}$ is

$$\big(\langle p\,,\dot{x}\rangle - \mathcal{L}_{\text{mod},2}\big)\Big|_{\dot{x}=p-\frac{h}{2}U'+\frac{h^2}{6}U''p+\mathcal{O}(h^3)} = \mathcal{H}_{\text{mod},2} + \mathcal{O}(h^3).$$

We see that, up to a truncation error, the modified Lagrangian (2.14) and the modified Hamiltonian (2.15) are obtained from one another by Legendre transformation.

2.8.3. A non-separable Lagrangian

Our approach is not limited to separable Lagrangians. It can be applied whenever the Lagrangian is regular.

As an example we consider an anisotropic harmonic oscillator, which has a Lagrangian of the form

$$\mathcal{L}(x,\dot{x}) = \frac{1}{2}\langle \dot{x}\,, M\dot{x}\rangle + \frac{1}{2}\langle x\,, (J_+ + J_-)\dot{x}\rangle + \frac{1}{2}\langle x\,, Ax\rangle\,,$$

where the matrices A, J_+, and M are symmetric, and J_- is antisymmetric. Its Euler-Lagrange equation is

$$-M\ddot{x} + J_-\dot{x} + Ax = 0.$$

We use the discrete Lagrangian from Example 1.1(d),

$$\begin{aligned} L_{\mathrm{disc}}(x_j, x_{j+1}, h) &= \mathcal{L}\left(x_{j+1}, \frac{x_{j+1} - x_j}{h} \right) \\ &= \frac{1}{2} \left\langle \frac{x_{j+1} - x_j}{h} , M \frac{x_{j+1} - x_j}{h} \right\rangle \\ &\quad + \frac{1}{2} \left\langle x_{j+1} , (J_+ + J_-) \frac{x_{j+1} - x_j}{h} \right\rangle + \frac{1}{2} \left\langle x_{j+1} , Ax_{j+1} \right\rangle . \end{aligned}$$

Its discrete Euler-Lagrange equation is

$$\left(M + \frac{h}{2} J_+ \right) \frac{-x_{j+1} + 2x_j - x_{j-1}}{h^2} + J_- \frac{x_{j+1} - x_{j-1}}{2h} + Ax_j = 0.$$

Note that this depends on J_+, even though the continuous Euler-Lagrange equation does not. We have

$$\begin{aligned} \mathcal{L}_{\mathrm{disc}}([x], h) &= \mathcal{L}\left(x + \frac{h}{2}\dot{x},\ \dot{x} \right) + \mathcal{O}(h^2) \\ &= \mathcal{L} + \frac{h}{2} \left\langle \frac{\partial \mathcal{L}}{\partial x} , \dot{x} \right\rangle + \mathcal{O}(h^2) \\ &= \frac{1}{2} \left\langle \dot{x} , M\dot{x} \right\rangle + \frac{1}{2} \left\langle x , (J_+ + J_-)\dot{x} \right\rangle + \frac{1}{2} \left\langle x , Ax \right\rangle \\ &\qquad + \frac{h}{2} \left\langle \frac{1}{2}(J_+ + J_-)\dot{x} + Ax , \dot{x} \right\rangle + \mathcal{O}(h^2). \end{aligned}$$

Up to first order, the meshed modified Lagrangian is equal to $\mathcal{L}_{\mathrm{disc}}$,

$$\begin{aligned} \mathcal{L}_{\mathrm{mesh}}([x], h) = \frac{1}{2} \left\langle \dot{x} , M\dot{x} \right\rangle &+ \frac{1}{2} \left\langle x , (J_+ + J_-)\dot{x} \right\rangle + \frac{1}{2} \left\langle x , Ax \right\rangle \\ &+ \frac{h}{2} \left(\frac{1}{2} \left\langle \dot{x} , J_+\dot{x} \right\rangle + \left\langle \dot{x} , Ax \right\rangle \right) + \mathcal{O}(h^2). \end{aligned}$$

Since second and higher derivatives of x do not occur in these terms, the classical modified Lagrangian $\mathcal{L}_{\mathrm{mod},1}(x, \dot{x}, h)$ is obtained by simply truncating $\mathcal{L}_{\mathrm{mesh}}([x], h)$ after the first order term. Its Euler-Lagrange equation is

$$-M\ddot{x} + J_-\dot{x} + Ax - \frac{h}{2} J_+\ddot{x} = 0.$$

Solving for $\ddot{x}$ we find the modified equation

$$\ddot{x} = M^{-1}(J_-\dot{x} + Ax) - \frac{h}{2}M^{-1}J_+M^{-1}(J_-\dot{x} + Ax) + \mathcal{O}(h^2)$$

We see that the first order term of the modified equation depends on J_+, even though the original Euler-Lagrange equation does not. This example illustrates how different but equivalent continuous Lagrangians lead to different discretizations and different modified Lagrangians. However, the leading order term of the modified equation is the same for all of them. This term is just the original Euler-Lagrange equation.

3. The case of degenerate Lagrangians linear in velocities

This chapter is an adaptation of [91]

In Chapter 2 we considered modified equations for variational integrators in the case of non-degenerate Lagrangians. We gave a construction for a modified Lagrangian, which produces the modified equation as its Euler-Lagrange equation up to a truncation error of arbitrarily high order. Although the construction was new, the claim that modified equations for variational integrators are Lagrangian was not. This follows by Legendre transformation from the well-known fact that modified equations for symplectic integrators are Hamiltonian. In this chapter we extend our previous construction to the case of degenerate Lagrangians that are linear in velocities. In this context the Legendre transformation is not invertible, so the fact that the modified equation is Lagrangian cannot be inferred in the same way from the theory of symplectic integrators.

We consider Lagrangians $\mathcal{L} : T\mathbb{R}^N \cong \mathbb{R}^{2N} \to \mathbb{R}$ of the form

$$\mathcal{L}(q, \dot{q}) = \langle \alpha(q) \, , \dot{q} \rangle - H(q), \tag{3.1}$$

where $\alpha : \mathbb{R}^N \to \mathbb{R}^N$, $H : \mathbb{R}^N \to \mathbb{R}$, and the brackets $\langle \, , \rangle$ denote the standard scalar product. Variational integrators for such Lagrangians were studied for example in [76] and [87]. An important role will be played by the matrices

$$A(q) = \alpha'(q) = \left(\frac{\partial \alpha_i(q)}{\partial q_j} \right)_{i,j=1,\ldots,N} \qquad \text{and} \qquad A_{\text{skew}}(q) = A(q)^T - A(q). \tag{3.2}$$

We assume that $A_{\text{skew}}(q)$ is invertible, then the Euler-Lagrange equation for $\mathcal{L}$ is given by

$$\dot{q} = A_{\text{skew}}(q)^{-1} H'(q)^T, \tag{3.3}$$

where $H'(q)$ is the row vector of partial derivatives of H with respect to the column vector $q = (q_1, \ldots, q_N)^T$. In contrast to the case of non-degenerate Lagrangians, this is a first order ODE.

A well-known example where a Lagrangian of the form (3.1) arises is the dynamics of point vortices in the plane. We will discuss this example in detail in Section

3.5.2. Another reason to study this class of Lagrangians is that its extension to PDEs covers several important equations. For example, the nonlinear Schrödinger equation [25, 78] is the Euler-Lagrange equation of a Lagrangian whose kinetic term is linear in the time-derivatives. Perhaps the most general application of Lagrangians that are linear in velocities is the variational formulation in phase space of mechanics, where $\mathcal{L} : TT^*\mathbb{R} \cong \mathbb{R}^{4N} \to \mathbb{R}$ is given by

$$\mathcal{L}(p, q, \dot{p}, \dot{q}) = \langle p \,, \dot{q} \rangle - H(p, q).$$

Its Euler-Lagrange equations are Hamilton's canonical equations

$$\dot{q} = \left(\frac{\partial H}{\partial p} \right)^T \qquad \text{and} \qquad \dot{p} = -\left(\frac{\partial H}{\partial q} \right)^T .$$

Note that even though $A = \left(\begin{smallmatrix} 0 & 0 \\ I & 0 \end{smallmatrix}\right)$ is singular in this case, the assumption that A_{skew} is invertible still holds. Like many concepts in classical mechanics, the variational principle in phase space dates back to the 19th century [72, Chapter XXIX]. A modern treatment can be found for example in [31, Section 8–5], and an application to geometric integration in [45].

The construction of modified Lagrangians for variational integrators from Chapter 2 carries over to the case of degenerate Lagrangians that are linear in velocities. However, there is a catch. The original differential equation is of first order for the Lagrangians considered here, but the difference equation produced by a variational integrator is of second order. Hence in this context variational integrators are two-step methods and parasitic solutions can occur. These are solutions of the difference equation that exhibit large spurious oscillations, see for example [35, Chapter XV]. To capture this behavior, an extended version of the notion of modified equations is needed.

In Section 3.1 we present two variational integrators which we will be the protagonists of all examples throughout this work. In Section 3.2 the essentials of the theory of modified equations for multi-step methods are presented. In Section 3.3 we summarize the construction of modified Lagrangians from Chapter 2 and in Section 3.4 we will present a method to extend it to the full system of modified equations. In Section 3.5 we look at some example systems.

A note on notation. As mentioned before we use the convention that the derivative of a scalar with respect to a column vector yields a row vector. In particular, this means that the derivative of the scalar product of two column vectors is calculated as

$$\langle x \,, y \rangle' = \left(x^T y \right)' = x^T y' + y^T x'.$$

Later on we will be taking higher derivatives of vectors with respect to other vectors, resulting in a zoo of tensors. We want to avoid heavy notations using indices, like

$$\sum_a A_a x^a, \qquad \sum_{a,b} B_{a,b} x^a y^b, \qquad \sum_{a,b,c} C_{a,b,c} x^a y^b z^c, \qquad \cdots . \tag{3.4}$$

If the tensor involved is symmetric, we will use the notations

$$A(x), \qquad B(x,y), \qquad C(x,y,z), \qquad \cdots$$

instead. If the tensor is of first or second order, we will often write these expressions as matrix multiplication,

$$Ax \qquad \text{and} \qquad x^T By.$$

We will also use the inner product notation $\left\langle A^T, x \right\rangle$ as an alternative to Ax. This allows us to emphasize one particular pairing in a product of more than two tensors.

Using these notations interchangeably allows us to write equations in an intuitive form and avoid the heavy notation of (3.4). The downside is that such inconsistent notation could be a source of confusion for the reader. We hope this note is enough to avoid that.

3.1. Variational integrators are 2-step methods

For Lagrangians that are linear in the velocities, the continuous Euler-Lagrange equation (3.3) is of first order. However, the discrete Euler-Lagrange equation,

$$\mathrm{D}_2 L_{\text{disc}}(q_{j-1}, q_j, h) + \mathrm{D}_1 L_{\text{disc}}(q_j, q_{j+1}, h) = 0,$$

still involves three points, i.e. it is of second order. This means that we are dealing with two-step methods. The difference equation needs one more point of initial data than the differential equation. Heuristically speaking, the initial data must be compatible with the underlying first order equation. If they are not compatible, the numerical solution will oscillate around the desired solution. For many methods these oscillations grow exponentially. Hence, even if we start with perfect initial data, rounding errors will explode over time and give rise to *parasitic* oscillations which destroy the numerical approximation. It is important that this behavior is captured by the modified equation. To do so, it has to grow into a *system of* modified equations, where the additional equations in the system encode the parasitic oscillations.

In the present context we will discuss two variational integrators in detail. Both are obtained by using a simple quadrature rule to approximate the *exact discrete Lagrangian*

$$L_{\text{exact}}(q_j, q_{j+1}, h) = \int_{jh}^{(j+1)h} \mathcal{L}(q, \dot{q}) \,\mathrm{d}t,$$

where $q(jh) = q_j$, $q((j+1)h) = q_{j+1}$, and q solves the Euler-Lagrange equations.

Midpoint rule

Using $\frac{q_{j+1}-q_j}{2}$ to approximate $\dot{q}$ and the average $\frac{q_j+q_{j+1}}{2}$ to approximate q in the integrand, we find the discrete Lagrangian

$$L_{\text{disc}}(q_j, q_{j+1}, h) = \left\langle \alpha\left(\frac{q_j + q_{j+1}}{2}\right), \frac{q_{j+1} - q_j}{h} \right\rangle - H\left(\frac{q_j + q_{j+1}}{2}\right) \tag{3.5}$$

with discrete Euler-Lagrange equation

$$\begin{aligned}&\frac{1}{2}\left(\frac{q_j - q_{j-1}}{h}\right)^T \alpha'\left(\frac{q_{j-1}+q_j}{2}\right) + \frac{1}{2}\left(\frac{q_{j+1}-q_j}{h}\right)^T \alpha'\left(\frac{q_j+q_{j+1}}{2}\right)\\ &- \frac{1}{h}\alpha\left(\frac{q_j+q_{j+1}}{2}\right)^T + \frac{1}{h}\alpha\left(\frac{q_{j-1}+q_j}{2}\right)^T - \frac{1}{2}H'\left(\frac{q_{j-1}+q_j}{2}\right) - \frac{1}{2}H'\left(\frac{q_j+q_{j+1}}{2}\right) = 0.\end{aligned}$$

In case α is linear, $\alpha(q) = Aq$, this simplifies to

$$\frac{q_{j+1}-q_{j-1}}{2h} = A_{\text{skew}}^{-1}\left(\frac{1}{2}H'\left(\frac{q_{j-1}+q_j}{2}\right)^T + \frac{1}{2}H'\left(\frac{q_j+q_{j+1}}{2}\right)^T\right),$$

where A_{skew} is defined in Equation (3.2). In the case of a non-degenerate Lagrangian this discretization would lead to a variational integrator that is equivalent to the implicit midpoint rule applied to the corresponding symplectic system. Also in the present context we will refer to it as the *midpoint rule*.

Trapezoidal rule

To obtain the second discretization we use the trapezoidal quadrature rule to approximate the exact discrete Lagrangian: we take the average of the integrand evaluated with $q = q_j$ and $q = q_{j+1}$, while still using $\frac{q_{j+1}-q_j}{2}$ to approximate the derivative $\dot{q}$. We find the discrete Lagrangian

$$L_{\text{disc}}(q_j, q_{j+1}, h) = \left\langle \frac{1}{2}\alpha(q_j) + \frac{1}{2}\alpha(q_{j+1}), \frac{q_{j+1}-q_j}{h}\right\rangle - \frac{1}{2}H(q_j) - \frac{1}{2}H(q_{j+1}) \qquad (3.6)$$

with discrete Euler-Lagrange equation

$$\left(\frac{q_{j+1}-q_{j-1}}{2h}\right)^T \alpha'(q_j) - \frac{\alpha(q_{j+1})^T - \alpha(q_{j-1})^T}{2h} - H'(q_j) = 0.$$

In case α is linear, $\alpha(q) = Aq$, this simplifies to

$$\frac{q_{j+1}-q_{j-1}}{2h} = A_{\text{skew}}^{-1} H'(q_j)^T.$$

This discretization is sometimes called the explicit midpoint rule, but we will not use this name to avoid confusion with the previous method. Instead we call this method the *trapezoidal rule*. In the case of a non-degenerate Lagrangian the trapezoidal rule would lead to the Störmer-Verlet method.

3.2. Modified equations for multistep methods

The classical theory of modified equations does not capture parasitic solutions of multistep methods. An extension of this theory for linear multistep methods was developed by Hairer [33]. (See also [35, Chapter XV].) Here we mention some of the main results, restricted to the case of two-step methods.

For a first order ODE $\dot{q} = f(q)$, consider the linear two-step method

$$\frac{a_0 q_j + a_1 q_{j+1} + a_2 q_{j+2}}{h} = b_0 f(q_j) + b_1 f(q_{j+1}) + b_2 f(q_{j+2}). \tag{3.7}$$

We call the method (3.7) *symmetric* if $a_0 = -a_2$, $a_1 = 0$, and $b_0 = b_2$. We say that it is *stable* if all roots of the polynomial $\rho(\zeta) = a_0 + a_1\zeta + a_2\zeta^2$ satisfy $|\zeta| \leq 1$ and the roots with $|\zeta| = 1$ are simple. A method is stable if and only if the numerical solution for $\dot{q} = 0$ is bounded for any initial condition. The method (3.7) is called *consistent* if $a_0 + a_1 + a_2 = 0$ and $a_2 - a_0 = b_0 + b_1 + b_2 \neq 0$, i.e. if it converges to the ODE $\dot{q} = f(q)$ as $h \to 0$. Note that the trapezoidal rule is a stable symmetric linear two-step method, but that the midpoint rule is not of the form (3.7).

The theory of modified equations for one-step methods is easily extended to yield the following.

Proposition 3.1 (Special case of [35, Theorem XV.3.1])**.** *Consider a consistent method of the form* (3.7)*. Then there exist unique h-independent functions $(f_n(q))_{n\in\mathbb{N}}$ such that for every* truncation index k*, every solution of*

$$\dot{q} = f(q) + hf_1(q) + h^2 f_2(q) + \ldots + h^k f_k(q) \tag{3.8}$$

satisfies

$$\frac{a_0 q(t) + a_1 q(t+h) + a_2 q(t+2h)}{h} = b_0 f(q(t)) + b_1 f(q(t+h)) + b_2 f(q(t+2h)) + \mathcal{O}(h^{k+1}).$$

We will call the formal differential equation

$$\dot{q} = f(q) + hf_1(q) + h^2 f_2(q) + \ldots \tag{3.9}$$

the *principal modified equation.* Up to truncation errors, every solution of the principal modified equation gives a solution of the difference equation when evaluated on a mesh $t_0 + h\mathbb{Z}$. However, not every solution of the difference equation can be obtained this way. The solutions that are missed are exactly the parasitic solutions.

Proposition 3.2 (Special case of [35, Theorem XV.3.5])**.** *Assume that the method* (3.7) *is stable, consistent, and symmetric. Then there exist h-independent functions*

$(f_n(x,y))_{n\in\mathbb{N}}$ and $(g_n(x,y))_{n\in\mathbb{N}}$ such that for every truncation index k, for every solution of

$$\dot{x} = f_0(x,y) + hf_1(x,y) + \ldots + h^k f_k(x,y) \tag{3.10}$$
$$\dot{y} = g_0(x,y) + hg_1(x,y) + \ldots + h^k g_k(x,y), \tag{3.11}$$

with $y(0) = \mathcal{O}(h)$, the function $q(t) = x(t) + e^{i\pi t/h} y(t)$ satisfies

$$\frac{a_0 q(t) + a_1 q(t+h) + a_2 q(t+2h)}{h} = b_0 f(q(t)) + b_1 f(q(t+h)) + b_2 f(q(t+2h)) + \mathcal{O}(h^{k+1}).$$

We will call the corresponding system of formal differential equations

$$\dot{x} = f_0(x,y) + hf_1(x,y) + h^2 f_2(x,y) + \ldots, \tag{3.12}$$
$$\dot{y} = g_0(x,y) + hg_1(x,y) + h^2 g_2(x,y) + \ldots, \tag{3.13}$$

the *full system of modified equations*. We call Equation (3.13) the *parasitic modified equation*.

If $y = 0$, then Equation (3.12) reduces to the principal modified equation (3.9) and Equation (3.13) reads $\dot{y} = 0$. Hence to determine whether parasitic solutions become dominant over time we need to determine the stability of the invariant manifold $y = 0$ of the system (3.12)–(3.13).

In general, even if the difference equation is not of the form (3.7), we have the following definition.

Definition 3.3. Let $\Phi(q_{j-1}, q_j, q_{j+1}, h)$ be a consistent discretization of some function $F(q, \dot{q})$.

(a) Equation (3.9) is the *principal modified equation* for the difference equation

$$\Phi(q_{j-1}, q_j, q_{j+1}, h) = 0 \tag{3.14}$$

if for every truncation index k, every solution of the truncated equation (3.8) satisfies

$$\Phi(q(t-h), q(t), q(t+h), h) = \mathcal{O}(h^{k+1}) \tag{3.15}$$

at all times t.

(b) The system of equations (3.12)–(3.13) is the *full system of modified equations* for the Equation (3.14) if for every truncation index k, for every solution (x, y) of the truncated system (3.10)–(3.11), the function $q(t) = x(t) + e^{i\pi t/h} y(t)$ satisfies Equation (3.15) at all times t.

3.3. A Lagrangian for the principal modified equation

In Chapter 2 we constructed a modified Lagrangian in the case of non-degenerate Lagrangian systems. A straightforward adaptation of this construction will give us a Lagrangian for the principal modified equation. We repeat the outline of the method, adapted to the present context of Lagrangians linear in velocities.

We identify points q_j of a numerical solution with step size h with evaluations $q(jh)$ of an interpolating curve. Using a Taylor expansion we can write the discrete Lagrangian $L_{\text{disc}} : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}_{>0} \to \mathbb{R}$ as a function of the interpolating curve q and its derivatives, all evaluated at the point $jh - \frac{h}{2}$,

$$\begin{aligned}\mathcal{L}_{\text{disc}}([q], h) &= L_{\text{disc}}\left(q - \frac{h}{2}\dot{q} + \frac{1}{2}\left(\frac{h}{2}\right)^2 \ddot{q} - \dots,\ q + \frac{h}{2}\dot{q} + \frac{1}{2}\left(\frac{h}{2}\right)^2 \ddot{q} + \dots, h \right) \\ &= L_{\text{disc}}(q_{j-1}, q_j, h),\end{aligned}$$

where the square brackets denote dependence on q and any number of its derivatives.

We want to write the discrete action

$$S_{\text{disc}}((q_j)_{j\in\{0,\dots,n\}}, h) = \sum_{j=1}^{n} h L_{\text{disc}}(q_{j-1}, q_j, h) = \sum_{j=1}^{n} h \mathcal{L}_{\text{disc}}\left(\left[q\left(jh - \tfrac{h}{2}\right)\right], h\right)$$

as an integral. This can be done using the Euler-Maclaurin formula. We obtain the *meshed modified Lagrangian*

$$\begin{aligned}\mathcal{L}_{\text{mesh}}([q(t)], h) &= \sum_{i=0}^{\infty} \left(2^{1-2i} - 1\right) \frac{h^{2i} B_{2i}}{(2i)!} \frac{\mathrm{d}^{2i}}{\mathrm{d}t^{2i}} \mathcal{L}_{\text{disc}}([q(t)], h) \\ &= \mathcal{L}_{\text{disc}}([q(t)], h) - \frac{h^2}{24} \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathcal{L}_{\text{disc}}([q(t)], h) + \frac{7h^4}{5760} \frac{\mathrm{d}^4}{\mathrm{d}t^4} \mathcal{L}_{\text{disc}}([q(t)], h) + \dots,\end{aligned}$$

where B_{2i} are the Bernoulli numbers. The power series defining $\mathcal{L}_{\text{mesh}}$ generally does not converge. Formally, it satisfies

$$S_{\text{disc}}((q(jh))_{j\in\{0,\dots,n\}}, h) = \int \mathcal{L}_{\text{mesh}}([q(t)], h)\, \mathrm{d}t.$$

In the *meshed variational problem*, non-differentiable curves are admissible as long as their singular points are consistent with the mesh, i.e. if they occur at times that are an integer multiple of h away from each other. This imposes additional conditions on critical curves, which we called *natural interior conditions*,

$$\forall \ell \geq 2 : \quad \frac{\partial \mathcal{L}}{\partial q^{(\ell)}}(t) = 0. \tag{3.16}$$

Because the action integral of $\mathcal{L}_{\text{mesh}}$ equals the discrete action, variations supported on a single mesh interval (i.e. in between consecutive points of the discrete curve) do not change the action integral of $\mathcal{L}_{\text{mesh}}$. This implies that the natural interior conditions are automatically satisfied on solutions of the Euler-Lagrange equation (for the particular Lagrangian $\mathcal{L}_{\text{mesh}}$, but not in general).

Consider the Euler-Lagrange equation of $\mathcal{L}_{\text{mesh}}$,

$$\sum_{j=0}^{\infty} (-1)^j \frac{\mathrm{d}^j}{\mathrm{d}t^j} \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial q^{(j)}} = 0.$$

Because the natural interior conditions (3.16) are automatically satisfied on critical curves, it is equivalent to

$$\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial \dot{q}} = 0.$$

This equation is of the form

$$\mathcal{E}_0(q,\dot{q}) + h\mathcal{E}_1(q,\dot{q},\ddot{q}) + h\mathcal{E}_2\big(q,\dot{q},\ddot{q},q^{(3)}\big) + \ldots = 0.$$

Assuming that the derivatives of q are bounded as $h \to 0$, we can write this as a first order differential equation, say

$$\dot{q} = F(q,h). \tag{3.17}$$

Then expressions for all higher derivatives follow by differentiation and substitution,

$$\ddot{q} = F_2(q,h), \qquad q^{(3)} = F_3(q,h), \qquad \cdots. \tag{3.18}$$

Using (3.17) and (3.18) we can replace second and higher derivatives in the meshed Lagrangian to find a first order *modified Lagrangian*,

$$\mathcal{L}_{\text{mod}}(q,\dot{q},h) = \mathcal{L}_{\text{mesh}}([q],h)\,\Big|_{q^{(j)}=F_j(q,h),\ \forall j \geq 2}.$$

Or, avoiding formal power series, a truncated modified Lagrangian

$$\mathcal{L}_{\text{mod},k}(q,\dot{q},h) = \mathcal{T}_k\bigg(\mathcal{L}_{\text{mesh}}([q],h)\,\Big|_{q^{(j)}=F_j(q,h),\ \forall j \geq 2} \bigg),$$

where $\mathcal{T}_k$ denotes truncation of the power series after order k. In general the replacements $q^{(j)} = F_j(q,h)$ would change the Euler-Lagrange equations, but because of the natural interior conditions (3.16) this is not the case here. Indeed, one finds

$$\frac{\partial \mathcal{L}_{\text{mod},k}}{\partial q} = \mathcal{T}_k\bigg(\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial q} + \sum_{\ell=2}^{\infty} \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial q^{(\ell)}} \frac{\partial F_\ell(q,h)}{\partial q} \bigg) = \mathcal{T}_k\bigg(\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial q} \bigg)$$

and

$$\frac{\partial \mathcal{L}_{\mathrm{mod},k}}{\partial \dot{q}} = \mathcal{T}_k\left(\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial \dot{q}}\right).$$

It follows that

$$\frac{\partial \mathcal{L}_{\mathrm{mod},k}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\mathrm{mod},k}}{\partial \dot{q}} = \mathcal{T}_k\left(\sum_{j=0}^{\infty}(-1)^j \frac{\mathrm{d}^j}{\mathrm{d}t^j}\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial q^{(j)}}\right),$$

so up to a truncation error, both Lagrangians yield the same Euler-Lagrange equations. Note that the natural interior conditions *do not* imply that $\partial \mathcal{L}_{\mathrm{mesh}}/\partial \dot{q} = 0$, so replacing first derivatives using $\dot{q} = F(q,h)$ is not allowed!

The details presented in Chapter 2 carry over to the degenerate case and yield the following result.

Theorem 3.4. *Consider a discrete Lagrangian that is a consistent discretization of a Lagrangian of the form (3.1). Let $\mathcal{L}$ be either $\mathcal{L}_{\mathrm{mesh}}$, or $\mathcal{L}_{\mathrm{mod},k}$, derived from this discrete Lagrangian. Solve the equation*

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} = 0$$

for $\dot{q}$, and truncate the resulting power series after order k. The result,

$$\dot{q} = f(q) + h f_1(q) + h^2 f_2(q) + \ldots + h^k f_k(q),$$

is a truncation of the principal modified equation.

Midpoint rule

From the discrete Lagrangian (3.5) we find

$$\begin{aligned}
\mathcal{L}_{\mathrm{disc}}([q],h) &= L_{\mathrm{disc}}\left(q - \frac{h}{2}\dot{q} + \frac{h^2}{8}\ddot{q} - \ldots\,,\, q + \frac{h}{2}\dot{q} + \frac{h^2}{8}\ddot{q} + \ldots, h\right) \\
&= \left\langle \alpha\left(q + \frac{h^2}{8}\ddot{q} + \ldots\right), \dot{q} + \frac{h^2}{24}q^{(3)} + \ldots \right\rangle - H\left(q + \frac{h^2}{8}\ddot{q} + \ldots\right) \\
&= \langle \alpha(q)\,,\dot{q}\rangle - H(q) + \frac{h^2}{24}\left(\left\langle \alpha(q)\,, q^{(3)}\right\rangle + 3\left\langle \alpha'(q)\ddot{q}\,,\dot{q}\right\rangle - 3H'(q)\ddot{q}\right) + \mathcal{O}(h^4).
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathcal{L}_{\mathrm{mesh}}([q],h) = \langle \alpha\,, \dot{q}\rangle - H& \\
&+ \frac{h^2}{24}\left(2\left\langle A_{\mathrm{skew}}\dot{q}\,,\ddot{q}\right\rangle - \left\langle \alpha''(\dot{q},\dot{q})\,,\dot{q}\right\rangle - 2H'\ddot{q} + H''(\dot{q},\dot{q})\right) + \mathcal{O}(h^4),
\end{aligned}$$

where the argument q of A_{skew}, α, H, and their derivatives are omitted. From this expression we obtain $\mathcal{L}_{\text{mod},3}$ by replacing all second derivatives of q using the derivative of the leading order equation,

$$\ddot{q} = \frac{\mathrm{d}}{\mathrm{d}t}\left(A_{\text{skew}}(q)^{-1} H'(q)^T\right) + \mathcal{O}(h^2).$$

In case that α is linear we have

$$\ddot{q} = A_{\text{skew}}^{-1} H'' \dot{q} + \mathcal{O}(h^2) \tag{3.19}$$

and we find the following expression for the modified Lagrangian (truncated after h^3):

$$\mathcal{L}_{\text{mod},3} = \dot{q}^T A q - H + \frac{h^2}{24}\left(-\dot{q}^T H'' \dot{q} - 2H' A_{\text{skew}}^{-1} H'' \dot{q}\right).$$

Trapezoidal rule

From the discrete Lagrangian (3.6) we find

$$\begin{aligned}
\mathcal{L}_{\text{disc}}([q],h) &= \left\langle \frac{1}{2}\alpha\left(q - \frac{h}{2}\dot{q} + \frac{h^2}{8}\ddot{q}\right) + \frac{1}{2}\alpha\left(q + \frac{h}{2}\dot{q} + \frac{h^2}{8}\ddot{q}\right), \dot{q} + \frac{h^2}{24}q^{(3)} \right\rangle \\
&\quad - \frac{1}{2}H\left(q - \frac{h}{2}\dot{q} + \frac{h^2}{8}\ddot{q}\right) - \frac{1}{2}H\left(q + \frac{h}{2}\dot{q} + \frac{h^2}{8}\ddot{q}\right) + \mathcal{O}(h^4) \\
&= \langle \alpha , \dot{q} \rangle - H \\
&\quad + \frac{h^2}{8}\left(\frac{1}{3}\left\langle \alpha , q^{(3)} \right\rangle + \left\langle \alpha' \ddot{q} , \dot{q} \right\rangle + \left\langle \alpha''(\dot{q},\dot{q}) , \dot{q} \right\rangle - H'\ddot{q} - H''(\dot{q},\dot{q})\right) + \mathcal{O}(h^4)
\end{aligned}$$

and

$$\mathcal{L}_{\text{mesh}}([q],h) = \langle \alpha , \dot{q} \rangle - H + \frac{h^2}{12}\left(\langle A_{\text{skew}}\dot{q} , \ddot{q} \rangle + \left\langle \alpha''(\dot{q},\dot{q}) , \dot{q} \right\rangle - H'\ddot{q} - H''(\dot{q},\dot{q})\right) + \mathcal{O}(h^4).$$

Again we assume that α is linear. Using Equation (3.19) we find the modified Lagrangian

$$\mathcal{L}_{\text{mod},3} = \dot{q}^T A q - H + \frac{h^2}{12}\left(-2\dot{q}^T H'' \dot{q} - H' A_{\text{skew}}^{-1} H'' \dot{q}\right).$$

3.4. The full system of modified equations

For linear symmetric two-step methods (3.7), Proposition 3.2 describes the full system of modified equations. Here we will show that for variational integrators, without assuming linearity, the full system of modified equations is of the same form. In order to construct the system of modified equations, we split the variable q_j of the discrete system in two parts,

$$q_j = x_j + (-1)^j y_j.$$

The motivation for this is that we want to use one variable, x_j, to encode the principal behavior and the other, y_j, for the parasitic behavior. This is inspired by the formula $q(t) = x(t) + e^{i\pi t/h} y(t)$ from Proposition 3.2.

3.4.1. The Lagrangian approach

A key property of this doubling of variables is that the extended system is still variational.

Proposition 3.5. *The discrete curve $(x_j, y_j)_{j\in\{0,\ldots,n\}}$ is critical for*

$$\widehat{L}(x_j, y_j, x_{j+1}, y_{j+1}, h) = \frac{1}{2} L(x_j + y_j, x_{j+1} - y_{j+1}, h) + \frac{1}{2} L(x_j - y_j, x_{j+1} + y_{j+1}, h),$$

if and only if the discrete curves $(q_j^+)_{j\in\{0,\ldots,n\}}$ and $(q_j^-)_{j\in\{0,\ldots,n\}}$, defined by $q_j^\pm = x_j \pm (-1)^j y_j$, are critical for $L(q_j, q_{j+1}, h)$.

Proof. The discrete Euler-Lagrange equations for $\widehat{L}(x_j, y_j, x_{j+1}, y_{j+1}, h)$ are

$$\begin{aligned} &\frac{1}{2}\mathrm{D}_2 L(x_{j-1} + y_{j-1}, x_j - y_j, h) + \frac{1}{2}\mathrm{D}_2 L(x_{j-1} - y_{j-1}, x_j + y_j, h) \\ &+\frac{1}{2}\mathrm{D}_1 L(x_j + y_j, x_{j+1} - y_{j+1}, h) + \frac{1}{2}\mathrm{D}_1 L(x_j - y_j, x_{j+1} + y_{j+1}, h) = 0 \end{aligned}$$

and

$$\begin{aligned} &-\frac{1}{2}\mathrm{D}_2 L(x_{j-1} + y_{j-1}, x_j - y_j, h) + \frac{1}{2}\mathrm{D}_2 L(x_{j-1} - y_{j-1}, x_j + y_j, h) \\ &+\frac{1}{2}\mathrm{D}_1 L(x_j + y_j, x_{j+1} - y_{j+1}, h) - \frac{1}{2}\mathrm{D}_1 L(x_j - y_j, x_{j+1} + y_{j+1}, h) = 0. \end{aligned}$$

Taking the sum resp. the difference of these equations we find

$$\begin{aligned} \mathrm{D}_2 L(x_{j-1} - y_{j-1}, x_j + y_j, h) + \mathrm{D}_1 L(x_j + y_j, x_{j+1} - y_{j+1}, h) = 0, \\ \mathrm{D}_2 L(x_{j-1} + y_{j-1}, x_j - y_j, h) + \mathrm{D}_1 L(x_j - y_j, x_{j+1} + y_{j+1}, h) = 0. \end{aligned}$$

Depending on the parity of j, either the first or the second of those equations is

$$\mathrm{D}_2 L(q_{j-1}^+, q_j^+, h) + \mathrm{D}_1 L(q_j^+, q_{j+1}^+, h) = 0.$$

The other one is

$$\mathrm{D}_2 L(q_{j-1}^-, q_j^-, h) + \mathrm{D}_1 L(q_j^-, q_{j+1}^-, h) = 0.$$

Hence $(x_j, y_j)_{j\in\{0,\ldots,n\}}$ satisfies the Euler-Lagrange equations for $\widehat{L}(x_j, y_j, x_{j+1}, y_{j+1}, h)$ if and only if $(q_j^+)_{j\in\{0,\ldots,n\}}$ and $(q_j^-)_{j\in\{0,\ldots,n\}}$ satisfy the Euler-Lagrange equation for $L(q_j, q_{j+1}, h)$. □

Theorem 3.6. *Let*

$$\begin{aligned}\dot{x} &= f_0(x,y) + hf_1(x,y) + \ldots + h^k f_k(x,y) \\ \dot{y} &= g_0(x,y) + hg_1(x,y) + \ldots + h^k g_k(x,y),\end{aligned} \tag{3.20}$$

be the k-th truncation of the principal modified equation for the difference equation described by the discrete Lagrangian $\widehat{L}$ from Proposition 3.5. Then (3.20) is the k-th truncation of the full system of modified equations for the variational integrator described by L.

Proof. Let $(x(t), y(t))$ be a solution of the system (3.20). By definition of the principal modified equation, the discrete curve

$$\big(x(jh), y(jh)\big)_{j \in \{0,\ldots,n\}}$$

satisfies the discrete Euler-Lagrange equations for $\widehat{L}$ up to a truncation error. Hence, by Proposition 3.5, the discrete curve

$$\big(x(jh) + (-1)^j y(jh)\big)_{j \in \{0,\ldots,n\}}$$

satisfies the discrete Euler-Lagrange equations for L up to a truncation error. This is exactly the defining property of the system of modified equations. □

Corollary 3.7. *Up to a truncation error of arbitrarily high order, the full system of modified equations (3.20) for a variational integrator is Lagrangian.*

Let us illustrate this construction by applying it to the two methods and, for comparison, to a non-degenerate Lagrangian.

Midpoint rule

We have

$$\begin{aligned}\widehat{L}_{\mathrm{disc}}(x_j, y_j, x_{j+1}, y_{j+1}, h) = {} & \frac{1}{2}\left\langle \alpha\left(\frac{x_j + y_j + x_{j+1} - y_{j+1}}{2}\right), \frac{x_{j+1} - y_{j+1} - x_j - y_j}{h} \right\rangle \\ & + \frac{1}{2}\left\langle \alpha\left(\frac{x_j - y_j + x_{j+1} + y_{j+1}}{2}\right), \frac{x_{j+1} + y_{j+1} - x_j + y_j}{h} \right\rangle \\ & - \frac{1}{2} H\left(\frac{x_j + y_j + x_{j+1} - y_{j+1}}{2}\right) \\ & - \frac{1}{2} H\left(\frac{x_j - y_j + x_{j+1} + y_{j+1}}{2}\right).\end{aligned}$$

Hence

$$\begin{aligned}\widehat{\mathcal{L}}_{\text{disc}}([x,y],h) &= \frac{1}{2}\left\langle \alpha\!\left(x - \frac{h}{2}\dot{y}\right), \dot{x} - \frac{2}{h}y \right\rangle + \frac{1}{2}\left\langle \alpha\!\left(x + \frac{h}{2}\dot{y}\right), \dot{x} + \frac{2}{h}y \right\rangle - H(x) + \mathcal{O}(h) \\ &= \langle \alpha(x)\,, \dot{x}\rangle + \langle \alpha'(x)\dot{y}\,, y\rangle - H(x) + \mathcal{O}(h).\end{aligned}$$

This is also the leading order term of the modified Lagrangian, $\widehat{\mathcal{L}}_{\text{mod},0}(x,y,\dot{x},\dot{y},h)$. If α is linear, its Euler-Lagrange equations are

$$\begin{aligned}\dot{x} &= A_{\text{skew}}^{-1} H'(x)^T + \mathcal{O}(h), \\ \dot{y} &= 0 + \mathcal{O}(h).\end{aligned} \tag{3.21}$$

Since y is constant in leading order, we need to look at higher order terms to determine whether parasitic solutions occur.

Proposition 3.8. *If α is linear, the higher order terms of the system of modified equations (3.21) do not contain y and are of even combined degree in the derivatives $\dot{y}, \ddot{y}, \ldots$.*

Proof. Assume for now that $H = 0$. Then the discrete Lagrangian is

$$\widehat{L}_{\text{disc}}(x_j, y_j, x_{j+1}, y_{j+1}, h) = \frac{1}{2}\langle A(x_j + x_{j+1})\,, x_{j+1} - x_j\rangle + \frac{1}{2}\langle A(y_{j+1} - y_j)\,, y_j + y_{j+1}\rangle\,,$$

leading to the Euler-Lagrange equations

$$x_{j+1} = x_{j-1} \qquad \text{and} \qquad y_{j+1} = y_{j-1}.$$

Hence in this case the system of modified equations is

$$\dot{x} = 0 \qquad \text{and} \qquad \dot{y} = 0.$$

This implies that also for nonvanishing H, the kinetic term of the Lagrangian does not contribute to the higher order terms of the system of modifies equations. It is easy to see that the series expansion of the H-terms in $\widehat{L}$,

$$-\frac{1}{2}H\!\left(x - \frac{h}{2}\dot{y} + \frac{h^2}{8}\ddot{x} - \ldots\right) - \frac{1}{2}H\!\left(x + \frac{h}{2}\dot{y} + \frac{h^2}{8}\ddot{x} + \ldots\right),$$

contains derivatives of y but not y itself. Furthermore, since the Lagrangian $\widehat{L}$ is invariant under the transformation $y \mapsto -y$, these derivatives must occur with an even combined degree in the expansion. □

In other words, no higher order terms of the modified Lagrangian contain y itself, and those terms that contain derivatives of y are at least quadratic in the derivatives of y. This implies that the Euler-Lagrange equation with respect to y contains a derivative of y in every higher order term. Since we already know that $\dot{y} = 0 + \mathcal{O}(h)$, we can now recursively deduce that $\dot{y} = 0 + \mathcal{O}(h^k)$ for any k.

In conclusion, the parasitic modified equation is $\dot{y} = 0$ to any order of accuracy. Hence if the initialization of the discrete system is close to a solution of the principal modified equation, then the discrete solution will remain close to it.

Trapezoidal rule

We have

$$\begin{aligned}
\widehat{L}_{\text{disc}}(x_j, y_j, x_{j+1}, y_{j+1}, h) = \frac{1}{4} &\left\langle \alpha(x_j + y_j) + \alpha(x_{j+1} - y_{j+1}) \,,\, \frac{x_{j+1} - y_{j+1} - x_j - y_j}{h} \right\rangle \\
+ \frac{1}{4} &\left\langle \alpha(x_j - y_j) + \alpha(x_{j+1} + y_{j+1}) \,,\, \frac{x_{j+1} + y_{j+1} - x_j + y_j}{h} \right\rangle \\
- \frac{1}{4}&H(x_j + y_j) - \frac{1}{4}H(x_{j+1} - y_{j+1}) \\
- \frac{1}{4}&H(x_j - y_j) - \frac{1}{4}H(x_{j+1} + y_{j+1}).
\end{aligned}$$

Hence

$$\begin{aligned}
\widehat{\mathcal{L}}_{\text{disc}}([x,y],h) &= \frac{1}{4} \left\langle \alpha\left(x + y - \frac{h}{2}\dot{x} - \frac{h}{2}\dot{y} \right) + \alpha\left(x - y + \frac{h}{2}\dot{x} - \frac{h}{2}\dot{y} \right), \dot{x} - \frac{2}{h}y \right\rangle \\
&\quad + \frac{1}{4} \left\langle \alpha\left(x - y - \frac{h}{2}\dot{x} + \frac{h}{2}\dot{y} \right) + \alpha\left(x + y + \frac{h}{2}\dot{x} + \frac{h}{2}\dot{y} \right), \dot{x} + \frac{2}{h}y \right\rangle \\
&\quad - \frac{1}{2}H(x+y) - \frac{1}{2}H(x-y) + \mathcal{O}(h) \\
&= \frac{1}{4} \left\langle \alpha(x{+}y) - \frac{h}{2}\alpha'(x{+}y)(\dot{x} + \dot{y}) + \alpha(x{-}y) + \frac{h}{2}\alpha'(x{-}y)(\dot{x} - \dot{y}) \,,\, \dot{x} - \frac{2}{h}y \right\rangle \\
&\quad + \frac{1}{4} \left\langle \alpha(x{-}y) - \frac{h}{2}\alpha'(x{-}y)(\dot{x} - \dot{y}) + \alpha(x{+}y) + \frac{h}{2}\alpha'(x{+}y)(\dot{x} + \dot{y}) \,,\, \dot{x} + \frac{2}{h}y \right\rangle \\
&\quad - \frac{1}{2}H(x+y) - \frac{1}{2}H(x-y) + \mathcal{O}(h) \\
&= \frac{1}{2} \left\langle \alpha(x+y) \,,\, \dot{x} \right\rangle + \frac{1}{2} \left\langle \alpha(x-y) \,,\, \dot{x} \right\rangle + \frac{1}{2} \left\langle \alpha'(x+y)(\dot{x} + \dot{y}) \,,\, y \right\rangle \\
&\quad - \frac{1}{2} \left\langle \alpha'(x-y)(\dot{x} - \dot{y}) \,,\, y \right\rangle - \frac{1}{2}H(x+y) - \frac{1}{2}H(x-y) + \mathcal{O}(h).
\end{aligned}$$

This is also the leading order term of the modified Lagrangian, $\widehat{\mathcal{L}}_{\text{mod},0}(x, y, \dot{x}, \dot{y}, h)$. If α is linear, $\alpha(q) = Aq$, then we find

$$\widehat{\mathcal{L}}_{\text{mod}}(x, y, \dot{x}, \dot{y}, h) = \langle Ax \,,\, \dot{x} \rangle + \langle A\dot{y} \,,\, y \rangle - \frac{1}{2}H(x+y) - \frac{1}{2}H(x-y) + \mathcal{O}(h).$$

Its Euler-Lagrange equations are

$$\begin{aligned}
\dot{x} &= A_{\text{skew}}^{-1} \left(\frac{1}{2}H'(x+y)^T + \frac{1}{2}H'(x-y)^T \right) + \mathcal{O}(h), \\
\dot{y} &= A_{\text{skew}}^{-1} \left(-\frac{1}{2}H'(x+y)^T + \frac{1}{2}H'(x-y)^T \right) + \mathcal{O}(h).
\end{aligned}$$

We linearize the second equation around $y = 0$ and find

$$\dot{y} = -A_{\text{skew}}^{-1} H''(x) y + \mathcal{O}(|y|^2 + h). \tag{3.22}$$

Heuristically we would expect exponentially growing parasitic solutions if the matrix $-A_{\text{skew}}^{-1} H''(x)$ has at least one eigenvalue with positive real part. However, since this matrix is not constant it is difficult to give a general condition for the occurrence of exponentially growing parasites. This has to be investigated on a case-by-case basis.

A non-degenerate Lagrangian

For comparison, consider a discretization of a non-degenerate mechanical Lagrangian,

$$L_{\text{disc}}(q_j, q_{j+1}, h) = \frac{1}{2}\left(\frac{q_{j+1} - q_j}{h}\right)^2 - U(q_j, q_{j+1}).$$

In this case, both the continuous and the discrete system are of second order. They both need two points of initial data, so unlike in the case of linear Lagrangians, there is no additional initial datum in the discrete case that could cause parasitic oscillations. Nevertheless, we could also double the variables in the current system, because Proposition 3.2 makes no assumption on the form of the Lagrangian. We find

$$\begin{aligned}\widehat{L}_{\text{disc}}(x_j, y_j, x_{j+1}, y_{j+1}, h) = \frac{1}{4}\left(\frac{x_{j+1} - y_{j+1} - x_j - y_j}{h}\right)^2 + \frac{1}{4}\left(\frac{x_{j+1} + y_{j+1} - x_j + y_j}{h}\right)^2 \\ - \frac{1}{2} U(x_j + y_j, x_{j+1} - y_{j+1}) - \frac{1}{2} U(x_j - y_j, x_{j+1} + y_{j+1}),\end{aligned}$$

hence

$$\widehat{\mathcal{L}}_{\text{disc}}([x, y], h) = \frac{1}{4}\left(\dot{x} - \frac{2}{h} y\right)^2 + \frac{1}{4}\left(\dot{x} + \frac{2}{h} y\right)^2 + \mathcal{O}(1) = \frac{2}{h^2} y^2 + \mathcal{O}(1).$$

and also $\widehat{\mathcal{L}}_{\text{mod},-1}([x, y], h) = \frac{2}{h^2} y^2$. In the leading order we find the Euler-Lagrange equation $y = 0 + \mathcal{O}(h^2)$. Similar to the case of the midpoint rule, we can iteratively increase the order of this equation to find that $y = 0$ to any order. In other words, the parasitic variable is identically zero, leaving only the principal variable x. Doubling the dimension was pointless in this case.

3.4.2. The direct approach

If one is not interested in the Lagrangian structure of the problem, it might be preferable to calculate the modified equation directly from the difference equation, ignoring the Lagrangian. We demonstrate this method in the case of linear α for our two integrators. For a more detailed discussion we refer to [33].

Midpoint rule

In the difference equation

$$\frac{q_{j+1}-q_{j-1}}{2h} = A_{\text{skew}}^{-1}\left(\frac{1}{2}H'\!\left(\frac{q_{j-1}+q_j}{2}\right)^T + \frac{1}{2}H'\!\left(\frac{q_j+q_{j+1}}{2}\right)^T\right)$$

we set $q_j = x(t) + (-1)^j y(t)$ and

$$\begin{aligned} q_{j\pm 1} &= x(t\pm h) + (-1)^{j\pm 1} y(t\pm h) \\ &= \left(x(t) \pm h\dot{x}(t) + \frac{h^2}{2}\ddot{x}(t) \pm \dots\right) - (-1)^j\left(y(t) \pm h\dot{y}(t) + \frac{h^2}{2}\ddot{y}(t) \pm \dots\right). \end{aligned}$$

It follows that

$$\frac{q_{j+1}-q_{j-1}}{2h} = \dot{x}(t) - (-1)^j\dot{y}(t) + \mathcal{O}(h^2)$$

and

$$\begin{aligned} H'\!\left(\frac{q_j+q_{j\pm 1}}{2}\right) &= H'\!\left(x \pm \frac{h}{2}\dot{x} \pm \frac{h}{2}(-1)^{j+1}\dot{y}\right) + \mathcal{O}(h^2) \\ &= H'(x) \pm \frac{h}{2}H''(x)\left(\dot{x} + (-1)^{j+1}\dot{y}\right) + \mathcal{O}(h^2). \end{aligned}$$

Hence

$$\begin{aligned} \dot{x} - (-1)^j\dot{y} &= A_{\text{skew}}^{-1}\left(H'(x) + \frac{h}{4}H''(x)\left(\dot{x} + (-1)^{j+1}\dot{y}\right) - \frac{h}{4}H''(x)\left(\dot{x} + (-1)^{j+1}\dot{y}\right)\right)^T \\ &\qquad + \mathcal{O}(h^2) \\ &= A_{\text{skew}}^{-1}H'(x)^T + \mathcal{O}(h^2). \end{aligned}$$

Separating the alternating terms from the rest, we find

$$\begin{aligned} \dot{x} &= A_{\text{skew}}^{-1}H'(x)^T + \mathcal{O}(h^2) \\ \dot{y} &= 0 + \mathcal{O}(h^2). \end{aligned}$$

Unsurprisingly, we find the same system of modified equations as with the Lagrangian method.

Trapezoidal rule

Now we consider the difference equation

$$\frac{q_{j+1}-q_{j-1}}{2h} = A_{\text{skew}}^{-1}H'(q_j)^T$$

and make the same identifications as before. We find

$$\begin{aligned}\dot{x} + -(-1)^j \dot{y} &= A_{\text{skew}}^{-1} H'(x + (-1)^j y)^T + \mathcal{O}(h^2) \\ &= A_{\text{skew}}^{-1} H'(x)^T + (-1)^j A_{\text{skew}}^{-1} H''(x) y + \mathcal{O}(|y|^2 + h^2).\end{aligned}$$

If we assume that $y = \mathcal{O}(h)$, then in the leading order the system of modified equations is

$$\begin{aligned}\dot{x} &= A_{\text{skew}}^{-1} H'(x)^T + \mathcal{O}(h^2) \\ \dot{y} &= -A_{\text{skew}}^{-1} H''(x) y + \mathcal{O}(h^2).\end{aligned}$$

3.5. Examples

To illustrate the theory above, we apply our two integrators to two examples. Since the calculations tend to be quite long in real-world problems, we start with a minimal toy problem. After that, we discuss the dynamics of point vortices in the plane.

3.5.1. Toy Problem

Consider the Lagrangian

$$\mathcal{L}(p, q, \dot{p}, \dot{q}) = \frac{1}{2}(p\dot{q} - q\dot{p}) - U(p) - V(q)$$

on $T\mathbb{R}^2$. Its Euler-Lagrange equations are

$$\dot{p} = -V'(q) \qquad \text{and} \qquad \dot{q} = U'(p).$$

As a concrete example, the choice $V(q) = -\cos(q)$ and $U(p) = \frac{1}{2}p^2$ describes the pendulum.

Midpoint rule

We have

$$\begin{aligned}L_{\text{disc}}(p_j, q_j, p_{j+1}, q_{j+1}, h) = \frac{1}{2}\left(\frac{p_j + p_{j+1}}{2} \frac{q_{j+1} - q_j}{h} - \frac{q_j + q_{j+1}}{2} \frac{p_{j+1} - p_j}{h} \right) \\ - U\left(\frac{p_j + p_{j+1}}{2} \right) - V\left(\frac{q_j + q_{j+1}}{2} \right).\end{aligned}$$

This corresponds to the following system of difference equations:

$$\begin{aligned}\frac{q_{j+1} - q_{j-1}}{2h} &= \frac{1}{2} U'\left(\frac{p_{j-1} + p_j}{2} \right) + \frac{1}{2} U'\left(\frac{p_j + p_{j+1}}{2} \right), \\ \frac{p_{j+1} - p_{j-1}}{2h} &= -\frac{1}{2} V'\left(\frac{q_{j-1} + q_j}{2} \right) - \frac{1}{2} V'\left(\frac{q_j + q_{j+1}}{2} \right).\end{aligned}$$

By Taylor expansion we obtain

$$\begin{aligned}\mathcal{L}_{\text{disc}}([p,q],h) &= \mathcal{L}(p,q,\dot{p},\dot{q}) \\ &\qquad + \frac{h^2}{24}\left(\frac{1}{2}\left(pq^{(3)} + 3\ddot{p}\dot{q} - 3\dot{p}\ddot{q} - p^{(3)}q\right) - 3U'\ddot{p} - 3V'\ddot{q}\right) + \mathcal{O}(h^4).\end{aligned}$$

It follows that

$$\mathcal{L}_{\text{mesh}}([p,q],h) = \mathcal{L}(p,q,\dot{p},\dot{q}) + \frac{h^2}{24}\left(2\ddot{p}\dot{q} - 2\dot{p}\ddot{q} - 2U'\ddot{p} + U''\dot{p}^2 - 2V'\ddot{q} + V''\dot{q}^2\right) + \mathcal{O}(h^4).$$

Hence the modified equations are

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial p} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial \dot{p}} = \dot{q} - U' + \frac{h^2}{24}\left(2q^{(3)} - U^{(3)}\dot{p}^2 - 4U''\ddot{p}\right) + \mathcal{O}(h^4),\\ 0 &= \frac{\partial \mathcal{L}_{\text{mesh}}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial \dot{q}} = -\dot{p} - V' + \frac{h^2}{24}\left(-2p^{(3)} - V^{(3)}\dot{q}^2 - 4V''\ddot{q}\right) + \mathcal{O}(h^4).\end{aligned}$$

These imply that

$$\begin{aligned}\dot{q} &= U' - \frac{h^2}{24}\left(U^{(3)}V'^2 + 2U''V''U'\right) + \mathcal{O}(h^4),\\ \dot{p} &= -V' + \frac{h^2}{24}\left(V^{(3)}U'^2 + 2V''U''V'\right) + \mathcal{O}(h^4).\end{aligned}$$

Eliminating higher derivatives in $\mathcal{L}_{\text{mesh}}$ we find

$$\mathcal{L}_{\text{mod}}(p,q,\dot{p},\dot{q},h) = \mathcal{L}(p,q,\dot{p},\dot{q}) + \frac{h^2}{24}\left(-V''\dot{q}^2 - U''\dot{p}^2 + 2U'V''\dot{q} - 2V'U''\dot{p}\right) + \mathcal{O}(h^4).$$

As discussed in the previous section we do not expect parasitic solutions with this method (see Figure 3.1).

Trapezoidal rule

We have

$$\begin{aligned}L_{\text{disc}}(p_j,q_j,p_{j+1},q_{j+1},h) &= \frac{1}{2}\left(\frac{p_j + p_{j+1}}{2}\frac{q_{j+1}-q_j}{h} - \frac{q_j+q_{j+1}}{2}\frac{p_{j+1}-p_j}{h}\right)\\ &\qquad - \frac{1}{2}U(p_j) - \frac{1}{2}U(p_{j+1}) - \frac{1}{2}V(q_j) - \frac{1}{2}V(q_{j+1}).\end{aligned}$$

The corresponding discrete Euler-Lagrange equations are

$$\frac{q_{j+1}-q_{j-1}}{2h} = U'(p_j), \qquad \frac{p_{j+1}-p_{j-1}}{2h} = -V'(q_j).$$

By Taylor expansion we obtain

$$
\begin{aligned}
\mathcal{L}_{\mathrm{disc}}([p,q],h) = \mathcal{L}(p,q,\dot{p},\dot{q}) \\
+ \frac{h^2}{24}\left(\frac{1}{2}\left(pq^{(3)} + 3\ddot{p}\dot{q} - 3\dot{p}\ddot{q} - p^{(3)}q\right) - 3U'\ddot{p} - 3U''\dot{p}^2 - 3V'\ddot{q} - 3V''\dot{q}^2\right) \\
+ \mathcal{O}(h^4).
\end{aligned}
$$

It follows that

$$
\mathcal{L}_{\mathrm{mesh}}([p,q],h) = \mathcal{L}(p,q,\dot{p},\dot{q}) + \frac{h^2}{12}\left(\ddot{p}\dot{q} - \dot{p}\ddot{q} - U'\ddot{p} - U''\dot{p}^2 - V'\ddot{q} - V''\dot{q}^2\right) + \mathcal{O}(h^4).
$$

Hence the modified equations are

$$
\begin{aligned}
0 &= \frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial p} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial \dot{p}} = \dot{q} - U' + \frac{h^2}{12}\left(q^{(3)} + U^{(3)}\dot{p}^2 + U''\ddot{p}\right) + \mathcal{O}(h^4), \\
0 &= \frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\mathrm{mesh}}}{\partial \dot{q}} = -\dot{p} - V' + \frac{h^2}{12}\left(-p^{(3)} + V^{(3)}\dot{q}^2 + V''\ddot{q}\right) + \mathcal{O}(h^4).
\end{aligned}
$$

These imply that

$$
\begin{aligned}
\dot{q} &= U' - \frac{h^2}{6}\left(U^{(3)}V'^2 - U''V''U'\right) + \mathcal{O}(h^4), \\
\dot{p} &= -V' + \frac{h^2}{6}\left(V^{(3)}U'^2 - V''U''V'\right) + \mathcal{O}(h^4).
\end{aligned}
$$

Eliminating higher derivatives in $\mathcal{L}_{\mathrm{mesh}}$ we find

$$
\mathcal{L}_{\mathrm{mod}}(p,q,\dot{p},\dot{q},h) = \mathcal{L}(p,q,\dot{p},\dot{q}) + \frac{h^2}{12}\left(-2V''\dot{q}^2 - 2U''\dot{p}^2 + U'V''\dot{q} - V'U''\dot{p}\right) + \mathcal{O}(h^4).
$$

For the pendulum, $V(q) = -\cos(q)$ and $U(p) = \frac{1}{2}p^2$, we have

$$
A = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{and} \qquad H'' = \begin{pmatrix} U''(p) & 0 \\ 0 & V''(q) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \cos(q) \end{pmatrix},
$$

hence the matrix in Equation (3.22) is

$$
-A_{\mathrm{skew}}^{-1}H'' = \begin{pmatrix} 0 & -\cos(q) \\ 1 & 0 \end{pmatrix}.
$$

This matrix has a pair of real eigenvalues if $\cos(q) < 0$ and a pair of purely imaginary eigenvalues if $\cos(q) > 0$. This suggests (but doesn't prove; q is not constant) that exponentially growing parasites occur in the regions where $\cos(q) < 0$.

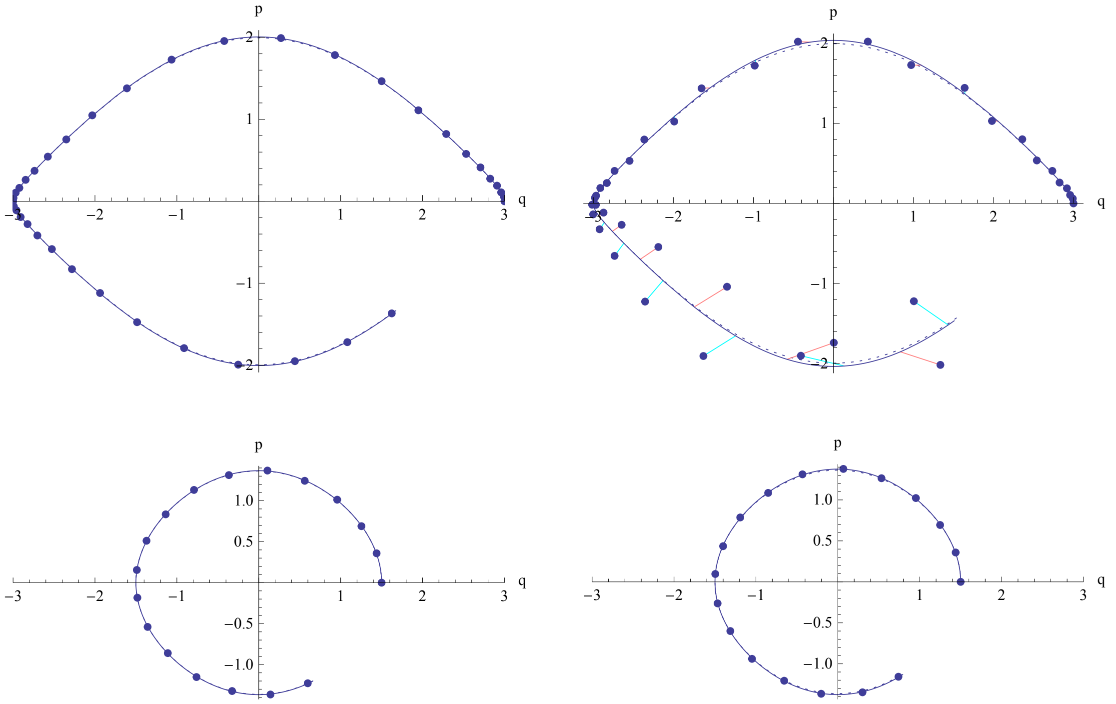


Figure 3.1. *Pendulum with midpoint rule (left) and trapezoidal rule (right), both with step size $h = 0.35$ and initial point either $(3, 0)$ (top) or $(1.5, 0)$ (bottom).*

Dashed curve: exact solution.

Bullets: discrete solution.

Solid curve: solution of the principal modified equation, truncated after the second order term.

Line segments: visualization of parasitic oscillations

In the top right image of Figure 3.1 one clearly observes parasitic solutions for this method. Note the growth of the parasites, is mostly concentrated around $|q| \approx 3$, where $\cos(q)$ is negative and near its minimum. In the region where $|q| < \frac{\pi}{2}$ there is no noticeable growth in the amplitude of the oscillations. Instead we observe a rotation in the direction of the oscillations, as expected when the eigenvalues are purely imaginary. This is visualized in Figure 3.1 by line segments connecting the points of the discrete solution with the corresponding points on the solution of the principal modified equation.

When the initial conditions are chosen such that q remains in the stable region $|q| < \frac{\pi}{2}$ no parasites are observed (bottom right image of Figure 3.1), even if the simulation is continued for many periods (not pictured).

3.5.2. Point vortices

Our second example involves vortices on a planar surface. If all vorticity is contained in a finite number of points, then the movement of those points is described by first order ODEs [59, 76]. To be precise, the dynamics of N point vortices in the (complex) plane is described by the Lagrangian

$$\begin{aligned}\mathcal{L}(z,\dot z) &= \sum_{j=1}^{N}\Gamma_j \operatorname{Im}(\overline{z}_j\dot z_j) - \frac{1}{\pi}\sum_{j=1}^{N}\sum_{k=1}^{j-1}\Gamma_j\Gamma_k\log\left|z_j - z_k\right| \\ &= \sum_{j=1}^{N}\frac{i\Gamma_j}{2}(z_j\dot{\overline{z}}_j - \overline{z}_j\dot z_j) - \frac{1}{2\pi}\sum_{j=1}^{N}\sum_{k=1}^{j-1}\Gamma_j\Gamma_k\log\big((z_j - z_k)(\overline{z}_j - \overline{z}_k)\big),\end{aligned}$$

where z_j and Γ_j are the position and circulation of the j-th vortex, and the bar denotes the complex conjugate. The equations of motion are

$$\dot z_j = \frac{i}{2\pi}\sum_{k\neq j}\frac{\Gamma_k}{\overline{z}_j - \overline{z}_k} \qquad \text{for } j = 1,\dots,N.$$

It follows that

$$\ddot z_j = \frac{i}{2\pi}\sum_{k\neq j}\frac{-\Gamma_k}{(\overline{z}_j - \overline{z}_k)^2}\left(\dot{\overline{z}}_j - \dot{\overline{z}}_k\right)$$

Midpoint rule

We have

$$\begin{aligned}&\mathcal{L}_{\mathrm{disc}}([z],h) = \mathcal{L}(z,\dot z) \\ &\quad + \frac{h^2}{24}\left[\sum_{j=1}^{N}\Gamma_j\operatorname{Im}\Big(3\dot z_j\ddot{\overline{z}}_j + z_j^{(3)}\overline{z}_j\Big) - \sum_{j=1}^{N}\sum_{k=1}^{j-1}\frac{3\Gamma_j\Gamma_k}{\pi}\operatorname{Re}\left(\frac{\ddot z_j - \ddot z_k}{z_j - z_k}\right)\right] + \mathcal{O}(h^4)\end{aligned}$$

and

$$\begin{aligned}&\mathcal{L}_{\mathrm{mesh}}([z],h) = \mathcal{L}(z,\dot z) \\ &\quad + \frac{h^2}{24}\left[4\sum_{j=1}^{N}\Gamma_j\operatorname{Im}\big(\dot z_j\ddot{\overline{z}}_j\big) - \sum_{j=1}^{N}\sum_{k=1}^{j-1}\frac{\Gamma_j\Gamma_k}{\pi}\operatorname{Re}\left(2\frac{\ddot z_j - \ddot z_k}{z_j - z_k} + \left(\frac{\dot z_j - \dot z_k}{z_j - z_k}\right)^2\right)\right] + \mathcal{O}(h^4).\end{aligned} \tag{3.23}$$

To obtain the modified Lagrangian we evaluate the second derivatives in $\mathcal{L}_{\text{mesh}}$ using the modified equation. We find

$$\begin{aligned}
\sum_{j=1}^{N} \Gamma_j \operatorname{Im}\big(\dot{z}_j \ddot{\overline{z}}_j\big) &= \sum_{j=1}^{N} \sum_{k \neq j} \Gamma_j \operatorname{Im}\left(\dot{z}_j \frac{i}{2\pi} \frac{\Gamma_k}{(z_j - z_k)^2} \big(\dot{z}_j - \dot{z}_k \big) \right) + \mathcal{O}(h^2) \\
&= \sum_{j=1}^{N} \sum_{k \neq j} \frac{\Gamma_j \Gamma_k}{2\pi} \operatorname{Re}\left(\dot{z}_j \frac{\dot{z}_j - \dot{z}_k}{(z_j - z_k)^2} \right) + \mathcal{O}(h^2) \\
&= \sum_{j=1}^{N} \sum_{k \neq j} \frac{\Gamma_j \Gamma_k}{4\pi} \operatorname{Re}\left(\frac{\big(\dot{z}_j - \dot{z}_k\big)^2}{(z_j - z_k)^2} \right) + \mathcal{O}(h^2)
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j=1}^{N} \sum_{k=1}^{j-1} \Gamma_j \Gamma_k \operatorname{Re}\left(2 \frac{\ddot{z}_j - \ddot{z}_k}{z_j - z_k} \right) &= 2 \sum_{j=1}^{N} \sum_{k \neq j} \Gamma_j \Gamma_k \operatorname{Re}\left(\frac{\ddot{z}_j}{z_j - z_k} \right) + \mathcal{O}(h^2) \\
&= 2 \sum_{j=1}^{N} \sum_{k \neq j} \sum_{\ell \neq j} \Gamma_j \Gamma_k \operatorname{Re}\left(\frac{-i}{2\pi} \frac{\Gamma_\ell \big(\dot{\overline{z}}_j - \dot{\overline{z}}_\ell \big)}{(z_j - z_k) \big(\overline{z}_j - \overline{z}_\ell \big)^2} \right) + \mathcal{O}(h^2) \\
&= \frac{1}{\pi} \sum_{j=1}^{N} \sum_{k \neq j} \sum_{\ell \neq j} \Gamma_j \Gamma_k \Gamma_\ell \operatorname{Im}\left(\frac{\big(\dot{\overline{z}}_j - \dot{\overline{z}}_\ell \big)}{(z_j - z_k) \big(\overline{z}_j - \overline{z}_\ell \big)^2} \right) + \mathcal{O}(h^2).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathcal{L}_{\text{mod}}(z, \dot{z}, h) = \mathcal{L}(z, \dot{z}) + \frac{h^2}{24} & \left[\frac{1}{2\pi} \sum_{j=1}^{N} \sum_{k \neq j} \Gamma_j \Gamma_k \operatorname{Re}\left(\left(\frac{\dot{z}_j - \dot{z}_k}{z_j - z_k} \right)^2 \right) \right. \\
& \left. - \frac{1}{\pi^2} \sum_{j=1}^{N} \sum_{k \neq j} \sum_{\ell \neq j} \Gamma_j \Gamma_k \Gamma_\ell \operatorname{Im}\left(\frac{\big(\dot{\overline{z}}_j - \dot{\overline{z}}_\ell \big)}{(z_j - z_k) \big(\overline{z}_j - \overline{z}_\ell \big)^2} \right) \right] + \mathcal{O}(h^4).
\end{aligned} \tag{3.24}$$

Trapezoidal rule

For the Trapezoidal rule, we find in the same way that

$$\begin{aligned}
&\mathcal{L}_{\text{mesh}}([z], h) = \mathcal{L}(z, \dot{z}) \\
&\quad + \frac{h^2}{24} \left[4 \sum_{j=1}^{N} \Gamma_j \operatorname{Im}\big(\dot{z}_j \ddot{\overline{z}}_j \big) - 2 \sum_{j=1}^{N} \sum_{k=1}^{j-1} \frac{\Gamma_j \Gamma_k}{\pi} \operatorname{Re}\left(\frac{\ddot{z}_j - \ddot{z}_k}{z_j - z_k} - \left(\frac{\dot{z}_j - \dot{z}_k}{z_j - z_k} \right)^2 \right) \right] + \mathcal{O}(h^4)
\end{aligned} \tag{3.25}$$

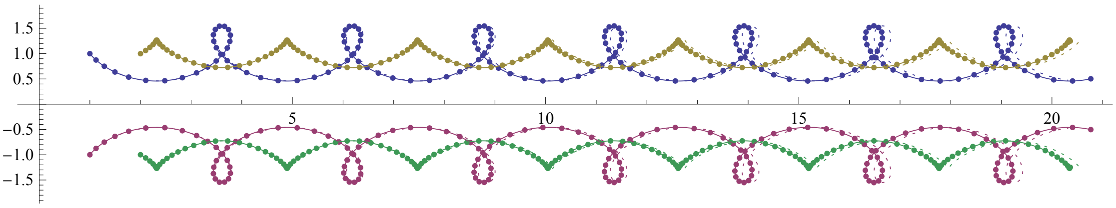


Figure 3.2. *Leapfrogging vortex pairs with the midpoint rule. No parasitic behavior is visible.*

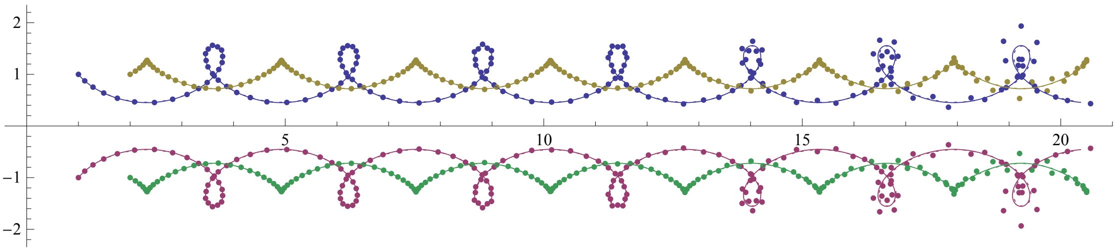


Figure 3.3. *Leapfrogging vortex pairs with the trapezoidal rule. One observes parasitic oscillations.*

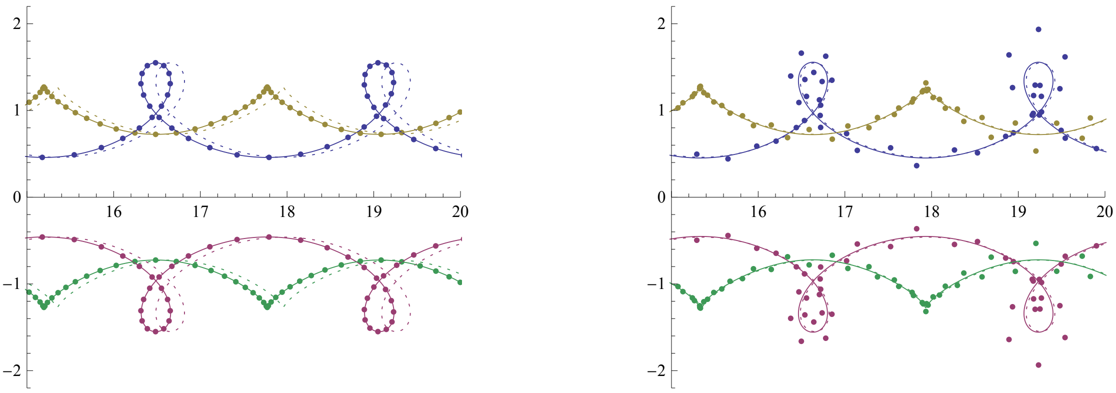


Figure 3.4. *Enlarged versions of the right hand sections of Figures 3.2 and 3.3: midpoint rule (left) and trapezoidal rule (right).*

Legend:	Dashed curves	*exact solution.*
	Bullets	*discrete solution.*
	Solid curves	*solution of the principal modified equation, truncated after the second order term.*
Parameters:	Initial positions	$(1,1)$, $(1,-1)$, $(2,1)$, *and* $(2,-1)$.
	Vortex strengths	1, -1, 2, -2, *respectively.*
	Time interval	$0 \leq t \leq 80$.

and

$$\begin{aligned}\mathcal{L}_{\text{mod}}(z,\dot{z},h) = \mathcal{L}(z,\dot{z}) + \frac{h^2}{24}\Bigg[&\frac{2}{\pi}\sum_{j=1}^{N}\sum_{k\neq j}\Gamma_j\Gamma_k \operatorname{Re}\left(\left(\frac{\dot{z}_j-\dot{z}_k}{z_j-z_k}\right)^2\right) \\ &-\frac{1}{\pi^2}\sum_{j=1}^{N}\sum_{k\neq j}\sum_{\ell\neq j}\Gamma_j\Gamma_k\Gamma_\ell \operatorname{Im}\left(\frac{(\dot{\overline{z}}_j-\dot{\overline{z}}_\ell)}{(z_j-z_k)(\overline{z}_j-\overline{z}_\ell)^2}\right)\Bigg] + \mathcal{O}(h^4).\end{aligned} \tag{3.26}$$

In Figures 3.2–3.4 we observe parasitic solutions for the trapezoidal rule, but not for the midpoint rule, where the solution of (the second-order truncation of) the principal modified equation shows excellent agreement with the discrete solution. In the numerical experiment the second truncations of the modified equations were calculated form Equations (3.23) and (3.25) using

$$\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_{\text{mesh}}}{\partial \dot{q}} = \mathcal{O}(h^4).$$

Note that Equations (3.23)–(3.26) are Lagrangians for the principal modified equation, not for the full system. In the context of Figures 3.2–3.4 they describe the solid curves as perturbation of the dashed curves. Whether or not the discrete solution oscillates around the solution of the principal modified equation cannot be seen form the calculations in this section, but it follows from the discussion in Section 3.4. There we showed that the midpoint rule does not suffer from parasitic solutions, but for the trapezoidal rule parasites can be expected, exactly as we see in Figures 3.2–3.4.

4. Application to numerical precession in the Kepler problem

This chapter is an adaptation of [88]

The Kepler problem models a point mass moving in a classical gravitational potential. Its Lagrangian is

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2}|\dot{x}|^2 + \frac{1}{|x|},$$

where $|x|$ denotes the Euclidean norm on $\mathbb{R}^N$. The equations of motion are

$$\ddot{x} = -\frac{x}{|x|^3}. \tag{4.1}$$

It is well known that the orbits of the Kepler problem with negative energy are ellipses with one of their foci at the origin (a property known as *Kepler's first law*). Since every orbit lies in a plane, it is sufficient to study this problem in $\mathbb{R}^2$.

Very good integrators for the Kepler problem are already available, see for example [18] and the references therein. The goal of this chapter is to illustrate how modified Lagrangians can be used to analyze and improve numerical integrators. We focus on ideas instead of performance and start with very simple methods. They are far from competitive compared to other known methods. Even the improved methods we construct will not be competitive compared to specialized methods available in the literature.

Central in our treatment will be the *precession* or *perihelion advance* of the numerical orbits, i.e. the slow rotation of the ellipse that the solution traces. For the exact solution there is no precession, but no common numerical method integrates the Kepler problem without precession. Using modified Lagrangians and a version of Noether's theorem to analyze the perturbation, we will provide leading order estimates of the precession for the Störmer-Verlet method and the implicit midpoint rule. We will use those estimates to construct some new methods which are superior for the Kepler problem. This procedure is similar in spirit to the concept of *modifying integrators* [16], which are numerical methods that start by perturbing the differential equation in a way that cancels the discretization error up to a certain order.

We start by mentioning a few well-known properties of the Kepler problem that will be useful later on. A thorough analytical study of the Kepler problem, including proofs of these properties, can be found for example in [31, Chapter 3].

Proposition 4.1. *The angular momentum* $\mathbb{L} = x_1\dot{x}_2 - \dot{x}_1 x_2$ *and the total energy* $\mathbb{E} = \frac{1}{2}|\dot{x}|^2 - \frac{1}{|x|}$ *are constants of motion of the Kepler problem in* $\mathbb{R}^2$*. Furthermore, the angular momentum satisfies*

$$\mathbb{L}^2 = |x|^2|\dot{x}|^2 - \langle x\,,\dot{x}\rangle^2\,,$$

where the brackets $\langle\cdot\,,\cdot\rangle$ *denote the standard scalar product on* $\mathbb{R}^2$*.*

Proposition 4.2. *Let* a *and* b *denote the lengths of the semimajor and semiminor axes of an orbit respectively. Let* θ *denote the angle between the major axis and the location* x *of the point mass. Then*

(a) *the eccentricity of the orbit is* $e = \sqrt{1 - \dfrac{b^2}{a^2}}$ *and there holds* $|x|^{-1} = \dfrac{a}{b^2}(1 + e\cos\theta)$,

(b) *the energy equals* $\mathbb{E} = \dfrac{-1}{2a}$,

(c) *the square of the angular momentum equals* $\mathbb{L}^2 = |x|^4\dot{\theta}^2 = \dfrac{b^2}{a}$ *(Kepler's second law),*

(d) *the period equals* $T = 2\pi a^{3/2}$ *(Kepler's third law).*

4.1. Noether's theorem with perturbations

The modified equation of a numerical integrator for the Kepler problem describes a perturbed Kepler problem. Perturbed Kepler problems are very relevant in celestial mechanics. In particular, one of the classical tests of general relativity is that its perturbation in the Kepler potential accounts for the precession of the orbit of the planet Mercury [95] (along with perturbations caused by the gravitational pull of the other planets). A Hamiltonian treatment of perturbed Kepler problems can be found for example in [31] or [18]. Here we will study it from the Lagrangian point of view.

The key observation in our study of the perturbed Kepler problem is that Noether's theorem [66, 68] can be extended to describe how perturbations affect conserved quantities.

Theorem 4.3. *Consider a Lagrange function* $\mathcal{L} : T\mathbb{R}^2 \to \mathbb{R}$ *and a horizontal vector field* ξ *on* $T\mathbb{R}^2$*, i.e.* $\xi = \xi_1 \frac{\partial}{\partial x_1} + \xi_2 \frac{\partial}{\partial x_2}$ *with coefficients* ξ_i *that are functions* $T\mathbb{R}^2 \to \mathbb{R}$*. Let*

$$\xi^{(1)} = \sum_{i=1}^{2} \left(\xi_i \frac{\partial}{\partial x_i} + \dot{\xi}_i \frac{\partial}{\partial \dot{x}_i} \right)$$

be the first prolongation of ξ, evaluated on solutions of the Euler-Lagrange equations, i.e. with

$$\dot{\xi}_i = \left\langle \frac{\partial \xi_i}{\partial x}\,, \dot{x} \right\rangle + \left\langle \frac{\partial \xi_i}{\partial \dot{x}}\,, \left(\frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2} \right)^{-1} \left(\frac{\partial \mathcal{L}}{\partial x} - \frac{\partial^2 \mathcal{L}}{\partial x \partial \dot{x}} \dot{x} \right) \right\rangle .$$

If

$$\xi^{(1)} \mathcal{L} = \frac{\mathrm{d}G}{\mathrm{d}t} + \varepsilon F$$

for some functions $F : T\mathbb{R}^2 \to \mathbb{R}$ and $G : \mathbb{R}^2 \to \mathbb{R}$, and a (small) parameter $\varepsilon \in \mathbb{R}$, then on solutions of the Euler-Lagrange equations we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\left\langle \frac{\partial \mathcal{L}}{\partial \dot{x}}\,, \xi \right\rangle - G \right) = \varepsilon F,$$

where by abuse of notation $\xi = (\xi_1, \xi_2)$. In particular, if $\varepsilon F = 0$, we have a conserved quantity $A = \frac{\partial \mathcal{L}}{\partial \dot{x}_1} \xi_1 + \frac{\partial \mathcal{L}}{\partial \dot{x}_2} \xi_2 - G$.

Proof. We have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \left(\left\langle \frac{\partial \mathcal{L}}{\partial \dot{x}}\,, \xi \right\rangle - G \right) &= \left\langle \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{x}}\,, \xi \right\rangle + \left(\xi^{(1)} \mathcal{L} - \left\langle \frac{\partial \mathcal{L}}{\partial x}\,, \xi \right\rangle \right) - \frac{\mathrm{d}G}{\mathrm{d}t} \\ &= \xi^{(1)} \mathcal{L} - \frac{\mathrm{d}G}{\mathrm{d}t} - \left\langle \frac{\partial \mathcal{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{x}}\,, \xi \right\rangle = \varepsilon F. \end{aligned}$$

□

4.1.1. The Laplace-Runge-Lenz vector

Following [46] we consider the Kepler problem and the vector field ξ defined by

$$\xi_1 = -\frac{1}{2} x_2 \dot{x}_2 \qquad \text{and} \qquad \xi_2 = x_1 \dot{x}_2 - \frac{1}{2} \dot{x}_1 x_2. \tag{4.2}$$

On solutions we have

$$\dot{\xi}_1 = -\frac{1}{2} \dot{x}_2^2 + \frac{1}{2} \frac{x_2^2}{|x|^3} \qquad \text{and} \qquad \dot{\xi}_2 = \frac{1}{2} \dot{x}_1 \dot{x}_2 - \frac{1}{2} \frac{x_1 x_2}{|x|^3}.$$

A straightforward calculation then shows that

$$\xi^{(1)} \mathcal{L} = \left\langle \frac{\partial \mathcal{L}}{\partial x}\,, \xi \right\rangle + \left\langle \frac{\partial \mathcal{L}}{\partial \dot{x}}\,, \dot{\xi} \right\rangle = \frac{\dot{x}_1}{|x|} - \frac{\langle x\,, \dot{x} \rangle\, x_1}{|x|^3} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{x_1}{|x|} \right).$$

Hence we can apply the unperturbed Noether theorem (i.e. $\varepsilon F = 0$) with $G(x) = \frac{x_1}{|x|}$ and find that

$$A(x, \dot{x}) = -\dot{x}_1 x_2 \dot{x}_2 + x_1 \dot{x}_2^2 - \frac{x_1}{|x|} = |\dot{x}|^2 x_1 - \langle x\,, \dot{x} \rangle\, \dot{x}_1 - \frac{x_1}{|x|}$$

is a conserved quantity.

The conserved quantity A is the first component of the *Laplace-Runge-Lenz (LRL) vector*, which points from the gravitational center to the perihelion and has a magnitude equal to the eccentricity e of the orbit. The second component of the LRL vector is

$$B(x, \dot{x}) = |\dot{x}|^2 x_2 - \langle x \,, \dot{x} \rangle \, \dot{x}_2 - \frac{x_2}{|x|}$$

and can be obtained by setting $\xi_1 = x_2 \dot{x}_1 - \frac{1}{2} x_1 \dot{x}_2$ and $\xi_2 = -\frac{1}{2} x_1 \dot{x}_1$. We denote by $\omega = \arctan\left(\frac{B}{A}\right)$ the angle of the LRL vector with the first coordinate axis.

Remark. The existence of this conserved quantity is related to the fact that the 3-dimensional Kepler problem possesses an $SO(4)$-symmetry, rather than just the obvious $SO(3)$-symmetry. In suitable coordinates a solution can be "rotated" into other solutions with the same energy but different angular momentum [57, 75].

4.1.2. Precession in the perturbed Kepler problem

Now consider the perturbed Kepler problem, $\mathcal{L} = \frac{1}{2}|\dot{x}|^2 + \frac{1}{|x|} + \varepsilon \overline{\mathcal{L}}(x, \dot{x})$. Note that this also induces a perturbation in the prolonged vector field, which now reads $\xi^{(1)} + \varepsilon \overline{\xi^{(1)}}$ because the quantities $\dot{\xi}_1$ and $\dot{\xi}_2$ contain second derivatives which are evaluated using the perturbed equations of motion. The perturbation term is

$$\overline{\xi^{(1)}} = \left\langle \frac{\partial \xi_1}{\partial \dot{x}} \,, \mathrm{EL}(\overline{\mathcal{L}}) \right\rangle \frac{\partial}{\partial \dot{x}_1} + \left\langle \frac{\partial \xi_2}{\partial \dot{x}} \,, \mathrm{EL}(\overline{\mathcal{L}}) \right\rangle \frac{\partial}{\partial \dot{x}_2} + \mathcal{O}(\varepsilon),$$

where

$$\mathrm{EL}(\overline{\mathcal{L}}) = \frac{\partial \overline{\mathcal{L}}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \overline{\mathcal{L}}}{\partial \dot{x}}$$

is the Euler-Lagrange expression for $\overline{\mathcal{L}}$. We call the change in angle of the LRL vector over one period of the unperturbed system the *precession rate*.

Proposition 4.4. *It the major axis of an orbit is $\mathcal{O}(\varepsilon)$-close to the x_2-axis, then the precession rate is given by*

$$\Delta\omega = -\frac{2\varepsilon T}{e} \left[\left\langle \mathrm{EL}(\overline{\mathcal{L}}) \,, \xi \right\rangle \right] + \mathcal{O}(\varepsilon^2), \tag{4.3}$$

where T is the period of the unperturbed orbit, $\xi = (\xi_1, \xi_2)$ is defined by Equation (4.2), and $[\,\cdot\,]$ denotes the average over one period.

Note that Equation (4.3) is not invariant under rotations because we use a fixed vector field ξ, corresponding to the first component of the LRL vector. Hence the condition on the orientation of the orbit.

Proof of Proposition 4.4. Set $G = \frac{x_1}{|x|}$, then

$$\left(\xi^{(1)} + \varepsilon\overline{\xi^{(1)}}\right)\left(\mathcal{L} + \varepsilon\overline{\mathcal{L}}\right) = \frac{\mathrm{d}G}{\mathrm{d}t} + \varepsilon\left(\overline{\xi^{(1)}}\mathcal{L} + \xi^{(1)}\overline{\mathcal{L}}\right) + \mathcal{O}(\varepsilon^2),$$

where $\xi^{(1)}+\varepsilon\overline{\xi^{(1)}}$ is the first prolongation of ξ on solutions of the Euler Lagrange equations of the perturbed Lagrangian $\mathcal{L} + \varepsilon\overline{\mathcal{L}}$. Hence by Theorem 4.3 it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\langle \frac{\partial(\mathcal{L} + \varepsilon\overline{\mathcal{L}})}{\partial \dot{x}}, \xi \right\rangle - G\right) = \varepsilon\left(\overline{\xi^{(1)}}\mathcal{L} + \xi^{(1)}\overline{\mathcal{L}}\right) + \mathcal{O}(\varepsilon^2),$$

from which we conclude that

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \varepsilon\left(\overline{\xi^{(1)}}\mathcal{L} + \xi^{(1)}\overline{\mathcal{L}} - \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\partial\overline{\mathcal{L}}}{\partial \dot{x}}, \xi \right\rangle\right) + \mathcal{O}(\varepsilon^2). \tag{4.4}$$

Now observe that

$$\begin{aligned}\xi^{(1)}\overline{\mathcal{L}} - \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\partial\overline{\mathcal{L}}}{\partial \dot{x}}, \xi \right\rangle &= \left\langle \frac{\partial\overline{\mathcal{L}}}{\partial x}, \xi \right\rangle + \left\langle \frac{\partial\overline{\mathcal{L}}}{\partial \dot{x}}, \dot{\xi} \right\rangle - \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\partial\overline{\mathcal{L}}}{\partial \dot{x}}, \xi \right\rangle \\ &= \left\langle \mathrm{EL}(\overline{\mathcal{L}}), \xi \right\rangle + \mathcal{O}(\varepsilon),\end{aligned}$$

where the error term comes from the fact that $\dot{\xi}$ is evaluated on the unperturbed system. We also have that

$$\overline{\xi^{(1)}}\mathcal{L} = \left\langle \frac{\partial \xi_1}{\partial \dot{x}}, \mathrm{EL}(\overline{\mathcal{L}}) \right\rangle \dot{x}_1 + \left\langle \frac{\partial \xi_2}{\partial \dot{x}}, \mathrm{EL}(\overline{\mathcal{L}}) \right\rangle \dot{x}_2 = \left\langle \frac{\partial \xi_1}{\partial \dot{x}}\dot{x}_1 + \frac{\partial \xi_2}{\partial \dot{x}}\dot{x}_2, \mathrm{EL}(\overline{\mathcal{L}}) \right\rangle + \mathcal{O}(\varepsilon).$$

For our choice of ξ, defined in Equation (4.2), we have $\frac{\partial \xi_1}{\partial \dot{x}}\dot{x}_1 + \frac{\partial \xi_2}{\partial \dot{x}}\dot{x}_2 = (\xi_1, \xi_2) = \xi$, hence Equation (4.4) simplifies to

$$\frac{\mathrm{d}A}{\mathrm{d}t} = 2\varepsilon \left\langle \mathrm{EL}(\overline{\mathcal{L}}), \xi \right\rangle + \mathcal{O}(\varepsilon^2).$$

The change in angle of the Laplace-Runge-Lenz vector is given by

$$\dot{\omega} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\arctan\frac{B}{A}\right) = \frac{1}{A^2 + B^2}\left(A\frac{\mathrm{d}B}{\mathrm{d}t} - B\frac{\mathrm{d}A}{\mathrm{d}t}\right).$$

Choose a coordinate system such that $A = \mathcal{O}(\varepsilon)$ and $B \geq 0$. Then B approximately equals the eccentricity e and the derivative of the angle of the LRL vector is

$$\dot{\omega} = -\frac{1}{B}\frac{\mathrm{d}A}{\mathrm{d}t} + \mathcal{O}(\varepsilon^2) = -\frac{2\varepsilon}{e}\left\langle \mathrm{EL}(\overline{\mathcal{L}}), \xi \right\rangle + \mathcal{O}(\varepsilon^2). \qquad \square$$

4.2. Modified Lagrangians for two common integrators

Let us make things more concrete. We look at two simple variational integrators, calculate their modified Lagrangian, and have a first look at their numerical performance.

4.2.1. Störmer-Verlet method

The Störmer-Verlet (SV) discretization with step size h of a second order differential equation $\ddot{x} = f(x)$ is

$$x_{k+1} - 2x_k + x_{k-1} = h^2 f(x_k).$$

If $f(x) = -\frac{\mathrm{d}}{\mathrm{d}x} U(x)$, this is the discrete Euler-Lagrange equation for

$$L_{SV}(x_k, x_{k+1}) = \frac{1}{2} \left| \frac{x_{k+1} - x_k}{h} \right|^2 - \frac{1}{2} U(x_k) - \frac{1}{2} U(x_{k+1}).$$

The modified Lagrangian of second order accuracy is given by Equation (2.13),

$$\mathcal{L}_{\mathrm{mod},2}(x, \dot{x}) = \frac{1}{2} |\dot{x}|^2 - U(x) + \frac{h^2}{24} \Big(|U'(x)|^2 - 2 \left\langle \dot{x} \, , U''(x) \dot{x} \right\rangle \Big).$$

In the particular case of the Kepler problem this becomes

$$\mathcal{L}_{\mathrm{mod},2}(x, \dot{x}) = \frac{1}{2} |\dot{x}|^2 + \frac{1}{|x|} + \frac{h^2}{24} \left(\frac{1}{|x|^4} - 2 \frac{|\dot{x}|^2}{|x|^3} + 6 \frac{\langle x \, , \dot{x} \rangle^2}{|x|^5} \right). \tag{4.5}$$

Its Euler-Lagrange equation agrees with the modified equation with a defect of order $\mathcal{O}(h^4)$. A comparison of the numerical solution and the solution of this truncation of the modified equation is shown in Figure 4.1.

4.2.2. Implicit midpoint rule

The second order formulation of the implicit midpoint rule (MP) applied to the differential equation $\ddot{x} = f(x)$ is

$$x_{k+1} - 2x_k + x_{k-1} = \frac{h^2}{2} f\left(\frac{x_k + x_{k+1}}{2} \right) + \frac{h^2}{2} f\left(\frac{x_{k-1} + x_k}{2} \right).$$

If $f(x) = -\frac{\mathrm{d}}{\mathrm{d}x} U(x)$, this is the discrete Euler-Lagrange equation for

$$L_{MP}(x_k, x_{k+1}) = \frac{1}{2} \left| \frac{x_{k+1} - x_k}{h} \right|^2 - U\left(\frac{x_k + x_{k+1}}{2} \right).$$

For this discrete Lagrangian we have

$$\begin{aligned}\mathcal{L}_{\mathrm{disc}}([x],h) &= \frac{1}{2}|\dot{x}|^2+\frac{h^2}{24}\Big(\big\langle x^{(3)}\,,\dot{x}\big\rangle - 3\,\big\langle U'(x)\,,\ddot{x}\big\rangle\Big) + \mathcal{O}(h^4)\\ &= \frac{1}{2}|\dot{x}|^2+\frac{h^2}{24}\Big(\big\langle x^{(3)}\,,\dot{x}\big\rangle + 3|U'(x)|^2\Big) + \mathcal{O}(h^4),\end{aligned}$$

from which we need to subtract

$$\begin{aligned}\frac{h^2}{24}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{L}(x,\dot{x}) &= \frac{h^2}{24}\Big(\big\langle x^{(3)}\,,\dot{x}\big\rangle + |\ddot{x}|^2 - \big\langle U'(x)\,,\ddot{x}\big\rangle - \big\langle \dot{x}\,,U''(x)\dot{x}\big\rangle\Big) + \mathcal{O}(h^4)\\ &= \frac{h^2}{24}\Big(\big\langle x^{(3)}\,,\dot{x}\big\rangle + 2|U'(x)|^2 - \big\langle \dot{x}\,,U''(x)\dot{x}\big\rangle\Big) + \mathcal{O}(h^4)\end{aligned}$$

to get the modified Lagrangian of second order accuracy. We find

$$\mathcal{L}_{\mathrm{mod},2}(x,\dot{x}) = \frac{1}{2}|\dot{x}|^2 - U(x) + \frac{h^2}{24}\Big(|U'(x)|^2 + \big\langle \dot{x}\,,U''(x)\dot{x}\big\rangle\Big).$$

For the Kepler problem we have

$$\mathcal{L}_{\mathrm{mod},2}(x,\dot{x}) = \frac{1}{2}|\dot{x}|^2+\frac{1}{|x|} + \frac{h^2}{24}\left(\frac{1}{|x|^4} + \frac{|\dot{x}|^2}{|x|^3} - 3\frac{\langle x\,,\dot{x}\rangle^2}{|x|^5}\right).$$

A comparison of the numerical solution and the solution of the modified equation of second order accuracy is shown in Figure 4.2.

4.3. Numerical precession

We now apply Proposition 4.4 to the modified Lagrangians from Section 4.2. This gives us a leading order estimate of the precession rates of the integrators.

4.3.1. Störmer-Verlet scheme

The perturbation term of the truncated modified Lagrangian (4.5) is

$$\varepsilon\overline{\mathcal{L}} = \frac{h^2}{24}\left(\frac{1}{|x|^4} - 2\frac{|\dot{x}|^2}{|x|^3} + 6\frac{\langle x\,,\dot{x}\rangle^2}{|x|^5}\right).$$

In the following we identify $\varepsilon = \frac{h^2}{24}$. We want to evaluate Equation (4.3). Using the leading order equations of motion (4.1), which are valid up to an error of order $\mathcal{O}(h^2)$, we find

$$\mathrm{EL}(\overline{\mathcal{L}}) = 4\frac{x}{|x|^6} - 6\frac{|\dot{x}|^2 x}{|x|^5} + 30\frac{\langle x\,,\dot{x}\rangle^2 x}{|x|^7} - 12\frac{\langle x\,,\dot{x}\rangle\,\dot{x}}{|x|^5} + \mathcal{O}(h^2).$$

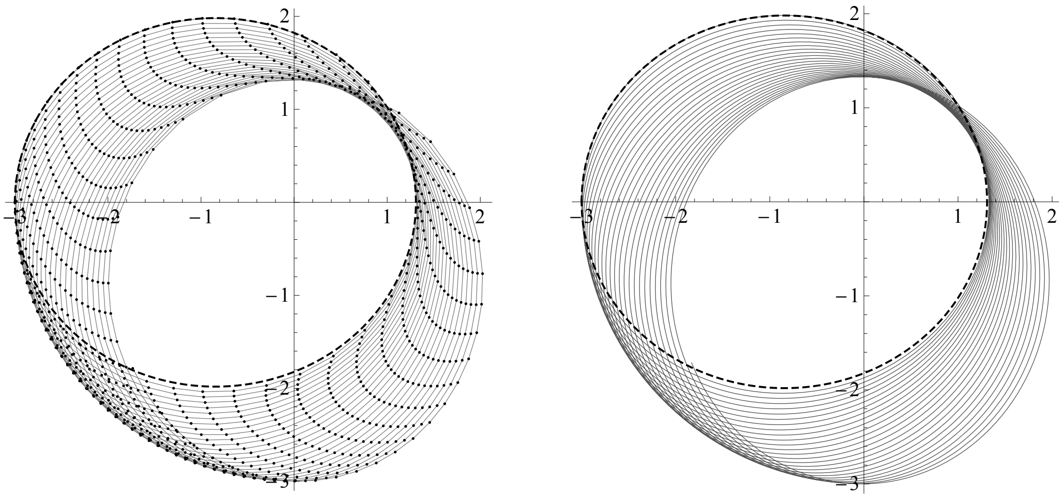


Figure 4.1. *Störmer-Verlet method with* 1000 *steps of size* $h = 0.5$.

Left: numerical solution.

Right: solution of the modified equation truncated after the second order term.

In each image the dashed ellipse is the exact solution. The initial values are chosen as described in Subsection 4.5.1.

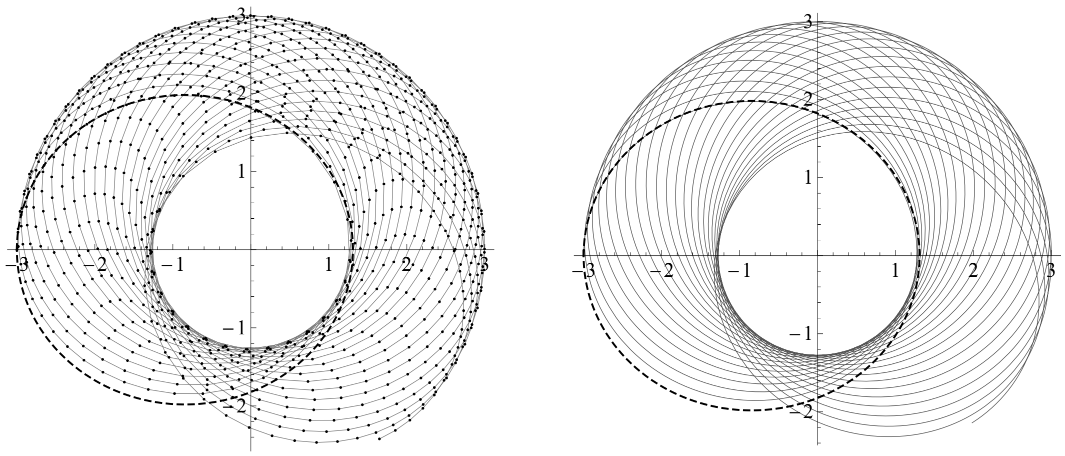


Figure 4.2. *Implicit midpoint rule with* 1000 *steps of size* $h = 0.5$.

Using the fact that $\langle x\,,\xi\rangle = \frac{1}{2}(x_1\dot{x}_2 - \dot{x}_1 x_2)x_2 = \frac{1}{2}\mathbb{L}x_2$ and $\langle \dot{x}\,,\xi\rangle = \mathbb{L}\dot{x}_2$, the leading order equations of motion, and Proposition 4.1 we obtain

$$\begin{aligned}[\langle \mathrm{EL}(\overline{\mathcal{L}})\,,\xi\rangle] &= \left[2\frac{x_2}{|x|^6} - 3\frac{|\dot{x}|^2 x_2}{|x|^5} + 15\frac{\langle x\,,\dot{x}\rangle^2 x_2}{|x|^7} - 12\frac{\langle x\,,\dot{x}\rangle\,\dot{x}_2}{|x|^5}\right]\mathbb{L} + \mathcal{O}(h^2)\\ &= \left[30\frac{x_2}{|x|^6} + 24\mathbb{E}\frac{x_2}{|x|^5} - 15\mathbb{L}^2\frac{x_2}{|x|^7} + 4\frac{\mathrm{d}}{\mathrm{d}t}\frac{\dot{x}_2}{|x|^3}\right]\mathbb{L} + \mathcal{O}(h^2). \end{aligned} \tag{4.6}$$

The average $[\cdot]$ is taken along the unperturbed orbit, which is periodic, so $\left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\dot{x}_2}{|x|^3}\right] = 0$. For the other terms we have the following Lemma, which corresponds to the computation of the $C_n(e)$ of [18].

Lemma 4.5. *On solutions of the unperturbed Kepler problem, with the LRL vector along the negative x_2-axis, there holds*

(*a*) $\left[\dfrac{x_2}{|x|^5}\right] = \dfrac{a}{b^5}e,$

(*b*) $\left[\dfrac{x_2}{|x|^6}\right] = \dfrac{a^2}{b^7}\left(\dfrac{3}{2}e + \dfrac{3}{8}e^3\right),$

(*c*) $\left[\dfrac{x_2}{|x|^7}\right] = \dfrac{a^3}{b^9}\left(2e + \dfrac{3}{2}e^3\right),$

where a and b are the lengths of the semimajor and semiminor axes of the orbit respectively, and e is the eccentricity.

Proof. Introduce polar coordinates $x_1 = -r\sin\theta$, $x_2 = r\cos\theta$, where $\theta = 0$ corresponds to the positive x_2-axis. We have

$$\left[\frac{x_2}{|x|^k}\right] = \left[\frac{cos\theta}{|x|^{k-1}}\right] = \frac{1}{T}\int_0^T \frac{cos\theta}{|x|^{k-1}}\,\mathrm{d}t.$$

Following [20] we use the identities from Proposition 4.2 ((*a*), (*c*), and (*d*)) to rewrite this as

$$\begin{aligned}\left[\frac{x_2}{|x|^k}\right] &= \frac{b^{5-2k}}{\pi a^{4-k}}\int_0^\pi (1 + e\cos\theta)^{k-3}\cos\theta\,\mathrm{d}\theta\\ &= \frac{b^{5-2k}}{\pi a^{4-k}}\int_0^\pi \sum_j \binom{k-3}{j} e^j \cos^{j+1}\theta\,\mathrm{d}\theta.\end{aligned}$$

Whenever, j is even, we have $\int_0^\pi \cos^{j+1}\theta\,\mathrm{d}\theta = 0$. For $j = 1$ and $j = 3$ we find $\int_0^\pi \cos^2\theta\,\mathrm{d}\theta = \frac{\pi}{2}$ and $\int_0^\pi \cos^4\theta\,\mathrm{d}\theta = \frac{3\pi}{8}$. Hence

$$\left[\frac{x_2}{|x|^k}\right] = \frac{b^{5-2k}}{\pi a^{4-k}}\left(\frac{\pi}{2}\binom{k-3}{1}e + \frac{3\pi}{8}\binom{k-3}{3}e^3 + \dots\right).$$

The claims now follow by evaluating this expression for $k = 5, 6, 7$. $\square$

We can use Proposition 4.2 ((*b*) and (*c*)) and Lemma 4.5 to write Equation (4.6) in terms of a and b. Using Proposition 4.4 we then find the precession per revolution:

$$
\begin{aligned}
&-4\pi a^{3/2}\frac{h^2}{24}\left(30\frac{a^2}{b^7}\left(\frac{3}{2}+\frac{3}{8}e^2\right)+24\frac{-1}{2a}\frac{a}{b^5}-15\frac{b^2}{a}\frac{a^3}{b^9}\left(2+\frac{3}{2}e^2\right)\right)\frac{b}{\sqrt{a}}\operatorname{sgn}(\mathbb{L})+\mathcal{O}(h^4)\\
&=-4\pi ab\frac{h^2}{24}\left(30\frac{a^2}{b^7}\left(\frac{15}{8}-\frac{3}{8}\frac{b^2}{a^2}\right)-\frac{12}{b^5}-15\frac{a^2}{b^7}\left(\frac{7}{2}-\frac{3}{2}\frac{b^2}{a^2}\right)\right)\operatorname{sgn}(\mathbb{L})+\mathcal{O}(h^4)\\
&=-\frac{\pi h^2}{24}\left(15\frac{a^3}{b^6}-3\frac{a}{b^4}\right)\operatorname{sgn}(\mathbb{L})+\mathcal{O}(h^4),
\end{aligned}
$$

assuming the major axis of the orbit is $\mathcal{O}(h^2)$-close to the x_2-axis. However, since both this expression and the perturbed Kepler problem are rotationally symmetric, we can conclude that statement holds regardless of the orientation of the major axis.

In summary we have the following:

Theorem 4.6. *The numerical precession rate of the Störmer-Verlet method with step size h is*

$$
-\operatorname{sgn}(\mathbb{L})\frac{\pi}{24}\left(15\frac{a^3}{b^6}-3\frac{a}{b^4}\right)h^2+\mathcal{O}(h^4),
$$

where a and b denote the semimajor and semiminor axes of the orbit of the exact solution and sgn *is the sign function. In particular, the precession and the motion are in opposite directions.*

For the example shown in Figure 4.1, the precession rate predicted by Theorem 4.6 is 0.067 radians per revolution and the observed numerical precession rate is 0.064 radians per revolution.

4.3.2. Implicit midpoint rule

In exactly the same way as for the Störmer-Verlet method, we obtain the following result:

Theorem 4.7. *The numerical precession rate of the midpoint rule with step size h is*

$$
\operatorname{sgn}(\mathbb{L})\frac{\pi}{12}\left(15\frac{a^3}{b^6}-3\frac{a}{b^4}\right)h^2+\mathcal{O}(h^4).
$$

In particular, the precession is in the same direction as the motion.

Note that in the leading order this expression differs by exactly a factor -2 from the expression for the Störmer-Verlet method. We will exploit this in the next section to construct new integrators.

For the example shown in Fig. 4.2, the precession rate predicted by Theorem 4.7 is -0.13 radians per revolution and the observed numerical precession rate is -0.16 radians per revolution.

4.4. New integrators

Based on Theorems 4.6 and 4.7 we propose three new integrators. They all have a precession rate of order $\mathcal{O}(h^4)$ instead of $\mathcal{O}(h^2)$.

4.4.1. Linear combination of the Lagrangians

Consider the discrete Lagrangian

$$\begin{aligned} L(x_j, x_{j+1}) &= \frac{2}{3} L_{SV}(x_j, x_{j+1}) + \frac{1}{3} L_{MP}(x_j, x_{j+1}) \\ &= \frac{1}{2} \left| \frac{x_{j+1} - x_j}{h} \right|^2 - \frac{1}{3} U(x_j) - \frac{1}{3} U(x_{j+1}) - \frac{1}{3} U\left(\frac{x_j + x_{j+1}}{2} \right). \end{aligned}$$

Its Euler-Lagrange equations define an implicit method,

$$x_{j+1} - 2x_j + x_{j-1} = -\frac{2h^2}{3} U'(x_j) - \frac{h^2}{6} U'\left(\frac{x_{j-1} + x_j}{2} \right) - \frac{h^2}{6} U'\left(\frac{x_j + x_{j+1}}{2} \right).$$

We refer to this integrator as the *mixed Lagrangian* (ML) method. By construction, this is a variational integrator.

4.4.2. Lagrangian Composition

Consider the discrete Lagrangians

$$L_j(x_k, x_{k+1}) = \begin{cases} L_{MP}(x_k, x_{k+1}) = \dfrac{1}{2} \left| \dfrac{x_{k+1} - x_k}{h} \right|^2 - U\left(\dfrac{x_k + x_{k+1}}{2} \right) & \text{if } 3|j, \\ L_{SV}(x_k, x_{k+1}) = \dfrac{1}{2} \left| \dfrac{x_{k+1} - x_k}{h} \right|^2 - \dfrac{1}{2} U(x_k) - \dfrac{1}{2} U(x_{k+1}) & \text{otherwise.} \end{cases}$$

We look for a discrete curve $(x_j)_{j \in \mathbb{Z}}$ that extremizes the action

$$\sum_{j=1}^{N} L_j(x_{j-1}, x_j) = L_{SV}(x_0, x_1) + L_{SV}(x_1, x_2) + L_{MP}(x_2, x_3) + \ldots.$$

This gives us three different Euler-Lagrange equations which are applied for different values of j mod 3. Indeed $\mathrm{D}_2 L_j(x_{j-1}, x_j) + \mathrm{D}_1 L_{j+1}(x_j, x_{j+1})$ simplifies to

$$\begin{cases} x_{j+1} - 2x_j + x_{j-1} = -\dfrac{h^2}{2} U'\left(\dfrac{x_{j-1} + x_j}{2} \right) - \dfrac{h^2}{2} U'(x_j) & \text{if } j \equiv 0 \bmod 3, \\ x_{j+1} - 2x_j + x_{j-1} = -h^2 U'(x_j) & \text{if } j \equiv 1 \bmod 3, \\ x_{j+1} - 2x_j + x_{j-1} = -\dfrac{h^2}{2} U'\left(\dfrac{x_j + x_{j+1}}{2} \right) - \dfrac{h^2}{2} U'(x_j) & \text{if } j \equiv 2 \bmod 3. \end{cases}$$

Hence to determine the evolution we alternate between the Störmer-Verlet method (for $j \equiv 1 \mod 3$) and two new difference equations. We refer to this integrator as the *Lagrangian composition* (LC) method. Strictly speaking the LC method should be considered as an integrator with step size $3h$, but for fair comparison with the other methods we will still refer to the internal step h as the step size.

This method of composing variational integrators is equivalent to composing the corresponding symplectic maps [52, Sect. 2.5].

4.4.3. Composition of the difference equations

Alternatively we can compose the difference equations obtained by the implicit midpoint rule and the Störmer-Verlet method respectively,

$$\begin{cases} x_{j+1} - 2x_j + x_{j-1} = -\dfrac{h^2}{2} U'\left(\dfrac{x_{j-1} + x_j}{2}\right) - \dfrac{h^2}{2} U'\left(\dfrac{x_j + x_{j+1}}{2}\right) & \text{if } j \equiv 2 \bmod 3, \\ x_{j+1} - 2x_j + x_{j-1} = -h^2 U'(x_j) & \text{otherwise.} \end{cases}$$

We refer to this integrator as the *difference equation composition* (DEC) method. Just like for the LC method, we will abuse terminology and call the internal step h the step size.

It is not clear if this construction yields a variational method, but numerical experiments show long-term near-conservation of energy and angular momentum. This seems to be a general phenomenon: also for other potentials U and other variational integrators, the corresponding DEC method shows the long-term behavior one expects from a variational integrator.

4.5. Numerical results

In this section we compare the new methods of Section 4.4 numerically with the Störmer-Verlet scheme, the implicit midpoint rule, and two fourth order symplectic methods: the well-known integrator of Forest and Ruth [26] and Chin's "C" algorithm which is especially well-suited for the Kepler problem [17, 19].

4.5.1. Choice of initial values

In all our examples we use the initial values

$$x(0) = (-3, 0) \qquad \text{and} \qquad \dot{x}(0) = (0, 0.45).$$

For the discretizations we need specify $x_0 = x(0)$ and $x_1 \approx x(h)$. Our convention is to choose x_1 such that the discrete momentum $p_0 = -\mathrm{D}_1 L(x_0, x_1)$ equals the initial velocity $\dot{x}(0)$.

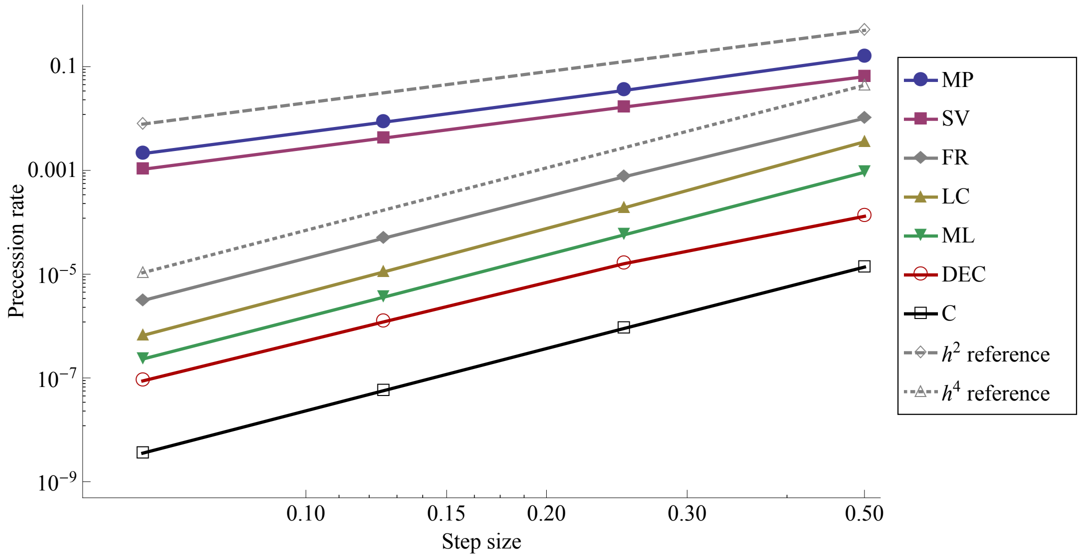


Figure 4.3. *Precession rate in radians per revolution for the different methods with step sizes* $h = 0.0625$*,* $h = 0.125$*,* $h = 0.25$ *and* $h = 0.5$*.*

For the composition of difference equations no discrete Lagrangian and hence no discrete momentum is known. To determine the second initial point x_1 in this case we use the momentum p_0 corresponding to the Störmer-Verlet method, because this is the method we would have used to calculate x_1 if x_0 was not the first point.

The choice of the initial value x_1 does not affect the precession behavior. However, it can have a significant effect on the error over time. If the initial condition has a slightly wrong energy, then the period of the numerical solution will have a slight error as well. This will cause a linearly growing phase shift.

4.5.2. Precession

Figure 4.3 shows the precession rates on a logarithmic scale for all five methods and a few choices of step size. It shows that the precession rates of the new methods behave like h^4, compared to h^2 for the methods from Section 4.2.

As for the three new methods, the mixed Lagrangian method beats the Lagrangian composition method, but the surprising winner is the composition of difference equations.

All our new methods have smaller precession rates than the fourth order symplectic integrator of Forest and Ruth [26]. On the other hand, Chin's fourth order symplectic "C" algorithm [17, 19] outperforms our methods.

4.5.3. Total error

The precession rate is not as closely related to the total error as one might expect. In many cases the numerical solution has a phase shift which contributes significantly to the total error. For the composition methods LC and DEC this phase shift is highly dependent on the step size and the initial conditions. Hence the total error growth for these methods is also sensitive to the choice of step size and initial condition. This can be seen by comparing Figures 4.4 and 4.5. In these figures we show a long time calculation with a large step size, leading to large errors. This means that the result is useless for practical purposes, but it allows us to visualize the rate of error growth of the different methods relative to each other.

4.5.4. Speed

To give a rough comparison of the computational effort required for the different methods, we list the relative running times of a long time calculation (20 000 steps):

Störmer-Verlet	(SV)	0.67s	Mixed Lagrangian	(ML)	23s
MidPoint rule	(MP)	22s	Difference Equation composition	(DEC)	7.9s
Forest-Ruth	(FR)	2.0s	Lagrangian Composition	(LC)	8.2s
Chin C	(C)	2.2s			

We made a limited effort towards optimizing our implementation, so the given running times should only be taken as a rough indication. As expected the explicit methods SV, FR, and C are the fastest. Between those, SV is about three times faster than the other two, which have internal steps increasing their computational complexity. For the composition methods DEC and LC only one out of every three steps is implicit, hence they are roughly three times faster than the fully implicit methods MP and ML.

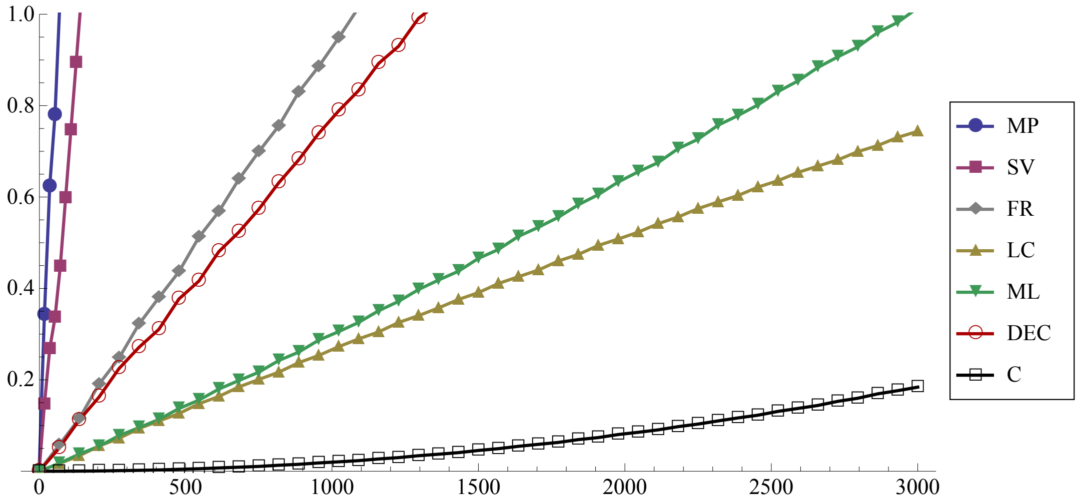


Figure 4.4. *Error in position over a time interval of length* 3000 *with step size* $h = 0.45$, *smoothed by taking a moving maximum over one period. The markers are only for the purpose of identifying the methods, they do not correspond to individual time steps.*

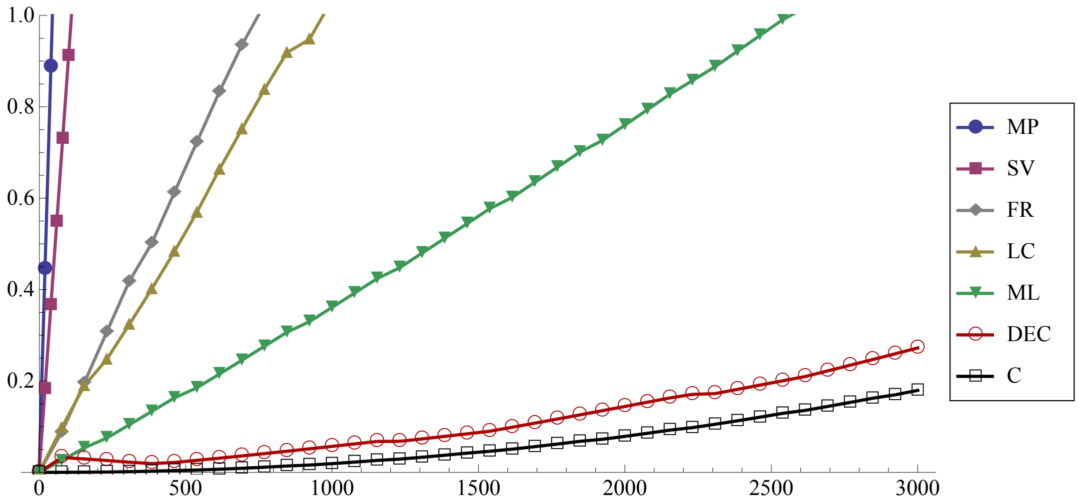


Figure 4.5. *Error in position with step size* $h = 0.5$. *Everything else as in Figure 4.4.*

5. Summary and outlook: variational principles in numerical integration

We addressed the question whether modified equations for variational integrators are Lagrangian. In a strict sense the answer is no: truncations of the modified equations are not Euler-Lagrange equations. However, they can be turned into Euler-Lagrange equations by adding higher-order corrections, or by considering the full formal power series. We have proved this constructively and without a detour to the Hamiltonian side, where the corresponding property is well-known.

Starting from the discrete Lagrangian of the numerical method, we first constructed continuous Lagrangians $\mathcal{L}_{\text{disc}}[x]$ and $\mathcal{L}_{\text{mesh}}[x]$ satisfying unconventional action principles and depending on higher derivatives. Using the peculiarities of the meshed action principle and the relation of $\mathcal{L}_{\text{mesh}}[x]$ to the discrete Lagrangian, we managed to eliminate all higher derivatives and obtained the modified Lagrangian $\mathcal{L}_{\text{mod}}(x, \dot{x})$. From each of these three Lagrangians the modified equation can be calculated.

Degenerate Lagrangians do not pose a fundamental obstruction to our method. In the case of Lagrangians linear in velocities, the main issue is that variational integrators are two-step methods and parasitic solutions can occur. To deal with this we used the trick of doubling the variables. Since this doubling preserves the variational nature of the system, it does not interfere with the construction of a modified Lagrangian.

As an application of modified Lagrangians we studied the precession rates of the implicit midpoint rule and the Störmer-Verlet method applied to the Kepler problem. The Lagrangian point of view lends itself perfectly to the use of a perturbed version of Noether's Theorem. The leading order estimates of the precession rate motivated the construction of three new integrators. They are significantly better than the methods we started from, but still they are outperformed by specialized methods. Our main goal was to elucidate methodology, rather than to obtain competitive methods. The techniques we used to analyze the integrators can be applied to any variational integrator and generalized to any order. However, it is not clear to which extend to constructions of improved integrators can be generalized. Hence further research is needed in order to convert these ideas into a scheme to produce competitive numerical methods.

Our discussion of the Kepler problem is one illustration of how modified equations can be used to construct better numerical integrators. More generally, there is a class of improved numerical methods known as *modifying integrators* [16]. These integrators involve a procedure which can be though of as the inverse of calculating a modified equation, where the differential equation is perturbed *before* applying the integrator to cancel the error up to a certain order. Building on our construction of a modified Lagrangian, the same idea has been applied in [22] on the level of Lagrangians. Based on the modified Lagrangian they construct a *surrogate Lagrangian*, which is a continuous Lagrangian with a perturbation that counteracts the discretization error. This approach looks especially promising in the context of optimal control [21].

On the theoretical side some interesting questions remain open. Can the method of modified Lagrangians be extended to systems subject to external forces or constraints, in particular nonholonomic ones? Can it be used for Lagrangian PDEs? In the end one would always like to get rigorous long-time conservation results from modified equations. Such results often rely on optimal truncation of the power series in the absence of fast oscillations. It is unclear whether the Lagrangian approach can improve known results of that type, but the heuristic wisdom that "criticality implies regularity" suggests that it might.

Part II.

Variational principles in discrete and continuous integrable systems

6. Pluri-Lagrangian systems

This chapter is mostly a review of the existing literature. This includes the paper [81] based on the author's master thesis. The only original result in this chapter is in Section 6.5.3.

Even more than geometric numerical integration, the literature on integrable systems tends to prefer the Hamiltonian point of view over the Lagrangian one. In mechanics, complete integrability is usually understood as Liouville-Arnold integrability, which is a quintessentially Hamiltonian condition [8] [7, Chapter 10]. Also in the context of soliton equations, and the KdV equation in particular, Hamiltonians have been present since the early days of the modern study of integrability [29, 99]. By virtue of bi-Hamiltonian structures [50] they have become one of the most important tools in the theory of integrable PDEs. A very rich source on Hamiltonian methods for integrable PDEs is the monograph [23].

Over the past decade, in [47, 97, 80, 81] and other works, a variational perspective of integrability has been developed. The notion of *pluri-Lagrangian* or *Lagrangian multiform* systems centers around an unconventional variational principle. A major advantage of the variational point of view is that it can be applied in the discrete as well as the continuous case. This raises the question of how to relate those two cases. How to discretize such systems? Or take continuum limits of them? The first question seems impossible to answer in general. Answering the second is exactly the aim of Part II of this thesis.

Consider a hierarchy of integrable Hamiltonian differential equations, like the KdV hierarchy. Usually, each individual equation will also have a Lagrangian description, even though the Legendre transformation might not be invertible or even well-defined. In such a case one needs to be creative, for example by replacing the field variable with a potential and some educated guesswork, but it seems that somehow a variational description can always be found. Also, the Lax formulation can help to find a Lagrangian [100].

From the Hamiltonian point of view, what makes a hierarchy integrable is the fact that the Hamilton functions of the individual equations are in involution with respect to a Poisson bracket. What could be the Lagrangian equivalent of this statement? Suppose we embed our hierarchy in a higher-dimensional space, where each equation has its own time variable. In the case of the KdV hierarchy, we have coordinates $t_1 = x, t_2, t_3, \ldots, t_N$ and for each t_i with $i \geq 2$ there is a PDE expressing the t_i-derivative in terms of x-derivatives. Each equation has a Lagrangian; we denote these by $\mathcal{L}_{12}, \ldots, \mathcal{L}_{1N}$. A solution to the hierarchy will deliver critical values of the action

$$\int_{\Gamma_{1j}} \mathcal{L}_{1j} \, \mathrm{d}t_1 \wedge \mathrm{d}t_j \tag{6.1}$$

for each plane Γ_{1j} that is tangent to the t_1 and t_j directions.

In the pluri-Lagrangian context the Lagrange-function is replaced by a 2-form $\mathcal{L} = \sum_{i,j} \mathcal{L}_{ij} \, \mathrm{d}t_i \wedge \mathrm{d}t_j$. The pluri-Lagrangian principle requires that the action

$$\int_\Gamma \mathcal{L} = \int_\Gamma \sum_{i,j} \mathcal{L}_{ij} \, \mathrm{d}t_i \wedge \mathrm{d}t_j \tag{6.2}$$

is critical on every 2-manifold Γ simultaneously. The coefficients $\mathcal{L}_{1j}$ usually are the Lagrangians of the individual equations, but the additional coefficients $\mathcal{L}_{ij}$ with $i, j > 1$ do not correspond to any classical action principle. By choosing a coordinate plane for Γ we can recover the actions (6.1), but for other surfaces the criticality of the action (6.2) leads to Euler-Lagrange equations that are not implied by the classical variational principle.

At this point one may wonder if the pluri-Lagrangian principle really characterizes integrability. One might also doubt that there are nontrivial examples of pluri-Lagrangian systems. Even if we know a Lagrangian for each individual equation, finding suitable $\mathcal{L}_{ij}$ with $i, j > 1$ is a nontrivial task. As we will see, the Euler-Lagrange equations describing a pluri-Lagrangian system are massively overdetermined. On the bright side, being overdetermined suggests that if there are nontrivial examples, they probably belong to the realm of integrable systems. For the KdV hierarchy, a pluri-Lagrangian structure was found in [82].

The relation between pluri-Lagrangian structures and other notions of integrability is a subject of active investigation. This is beyond the scope of this thesis. Let us just mention a few relevant works on such connections. The relation of pluri-Lagrangian systems to variational symmetries is explored in [71] and [81], some initial links to the Hamiltonian theory of integrability are discussed in [80] and at the end of [82], and it is worth pointing out the example of pluri-Harmonic functions [11] [77, Section 4.4], which inspired the name.

Crucially, the pluri-Lagrangian principle can also be used for lattice systems. In fact the notion was born in the context of difference equations on quadrilateral graphs. The idea is perfectly analogous to the continuous one. The role of the Lagrangian is played by a discrete differential form, which is integrated over an arbitrary discrete surface in a higher-dimensional lattice. A solution must be a critical point of the action for all choices of the discrete surface.

Respecting the history of the pluri-Lagrangian concept – young as it may be – and general theme of this thesis, we will start our discussion with the discrete version of the theory before we move on to its continuous counterpart. First, let us have a look at the kind of equations that it applies to.

6.1. Meet our protagonists: the ABS list

Multidimensionally consistent *quad equations*, equations on quadrilateral graphs, were classified by Adler Bobenko and Suris in [4]. The list of equations they found is widely known as the ABS list. It classifies equations of the form

$$Q(U, U_1, U_2, U_{12}, \alpha_1, \alpha_2) = 0,$$

where subscript of the field $U : \mathbb{Z}^2 \to \mathbb{C}$ denote lattice shifts,

$$U = U(m, n), \qquad U_1 = U(m+1, n), \qquad U_2 = U(m, n+1), \qquad U_{12} = U(m+1, n+1),$$

and $\alpha_i \in \mathbb{C}$ are parameters associated to the lattice directions. Two quad equations are considered equivalent if they are related by a transformation of the parameters and a Möbius transformation of the fields.

The following properties are imposed on Q.

Linearity. We require that Q is affine in each of the fields, i.e.

$$\frac{\partial^2 Q}{\partial U^2} = \frac{\partial^2 Q}{\partial U_1^2} = \frac{\partial^2 Q}{\partial U_2^2} = \frac{\partial^2 Q}{\partial U_{12}^2} = 0.$$

This guarantees that the equation can be solved for any of the fields, given the other three.

Symmetry. The equation $Q = 0$ should be invariant under the symmetries of the square,

$$Q(U, U_1, U_2, U_{12}, \alpha_1, \alpha_2) = \pm Q(U, U_2, U_1, U_{12}, \alpha_2, \alpha_1) = \pm Q(U_1, U, U_{12}, U_2, \alpha_1, \alpha_2).$$

Multidimensional consistency. This is the property that makes our quad equations integrable. Even though the equations live on $\mathbb{Z}^2$ (or more generally on a planar quad graph), we require that we can consistently implement them on every square in a higher-dimensional lattice $\mathbb{Z}^d$. A necessary and sufficient condition for this is that the equation is *consistent around the cube*:

Given lattice parameters α_1, α_2, and α_3 and field values U, U_1, U_2, and U_3, we can use the equations

$$Q(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = 0, \qquad 1 \leq i < j \leq 3$$

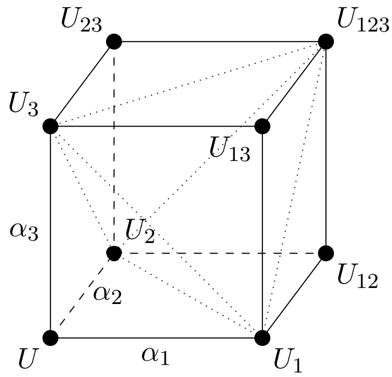


Figure 6.1. *A quad equation is consistent around the cube if U_{123} can be uniquely determined from U, U_1, U_2 and U_3. If in addition U_{123} is independent of U, then the equation satisfies the tetrahedron property.*

to determine U_{12}, U_{13}, and U_{23}. Then we can use each of the three equations

$$Q(U_i, U_{ij}, U_{ik}, U_{ijk}, \alpha_j, \alpha_k) = 0, \qquad (i,j,k) \equiv (i, i+1, i+2) \bmod 3$$

to determine U_{123}. If these three values agree (for all initial conditions U, U_1, U_2, and U_3 and all parameters α_1, α_2, and α_3), then the equation is called consistent around the cube.

Tetrahedron property. If a quad equation is consistent around the cube, the value of U_{123} can be considered as a function of U, U_1, U_2, U_3, and the lattice parameters,

$$U_{123} = z(U, U_1, U_2, U_3, \alpha_1, \alpha_2, \alpha_3).$$

If this function z does not depend on U we say that the equation has the tetrahedron property. (The corners of the cube corresponding to U_{123}, U_1, U_2, and U_3 form a regular tetrahedron.)

A sufficient condition for the tetrahedron property is that the equation can be written in a *three-leg form*,

$$Q(U, U_1, U_2, U_{12}, \alpha_1, \alpha_2) = \Psi(U, U_1, \alpha_1) - \Psi(U, U_2, \alpha_2) + \Phi(U, U_{12}, \alpha_1, \alpha_2).$$

Indeed, by symmetry this implies that up to sign

$$Q(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = \Psi(U_{ij}, U_j, \alpha_i) - \Psi(U_{ij}, U_i, \alpha_j) + \Phi(U_{ij}, U, \alpha_i, \alpha_j)$$

and summing up the three instances on the cube of this equation that contain U_{123}, we find

$$\Phi(U_{123}, U_3, \alpha_1, \alpha_2) + \Phi(U_{123}, U_1, \alpha_2, \alpha_3) + \Phi(U_{123}, U_2, \alpha_3, \alpha_1) = 0.$$

Assuming we can solve this for U_{123} we obtain the tetrahedron property.

Some years after the original ABS paper, the same authors published a classification of type Q equations (see below) where the assumptions symmetry and the tetrahedron property were dropped [5]. In addition it was no longer assumed that the equations on different faces of the cube are the same up to the parameter values.

Already in the original ABS paper [4] Lagrangians were given for each of the equations in the classification. However, it was some years later that Lobb and Nijhoff [47] formulated the variational structure in a multidimensional setting, giving rise to what we here call pluri-Lagrangian structures.

Below we list all the ABS equations. We postpone the presentation of their Lagrangians until we are ready to discuss their continuum limits.

6.1.1. Type Q

An important role in the classification of integrable quad equations is played by the six *biquadratics* corresponding to the equation $Q = 0$:

$$\begin{aligned}
&\frac{\partial Q}{\partial U}\frac{\partial Q}{\partial U_1} - Q\frac{\partial^2 Q}{\partial U \partial U_1}, &&\frac{\partial Q}{\partial U}\frac{\partial Q}{\partial U_2} - Q\frac{\partial^2 Q}{\partial U \partial U_2},\\
&\frac{\partial Q}{\partial U}\frac{\partial Q}{\partial U_{12}} - Q\frac{\partial^2 Q}{\partial U \partial U_{12}}, &&\frac{\partial Q}{\partial U_1}\frac{\partial Q}{\partial U_2} - Q\frac{\partial^2 Q}{\partial U_1 \partial U_2},\\
&\frac{\partial Q}{\partial U_1}\frac{\partial Q}{\partial U_{12}} - Q\frac{\partial^2 Q}{\partial U_1 \partial U_{12}}, &&\frac{\partial Q}{\partial U_2}\frac{\partial Q}{\partial U_{12}} - Q\frac{\partial^2 Q}{\partial U_2 \partial U_{12}}.
\end{aligned}$$

Due the the linearity assumption on Q, each of these is a quadratic function of two of the four field variables. A quad equation is said to be of *type Q* if all its biquadratics are nondegenerate, i.e. if they do not have a linear factor depending on only one of the variables. Furthermore, it is noteworthy that for all ABS equations of type Q a three-leg form can be found where the functions of the long and short legs coincide, $\Psi = \Phi$.

The ABS equations of type Q are:

Q1
$$\alpha_1(U - U_2)(U_1 - U_{12}) - \alpha_2(U - U_1)(U_2 - U_{12}) + \delta^2\alpha_1\alpha_2(\alpha_1 - \alpha_2) = 0. \tag{6.3}$$

With $\delta = 0$, it is known as the *cross-ratio equation* or the *lattice Schwarzian KdV equation*, see for example [60].

Q2
$$\begin{aligned}\alpha_1\alpha_2(\alpha_1 - \alpha_2)(U + U_1 + U_2 + U_{12}) - \alpha_1\alpha_2(\alpha_1 - \alpha_2)(\alpha_1^2 - \alpha_1\alpha_2 + \alpha_2^2)\\
+\alpha_1(U - U_2)(U_1 - U_{12}) - \alpha_2(U - U_1)(U_2 - U_{12}) = 0.\end{aligned} \tag{6.4}$$

It seems that this equation has only appeared in the literature in the context of the ABS list.

Q3
$$\begin{aligned}(\alpha_2^2 - \alpha_1^2)(UU_{12} + U_1U_2) + \alpha_2(\alpha_1^2 - 1)(UU_1 + U_2U_{12})\\
- \alpha_1(\alpha_2^2 - 1)(UU_2 - U_1U_{12}) - \frac{\delta^2}{4\alpha_1\alpha_2}(\alpha_1^2 - \alpha_2^2)(\alpha_1^2 - 1)(\alpha_2^2 - 1) = 0.\end{aligned} \tag{6.5}$$

A closely related lattice equation, now known as the *NQC equation*, first appeared in [64]. An explicit transformation between the NQC equation and Q3 with $\delta = 0$ is given in [62].

Q4
$$
\begin{aligned}
&A\big((U-b)(U_2-b)-(a-b)(c-b)\big)\big((U_1-b)(U_{12}-b)-(a-b)(c-b)\big)\\
&+B\big((U-a)(U_1-a)-(b-a)(c-a)\big)\big((U_2-a)(U_{12}-a)-(b-a)(c-a)\big)\\
&=ABC(a-b),
\end{aligned}
\tag{6.6}
$$

where

$$
\begin{aligned}
(a,A) &= \big(\wp(\alpha_1),\wp'(\alpha_1)\big)\\
(b,B) &= \big(\wp(\alpha_2),\wp'(\alpha_2)\big)\\
(c,C) &= \big(\wp(\alpha_2-\alpha_1),\wp'(\alpha_2-\alpha_1)\big)
\end{aligned}
$$

are points on the elliptic curve $\{A^2 = 4a^3 - g_2 a - g_3\}$, i.e. $\wp$ is the Weierstrass elliptic function. This equation first appeared in [2] in a rational parameterization. The formulation (6.6) in terms of an elliptic curve was introduced in [63].

Q4 is variously known as the *lattice Krichever-Novikov equation*, the *Adler equation*, and the *integrable master equation*. The last name is justified the following diagram of degenerations, as well as by its connection to several semi-discrete and continuous systems with a parameter on an elliptic curve [3].

$$
\begin{tikzcd}
 & \mathrm{Q3}_{\delta=1} \arrow[r] \arrow[dd] & \mathrm{Q3}_{\delta=0} \arrow[dd] & \\
\mathrm{Q4} \arrow[ur] \arrow[dr] & & & \\
 & \mathrm{Q2} \arrow[r] & \mathrm{Q1}_{\delta=1} \arrow[r] & \mathrm{Q1}_{\delta=0}
\end{tikzcd}
$$

6.1.2. Type H

H1
$$
(U-U_{12})(U_1-U_2)+\alpha_2-\alpha_1=0. \tag{6.7}
$$

Known as the *lattice potential KdV equation*, H1 is one of the oldest and most widespread of the ABS equations, going back at least as far as [64, 73]. Even earlier, a lattice version of the non-potential KdV equation was given by Hirota [39]. Using a non-autonomous transformation $V(n,m) = U(n,m)+n\lambda_1+m\lambda_2$ and the reparameterization $\alpha_i = \lambda_i^2$, it takes the form

$$
(\lambda_1+\lambda_2+V-V_{12})(\lambda_2-\lambda_1+V_1-V_2)-\lambda_1^2+\lambda_2^2=0,
$$

in which it was studied in detail for example in [94] and [60].

H2 $$(U - U_{12})(U_1 - U_2) + (\alpha_2 - \alpha_1)(U + U_1 + U_2 + U_{12}) + \alpha_2^2 - \alpha_1^2 = 0. \tag{6.8}$$

It seems that this equation has only appeared in the literature in the context of the ABS list.

H3 $$\alpha_1(UU_1 + U_2U_{12}) - \alpha_2(UU_2 + U_1U_{12}) + \delta(\alpha_1^2 - \alpha_2^2) = 0. \tag{6.9}$$

The full equation H3 first appeared in the ABS list. For $\delta = 0$ we can do a nonautonomous transformation $U(n,m) \mapsto i^{n+m}U(n,m)$ to obtain

$$\alpha_1(UU_1 - U_2U_{12}) - \alpha_2(UU_2 - U_1U_{12}) = 0, \tag{6.10}$$

which is known as the *lattice modified KdV equation* [60].

An additional transformation, $U(n,m) \mapsto U(n,m)^{(-1)^m}$, which breaks the symmetry, turns Equation (6.10) into the lattice sine-Gordon equation, which dates back to [40].

6.1.3. Type A

For the sake of completeness, we include the equations of type A. However, both of them can be reduced to a type Q equation by a nonautonomous change of variables. For this reason we will not consider A1 and A2 in the rest of this work

A1 $$\alpha_1(U + U_2)(U_1 + U_{12}) - \alpha_2(U + U_1)(U_2 + U_{12}) - \delta^2\alpha_1\alpha_2(\alpha_1 - \alpha_2) = 0.$$

A1 is related to Q1 by $U(n,m) \mapsto (-1)^{n+m}U(n,m)$.

A2 $$\begin{aligned}(\alpha_2^2 - \alpha_1^2)(UU_1U_2U_{12} + 1) + \alpha_2(\alpha_1^2 - 1)(UU_2 + U_1U_{12})&\\ - \alpha_1(\alpha_2^2 - 1)(UU_1 + U_2U_{12}) &= 0.\end{aligned}$$

A2 is related to Q3 with $\delta = 0$ by $U(n,m) \mapsto U(n,m)^{(-1)^{n+m}}$.

6.2. The discrete pluri-Lagrangian principle

Consider the lattice $\mathbb{Z}^N$ with basis vectors $\mathfrak{e}_1, \ldots, \mathfrak{e}_N$. To each lattice direction we associate a parameter $\alpha_i \in \mathbb{C}$. The equations we are interested in involve the values of a field $U : \mathbb{Z}^N \to \mathbb{C}$ on elementary squares in this lattice, or more generally, on d-dimensional *plaquettes*. Such a plaquette is a 2^d-tuple of lattice points that form an elementary hypercube. We denote it by

$$\square_{i_1,\ldots,i_d}(\mathbf{n}) = \left\{ \mathbf{n} + \varepsilon_1\mathfrak{e}_{i_1} + \ldots + \varepsilon_d\mathfrak{e}_{i_d} \,\middle|\, \varepsilon_k \in \{0,1\} \right\} \subset \mathbb{Z}^N,$$

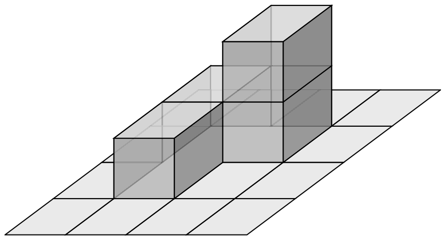


Figure 6.2. *Visualization of a discrete* 2*-surface in* $\mathbb{Z}^3$.

where $\mathbf{n} = (n_1, \ldots, n_N)$. Plaquettes are considered to be oriented; an odd permutation of the directions $i_1, \ldots, i_d$ reverses the orientation of the plaquette. We will write $U(\square_{i_1,\ldots,i_d}(\mathbf{n}))$ for the 2^d-tuple

$$U(\square_{i_1,\ldots,i_d}(\mathbf{n})) = \Big(U(\mathbf{n}), U(\mathbf{n} + \mathfrak{e}_{i_1}), U(\mathbf{n} + \mathfrak{e}_{i_2}), \ldots, U(\mathbf{n} + \mathfrak{e}_{i_1} \ldots + \mathfrak{e}_{i_d}) \Big).$$

Occasionally we will also consider the corresponding "filled-in" hypercubes in $\mathbb{R}^N$,

$$\blacksquare_{i_1,\ldots,i_d}(\mathbf{n}) = \Big\{ \mathbf{n} + \eta_1 \mathfrak{e}_{i_1} + \ldots + \eta_d \mathfrak{e}_{i_d} \,\Big|\, \eta_k \in [0,1] \Big\} \subset \mathbb{R}^N,$$

on which we consider the orientation defined by the volume form $\mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d}$.

The role of a Lagrange function is played by a discrete d-form

$$L(U(\square_{i_1,\ldots,i_d}(\mathbf{n})), \alpha_{i_1}, \ldots, \alpha_{i_d}),$$

which is a function of the values of the field $U : \mathbb{Z}^N \to \mathbb{C}$ on a plaquette and of the corresponding lattice parameters, where

$$L\big(U\left(\square_{\sigma(i_1),\ldots,\sigma(i_d)}(\mathbf{n})\right), \alpha_{\sigma(i_1)}, \ldots, \alpha_{\sigma(i_d)}\big) = \operatorname{sgn}(\sigma) L(\square_{i_1,\ldots,i_d}(\mathbf{n}), \alpha_{i_1}, \ldots, \alpha_{i_d})$$

for any permutation σ of $\{i_1, \ldots, i_d\}$. In other words, L is skew-symmetric with respect to the orientation of the plaquette.

Consider a discrete d-surface $\Gamma = \{\square_\alpha\}$ in the lattice, i.e. a set of d-dimensional plaquettes, such that the union of the corresponding filled-in plaquettes $\bigcup_\alpha \blacksquare_\alpha$ is an oriented topological d-manifold (possibly with boundary). The action over Γ is given by

$$S_\Gamma = \sum_{\square_{i_1,\ldots,i_d}(\mathbf{n}) \in \Gamma} L(U(\square_{i_1,\ldots,i_d}(\mathbf{n})), \alpha_{i_1}, \ldots, \alpha_{i_d}).$$

The field U is a solution to the *discrete pluri-Lagrangian problem* if it is a critical point of S_Γ (with respect to variations that are zero on the boundary of Γ) for all discrete d-surfaces Γ simultaneously.

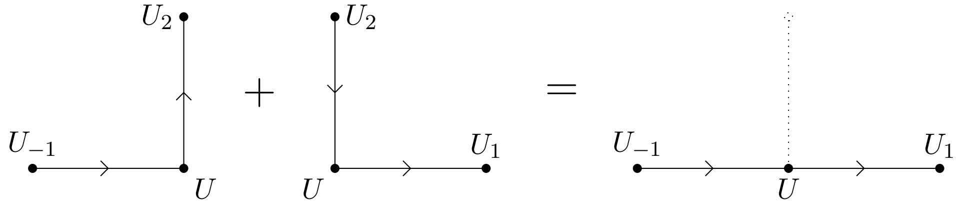


Figure 6.3. *A straight segment of a discrete curve is the sum of two corners.*

For $d = 1$ we have

$$S_\Gamma = \sum_{\{\mathbf{n},\mathbf{n}+\mathfrak{e}_i\}\in\Gamma} L(U(\mathbf{n}), U(\mathbf{n}+\mathfrak{e}_i), \alpha_i).$$

The Euler-Lagrange equations at general elementary corners,

$$\frac{\partial}{\partial U(\mathbf{n})}\Big(L(U(\mathbf{n}\pm\mathfrak{e}_i), U(\mathbf{n}), \alpha_i) + L(U(\mathbf{n}), U(\mathbf{n}\pm\mathfrak{e}_j), \alpha_j)\Big) = 0,$$

are sufficient conditions for U to be a solution to the pluri-Lagrangian problem. Indeed, any discrete curve locally consists of such corners. We can even restrict our attention to true corners, with $i \neq j$, because a straight segment can be considered as the sum of oriented corners, as in Figure 6.3.

For $d = 2$ we have

$$S_\Gamma = \sum_{\{\mathbf{n},\mathbf{n}+\mathfrak{e}_i,\mathbf{n}+\mathfrak{e}_j,\mathbf{n}+\mathfrak{e}_i+\mathfrak{e}_j\}\in\Gamma} L(U(\mathbf{n}), U(\mathbf{n}+\mathfrak{e}_i), U(\mathbf{n}+\mathfrak{e}_j), U(\mathbf{n}+\mathfrak{e}_i+\mathfrak{e}_j), \alpha_i, \alpha_j).$$

Since every discrete surface can be constructed out of corners of cubes, as illustrated in Figure 6.4, it is sufficient to determine the Euler-Lagrange equations on these elementary building blocks. They are

$$\frac{\partial}{\partial U}\Big(L(U,U_i,U_j,U_{ij},\alpha_i,\alpha_j) + L(U,U_j,U_k,U_{jk},\alpha_j,\alpha_k) + L(U,U_k,U_i,U_{ik},\alpha_k,\alpha_i) \Big) = 0, \tag{6.11}$$

$$\frac{\partial}{\partial U_i}\Big(L(U,U_i,U_j,U_{ij},\alpha_i,\alpha_j) - L(U_i,U_{ij},U_{ik},U_{ijk},\alpha_j,\alpha_k) + L(U,U_k,U_i,U_{ik},\alpha_k,\alpha_i) \Big) = 0, \tag{6.12}$$

$$\frac{\partial}{\partial U_{ij}}\Big(L(U,U_i,U_j,U_{ij},\alpha_i,\alpha_j) - L(U_i,U_{ij},U_{ik},U_{ijk},\alpha_j,\alpha_k) - L(U_j,U_{jk},U_{ij},U_{ijk},\alpha_k,\alpha_i)\Big) = 0, \tag{6.13}$$

$$\frac{\partial}{\partial U_{ijk}}\Big(-L(U_k,U_{ik},U_{jk},U_{ijk},\alpha_i,\alpha_j) - L(U_i,U_{ij},U_{ik},U_{ijk},\alpha_j,\alpha_k) - L(U_j,U_{jk},U_{ij},U_{ijk},\alpha_k,\alpha_i)\Big) = 0. \tag{6.14}$$

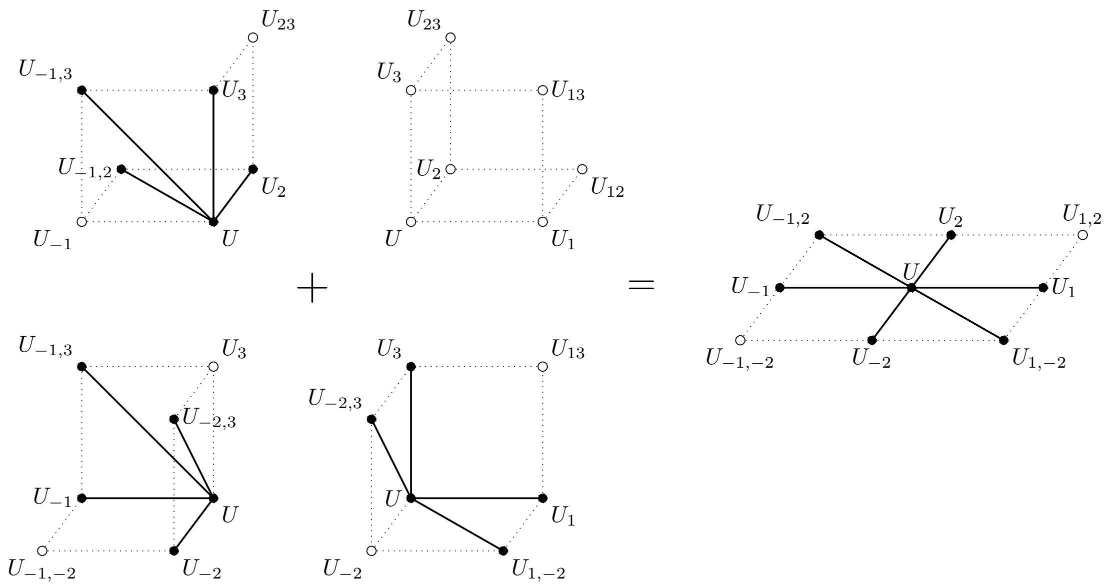


Figure 6.4. *A planar segment of a discrete surface is the sum of four corners. The marked vertices are those entering the Euler-Lagrange equations around U for a Lagrangian in triangle form, which leads to two copies of the quad equation in three-leg form.*

These *corner equations* are necessary and sufficient conditions for U to be a solution to the pluri-Lagrangian problem. Often, L can be written in a *triangle form*

$$L(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = A(U, U_i, \alpha_i) - A(U, U_j, \alpha_j) + B(U_i, U_j, \alpha_i - \alpha_j),$$

which renders the first and last corner equations, Equations (6.11) and (6.14), trivial. The triangle form of the Lagrangian is closely related to the three-leg form of the quad equation.

Example 6.1. Consider the following Lagrangian for H1, aka. the lattice potential KdV equation, aka. Equation (6.7),

$$L(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = UU_i - UU_j - (\alpha_i - \alpha_j)\log(U_i - U_j).$$

This Lagrangian was found in [15], but its multi-dimensional interpretation came much later [47]. Because of the triangle form, we only need to look at two of the corner equations. Equation (6.12) yields

$$\left(U - \frac{\alpha_i - \alpha_j}{U_i - U_j} - U_{ij} \right) + \left(U_{ik} - U + \frac{\alpha_k - \alpha_i}{U_k - U_i} \right) = 0$$

and Equation (6.13) gives

$$\left(-U_i + \frac{\alpha_j - \alpha_k}{U_{ij} - U_{ik}} + U_{ijk}\right) - \left(U_{ijk} - U_j + \frac{\alpha_k - \alpha_i}{U_{jk} - U_{ij}}\right) = 0.$$

In both of these corner equations we recognize two copies of Equation (6.7). Hence the Euler-Lagrange equations follow from the quad equation, but are not equivalent to it. This is the case for all ABS equations.

For the sake of brevity we will not give a more detailed overview of the theory of discrete pluri-Lagrangian systems here. Instead we refer to the groundbreaking paper [47], which introduced the pluri-Lagrangian (or Lagrangian multiform) idea, and to the reviews [13], [38, Chapter 12], and the references therein.

6.3. Continuous pluri-Lagrangian systems

6.3.1. A typical integrable hierarchy of PDEs

At first glance it is far from obvious what the continuous counterparts of the quad equations from the ABS list might be. One way to connect the discrete and continuous worlds is the observation that many quad equations arise as compatibility conditions of Bäcklund transformations for integrable differential equations, see for example [38, Section 2.5]. On the other hand, for some equations of the ABS list it is well-known that they produce certain integrable hierarchies of PDEs in a suitable continuum limit, but a systematic study of these continuum limits seems to be absent from the literature. Chapter 7 will provide such an overview, although its main goal will be to provide continuum limits of the pluri-Lagrangian structures involved.

The hierarchies we encounter are typically of the form

$$\begin{aligned} u_{t_2} &= f_2(u, u_x, u_{xx}, \ldots), \\ u_{t_3} &= f_3(u, u_x, u_{xx}, \ldots), \\ &\vdots \end{aligned}$$

where we identify $x = t_1$. A typical example is the potential Korteweg-de Vries hierarchy

$$\begin{aligned} u_{t_2} &= 0, \\ u_{t_3} &= 3u_x^2 + u_{xxx}, \\ u_{t_4} &= 0, \\ u_{t_5} &= 10u_x^3 + 5u_{xx}^2 + 10u_x u_{xxx} + u_{xxxxx}, \\ &\vdots \end{aligned}$$

In this hierarchy all equations for even-numbered time-variables are trivial and one could choose to omit them. However, they reflect the construction of the hierarchy through Lax pairs of pseudo-differential operators (see Section 9.1), and we will see the same pattern come out of the continuum limit.

The Korteweg-de Vries hierarchy has been central in the development of the modern theory of integrable systems. Solitons where first observed in the context of this equation, it was the first system the inverse scattering method was applied to, it was one of the first for which a bi-Hamiltonian structure was given, and so on. The history of the KdV equation has been discussed on every street corner of the integrable systems literature, for example in [42, 58, 70], so we will not go into detail here.

6.3.2. The continuous pluri-Lagrangian principle

In analogy to the discrete case, we define a continuous pluri-Lagrangian system as follows.

Definition 6.2. Let

$$\mathcal{L}[u] = \sum_{i_1 < \ldots < i_d} \mathcal{L}_{i_1,\ldots,i_d}[u] \, \mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d}.$$

be a d-form in $\mathbb{R}^N$, depending on a field $u : \mathbb{R}^N \to \mathbb{C}$ and any number of its derivatives. The field u solves the *continuous pluri-Lagrangian problem* for $\mathcal{L}$ if for any d-dimensional oriented submanifold $\Gamma \subset \mathbb{R}^N$ and any variation δu that vanishes near the boundary $\partial\Gamma$, there holds

$$\delta \int_\Gamma \mathcal{L}[u] = 0.$$

Some authors include in the definition that the pluri-Lagrangian form must be closed when evaluated on solutions. That would be equivalent to requiring that the action is not just critical on every surface, but even takes the same value on every surface with the same boundary and topology. In this perspective, one can take variations of the geometry as well as of the fields. We choose not to include the closedness in our definition, because we will obtain a slightly weaker property as a consequence of our definition in Proposition 6.4. Most of the authors that include closedness in the definition use the term "Lagrangian multiform", so one could argue that a Lagrangian multiform system is a pluri-Lagrangian system that satisfies the closedness property, giving the two names a marginally different meaning.

The first question about continuous pluri-Lagrangian systems is to find a set of equations characterizing criticality in the pluri-Lagrangian sense. We will call these equations the *multi-time Euler-Lagrange equations*. They were derived in [82] for $d = 1$ and $d = 2$ by approximating an arbitrary given curve or surface by a stepped curve or stepped surface, which are piecewise flat with tangent spaces spanned by coordinate directions. In Section 6.5 we will take a more intrinsic approach, which is less suited to deriving the equations from scratch, but very powerful to verify that they imply criticality. This approach relies on the variational bicomplex, which we introduce below.

6.4. The variational bicomplex

We allow the Lagrangian form $\mathcal{L}[u]$ to depend on any combination of partial derivatives of the field u, i.e. on the *infinite jet* of u. We will use a multi-index notation for partial derivatives of u. An *N-index* I is a N-tuple of non-negative integers. There is a natural bijection between N-indices and partial derivatives of $u : \mathbb{R}^N \to \mathbb{C}$. We denote by u_I the mixed partial derivative of u, where the number of derivatives with respect to each t_i is given by the entries of I. Note that if $I = (0, \dots, 0)$, then $u_I = u$.

We will often denote a multi-index suggestively by a string of t_i-variables, but it should be noted that this representation is not always unique. For example,

$$t_1 = (1, 0, \dots, 0), \qquad t_N = (0, \dots, 0, 1), \qquad t_1 t_2 = t_2 t_1 = (1, 1, 0, \dots, 0).$$

In this notation we will also make use of exponents to compactify the expressions, for example

$$t_2^3 = t_2 t_2 t_2 = (0, 3, 0, \dots, 0).$$

The notation It_j should be interpreted as concatenation in the string representation, hence it denotes the multi-index obtained from I by increasing the j-th entry by one. Finally, if the j-th entry of I is nonzero we say that I contains t_j, and write $I \ni t_j$.

We consider the function $u : \mathbb{R}^N \to \mathbb{C}$ as a section of the trivial fiber bundle $\mathbb{R}^N \times \mathbb{C}$. This bundle has coordinates $(t_1, \dots, t_N, u)$. The derivatives of u are contained in the infinite jet bundle $J^\infty(\mathbb{R}^N \times \mathbb{C})$, with coordinates $(t_i, u, u_{t_i}, u_{t_i t_j}, \dots)_{i,j,\dots \in \{1,\dots,N\}}$. To facilitate the variational calculus in the pluri-Lagrangian setting, it is useful to consider the variation operator δ as an exterior derivative, acting in the fibers of the jet bundle. We call δ the *vertical exterior derivative* and d, which acts in $\mathbb{R}^N$, the *horizontal exterior derivative*. Together they provide a double grading of the space $\Omega(J^\infty(\mathbb{R}^N \times \mathbb{C}))$ of differential forms on the jet bundle. An *(a,b)-form* is a differential $(a+b)$-form structured as

$$\omega^{a,b} = \sum f_{I_1,\dots,I_a,j_1,\dots,j_b}[u] \, \delta u_{I_1} \wedge \ldots \wedge \delta u_{I_a} \wedge \mathrm{d}t_{j_1} \ldots \wedge \mathrm{d}t_{j_b}.$$

We denote the space of (a,b)-forms by $\Omega^{(a,b)} \subset \Omega^{a+b}(J^\infty(\mathbb{R}^N \times \mathbb{C}))$. We call elements of $\Omega^{(0,b)}$ horizontal forms and elements of $\Omega^{(a,0)}$ vertical forms. The Lagrangian is a horizontal d-form, $\mathcal{L} \in \Omega^{(0,d)} \subset \Omega^d(J^\infty(\mathbb{R}^N \times \mathbb{C}))$.

The spaces $\Omega^{(a,b)}$ are related to each other by d and δ as in the following diagram, known as the *variational bicomplex*:

$$\begin{array}{ccccccccc}
\vdots & & \vdots & & & & \vdots & & \vdots \\
\uparrow{\scriptstyle\delta} & & \uparrow{\scriptstyle\delta} & & & & \uparrow{\scriptstyle\delta} & & \uparrow{\scriptstyle\delta} \\
\Omega^{(2,0)} & \xrightarrow{\mathrm{d}} & \Omega^{(2,1)} & \xrightarrow{\mathrm{d}} & \cdots & \xrightarrow{\mathrm{d}} & \Omega^{(2,n-1)} & \xrightarrow{\mathrm{d}} & \Omega^{(2,n)} \\
\uparrow{\scriptstyle\delta} & & \uparrow{\scriptstyle\delta} & & & & \uparrow{\scriptstyle\delta} & & \uparrow{\scriptstyle\delta} \\
\Omega^{(1,0)} & \xrightarrow{\mathrm{d}} & \Omega^{(1,1)} & \xrightarrow{\mathrm{d}} & \cdots & \xrightarrow{\mathrm{d}} & \Omega^{(1,n-1)} & \xrightarrow{\mathrm{d}} & \Omega^{(1,n)} \\
\uparrow{\scriptstyle\delta} & & \uparrow{\scriptstyle\delta} & & & & \uparrow{\scriptstyle\delta} & & \uparrow{\scriptstyle\delta} \\
\Omega^{(0,0)} & \xrightarrow{\mathrm{d}} & \Omega^{(0,1)} & \xrightarrow{\mathrm{d}} & \cdots & \xrightarrow{\mathrm{d}} & \Omega^{(0,n-1)} & \xrightarrow{\mathrm{d}} & \Omega^{(0,n)}
\end{array}$$

The horizontal and vertical exterior derivatives are characterized by the anti-derivation property,

$$\begin{aligned}
\mathrm{d}\left(\omega_1^{p_1,q_1} \wedge \omega_2^{p_2,q_2}\right) &= \mathrm{d}\omega_1^{p_1,q_1} \wedge \omega_2^{p_2,q_2} + (-1)^{p_1+q_1}\, \omega_1^{p_1,q_1} \wedge \mathrm{d}\omega_2^{p_2,q_2}, \\
\delta\left(\omega_1^{p_1,q_1} \wedge \omega_2^{p_2,q_2}\right) &= \delta\omega_1^{p_1,q_1} \wedge \omega_2^{p_2,q_2} + (-1)^{p_1+q_1}\, \omega_1^{p_1,q_1} \wedge \delta\omega_2^{p_2,q_2},
\end{aligned}$$

and by the way they act on $(0,0)$-forms, and basic $(1,0)$ and $(0,1)$-forms:

$$\begin{aligned}
\mathrm{d}f &= \sum_j \mathrm{D}_j f \,\mathrm{d}t_j, & \delta f &= \sum_I \frac{\partial f}{\partial u_I}\delta u_I, \\
\mathrm{d}(\delta u_I) &= -\sum_j \delta u_{Ij} \wedge \mathrm{d}t_j, & \delta(\delta u_I) &= 0, \\
\mathrm{d}(\mathrm{d}t_j) &= 0, & \delta(\mathrm{d}t_j) &= 0,
\end{aligned}$$

where $\mathrm{D}_j = \frac{\mathrm{d}}{\mathrm{d}t_j}$. As a simple but important example, note that

$$\mathrm{d}(f[u]\,\delta u_I) = \sum_{j=1}^N \mathrm{D}_j f[u]\,\mathrm{d}t_j \wedge \delta u_I - f[u]\,\delta u_{It_j} \wedge \mathrm{d}t_j = \sum_{j=1}^N -\mathrm{D}_j(f[u]\,\delta u_I) \wedge \mathrm{d}t_j.$$

One can verify that $\mathrm{d} + \delta : \Omega^{a+b} \to \Omega^{a+b+1}$ is the usual exterior derivative and that

$$\delta^2 = \mathrm{d}^2 = \delta\mathrm{d} + \mathrm{d}\delta = 0.$$

Furthermore, for any vertical vector field $V = \sum_I v^{[I]} \frac{\partial}{\partial u_I}$, there holds

$$\iota_V \mathrm{d} + \mathrm{d}\iota_V = 0.$$

More on the variational bicomplex can be found in [23, Chapter 19]. A slightly different version of the variational bicomplex – using contact forms instead of vertical forms –

is discussed in [6]. We will not discuss the rich algebraic structure of the variational bicomplex here.

For a horizontal $(0,d)$-form $\mathcal{L}[u]$, the variational principle

$$\delta \int_\Gamma \mathcal{L}[u] = \delta \int_\Gamma \sum_{i_1 < \ldots < i_d} \mathcal{L}_{i_1,\ldots,i_d}[u] \, \mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d} = 0$$

can be understood as follows. For every vector field $V = v(t_1,\ldots,t_a)\frac{\partial}{\partial u}$ that vanishes near the boundary $\partial\Gamma$, its *prolongation*

$$\operatorname{pr} V = \sum_I v_I \frac{\partial}{\partial u_I}$$

must satisfy

$$\int_\Gamma \iota_{\operatorname{pr} V} \delta \mathcal{L} = \int_\Gamma \sum_{i_1 < \ldots < i_d} \iota_{\operatorname{pr} V} \big(\delta \mathcal{L}_{i_1,\ldots,i_d}[u] \wedge \mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d} \big) = 0.$$

Note that the integrand is a horizontal form, so the integration takes place on $\Gamma \subset \mathbb{R}^N$, independent of the bundle structure. This formulation in itself might not be a terribly convenient characterization of criticality, but it provides a stepping stone toward a powerful tool, which we call *pluri-variational calculus.*

6.4.1. Pluri-Variational calculus

Just like total derivatives play in important role in the classical variational principle, exterior derivatives can be used to characterize criticality in the pluri-Lagrangian sense.

Proposition 6.3. *The field u is a critical point of the action $\int_\Gamma \mathcal{L}[u]$ for all Γ if and only if locally there exists a $(1,d-1)$-form Θ such that $\delta\mathcal{L}[u] = \mathrm{d}\Theta$.*

Proof. Assume such a $(1,d-1)$-form Θ exists. Since the horizontal exterior derivative d anti-commutes with the interior product operator $\iota_{\operatorname{pr} V}$ for the prolongation of a vertical vector field V, it follows that for any variation V of the field u that is zero near the boundary of a manifold Γ:

$$\int_\Gamma \iota_{\operatorname{pr} V} \delta \mathcal{L} = -\int_\Gamma \mathrm{d}\left(\iota_{\operatorname{pr} V} \Theta\right) = -\int_{\partial\Gamma} \iota_{\operatorname{pr} V} \Theta = 0.$$

Conversely, let $\mathcal{L}$ be evaluated on a critical field, and let $\Gamma = \partial B^{d+1}$ be a small d-sphere. Then for any prolonged vector field $\operatorname{pr} V$ there holds

$$\int_{B^{d+1}} \mathrm{d}(\iota_{\operatorname{pr} V} \delta \mathcal{L}) = \int_\Gamma \iota_{\operatorname{pr} V} \delta \mathcal{L} = 0.$$

Since the ball B^{d+1} can be arbitrarily chosen, it follows that $\mathrm{d}(\iota_{\mathrm{pr}\,V}\delta\mathcal{L}) = 0$. Hence, locally, we can find a $(0, d-1)$-form $\Omega(\mathrm{pr}\,V)$, depending on the vector field, such that

$$\iota_{\mathrm{pr}\,V}\delta\mathcal{L} = \mathrm{d}\Omega(\mathrm{pr}\,V).$$

Since $\Omega(\mathrm{pr}\,V)$ is linear in $\mathrm{pr}\,V$, it has to be of the form $\Omega(\mathrm{pr}\,V) = \iota_{\mathrm{pr}\,V}\Theta$ for some $(1, d-1)$-form Θ. It follows that

$$\iota_{\mathrm{pr}\,V}\delta\mathcal{L} = \mathrm{d}(\iota_{\mathrm{pr}\,V}\Theta) = -\iota_{\mathrm{pr}\,V}\mathrm{d}\Theta,$$

hence $\delta\mathcal{L} = -\mathrm{d}\Theta$. □

If we are dealing with a classical Lagrangian problem from mechanics, $\mathcal{L} = L(u, u_t)\mathrm{d}t$, we have $\Theta = \frac{\partial L}{\partial u_t}\delta u$, which is the pull back to the tangent bundle of the canonical 1-form $\sum_i p_i \mathrm{d}q_i$ on the cotangent bundle. In first order field theory Θ is the Cartan form [53]. In the pluri-Lagrangian context the form Θ can be used to show that multi-time Euler-Lagrange equations are indeed sufficient for criticality.

As mentioned in Section 6.3, we would like the Lagrangian form to be closed when evaluated on solutions. We did not include this in the definition of a pluri-Lagrangian system, because our definition already implies a slightly weaker property.

Proposition 6.4. *The horizontal exterior derivative $\mathrm{d}\mathcal{L}$ of a pluri-Lagrangian form is constant on connected components of the set of critical fields for $\mathcal{L}$.*

Proof. Critical points satisfy locally

$$\delta\mathcal{L} = \mathrm{d}\Theta \qquad \Rightarrow \qquad \mathrm{d}\delta\mathcal{L} = 0 \qquad \Rightarrow \qquad \delta\mathrm{d}\mathcal{L} = 0.$$

Hence for any vertical variation V the infinitesimal change of $\mathrm{d}\mathcal{L}$ along V is $\iota_{\mathrm{pr}\,V}\delta(\mathrm{d}\mathcal{L}) = 0$. □

Often the well-posedness of an initial or boundary value problem for the system of multi-time Euler-Lagrange equations implies that the set of critical points is connected. Then $\mathrm{d}\mathcal{L}$ is constant on the set of all critical fields. Furthermore, for many examples we have a simple vacuum solution, which we can use to verify that this constant is zero.

6.4.2. Discrete analogues

Let us have a closer look at discrete differential forms. Our discussion can be considered as a minimal version of the theory presented in [41, 51].

A *discrete d-form* is a skew-symmetric function η of the field variables and lattice parameters on a d-dimensional plaquette,

$$\eta\big(\square_{\sigma(i_1),\ldots,\sigma(i_d)}(\mathbf{n}), \alpha_{\sigma(i_1)}, \ldots, \alpha_{\sigma(i_d)}\big) = \mathrm{sgn}(\sigma)\, \eta(\square_{i_1,\ldots,i_d}(\mathbf{n}), \alpha_{i_1}, \ldots, \alpha_{i_d}).$$

The analogy between discrete and continuous differential forms goes much deeper than this. Also for discrete differential forms there is an exterior derivative. Let $\Omega^d(\mathbb{Z}^N)$ denote the space of discrete d-forms in the lattice $\mathbb{Z}^N$. Then $\Delta : \Omega^{d-1}(\mathbb{Z}^N) \to \Omega^d(\mathbb{Z}^N)$ is defined by

$$\begin{aligned}
&\Delta\eta(\square_{i_1,\ldots,i_d}(\mathbf{n}), \alpha_{i_1}, \ldots, \alpha_{i_d}) \\
&\quad = \sum_{k=1}^{d} (-1)^k \left(\eta\Big(\square_{j_1,\ldots,\widehat{j_k},\ldots,j_d}(\mathbf{n})\Big) - \eta\Big(\square_{j_1,\ldots,\widehat{j_k},\ldots,j_d}(\mathbf{n} + \mathfrak{e}_k)\Big) \right),
\end{aligned}$$

where the hat denotes and missing index and the parameters $\alpha_{i_1}, \ldots, \widehat{\alpha_{i_k}}, \ldots, \alpha_{i_d}$ have been omitted from the right hand side. Note that the sum is taken over set of $(d-1)$-faces of the of the d-dimensional cube $\square_{i_1,\ldots,i_d}(\mathbf{n})$, with the induced orientation. Applying the discrete exterior derivative twice always yields zero. Indeed, we have

$$\begin{aligned}
&\Delta\Delta\eta(\square_{i_1,\ldots,i_d}(\mathbf{n}), \alpha_{i_1}, \ldots, \alpha_{i_d}) \\
&= \sum_{k=1}^{d} (-1)^k \left(\Delta\eta\Big(\square_{j_1,\ldots,\widehat{j_k},\ldots,j_d}(\mathbf{n})\Big) - \Delta\eta\Big(\square_{j_1,\ldots,\widehat{j_k},\ldots,j_d}(\mathbf{n} + \mathfrak{e}_k)\Big) \right) \\
&= \sum_{k=1}^{d} \sum_{\ell=1}^{d} (-1)^{k+\ell} \operatorname{sgn}(k-\ell) \left(\eta\Big(\square_{j_1,\ldots,\widehat{j_k},\widehat{j_\ell},\ldots,j_d}(\mathbf{n})\Big) - \eta\Big(\square_{j_1,\ldots,\widehat{j_k},\widehat{j_\ell},\ldots,j_d}(\mathbf{n} + \mathfrak{e}_k)\Big) \right. \\
&\qquad \left. - \eta\Big(\square_{j_1,\ldots,\widehat{j_k},\widehat{j_\ell},\ldots,j_d}(\mathbf{n} + \mathfrak{e}_\ell)\Big) + \eta\Big(\square_{j_1,\ldots,\widehat{j_k},\widehat{j_\ell},\ldots,j_d}(\mathbf{n} + \mathfrak{e}_k + \mathfrak{e}_\ell)\Big) \right).
\end{aligned}$$

If we interchange k and ℓ in this expression, only the sign function $\operatorname{sgn}(k-\ell)$ changes, so the double sum evaluates to zero. Geometrically speaking, each $(d-2)$-face forms the boundary between two of the $(d-1)$-faces and hence occurs twice with opposite orientation.

Proposition 6.5. *For every discrete $(d-1)$-form η, the discrete d-form $L = \Delta\eta$ is a null Lagrangian, i.e. every field $\mathbb{Z}^N \mapsto \mathbb{C}$ solves the pluri-Lagrangian problem for L.*

Proof. Just like in the continuous case, the discrete integral of $\Delta\eta$ only depends on the boundary. Indeed, for any discrete d-surface Γ the contributions on interior $(d-1)$-faces cancel in the sum

$$\sum_{\Gamma} \Delta\eta = \sum_{\square_{j_1,\ldots,j_d}(\mathbf{n}) \in \Gamma} \sum_{k=1}^{d} (-1)^k \left(\eta\Big(\square_{j_1,\ldots,\widehat{j_k},\ldots,j_d}(\mathbf{n})\Big) - \eta\Big(\square_{j_1,\ldots,\widehat{j_k},\ldots,j_d}(\mathbf{n} + \mathfrak{e}_k)\Big) \right),$$

so the action $\sum_\Gamma \Delta\eta$ is always critical with respect to variations that vanish on the boundary of Γ. $\square$

In the discrete setting we have sums instead of integrals, so to arrive at a discrete analogue of Proposition 6.3 we would have to discretize the horizontal differential forms in the variational bicomplex and replace the horizontal operator d by Δ. However, a discrete version of this result would not be very useful. The discrete variational problem is much easier than the continuous one, because in the discrete case it is possible to take partial derivatives with respect to the value of the field at each lattice site independently. Hence using a discrete version of Proposition 6.3 to derive Euler-Lagrange equations would be much more complicated than a direct calculation. On the other hand, Proposition 6.5, which can be considered as a special case of such a result, will be useful in Chapter 8 to prepare discrete Lagrangians for the continuum limit procedure.

6.5. Multi-time Euler-Lagrange equations

In this section we discuss the multi-time Euler-Lagrange equations, which characterize fields that are critical in the pluri-Lagrangian sense. We start with the cases of 1-forms and 2-forms, before generalizing the discussion to d-forms for arbitrary d. We will only prove that the proposed equations are sufficient for criticality. For $d = 1$ and $d = 2$, a proof that they are necessary was given in [82].

6.5.1. One-forms

Consider the pluri-Lagrangian 1-form

$$\mathcal{L} = \sum_{j=1}^{N} \mathcal{L}_j[u] \, \mathrm{d}t_j,$$

depending on an arbitrary but finite number of derivatives of u.

Theorem 6.6. *The multi-time Euler-Lagrange equations are*

$$\frac{\delta_j \mathcal{L}_j}{\delta u_I} = 0 \qquad \forall j, \forall I \not\ni t_j, \tag{6.15}$$

$$\frac{\delta_j \mathcal{L}_j}{\delta u_{It_j}} - \frac{\delta_1 \mathcal{L}_1}{\delta u_{It_1}} = 0 \qquad \forall j \geq 2, \forall I, \tag{6.16}$$

where $\frac{\delta_j}{\delta u_I}$ *denotes the variational derivatives in the direction of* t_j *with respect to* u_I*,*

$$\frac{\delta_j}{\delta u_I} = \frac{\partial}{\partial u_I} - \mathrm{D}_j \frac{\partial}{\partial u_{It_j}} + \mathrm{D}_j^2 \frac{\partial}{\partial u_{It_j t_j}} - \cdots .$$

Equations (6.15) and (6.16) where originally derived in [82] by approximating any given curve by a stepped curve, i.e. a piecewise straight curve with all pieces in coordinate directions. Here we will only prove that they are sufficient for criticality in the pluri-Lagrangian sense, using Proposition 6.3 instead of a stepped surface approximation.

Proof of sufficiency. We calculate the vertical exterior derivative $\delta\mathcal{L}$ of the Lagrangian 1-form modulo Equations (6.15) and (6.16),

$$\begin{aligned}
\delta\mathcal{L} &= \sum_{j=1}^{N} \sum_{I} \frac{\partial \mathcal{L}_j}{\partial u_I} \delta u_I \wedge \mathrm{d}t_j \\
&= \sum_{j=1}^{N} \sum_{I} \left(\frac{\delta_j \mathcal{L}_j}{\delta u_I} + \mathrm{D}_j \frac{\delta_j \mathcal{L}_j}{\delta u_{It_j}} \right) \delta u_I \wedge \mathrm{d}t_j \\
&= \sum_{j=1}^{N} \left[\sum_{I \not\ni t_j} \frac{\delta_j \mathcal{L}_j}{\delta u_I} \delta u_I \wedge \mathrm{d}t_j + \sum_{I} \left(\frac{\delta_j \mathcal{L}_j}{\delta u_{It_j}} \delta u_{It_j} \wedge \mathrm{d}t_j + \left(\mathrm{D}_j \frac{\delta_j \mathcal{L}_j}{\delta u_{It_j}} \right) \delta u_I \wedge \mathrm{d}t_j \right) \right].
\end{aligned}$$

On solutions of Equation (6.16), we can define the generalized momenta

$$p^I = \frac{\delta_j L_j}{\delta u_{It_j}}.$$

Using Equations (6.15) and (6.16) it follows that

$$\delta\mathcal{L} = \sum_{j=1}^{N} \sum_{I} \left(p^I \delta u_{It_j} \wedge \mathrm{d}t_j + \left(\mathrm{D}_j p^I \right) \delta u_I \wedge \mathrm{d}t_j \right) = -d \left(\sum_{I} p^I \delta u_I \right).$$

This implies by Proposition 6.3 that Equations (6.15) and (6.16) are indeed sufficient for the action to be critical. □

6.5.2. Two-forms

Consider the pluri-Lagrangian 2-form

$$\mathcal{L} = \sum_{i<j} \mathcal{L}_{ij}[u] \, \mathrm{d}t_i \wedge \mathrm{d}t_j,$$

depending on an arbitrary but finite number of derivatives of u.

Theorem 6.7. *The multi-time Euler-Lagrange equations are*

$$\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_I} = 0 \qquad \forall I \not\ni t_i, t_j, \tag{6.17}$$

$$\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{It_j}} - \frac{\delta_{ik} \mathcal{L}_{ik}}{\delta u_{It_k}} = 0 \qquad \forall I \not\ni t_i, \tag{6.18}$$

$$\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{It_i t_j}} + \frac{\delta_{jk} \mathcal{L}_{jk}}{\delta u_{It_j t_k}} + \frac{\delta_{ki} \mathcal{L}_{ki}}{\delta u_{It_k t_i}} = 0 \qquad \forall I, \tag{6.19}$$

for i, j, and k distinct, where

$$\frac{\delta_{ij}}{\delta u_I} = \sum_{\alpha,\beta=0}^{\infty} (-1)^{\alpha+\beta} \mathrm{D}_i^\alpha \mathrm{D}_j^\beta \frac{\partial}{\partial u_{It_i^\alpha t_j^\beta}}$$

This theorem was proved in [82], analogous to the 1-form case, by approximating any given surface with a stepped surface.

Proof of sufficiency. We calculate the vertical exterior derivative $\delta\mathcal{L}$ modulo Equations (6.17)–(6.19). We temporarily break the symmetry and make t_1 a distinguished coordinate. Set

$$\begin{aligned} p_j^I &= \frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{It_1}} \qquad \text{for } I \not\ni t_j, \\ \pi_j^I &= \frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{It_1t_j}}. \end{aligned}$$

Since the coefficients L_{ij} are anti-symmetric, we have that $p_1^I = \pi_1^I = 0$.

We could also leave out the restriction $I \not\ni t_j$ in the definition of p_j^I, and write $p_j^{It_j}$ instead of π_j^I. However, in our notation the p_j^I are variational derivatives that occur in Equation (6.18), and the π_j^I occur in Equation (6.19). It is helpful to reflect this distinction in the notation.

We have

$$\begin{aligned} \delta\mathcal{L} &= \sum_{i<j} \sum_I \frac{\partial \mathcal{L}_{ij}}{\partial u_I} \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \\ &= \sum_{i<j} \sum_I \left(\frac{\delta_{ij}\mathcal{L}_{ij}}{\delta u_I} + \mathrm{D}_i \frac{\delta_{ij}\mathcal{L}_{ij}}{\delta u_{It_i}} + \mathrm{D}_j \frac{\delta_{ij}\mathcal{L}_{ij}}{\delta u_{It_j}} + \mathrm{D}_i\mathrm{D}_j \frac{\delta_{ij}\mathcal{L}_{ij}}{\delta u_{It_it_j}} \right) \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \end{aligned}$$

Modulo the multi-time Euler-Lagrange equations this becomes

$$\begin{aligned} \delta\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^{N} \Bigg[&\sum_{I \not\ni t_i,t_j} \frac{\delta_{ij}\mathcal{L}_{ij}}{\delta u_I} \, \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \\ &+ \sum_{I \not\ni t_j} \left(p_j^I \, \delta u_{It_i} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j + \mathrm{D}_i p_j^I \, \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \right) \\ &- \sum_{I \not\ni t_i} \left(p_i^I \, \delta u_{It_j} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j + \mathrm{D}_j p_i^I \, \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \right) \\ &+ \sum_I \Big((\pi_j^I - \pi_i^I) \, \delta u_{It_it_j} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j + \mathrm{D}_j(\pi_j^I - \pi_i^I) \, \delta u_{It_i} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \\ &\qquad + \mathrm{D}_i(\pi_j^I - \pi_i^I) \, \delta u_{It_j} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j + \mathrm{D}_i\mathrm{D}_j(\pi_j^I - \pi_i^I) \, \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \Big) \Bigg]. \end{aligned}$$

Using Equation (6.17) and the anti-symmetry, we can write this as

$$
\begin{aligned}
\delta \mathcal{L} = \sum_{i,j=1}^{N} \Bigg[& \sum_{I \not\ni t_j} \left(p_j^I \, \delta u_{I t_i} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j + \mathrm{D}_i p_j^I \, \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \right) \\
& + \sum_{I} \Big(\pi_j^I \, \delta u_{I t_i t_j} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j + \mathrm{D}_j \pi_j^I \, \delta u_{I t_i} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \\
& \qquad + \mathrm{D}_i \pi_j^I \, \delta u_{I t_j} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j + \mathrm{D}_i \mathrm{D}_j \pi_j^I \, \delta u_I \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \Big) \Bigg] \\
= \sum_{j=1}^{N} \Bigg[& \sum_{I \not\ni t_j} -\mathrm{d} \left(p_j^I \, \delta u_I \wedge \mathrm{d}t_j \right) + \sum_{I} -d \left(\pi_j^I \, \delta u_{I t_j} \wedge \mathrm{d}t_j + \mathrm{D}_j \pi_j^I \, \delta u_I \wedge \mathrm{d}t_j \right) \Bigg].
\end{aligned}
$$

The claim now follows by Proposition 6.3. □

Example 6.8. A pluri-Lagrangian 2-form for the potential KdV hierarchy was first given in [82] and an equivalent form will be derived in Chapter 8. Restricting the latter form to a minimal example we have

$$
\begin{aligned}
\mathcal{L}_{13} &= -2u_1^3 - u_1 u_{111} + u_1 u_3, \\
\mathcal{L}_{15} &= -5u_1^4 + 10 u_1 u_{11}^2 - u_{111}^2 + u_1 u_5,
\end{aligned}
$$

which are the classical Lagrangians of the potential KdV hierarchy. However, their classical Euler-Lagrange equations give the hierarchy only in a differentiated form,

$$
\begin{aligned}
u_{13} &= 6 u_1 u_{11} + u_{1111}, \\
u_{15} &= 30 u_1^2 u_{11} + 20 u_{11} u_{111} + 10 u_1 u_{1111} + u_{111111},
\end{aligned}
$$

The pluri-Lagrangian 2-form also contains a coefficient $\mathcal{L}_{35}$, which does not have a classical interpretation. (Skip ahead to Table 8.14 for an expression for $\mathcal{L}_{35}$.) However, it contributes meaningfully in the pluri-Lagrangian formalism. In particular, the multi-time Euler-Lagrange equations

$$
\frac{\delta_{13} \mathcal{L}_{13}}{\delta u_1} + \frac{\delta_{35} \mathcal{L}_{35}}{\delta u_5} = 0 \qquad \text{and} \qquad \frac{\delta_{15} \mathcal{L}_{15}}{\delta u_1} - \frac{\delta_{35} \mathcal{L}_{35}}{\delta u_3} = 0
$$

yield the KdV equations in their evolutionary form,

$$
\begin{aligned}
u_3 &= 3u_1^2 + u_{111}, \\
u_5 &= 10 u_1^3 + 5 u_{11}^2 + 10 u_1 u_{111} + u_{11111}.
\end{aligned}
$$

All other multi-time Euler-Lagrange equations are consequences of the hierarchy in this evolutionary form.

6.5.3. d-forms

To state the multi-time Euler-Lagrange equations for general d, we introduce some combinatorical notations. Consider the set

$$\mathcal{C}_{n,m} = \{\sigma : \mathbb{Z}^m \to \mathbb{Z}^n \mid \sigma(1) < \ldots < \sigma(m)\}$$

of combinations of m elements from $\{1, 2, \ldots, n\}$. We will also need nested combinations, where we first choose m elements from n and then ℓ from those m:

$$\mathcal{C}_{n,m,\ell} = \{\sigma : \mathbb{Z}^m \to \mathbb{Z}^n \mid \sigma(1) < \ldots < \sigma(\ell) \text{ and } \sigma(\ell+1) < \ldots < \sigma(m)\}.$$

Consider the pluri-Lagrangian d-form

$$\mathcal{L} = \sum_{i_1 < \ldots < i_d} \mathcal{L}_{i_1 \ldots i_d}[u] \, \mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d},$$

depending on an arbitrary number of derivatives of u. A multi-time Euler-Lagrange equation involves $d+1$ indices, $i_1 \ldots i_{d+1}$, and $\ell \leq d+1$ terms, which each contain exactly d of the indices,

$$i_1, \ldots i_{k-1}, \widehat{i_k}, i_{k+1}, \ldots, i_{d+1} \qquad \text{for } k \leq \ell,$$

where the hat denotes the missing index. We will use the following abbreviated notation for the relevant variational derivatives:

$$\delta^{i_1 \ldots i_d}_{I j_1 \ldots j_k} = \frac{\delta_{i_1 \ldots i_d} \mathcal{L}_{i_1 \ldots i_d}}{\delta u_{I t_{j_1} \ldots t_{j_k}}}.$$

Theorem 6.9. *For all $\ell \in \{1, \ldots d+1\}$ and for all $\sigma \in \mathcal{C}_{N,d+1,\ell}$ consider the equations*

$$\sum_{k=1}^{\ell} (-1)^k \delta^{\sigma(1) \ldots \widehat{\sigma(k)} \ldots \sigma(d+1)}_{I \sigma(1) \ldots \widehat{\sigma(k)} \ldots \sigma(\ell)} = 0 \qquad \forall I \not\ni t_{\sigma(\ell+1)}, \ldots, t_{\sigma(d+1)}. \tag{6.20}$$

All together, these equations are sufficient for criticality in the pluri-Lagrangian sense.

Before we proceed with the proof, a few comments are in order.

- To state one of the multi-time Euler-Lagrange equations (6.20), first fix the number of terms $\ell \leq d+1$, then choose $d+1$ indices that will occur in the equation, and finally pick ℓ of those that will be absent in one of the terms. This choice is represented by $\sigma \in \mathcal{C}_{N,d+1,\ell}$.
- The condition on I can be memorized as follows: if an index is present both in I and in the coefficients of the Lagrangian, then it must be one of the indices $\sigma(1), \ldots, \sigma(\ell)$ that do not occur in every term.

- If $\ell = 1$ we have only one term and the index $\sigma(1)$ is absent. This yields an Euler-Lagrange equation of classical type,

$$\delta_I^{\sigma(2)\dots\sigma(d+1)} = \frac{\delta \mathcal{L}_{\sigma(2)\dots\sigma(d+1)}}{\delta u_I} = 0, \qquad \forall I \not\ni t_{\sigma(2)}, \dots, t_{\sigma(d+1)}.$$

- Conjecturally, Equations (6.20) are also necessary for criticality. In any case they are a natural generalization of the multi-time Euler-Lagrange equations for $d = 1$ and $d = 2$.

Proof of Theorem 6.9. First observe that partial derivatives can be represented as sums of variational derivatives:

$$\frac{\partial \mathcal{L}_{i_1\dots i_d}}{\partial u_I} = \sum_{\varepsilon_1,\dots,\varepsilon_d=0}^{1} \mathrm{D}_{i_1}^{\varepsilon_1} \dots \mathrm{D}_{i_d}^{\varepsilon_d}\, \delta_{I i_1^{\varepsilon_1}\dots i_d^{\varepsilon_d}}^{i_1\dots i_d},$$

where the index i^ε is ignored if $\varepsilon = 0$ (and of course equal to i if $\varepsilon = 1$). Alternatively, we can write this as

$$\frac{\partial \mathcal{L}_{i_1\dots i_d}}{\partial u_I} = \sum_{\ell=0}^{d} \sum_{\sigma \in \mathcal{C}_{d,\ell}} \mathrm{D}_{i_{\sigma(1)}} \dots \mathrm{D}_{i_{\sigma(\ell)}}\, \delta_{I i_{\sigma(1)}\dots i_{\sigma(\ell)}}^{i_1\dots i_d}.$$

It follows that the vertical derivative of the Lagrangian d-form can be written as

$$\begin{aligned}
\delta \mathcal{L} &= \sum_{\sigma \in \mathcal{C}_{N,d}} \sum_I \frac{\partial \mathcal{L}_{\sigma(1)\dots\sigma(d)}}{\partial u_I}\, \delta u_I \wedge \mathrm{d}t_{\sigma(1)} \wedge \dots \wedge \mathrm{d}t_{\sigma(d)} \\
&= \sum_{\ell=0}^{d} \sum_{\sigma \in \mathcal{C}_{N,d,\ell}} \sum_I \left(\mathrm{D}_{\sigma(1)} \dots \mathrm{D}_{\sigma(\ell)}\, \delta_{I\sigma(1)\dots\sigma(\ell)}^{\sigma(1)\dots\sigma(d)} \right) \delta u_I \wedge \mathrm{d}t_{\sigma(1)} \wedge \dots \wedge \mathrm{d}t_{\sigma(d)}.
\end{aligned} \tag{6.21}$$

We claim that this is equivalent to

$$\delta \mathcal{L} = \sum_{\ell=0}^{d} \sum_{\sigma \in \mathcal{C}_{N,d,\ell}} \sum_{I \not\ni t_{\sigma(\ell+1)},\dots,t_{\sigma(d)}} \mathrm{D}_{\sigma(1)} \dots \mathrm{D}_{\sigma(\ell)} \left(\delta_{I\sigma(1)\dots\sigma(\ell)}^{\sigma(1)\dots\sigma(d)}\, \delta u_I \right) \wedge \mathrm{d}t_{\sigma(1)} \wedge \dots \wedge \mathrm{d}t_{\sigma(d)}. \tag{6.22}$$

To see the equivalence of Equations (6.21) and (6.22), observe that the expansion using the Leibniz rule of $\mathrm{D}_{\sigma(1)} \dots \mathrm{D}_{\sigma(\ell)} \left(\delta_{I\sigma(1)\dots\sigma(\ell)}^{\sigma(1)\dots\sigma(d)}\, \delta u_I \right)$ in Equation (6.22) yields terms from Equation (6.21), that none of those terms occur more than once in Equation (6.22), and that all of the terms in Equation (6.21) occur in the expansion of Equation (6.22).

Consider the special case of Equation (6.20), where $\sigma(1) = 1$, with ℓ replaced by $\ell + 1$:

$$-\delta_{I\sigma(2)\dots\sigma(\ell+1)}^{\sigma(2)\dots\sigma(d+1)} + \sum_{k=2}^{\ell+1} (-1)^k \delta_{I1\sigma(2)\dots\widehat{\sigma(k)}\dots\sigma(\ell+1)}^{1\sigma(2)\dots\widehat{\sigma(k)}\dots\sigma(d+1)} = 0.$$

After relabeling, it follows that

$$\delta^{\sigma(1)\ldots\sigma(d)}_{I\sigma(1)\ldots\sigma(\ell)} = \sum_{k=1}^{\ell} (-1)^{k+1} \delta^{1\sigma(1)\ldots\widehat{\sigma(k)}\ldots\sigma(d)}_{I1\sigma(1)\ldots\widehat{\sigma(k)}\ldots\sigma(\ell)}, \tag{6.23}$$

where $\sigma \in \mathcal{C}_{N,d,\ell}$ and $1 \notin \{\sigma(1), \ldots, \sigma(d)\}$. However, if $1 \in \{\sigma(1), \ldots, \sigma(d)\}$ in Equation (6.23), only the term where $\sigma(k) = 1$ will be nonzero in the right hand side and the identity holds trivially.

By virtue of equation (6.23), Equation (6.22) is equivalent to

$$\begin{aligned} \delta\mathcal{L} = -\sum_{\ell=0}^{d} \sum_{\sigma \in \mathcal{C}_{N,d,\ell}} \sum_{I \not\ni t_{\sigma(\ell+1)},\ldots,t_{\sigma(d)}} \\ \mathrm{D}_{\sigma(1)} \ldots \mathrm{D}_{\sigma(\ell)} \left(\sum_{k=1}^{\ell} (-1)^{k+1} \delta^{1\sigma(1)\ldots\widehat{\sigma(k)}\ldots\sigma(d)}_{I1\sigma(1)\ldots\widehat{\sigma(k)}\ldots\sigma(\ell)} \,\delta u_I \right) \wedge \mathrm{d}t_{\sigma(1)} \wedge \ldots \wedge \mathrm{d}t_{\sigma(d)}. \end{aligned}$$

In each term there is now a special index $j = \sigma(k)$, which occurs in the sequence of derivatives $\mathrm{D}_{\sigma(1)} \ldots \mathrm{D}_{\sigma(\ell)}$, but not in the string of indices $\sigma(1) \ldots \sigma(d)$. We rearrange the sum to highlight the role of this index,

$$\begin{aligned} \delta\mathcal{L} = \sum_{\ell=0}^{d} \sum_{\sigma \in \mathcal{C}_{N,d-1,\ell-1}} \sum_{I \not\ni t_{\sigma(\ell)},\ldots,t_{\sigma(d-1)}} \sum_{j=1}^{N} \\ - \mathrm{D}_j \mathrm{D}_{\sigma(1)} \ldots \mathrm{D}_{\sigma(\ell-1)} \left(\delta^{1\sigma(1)\ldots\sigma(d-1)}_{I1\sigma(1)\ldots\sigma(\ell-1)} \,\delta u_I \right) \wedge \mathrm{d}t_j \wedge \mathrm{d}t_{\sigma(1)} \wedge \ldots \wedge \mathrm{d}t_{\sigma(d)}. \end{aligned}$$

There is no need to specify in this sum that $j \notin \{\sigma(1), \ldots, \sigma(d)\}$, because terms that violate this condition vanish due to the skew-symmetry of the wedge product. We now recognize that $\delta\mathcal{L}$ is a horizontal exterior derivative,

$$\begin{aligned} \delta\mathcal{L} = \sum_{\ell=0}^{d} \sum_{\sigma \in \mathcal{C}_{N,d-1,\ell-1}} \sum_{I \not\ni t_{\sigma(\ell)},\ldots,t_{\sigma(d)}} \\ \mathrm{d} \left(\mathrm{D}_{\sigma(1)} \ldots \mathrm{D}_{\sigma(\ell-1)} \left(\delta^{1\sigma(1)\ldots\sigma(d-1)}_{I1\sigma(1)\ldots\sigma(\ell-1)} \,\delta u_I \right) \wedge \mathrm{d}t_{\sigma(1)} \wedge \ldots \wedge \mathrm{d}t_{\sigma(d)} \right), \end{aligned}$$

hence Proposition 6.3 implies that the pluri-Lagrangian principle is satisfied. □

7. Continuum limits of pluri-Lagrangian systems

This chapter is an adaptation of [89]

7.1. Miwa variables

To motivate our approach to the continuum limit, we start by considering the opposite direction[1]. The problem of integrable discretization has been studied at impressive length in the monograph [79]. Let us briefly summarize the "recipe" for discretizing Toda-type systems from Section 2.9 of that work. It starts from an integrable ODE with a Lax representation of the form

$$L_t = [L, \pi_+(f(L))] \tag{7.1}$$

in a Lie algebra $\mathfrak{g} = \mathfrak{g}_+ \oplus \mathfrak{g}_-$, where π_+ denotes projection onto $\mathfrak{g}_+$ and $f : \mathfrak{g} \to \mathfrak{g}$ is an Ad-covariant function. Here L denotes the Lax operator, not to be confused with a Lagrangian. Such an equation is part of an integrable hierarchy, given by

$$L_{t_k} = \Big[L, \pi_+\big(f(L)^k\big)\Big] \,. \tag{7.2}$$

A related integrable difference equation can be formulated in the corresponding Lie group G, with subgroups G_+ and G_- having Lie algebras $\mathfrak{g}_+$ and $\mathfrak{g}_-$ respectively. Any element $x \in G$ close to the unit $\mathrm{Id} \in G$ can be factorized as $x = \Pi_+(x)\Pi_-(x)$, where $\Pi_\pm(x) \in G_\pm$. The difference equation is given by

$$\widetilde{L} = \Pi_+(F(L))^{-1} \, L \, \Pi_+(F(L)), \tag{7.3}$$

where the tilde $\widetilde{\cdot}$ denotes a discrete time step and $F : \mathfrak{g} \to G$ approximates $L \mapsto \exp(\lambda L)$ for some small parameter λ. In particular, if G is a matrix group we can set

$$F(L) = \mathrm{Id} + \lambda f(L).$$

[1] This motivation was suggested by Yuri Suris.

Solutions of the differential equation (7.1) are given by

$$L(t) = \Pi_+\left(e^{tf(L_0)}\right)^{-1} L_0 \, \Pi_+\left(e^{tf(L_0)}\right).$$

A simultaneous solution to the whole hierarchy (7.2) takes the form

$$L(t_1, t_2, \ldots) = \Pi_+\left(e^{t_1 f(L_0) + t_2 f(L_0)^2 + \ldots}\right)^{-1} L_0 \, \Pi_+\left(e^{t_1 f(L_0) + t_2 f(L_0)^2 + \ldots}\right). \tag{7.4}$$

A solution of the discretization (7.3) is given by

$$\begin{aligned} L(n) &= \Pi_+(F^n(L_0))^{-1} L_0 \, \Pi_+(F^n(L_0)) \\ &= \Pi_+\left(e^{n \log(1+\lambda f(L_0))}\right)^{-1} L_0 \, \Pi_+\left(e^{n \log(1+\lambda f(L_0))}\right) \\ &= \Pi_+\left(e^{n\lambda f(L_0) - \frac{n}{2}\lambda^2 f(L_0)^2 + \ldots}\right)^{-1} L_0 \, \Pi_+\left(e^{n\lambda f(L_0) - \frac{n}{2}\lambda^2 f(L_0)^2 + \ldots}\right). \end{aligned} \tag{7.5}$$

Comparing equations (7.4) and (7.5), it is natural to identify a discrete step $n \mapsto n+1$ with the continuous time shift

$$(t_1, t_2, \ldots, t_i, \ldots) \mapsto \left(t_1 + \lambda, t_2 - \frac{\lambda^2}{2}, \ldots, t_i + (-1)^{i+1}\frac{\lambda^i}{i}, \ldots\right).$$

This gives us a map from the discrete space $\mathbb{Z}^N(n_1, \ldots, n_N)$ into the continuous multi-time $\mathbb{R}^N(t_1, \ldots, t_N)$: we associate a parameter λ_i to each lattice direction and set

$$t_i = (-1)^{i+1}\left(n_1 \frac{\lambda_1^i}{i} + \ldots + n_N \frac{\lambda_N^i}{i}\right).$$

Note that a single step in the lattice (changing one n_j) affects all the times t_i, hence we are dealing with a very skew embedding of the lattice. We will also consider a slightly more general correspondence,

$$t_i = (-1)^{i+1}\left(n_1 \frac{c\lambda_1^i}{i} + \ldots + n_N \frac{c\lambda_N^i}{i}\right) + \tau_i, \tag{7.6}$$

where constants $c, \tau_1, \ldots, \tau_N$ describe a scaling and a shift of the lattice. The variables n_j and λ_j are known in the literature as Miwa variables and have their origin in [55]. In the context of continuum limits we call the λ_j *lattice parameters*.

We call Equation (7.6) the *Miwa correspondence*. Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$ and consider the $N \times N$ matrix

$$M_{\boldsymbol{\lambda}} = \left((-1)^{i+1}\frac{\lambda_j^i}{i}\right)_{i,j=1}^N.$$

Then we can write the Miwa correspondence as

$$\mathbf{t} = cM_{\boldsymbol{\lambda}}\mathbf{n} + \boldsymbol{\tau},$$

where $\mathbf{t} = (t_1, \dots, t_N)^T$, $\mathbf{n} = (n_1, \dots, n_N)^T$, and $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)^T$. In other words, we consider the mesh $\mathbb{Z}^N$ under the affine transformation

$$A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}} : \mathbb{R}^N \to \mathbb{R}^N : \mathbf{t} \mapsto cM_{\boldsymbol{\lambda}}\mathbf{t} + \boldsymbol{\tau}. \tag{7.7}$$

We will use the Miwa correspondence (7.6) even if the discrete system is not generated by the recipe described above. In many cases one can justify this in a similar way by considering *plane wave factors*, which are solutions of the linearized system. For more on this perspective, see e.g. [60, 64, 94] and [38, Chapter 5]. For a completely different motivation for Miwa variables, note that for N distinct parameter values $\lambda_1, \dots, \lambda_N$ the corresponding vectors

$$\nu(\lambda) = \left(c\lambda, -\frac{c\lambda^2}{2}, \dots, (-1)^{N+1}\frac{c\lambda^N}{N} \right)$$

are linearly independent. Up to projective transformations, ν is the only curve with that property. It is known as the *rational normal curve* [36].

To perform the continuum limit of a difference equation involving $U : \mathbb{Z}^N \to \mathbb{C}$, we associate to it a function $u : \mathbb{R}^N \to \mathbb{C}$ that interpolates it:

$$U(\mathbf{n}) = u(A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\mathbf{n})) \qquad \forall \mathbf{n} \in \mathbb{Z}^N.$$

We denote the shift of U in the i-th lattice direction by U_i. If $U(\mathbf{n}) = u(t_1, \dots, t_N)$, it is given by

$$U_i = U(\mathbf{n} + \mathfrak{e}_i) = u\left(t_1 + c\lambda_i, t_2 - \frac{c\lambda_i^2}{2}, \dots, t_n - (-1)^N \frac{c\lambda_i^N}{N} \right),$$

which we can expand as a power series in λ_i. The difference equation thus turns into a power series in the lattice parameters. If all goes well, its coefficients will define differential equations that form an integrable hierarchy.

Note that such a procedure is strictly speaking not a continuum *limit*; sending $\lambda_i \to 0$ would only leave the leading order term of the power series. A more precise formulation is that the continuous u interpolates the discrete U (with a defect of sufficiently high order in λ_i), where U is defined on a mesh that is embedded in $\mathbb{R}^N$ using the Miwa correspondence. Since λ_i is assumed to be small, it makes sense to think of the outcome as a limit, but it is essential that higher order terms are not disregarded.

7.2. Continuum limits of Lagrangian forms

7.2.1. Modified Lagrangians in the classical variational problem

In Chapter 2 we performed a continuum limit on Lagrangian systems in the context of variational integrators for ODEs. Given a discrete Lagrangian, we constructed a continuous *modified Lagrangian* whose critical curves interpolate solutions of the discrete problem. A similar approach can be used in the context of pluri-Lagrangian systems, but first we present the relevant ideas in the context of the classical variational formulation of a partial difference equation. Here we use a fixed lattice $\mathbb{Z}^d \subset \mathbb{R}^d$. In Section 7.2.2 we will introduce parameters in the Lagrangian, transform the lattice according to the corresponding Miwa embedding, and consider the pluri-Lagrangian problem.

In the classical discrete variational principle we consider elementary plaquettes of full dimension, so it is sufficient to label them only by position, $\square(\mathbf{n})$, leaving out the subscripts denoting the direction. We consider Lagrangians $L_{\text{disc}}(U(\square(\mathbf{n})))$ depending on the values of the field $U : \mathbb{Z}^d \to \mathbb{C}$ on a plaquette $\square(\mathbf{n})$.

We identify points of a discrete solution with evaluations of an interpolating field $u : \mathbb{R}^d \to \mathbb{C}$, where we use a mesh size of 1 in all directions, i.e. $U(\mathbf{n}) = u(\mathbf{n})$ for $\mathbf{n} \in \mathbb{Z}^d$. Using a Taylor expansion we can write the discrete Lagrangian $L_{\text{disc}}(U(\square(\mathbf{n})))$ as a function of the interpolating field u and its derivatives,

$$\begin{aligned}
\mathcal{L}_{\text{disc}}[u]\Big|_{\mathbf{t}} &= L_{\text{disc}}\Big(\Big\{u(\mathbf{t} + \varepsilon_1 \mathfrak{e}_1 + \ldots + \varepsilon_d \mathfrak{e}_d) \,\Big|\, \varepsilon_1, \ldots, \varepsilon_d \in \{0,1\}\Big\}\Big) \\
&= L_{\text{disc}}\left(\left\{u + \sum_{k=1}^{d} \varepsilon_k u_{t_k} + \frac{1}{2}\sum_{k=1}^{d}\sum_{\ell=1}^{d} \varepsilon_k \varepsilon_\ell u_{t_k t_\ell} + \ldots \,\middle|\, \varepsilon_1, \ldots, \varepsilon_d \in \{0,1\}\right\}\right)\Bigg|_{\mathbf{t}},
\end{aligned}$$

where the square brackets denote dependence on u and any number of its partial derivatives, $\mathfrak{e}_1, \ldots, \mathfrak{e}_N$ are the unit vectors in the lattice $\mathbb{Z}^N$, and $\big|_{\mathbf{t}}$ indicates that all fields in the expression are evaluated at times $\mathbf{t} = (t_1, \ldots, t_d)$.

So far we have only written the discrete Lagrangian as a function of the continuous field. The corresponding action is still a sum:

$$S = \sum_{\mathbf{n} \in \mathbb{Z}^d} L_{\text{disc}}(U(\square(\mathbf{n}))) = \sum_{\mathbf{n} \in \mathbb{Z}^d} \mathcal{L}_{\text{disc}}[u(\mathbf{n})].$$

We want to write the action as an integral. This can be done using the Euler-Maclaurin formula [1, Sec. 23.1] (compare to Lemma 2.3),

$$\begin{aligned}
\sum_{k=0}^{m-1} F(a+k) &= \int_a^{a+m} F(t)\,\mathrm{d}t + \sum_{i=1}^{\infty} \frac{B_i}{i!}\left(F^{(i-1)}(a+m) - F^{(i-1)}(a)\right) \\
&= \int_a^{a+m} \left(\sum_{i=0}^{\infty} \frac{B_i}{i!} F^{(i)}(t)\right) \mathrm{d}t,
\end{aligned}$$

where B_i denote the Bernoulli numbers $1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \cdots$. This idea of turning sums into integrals was dubbed *anti-Taylor expansion* in [97]. Applying this to $\mathcal{L}_{\mathrm{disc}}$ in each of the lattice directions, we obtain the *meshed modified Lagrangian*

$$\mathcal{L}_{\mathrm{mesh}}[u] = \sum_{i_1,\ldots,i_d=0}^{\infty} \frac{B_{i_1}\ldots B_{i_d}}{i_1!\ldots i_d!} \mathrm{D}_{t_1}^{i_1}\ldots\mathrm{D}_{t_d}^{i_d}\mathcal{L}_{\mathrm{disc}}[u].$$

The series in the Euler-Maclaurin Formula generally does not converge. The same is true for the series defining $\mathcal{L}_{\mathrm{mesh}}$. Formally, it satisfies

$$S = \int_{\mathbb{R}^d} \mathcal{L}_{\mathrm{mesh}}[u(\mathbf{t})]\,\mathrm{d}t_1 \wedge \ldots \wedge \mathrm{d}t_d.$$

This property also holds locally,

$$L_{\mathrm{disc}}(U(\square(\mathbf{n}))) = \int_{\blacksquare(\mathbf{n})} \mathcal{L}_{\mathrm{mesh}}[u(\mathbf{t})]\,\mathrm{d}t_1 \wedge \ldots \wedge \mathrm{d}t_d. \tag{7.8}$$

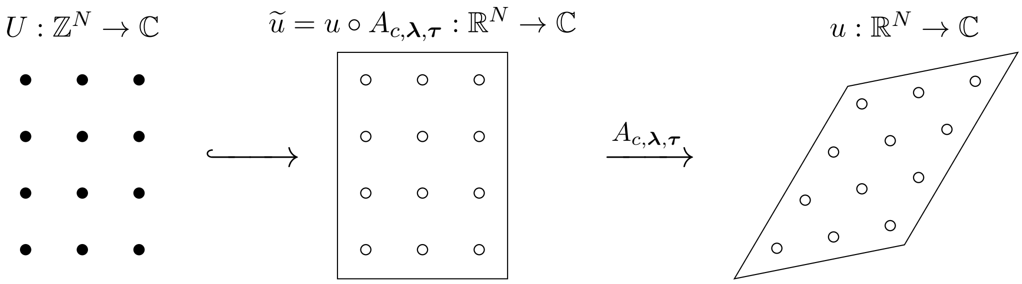
Due to the absence of lattice parameters, it is not straightforward to use truncation and the notion of k-criticality to avoid convergence issues, as we did in Chapter 2. However, once we switch from the straight lattice to a Miwa embedding, this strategy will be available again.

The word *meshed* refers to the fact that the discrete system provides additional structure for the continuous variational problem. In the *meshed variational problem*, non-differentiable fields are admissible as long as their singular points are consistent with the mesh, i.e. if they only occur on the boundaries of mesh cells. This imposes additional conditions on critical curves. In Chapter 2 these conditions were used to turn the meshed modified Lagrangian into a true modified Lagrangian which does not depend on higher derivatives. In the present context these conditions are less important, because we will find that the pluri-Lagrangian structure provides us with simpler tools to eliminate unwanted derivatives.

7.2.2. From discrete to continuous pluri-Lagrangian structures

In the pluri-Lagrangian context we consider a discrete Lagrangian d-form in a higher dimensional lattice $\mathbb{Z}^N$, $N > d$. The Lagrangian depends on lattice parameters, which are interpreted as Miwa variables. Consider N pairwise distinct lattice parameters $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$ and denote by $\mathfrak{e}_1, \ldots, \mathfrak{e}_N$ the unit vectors in the lattice $\mathbb{Z}^N$. The (differential of) the Miwa correspondence (7.6) maps them to linearly independent vectors in $\mathbb{R}^N$:

$$\mathfrak{e}_i \mapsto \mathfrak{v}_i = cM_{\boldsymbol{\lambda}}\mathfrak{e}_i = \left(c\lambda_i, -\frac{c\lambda_i^2}{2}, \ldots, (-1)^{N+1}\frac{c\lambda_i^N}{N} \right)^T.$$



$$L(U(\square(\boldsymbol{n}))) \quad = \quad \cdots \quad = \quad \mathcal{L}_{\mathrm{disc}}[u]\Big|_{A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\boldsymbol{n})}$$

$$\sum L(U(\square)) \quad = \quad \int_\Gamma \mathcal{L}_{\mathrm{mesh}}[\widetilde{u}]\,\mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d} \quad = \quad \int_{A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\Gamma)} \mathcal{L}_{\mathrm{Miwa}}[u]\,\eta_{i_1} \wedge \ldots \wedge \eta_{i_d}$$

Figure 7.1. *Visualisation of the lattice, the straight embedding from Section 7.2.1, and the skew embedding by the Miwa correspondence.*

We calculate the modified Lagrangian in the transformed coordinate system. We have

$$\begin{aligned}
&\mathcal{L}_{\mathrm{disc}}([u], \lambda_1, \ldots, \lambda_d)\Big|_{\mathbf{t}} \\
&= L_{\mathrm{disc}}\left(\left\{ u(\mathbf{t} + \varepsilon_1 \mathfrak{v}_1 + \ldots + \varepsilon_d \mathfrak{v}_d) \,\middle|\, \varepsilon_1, \ldots, \varepsilon_d \in \{0,1\} \right\}, \lambda_1, \ldots, \lambda_d \right) \\
&= L_{\mathrm{disc}}\left(\left\{ u + \sum_{k=1}^d \varepsilon_k \partial_k u + \frac{1}{2} \sum_{k=1}^d \sum_{\ell=1}^d \varepsilon_k \varepsilon_\ell \partial_k \partial_\ell u + \ldots \,\middle|\, \varepsilon_i \in \{0,1\} \right\}, \lambda_1, \ldots, \lambda_d \right) \Bigg|_{\mathbf{t}},
\end{aligned}$$

where now the differential operators correspond to the lattice directions under the Miwa correspondence,

$$\partial_k = \sum_{j=1}^N (-1)^{j+1} \frac{c\lambda_k^j}{j} \mathrm{D}_{t_j}.$$

Interpreted as vector fields, we can identify $\partial_k = \mathfrak{v}_k$ and $\mathrm{D}_{t_k} = \mathfrak{e}_k$, the former being the pushforward by $A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}$ of the latter.

The meshed modified Lagrangian in Miwa coordinates is given by

$$\begin{aligned}
\mathcal{L}_{\mathrm{Miwa}}([u], \lambda_1, \ldots, \lambda_d)\Big|_{\mathbf{t}} &= \mathcal{L}_{\mathrm{mesh}}([\widetilde{u}], \lambda_1, \ldots, \lambda_d)\Big|_{A^{-1}_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\mathbf{t})} \\
&= \sum_{i_1, \ldots, i_d = 0}^{\infty} \frac{B_{i_1} \ldots B_{i_d}}{i_1! \ldots i_d!} \partial_1^{i_1} \ldots \partial_d^{i_d} \mathcal{L}_{\mathrm{disc}}([u], \lambda_1, \ldots, \lambda_d)\Big|_{\mathbf{t}}.
\end{aligned}$$

The relation between U, $\widetilde{u}$ and u is illustrated in Figure 7.1. Note that, although

$\widetilde{u}\big|_{A^{-1}_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\mathbf{t})} = u\big|_{\mathbf{t}}$, no such equality holds for derivatives, hence

$$[\widetilde{u}]\Big|_{A^{-1}_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\mathbf{t})} \neq [u]\Big|_{\mathbf{t}}.$$

Due to the $\boldsymbol{\lambda}$-dependence of the operators $\partial_1, \ldots, \partial_N$, we can again use truncation of the power series in $\boldsymbol{\lambda}$ to obtain analytically meaningful results. However, it is convenient to keep working with formal series for the time being.

Lemma 7.1. *Consider a filled-in plaquette of the embedded lattice, $A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\blacksquare_{i_1,\ldots,i_d}(\mathbf{n}))$, and let η_k be the 1-forms dual to the Miwa shifts,*

$$\eta_k = (A^{-1}_{c,\boldsymbol{\lambda},\boldsymbol{\tau}})^* \mathrm{d}t_k,$$

where * denotes the pullback. Then $\mathcal{L}_{\mathrm{Miwa}}$ formally satisfies

$$\int_{A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\blacksquare_{i_1,\ldots,i_d}(\mathbf{n}))} \mathcal{L}_{\mathrm{Miwa}}([u], \lambda_{i_1}, \ldots, \lambda_{i_d}) \eta_{i_1} \wedge \ldots \wedge \eta_{i_d} = L_{\mathrm{disc}}(U(\square_{i_1,\ldots,i_d}(\mathbf{n})), \lambda_{i_1}, \ldots, \lambda_{i_d}).$$

Proof. In Equation (7.8) we have the corresponding result for $\mathcal{L}_{\mathrm{mesh}}$, so the proof is a simple change of variables. With $\widetilde{u} = u \circ A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}$, we have

$$\begin{aligned}
&\int_{A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\blacksquare_{i_1,\ldots,i_d}(\mathbf{n}))} \mathcal{L}_{\mathrm{Miwa}}([u], \lambda_{i_1}, \ldots, \lambda_{i_d})\Big|_{\mathbf{t}} \eta_{i_1} \wedge \ldots \wedge \eta_{i_d} \\
&= \int_{A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\blacksquare_{i_1,\ldots,i_d}(\mathbf{n}))} \mathcal{L}_{\mathrm{mesh}}([\widetilde{u}], \lambda_{i_1}, \ldots, \lambda_{i_d})\Big|_{A^{-1}_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\mathbf{t})} (A^{-1}_{c,\boldsymbol{\lambda},\boldsymbol{\tau}})^*(\mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d}) \\
&= \int_{\blacksquare_{i_1,\ldots,i_d}(\mathbf{n})} \mathcal{L}_{\mathrm{mesh}}([\widetilde{u}], \lambda_{i_1}, \ldots, \lambda_{i_d})\Big|_{\mathbf{t}} \mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d} \\
&= L_{\mathrm{disc}}(U(\square_{i_1,\ldots,i_d}(\mathbf{n})), \lambda_{i_1}, \ldots, \lambda_{i_d}). \qquad \square
\end{aligned}$$

We want to use this result for plaquettes in arbitrary directions. This suggests the Lagrangian d-form

$$\sum_{1 \leq i_1 < \ldots < i_d \leq N} \mathcal{L}_{\mathrm{Miwa}}([u], \lambda_{i_1}, \ldots, \lambda_{i_d})\, \eta_{i_1} \wedge \ldots \wedge \eta_{i_d}.$$

Up to a truncation error, this d-form can be written in a much more convenient way. Let $\mathcal{T}_N$ denote truncation of a power series after degree N in each variable,

$$\mathcal{T}_N\left(\sum_{i_1,\ldots,i_d=1}^{\infty} \lambda_1^{i_1} \ldots \lambda_d^{i_d} f_{i_1,\ldots,i_d} \right) = \sum_{i_1,\ldots,i_d=1}^{N} \lambda_1^{i_1} \ldots \lambda_d^{i_d} f_{i_1,\ldots,i_d}.$$

Lemma 7.2. *Assume that every term in the power series $\mathcal{L}_{\mathrm{Miwa}}$ is of strictly positive degree in each λ_i,*

$$\mathcal{L}_{\mathrm{Miwa}}([u],\lambda_1,\ldots,\lambda_d) = \sum_{i_1,\ldots,i_d=1}^{\infty} (-1)^{i_1+\ldots+i_d} c^d \frac{\lambda_1^{i_1}}{i_1} \ldots \frac{\lambda_d^{i_d}}{i_d} \mathcal{L}_{i_1,\ldots,i_d}[u], \tag{7.9}$$

then

$$\sum_{\substack{1\leq j_1<\ldots\\<j_d\leq N}} \mathcal{T}_N\big(\mathcal{L}_{\mathrm{Miwa}}([u],\lambda_{j_1},\ldots,\lambda_{j_d})\big)\, \eta_{j_1}\wedge\ldots\wedge\eta_{j_d} = \sum_{\substack{1\leq i_1<\ldots\\<i_d\leq N}} \mathcal{L}_{i_1,\ldots,i_d}[u]\,\mathrm{d}t_{i_1}\wedge\ldots\wedge\mathrm{d}t_{i_d}.$$

Note in Equation (7.9) that the factors $(-1)^{i_1+\ldots+i_d} c^d \frac{\lambda_1^{i_1}}{i_1} \ldots \frac{\lambda_d^{i_d}}{i_d}$ are terms of $(d\times d)$-minors of the transformation matrix $cM_{\boldsymbol{\lambda}}$.

Proof of Lemma 7.2. First observe that, just like the original discrete Lagrangian, the Lagrangian $\mathcal{L}_{\mathrm{Miwa}}([u],\lambda_{i_1},\ldots,\lambda_{i_d})$ is skew-symmetric as a function of $(\lambda_{i_1},\ldots,\lambda_{i_d})$. Therefore, the coefficients $\mathcal{L}_{i_1,\ldots,i_d}[u]$ are skew-symmetric as a function of $(i_1,\ldots,i_d)$.

We pair the form

$$\begin{aligned}\mathcal{L} &= \sum_{1\leq i_1<\ldots<i_d\leq N} \mathcal{L}_{i_1,\ldots,i_d}[u]\,\mathrm{d}t_{i_1}\wedge\ldots\wedge\mathrm{d}t_{i_d}\\ &= \frac{1}{d!}\sum_{i_1,\ldots,i_d=1}^{N} \mathcal{L}_{i_1,\ldots,i_d}[u]\,\mathrm{d}t_{i_1}\wedge\ldots\wedge\mathrm{d}t_{i_d}\end{aligned}$$

with a d-tuple of vectors $(\mathfrak{v}_{j_1},\ldots,\mathfrak{v}_{j_d}) = (cM_{\boldsymbol{\lambda}}\mathfrak{e}_{j_1},\ldots,cM_{\boldsymbol{\lambda}}\mathfrak{e}_{j_d})$:

$$\big\langle \mathcal{L},(\mathfrak{v}_{j_1},\ldots,\mathfrak{v}_{j_d})\big\rangle = \frac{1}{d!}\sum_{i_1,\ldots,i_d=1}^{N}\left(\mathcal{L}_{i_1,\ldots,i_d}[u]\sum_{\sigma\in S_d}\left(\mathrm{sgn}(\sigma)\prod_{k=1}^{d}\left\langle \mathrm{d}t_{i_{\sigma(k)}},\mathfrak{v}_{j_k}\right\rangle\right)\right).$$

Due to the skew-symmetry of $\mathcal{L}_{i_1,\ldots,i_d}[u]$, this can be written as

$$\big\langle \mathcal{L},(\mathfrak{v}_{j_1},\ldots,\mathfrak{v}_{j_d})\big\rangle = \frac{1}{d!}\sum_{i_1,\ldots,i_d=1}^{N}\sum_{\sigma\in S_d}\left(\mathcal{L}_{i_{\sigma(1)},\ldots,i_{\sigma(d)}}[u]\prod_{k=1}^{d}\left\langle \mathrm{d}t_{i_{\sigma(k)}},\mathfrak{v}_{j_k}\right\rangle\right).$$

Since the first sum is over all d-tuples $(i_1,\ldots,i_d)$ with strictly positive integer entries, permuting $(i_1,\ldots,i_d)$ yields a different term of this sum. Hence the additional summation over permutations $\sigma \in S_d$ amounts to multiplication by $d!$. We find

$$\begin{aligned}\left\langle \mathcal{L}, (\mathfrak{v}_{j_1}, \ldots, \mathfrak{v}_{j_d}) \right\rangle &= \sum_{i_1,\ldots,i_d=1}^{N} \mathcal{L}_{i_1,\ldots,i_d}[u] \prod_{k=1}^{d} \langle \mathrm{d}t_{i_k}, \mathfrak{v}_{j_k} \rangle \\ &= \sum_{i_1,\ldots,i_d=1}^{N} \mathcal{L}_{i_1,\ldots,i_d}[u] \prod_{k=1}^{d} (-1)^{i_k} c \frac{\lambda_{j_k}^{i_k}}{i_k} \\ &= \mathcal{T}_N\big(\mathcal{L}_{\mathrm{Miwa}}([u], \lambda_{j_1}, \ldots, \lambda_{j_d})\big). \qquad \square\end{aligned}$$

In contrast to Chapter 2, the discrete Lagrangian is not a consistent numerical discretization of some continuous function. Indeed, its leading terms are of the order $\lambda_{j_1}\lambda_{j_2}^2 \ldots \lambda_{j_d}^d$. Therefore we need a slightly different definition of k-criticality. Instead of requiring consistency, we now state the order condition relative to the magnitude of the action

Definition 7.3. Let Γ be a finite discrete surface in $\mathbb{Z}^N$. A discrete field $U : \mathbb{Z}^N \mapsto \mathbb{C}$ is *k-critical* for the action

$$S = \sum_{\square_{i_1,\ldots,i_d}(\mathbf{n}) \in \Gamma} L\big(U(\square_{i_1,\ldots,i_d}(\mathbf{n})), \lambda_{i_1}, \ldots, \lambda_{i_d}\big).$$

if for any $\mathbf{n} \in \mathbb{Z}^N$ there holds

$$\frac{\partial S}{\partial U(\mathbf{n})} = \mathcal{O}\big((\lambda_1^{k+1} + \ldots + \lambda_N^{k+1})|S|\big).$$

On the continuous side we do not need to rephrase the definition of k-criticality. In fact we will not need any notion of k-criticality on the continuous side, because we do not want the lattice parameters to survive in the continuum limit. And if u does not depend on the λ_j, then being k-critical for any nonnegative k implies being exactly critical.

Now that we have updated our understanding of k-criticality, we are finally ready to prove that a pluri-Lagrangian d-form can be constructed from the coefficients of the power series $\mathcal{L}_{\mathrm{Miwa}}$.

Theorem 7.4. *Let L_{disc} be a discrete Lagrangian d-form, such that every term in the corresponding power series $\mathcal{L}_{\mathrm{Miwa}}$ is of strictly positive degree in each λ_i, i.e. such that $\mathcal{L}_{\mathrm{Miwa}}$ is of the form (7.9). Consider the differential d-form*

$$\mathcal{L}[u] = \sum_{1 \leq i_1 < \ldots < i_d \leq N} \mathcal{L}_{i_1,\ldots,i_d}[u] \,\mathrm{d}t_{i_1} \wedge \ldots \wedge \mathrm{d}t_{i_d},$$

built out of the coefficients of $\mathcal{L}_{\mathrm{Miwa}}$. Then a field $u : \mathbb{R}^N \to \mathbb{C}$ is a solution to the continuous pluri-Lagrangian problem for $\mathcal{L}[u]$ if and only if the corresponding discrete fields

$$U_{\boldsymbol{\tau}} : \mathbb{Z}^N \to \mathbb{C} : \mathbf{n} \mapsto u(A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\mathbf{n})), \qquad \boldsymbol{\tau} \in \mathbb{R}^N,$$

are N-critical for the discrete pluri-Lagrangian problem for L_{disc}.

Proof. Consider a bounded d-surface Γ in $\mathbb{R}^N$ that does not depend on $\boldsymbol{\lambda}$. We can approximate it by an image of a discrete surface $\overline{\Gamma}$ under the Miwa embedding, with an error of order $\mathcal{O}(\lambda_1 + \ldots + \lambda_N)$,

$$\int_\Gamma \mathcal{L}[u] = \int_{A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\overline{\Gamma})} \mathcal{L}[u] + \mathcal{O}(\lambda_1 + \ldots + \lambda_N). \tag{7.10}$$

This idea of approximating any given surface by a *stepped surface* was used in [82] to derive the multi-time Euler-Lagrange equations.

We have

$$\int_{A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\overline{\Gamma})} \mathcal{L}[u] = \sum_{\square_{i_1,\ldots,i_d}(\mathbf{n})\in\overline{\Gamma}} L_{\mathrm{disc}}(U_{\boldsymbol{\tau}}(\square_{i_1,\ldots,i_d}(\mathbf{n})), \lambda_{i_1}, \ldots, \lambda_{i_d}) + \mathcal{O}\big(\lambda_1^{N+1} + \ldots + \lambda_N^{N+1}\big), \tag{7.11}$$

hence if the continuous field u is critical, then the discrete field $U_{\boldsymbol{\tau}}$ is N-critical.

From Equations (7.10) and (7.11) it follows that

$$\int_\Gamma \mathcal{L}[u] = L_{\mathrm{disc}}(U_{\boldsymbol{\tau}}(\square_{i_1,\ldots,i_d}(\mathbf{n})), \lambda_{i_1}, \ldots, \lambda_{i_d}) + \mathcal{O}(\lambda_1 + \ldots + \lambda_N).$$

Now assume that the discrete field is 0-critical. Then for any $\boldsymbol{\lambda}$-independent variation of u that is zero near the boundary of Γ, we have that

$$\delta \int_\Gamma \mathcal{L}[u] = \mathcal{O}(\lambda_1 + \ldots + \lambda_N).$$

Since the left hand side is independent of $\boldsymbol{\lambda}$, it must be exactly zero. Hence u is a critical field. □

Note that we did not just prove that discrete N-criticality is equivalent to continuous criticality, but also that discrete 0-criticality *implies* continuous criticality. Hence if a discrete field, obtained from a $\boldsymbol{\lambda}$-independent continuous field u by the relation $U_{\boldsymbol{\tau}}(\mathbf{n}) = u(A_{c,\boldsymbol{\lambda},\boldsymbol{\tau}}(\mathbf{n}))$, is just 0-critical, then it is automatically N-critical. Of course this does not hold for arbitrary discrete fields.

7.2.3. Eliminating alien derivatives

Unlike in the classical Lagrangian framework, Euler-Lagrange equations in the pluri-Lagrangian context are often *evolutionary*, i.e. of the form

$$u_{t_k} = f_k[u] \qquad \text{for } k \in \{d, d+1, \ldots, N\}$$

for a pluri-Lagrangian d-form in $\mathbb{R}^N$, where the f_k only depend on derivatives with respect to $t_1, \ldots, t_{d-1}$. If this is the case, then the differential consequences of the multi-time Euler-Lagrange equations can be written in a similar form,

$$u_I = f_I[u] \qquad \text{with } I \ni t_k \text{ for some } k \in \{d, d+1, \ldots, N\}, \tag{7.12}$$

where the I in f_I is a label, not a partial derivative. In this context it is natural to consider the first $d-1$ coordinates $t_1, \ldots, t_{d-1}$ as space coordinates and the others as time coordinates. If the multi-time Euler-Lagrange equations are not evolutionary, Equation (7.12) still holds for a slightly smaller set of multi-indices I.

Definition 7.5. A mixed partial derivative u_I is called $\{i_1, \ldots, i_d\}$*-native* if each individual derivative is taken with respect to one of the $t_{i_1}, \ldots, t_{i_d}$ or with respect to one of the space coordinates $t_1, \ldots, t_{d-1}$, i.e. if

$$t_k \in I \quad \Rightarrow \quad k \in \{1, \ldots, d-1, i_1, \ldots, i_d\}.$$

If u_I is not $\{i_1, \ldots, i_d\}$-native, i.e. if there is a $k \notin \{1, \ldots, d-1, i_1, \ldots, i_d\}$ such that $t_k \in I$, then we say u_I is $\{i_1, \ldots, i_d\}$*-alien*.

If it is clear what the relevant indices are, for example when discussing a coefficient $\mathcal{L}_{i_1,\ldots,i_d}$, we will use *native* and *alien* without mentioning the indices.

We would like the coefficient $\mathcal{L}_{i_1,\ldots,i_d}$ to contain only native derivatives. A naive approach would be to use the multi-time Euler-Lagrange equations (7.12) to eliminate alien derivatives. Let $R_{i_1,\ldots,i_d}$ denote the operator that replaces all $\{i_1, \ldots, i_d\}$-alien derivatives for which the multi-time Euler-Lagrange equations provide an expression. We denote the native version of the pluri-Lagrangian coefficients by

$$\overline{\mathcal{L}}_{i_1,\ldots,i_d} = R_{i_1,\ldots,i_d}(\mathcal{L}_{i_1,\ldots,i_d})$$

and the d-form with these coefficients by $\overline{\mathcal{L}}$. A priori there is no reason to believe that the d-form $\overline{\mathcal{L}}$ will be equivalent to the original pluri-Lagrangian d-form $\mathcal{L}$. For example, the 1-dimensional Lagrangian $\mathcal{L}(u, u_t, u_{tt}) = \frac{1}{2} u u_{tt}$ leads to the Euler-Lagrange equation $u_{tt} = 0$, but any curve is critical for the Lagrangian $\overline{\mathcal{L}}(u, u_t, u_{tt}) = 0$. However, in many cases the pluri-Lagrangian structure guarantees that $\mathcal{L}$ and $\overline{\mathcal{L}}$ have the same critical fields.

Theorem 7.6. *Assume that $R_{i_1,\ldots,i_d}$ commutes with the operators $\mathrm{D}_{t_{i_1}}, \ldots, \mathrm{D}_{t_{i_d}}$. If either*

- *$d = 1$ and $\mathcal{L}_1[u]$ does not depend on any alien derivatives, or*
- *$d = 2$ and for all j the coefficient $\mathcal{L}_{1j}[u]$ does not contain any alien derivatives,*

then every critical field u for the pluri-Lagrangian d-form $\mathcal{L}$ is also critical for $\overline{\mathcal{L}}$.

In particular, the commutativity condition holds if the equations in the hierarchy are evolutionary, or more generally, if none of their left hand sides are a mixed partial derivative. The condition for $d = 2$ might seem restrictive, but given a Lagrangian 2-form, we can often find an equivalent one with coefficients $\mathcal{L}_{1j}[u]$ that satisfy this condition by inspection.

Proof of Theorem 7.6. In this proof we consider the variation operator δ as the vertical exterior derivative in the variational bicomplex, see Section 6.4.

First we consider the case $d = 1$. Let

$$F_{i,J}[u] = R_i(u_J),$$

i.e. $F_{i,J}$ is obtained from u_j by eliminating as many alien derivatives as possible. Note that D_{t_i} and R_i commute, hence $\mathrm{D}_{t_i} F_{i,J} = F_{i,Jt_i}$. We have

$$\begin{aligned}
\delta\overline{\mathcal{L}} &= \sum_{1\le i\le N} \sum_J R_i\left(\frac{\partial \mathcal{L}_i}{\partial u_J}\right) \delta F_{i,J} \wedge \mathrm{d}t_i \\
&= \sum_{1\le i\le N} \sum_J R_i\left(\frac{\delta_i \mathcal{L}_i}{\delta u_J} + \mathrm{D}_{t_i} \frac{\delta_i \mathcal{L}_i}{\delta u_{Jt_i}}\right) \delta F_{i,J} \wedge \mathrm{d}t_i \\
&= \sum_{1\le i\le N} \left(\sum_{J\not\ni t_i} R_i\left(\frac{\delta_i \mathcal{L}_i}{\delta u_J}\right) \delta F_{i,J} + \sum_J \mathrm{D}_{t_i}\left(R_i\left(\frac{\delta_i \mathcal{L}_i}{\delta u_{Jt_i}}\right) \delta F_{i,J} \right)\right) \wedge \mathrm{d}t_i .
\end{aligned}$$

Hence on solutions of the pluri-Lagrangian problem for $\mathcal{L}$ there holds that

$$\delta\overline{\mathcal{L}} = \sum_{1\le i\le N} \left(\mathrm{D}_{t_i} \sum_J \frac{\delta_1 \mathcal{L}_1}{\delta u_{Jt_1}} \delta F_{i,J} \right) \wedge \mathrm{d}t_i .$$

Using the assumption that no alien derivatives occur in $\mathcal{L}_1$, we can simplify this to

$$\delta\overline{\mathcal{L}} = \sum_{1\le i\le N} \mathrm{D}_{t_i} \left(\sum_{\alpha=0}^{\infty} \frac{\partial \mathcal{L}_1}{\partial u_{t_1^{\alpha+1}}} \delta u_{t_1^\alpha} \right) \wedge \mathrm{d}t_i = \mathrm{d}\left(- \sum_{\alpha=0}^{\infty} \frac{\partial \mathcal{L}_1}{\partial u_{t_1^{\alpha+1}}} \delta u_{t_1^\alpha} \right).$$

By Proposition 6.3, the fact that $\delta\overline{\mathcal{L}}$ is exact with respect to d implies that u is a solution to the pluri-Lagrangian problem for $\overline{\mathcal{L}}$.

Now we consider the case $d = 2$. Let $F_{ij,J} = R_{ij}(u_J)$, then

$$\begin{aligned}
\delta\overline{\mathcal{L}} &= \sum_{1\le i<j\le N} \sum_J R_{ij}\left(\frac{\partial \mathcal{L}_{ij}}{\partial u_J}\right) \delta F_{ij,J} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \\
&= \sum_{1\le i<j\le N} \sum_J R_{ij}\left(\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_J} + \mathrm{D}_{t_i} \frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{Jt_i}} + \mathrm{D}_{t_j} \frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{Jt_j}} + \mathrm{D}_{t_i}\mathrm{D}_{t_j} \frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{Jt_it_j}}\right) \delta F_{ij,J} \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \\
&= \sum_{1\le i<j\le N} \Bigg(\sum_{J\not\ni t_i,t_j} R_{ij}\left(\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_J}\right) \delta F_{ij,J} + \sum_{J\not\ni t_j} \mathrm{D}_{t_i}\left(R_{ij}\left(\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{Jt_i}}\right) \delta F_{ij,J}\right) \\
&\qquad + \sum_{J\not\ni t_i} \mathrm{D}_{t_j}\left(R_{ij}\left(\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{Jt_j}}\right) \delta F_{ij,J}\right) + \sum_J \mathrm{D}_{t_i}\mathrm{D}_{t_j}\left(R_{ij}\left(\frac{\delta_{ij} \mathcal{L}_{ij}}{\delta u_{Jt_it_j}}\right) \delta F_{ij,J}\right) \Bigg) \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j .
\end{aligned}$$

On solutions of the pluri-Lagrangian problem for $\mathcal{L}$ there holds that

$$\delta\overline{\mathcal{L}} = \sum_{1 \leq i < j \leq N} \left(\sum_{J \not\ni t_j} \mathrm{D}_{t_i} \left(\frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{Jt_1}} \delta F_{ij,J} \right) - \sum_{J \not\ni t_i} \mathrm{D}_{t_j} \left(\frac{\delta_{1i}\mathcal{L}_{1i}}{\delta u_{Jt_1}} \delta F_{ij,J} \right) \right.$$
$$\left. + \sum_J \mathrm{D}_{t_i}\mathrm{D}_{t_j} \left(\left(\frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{Jt_1t_j}} - \frac{\delta_{1j}\mathcal{L}_{1i}}{\delta u_{Jt_1t_i}} \right) \delta F_{ij,J} \right) \right) \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j.$$

Since the $\mathcal{L}_{1j}$ do not contain any alien derivatives, only terms where J is $\{i,j\}$-native can be nonzero, so in all nonvanishing terms we find $F_{ij,J} = u_J$. Therefore,

$$\begin{aligned}
\delta\overline{\mathcal{L}} = \sum_{1 \leq i < j \leq N} & \left(\mathrm{D}_{t_i} \left(\sum_{J \not\ni t_j} \frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{Jt_1}} \delta u_J + \sum_J \mathrm{D}_{t_j} \left(\frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{Jt_1t_j}} \delta u_J \right) \right) \right. \\
& \left. - \mathrm{D}_{t_j} \left(\sum_{J \not\ni t_i} \frac{\delta_{1i}\mathcal{L}_{1i}}{\delta u_{Jt_1}} \delta u_J + \sum_J \mathrm{D}_{t_j} \left(\frac{\delta_{1i}\mathcal{L}_{1i}}{\delta u_{Jt_1t_i}} \delta u_J \right) \right) \right) \wedge \mathrm{d}t_i \wedge \mathrm{d}t_j \\
= \mathrm{d} & \left(\sum_{1 \leq j \leq N} \left(\sum_{J \not\ni t_j} \frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{Jt_1}} \delta u_J + \sum_J \mathrm{D}_{t_j} \left(\frac{\delta_{1j}\mathcal{L}_{1j}}{\delta u_{Jt_1t_j}} \delta u_J \right) \right) \wedge \mathrm{d}t_j \right).
\end{aligned}$$

Hence u is a solution to the pluri-Lagrangian problem for $\overline{\mathcal{L}}$. $\square$

7.3. Example: the Toda lattice

We begin our list of examples with the 1-form case and discuss the continuum limit of the discrete Toda lattice. The Toda lattice [84, 85] consists of a number of particles on a line with an exponential nearest-neighbor force. We denote the displacement of the particles from their equilibrium positions by

$$q(t) = \left(q^{[0]}(t), q^{[1]}(t), \dots, q^{[N]}(t) \right).$$

Their motion is described by the equation

$$\frac{\mathrm{d}^2 q^{[k]}}{\mathrm{d}t^2} = \exp\!\left(q^{[k+1]} - q^{[k]} \right) - \exp\!\left(q^{[k]} - q^{[k-1]} \right).$$

There are two common conventions regarding boundary conditions: periodic (formally $q^{[N+1]} \equiv q^{[1]}$) and open-end (formally $q^{[0]} \equiv +\infty$ and $q^{[N+1]} \equiv -\infty$). An integrable discretization of the Toda lattice is given by (see e.g. [79, Chapter 5])

$$\begin{aligned}
&\frac{1}{\lambda_i} \left(\exp\!\left(Q_i^{[k]} - Q^{[k]} \right) - \exp\!\left(Q^{[k]} - Q_{-i}^{[k]} \right) \right) \\
&\qquad + \lambda_i \left(\exp\!\left(Q^{[k]} - Q_i^{[k-1]} \right) - \exp\!\left(Q_{-i}^{[k+1]} - Q^{[k]} \right) \right) = 0,
\end{aligned} \tag{7.13}$$

where the subscripts i and $-i$ denote forward and backward shifts respectively and λ_i is a lattice parameter.

The second order difference equation (7.13) can be written as a first order difference equation in a position-momentum formulation. Hence it defines a map from the phase space to itself. Two such maps for different paramters λ_i and λ_j commute, hence Equation (7.13) can be called a multidimensionally consistent equation.

We use the Miwa correspondence (7.6) with $c = 1$ to identify discrete steps with continuous time shifts

$$\begin{aligned}
Q^{[k]} &= q^{[k]}(t_1, t_2, t_3, \dots), \\
Q^{[k]}_i &= q^{[k]}\left(t_1 + \lambda_i, t_2 - \frac{\lambda_i^2}{2}, t_3 + \frac{\lambda_i^3}{3}, \dots \right), \\
Q^{[k]}_{-i} &= q^{[k]}\left(t_1 - \lambda_i, t_2 + \frac{\lambda_i^2}{2}, t_3 - \frac{\lambda_i^3}{3}, \dots \right).
\end{aligned}$$

We plug these identifications into Equation (7.13) and perform a Taylor expansion in λ_i:

$$\begin{aligned}
&\left(-\exp\left(q^{[k+1]} - q^{[k]}\right) + \exp\left(q^{[k]} - q^{[k-1]}\right) + q^{[k]}_{11} \right) \lambda_i \\
&\quad + \left(\exp\left(q^{[k+1]} - q^{[k]}\right) q_1^{[k+1]} - \exp\left(q^{[k]} - q^{[k-1]}\right) q_1^{[k-1]} + q_1^{[k]} q_{11}^{[k]} - q_{12}^{[k]} \right) \lambda_i^2 = \mathcal{O}(\lambda_i^3),
\end{aligned}$$

where the subscripts 1 and 2 are a shorthand for t_1 and t_2 and denote partial derivatives. As long as one remembers that discrete fields are printed in upper case, continuous fields in lower case, and parameters are denoted by greek letters, then there should be no confusion between partial derivatives, lattice shifts, and labels. In the leading order term we recognize the first Toda equation

$$q^{[k]}_{11} = \exp\left(q^{[k+1]} - q^{[k]}\right) - \exp\left(q^{[k]} - q^{[k-1]}\right). \tag{7.14}$$

Using this equation, we find that the coefficient of λ_i^2 is

$$\begin{aligned}
&\exp\left(q^{[k+1]} - q^{[k]}\right) q_1^{[k+1]} - \exp\left(q^{[k]} - q^{[k-1]}\right) q_1^{[k-1]} + q_1^{[k]} q_{11}^{[k]} - q_{12}^{[k]} \\
&= \exp\left(q^{[k+1]} - q^{[k]}\right) \left(q_1^{[k+1]} - q_1^{[k]}\right) - \exp\left(q^{[k]} - q^{[k-1]}\right) \left(q_1^{[k-1]} - q_1^{[k]}\right) + 2 q_1^{[k]} q_{11}^{[k]} - q_{12}^{[k]} \\
&= \mathrm{D}_{t_1} \left(\exp\left(q^{[k+1]} - q^{[k]}\right) + \exp\left(q^{[k]} - q^{[k-1]}\right) + \left(q_1^{[k]}\right)^2 - q_2^{[k]} \right).
\end{aligned}$$

Under the differentiation one can recognize the second Toda equation

$$q_2^{[k]} = \left(q_1^{[k]}\right)^2 + \exp\left(q^{[k+1]} - q^{[k]}\right) + \exp\left(q^{[k]} - q^{[k-1]}\right). \tag{7.15}$$

Similarly, the higher order terms correspond to the subsequent equations of the Toda hierarchy.

It is not a coincidence that the equations found at second and higher orders can be integrated with respect to t_1. Without such integration we would get a hierarchy of the form

$$q_{11} = F(q), \qquad q_{12} = G(q, q_1), \qquad \ldots. \tag{7.16}$$

On solutions of such a hierarchy we have

$$\frac{\partial F(q)}{\partial q} q_2 = q_{112} = \frac{\partial G(q,q_1)}{\partial q} q_1 + \frac{\partial G(q,q_1)}{\partial q_1} F(q).$$

Hence, if $\frac{\partial F(q)}{\partial q}$ is invertible,

$$q_2 = \left(\frac{\partial F(q)}{\partial q} \right)^{-1} \left(\frac{\partial G(q,q_1)}{\partial q} q_1 + \frac{\partial G(q,q_1)}{\partial q_1} F(q) \right).$$

Thus, under mild nondegeneracy conditions, a hierarchy of the form (7.16) can be reformulated as a hierarchy where only the first equation is of second order.

A pluri-Lagrangian structure for the discrete Toda equation was studied in [12]. The Lagrangian is given by

$$\begin{aligned} L\big(Q, Q_i, \lambda_i\big) = \frac{1}{\lambda_i} \sum_k \Big(\exp\big(Q_i^{[k]} - Q^{[k]}\big) - 1 - \big(Q_i^{[k]} - Q^{[k]}\big) \Big) \\ - \lambda_i \sum_k \exp\big(Q^{[k]} - Q_i^{[k-1]}\big) . \end{aligned} \tag{7.17}$$

Performing a Taylor expansion and applying the Euler-Maclaurin formula as in Section 7.2.2, we obtain

$$\mathcal{L}_{\text{Miwa}}([q], \lambda) = \sum_{j=1}^{\infty} (-1)^{j+1} \frac{\lambda^j}{j} \mathcal{L}_j[q]$$

with coefficients

$$\begin{aligned}
\mathcal{L}_1 &= \sum_k \left(\frac{1}{2} \big(q_1^{[k]}\big)^2 - \exp\big(q^{[k]} - q^{[k-1]}\big) \right), \\
\mathcal{L}_2 &= \sum_k \left(q_1^{[k]} q_2^{[k]} - \frac{1}{3} \big(q_1^{[k]}\big)^3 - \big(q_1^{[k]} + q_1^{[k-1]}\big) \exp\big(q^{[k]} - q^{[k-1]}\big) \right), \\
\mathcal{L}_3 &= \sum_k \bigg(- \frac{1}{4} \Big(\big(q_1^{[k+1]}\big)^2 + 4 q_1^{[k+1]} q_1^{[k]} + \big(q_1^{[k]}\big)^2 + q_{11}^{[k+1]} \Big) \exp\big(q^{[k+1]} - q^{[k]}\big) \\
&\qquad + \frac{1}{4} \big(-q_{11}^{[k+1]} + q_{11}^{[k]} - 3 q_2^{[k]} - 3 q_2^{[k+1]} \big) \exp\big(q^{[k+1]} - q^{[k]}\big) \\
&\qquad + \frac{1}{8} \big(q_1^{[k]}\big)^4 - \frac{3}{4} \big(q_1^{[k]}\big)^2 q_2^{[k]} - \frac{1}{8} \big(q_{11}^{[k]}\big)^2 + \frac{3}{8} \big(q_2^{[k]}\big)^2 + q_1^{[k]} q_3^{[k]} \bigg), \\
&\vdots
\end{aligned}$$

By Theorem 7.4, these are the coefficients of a pluri-Lagrangian 1-form $\mathcal{L} = \sum_i \mathcal{L}_i \,\mathrm{d}t_i$ for the Toda hierarchy (7.14), (7.15), $\cdots$.

Note that $\mathcal{L}_2$ contains derivatives with respect to t_1, which are alien. However, there is no equation in the hierarchy that can be used to eliminate these derivatives. In this example we have to tolerate the alien derivative $q_1^{[k]}$. The next coefficient, $\mathcal{L}_3$, contains second derivatives with respect to t_1 and derivatives with respect to t_2. We replace these using the first and second Toda equation and find

$$\overline{\mathcal{L}}_3 = \sum_k \bigg(-\frac{1}{4}\left(q_1^{[k]}\right)^4 - \left(\left(q_1^{[k+1]}\right)^2 + q_1^{[k+1]} q_1^{[k]} + \left(q_1^{[k]}\right)^2\right) \exp\left(q^{[k+1]} - q^{[k]}\right) \\ + q_1^{[k]} q_3^{[k]} - \exp\left(q^{[k+2]} - q^{[k]}\right) - \frac{1}{2} \exp\left(2(q^{[k+1]} - q^{[k]})\right) \bigg).$$

Similarly one can obtain $\overline{\mathcal{L}_i}$ for $i \geq 4$. By Theorem 7.6, the corresponding 1-form $\overline{\mathcal{L}}$ is equivalent to $\mathcal{L}$. The Lagrangian 1-form $\overline{\mathcal{L}}$ is identical to the one that was found in [71] using the variational symmetries of the Toda lattice.

7.4. Example: a linear quad equation

As the first example for $d = 2$ we discuss a linear quad equation. This will help us understand how to proceed for the non-linear quad equations that follow in Chapter 8. Consider the equation

$$(\alpha_1 - \alpha_2)(U - U_{12}) = (\alpha_1 + \alpha_2)(U_1 - U_2). \tag{7.18}$$

It is a discrete analogue of the Cauchy-Riemann equations [11] and also the linearization of the lattice potential KdV equation, which will be discussed in Section 8.3.1. Therefore all the results in this section are consequences of those in Section 8.3.1. Nevertheless, this simple quad equation is a good subject to illustrate some of the subtleties of the continuum limit procedure.

To get meaningful equations in the continuum limit, we need to write the quad equation in a suitable form. Since in the Miwa correspondence the parameter enters linearly in the t_1-coordinate and with higher powers in the other coordinates, the leading order of the expansion of the shifts of U will only contain derivatives with respect to t_1. Other derivatives enter at higher orders. Since we want to obtain PDEs in the continuum limit, not ODEs, we require that the leading order of the expansion yields a trivial equation.

Written in terms of difference quotients, Equation (7.18) reads

$$\frac{U_1 - U_2}{\alpha_1 - \alpha_2} = \frac{U - U_{12}}{\alpha_1 + \alpha_2}.$$

Setting $U = u(t_1,\ldots)$, $U_i = u(t_1+\alpha_i,\ldots)$, etc., this would yield $u_{t_1} = -u_{t_1}$ in the leading order of the expansion. In order to avoid this, we introduce new parameters $\lambda_i = \alpha_i^{-1}$. Then Equation (7.18) reads

$$\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right)(U - U_{12}) - \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right)(U_1 - U_2) = 0. \tag{7.19}$$

or, equivalently,

$$\frac{\lambda_1^2 - \lambda_2^2}{\lambda_1\lambda_2}\left(\frac{U_1 - U_2}{\lambda_1 - \lambda_2} - \frac{U_{12} - U}{\lambda_1 + \lambda_2}\right) = 0.$$

Inside the brackets we find $u_{t_1} = u_{t_1}$ in the leading order if we set $U = u(t_1,\ldots)$, $U_i = u(t_1 + \lambda_i,\ldots)$, etc., which is trivial as desired.

We use the Miwa correspondence (7.6) with $c = -2$. This choice will give us a nice normalization of the resulting differential equations. We apply the Miwa correspondence to Equation (7.19) and expand to find a double power series in λ_1 and λ_2,

$$\sum_{i,j} \frac{4(-1)^{i+j}}{ij}\mathcal{F}_{ij}[u]\lambda_1^i\lambda_2^j = 0,$$

where $\mathcal{F}_{ji} = -\mathcal{F}_{ij}$. The factor $(-1)^{i+j}\frac{4}{ij}$ is chosen to normalize the $\mathcal{F}_{0j}$, but does not influence the final result. The first few of these coefficients are

$$\begin{aligned}
\mathcal{F}_{01} &= u_{t_2},\\
\mathcal{F}_{02} &= -u_{t_1t_1t_1} + \frac{3}{2}u_{t_1t_2} + u_{t_3},\\
\mathcal{F}_{03} &= -\frac{4}{3}u_{t_1t_1t_1t_1} + \frac{4}{3}u_{t_1t_3} + u_{t_2t_2} + u_{t_4},\\
\mathcal{F}_{04} &= -u_{t_1t_1t_1t_1t_1} - \frac{5}{3}u_{t_1t_1t_1t_2} + \frac{5}{4}u_{t_1t_2t_2} + \frac{5}{4}u_{t_1t_4} + \frac{5}{3}u_{t_2t_3} + u_{t_5},\\
&\ \vdots
\end{aligned}$$

Setting all coefficients equal to zero, we obtain the continuum limit hierarchy. We see that the equations corresponding to even times are trivial. In the odd orders we find a hierarchy of linear equations,

$$u_{t_2} = 0, \qquad u_{t_3} = u_{t_1t_1t_1}, \qquad u_{t_4} = 0, \qquad u_{t_5} = u_{t_1t_1t_1t_1t_1}, \qquad \ldots.$$

For $i \geq 1$, the equations $\mathcal{F}_{ij} = 0$ are consequences of these equations.

The linear quad equation (7.18) possesses a pluri-Lagrangian structure [11, 43],

$$L(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = U(U_i - U_i) - \frac{1}{2}\frac{\alpha_i + \alpha_j}{\alpha_i - \alpha_j}(U_i - U_j)^2. \tag{7.20}$$

The following Lemma will help us put this Lagrangian in a more convenient form.

Lemma 7.7. *$L_0(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = (U + U_{ij})(U_i - U_j)$ is a null Lagrangian (i.e. its multi-time Euler-Lagrange equations are trivially satisfied)*

Proof. Consider the discrete 1-form given by $\eta(U, U_i) = UU_i$ and $\eta(U_i, U) = -UU_i$. Its discrete exterior derivative is

$$\Delta\eta(U, U_i, U_{ij}, U_j) = UU_i + U_iU_{ij} - U_{ij}U_j - U_jU = L_0.$$

By Proposition 6.5, this implies that L_0 is a null Lagrangian. □

Using Lemma 7.7, we see that the Lagrangian (7.20) is equivalent to (denoted with = by abuse of notation)

$$L(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = \frac{1}{2}(U_i - U_j)(U - U_{ij}) - \frac{1}{2}\frac{\alpha_i + \alpha_j}{\alpha_i - \alpha_j}(U_i - U_j)^2,$$

or, in terms of the parameters $\lambda_k = \alpha_k^{-1}$,

$$L(U, U_i, U_j, U_{ij}, \lambda_i, \lambda_j) = \frac{1}{2}(U_i - U_j)(U - U_{ij}) + \frac{1}{2}\frac{\lambda_i + \lambda_j}{\lambda_i - \lambda_j}(U_i - U_j)^2.$$

Since the Taylor expansion of $(U_i - U_j)^2$ contains a factor $\lambda_i - \lambda_j$, the expansion of the Lagrangian does not contain any negative order terms. In fact all zeroth order terms vanish as well, which can easily be checked by setting one of the parameters equal to zero. Hence Theorem 7.4 applies: the coefficients of the power series

$$\mathcal{L}_{\mathrm{Miwa}}([u], \lambda_1, \lambda_2) = \sum_{i,j=1}^{\infty} \frac{4(-1)^{i+j}}{ij} \mathcal{L}_{ij}[u] \lambda_1^i \lambda_2^j$$

define a pluri-Lagrangian 2-form

$$\mathcal{L} = \sum_{1 \leq i < j \leq N} \mathcal{L}_{ij}\, \mathrm{d}t_i \wedge \mathrm{d}t_j.$$

We find

$$\begin{aligned}
\mathcal{L}_{12} &= u_{t_1} u_{t_2}, \\
\mathcal{L}_{13} &= -u_{t_1} u_{t_1 t_1 t_1} + \frac{3}{4} u_{t_2}^2 + u_{t_1} u_{t_3}, \\
\mathcal{L}_{23} &= -u_{t_1} u_{t_1 t_1 t_2} + u_{t_1 t_1} u_{t_1 t_2} - 2 u_{t_1 t_1 t_1} u_{t_2} - 3 u_{t_1 t_2} u_{t_2} - 3 u_{t_1} u_{t_2 t_2} + u_{t_2} u_{t_3}, \\
&\ \vdots
\end{aligned}$$

We will not study this example in more detail. Instead we move on to its non-linear cousins in the ABS list. They make up Chapter 8.

8. Limit hierarchies of the ABS equations

Most of the material in this chapter has not yet been published, except the results on H1 and Q1, which are contained in [89].

The computations in this chapter were performed in the SageMath software system [83]. The code is available at `https://github.com/mvermeeren/pluri-lagrangian-clim`.

8.1. A bit of clairvoyance: the limit equations

The continuum limit of a single quad equation is a hierarchy of differential equations. To identify such a hierarchy, it is sufficient to look at its leading equation. As we will see, the most general equation occurring in the leading order of the continuum limit of ABS equations is the Krichever-Novikov (KN) equation [44, 63]

$$v_t = v_{xxx} - \frac{3}{2}\frac{v_{xx}^2}{v_x} + \frac{3}{8}\frac{Q(v)}{v_x}, \tag{8.1}$$

where $Q(v) = 4v^3 + g_2 v + g_3$. Even more generally, one can replace Q by any fourth degree polynomial, but this can be reduced by a Möbius transformation to the cubic form. Another useful form of this equation is found by setting $v = \wp(u)$, where $\wp$ is the Weierstrass elliptic function, which satisfies $(\wp)'^2 = Q(\wp)$. The transformed equation reads

$$u_t = u_{xxx} - \frac{3}{2}\frac{u_{xx}^2 - \frac{1}{4}}{u_x} - \frac{3}{2}\wp(2u)u_x^3. \tag{8.2}$$

An introduction to elliptic functions tailored to their use in (discrete) integrable systems can be found in [38, Appendix B].

In the leading order of the continuum limit of all ABS equations of type Q we will find the KN equation or degenerations thereof, by which we mean a particular choice of Q in equation (8.1) or of the periods of $\wp$ in Equation (8.2). For example, for $Q = 0$ we find the Schwarzian KdV equation

$$v_t = v_{xxx} - \frac{3}{2}\frac{v_{xx}^2}{v_x}. \tag{8.3}$$

Alternatively the SKdV equation can be written as $\frac{v_t}{v_x} = Sv$, where S denotes the *Schwarzian derivative*, which is invariant under Möbius transformations. See [69] for an introduction to this lovely little operator.

For the equations of type H we will encounter the potential KdV equation

$$v_t = v_{xxx} - 3v_x^2 \tag{8.4}$$

and its modified version.

Table 8.1 provides a sneak preview of what is to come in this chapter. We list the leading order equations of the continuum limits of (most of) the ABS equations. In the rest of this chapter we look at the quad equations individually and study their continuum limits as a hierarchy with a pluri-Lagrangian structure. In the cases H2 and H3$_{\delta=1}$ we will give an informal discussion of the obstructions facing a continuum limit.

8.2. Type Q

All equations of type Q can be prepared for the continuum limit in the same way, based on their particularly symmetric three leg form. In the following subsection we present this general strategy, but on first reading it might be preferable to skip ahead to Subsection 8.2.2, where we give a presentation of the continuum limit of Q1$_{\delta=0}$ independent of this general framework.

8.2.1. Three leg forms and Lagrangians

All quad equations $Q(V, V_1, V_2, V_{12}, \alpha_1, \alpha_2) = 0$ from the ABS list have a three leg form:

$$Q(V, V_1, V_2, V_{12}, \alpha_1, \alpha_2) = \Psi(V, V_1, \alpha_1) - \Psi(V, V_2, \alpha_2) - \Phi(V, V_{12}, \alpha_1 - \alpha_2).$$

For the equations of type Q, the function Φ on the long (diagonal) leg is the same as the function Ψ on the short legs:

$$Q(V, V_1, V_2, V_{12}, \lambda_1, \lambda_2) = \Psi(V, V_1, \lambda_1^2) - \Psi(V, V_2, \lambda_2^2) - \Psi(V, V_{12}, \lambda_1^2 - \lambda_2^2),$$

where we introduced new parameters by $\lambda_i^2 = \alpha_i$. Suitable leg functions Ψ were listed in [10]. For the purposes of a continuum limit, it is useful to reverse one of the time directions, i.e. to consider

$$Q(V, V_{-1}, V_2, V_{-1,2}, \lambda_1, \lambda_2) = \Psi(V, V_{-1}, \lambda_1^2) - \Psi(V, V_2, \lambda_2^2) - \Psi(V, V_{-1,2}, \lambda_1^2 - \lambda_2^2).$$

Quad equation	**Leading equation of the continuum limit**	**Identification**
$Q1_{\delta=0}$	$v_3 = v_{111} - \frac{3}{2}\frac{v_{11}^2}{v_1}$	SKdV
$Q1_{\delta=1}$	$v_3 = v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1}$	KN with $\wp = 0$
Q2	$v_3 = v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1} - \frac{3}{2}\frac{v_1^3}{v^2}$	KN with $\wp = \frac{1}{v^2}$
$Q3_{\delta=0}$	$v_3 = v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1} + \frac{1}{2}v_1^3$	KN with $\wp = -\frac{1}{3}$
$Q3_{\delta=1}$	$v_3 = v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1} + \frac{1}{2}v_1^3 - \frac{3}{2}\frac{v_1^3}{\sin(v)^2}$	KN with $\wp = \frac{1}{\sin(v)^2} - \frac{1}{3}$
Q4	$v_3 = v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1} - \frac{3}{2}\wp(2v)v_1^3$	KN – general case
H1	$v_3 = v_{111} + 3v_1^2$	pKdV
H2	–	
$H3_{\delta=0}$	$v_3 = v_{111} + \frac{1}{2}v_1^3$	pmKdV
$H3_{\delta=1}$	–	

Table 8.1. *Overview of limit equations for the ABS list. We use the subscript i as a shorthand for t_i to denote partial derivatives of v.*

Towards the continuum limit it is more suitable to write the Ψ in terms of difference quotients. We will introduce a function

$$\psi(v, v', \lambda, \mu) = \psi_1(v, \lambda, \mu) + \psi_2(v', \lambda, \mu)$$

of the continuous variables, from which we can recover $\Psi(V, W, \lambda\mu)$ by plugging in suitable approximations to v and v'. Note that Ψ takes only one parameter, which is the product of the two parameters of ψ. For all of the ABS equations we will use $c = -2$ in the Miwa correspondence, which means that the derivative $v' = v_1$ is approximated by difference quotients such as $\frac{V_{-1} - V}{2\lambda_1}$ and $\frac{V - V_2}{2\lambda_2}$. We identify

$$\Psi(V, W, \lambda\mu) = \psi\left(\frac{V + W}{2}, \frac{V - W}{2\lambda}, \lambda, \mu\right).$$

All equations of the Q-list can be written in the form

$$
\begin{aligned}
Q(V,V_{-1},V_2,V_{-1,2},\lambda_1,\lambda_2) = {}& \psi\left(\frac{V+V_{-1}}{2},\frac{V-V_{-1}}{2\lambda_1},\lambda_1,\lambda_1\right)\\
&-\psi\left(\frac{V+V_2}{2},\frac{V-V_2}{2\lambda_2},\lambda_2,\lambda_2\right)\\
&-\psi\left(\frac{V+V_{-1,2}}{2},\frac{V-V_{-1,2}}{2(\lambda_1-\lambda_2)},\lambda_1-\lambda_2,\lambda_1+\lambda_2\right).
\end{aligned}
\tag{8.5}
$$

As suggested by the symmetry of the quad equation, we require that

$$
\begin{aligned}
\psi_1(v,-\lambda,\mu) &= -\psi_1(v,\lambda,\mu),\\
\psi_2(-v',\lambda,\mu) &= -\psi_2(v',\lambda,\mu),\\
\psi_2(v',-\lambda,\mu) &= \psi_2(v',\lambda,\mu).
\end{aligned}
$$

Furthermore, we require that $\psi(v,v',0,0)=0$.

We would expect the first nonzero terms in the series expansion at first order in λ_1,λ_2, but we find

$$
\begin{aligned}
&Q(V,V_{-1},V_2,V_{-1,2},\lambda_1,\lambda_2)\\
&= -\psi\left(\frac{V+V_{-1}}{2},\frac{V_{-1}-V}{2\lambda_1},-\lambda_1,\lambda_1\right)-\psi\left(\frac{V+V_2}{2},\frac{V-V_2}{2\lambda_2},\lambda_2,\lambda_2\right)\\
&\quad+\psi\left(\frac{V+V_{-1,2}}{2},\frac{V_{-1,2}-V}{2(\lambda_1-\lambda_2)},-\lambda_1+\lambda_2,\lambda_1+\lambda_2\right)\\
&= -\psi\left(v+\lambda_1 v_1, v_1+\frac{\lambda_1}{2}(v_{11}-v_2),-\lambda_1,\lambda_1\right)-\psi\left(v-\lambda_2 v_1, v_1-\frac{\lambda_2}{2}(v_{11}+v_2),\lambda_2,\lambda_2\right)\\
&\quad+\psi\left(v+(\lambda_1-\lambda_2)v_1, v_1+\frac{\lambda_1-\lambda_2}{2}v_{11}-\frac{\lambda_1+\lambda_2}{2}v_2,-\lambda_1+\lambda_2,\lambda_1+\lambda_2\right)\\
&\quad+\mathcal{O}\big(\lambda_1^2+\lambda_2^2\big)\\
&= \mathcal{O}\big(\lambda_1^2+\lambda_2^2\big).
\end{aligned}
$$

This is the leading order cancellation required to obtain PDEs in the continuum limit: at the first order, where generically we would get only derivatives with respect to t_1, we get nothing at all.

Equation (8.5) also reveals a reason for considering the three-leg form with a "downward" diagonal leg, as in Figure 8.1(b): the difference quotient $\frac{V_{-1,2}-V}{2(\lambda_2-\lambda_1)}$ can be expanded in a double power series, but its "upward" analogue $\frac{V_{1,2}-V}{2(\lambda_1+\lambda_2)}$ cannot.

To find a Lagrangian for Equation (8.5), we follow [4, 10] and integrate the leg function ψ. We take

$$
\chi_1(v,\lambda,\mu)=\frac{2}{\lambda}\int\psi_1(v,\lambda,\mu)\,\mathrm{d}v \qquad\text{and}\qquad \chi_2(v',\lambda,\mu)=2\int\psi_2(v',\lambda,\mu)\,\mathrm{d}v'.
$$

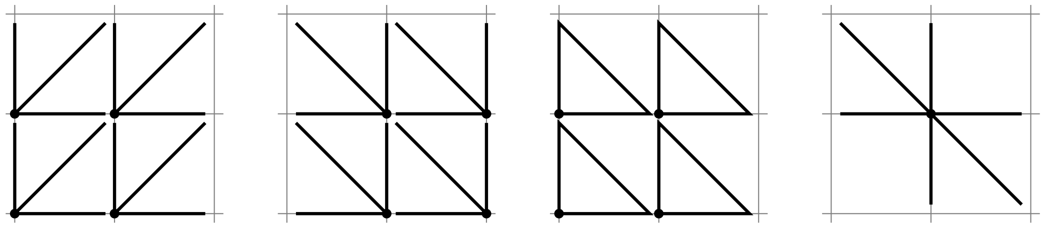


Figure 8.1. *The stencils on four adjacent quads for* (*a*) *the three-leg form in the usual orientation,* (*b*) *the three-leg form after time-reversal,* (*c*) *the triangle form for the Lagrangian, and* (*d*) *the Euler-Lagrange equations in a planar lattice.*

Then

$$\chi_1(v,-\lambda,\mu)=\chi_1(v,\lambda,\mu),$$
$$\chi_2(-v',\lambda,\mu)=\chi_2(v',-\lambda,\mu)=\chi_2(v',\lambda,\mu),$$

and $\chi=\chi_1+\chi_2$ satisfies

$$\frac{\lambda}{2}\frac{\partial}{\partial v}\chi(v,v',\lambda,\mu)+\frac{1}{2}\frac{\partial}{\partial v'}\chi(v,v',\lambda,\mu)=\psi(v,v',\lambda,\mu).$$

Now

$$\Lambda(V,W,\lambda,\mu)=\lambda\,\chi\left(\frac{V+W}{2},\frac{V-W}{2\lambda},\lambda,\mu\right)$$

gives the terms of the Lagrangian in triangle form:

$$\begin{aligned}L(V,V_1,V_2,\lambda_1,\lambda_2)&=\Lambda(V,V_1,\lambda_1,\lambda_1)-\Lambda(V,V_2,\lambda_2,\lambda_2)-\Lambda(V_1,V_2,\lambda_1-\lambda_2,\lambda_1+\lambda_2)\\&=\lambda_1\chi\left(\frac{V+V_1}{2},\frac{V-V_1}{2\lambda_1},\lambda_1,\lambda_1\right)-\lambda_2\chi\left(\frac{V+V_2}{2},\frac{V-V_2}{2\lambda_2},\lambda_2,\lambda_2\right)\\&\quad-(\lambda_1-\lambda_2)\chi\left(\frac{V_1+V_2}{2},\frac{V_1-V_2}{2(\lambda_1-\lambda_2)},\lambda_1-\lambda_2,\lambda_1+\lambda_2\right).\end{aligned}\tag{8.6}$$

Note the symmetries of Λ:

$$\Lambda(V,W,\lambda,\mu)=\Lambda(W,V,\lambda,\mu)=\Lambda(V,W,-\lambda,\mu).$$

In some cases we will rescale Λ and hence L by a constant factor. This is purely for esthetic reasons and does not affect the multi-time Euler-Lagrange equations.

Proposition 8.1. *Solutions of the quad equation $Q=0$ in the plane, with Q given by Equation* (8.5)*, are critical fields for the action of the Lagrangian given by Equation* (8.6)*.*

Proof. We have

$$\begin{aligned}
\frac{\partial}{\partial V} L(V, V_1, V_2, \lambda_1, \lambda_2) &= \lambda_1 \frac{\partial}{\partial V} \chi\left(\frac{V+V_1}{2}, \frac{V-V_1}{2\lambda_1}, \lambda_1, \lambda_1\right) \\
&\quad - \lambda_2 \frac{\partial}{\partial V} \chi\left(\frac{V+V_2}{2}, \frac{V-V_2}{2\lambda_2}, \lambda_2, \lambda_2\right) \\
&= \frac{\lambda_1}{2} \chi_1'\left(\frac{V+V_1}{2}, \lambda_1, \lambda_1\right) + \frac{1}{2} \chi_2'\left(\frac{V-V_1}{2\lambda_1}, \lambda_1, \lambda_1\right) \\
&\quad - \frac{\lambda_1}{2} \chi_1'\left(\frac{V+V_2}{2}, \lambda_2, \lambda_2\right) - \frac{1}{2} \chi_2'\left(\frac{V-V_2}{2\lambda_2}, \lambda_2, \lambda_2\right) \\
&= \psi\left(\frac{V+V_1}{2}, \frac{V-V_1}{2\lambda_1}, \lambda_1, \lambda_1\right) - \psi\left(\frac{V+V_2}{2}, \frac{V-V_2}{2\lambda_2}, \lambda_2, \lambda_2\right) \\
&= \Psi(V, V_1, \lambda_1^2) - \Psi(V, V_2, \lambda_2^2).
\end{aligned}$$

Similarly, we have

$$\frac{\partial}{\partial V_1} L(V, V_1, V_2, \lambda_1, \lambda_2) = \Psi(V_1, V, \lambda_1^2) - \Psi(V_1, V_2, \lambda_1^2 - \lambda_2^2)$$

and

$$\frac{\partial}{\partial V_2} L(V, V_1, V_2, \lambda_1, \lambda_2) = -\Psi(V_2, V, \lambda_1^2) - \Psi(V_2, V_1, \lambda_1^2 - \lambda_2^2).$$

Summing up all derivatives of the action in the plane with respect to the field at one vertex, and using the symmetry of the quad equation, we find two shifted copies of Equation (8.5), arranged as in Figure 8.1(d). □

Note that the opposite implication does not hold: not all solutions of the Euler-Lagrange equations solve the quad equation.

The Lagrangian constructed this way is suitable for the continuum limit procedure, as the following proposition establishes.

Proposition 8.2. *Every term of L is of at least first order in both parameters, $L = \mathcal{O}(\lambda_1 \lambda_2)$.*

Proof. Taking the limit $\lambda_1 \to 0$ the Lagrangian vanishes:

$$\begin{aligned}
L(V, V, V_2, 0, \lambda_2) &= -\lambda_2 \chi\left(\frac{V+V_2}{2}, \frac{V-V_2}{2\lambda_2}, \lambda_2, \lambda_2\right) + \lambda_2 \chi\left(\frac{V+V_2}{2}, \frac{V-V_2}{-2\lambda_2}, -\lambda_2, \lambda_2\right) \\
&= 0.
\end{aligned}$$

Similarly, for $\lambda_2 \to 0$ we have

$$\begin{aligned}
L(V, V_1, V, \lambda_1, 0) &= \lambda_1 \chi\left(\frac{V+V_1}{2}, \frac{V-V_1}{2\lambda_1}, \lambda_1, \lambda_1\right) - \lambda_1 \chi\left(\frac{V_1+V}{2}, \frac{V_1-V}{2\lambda_1}, \lambda_1, \lambda_1\right) \\
&= 0.
\end{aligned}$$

□

8.2.2. Cross-ratio equation ($Q1_{\delta=0}$)

The general procedure outlined above can be carried out for the cross-ratio equation with the specific choices listed in Table 8.2. However, it is instructive to forget about the ideas of Section 8.2.1 for a moment and try to figure out the continuum limit starting from Equation (6.3) itself. We would like to view it as a consistent numerical discretization of some differential equation. To achieve this, we identify $\alpha_1 = \lambda_1^2$ and $\alpha_2 = \lambda_2^2$. Then Equation (6.3) is equivalent to

$$\frac{V_1 - V}{\lambda_1} \frac{V_{12} - V_2}{\lambda_1} - \frac{V_2 - V}{\lambda_2} \frac{V_{12} - V_1}{\lambda_2} = 0, \tag{8.7}$$

which is equivalent to the three-leg form

$$\frac{\lambda_1^2}{V - V_1} - \frac{\lambda_2^2}{V - V_2} - \frac{\lambda_1^2 - \lambda_2^2}{V - V_{12}} = 0.$$

For the continuum limit we use the the Miwa correspondence (7.6) with $c = -2$. A Taylor expansion of (8.7) yields

$$\sum_{i,j} \frac{4(-1)^{i+j}}{ij} \mathcal{F}_{ij}[v] \lambda_1^i \lambda_2^j = 0$$

with

$$\begin{aligned}
\mathcal{F}_{01} &= v_1 v_2, \\
\mathcal{F}_{02} &= \frac{3}{2} v_{11}^2 - v_1 v_{111} + \frac{3}{2} v_1 v_{12} + \frac{3}{2} v_{11} v_2 + \frac{3}{8} v_2^2 + v_1 v_3, \\
\mathcal{F}_{03} &= \frac{8}{3} v_{11} v_{111} - \frac{4}{3} v_1 v_{1111} + 4 v_{11} v_{12} + \frac{4}{3} v_1 v_{13} + \frac{4}{3} v_{111} v_2 + 2 v_{12} v_2 + v_1 v_{22} + \frac{4}{3} v_{11} v_3 \\
&\quad + \frac{2}{3} v_2 v_3 + v_1 v_4, \\
\mathcal{F}_{04} &= -\frac{10}{9} v_{111}^2 - \frac{5}{3} v_{11} v_{1111} + v_1 v_{11111} + \frac{5}{3} v_1 v_{1112} - 5 v_{11} v_{112} - \frac{10}{3} v_{111} v_{12} - \frac{5}{2} v_{12}^2 \\
&\quad - \frac{5}{4} v_1 v_{122} - \frac{10}{3} v_{11} v_{13} - \frac{5}{4} v_1 v_{14} - \frac{5}{6} v_{1111} v_2 - \frac{5}{2} v_{112} v_2 - \frac{5}{3} v_{13} v_2 - \frac{5}{4} v_{11} v_{22} \\
&\quad - \frac{5}{8} v_2 v_{22} - \frac{5}{3} v_1 v_{23} - \frac{10}{9} v_{111} v_3 - \frac{5}{3} v_{12} v_3 - \frac{5}{18} v_3^2 - \frac{5}{4} v_{11} v_4 - \frac{5}{8} v_2 v_4 - v_1 v_5, \\
&\ \ \vdots
\end{aligned}$$

where once again we use the subscript i rather than t_i to denote partial derivatives of v. We assume that $v_1 \neq 0$. Then we see that the flows corresponding to even times are

$$Q = \lambda_1^2(V_2 - V)(V_{12} - V_1) - \lambda_2^2(V_1 - V)(V_{12} - V_2)$$
$$\Psi(V, W, \lambda^2) = \frac{\lambda^2}{V - W}$$
$$\psi(v, v', \lambda, \mu) = \frac{\mu}{2v'}$$
$$\chi(v, v', \lambda, \mu) = \mu \log(v')$$
$$\Lambda(V, W, \lambda, \mu) = \lambda\mu \log\left(\frac{V - W}{2\lambda}\right)$$

Table 8.2. *$Q1_{\delta=0}$ fact sheet. See Section 8.2.1 for the meaning of these functions.*

trivial and in the odd orders we find the hierarchy of Schwarzian KdV equations,

$$\begin{aligned}
v_2 &= 0, \\
\frac{v_3}{v_1} &= -\frac{3v_{11}^2}{2v_1^2} + \frac{v_{111}}{v_1}, \\
v_4 &= 0, \\
\frac{v_5}{v_1} &= -\frac{45v_{11}^4}{8v_1^4} + \frac{25v_{11}^2 v_{111}}{2v_1^3} - \frac{5v_{111}^2}{2v_1^2} - \frac{5v_{11}v_{1111}}{v_1^2} + \frac{v_{11111}}{v_1}, \\
&\vdots
\end{aligned} \tag{8.8}$$

For $i \geq 1$, the equations $\mathcal{F}_{ij} = 0$ are differential consequences of these equations.

A Pluri-Lagrangian description of Equation (6.3) was found in [47],

$$L = \alpha_i \log(V - V_i) - \alpha_j \log(V - V_i) - (\alpha_i - \alpha_j)\log(V_i - V_j). \tag{8.9}$$

It is equivalent to

$$L = \lambda_i^2 \log\left(\frac{V - V_i}{\lambda_i}\right) - \lambda_j^2 \log\left(\frac{V - V_j}{\lambda_j}\right) - (\lambda_i^2 - \lambda_j^2)\log\left(\frac{V_i - V_j}{\lambda_i - \lambda_j}\right).$$

Each term of the series $\mathcal{L}_{\text{Miwa}}$ constructed form this discrete Lagrangian contains strictly positive powers of both λ_i and λ_j. Thus by Theorem 7.4 we can identify the coefficients of this power series with the coefficients of a pluri-Lagrangian 2-form. The first few coefficients of the form $\mathcal{L}_{1j}$ are

$$
\begin{aligned}
\mathcal{L}_{12} &= -\frac{v_{11}}{2v_1} - \frac{v_2}{4v_1}, \\
\mathcal{L}_{13} &= \frac{v_{111}}{4v_1} - \frac{3v_{12}}{8v_1} + \frac{3v_{11}v_2}{8v_1^2} + \frac{3v_2^2}{16v_1^2} - \frac{v_3}{4v_1}, \\
\mathcal{L}_{14} &= \frac{v_{11}^3}{3v_1^3} - \frac{v_{11}v_{111}}{3v_1^2} + \frac{v_{112}}{3v_1} - \frac{v_{11}v_{12}}{6v_1^2} - \frac{v_{13}}{3v_1} + \frac{v_{11}^2 v_2}{6v_1^3} - \frac{v_{111}v_2}{6v_1^2} + \frac{v_{12}v_2}{4v_1^2} - \frac{v_{11}v_2^2}{4v_1^3} - \frac{v_2^3}{8v_1^3} \\
&\quad + \frac{v_{11}v_3}{3v_1^2} + \frac{v_2 v_3}{3v_1^2} - \frac{v_4}{4v_1}.
\end{aligned}
$$

In the final step we eliminate the alien derivatives. We are free to add terms that attain a double zero on solutions. This allows us to eliminate all products of time derivatives, e.g. $v_2^2, v_{12}v_2, v_2v_3, \cdots$. The remaining alien derivatives (which must occur linearly) can be eliminated by adding a suitable exact form $\mathrm{d}\left(\sum_j c_j \mathrm{d}t_j\right)$. In this case we take

$$
\sum_j c_j \mathrm{d}t_j = \frac{1}{2}\log(v_1)\mathrm{d}t_2 - \frac{2v_{11} - 3v_2}{8v_1}\mathrm{d}t_3 + \left(\frac{v_{11}^2 - v_{11}v_2}{6v_1^2} - \frac{v_{12} - v_3}{3v_1}\right)\mathrm{d}t_4 + \cdots.
$$

This was chosen to eliminate alien derivatives from the first row of coefficients $\mathcal{L}_{1j}$. Then by Theorem 7.6 all alien derivatives in other coefficients can be killed by adding double zeros. Some of the coefficients, after eliminating alien derivatives, are given in Table 8.3.

There are too many multi-time Euler-Lagrange equations to list them all. Arguably the most interesting ones are those of the form

$$
\frac{\delta_{1j}\mathcal{L}_{1j}}{\delta v_1} = \frac{\delta_{ij}\mathcal{L}_{ij}}{\delta v_i} \qquad \text{for } i \neq j \neq 1.
$$

They require that, ignoring the diagonal entries, all rows are equal in the infinite matrix

$$
\left(\frac{\delta_{ij}\mathcal{L}_{ij}}{\delta v_i}\right)_{i,j\in\mathbb{N}} =
$$
$$
\begin{pmatrix}
0 & \frac{v_2}{4v_1^2} & \frac{v_{11}^2}{2v_1^3} - \frac{v_{111}}{2v_1^2} + \frac{v_3}{4v_1^2} & \frac{v_4}{4v_1^2} & \frac{9v_{11}^4}{4v_1^5} - \frac{21v_{11}^2 v_{111}}{4v_1^4} - \frac{2v_{11111} - v_5}{4v_1^2} + \frac{3v_{111}^2 + 4v_{11}v_{1111}}{2v_1^3} & \\
\frac{1}{4v_1} & 0 & \frac{v_{11}^2}{8v_1^3} - \frac{v_{111}}{4v_1^2} & 0 & \frac{27v_{11}^4}{32v_1^5} - \frac{17v_{11}^2 v_{111}}{8v_1^4} - \frac{v_{11111}}{4v_1^2} + \frac{7v_{111}^2 + 6v_{11}v_{1111}}{8v_1^3} & \\
\frac{1}{4v_1} & 0 & 0 & 0 & \frac{27v_{11}^4}{32v_1^5} - \frac{17v_{11}^2 v_{111}}{8v_1^4} - \frac{v_{11111}}{4v_1^2} + \frac{7v_{111}^2 + 6v_{11}v_{1111}}{8v_1^3} & \\
\frac{1}{4v_1} & 0 & \frac{v_{11}^2}{8v_1^3} - \frac{v_{111}}{4v_1^2} & 0 & \frac{27v_{11}^4}{32v_1^5} - \frac{17v_{11}^2 v_{111}}{8v_1^4} - \frac{v_{11111}}{4v_1^2} + \frac{7v_{111}^2 + 6v_{11}v_{1111}}{8v_1^3} & \\
\frac{1}{4v_1} & 0 & \frac{v_{11}^2}{8v_1^3} - \frac{v_{111}}{4v_1^2} & 0 & 0 & \\
 & & & & & \ddots
\end{pmatrix}
$$

$$\mathcal{L}_{12} = -\frac{v_2}{4v_1}$$

$$\mathcal{L}_{13} = \frac{v_{11}^2}{4v_1^2} - \frac{v_3}{4v_1}$$

$$\mathcal{L}_{14} = -\frac{v_4}{4v_1}$$

$$\mathcal{L}_{15} = \frac{3v_{11}^4}{16v_1^4} - \frac{v_{111}^2}{4v_1^2} - \frac{v_5}{4v_1}$$

$$\mathcal{L}_{23} = \frac{v_{11}^2 v_2}{8v_1^3} + \frac{2v_{11}v_{12} - v_{111}v_2}{4v_1^2}$$

$$\mathcal{L}_{24} = 0$$

$$\begin{aligned}\mathcal{L}_{25} = {} & \frac{27v_{11}^4 v_2}{32v_1^5} - \frac{2v_{111}v_{112} - 2v_{1111}v_{12} + v_{11111}v_2}{4v_1^2} \\ & - \frac{8v_{11}v_{111}v_{12} - 7v_{111}^2 v_2 - 6v_{11}v_{1111}v_2}{8v_1^3} + \frac{6v_{11}^3 v_{12} - 17v_{11}^2 v_{111} v_2}{8v_1^4}\end{aligned}$$

$$\mathcal{L}_{34} = -\frac{v_{11}^2 v_4}{8v_1^3} - \frac{2v_{11}v_{14} - v_{111}v_4}{4v_1^2}$$

$$\begin{aligned}&\mathcal{L}_{35} \\ &= -\frac{v_{1111}^2 - v_{111}v_{11111} + 2v_{111}v_{113} - 2v_{1111}v_{13} + 2v_{11}v_{15} + v_{11111}v_3 - v_{111}v_5}{4v_1^2} \\ &\quad - \frac{7v_{111}^3 - 6v_{11}v_{111}v_{1111} + 3v_{11}^2 v_{11111} + 8v_{11}v_{111}v_{13} - 7v_{111}^2 v_3 - 6v_{11}v_{1111}v_3 + v_{11}^2 v_5}{8v_1^3} \\ &\quad + \frac{19v_{11}^2 v_{111}^2 + 6v_{11}^3 v_{1111} + 12v_{11}^3 v_{13} - 34v_{11}^2 v_{111} v_3}{16v_1^4} - \frac{3\left(19v_{11}^4 v_{111} - 9v_{11}^4 v_3\right)}{32v_1^5} + \frac{45v_{11}^6}{64v_1^6}\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{45} = {} & \frac{27v_{11}^4 v_4}{32v_1^5} - \frac{2v_{111}v_{114} - 2v_{1111}v_{14} + v_{11111}v_4}{4v_1^2} \\ & - \frac{8v_{11}v_{111}v_{14} - 7v_{111}^2 v_4 - 6v_{11}v_{1111}v_4}{8v_1^3} + \frac{6v_{11}^3 v_{14} - 17v_{11}^2 v_{111} v_4}{8v_1^4}\end{aligned}$$

Table 8.3. *Coefficients $\mathcal{L}_{ij}$ for $Q1_{\delta=0}$, after eliminating alien derivatives.*

In each column, except the first one, this condition immediately gives a member of the SKdV hierarchy (8.8). All other multi-time Euler-Lagrange equations are differential consequences of the hierarchy. In particular, the classical variational principle,

$$\frac{\delta_{1j}\mathcal{L}_{1j}}{\delta v} = 0,$$

yields only a consequence of the hierarchy:

$$\begin{aligned}
0 &= -\frac{v_{12}}{2v_1^2} + \frac{v_{11}v_2}{2v_1^3} \\
0 &= \frac{3v_{11}^3}{2v_1^4} + \frac{v_{1111} - v_{13}}{2v_1^2} - \frac{4v_{11}v_{111} - v_{11}v_3}{2v_1^3} \\
0 &= -\frac{v_{14}}{2v_1^2} + \frac{v_{11}v_4}{2v_1^3} \\
0 &= \frac{45v_{11}^5}{4v_1^6} - \frac{30v_{11}^3 v_{111}}{v_1^5} + \frac{v_{111111} - v_{15}}{2v_1^2} - \frac{10v_{111}v_{1111} + 6v_{11}v_{11111} - v_{11}v_5}{2v_1^3} \\
&\quad + \frac{15\big(4v_{11}v_{111}^2 + 3v_{11}^2 v_{1111}\big)}{4v_1^4} \\
&\;\vdots
\end{aligned}$$

The even-numbered times correspond to trivial equations, $v_{t_{2i}} = 0$, restricting the dynamics to a space of half the dimension. We can also restrict the pluri-Lagrangian formulation to this space:

$$\mathcal{L} = \sum_{i<j} \mathcal{L}_{2i+1,2j+1}\, \mathrm{d}t_{2i+1} \wedge \mathrm{d}t_{2j+1}$$

is a pluri-Lagrangian 2-form for the hierarchy of nontrivial SKdV equations,

$$\begin{aligned}
\frac{v_3}{v_1} &= -\frac{3v_{11}^2}{2v_1^2} + \frac{v_{111}}{v_1}, \\
\frac{v_5}{v_1} &= -\frac{45v_{11}^4}{8v_1^4} + \frac{25v_{11}^2 v_{111}}{2v_1^3} - \frac{5v_{111}^2}{2v_1^2} - \frac{5v_{11}v_{1111}}{v_1^2} + \frac{v_{11111}}{v_1}, \\
&\;\vdots
\end{aligned}$$

On the level of equations we could have restricted to the odd-numbered coordinates $t_1, t_3, \ldots$ from the beginning. However, on the level of Lagrangians we need to consider the even-numbered coordinates as well, at least in the theoretical arguments, because otherwise there is no interpretation for the (generally nonzero) coefficients of $\lambda_1^{2i}\lambda_2^j$ and $\lambda_1^i\lambda_2^{2j}$ in the power series $\mathcal{L}_{\mathrm{Miwa}}$.

8.2.3. $\mathrm{Q1}_{\delta=1}$

We apply the procedure of Section 8.2.1 to find a suitable Lagrangian for $\mathrm{Q1}_{\delta=1}$. The intermediate steps are given in Table 8.4. We find the discrete Lagrangian

$$\begin{aligned}
L(V,&V_1,V_2,\lambda_1,\lambda_2)\\
&= \Lambda(V,V_1,\lambda_1,\lambda_1) - \Lambda(V,V_2,\lambda_2,\lambda_2) - \Lambda(V_1,V_2,\lambda_1-\lambda_2,\lambda_1+\lambda_2)\\
&= \left(V - V_1 + \lambda_1^2\right)\log\left(V - V_1 + \lambda_1^2\right) - \left(V - V_1 - \lambda_1^2\right)\log\left(V - V_1 - \lambda_1^2\right)\\
&\quad - \left(V - V_2 + \lambda_2^2\right)\log\left(V - V_2 + \lambda_2^2\right) + \left(V - V_2 - \lambda_2^2\right)\log\left(V - V_2 - \lambda_2^2\right)\\
&\quad - \left(V_1 - V_2 + \lambda_1^2 - \lambda_2^2\right)\log\left(V_1 - V_2 + \lambda_1^2 - \lambda_2^2\right)\\
&\quad + \left(V_1 - V_2 - \lambda_1^2 + \lambda_2^2\right)\log\left(V_1 - V_2 - \lambda_1^2 + \lambda_2^2\right).
\end{aligned}$$

In the continuum limit of the equation we find

$$\begin{aligned}
v_2 &= 0,\\
v_3 &= v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1},\\
v_4 &= 0,\\
v_5 &= -\frac{45v_{11}^4}{8v_1^3} + \frac{25v_{11}^2 v_{111}}{2v_1^2} - \frac{5v_{111}^2}{2v_1} - \frac{5v_{11}v_{1111}}{v_1} + v_{11111} + \frac{25v_{11}^2}{16v_1^3} - \frac{5v_{111}}{8v_1^2} - \frac{5}{128v_1^3},\\
&\vdots
\end{aligned}$$

The first nontrivial equation of this hierarchy is the Krichever-Novikov equation (8.2) with $\wp = 0$.

Some coefficients of the continuous pluri-Lagrangian 2-form are given in Table 8.5. We have only included coefficients corresponding to odd-numbered coordinates.

$$
\begin{aligned}
Q &= \lambda_1^2(V_2 - V)(V_{12} - V_1) - \lambda_2^2(V_1 - V)(V_{12} - V_2) + \lambda_1^2\lambda_2^2(\lambda_1^2 - \lambda_2^2) \\
\Psi(V, W, \lambda^2) &= \log\left(\frac{V - W + \lambda^2}{V - W - \lambda^2}\right) \\
\psi(v, v', \lambda, \mu) &= \log\left(\frac{2v' + \mu}{2v' - \mu}\right) \\
\chi(v, v', \lambda, \mu) &= (2v' + \mu)\log(2v' + \mu) - (2v' - \mu)\log(2v' - \mu) \\
\Lambda(V, W, \lambda, \mu) &= (V - W + \lambda\mu)\log(V - W + \lambda\mu) - (V - W - \lambda\mu)\log(V - W - \lambda\mu)
\end{aligned}
$$

Table 8.4. *$Q1_{\delta=1}$ fact sheet. See Section 8.2.1 for the meaning of these functions.*

$$
\begin{aligned}
\mathcal{L}_{13} &= -\frac{v_3}{2v_1} + \frac{4v_{11}^2 + 1}{8v_1^2} \\
\mathcal{L}_{15} &= -\frac{v_{111}^2}{2v_1^2} - \frac{v_5}{2v_1} + \frac{48v_{11}^4 - 40v_{11}^2 - 1}{128v_1^4} \\
\mathcal{L}_{35} &= -\frac{v_{1111}^2 - v_{111}v_{11111} + 2v_{111}v_{113} - 2v_{1111}v_{13} + 2v_{11}v_{15} + v_{11111}v_3 - v_{111}v_5}{2v_1^2} \\
&\quad - \frac{28v_{111}^3 - 24v_{11}v_{111}v_{1111} + 12v_{11}^2v_{11111} + 32v_{11}v_{111}v_{13} - 28v_{111}^2v_3 - 24v_{11}v_{1111}v_3 + 4v_{11}^2v_5 - 3v_{11111} - v_5}{16v_1^3} \\
&\quad + \frac{76v_{11}^2v_{111}^2 + 24v_{11}^3v_{1111} + 48v_{11}^3v_{13} - 136v_{11}^2v_{111}v_3 - 27v_{111}^2 - 6v_{11}v_{1111} - 20v_{11}v_{13} + 10v_{111}v_3}{32v_1^4} \\
&\quad - \frac{912v_{11}^4v_{111} - 432v_{11}^4v_3 - 328v_{11}^2v_{111} + 120v_{11}^2v_3 + 25v_{111} - 3v_3}{256v_1^5} + \frac{5\left(576v_{11}^6 - 304v_{11}^4 + 44v_{11}^2 - 1\right)}{2048v_1^6}
\end{aligned}
$$

Table 8.5. *Coefficients $\mathcal{L}_{ij}$ for $Q1_{\delta=1}$, in the space spanned by odd-numbered times, after eliminating alien derivatives.*

8.2.4. Q2

The general strategy of Section 8.2.1 works with the choices in Table 8.6. The continuum limit hierarchy is

$$
\begin{aligned}
v_2 &= 0, \\
v_3 &= v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1} - \frac{3}{2}\frac{v_1^3}{v^2}, \\
v_4 &= 0, \\
v_5 &= -\frac{45v_1^5}{8v^4} + \frac{15v_1^3 v_{11}}{v^3} + \frac{15v_1 v_{11}^2}{4v^2} - \frac{45v_{11}^4}{8v_1^3} - \frac{15v_1^2 v_{111}}{2v^2} + \frac{25v_{11}^2 v_{111}}{2v_1^2} - \frac{5v_{111}^2}{2v_1} \\
&\quad - \frac{5v_{11}v_{1111}}{v_1} + v_{11111} + \frac{5v_1}{16v^2} + \frac{25v_{11}^2}{16v_1^3} - \frac{5v_{111}}{8v_1^2} - \frac{5}{128v_1^3}, \\
&\ \ \vdots
\end{aligned}
$$

The first nontrivial equation is the Krichever-Novikov equation with $\wp = \frac{1}{v^2}$. A few coefficients of the pluri-Lagrangian 2-form are given in Table 8.7.

$$
\begin{aligned}
Q &= \lambda_1^2\big(V_2^2-V^2\big)\big(V_{12}^2-V_1^2\big) - \lambda_2^2\big(V_1^2-V^2\big)\big(V_{12}^2-V_2^2\big) \\
&\quad + \lambda_1^2\lambda_2^2\big(\lambda_1^2-\lambda_2^2\big)\big(V^2+V_1^2+V_2^2+V_{12}^2-\lambda_1^4+\lambda_1^2\lambda_2^2-\lambda_2^4\big) \\
\Psi(V,W,\lambda^2) &= \log\left(\frac{(V+W+\lambda^2)(V-W+\lambda^2)}{(V+W-\lambda^2)(V-W-\lambda^2)}\right) \\
\psi(v,v',\lambda,\mu) &= \log\left(\frac{(2v+\lambda\mu)(2v'+\mu)}{(2v-\lambda\mu)(2v'-\mu)}\right) \\
\chi(v,v',\lambda,\mu) &= \frac{1}{\lambda}(2v+\lambda\mu)\log(2v+\lambda\mu) + (2v'+\mu)\log(2v'+\mu) \\
&\quad - \frac{1}{\lambda}(2v-\lambda\mu)\log(2v-\lambda\mu) - (2v'-\mu)\log(2v'-\mu) \\
\Lambda(V,W,\lambda,\mu) &= (V+W+\lambda\mu)\log(V+W+\lambda\mu) + (V-W+\lambda\mu)\log\left(\frac{V-W}{\lambda}+\mu\right) \\
&\quad - (V+W-\lambda\mu)\log(V+W-\lambda\mu) - (V-W-\lambda\mu)\log\left(\frac{V-W}{\lambda}-\mu\right)
\end{aligned}
$$

Table 8.6. *Q2 fact sheet. See Section 8.2.1 for the meaning of these functions.*

$$
\mathcal{L}_{13} = \frac{3v_1^2}{2v^2} - \frac{v_3}{2v_1} + \frac{4v_{11}^2+1}{8v_1^2}
$$

$$
\mathcal{L}_{15} = \frac{15v_1^4}{8v^4} - \frac{v_{111}^2}{2v_1^2} - \frac{v_5}{2v_1} - \frac{5\big(12v_{11}^2-1\big)}{16v^2} + \frac{48v_{11}^4-40v_{11}^2-1}{128v_1^4}
$$

$$
\begin{aligned}
\mathcal{L}_{35} = {} & \frac{45v_1^6}{32v^6} - \frac{9v_1^4v_{11}}{4v^5} + \frac{3\big(v_{111}^2+18v_{11}v_{1111}-2v_1v_{11111}-20v_{11}v_{13}+10v_{111}v_3-6v_1v_5\big)}{8v^2} \\
& - \frac{v_{1111}^2-v_{111}v_{11111}+2v_{111}v_{113}-2v_{1111}v_{13}+2v_{11}v_{15}+v_{11111}v_3-v_{111}v_5}{2v_1^2} \\
& - \frac{324v_{11}^3-312v_1v_{11}v_{111}+48v_1^2v_{1111}+120v_1v_{11}v_3-65v_{11}}{16v^3} \\
& - \frac{28v_{111}^3-24v_{11}v_{111}v_{1111}+12v_{11}^2v_{11111}+32v_{11}v_{111}v_{13}-28v_{111}^2v_3-24v_{11}v_{1111}v_3+4v_{11}^2v_5-3v_{11111}-v_5}{16v_1^3} \\
& - \frac{5\big(132v_{11}^2v_{111}-12v_{11}^2v_3+7v_{111}-v_3\big)}{32v^2v_1} + \frac{3\big(84v_1^2v_{11}^2-24v_1^3v_{111}+200v_1^3v_3-45v_1^2\big)}{128v^4} \\
& + \frac{76v_{11}^2v_{111}^2+24v_{11}^3v_{1111}+48v_{11}^3v_{13}-136v_{11}^2v_{111}v_3-27v_{111}^2-6v_{11}v_{1111}-20v_{11}v_{13}+10v_{111}v_3}{32v_1^4} \\
& + \frac{15\big(528v_{11}^4-136v_{11}^2+1\big)}{512v^2v_1^2} - \frac{912v_{11}^4v_{111}-432v_{11}^4v_3-328v_{11}^2v_{111}+120v_{11}^2v_3+25v_{111}-3v_3}{256v_1^5} \\
& + \frac{5\big(576v_{11}^6-304v_{11}^4+44v_{11}^2-1\big)}{2048v_1^6}
\end{aligned}
$$

Table 8.7. *Coefficients $\mathcal{L}_{ij}$ for Q2 after eliminating alien derivatives.*

8.2.5. $\mathrm{Q3}_{\delta=0}$

Starting from Q3 in the form of Equation (6.5), we need to do some preparatory work before unleashing the procedure of Subsection 8.2.1. The reason is that a three-leg form is only known for versions of Q3 where the field has been transformed. We choose to set $U = \exp(iV)$. In addition to that transformation of the field, we simplify the parameterization of the equation. If we divide it by $\alpha_1\alpha_2$ its coefficients become

$$\alpha_1 - \frac{1}{\alpha_1}, \qquad \alpha_2 - \frac{1}{\alpha_2}, \qquad \text{and} \qquad \frac{\alpha_1}{\alpha_2} - \frac{\alpha_2}{\alpha_1}.$$

Now set $\alpha_i - \frac{1}{\alpha_i} = \sin(\lambda_i^2)$. Then by the addition formula for the sine function one quickly verifies for $\lambda_1^2, \lambda_2^2 < \frac{\pi}{2}$ that $\frac{\alpha_1}{\alpha_2} - \frac{\alpha_2}{\alpha_1} = \sin(\lambda_1^2 - \lambda_2^2)$, hence the equation can be written as

$$\sin(\lambda_1^2)\left(e^{iV}e^{iV_1} + e^{iV_2}e^{iV_{12}}\right) - \sin(\lambda_2^2)\left(e^{iV}e^{iV_2} + e^{iV_1}e^{iV_{12}}\right) \\ - \sin(\lambda_1^2 - \lambda_2^2)\left(e^{iV}e^{iV_{12}} + e^{iV_1}e^{iV_2}\right) = 0.$$

Starting from this form of the equation, and transforming the three-leg form found in [4] accordingly, the strategy of Section 8.2.1 works as outlined in Table 8.8. The continuum limit hierarchy is

$$\begin{aligned}
v_2 &= 0,\\
v_3 &= v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1} + \frac{1}{2}v_1^3,\\
v_4 &= 0,\\
v_5 &= \frac{3}{8}v_1^5 - \frac{5}{4}v_1v_{11}^2 + \frac{5}{2}v_1^2v_{111} - \frac{5}{48}v_1 - \frac{45v_{11}^4}{8v_1^3} + \frac{25v_{11}^2v_{111}}{2v_1^2} - \frac{5v_{111}^2}{2v_1} - \frac{5v_{11}v_{1111}}{v_1}\\
&\quad + v_{11111} + \frac{25v_{11}^2}{16v_1^3} - \frac{5v_{111}}{8v_1^2} - \frac{5}{128v_1^3},\\
&\ \ \vdots
\end{aligned}$$

The first nontrivial equation is the Krichever-Novikov equation with the $\wp = -\frac{1}{3}$.

Note on dilogarithms

The expression for the function χ contains dilogarithms. The dilogarithm is given by

$$\mathrm{Li}_2(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2} \qquad \text{for } |z| < 1$$

and by analytic continuation for other $z \in \mathbb{C} \setminus [1, \infty)$. There exists a staggering amount of beautiful identities for the dilogarithm, see for example [98]. We are mostly interested

$$
\begin{aligned}
Q &= \sin(\lambda_1^2)\left(e^{iV}e^{iV_1}+e^{iV_2}e^{iV_{12}}\right)-\sin(\lambda_2^2)\left(e^{iV}e^{iV_2}+e^{iV_1}e^{iV_{12}}\right)\\
&\quad -\sin(\lambda_1^2-\lambda_2^2)\left(e^{iV}e^{iV_{12}}+e^{iV_1}e^{iV_2}\right)\\
\Psi(V,W,\lambda^2) &= \log\left(\frac{e^{i\lambda^2}e^{iV}-e^{iW}}{e^{iV}-e^{i\lambda^2}e^{iW}}\right)\\
&= \log\left(\sin\left(\frac{V-W+\lambda^2}{2}\right)\right)-\log\left(\sin\left(\frac{V-W-\lambda^2}{2}\right)\right)\\
\psi(v,v',\lambda,\mu) &= \log\left(\sin\left(\lambda v'+\frac{\lambda\mu}{2}\right)\right)-\log\left(\sin\left(\lambda v'-\frac{\lambda\mu}{2}\right)\right)\\
\chi(v,v',\lambda,\mu) &= \frac{i}{\lambda}\left(-2\lambda^2\mu v'+\mathrm{Li}_2\left(e^{i(2\lambda v'+\lambda\mu)}\right)-\mathrm{Li}_2\left(e^{i(2\lambda v'-\lambda\mu)}\right)\right)\\
\Lambda(V,W,\lambda,\mu) &= \lambda\mu(V-W)-\mathrm{Li}_2\left(e^{i(V-W+\lambda\mu)}\right)+\mathrm{Li}_2\left(e^{i(V-W-\lambda\mu)}\right)
\end{aligned}
$$

Table 8.8. *$Q3_{\delta=0}$ fact sheet. See Section 8.2.1 for the meaning of these functions.*

$$
\mathcal{L}_{13} = -\frac{1}{4}v_1^2-\frac{v_3}{4v_1}+\frac{4v_{11}^2+1}{16v_1^2}
$$

$$
\mathcal{L}_{15} = -\frac{1}{16}v_1^4+\frac{5}{8}v_{11}^2-\frac{v_{111}^2}{4v_1^2}-\frac{v_5}{4v_1}+\frac{48v_{11}^4-40v_{11}^2-1}{256v_1^4}-\frac{5}{96}
$$

$$
\begin{aligned}
\mathcal{L}_{35} = {}& \tfrac{1}{64}v_1^6-\tfrac{17}{64}v_1^2v_{11}^2+\tfrac{7}{32}v_1^3v_{111}-\tfrac{5}{32}v_1^3v_3+\tfrac{91}{768}v_1^2-\tfrac{1}{16}v_{111}^2-\tfrac{9}{8}v_{11}v_{1111}+\tfrac{1}{8}v_1v_{11111}\\
&+\tfrac{5}{4}v_{11}v_{13}-\tfrac{5}{8}v_{111}v_3+\tfrac{3}{8}v_1v_5+\tfrac{660v_{11}^2v_{111}-60v_{11}^2v_3+35v_{111}-5v_3}{192v_1}+\tfrac{5}{1024v_1^2}\\
&-\tfrac{330v_{11}^4-85v_{11}^2+32v_{1111}^2-32v_{111}v_{11111}+64v_{111}v_{113}-64v_{1111}v_{13}+64v_{11}v_{15}+32v_{11111}v_3-32v_{111}v_5}{128v_1^2}\\
&-\tfrac{28v_{111}^3-24v_{11}v_{111}v_{1111}+12v_{11}^2v_{11111}+32v_{11}v_{111}v_{13}-28v_{111}^2v_3-24v_{11}v_{1111}v_3+4v_{11}^2v_5-3v_{11111}-v_5}{32v_1^3}\\
&+\tfrac{76v_{11}^2v_{111}^2+24v_{11}^3v_{1111}+48v_{11}^3v_{13}-136v_{11}^2v_{111}v_3-27v_{111}^2-6v_{11}v_{1111}-20v_{11}v_{13}+10v_{111}v_3}{64v_1^4}\\
&-\tfrac{912v_{11}^4v_{111}-432v_{11}^4v_3-328v_{11}^2v_{111}+120v_{11}^2v_3+25v_{111}-3v_3}{512v_1^5}+\tfrac{2880v_{11}^6-1520v_{11}^4+220v_{11}^2-5}{4096v_1^6}
\end{aligned}
$$

Table 8.9. *Coefficients $\mathcal{L}_{ij}$ for $Q3_{\delta=0}$, in the space spanned by odd-numbered times, after eliminating alien derivatives.*

in its derivative, which is

$$\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_2(z) = -\frac{\log(1-z)}{z}.$$

This can be used to integrate $\log \circ \sin$, which occurs in the leg function Ψ,

$$\int \log(\sin z)\,\mathrm{d}z = \frac{i}{2}\left(-z^2 + \operatorname{Li}_2\big(e^{2iz}\big)\right) - z\log(-2i) + c.$$

In the continuum limit procure we need a series expansion of $\operatorname{Li}_2(e^x)$. For small x this seems problematic, since $\operatorname{Li}_2(z)$ has a branch point at $z = 1$. The way to handle this is to carry along this branch point in the formal series expansion

$$\operatorname{Li}_2(e^x) = \operatorname{Li}_2(1) - x\log(-x) + x - \frac{x^2}{4} - \frac{x^3}{72} + \dots, \tag{8.10}$$

where the first term famously equals $\operatorname{Li}_2(1) = \frac{\pi^2}{6}$. To derive this expansion, calculate

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{Li}_2(e^x) = -\log(1-e^x) &= -\log\left(-x - \frac{x^2}{2} - \frac{x^3}{6} - \dots\right) \\
&= -\left(\log(-x) + \log\left(1 + \frac{x}{2} + \frac{x^2}{6} + \dots\right)\right) \\
&= -\left(\log(-x) + \frac{x}{2} + \frac{x^2}{24} + \dots\right)
\end{aligned}$$

and integrate this to obtain Equation (8.10).

8.2.6. Q3$_{\delta=1}$

Starting from Equation (6.5) with $\delta = 1$, we perform the same transformation on the parameters as for $\delta = 0$ above, but we now transform the fields as $U = \cos(V)$. The general strategy of Section 8.2.1 applies to this transformed equation with the choices in Table 8.10. The continuum limit hierarchy is

$$
\begin{aligned}
v_3 &= v_{111} - \frac{3}{2}\frac{v_{11}^2 - \frac{1}{4}}{v_1} + \frac{1}{2}v_1^3 - \frac{3}{2}\frac{v_1^3}{\sin(v)^2},\\
v_5 &= \frac{3}{8}v_1^5 - \frac{5}{4}v_1 v_{11}^2 + \frac{5}{2}v_1^2 v_{111} + \frac{15 v_1^5}{4\sin(v)^2} + \frac{15 v_1^3 v_{11}\cos(v)}{\sin(v)^3} - \frac{5}{48}v_1 - \frac{45 v_{11}^4}{8 v_1^3} + \frac{25 v_{11}^2 v_{111}}{2 v_1^2}\\
&\quad - \frac{5 v_{111}^2}{2 v_1} - \frac{5 v_{11} v_{1111}}{v_1} + v_{11111} - \frac{45 v_1^5}{8\sin(v)^4} + \frac{15 v_1 v_{11}^2}{4\sin(v)^2} - \frac{15 v_1^2 v_{111}}{2\sin(v)^2} + \frac{25 v_{11}^2}{16 v_1^3}\\
&\quad - \frac{5 v_{111}}{8 v_1^2} + \frac{5 v_1}{16\sin(v)^2} - \frac{5}{128 v_1^3},\\
&\ \ \vdots
\end{aligned}
$$

and $v_{2k} = 0$. The leading equation is the Krichever-Novikov equation with $\wp = \frac{1}{\sin(v)^2} - \frac{1}{3}$.

$$
\begin{aligned}
Q &= \sin(\lambda_1^2)\left(\cos(V)\cos(V_1) + \cos(V_2)\cos(V_{12})\right)\\
&\quad - \sin(\lambda_2^2)\left(\cos(V)\cos(V_2) + \cos(V_1)\cos(V_{12})\right)\\
&\quad - \sin(\lambda_1^2 - \lambda_2^2)\left(\cos(V)\cos(V_{12}) + \cos(V_1)\cos(V_2) - \sin(\lambda_1^2)\sin(\lambda_2^2)\right)\\
\Psi(V, W, \lambda^2) &= \log\left(\sin\left(\frac{V - W + \lambda^2}{2}\right)\sin\left(\frac{V + W + \lambda^2}{2}\right)\right)\\
&\quad - \log\left(\sin\left(\frac{V - W - \lambda^2}{2}\right)\sin\left(\frac{V + W - \lambda^2}{2}\right)\right)\\
\psi(v, v', \lambda, \mu) &= \log\left(\sin\left(\lambda v' + \frac{\lambda\mu}{2}\right)\sin\left(v + \frac{\lambda\mu}{2}\right)\right)\\
&\quad - \log\left(\sin\left(\lambda v' - \frac{\lambda\mu}{2}\right)\sin\left(v - \frac{\lambda\mu}{2}\right)\right)\\
\chi(v, v', \lambda, \mu) &= \frac{i}{\lambda}\left(-2\lambda^2\mu v' + \mathrm{Li}_2\left(e^{i(2\lambda v' + \lambda\mu)}\right) - \mathrm{Li}_2\left(e^{i(2\lambda v' - \lambda\mu)}\right)\right)\\
&\quad - \frac{i}{\lambda}\left(-2\lambda\mu v + \mathrm{Li}_2\left(e^{i(2v + \lambda\mu)}\right) - \mathrm{Li}_2\left(e^{i(2v - \lambda\mu)}\right)\right)\\
\Lambda(V, W, \lambda, \mu) &= \lambda\mu(V - W) - \mathrm{Li}_2\left(e^{i(V - W + \lambda\mu)}\right) + \mathrm{Li}_2\left(e^{i(V - W - \lambda\mu)}\right)\\
&\quad + \lambda\mu(V + W) - \mathrm{Li}_2\left(e^{i(V + W + \lambda\mu)}\right) + \mathrm{Li}_2\left(e^{i(V + W - \lambda\mu)}\right)
\end{aligned}
$$

Table 8.10. *Q3$_{\delta=1}$ fact sheet. See Section 8.2.1 for the meaning of these functions.*

$$\mathcal{L}_{13} = -\frac{1}{4}v_1^2 - \frac{v_3}{4v_1} + \frac{3v_1^2}{4\sin(v)^2} + \frac{4v_{11}^2+1}{16v_1^2}$$

$$\begin{aligned}\mathcal{L}_{15} = &-\frac{1}{16}v_1^4 + \frac{5}{8}v_{11}^2 - \frac{5}{12}vv_{13} - \frac{5}{12}v_1v_3 - \frac{v_{111}^2}{4v_1^2} - \frac{v_5}{4v_1} \\ &+ \frac{15v_1^4}{16\sin(v)^4} - \frac{20v_1^4+60v_{11}^2-5}{32\sin(v)^2} + \frac{48v_{11}^4-40v_{11}^2-1}{256v_1^4}\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{35} = &\tfrac{23}{192}v_1^6 - \tfrac{57}{64}v_1^2v_{11}^2 + \tfrac{61}{96}v_1^3v_{111} - \tfrac{55}{96}v_1^3v_3 - \frac{v_1^6}{64\sin(v)^2} + \frac{3v_1^4v_{11}\cos(v)}{8\sin(v)^3} + \tfrac{211}{768}v_1^2 - \frac{105v_{11}^4}{64v_1^2} \\ &+ \frac{35v_{11}^2v_{111}}{16v_1} + \tfrac{17}{48}v_{111}^2 - \tfrac{9}{8}v_{11}v_{1111} + \tfrac{1}{8}v_1v_{11111} + \tfrac{5}{4}v_{11}v_{13} + \frac{15v_{11}^2v_3}{16v_1} - \tfrac{35}{24}v_{111}v_3 - \tfrac{5}{12}vv_{33} \\ &+ \tfrac{3}{8}v_1v_5 + \frac{15v_1^6}{64\sin(v)^4} + \frac{39v_1^2v_{11}^2}{32\sin(v)^2} - \frac{17v_1^3v_{111}}{16\sin(v)^2} - \frac{5v_1^3v_3}{16\sin(v)^2} - \frac{9v_1^4v_{11}\cos(v)}{8\sin(v)^5} - \frac{81v_{11}^3\cos(v)}{8\sin(v)^3} \\ &+ \frac{39v_1v_{11}v_{111}\cos(v)}{4\sin(v)^3} - \frac{3v_1^2v_{1111}\cos(v)}{2\sin(v)^3} - \frac{15v_1v_{11}v_3\cos(v)}{4\sin(v)^3} + \frac{25v_{11}^2}{128v_1^2} + \frac{45v_{11}^6}{64v_1^6} + \frac{95v_{111}}{192v_1} - \frac{57v_{11}^4v_{111}}{32v_1^5} \\ &+ \frac{19v_{11}^2v_{111}^2}{16v_1^4} - \frac{7v_{111}^3}{8v_1^3} + \frac{3v_{11}^3v_{1111}}{8v_1^4} + \frac{3v_{11}v_{111}v_{1111}}{4v_1^3} - \frac{v_{1111}^2}{4v_1^2} - \frac{3v_{11}^2v_{11111}}{8v_1^3} + \frac{v_{111}v_{11111}}{4v_1^2} - \frac{v_{111}v_{113}}{2v_1^2} \\ &+ \frac{3v_{11}^3v_{13}}{4v_1^4} - \frac{v_{11}v_{111}v_{13}}{v_1^3} + \frac{v_{1111}v_{13}}{2v_1^2} - \frac{v_{11}v_{15}}{2v_1^2} - \frac{65v_3}{192v_1} + \frac{27v_{11}^4v_3}{32v_1^5} - \frac{17v_{11}^2v_{111}v_3}{8v_1^4} + \frac{7v_{111}^2v_3}{8v_1^3} \\ &+ \frac{3v_{11}v_{1111}v_3}{4v_1^3} - \frac{v_{11111}v_3}{4v_1^2} - \frac{v_{11}^2v_5}{8v_1^3} + \frac{v_{111}v_5}{4v_1^2} + \frac{45v_1^6}{64\sin(v)^6} + \frac{63v_1^2v_{11}^2}{64\sin(v)^4} - \frac{9v_1^3v_{111}}{32\sin(v)^4} + \frac{75v_1^3v_3}{32\sin(v)^4} \\ &- \frac{15v_1^2}{128\sin(v)^2} + \frac{495v_{11}^4}{64v_1^2\sin(v)^2} - \frac{165v_{11}^2v_{111}}{16v_1\sin(v)^2} + \frac{3v_{111}^2}{16\sin(v)^2} + \frac{27v_{11}v_{1111}}{8\sin(v)^2} - \frac{3v_1v_{11111}}{8\sin(v)^2} - \frac{15v_{11}v_{13}}{4\sin(v)^2} \\ &+ \frac{15v_{11}^2v_3}{16v_1\sin(v)^2} + \frac{15v_{111}v_3}{8\sin(v)^2} - \frac{9v_1v_5}{8\sin(v)^2} + \frac{65v_{11}\cos(v)}{32\sin(v)^3} + \frac{55}{1024v_1^2} - \frac{95v_{11}^4}{256v_1^6} + \frac{41v_{11}^2v_{111}}{64v_1^5} - \frac{27v_{111}^2}{64v_1^4} \\ &- \frac{3v_{11}v_{1111}}{32v_1^4} + \frac{3v_{11111}}{32v_1^3} - \frac{5v_{11}v_{13}}{16v_1^4} - \frac{15v_{11}^2v_3}{64v_1^5} + \frac{5v_{111}v_3}{32v_1^4} + \frac{v_5}{32v_1^3} - \frac{135v_1^2}{256\sin(v)^4} - \frac{255v_{11}^2}{128v_1^2\sin(v)^2} \\ &- \frac{35v_{111}}{64v_1\sin(v)^2} + \frac{5v_3}{64v_1\sin(v)^2} + \frac{55v_{11}^2}{1024v_1^6} - \frac{25v_{111}}{512v_1^5} + \frac{3v_3}{512v_1^5} + \frac{15}{1024v_1^2\sin(v)^2} - \frac{5}{4096v_1^6}\end{aligned}$$

Table 8.11. *Coefficients $\mathcal{L}_{ij}$ for $Q3_{\delta=1}$, in the space spanned by odd-numbered times, after eliminating alien derivatives.*

8.2.7. Q4

We make the transformation $U = \wp(V)$, which turns Equation (6.6) into

$$
\begin{aligned}
&A\big((\wp(V) - b)(\wp(V_2) - b) - (a - b)(c - b)\big)\big((\wp(V_1) - b)(\wp(V_{12}) - b) - (a - b)(c - b)\big) \\
&+ B\big((\wp(V) - a)(\wp(V_1) - a) - (b - a)(c - a)\big)\big((\wp(V_2) - a)(\wp(V_{12}) - a) - (b - a)(c - a)\big) \\
&= ABC(a - b),
\end{aligned}
$$

where

$$
\begin{aligned}
(a, A) &= \left(\wp(\lambda_1^2), \wp'(\lambda_1^2)\right), \\
(b, B) &= \left(\wp(\lambda_2^2), \wp'(\lambda_2^2)\right), \\
(c, C) &= \left(\wp(\lambda_2^2 - \lambda_1^2), \wp'(\lambda_2^2 - \lambda_1^2)\right).
\end{aligned}
$$

The continuum limit hierarchy is

$$
\begin{aligned}
v_3 &= -6v_1^3\wp(2v) - \frac{3v_{11}^2}{2v_1} + v_{111} + \frac{3}{8v_1} \\
v_5 &= -90v_1^5\wp(2v)^2 + 144v_1^5\wp(v)^2 - 24v_1^5\wp''(v) + \frac{60v_1^3v_{11}\wp(2v)\wp''(v)}{\wp'(v)} - \frac{60v_1^3v_{11}\wp(v)\wp''(v)}{\wp'(v)} \\
&\quad + 60v_1^3v_{11}\wp'(v) + 15v_1v_{11}^2\wp(2v) - 30v_1^2v_{111}\wp(2v) + \frac{5}{4}v_1\wp(2v) - \frac{45v_{11}^4}{8v_1^3} + \frac{25v_{11}^2v_{111}}{2v_1^2} \\
&\quad - \frac{5v_{111}^2}{2v_1} - \frac{5v_{11}v_{1111}}{v_1} + v_{11111} + \frac{25v_{11}^2}{16v_1^3} - \frac{5v_{111}}{8v_1^2} - \frac{5}{128v_1^3} \\
&\vdots
\end{aligned}
$$

and $v_{2k} = 0$. As the first nontrivial equation we recognize the Krichever-Novikov equation in its full generality. To simplify these equations we used the doubling formula for the Weierstrass function,

$$
\wp(2v) = -2\wp(v) + \left(\frac{\wp''(v)}{2\wp'(v)}\right)^2.
$$

In Table 8.12, which gives an overview of the construction of a suitable discrete Lagrangian, a few additional functions and constants appear. These are the Weierstrass functions σ and ζ, with the same periods as $\wp$, and the invariants g_2 and g_3 of the $\wp$-function. They satisfy

$$
\begin{aligned}
\zeta &= \frac{\sigma'}{\sigma}, & g_2 &= 12\wp^2 - 2\wp'', \\
\zeta' &= -\wp, & g_3 &= -8\wp^3 + 2\wp''\wp - (\wp')^2.
\end{aligned}
$$

Additionally, for the series expansion of ψ we use

$$\sigma(z) = \sum_{m,n=0}^{\infty} a_{m,n} \left(\frac{g_2}{2}\right)^m (2g_3)^n \frac{z^{4m+6n+1}}{(4m+6n+1)!},$$

where the first coefficients are $a_{0,0} = 1$, $a_{1,0} = -1$, and $a_{0,1} = -3$. These and many more identities involving the Weierstrass elliptic functions can be found in [1, Chapter 18].

$$
\begin{aligned}
Q &= A\big((\wp(V)-b)(\wp(V_2)-b)-(a-b)(c-b)\big) \\
&\qquad \big((\wp(V_1)-b)(\wp(V_{12})-b)-(a-b)(c-b)\big) \\
&\quad + B\big((\wp(V)-a)(\wp(V_1)-a)-(b-a)(c-a)\big) \\
&\qquad \big((\wp(V_2)-a)(\wp(V_{12})-a)-(b-a)(c-a)\big) \\
&\quad - ABC(a-b), \\
\Psi(V,W,\lambda^2) &= \log\left(\frac{\sigma(V+W+\lambda^2)\sigma(V-W+\lambda^2)}{\sigma(V+W-\lambda^2)\sigma(V-W-\lambda^2)}\right) \\
\psi(v,v',\lambda,\mu) &= \log\left(\frac{\sigma(2v+\lambda\mu)\sigma(2\lambda v'+\lambda\mu)}{\sigma(2v-\lambda\mu)\sigma(2\lambda v'-\lambda\mu)}\right) \\
&= \left(2\zeta(2v)\lambda\mu - \frac{1}{3}\wp'(2v)\lambda^3\mu^3 + \frac{1}{5}(\wp(2v)\wp'(2v))\lambda^5\mu^5 + \dots\right) \\
&\quad + \left(\log(2v'+\mu) - \frac{g_2(2\lambda v'+\lambda\mu)^4}{240} - \frac{g_3(2\lambda v'+\lambda\mu)^6}{840} + \dots\right) \\
&\quad - \left(\log(2v'-\mu) - \frac{g_2(2\lambda v'-\lambda\mu)^4}{240} - \frac{g_3(2\lambda v'-\lambda\mu)^6}{840} + \dots\right) \\
\chi(v,v',\lambda,\mu) &= \frac{1}{\lambda}\left(2\log(\sigma(2v))\lambda\mu - \frac{1}{3}\wp(2v)\lambda^3\mu^3 - \frac{1}{10}\wp(2v)^2\lambda^5\mu^5 + \dots\right) \\
&\quad + \left((2v'+\mu)(\log(2v'+\mu)-1) - \frac{g_2(2\lambda v'+\lambda\mu)^5}{1200\lambda} + \dots\right) \\
&\quad - \left((2v'-\mu)(\log(2v'-\mu)-1) - \frac{g_2(2\lambda v'-\lambda\mu)^5}{1200\lambda} + \dots\right) \\
\Lambda(V,W,\lambda,\mu) &= 2\log(\sigma(V+W))\lambda\mu - \frac{1}{3}\wp(V+W)\lambda^3\mu^3 - \frac{1}{10}\wp(V+W)^2\lambda^5\mu^5 + \dots \\
&\quad + (V-W+\mu\lambda)\left(\log\left(\frac{V-W}{\lambda}+\mu\right)-1\right) - \frac{g_2(V-W+\lambda\mu)^5}{1200} + \dots \\
&\quad - (V-W-\mu\lambda)\left(\log\left(\frac{V-W}{\lambda}-\mu\right)-1\right) + \frac{g_2(V-W-\lambda\mu)^5}{1200} + \dots
\end{aligned}
$$

Table 8.12. *Q4 fact sheet. See Section 8.2.1 for the meaning of these functions.*

$$
\begin{aligned}
\mathcal{L}_{13} &= 3v_1^2\wp(2v) - \frac{v_3}{4v_1} + \frac{4v_{11}^2+1}{16v_1^2} \\[1em]
\mathcal{L}_{15} &= 15v_1^4\wp(2v)^2 - 24v_1^4\wp(v)^2 + 4v_1^4\wp''(v) - \frac{15}{2}v_{11}^2\wp(2v) \\
&\quad - \frac{v_{111}^2}{4v_1^2} - \frac{v_5}{4v_1} + \frac{48v_{11}^4 - 40v_{11}^2 - 1}{256v_1^4} + \frac{5}{8}\wp(2v) \\[1em]
\mathcal{L}_{35} &= 45v_1^6\wp(2v)^3 + 216v_1^6\wp(2v)\wp(v)^2 - 288v_1^6\wp(v)^3 - 36v_1^6\wp(2v)\wp''(v) \\
&\quad + 72v_1^6\wp(v)\wp''(v) - 36v_1^6\wp'(v)^2 - 18v_1^4v_{11}\wp(2v)\wp'(v) + \tfrac{63}{4}v_1^2v_{11}^2\wp(2v)^2 \\
&\quad - \tfrac{9}{2}v_1^3v_{111}\wp(2v)^2 + \tfrac{75}{2}v_1^3v_3\wp(2v)^2 - 54v_1^2v_{11}^2\wp(v)^2 + 36v_1^3v_{111}\wp(v)^2 - 60v_1^3v_3\wp(v)^2 \\
&\quad + 9v_1^2v_{11}^2\wp''(v) - 6v_1^3v_{111}\wp''(v) + 10v_1^3v_3\wp''(v) - \tfrac{135}{16}v_1^2\wp(2v)^2 + \tfrac{51}{2}v_1^2\wp(v)^2 \\
&\quad - \tfrac{81}{2}v_{11}^3\wp'(v) + 39v_1v_{11}v_{111}\wp'(v) - 6v_1^2v_{1111}\wp'(v) - 15v_1v_{11}v_3\wp'(v) + \tfrac{3}{4}v_{111}^2\wp(2v) \\
&\quad + \tfrac{27}{2}v_{11}v_{1111}\wp(2v) - \tfrac{3}{2}v_1v_{11111}\wp(2v) - 15v_{11}v_{13}\wp(2v) + \tfrac{15}{2}v_{111}v_3\wp(2v) \\
&\quad - \tfrac{9}{2}v_1v_5\wp(2v) - \tfrac{17}{4}v_1^2\wp''(v) + \tfrac{65}{8}v_{11}\wp'(v) - \tfrac{5\left(132v_{11}^2v_{111}-12v_{11}^2v_3+7v_{111}-v_3\right)\wp(2v)}{16v_1} \\
&\quad - \tfrac{\left(144v_1^4v_{11}\wp(2v)-144v_1^4v_{11}\wp(v)+324v_{11}^3-312v_1v_{11}v_{111}+48v_1^2v_{1111}+120v_1v_{11}v_3-65v_{11}\right)\wp(2v)\wp''(v)}{8\wp'(v)} \\
&\quad - \tfrac{\left(-324v_{11}^3+312v_1v_{11}v_{111}-48v_1^2v_{1111}-120v_1v_{11}v_3+65v_{11}\right)\wp(v)\wp''(v)}{8\wp'(v)} \\
&\quad + \tfrac{\left(7920v_{11}^4-2040v_{11}^2+15\right)\wp(2v)}{256v_1^2} \\
&\quad + \tfrac{-v_{1111}^2+v_{111}v_{11111}-2v_{111}v_{113}+2v_{1111}v_{13}-2v_{11}v_{15}-v_{11111}v_3+v_{111}v_5}{4v_1^2} \\
&\quad - \tfrac{28v_{111}^3-24v_{11}v_{111}v_{1111}+12v_{11}^2v_{11111}+32v_{11}v_{111}v_{13}-28v_{111}^2v_3-24v_{11}v_{1111}v_3+4v_{11}^2v_5-3v_{11111}-v_5}{32v_1^3} \\
&\quad + \tfrac{76v_{11}^2v_{111}^2+24v_{11}^3v_{1111}+48v_{11}^3v_{13}-136v_{11}^2v_{111}v_3-27v_{111}^2-6v_{11}v_{1111}-20v_{11}v_{13}+10v_{111}v_3}{64v_1^4} \\
&\quad - \tfrac{912v_{11}^4v_{111}-432v_{11}^4v_3-328v_{11}^2v_{111}+120v_{11}^2v_3+25v_{111}-3v_3}{512v_1^5} + \tfrac{5\left(576v_{11}^6-304v_{11}^4+44v_{11}^2-1\right)}{4096v_1^6}
\end{aligned}
$$

Table 8.13. *Coefficients $\mathcal{L}_{ij}$ for Q4, in the space spanned by odd-numbered times, after eliminating alien derivatives.*

8.3. Type H

For quad equations of type H we do not have a general strategy to find a suitable form. This has to be investigated on a case-by-case basis.

8.3.1. Lattice potential KdV (H1)

We would like write the lpKdV equation in terms of difference quotients. To achieve this, we identify $\alpha_1 = -\lambda_1^{-2}$ and $\alpha_2 = -\lambda_2^{-2}$. Then Equation (6.7) is equivalent to

$$\frac{U_{12} - U}{\lambda_1 + \lambda_2} \frac{U_2 - U_1}{\lambda_2 - \lambda_1} = \frac{1}{\lambda_1^2 \lambda_2^2}.$$

The left hand side is now a product of meaningful difference quotients, but the right hand side explodes as the parameters tend to zero. (Setting $\alpha_i = -\lambda_i^2$ instead would cause a contradiction in the leading order, just like in the first attempt of Section 7.4.) To avoid this we make a nonautonomous change of variables

$$U(n_1, \ldots, n_N) = V(n_1, \ldots, n_N) + \frac{n_1}{\lambda_1} + \ldots \frac{n_N}{\lambda_N}.$$

Then the lpKdV equation takes the form

$$\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + V_{12} - V \right) \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1} + V_2 - V_1 \right) = \frac{1}{\lambda_2^2} - \frac{1}{\lambda_1^2}. \tag{8.11}$$

Up to a transformation of the paramters (usually $p = \pm\lambda_1^{-1}$ and $q = \pm\lambda_2^{-1}$), this is the form in which the lpKdV equation is most commonly stated, see [60] for an overview.

To see that the necessary leading order cancellation occurs, we write the equation in terms of difference quotients,

$$\frac{V_{12} - V}{\lambda_1 + \lambda_2} - \frac{V_2 - V_1}{\lambda_2 - \lambda_1} - \lambda_1 \lambda_2 \frac{V_{12} - V}{\lambda_1 + \lambda_2} \frac{V_2 - V_1}{\lambda_2 - \lambda_1} = 0.$$

If we identify $V = v(t_1, \ldots)$, $V_i = v(t_1 + c\lambda_i, \ldots)$, etc., in the leading we find the tautological equation $v_{t_1} - v_{t_1} = 0$. The last equation is only stated to check the leading order cancellation. Because the quotient $\frac{V_{12} - V}{\lambda_1 + \lambda_2}$ does not allow a double series expansion, we use Equation (8.11) to calculate the continuum limit.

Again we use the Miwa correspondence (7.6) with $c = -2$. From Equation (8.11) we find a double power series in λ_1 and λ_2,

$$\sum_{i,j} \frac{4(-1)^{i+j}}{ij} \mathcal{F}_{ij}[v] \lambda_1^i \lambda_2^j = 0,$$

where $\mathcal{F}_{ji} = -\mathcal{F}_{ij}$. The first few of these coefficients are

$$\begin{aligned}
\mathcal{F}_{01} &= v_2, \\
\mathcal{F}_{02} &= -3v_1^2 - v_{111} + \frac{3}{2}v_{12} + v_3, \\
\mathcal{F}_{03} &= -8v_1v_{11} - 4v_1v_2 - \frac{4}{3}v_{1111} + \frac{4}{3}v_{13} + v_{22} + v_4, \\
\mathcal{F}_{04} &= -5v_{11}^2 - \frac{20}{3}v_1v_{111} - 10v_1v_{12} - 5v_{11}u_2 - \frac{5}{4}v_2^2 + \frac{10}{3}v_1v_3 - v_{11111} \\
&\qquad - \frac{5}{3}v_{1112} + \frac{5}{4}v_{122} + \frac{5}{4}v_{14} + \frac{5}{3}v_{23} + v_5, \\
&\vdots
\end{aligned}$$

We see that the flows corresponding to even times are trivial. In the odd orders we find the potential KdV equations,

$$\begin{aligned}
v_2 &= 0, \\
v_3 &= 3v_1^2 + v_{111}, \\
v_4 &= 0, \\
v_5 &= 10v_1^3 + 5v_{11}^2 + 10v_1v_{111} + v_{11111}, \\
&\vdots
\end{aligned}$$

For $i \geq 1$, the equations $\mathcal{F}_{ij} = 0$ are consequences of these equations.

A pluri-Lagrangian description of Equation (6.7) was found in [47], the Lagrange function itself goes back to [15]. It reads

$$L(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = U(U_i - U_j) - (\alpha_i - \alpha_j)\log(U_i - U_j).$$

Using Lemma 7.7, we see that this Lagrangian is equivalent to (denoted with "=" by abuse of notation)

$$L(U, U_i, U_j, U_{ij}, \alpha_i, \alpha_j) = \frac{1}{2}(U - U_{ij})(U_i - U_j) + (\alpha_i - \alpha_j)\log(U_i - U_j).$$

In terms of V and λ it is (up to a constant)

$$\begin{aligned}
L(V, V_i, V_j, V_{ij}, \lambda_i, \lambda_j) = \frac{1}{2}&\left(V - V_{ij} - \lambda_i^{-1} - \lambda_j^{-1}\right)\left(V_i - V_j + \lambda_i^{-1} - \lambda_j^{-1}\right) \\
&+ \left(\lambda_i^{-2} - \lambda_j^{-2}\right)\log\left(1 + \frac{V_i - V_j}{\lambda_i^{-1} - \lambda_j^{-1}}\right).
\end{aligned}$$

Lemma 8.3. *$L_0(V, V_i, V_j, V_{ij}, \alpha_i, \alpha_j) = (\lambda_i^{-1} + \lambda_j^{-1})(V_i - V_j) + (\lambda_i^{-1} - \lambda_j^{-1})(V - V_{ij})$ is a null Lagrangian.*

Proof. Consider the discrete 1-form η defined by $\eta(V, V_i, \lambda_i) = \lambda_i^{-1}(V + V_i)$, hence by skew-symmetry $\eta(V_i, V, \lambda_i) = -\lambda_i^{-1}(V + V_i)$. Its discrete exterior derivative is

$$\Delta\eta(V, V_i, V_{ij}, V_j, \lambda_i, \lambda_j) = \frac{V + V_i}{\lambda_i} + \frac{V_i + V_{ij}}{\lambda_j} - \frac{V_{ij} + V_j}{\lambda_i} - \frac{V_j + V}{\lambda_j} = L_0. \qquad \square$$

Lemma 8.3 implies that L is equivalent to

$$\begin{aligned} L(V, V_i, V_j, V_{ij}, \lambda_i, \lambda_j) = \frac{1}{2}\left(V - V_{ij} - 2\lambda_i^{-1} - 2\lambda_j^{-1}\right)(V_i - V_j) \\ + \left(\lambda_i^{-2} - \lambda_j^{-2}\right)\log\left(1 + \frac{V_i - V_j}{\lambda_i^{-1} - \lambda_j^{-1}}\right). \end{aligned} \tag{8.12}$$

To see why this Lagrangian is preferable, do a first order Taylor expansion of the logarithm and admire the cancellation. Thanks to this cancellation we avoid terms of nonpositive order in the series expansion.

Applying the Miwa correspondence (7.6) with $c = -2$, a Taylor expansion, and the Euler-Maclaurin formula to the Lagrangian (8.12), we obtain a power series

$$\mathcal{L}_{\mathrm{Miwa}}([v], \lambda_1, \lambda_2) = \sum_{ij=1}^{\infty} \frac{4(-1)^{i+j}}{ij} \mathcal{L}_{ij}[v] \lambda_1^i \lambda_2^j,$$

whose coefficients define a continuous pluri-Lagrangian 2-form for the KdV hierarchy. The first row of coefficients reads:

$$\begin{aligned}
\mathcal{L}_{12} &= v_1 v_2, \\
\mathcal{L}_{13} &= -2v_1^3 - v_1 u_{111} + \frac{3}{4} v_2^2 + v_1 v_3, \\
\mathcal{L}_{14} &= -4v_1^2 v_2 - \frac{4}{3} v_1 v_{112} - \frac{2}{3} v_{11} v_{12} - \frac{2}{3} v_{111} v_2 + \frac{4}{3} v_2 v_3 + v_1 v_4, \\
\mathcal{L}_{15} &= \frac{10}{3} v_1 v_{11}^2 - \frac{5}{2} v_1 v_2^2 - \frac{10}{3} v_1^2 v_3 + \frac{5}{9} v_{11} v_{1111} + \frac{1}{9} v_1 v_{11111} - \frac{10}{9} v_1 v_{113} - \frac{5}{6} v_{12}^2 \\
&\quad - \frac{5}{12} v_1 v_{122} - \frac{5}{9} v_{11} v_{13} - \frac{5}{6} v_{112} v_2 - \frac{5}{12} v_{11} v_{22} - \frac{5}{9} v_{111} v_3 + \frac{5}{9} v_3^2 + \frac{5}{4} v_2 v_4 + v_1 v_5, \\
&\ \vdots
\end{aligned}$$

Note that we can get rid of the alien derivatives in each $\mathcal{L}_{1j}$ by adding a total derivative $\mathrm{D}_{t_1} c_j$ and discarding terms that have a double zero on solutions. To make sure we get an equivalent Lagrangian 2-form, we also add $\mathrm{D}_{t_i} c_j$ to the coefficients $\mathcal{L}_{ij}$, which amounts

$$
\begin{aligned}
\mathcal{L}_{12} &= v_1 v_2 \\
\mathcal{L}_{13} &= -2v_1^3 - v_1 v_{111} + v_1 v_3 \\
\mathcal{L}_{14} &= v_1 v_4 \\
\mathcal{L}_{15} &= -5v_1^4 + 10 v_1 v_{11}^2 - v_{111}^2 + v_1 v_5 \\
\mathcal{L}_{23} &= -3v_1^2 v_2 - v_1 v_{112} + v_{11} v_{12} - v_{111} v_2 \\
\mathcal{L}_{24} &= 0 \\
\mathcal{L}_{25} &= -10 v_1^3 v_2 + 20 v_1 v_{11} v_{12} - 5 v_{11}^2 v_2 - 10 v_1 v_{111} v_2 - 2 v_{111} v_{112} + 2 v_{1111} v_{12} - v_{11111} v_2 \\
\mathcal{L}_{34} &= 3v_1^2 v_4 + v_1 v_{114} - v_{11} v_{14} + v_{111} v_4 \\
\mathcal{L}_{35} &= 6v_1^5 - 15 v_1^2 v_{11}^2 + 20 v_1^3 v_{111} - 10 v_1^3 v_3 + 7 v_{11}^2 v_{111} + 6 v_1 v_{111}^2 - 12 v_1 v_{11} v_{1111} \\
&\quad + 3v_1^2 v_{11111} + 20 v_1 v_{11} v_{13} - 5 v_{11}^2 v_3 - 10 v_1 v_{111} v_3 + 3 v_1^2 v_5 - v_{1111}^2 \\
&\quad + v_{111} v_{11111} - 2 v_{111} v_{113} + v_1 v_{115} + 2 v_{1111} v_{13} - v_{11} v_{15} - v_{11111} v_3 + v_{111} v_5 \\
\mathcal{L}_{45} &= -10 v_1^3 v_4 + 20 v_1 v_{11} v_{14} - 5 v_{11}^2 v_4 - 10 v_1 v_{111} v_4 - 2 v_{111} v_{114} + 2 v_{1111} v_{14} - v_{11111} v_4
\end{aligned}
$$

Table 8.14. *Coefficients $\mathcal{L}_{ij}$ for H1, after eliminating alien derivatives.*

to adding the exact form $\mathrm{d}\left(\sum_j c_j \mathrm{d}t_j\right)$ to $\mathcal{L}$. In this particular example we take

$$
\begin{aligned}
\sum_j c_j \mathrm{d}t_j = {} & \left(\frac{4}{3} v_1 v_{12} - \frac{2}{3} v_{11} v_2 \right) \mathrm{d}t_4 \\
& + \left(\frac{10}{3} v_1^2 v_{11} - \frac{4}{9} v_{11} v_{111} - \frac{1}{9} v_1 v_{1111} + \frac{10}{9} v_1 v_{13} + \frac{5}{12} v_1 v_{22} - \frac{5}{9} v_{11} v_3 \right) \mathrm{d}t_5 \\
& + \dots .
\end{aligned}
$$

Now that we have disposed of the alien derivatives in the $\mathcal{L}_{1j}$, we can use Theorem 7.6 to eliminate the remaining alien derivatives in all other $\mathcal{L}_{ij}$. For $i < j \leq 5$, the coefficients obtained this way are displayed in Table 8.14.

Again we can restrict the pluri-Lagrangian formulation to a space of half the dimension: the 2-form

$$
\sum_{i<j} \mathcal{L}_{2i+1,2j+1} \, \mathrm{d}t_{2i+1} \wedge \mathrm{d}t_{2j+1}
$$

is a pluri-Lagrangian structure for the hierarchy of nontrivial pKdV equations,

$$\begin{aligned} v_3 &= 3v_1^2 + v_{111}, \\ v_5 &= 10v_1^3 + 5v_{11}^2 + 10v_1v_{111} + v_{11111}, \\ &\vdots \end{aligned}$$

8.3.2. H2

No continuum limit for this equation is known. In this subsection we give a heuristic explanation of the obstruction one encounters when trying to pass to the continuum limit.

After the change of parameters $\alpha_i = \lambda_i^2$, the equation H2 reads

$$(U - U_{12})(U_1 - U_2) + (\lambda_2^2 - \lambda_1^2)(U + U_1 + U_2 + U_{12}) + \lambda_2^4 - \lambda_1^4 = 0. \tag{8.13}$$

A tempting trick would be to change the sign of the field at every other vertex, $V(n,m) = (-1)^{n+m}U(n,m)$. The resulting equation

$$-(V - V_{12})(V_1 - V_2) \pm (\lambda_2^2 - \lambda_1^2)(V - V_1 - V_2 + V_{12}) + \lambda_2^4 - \lambda_1^4 = 0.$$

has a suitable power series expansion, but the sign of the second term depends on the location in the lattice. In other words, the equation has become nonautonomous. Though there should not be any fundamental objection to this, the pluri-Lagrangian theory for nonautonomous systems has not yet been developed. We mark this as a topic for future research.

Of course the fact that one particular change of variables fails does not imply that there is no continuum limit. A better perspective on the issue is given by *background solutions*. When the zero field $U \equiv 0$ is a solution to the quad equation, as was the case for H1 in the form of Equation (8.11), the continuum limit assumes the field to be small compared to the inverse of the parameters. Fast growth of the field could lead to mixing of orders in the power series expansion, which must be avoided. If $U \equiv 0$ is not a solution, as is the case for Equation (8.13), one can look for a different simple and well-behaved solution to expand around. Such solutions are known as background solutions. There are two kinds of background solutions for H2 that treat both lattice directions in the same way [37]:

$$U(n,m) = (\lambda_1 n + \lambda_2 m + c)^2$$

and

$$U(n,m) = \left(\frac{1}{2}(-1)^n \lambda_1 + \frac{1}{2}(-1)^m \lambda_2 + c \right)^2 .$$

Expanding around either of these solutions is equivalent to expanding around 0 after a nonautonomous changes of variables. As in our first attempt above, and unlike in the case of H1, these changes of variables make the equation nonautonomous. Lacking the corresponding pluri-Lagrangian theory, we do not consider these to be suitable candidates for the continuum limit.

Since we know a form of equation Q2 which is suitable for the continuum limit, we could also try to use the degeneration from Q2 to H2 [62] to find a suitable formulation of H2. This too runs into the problem of nonautonomy. If we start with a well-behaved solution of Q2, the transformed solution of H2 will be nonautonomous. In particular, one can see in [62, Equation (5.22)] that the plane wave factors pick up an alternating factor $(-1)^{n+m}$ in this degeneration. Trying to expand around such a solution will again yield a nonautonomous quad equation.

8.3.3. H3$_{\delta=0}$

We start from the transformed version of H3 given by Equation (6.10). With $\delta = 0$ it reads

$$\lambda_1(UU_1 - U_2U_{12}) - \lambda_2(UU_2 - U_1U_{12}) = 0. \tag{8.14}$$

If one would start instead from Equation (6.9), with plus signs within the parentheses, there seems to be little hope of performing the continuum limit. The continuum limit of Equation (8.14) can be taken immediately. We find

$$\begin{aligned}
u_2 &= 0,\\
u_3 &= u_{111} - 3\frac{u_1u_{11}}{u},\\
u_4 &= 0,\\
u_5 &= -10\frac{u_1^3u_{11}}{u^3} + 10\frac{2u_1u_{11}^2 + u_1^2u_{111}}{u^2} - 5\frac{2u_{11}u_{111} + u_1u_{1111}}{u} + u_{11111},\\
&\ \vdots
\end{aligned}$$

However, we run in to trouble on the Lagrangian level. The Lagrangian given in [47] for H3, adapted to the case $\delta = 0$, is

$$\begin{aligned}
L = \frac{1}{2}\log\left(\frac{UU_1}{\lambda_1}\right)^2 - \frac{1}{2}\log\left(\frac{UU_2}{\lambda_2}\right)^2 + \mathrm{Li}_2\left(\frac{\lambda_2U_1}{\lambda_1U_2}\right) - \mathrm{Li}_2\left(\frac{\lambda_1U_1}{\lambda_2U_2}\right)\\
+ 2\big(\log(\lambda_1) - \log(\lambda_2)\big)\log(U) + 2\log(\lambda_2)\big(\log(U_1) - \log(U_2)\big),
\end{aligned}$$

where Li_2 is the dilogarithm function. Unfortunately, the occurrence of expressions like $\log(\lambda_i)$ prohibits a power series expansion in the parameters λ_i.

In order to find a better form of H3$_{\delta=0}$, with an expandable Lagrangian, we make the transformation $U = \exp\big(\frac{i}{2}V\big)$. Then Equation (8.14) turns in to

$$\lambda_1\left(e^{\frac{i}{2}(V+V_1)} - e^{\frac{i}{2}(V_2+V_{12})}\right) - \lambda_2\left(e^{\frac{i}{2}(V+V_2)} - e^{\frac{i}{2}(V_1+V_{12})}\right) = 0.$$

Multiplying with $\exp\left(-\frac{i}{4}(V+V_1+V_2+V_{12})\right)$ turns this into

$$\lambda_1 \sin\left(\frac{1}{4}(V+V_1-V_2-V_{12}\right) - \lambda_2 \sin\left(\frac{1}{4}(V-V_1+V_2-V_{12}\right) = 0, \tag{8.15}$$

which is the form in which $\mathrm{H3}_{\delta=0}$ arises from Bäcklund transformations for the Sine Gordon equation [38, p60]. Additionally, it is only a simple transformation removed from the discrete Sine-Gordon equation of Hirota [40]. Using standard trigonometric identities we can rewrite the equation as

$$(\lambda_1-\lambda_2)\tan\left(\frac{V_{12}-V}{4}\right) - (\lambda_1+\lambda_2)\tan\left(\frac{V_1-V_2}{4}\right) = 0,$$

which we can put in a three-leg form by inspection:

$$\arctan\left(\frac{\lambda_1+\lambda_2}{\lambda_1-\lambda_2}\tan\left(\frac{V_1-V_2}{4}\right)\right) = \frac{V_{12}-V_1}{4} - \frac{V-V_1}{4}.$$

From the three-leg form we can derive a Lagrangian as in [10]. It takes the implicit form

$$L = \frac{1}{8}(V_1-V)^2 - \frac{1}{8}(V_2-V)^2 - \mathcal{I}_{\lambda_1,\lambda_2}(V_1-V_2),$$

where

$$\mathcal{I}_{\lambda_1,\lambda_2}(x) = \int_0^x \arctan\left(\frac{\lambda_1+\lambda_2}{\lambda_1-\lambda_2}\tan\left(\frac{y}{4}\right)\right) \mathrm{d}y.$$

There is no need to evaluate this integral exactly. Instead one can expand the integrand as a power series up to any desired order in y and integrate this series. This is sufficient to write L as a power series up to the corresponding order in the parameters λ_i. A leading order calculation shows that this power series does not contain any terms of nonpositive order in either of the parameters.

In the continuum limit of Equation (8.15) we find the potential modified KdV hierarchy,

$$\begin{aligned}
v_2 &= 0, \\
v_3 &= v_{111} + \frac{1}{2}v_1^3, \\
v_4 &= 0, \\
v_5 &= \frac{3}{8}v_1^5 + \frac{5}{2}v_1 v_{11}^2 + \frac{5}{2}v_1^2 v_{111} + v_{11111}, \\
&\ \vdots
\end{aligned}$$

Some coefficients of the pluri-Lagrangian 2-form are given in Table 8.15.

$$\mathcal{L}_{13} = \frac{1}{16}v_1^4 + \frac{1}{4}v_1v_{111} - \frac{1}{4}v_1v_3$$

$$\mathcal{L}_{15} = \frac{1}{32}v_1^6 - \frac{5}{24}v_1^2v_{11}^2 + \frac{5}{36}v_1^3v_{111} + \frac{5}{36}v_{111}^2 - \frac{5}{36}v_{11}v_{1111} - \frac{1}{36}v_1v_{11111} - \frac{1}{4}v_1v_5$$

$$\begin{aligned}\mathcal{L}_{35} = &-\frac{3}{256}v_1^8 + \frac{5}{32}v_1^4v_{11}^2 - \frac{7}{32}v_1^5v_{111} + \frac{3}{32}v_1^5v_3 + \frac{1}{16}v_{11}^4 - \frac{7}{8}v_1v_{11}^2v_{111}\\ &- \frac{3}{8}v_1^2v_{111}^2 + \frac{3}{4}v_1^2v_{11}v_{1111} - \frac{1}{8}v_1^3v_{11111} + \frac{5}{36}v_1^3v_{113} - \frac{5}{6}v_1^2v_{11}v_{13} + \frac{5}{8}v_1v_{11}^2v_3\\ &+ \frac{5}{8}v_1^2v_{111}v_3 - \frac{1}{8}v_1^3v_5 + \frac{1}{4}v_{1111}^2 - \frac{1}{4}v_{111}v_{11111} - \frac{1}{36}v_1v_{11113} - \frac{1}{9}v_{11}v_{1113}\\ &+ \frac{7}{18}v_{111}v_{113} - \frac{1}{4}v_1v_{115} - \frac{19}{36}v_{1111}v_{13} + \frac{1}{4}v_{11}v_{15} + \frac{1}{4}v_{11111}v_3 - \frac{1}{4}v_{111}v_5\end{aligned}$$

Table 8.15. *Coefficients $\mathcal{L}_{ij}$ for H3$_{\delta=0}$, in the space spanned by odd-numbered times, after eliminating alien derivatives.*

8.3.4. H3$_{\delta=1}$

Due to difficulties analogous to those of H2, no continuum limit is known for H3$_{\delta=1}$. In particular, the transformation $U(n,m) \mapsto i^{n+m}U(n,m)$ turns H3 into a nonautonomous equation if $\delta \neq 0$. Hence for H3 with nonzero parameter it seems impossible to get the convenient minus signs in the parentheses as in Equation (8.14).

8.4. Comments and connections

8.4.1. About the even-numbered times

In all the examples from this chapter, only the odd-numbered times have nontrivial equations associated to them. This is consistent with several purely continuous descriptions of integrable hierarchies. From the perspective of continuum limits, it follows from the fact that for all equations we dealt with, Q is an even or odd function of the parameters.

Proposition 8.4. *If the difference equation $Q = 0$ satisfies*

$$Q(U, U_1, U_2, U_{12}, \lambda_1, \lambda_2) = \pm Q(U, U_1, U_2, U_{12}, -\lambda_1, -\lambda_2)$$

then the continuum limit hierarchy is invariant under simultaneous reversal of all even-numbered times, $t_{2k} \mapsto -t_{2k}$.

Proof. Due to the symmetry assumptions on Q, we have

$$Q(U, U_1, U_2, U_{12}, \lambda_1, \lambda_2) = \pm Q(U_{12}, U_2, U_1, U, -\lambda_1, -\lambda_2).$$

Let $u(t_1, t_2, \ldots)$ be a solution to the continuum limit equations. Then

$$\begin{aligned} Q\bigg(u(t_1, t_2, \ldots), &u\Big(t_1 + c\lambda_1, t_2 - c\frac{\lambda_1^2}{2}, \ldots\Big), u\Big(t_1 + c\lambda_2, t_2 - c\frac{\lambda_2^2}{2}, \ldots\Big), \\ &u\Big(t_1 + c\lambda_1 + c\lambda_2, t_2 - c\frac{\lambda_1^2}{2} - c\frac{\lambda_2^2}{2}, \ldots\Big), \lambda_1, \lambda_2 \bigg) = 0 \end{aligned} \tag{8.16}$$

Now consider the same equation at times shifted once in both lattice directions, $\tau_i = t_i + (-1)^{i+1} c \left(\frac{\lambda_1^i}{i} + \frac{\lambda_2^i}{i} \right)$. Due to the symmetry of Q we have that

$$\begin{aligned} Q\bigg(u(\tau_1, \tau_2, \ldots), &u\Big(\tau_1 - c\lambda_1, \tau_2 + c\frac{\lambda_1^2}{2}, \ldots\Big), u\Big(\tau_1 - c\lambda_2, \tau_2 + c\frac{\lambda_2^2}{2}, \ldots\Big), \\ &u\Big(\tau_1 - c\lambda_1 - c\lambda_2, \tau_2 + c\frac{\lambda_1^2}{2} + c\frac{\lambda_2^2}{2}, \ldots\Big), -\lambda_1, -\lambda_2 \bigg) = 0 \end{aligned}$$

Introducing the parameters $\mu_1 = -\lambda_1$ and $\mu_2 = -\lambda_2$ we find

$$\begin{aligned} Q\bigg(u(\tau_1, \tau_2, \ldots), &u\Big(\tau_1 + c\mu_1, \tau_2 + c\frac{\mu_1^2}{2}, \ldots\Big), u\Big(\tau_1 + c\mu_2, \tau_2 + c\frac{\mu_2^2}{2}, \ldots\Big), \\ &u\Big(\tau_1 + c\mu_1 + c\mu_2, \tau_2 + c\frac{\mu_1^2}{2} + c\frac{\mu_2^2}{2}, \ldots\Big), \mu_1, \mu_2 \bigg) = 0. \end{aligned} \tag{8.17}$$

Comparing Equations (8.16) and (8.17), we immediately see that their series expansions in λ_1, λ_2 respectively μ_1, μ_2 only differ by a minus sign for each derivative with respect to an even-numbered time t_{2k}. In other words, if we have a solution u to the continuum limit hierarchy, then reversing all even times gives a new solution. □

Corollary 8.5. *If the continuum limit hierarchy of the difference equation $Q = 0$, as in Proposition 8.4, consists of evolutionary equations $u_k = f_k(u, u_1, u_{11}, \ldots)$, then for all even k we have $f_k = 0$.*

Unfortunately, as we will see in Chapter 9, there are situations where the continuum limit contains PDEs that are not evolutionary. Here we will see a different pattern of trivial equations. This pattern can be understood in the framework of reductions of the lattice KP system [27, 28], or in the context of pseudodifferential operators for the Gelfand-Dickey hierarchy. The latter will be briefly discussed in Chapter 9.

8.4.2. The double continuum limit of Wiersma and Capel

In [94] Wiersma and Capel presented a continuum limit of the equation

$$(p + q + U_{112} - U)(p - q + U_{12} - U_1) = p^2 - q^2,$$

which is equivalent to lpKdV equation (8.11) under the transformation

$$U(n, m) = V(n - m, m)$$

and with $p = \lambda_1^{-1}$ and $q = \lambda_2^{-1}$. Their procedure consists of two steps. First they obtain a hierarchy of differential-difference equations. A second continuum limit, applied to any single equation of this hierarchy, then yields the potential KdV hierarchy. Some ideas concerning this limit procedure were already developed in [64, 73]. Here we will summarize both limits in one step.

The limit procedure from [94] uses the lattice parameters $\nu = q - p$ and p itself. Consider an interpolating function u. If

$$U(n, m) = V(n - m, m) = v(t_1, t_3, t_5, \ldots),$$

then after the double limit of [94], lattice shifts correspond to multi-time shifts as follows:

$$U_1 = v\left(t_1 - \frac{2}{p}, t_3 - \frac{2}{3p^3}, t_5 - \frac{2}{5p^5}, \ldots\right)$$

and

$$U_2 = v\Bigg(t_1 + \nu\frac{2}{p^2} - \frac{\nu^2}{2}\frac{2}{p^3} + \frac{\nu^3}{3}\frac{2}{p^4} - \ldots, t_3 + \nu\frac{2}{p^4} - \frac{\nu^2}{2}\frac{4}{p^5} + \frac{\nu^3}{3}\frac{20}{3p^6} - \ldots,$$
$$t_5 + \nu\frac{2}{p^6} - \frac{\nu^2}{2}\frac{6}{p^7} + \frac{\nu^3}{3}\frac{14}{p^8} - \ldots, \ldots\Bigg).$$

The series occurring here can be recognized as Taylor expansions:

$$U_2 = v\bigg(t_1 - \left(\frac{2}{p+\nu} - \frac{2}{p} \right), t_3 - \frac{1}{3}\left(\frac{2}{(p+\nu)^3} - \frac{2}{p^3} \right), \\ t_5 - \frac{1}{5}\left(\frac{2}{(p+\nu)^5} - \frac{2}{p^5} \right), \dots \bigg).$$

Going back to the straight lattice coordinates and the original lattice parameters p and $q = p + \nu$, we find

$$\begin{aligned} V_2 = U_{12} &= v\bigg(t_1 - \frac{2}{q}, t_3 - \frac{2}{3q^3}, t_5 - \frac{2}{5q^5}, \dots \bigg), \\ V_1 = U_1 \ &= v\bigg(t_1 - \frac{2}{p}, t_3 - \frac{2}{3p^3}, t_5 - \frac{2}{5p^5}, \dots \bigg). \end{aligned}$$

Hence the end result of the double limit of Wiersma and Capel is the same as the limit we obtain using the odd-numbered Miwa variables only.

The paper [94] also presents a continuum limit for the lattice modified KdV equation. Its treatment is analogous to that of the nonmodified version. In particular, it is related to our approach in the same way.

8.4.3. The generating PDE of Nijhoff, Hone, and Joshi

Nijhoff, Hone, and Joshi [61] introduced a nonautonomous PDE for a function $\mathsf{z}_{n,m}(t,s)$ depending on a pair of continuous variables (s,t), and a pair of parameters (m,n). They noted that the flow of this PDE in continuous (s,t)-coordinates commutes with the difference equations

$$\frac{(\mathsf{z}_{n,m} - \mathsf{z}_{n+1,m})(\mathsf{z}_{n,m+1} - \mathsf{z}_{n+1,m+1})}{(\mathsf{z}_{n,m} - \mathsf{z}_{n,m+1})(\mathsf{z}_{n+1,m} - \mathsf{z}_{n+1,m+1})} = \frac{s}{t}. \tag{8.18}$$

Equation (8.18) is nothing but equation $Q1_{\delta=0}$. Hence it is possible to switch between the continuous and discrete picture by reversing the roles of parameters and independent variables.

The main feature of the PDE in question is that it generates the SKdV hierarchy[1] through the identification

$$\mathsf{z}_{n,m}(t,s) = v\bigg(x_1 + \frac{2n}{t^{\frac{1}{2}}} + \frac{2m}{s^{\frac{1}{2}}}, x_3 + \frac{2n}{3t^{\frac{3}{2}}} + \frac{2m}{3s^{\frac{3}{2}}}, \dots, \\ x_{2j+1} + \frac{2n}{(2j+1)t^{\frac{2j+1}{2}}} + \frac{2m}{(2j+1)s^{\frac{2j+1}{2}}}, \dots \bigg). \tag{8.19}$$

[1] Note that there is an error in the second SKdV equation as stated in [61]: the Lagrangian is missing the term $-\mathsf{z}_{x_2}^2/\mathsf{z}_{x_1}^2$ at the corresponding order and in the equation itself the factor 2 of the first term should be removed.

Because of this it has become known as the *generating PDE* [47, 96] for the SKdV hierarchy. Renaming the parameters $t = \lambda_1^2$ and $s = \lambda_2^2$ we obtain once again the odd order Miwa shifts

$$\begin{aligned}\mathsf{z}_{n,m} = v\bigg(& x_1 + \frac{2n}{\lambda_1} + \frac{2m}{\lambda_2}, x_3 + \frac{2n}{3\lambda_1^3} + \frac{2m}{3\lambda_2^3}, \dots, \\ & x_{2j+1} + \frac{2n}{(2j+1)\lambda_1^{2j+1}} + \frac{2m}{(2j+1)\lambda_2^{2j+1}}, \dots\bigg),\end{aligned}$$

hence our continuum limit of $\mathrm{Q1}_{\delta=0}$ is implicitly present in [61]. The relation between the (nonautonomous) generating PDE, the quad equation, and the hierarchy of (autonomous) PDEs is illustrated in the following diagram:

$$\begin{tikzcd} & \boxed{\text{Generating PDE for } \mathsf{z}_{n,m}(t,s)} \arrow[dl, "\text{role reversal } (n,m) \leftrightarrow (t,s)"'] \arrow[dr, "\text{expansion by (8.19)}"] & \\ \boxed{\text{Quad equation (8.18)}} \arrow[rr, "\text{continuum limit}"] & & \boxed{\text{SKdV hierarchy}} \end{tikzcd}$$

9. Gelfand-Dickey hierarchies

The material in this chapter has not yet been published.

The computations in this chapter were performed in the SageMath software system [83]. The code is available at https://github.com/mvermeeren/pluri-lagrangian-clim.

There are many integrable lattice equations that do not fit in the ABS list. There are 1-dimensional equations, i.e. integrable maps, like the discrete toda Lattice, and equations of dimension greater than two, for example the lattice KP systems. There are also lattice equations on a larger stencil and multi-component systems, both of which we will encounter in this chapter, where we study continuum limits of a few Gelfand-Dickey (GD) equations. The continuous GD hierarchy can be seen as a hierarchy of hierarchies, the first two of which are the KdV and the Boussinesq hierarchy. For each continuous GD hierarchy there is a corresponding lattice equation.

9.1. Continuous GD hierarchies

The Gelfand-Dickey hierarchies can be constructed by considering a differential operator

$$L_N = \partial^N + \phi^{[N-2]}\partial^{N-2} + \phi^{[N-3]}\partial^{N-3} + \ldots + \phi^{[1]}\partial + \phi^{[0]},$$

where ∂ acts on a function as $\partial f = f_x + f\partial$. The letter L is the usual choice for this operator, honoring Peter Lax. It is not to be confused with a Lagrangian. By pairing L_N with a suitable operator $P_{N,k}$ we can form a Lax representation

$$\frac{\mathrm{d}}{\mathrm{d}t_k} L_N = [P_{N,k}, L_N]$$

of a system of equations for the fields $\phi^{[0]}, \ldots, \phi^{[N-2]}$. Of course this only works if the commutator is of the form

$$[P_{N,k}, L_N] = \partial^N + f_{N-2}\partial^{N-2} + f_{N-3}\partial^{N-3} + \ldots + f_1\partial + f_0,$$

which is not the case for a generic operator P.

Suitable operators $P_{N,k}$ can be constructed using *pseudodifferential* operators, which are formal series of the form

$$\sum_{i=-\infty}^{N} f_i \partial^i$$

The powers of ∂ satisfy a generalized Leibniz rule,

$$\partial^k f = \sum_{i=0}^{\infty} \binom{k}{i} f_{x^i} \partial^{k-i},$$

where we have for negative as well as positive k that $\binom{k}{i} = \frac{k(k-1)\dots(k-i+1)}{i!}$. In the ring of pseudodifferential operators, it is possible to consider the N-th root of L_N,

$$L_N^{1/N} = \partial + \frac{1}{N}\phi^{[N-2]}\partial^{-1} + \cdots .$$

Then $P_{N,k}$ is defined as the nonnegetive order part of $L_N^{k/N}$,

$$P_{N,k} = \left(\left(L_N^{1/N} \right)^k \right)_+ ,$$

i.e. the differential operator obtained by deleting all terms with negative powers of ∂ from $L_N^{k/N}$.

This construction does not give a nontrivial equation for all k. For $k = 1$ we find $P_{N,1} = \partial$, which leads to the identification of x and t_1:

$$\frac{\mathrm{d}}{\mathrm{d}t_1} L_N = [\partial, L_N] = \phi_x^{[N-2]}\partial^{N-2} + \phi_x^{[N-3]}\partial^{N-3} + \ldots + \phi_x^{[1]}\partial + \phi_x^{[0]} = \frac{\mathrm{d}}{\mathrm{d}x} L_N .$$

Furthermore, if k is a multiple of N, then $P_{N,k}$ is a power of L_N and the commutator $[P_{N,k}, L_N]$ vanishes. Hence the equations corresponding to such t_k are trivial.

- For $N = 2$ we find the Schrödinger operator $L_2 = \partial^2 + \phi$. Its square root is

  $$L_2^{1/2} = \partial + \frac{1}{2}\phi\partial^{-1} - \frac{1}{4}\phi_x\partial^{-2} + \cdots .$$

  The first operator P which leads to a nontrivial equation is

  $$P_{2,3} = \left(L_2^{3/2} \right)_+ = \partial^3 + \frac{3}{2}\phi\partial + \frac{3}{4}\phi_x .$$

  We find

  $$\phi_{t_3} = [P_{2,3}, L_2] = \frac{1}{4}\phi_{xxx} + \frac{3}{2}\phi\phi_x ,$$

  which is the KdV equation. With the operators $(P_{2,2k+1})_{k\in\mathbb{N}}$ we recover the whole KdV hierarchy.

- For $N = 3$ we have $L_3 = \partial^3 + \phi\partial + \chi$ and
$$L_3^{1/3} = \partial + \frac{1}{3}\phi\partial^{-1} + \cdots.$$
In this case the first relevant operator is
$$P_{3,2} = \left(L_3^{2/3}\right)_+ = \partial^2 + \frac{2}{3}\phi.$$
We find
$$[P_{3,2}, L_3] = (2\chi_x - \phi_{xx})\partial + \chi_{xx} - \frac{2}{3}\phi\phi_x - \frac{2}{3}\phi_{xxx},$$
hence
$$\begin{aligned}\phi_{t_2} &= 2\chi_x - \phi_{xx},\\ \chi_{t_2} &= \chi_{xx} - \frac{2}{3}\phi\phi_x - \frac{2}{3}\phi_{xxx}.\end{aligned}$$
We can eliminate χ to find a single second order equation
$$\phi_{t_2 t_2} = -\frac{1}{3}\phi_{xxxx} - \frac{4}{3}\phi_x^2 - \frac{4}{3}\phi\phi_{xx}, \tag{9.1}$$
which is the Boussinesq equation.
- For $N = 4$ we have $L_4 = \partial^4 + \phi\partial^2 + \chi\partial + \psi$, hence
$$L_4^{1/4} = \partial + \frac{1}{4}\phi\partial^{-1} + \cdots$$
and
$$P_{4,2} = \left(L^{1/2}\right)_+ = \partial^2 + \frac{1}{2}\phi.$$
The commutator is
$$\begin{aligned}[P_{4,2}, L_4] = (2\chi_x - 2\phi_{xx})\partial^2 + (\chi_{xx} + 2\psi_x - 2\phi_{xxx} - \phi\phi_x)\partial\\ + \psi_{xx} - \frac{1}{2}\phi_{xxxx} - \frac{1}{2}\phi\phi_{xx} - \frac{1}{2}\chi\phi_x,\end{aligned}$$
yielding the system of PDEs
$$\phi_{t_2} = 2\chi_x - 2\phi_{xx}, \tag{9.2}$$
$$\chi_{t_2} = \chi_{xx} + 2\psi_x - 2\phi_{xxx} - \phi\phi_x, \tag{9.3}$$
$$\psi_{t_2} = \psi_{xx} - \frac{1}{2}\phi_{xxxx} - \frac{1}{2}\phi\phi_{xx} - \frac{1}{2}\chi\phi_x. \tag{9.4}$$

More on pseudodifferential operators and the continuous GD hierarchy can be found for example in the book [23] or in the lecture notes [9] (which are mostly based on the former).

9.2. Discrete GD hierarchy

A discrete counterpart of the Gelfand-Dickey hierarchy was introduced in [65], its pluri-Lagrangian structure in [48]. The N-th member of the hierarchy is a system of quad equations with $2N-3$ components, which we denote by $V^{[1]}, \dots, V^{[N-2]}$, $W^{[1]}, \dots, W^{[N-2]}$, and $U = V^{[0]} = W^{[0]}$. The equations are

$$V_2^{[j+1]} - V_1^{[j+1]} = \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1 \right) V_{12}^{[j]} - \frac{1}{\lambda_1} V_2^{[j]} + \frac{1}{\lambda_2} V_1^{[j]}, \tag{9.5}$$

$$W_2^{[j+1]} - W_1^{[j+1]} = -\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1 \right) W^{[j]} - \frac{1}{\lambda_2} W_2^{[j]} + \frac{1}{\lambda_1} W_1^{[j]}, \tag{9.6}$$

for $j = 0, \dots, N-3$, and

$$\begin{aligned} V_{12}^{[N-2]} - W^{[N-2]} = \frac{\frac{1}{\lambda_1^N} - \frac{1}{\lambda_2^N}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1} - \gamma_{N-1} + \sum_{i=0}^{N-3} \sum_{j=0}^{N-3-i} \gamma_{N-3-i-j} V_{12}^{[j]} W^{[i]} \\ - \sum_{j=0}^{N-3} \gamma_{N-2-j} \big(V_{12}^{[j]} - W^{[j]} \big), \end{aligned} \tag{9.7}$$

where

$$\gamma_j = (-1)^j \left(\frac{1}{\lambda_1^j} + \frac{1}{\lambda_1^{j-1} \lambda_2} + \ldots + \frac{1}{\lambda_1 \lambda_2^{j-1}} + \frac{1}{\lambda_2^j} \right).$$

Equation (9.7) is presented here in the form it appeared in [48].

It is important to note that the variables $V^{[N-2]}$ and $W^{[N-2]}$ can be eliminated. The reduced system consists of equations (9.5)–(9.6) for $j = 0, \dots, N-4$ and the 9-point equation

$$\begin{aligned} &\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{122} - U_{112} \right) V_{1122}^{[N-3]} - \frac{1}{\lambda_1} V_{122}^{[N-3]} + \frac{1}{\lambda_2} V_{112}^{[N-3]} \\ &+ \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1 \right) W^{[N-3]} + \frac{1}{\lambda_2} W_2^{[N-3]} - \frac{1}{\lambda_1} W_1^{[N-3]} \\ &= \frac{\frac{1}{\lambda_1^N} - \frac{1}{\lambda_2^N}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{22} - U_{12}} + \sum_{i=0}^{N-3} \sum_{j=0}^{N-3-i} \gamma_{N-3-i-j} \big(V_{122}^{[j]} W_2^{[i]} - V_{112}^{[j]} W_1^{[i]} \big) \\ &\quad - \frac{\frac{1}{\lambda_1^N} - \frac{1}{\lambda_2^N}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{12} - U_{11}} - \sum_{j=0}^{N-3} \gamma_{N-2-j} \big(V_{122}^{[j]} - W_2^{[j]} - V_{112}^{[j]} + W_1^{[j]} \big), \end{aligned} \tag{9.8}$$

obtained by evaluating $V_{122}^{[N-2]} - W_2^{[N-2]} - V_{112}^{[N-2]} + W_1^{[N-2]}$ once with Equations (9.5)–(9.6) and once with Equation (9.7).

The Lagrangian given in [48] for the N-th lattice Gelfand-Dickey equation is

$$
\begin{aligned}
L = {} & (-1)^{N+1} \left(\tfrac{1}{\lambda_1^N} - \tfrac{1}{\lambda_2^N} \right) \log\!\left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} - U_1 + U_2 \right) \\
& - \gamma_{N-1}(U_2 - U_1) - \sum_{j=0}^{N-2} \gamma_{N-2-j}(U_2 - U_1) V_{12}^{[j]} \\
& - \sum_{j=1}^{N-2} \sum_{i=0}^{N-2-j} \gamma_{N-2-i-j} W^{[i]} \\
& \qquad \cdot \left(V_2^{[j]} - V_1^{[j]} - \left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} - U_1 + U_2 \right) V_{12}^{[j-1]} + \tfrac{1}{\lambda_1} V_2^{[j-1]} - \tfrac{1}{\lambda_2} V_1^{[j-1]} \right).
\end{aligned}
$$

Or, rearranging terms,

$$
\begin{aligned}
L = {} & (-1)^{N+1} \left(\tfrac{1}{\lambda_1^N} - \tfrac{1}{\lambda_2^N} \right) \log\!\left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} - U_1 + U_2 \right) - \gamma_{N-1}(U_2 - U_1) \\
& - \sum_{j=0}^{N-2} \gamma_{N-2-j}(U_2 - U_1) V_{12}^{[j]} - \sum_{j=1}^{N-2} \sum_{i=0}^{N-2-j} \gamma_{N-2-i-j} W^{[i]} \left(V_2^{[j]} - V_1^{[j]} \right) \\
& + \sum_{j=0}^{N-3} \sum_{i=0}^{N-3-j} \gamma_{N-3-i-j} W^{[i]} \left(\left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} - U_1 + U_2 \right) V_{12}^{[j]} - \tfrac{1}{\lambda_1} V_2^{[j]} + \tfrac{1}{\lambda_2} V_1^{[j]} \right).
\end{aligned}
$$

Note that this Lagrangian depends on the field $V^{[N-2]}$ but not on $W^{[N-2]}$. A more symmetric equivalent Lagrangian is obtained by adding the exact discrete differential form

$$
-UV_1^{[N-2]} - U_1 V_{12}^{[N-2]} + U_2 V_{12}^{[N-2]} + UV_2^{[N-2]} = \Delta\!\left(-UV_i^{[N-2]} \right),
$$

which does not change the Euler-Lagrange equations (see Proposition 6.5). The resulting Lagrangian does not depend on $V^{[N-2]}$ either,

$$
\begin{aligned}
L = {} & (-1)^{N+1} \left(\tfrac{1}{\lambda_1^N} - \tfrac{1}{\lambda_2^N} \right) \log\!\left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} - U_1 + U_2 \right) - \gamma_{N-1}(U_2 - U_1) \\
& - \sum_{j=0}^{N-3} \gamma_{N-2-j}(U_2 - U_1) V_{12}^{[j]} - \sum_{j=1}^{N-3} \sum_{i=0}^{N-2-j} \gamma_{N-2-i-j} W^{[i]} \left(V_2^{[j]} - V_1^{[j]} \right) \\
& + \sum_{j=0}^{N-3} \sum_{i=0}^{N-3-j} \gamma_{N-3-i-j} W^{[i]} \left(\left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} - U_1 + U_2 \right) V_{12}^{[j]} - \tfrac{1}{\lambda_1} V_2^{[j]} + \tfrac{1}{\lambda_2} V_1^{[j]} \right).
\end{aligned}
\tag{9.9}
$$

Now that both variables $V^{[N-2]}$ and $W^{[N-2]}$ are absent from the Lagrangian, it becomes clear that the variational formulation does not produce the quad version of the Equation (9.5)–(9.7), but rather the 9-point version (9.8). In particular this means that, from the Lagrangian perspective, the scalar form of the Boussinesq equation ($N = 3$) is the most natural. The first truly multi-component Lagrangian equation of the hierarchy is found for $N = 4$. The continuum limit of this equation will allow us to formulate a continuous multi-component pluri-Lagrangian system.

Note that the Lagrangian (9.9) depends on fields on a single quad and on the corresponding lattice parameters. Hence it fits the pluri-Lagrangian theory. The multidimensional consistency of the equations can either be checked as consistency around an elementary cube for the quad version of the equation, or in a 27-point cube for the 9-point formulation. For the Boussinesq equation the former was explicitly done in [86] and the latter was numerically verified in [93, Section 5.7]. The multidimensional consistency also follows from the construction of the GD hierarchy using the direct linearization method [65]. In this construction, lattice shifts are identified with Bäcklund transformations, hence the permutability property implies multidimensional consistency.

9.3. Continuum limit of the lattice Boussinesq equation (GD3)

We consider the lattice Boussinesq equation in its 9-point scalar form, i.e. Equation (9.8) for $N = 3$,

$$\begin{aligned}
\frac{\frac{1}{\lambda_1^3} - \frac{1}{\lambda_2^3}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{12} - U_{11}} - \frac{\frac{1}{\lambda_1^3} - \frac{1}{\lambda_2^3}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{22} - U_{12}} - U_2 U_{122} + U_1 U_{112} \\
+ \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{122} - U_{112} \right) U_{1122} + \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1 \right) U \\
- \left(\frac{2}{\lambda_1} + \frac{1}{\lambda_2} \right) (U_1 + U_{122}) + \left(\frac{1}{\lambda_1} + \frac{2}{\lambda_2} \right) (U_2 + U_{112}) = 0.
\end{aligned}$$

For this equation we use the Miwa correspondence (7.6) with $c = -3$. As always, we perform a double series expansion of the lattice equation,

$$\sum_{i,j=0}^{\infty} \mathcal{F}_{ij} \lambda_1^i \lambda_2^j = 0.$$

The first column of coefficients of this expansion is

$$\begin{aligned}
\mathcal{F}_{00} &= 0, \\
\mathcal{F}_{10} &= 18 v_1 v_{11} - \tfrac{9}{2} v_{1111} - \tfrac{3}{2} v_{22}, \\
\mathcal{F}_{20} &= 81 v_{11}^2 + 81 v_1 v_{111} - \tfrac{81}{4} v_{11111} - \tfrac{27}{4} v_{122} - 3 v_{23}, \\
\mathcal{F}_{30} &= -108 v_1^2 v_{11} + 648 v_{11} v_{111} + 243 v_1 v_{1111} + 54 v_1 v_{112} + 54 v_{11} v_{12} - 18 v_{12} v_2 \\
&\quad - 9 v_1 v_{22} - 54 v_{111111} - \tfrac{27}{2} v_{11112} - 9 v_{1122} - 12 v_{123} - \tfrac{9}{2} v_{222} - 3 v_{24} - \tfrac{4}{3} v_{33},
\end{aligned}$$

$$
\begin{aligned}
\mathcal{F}_{40} = &-810v_1v_{11}^2 - 405v_1^2v_{111} - 135v_1^2v_{12} - 135v_1v_{11}v_2 + 1215v_{111}^2 + \tfrac{6885}{4}v_{11}v_{1111}\\
&+ \tfrac{2025}{4}v_1v_{11111} + \tfrac{675}{2}v_1v_{1112} + \tfrac{1215}{2}v_{11}v_{112} + 45v_1v_{113} + \tfrac{675}{2}v_{111}v_{12}\\
&- \tfrac{135}{2}v_{12}^2 - \tfrac{135}{4}v_1v_{122} + 45v_{11}v_{13} + \tfrac{135}{4}v_{1111}v_2 - \tfrac{135}{2}v_{112}v_2 - 15v_{13}v_2\\
&- \tfrac{135}{4}v_{11}v_{22} - \tfrac{45}{4}v_2v_{22} - 15v_1v_{23} - 15v_{12}v_3 - \tfrac{405}{4}v_{1111111} - \tfrac{621}{8}v_{111112}\\
&- \tfrac{45}{4}v_{11113} - 15v_{1123} - \tfrac{135}{8}v_{1222} - \tfrac{45}{4}v_{124} - 5v_{133} - \tfrac{45}{4}v_{223} - 3v_{25} - \tfrac{5}{2}v_{34},\\
&\vdots
\end{aligned}
$$

From $\mathcal{F}_{10} = 0$ we get the equation

$$v_{22} = 12v_1v_{11} - 3v_{1111}.$$

Using this equation we get $v_{23} = 0$ from $\mathcal{F}_{20} = 0$. We take the liberty of integrating this without a constant and take $v_3 = 0$ as the second equation in the hierarchy. We can proceed iteratively and at each step integrate with respect to t_2 to find the hierarchy

$$
\begin{aligned}
v_{22} &= 12v_1v_{11} - 3v_{1111},\\
v_3 &= 0,\\
v_4 &= -6v_1v_2 + 3v_{112},\\
v_5 &= -15v_1^3 + \frac{135}{4}v_{11}^2 + 45v_1v_{111} - \frac{15}{4}v_2^2 - 9v_{11111},\\
&\vdots
\end{aligned}
$$

The first equation of the hierarchy is the potential Boussinesq equation. We observe that every third equation is trivial, $v_{3k} = 0$, as expected from the construction of the hierarchy using pseudodifferential operators. Note that in this hierarchy all equations except the first are evolutionary, analogous to the Toda hierarchy of ODEs.

The Lagrangian is

$$
\begin{aligned}
L(U, U_1, U_2, U_{12}, \lambda_1, \lambda_2) = &\left(\frac{1}{\lambda_1^3} - \frac{1}{\lambda_2^3}\right)\log\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - U_1 + U_2\right)\\
&- \left(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_1\lambda_2} + \frac{1}{\lambda_2^2}\right)(U_2 - U_1) + \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right)(U_2 - U_1)U_{12}\\
&+ \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1\right)UU_{12} - \frac{1}{\lambda_1}UU_2 + \frac{1}{\lambda_2}UU_1.
\end{aligned}
$$

In order to get the necessary leading order cancellation, we add the following terms to the Lagrangian,

$$
\begin{aligned}
&\frac{1}{2}\left(\frac{1}{\lambda_1}UU_1 + \frac{1}{\lambda_2}U_1U_{12} - \frac{1}{\lambda_1}U_2U_{12} - \frac{1}{\lambda_2}UU_2\right) + \frac{1}{3}(U_1^3 - U_2^3)\\
&+ \frac{1}{4}\left(\frac{1}{\lambda_1}(U_1^2 - U^2) + \frac{1}{\lambda_2}(U_{12}^2 - U_1^2) + \frac{1}{\lambda_1}(U_2^2 - U_{12}^2) + \frac{1}{\lambda_2}(U^2 - U_2^2)\right).
\end{aligned}
$$

$$\mathcal{L}_{12} = 2v_1^3 + v_{11}^2 - \tfrac{3}{4}v_2^2$$

$$\mathcal{L}_{13} = -\tfrac{3}{2}v_2v_3$$

$$\mathcal{L}_{14} = -4v_1^4 - 12v_1v_{11}^2 + 6vv_{12}v_2 - \tfrac{4}{3}v_{111}^2 - v_{12}^2 - \tfrac{3}{2}v_2v_4$$

$$\mathcal{L}_{15} = -10v_1^3v_2 + 10v_1^2v_{112} - 5v_{11}^2v_2 - \tfrac{5}{4}v_2^3 + \tfrac{8}{3}v_{11}v_{1112} - \tfrac{3}{2}v_2v_5$$

$$\mathcal{L}_{23} = -6v_1^2v_3 - 2v_{11}v_{13} + 2v_{111}v_3$$

$$\begin{aligned}\mathcal{L}_{24} = {} & -16v_1^3v_2 - 8v_1v_{11}v_{12} - 4v_{11}^2v_2 + 8v_1v_{111}v_2 + 2v_2^3 + 6vv_2v_{22} \\ & - 6v_1^2v_4 - \tfrac{8}{3}v_{111}v_{112} - 2v_{11}v_{14} - 2v_{12}v_{22} + 2v_{111}v_4\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{25} = {} & -16v_1^5 + 60v_1^2v_{11}^2 + \tfrac{40}{3}v_1^3v_{111} - 15v_1^2v_2^2 + \tfrac{4}{3}v_{11}^2v_{111} - \tfrac{8}{3}v_1v_{111}^2 - \tfrac{64}{3}v_1v_{11}v_{1111} \\ & + 10v_1v_{12}^2 + 10v_1^2v_{122} - 10v_{11}v_{12}v_2 + 5v_{111}v_2^2 - 20v_1v_{11}v_{22} - 6v_1^2v_5 \\ & + \tfrac{16}{9}v_{1111}^2 + \tfrac{4}{3}v_{112}^2 + \tfrac{8}{3}v_{11}v_{1122} - \tfrac{8}{3}v_{111}v_{122} - 2v_{11}v_{15} + \tfrac{8}{3}v_{1111}v_{22} + 2v_{111}v_5\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{34} = {} & -16v_1^3v_3 - 24v_1v_{11}v_{13} + 6vv_2v_{23} + 12v_{11}^2v_3 + 24v_1v_{111}v_3 \\ & - \tfrac{8}{3}v_{111}v_{113} + \tfrac{8}{3}v_{1111}v_{13} - 2v_{12}v_{23} - \tfrac{8}{3}v_{11111}v_3\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{35} = {} & -30v_1^2v_2v_3 + 10v_1^2v_{123} - 10v_{11}v_{13}v_2 - 20v_1v_{11}v_{23} + 20v_1v_{112}v_3 + 10v_{11}v_{12}v_3 \\ & + 10v_{111}v_2v_3 + \tfrac{8}{3}v_{11}v_{1123} - \tfrac{8}{3}v_{111}v_{123} + \tfrac{8}{3}v_{1112}v_{13} + \tfrac{8}{3}v_{1111}v_{23} - \tfrac{8}{3}v_{11112}v_3\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{45} = {} & \tfrac{160}{3}v_1^6 - 160v_1^3v_{11}^2 - 160v_1^4v_{111} - 40v_1^3v_2^2 + 76v_{11}^4 + 176v_1v_{11}^2v_{111} + 144v_1^2v_{111}^2 \\ & + 32v_1^2v_{11}v_{1111} + \tfrac{160}{9}v_1^3v_{11111} - 20v_1^2v_{12}^2 + 80v_1^2v_{112}v_2 + 40v_1v_{11}v_{12}v_2 \\ & - 50v_{11}^2v_2^2 - 20v_1v_{111}v_2^2 + \tfrac{15}{4}v_2^4 - 30v_1^2v_2v_4 + 16v_1^3v_5 - \tfrac{32}{27}v_{111}^3 + \tfrac{32}{9}v_{11}v_{111}v_{1111} \\ & - \tfrac{32}{9}v_1v_{1111}^2 - \tfrac{80}{3}v_{11}^2v_{11111} - \tfrac{320}{9}v_1v_{111}v_{11111} - \tfrac{56}{3}v_1v_{112}^2 + \tfrac{32}{3}v_1v_{1112}v_{12} \\ & - \tfrac{64}{3}v_{11}v_{112}v_{12} + \tfrac{8}{3}v_{111}v_{12}^2 + 10v_1^2v_{124} + 24v_1v_{11}v_{15} - \tfrac{32}{3}v_1v_{11112}v_2 + \tfrac{32}{3}v_{11}v_{1112}v_2 \\ & - \tfrac{32}{3}v_{111}v_{112}v_2 - \tfrac{8}{3}v_{1111}v_{12}v_2 - 10v_{11}v_{14}v_2 + \tfrac{20}{3}v_{11111}v_2^2 - 20v_1v_{11}v_{24} \\ & - 6vv_2v_{25} + 20v_1v_{112}v_4 + 10v_{11}v_{12}v_4 + 10v_{111}v_2v_4 - 12v_{11}^2v_5 - 24v_1v_{111}v_5 \\ & + \tfrac{64}{27}v_{11111}^2 - \tfrac{16}{9}v_{1112}^2 + \tfrac{32}{9}v_{11112}v_{112} + \tfrac{8}{3}v_{11}v_{1124} + \tfrac{8}{3}v_{111}v_{115} - \tfrac{8}{3}v_{111}v_{124} \\ & + \tfrac{8}{3}v_{1112}v_{14} - \tfrac{8}{3}v_{1111}v_{15} + \tfrac{8}{3}v_{1111}v_{24} + 2v_{12}v_{25} - \tfrac{8}{3}v_{11112}v_4 + \tfrac{8}{3}v_{11111}v_5\end{aligned}$$

Table 9.1. *Coefficients $\mathcal{L}_{ij}$ for the Boussinesq hierarchy, after eliminating alien derivatives.*

These terms do not contribute to the Euler-Lagrange equations because they are the discrete exterior derivative (see Proposition 6.5) of the 1-form

$$\eta(U, U_i, \lambda_i) = \frac{1}{2\lambda_i} U U_i + \frac{1}{3} U^3 + \frac{1}{4\lambda_i} (U_i^2 - U^2).$$

Some coefficients of the Lagrangian 2-form are given in Table 9.1. Unfortunately, not all alien derivatives can be eliminated. Since we do not have any equation in the hierarchy to eliminate first derivatives with respect to t_2, the derivatives $v_2, v_{12}, \ldots$ remain in place. All other derivatives are eliminated from the coefficients where they are alien.

9.4. Continuum limit of lattice GD4

In order to get a truly multicomponent pluri-Lagrangian system we move to the next equation of the GD hierarchy, with $N = 4$. In its 3-component form the discrete GD4 equation reads

$$\begin{aligned}
&\left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} + U_{122} - U_{112}\right) V_{1122} + \left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} + U_2 - U_1\right) W \\
&- \tfrac{1}{\lambda_1} V_{122} + \tfrac{1}{\lambda_2} V_{112} + \tfrac{1}{\lambda_2} W_2 - \tfrac{1}{\lambda_1} W_1 - (V_{122} U_2 - V_{112} U_1 + U_{122} W_2 - U_{112} W_1) \\
&- \left(\tfrac{1}{\lambda_1} + \tfrac{1}{\lambda_2}\right) (V_{122} - V_{112} - W_2 + W_1) \\
&\quad = \frac{\frac{1}{\lambda_1^4} - \frac{1}{\lambda_2^4}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{12} - U_{11}} - \frac{\frac{1}{\lambda_1^4} - \frac{1}{\lambda_2^4}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_{22} - U_{12}} - \left(\tfrac{1}{\lambda_1} + \tfrac{1}{\lambda_2}\right) (U_2 U_{122} + U_1 U_{112}) \\
&\qquad - \left(\tfrac{1}{\lambda_1^2} + \tfrac{1}{\lambda_1 \lambda_2} + \tfrac{1}{\lambda_2^2}\right) (U_{122} - U_{112} - U_2 + U_1),
\end{aligned} \tag{9.10}$$

$$V_2 - V_1 = \left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} + U_2 - U_1\right) U_{12} - \tfrac{1}{\lambda_1} U_2 + \tfrac{1}{\lambda_2} U_1, \tag{9.11}$$

$$W_2 - W_1 = -\left(\tfrac{1}{\lambda_1} - \tfrac{1}{\lambda_2} + U_2 - U_1\right) U - \tfrac{1}{\lambda_2} U_2 + \tfrac{1}{\lambda_1} U_1. \tag{9.12}$$

Equations (9.11) and (9.12) do not just look similar, also their expansions are nearly identical in leading order:

$$\begin{aligned}
(\lambda_2 - \lambda_1) v_1 &= (\lambda_2 - \lambda_1) u_1 u + \frac{1}{2} (\lambda_2 - \lambda_1) u_{11} + \frac{1}{2} (\lambda_2 - \lambda_1) u_2 + \mathcal{O}\big((\lambda_1 + \lambda_2)^2\big), \\
(\lambda_2 - \lambda_1) w_1 &= -(\lambda_2 - \lambda_1) u_1 u - \frac{1}{2} (\lambda_2 - \lambda_1) u_{11} + \frac{1}{2} (\lambda_2 - \lambda_1) u_2 + \mathcal{O}\big((\lambda_1 + \lambda_2)^2\big),
\end{aligned}$$

where we used the constant $c = 1$ in the Miwa correspondence (7.6). This gives us the ODE $v_1 - w_1 = 2 u_1 u + u_{11}$, which integrates to

$$v - w = u^2 + u_1. \tag{9.13}$$

We have omitted the integration constant because higher order terms in the expansion force it to be zero. This relation allows us the eliminate either v or w from the continuous system. Hence we lose one of the components in the continuum limit. For convenience we make a change of variables in the discrete system,

$$V = \frac{X+Y}{2}, \qquad W = \frac{X-Y}{2}.$$

Then in the continuum limit the variable y can be eliminated by Equation (9.13), which in the new variables reads

$$y = u^2 + u_1. \tag{9.14}$$

One recognizes the Miura transformation [54]. For the remaining two variables we find the hierarchy

$$\begin{aligned} u_2 &= x_1, & x_{22} &= -4u_{11}x_1 - 8u_1x_{11} - x_{1111}, \\ u_3 &= \frac{3}{2}u_1^2 + \frac{1}{4}u_{111} + \frac{3}{4}x_2, & x_3 &= -3u_1x_1 - \frac{1}{2}x_{111}, \\ u_4 &= 0, & x_4 &= 0, \\ &\vdots & &\vdots \end{aligned} \tag{9.15}$$

where, presumably, every fourth pair of equations is trivial.

9.4.1. Comparison with Section 9.1

Just as for the KdV and Boussinesq equations, the leading equation of the continuum limit is a potential version of the GD4 equation. Additionally, we need to eliminate one of the three variables from the system (9.2)–(9.4) to connect it to the result of the continuum limit. Making an educated guess, we introduce the variables x and u by

$$\phi = 4u_1, \qquad \chi = 2x_1 + 4u_{11}.$$

From Equation (9.2) we find $u_{12} = x_{11}$, which integrates to

$$u_2 = x_1 \tag{9.16}$$

From Equations (9.2)–(9.3) we find

$$x_{12} = -\chi_{11} + 2\psi_1 - \phi\phi_1.$$

Integrating with respect to t_1 and differentiating with respect to t_2 we find

$$x_{22} = -x_{1111} - 4x_1u_{11} - 8x_{11}u_1. \tag{9.17}$$

Equations (9.16)–(9.17) are exactly the leading order equations of the continuum limit.

There is no obvious reason in pseudodifferential approach to eliminate one of the three variables. On the contrary, it seems more natural to leave the system in its first-order evolutionary form. The reduction to two variables is forced upon us by the continuum limit, because there does not seem to be any way of performing the limit without getting an ODE relation between the variables, as in Equation (9.13).

9.4.2. Pluri-Lagrangian structure

For $N = 4$ we have the Lagrangian

$$\begin{aligned}
&L(U, V, W, U_1, V_1, W_1, U_2, V_2, W_2, U_{12}, V_{12}, W_{12}, \lambda_1, \lambda_2) \\
&= \left(\frac{1}{\lambda_1^4} - \frac{1}{\lambda_2^4} \right) \log \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - U_1 + U_2 \right) - \left(\frac{1}{\lambda_1^3} + \frac{1}{\lambda_1^2 \lambda_2} + \frac{1}{\lambda_1 \lambda_2^2} + \frac{1}{\lambda_2^3} \right) (U_2 - U_1) \\
&\quad + \left(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_1 \lambda_2} + \frac{1}{\lambda_2^2} \right) (U_2 - U_1) U_{12} - \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} \right) (U_2 - U_1) V_{12} \\
&\quad - U \left(\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1 \right) V_{12} - \frac{1}{\lambda_1} V_2 + \frac{1}{\lambda_2} V_1 \right) \\
&\quad - \left(\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} \right) U - W \right) \left(V_2 - V_1 - \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} + U_2 - U_1 \right) U_{12} + \frac{1}{\lambda_1} U_2 - \frac{1}{\lambda_2} U_1 \right).
\end{aligned} \tag{9.18}$$

As before, we make the change of variables

$$V = \frac{X + Y}{2}, \qquad W = \frac{X - Y}{2}.$$

In order to achieve the necessary leading order cancellation, we add to the Lagrangian the discrete exterior derivative of the following discrete 1-form:

$$\begin{aligned}
\eta(U, X, Y, U_i, X_i, Y_i, \lambda_i) &= \frac{1}{2\lambda_i^2} U U_i - \frac{1}{4\lambda_i^2} (U^2 - U_i^2) + \frac{1}{6\lambda_i} (U U_i^2 + U^2 U_i) \\
&\quad - \frac{1}{2\lambda_i} (U X_i + U_i X) + \frac{1}{2\lambda_i} U X - \frac{1}{2\lambda_i} U Y_i \\
&\quad - \frac{1}{8} (Y Y_i - X X_i - X Y_i - X_i Y) - \frac{1}{4} (U_i^2 X + U^2 X_i).
\end{aligned}$$

This 1-form was found by trial and error. It would be interesting to establish a general strategy to find a Lagrangian within a given equivalence class that provides the required cancellation. Without such a strategy, it seems infeasible to apply the continuum limit to higher members of the Gelfand-Dickey hierarchy. Then again, judging from the complexity of the pluri-Lagrangian 2-form for the continuous GD4 hierarchy, with just three of its coefficients filling Table 9.2, it might not be very useful to find explicit expressions for pluri-Lagrangian structures for the higher Gelfand-Dickey hierarchies.

$$
\begin{aligned}
\mathcal{L}_{12} = {} & \tfrac{1}{3}u^3u_{12} - \tfrac{1}{3}u^3x_{11} - \tfrac{1}{3}u^2u_{112} - \tfrac{7}{6}u_1^2u_2 + u_1u_{11}x - \tfrac{1}{2}u_{11}xy \\
& - \tfrac{1}{2}u_1x_1y + \tfrac{1}{2}uxy_{11} - \tfrac{1}{4}u^2y_{12} - \tfrac{1}{4}u_1u_{112} + \tfrac{1}{4}u_{112}x - \tfrac{1}{4}u_1x_{111} \\
& + \tfrac{1}{4}ux_{112} - \tfrac{1}{2}xx_{12} - u_2x_2 - \tfrac{1}{4}x_{12}y - \tfrac{1}{4}x_{11}y_1 + \tfrac{1}{4}uy_{112} + \tfrac{1}{4}xy_{12}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{L}_{13} = {} & \tfrac{3}{2}u^2u_1^3 + \tfrac{23}{12}u_1^4 - \tfrac{1}{4}u^2u_{11}^2 + \tfrac{1}{3}u^3u_{13} - \tfrac{3}{2}u_1^3x - 3uu_1u_{11}x - \tfrac{1}{4}u^3x_{12} + \tfrac{3}{4}u_1^3y + \tfrac{3}{2}uu_1u_{11}y \\
& - \tfrac{3}{2}u_1u_{11}^2 - \tfrac{1}{3}u^2u_{113} - \tfrac{7}{6}u_1^2u_3 + \tfrac{3}{4}u_1u_{11}x_1 + 3ux_1x_{11} + \tfrac{1}{4}uu_1x_{111} - \tfrac{1}{4}u^2x_{13} \\
& + \tfrac{9}{8}u_1^2x_2 + \tfrac{3}{8}ux_{12}y + \tfrac{3}{8}u_1x_2y - \tfrac{3}{8}u_1u_{11}y_1 + \tfrac{3}{4}ux_{11}y_1 + \tfrac{3}{4}ux_1y_{11} - \tfrac{1}{8}uu_1y_{111} \\
& - \tfrac{1}{4}u^2y_{13} + \tfrac{1}{16}u_{111}^2 - \tfrac{1}{4}u_1u_{113} + \tfrac{1}{4}u_{113}x - \tfrac{1}{8}u_{111}x_{11} + \tfrac{1}{4}x_{11}^2 + \tfrac{1}{16}u_1x_{112} + \tfrac{1}{2}ux_{113} \\
& + \tfrac{1}{2}xx_{13} - u_3x_2 + \tfrac{3}{8}x_2^2 + \tfrac{3}{16}x_{112}y + \tfrac{1}{16}u_{111}y_{11} - \tfrac{1}{8}x_1y_{111} + \tfrac{1}{4}uy_{113} + \tfrac{1}{4}xy_{13}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{L}_{23} = {} & \tfrac{3}{2}u^3u_1u_{12} + \tfrac{3}{4}u^2u_1^2u_2 + \tfrac{3}{2}uu_1^3x - \tfrac{3}{2}u^2u_1u_{11}x - \tfrac{9}{4}u^2u_1^2x_1 + \tfrac{3}{2}u^3u_{11}x_1 + \tfrac{3}{2}u^3u_1x_{11} \\
& + \tfrac{7}{4}uu_1^2u_{12} - \tfrac{1}{2}u^2u_{11}u_{12} + \tfrac{14}{3}u_1^3u_2 - \tfrac{5}{2}uu_1u_{11}u_2 - \tfrac{3}{4}u_1^2u_{11}x + \tfrac{1}{4}uu_1u_{111}x \\
& - \tfrac{1}{8}u^2u_{1111}x - 3uu_1u_{12}x + \tfrac{1}{2}u^2u_{13}x - \tfrac{3}{2}u_1^2u_2x - uu_1u_3x - 5u_1^3x_1 + 2uu_1u_{11}x_1 \\
& + \tfrac{3}{4}u^2u_{111}x_1 + 3uu_1u_2x_1 + u^2u_3x_1 - 3uu_1x_1^2 + \tfrac{1}{4}uu_1^2x_{11} + \tfrac{5}{2}u^2u_{11}x_{11} + \tfrac{3}{4}u^2x_1x_{11} \\
& + u^2u_1x_{111} + \tfrac{3}{8}u^3x_{1111} + \tfrac{3}{4}u^2u_1x_{12} - \tfrac{3}{8}u^2xx_{12} + \tfrac{1}{3}u^3x_{13} - \tfrac{3}{2}uu_1^2x_2 - \tfrac{3}{8}u^2u_2x_2 \\
& + \tfrac{3}{4}uu_1xx_2 - \tfrac{3}{8}u^2x_1x_2 + \tfrac{1}{8}u^3x_{22} - \tfrac{1}{2}u^2u_1x_3 + \tfrac{3}{2}u_1u_{11}xy - \tfrac{3}{4}u_1^2x_1y - \tfrac{3}{2}uu_{11}x_1y \\
& - \tfrac{3}{2}uu_1x_{11}y - \tfrac{3}{4}u_1^2xy_1 + \tfrac{3}{2}uu_1x_1y_1 + \tfrac{3}{4}uu_1^2y_2 + \tfrac{1}{8}u_1^2u_{112} + \tfrac{1}{8}uu_{11}u_{112} - 3u_1u_{11}u_{12} \\
& + \tfrac{1}{24}uu_{111}u_{12} + \tfrac{3}{8}u_{11}^2u_2 + \tfrac{13}{8}u_1u_{111}u_2 - \tfrac{1}{6}uu_{1111}u_2 + \tfrac{1}{8}u_{11}u_{111}x - \tfrac{1}{8}u_1u_{1111}x \\
& - \tfrac{1}{2}u_1u_{13}x - \tfrac{1}{2}u_{11}u_3x - \tfrac{7}{8}u_{11}^2x_1 - \tfrac{7}{8}u_1u_{111}x_1 + \tfrac{1}{6}uu_{1111}x_1 + \tfrac{3}{2}u_1u_{12}x_1 - \tfrac{2}{3}uu_{13}x_1 \\
& - \tfrac{3}{4}u_{11}u_2x_1 + 4u_1u_3x_1 + \tfrac{3}{4}u_{11}x_1^2 + x_1^3 + \tfrac{7}{2}u_1u_{11}x_{11} - \tfrac{1}{24}uu_{111}x_{11} - \tfrac{3}{2}u_1u_2x_{11} \\
& + \tfrac{2}{3}uu_3x_{11} + \tfrac{3}{4}u_1x_1x_{11} + \tfrac{5}{8}u_1^2x_{111} - \tfrac{3}{8}uu_{11}x_{111} + \tfrac{1}{4}uu_2x_{111} - \tfrac{1}{4}ux_1x_{111} + \tfrac{1}{4}uu_1x_{1111} \\
& + \tfrac{3}{16}u^2x_{11111} + \tfrac{1}{8}u^2x_{1112} + \tfrac{1}{2}uu_{12}x_2 + \tfrac{3}{4}u_1^2x_{12} - \tfrac{1}{2}uu_2x_{12} - \tfrac{3}{8}u_1xx_{12} + \tfrac{7}{2}ux_1x_{12} \\
& + \tfrac{3}{16}u^2x_{122} + \tfrac{1}{2}uu_1x_{13} + \tfrac{3}{8}u_1u_2x_2 + \tfrac{3}{8}u_{11}xx_2 - \tfrac{27}{8}u_1x_1x_2 - \tfrac{1}{2}ux_{11}x_2 - \tfrac{1}{2}uu_{11}x_3 \\
& - \tfrac{1}{2}ux_1x_3 + \tfrac{1}{8}u_{1111}xy - \tfrac{3}{4}u_{111}x_1y - \tfrac{9}{4}u_{11}x_{11}y - \tfrac{3}{4}x_1x_{11}y - \tfrac{7}{4}u_1x_{111}y - \tfrac{3}{8}ux_{1111}y \\
& - \tfrac{3}{4}u_1x_{12}y + \tfrac{3}{8}xx_{12}y + \tfrac{3}{8}x_1x_2y - \tfrac{1}{8}u_{111}xy_1 + \tfrac{1}{8}u_{11}x_1y_1 + \tfrac{3}{4}x_1^2y_1 + \tfrac{1}{4}u_1x_{11}y_1 \\
& + \tfrac{3}{8}ux_{111}y_1 + \tfrac{3}{4}ux_{12}y_1 - \tfrac{3}{8}xx_2y_1 - \tfrac{1}{8}u_1x_1y_{11} - \tfrac{1}{8}ux_{11}y_{11} - \tfrac{1}{8}uu_1y_{112} - \tfrac{1}{4}u_1^2y_{12} \\
& + \tfrac{1}{8}uu_{11}y_{12} + \tfrac{3}{4}ux_1y_{12} - \tfrac{1}{2}uxy_{13} + \tfrac{3}{8}ux_2y_2 + \tfrac{1}{2}u_1xy_3 + \tfrac{3}{16}u_{111}u_{112} - \tfrac{1}{8}u_{1111}u_{12} \\
& + \tfrac{1}{16}u_{11111}u_2 - \tfrac{1}{16}u_{11111}x_1 + \tfrac{1}{8}u_{1111}x_{11} - \tfrac{1}{8}u_{112}x_{11} - \tfrac{1}{4}u_{13}x_{11} - \tfrac{3}{16}u_{111}x_{111} \\
& + \tfrac{1}{4}u_{12}x_{111} + \tfrac{3}{4}u_3x_{111} - \tfrac{1}{4}x_{11}x_{111} - \tfrac{1}{16}u_{11}x_{1111} - \tfrac{1}{4}u_2x_{1111} + \tfrac{1}{4}x_1x_{1111} + \tfrac{3}{16}u_1x_{11111} \\
& + \tfrac{1}{8}u_1x_{1112} - \tfrac{1}{8}u_{11}x_{112} - \tfrac{5}{16}u_2x_{112} + \tfrac{5}{16}x_1x_{112} + \tfrac{1}{4}u_1x_{113} + \tfrac{3}{16}u_{12}x_{12} - \tfrac{1}{4}u_3x_{12} \\
& + \tfrac{3}{8}x_{11}x_{12} + \tfrac{1}{4}u_1x_{122} + \tfrac{1}{4}ux_{123} + \tfrac{1}{2}x_1x_{13} + \tfrac{1}{4}u_{13}x_2 - \tfrac{9}{16}x_{111}x_2 - \tfrac{1}{4}u_{11}x_{22} \\
& + xx_{23} - \tfrac{1}{2}x_{11}x_3 + \tfrac{1}{2}x_2x_3 - \tfrac{3}{16}x_{11111}y - \tfrac{1}{8}x_{1112}y + \tfrac{1}{16}x_{1111}y_1 + \tfrac{1}{8}x_{112}y_1 \\
& - \tfrac{1}{8}x_{12}y_{11} - \tfrac{1}{8}x_1y_{112} + \tfrac{1}{8}x_{11}y_{12} - \tfrac{3}{16}x_2y_{12} + \tfrac{1}{4}x_1y_{13} + \tfrac{3}{16}x_{12}y_2 + \tfrac{1}{4}x_2y_3
\end{aligned}
$$

Table 9.2. *Coefficients $\mathcal{L}_{ij}$ for the 4th member of the GD hierarchy*

Note that the Lagrangian 2-form depends on all three fields, u, x, and y. The multi-time Euler-Lagrange equations are equivalent to Equations (9.14) and (9.15). That is, they contain both the constraint on y and the hierarchy in u and x.

As with the Boussinesq hierarchy, the elimination of alien derivatives needs a comment. We do not have any equation in the hierarchy to eliminate the derivatives $x_2, x_{12}, \ldots$, so these have to be tolerated in the coefficient $\mathcal{L}_{13}$. All other alien derivatives, in particular those of u and y, are eliminated as usual.

10. Summary and outlook: variational principles in integrable systems

Variational principles and integrability are rarely seen together in the literature. Even though many integrable systems do posses a variational structure, this fact alone is not nearly enough to make a system integrable. To address this missing link of variational integrability, an exotic variational principle has been developed in the last decade, mainly by the groups of Frank Nijhoff in Leeds and Yuri Suris in Berlin. It has become known as the theory of Lagrangian multiforms, or, of pluri-Lagrangian systems.

One of the appealing features of the pluri-Lagrangian framework is that exactly the same idea applies both in the discrete and the continuous world. A natural question is how to connect these two worlds. Since integrable discretization is notoriously difficult, we opted for the route of continuum limits. For many lattice equations, continuum limits are known, though not always easily found in the literature. One aim of this work was to provide an accessible discussion of some of these limits, but its scientific contribution lies in the fact that we applied the continuum limit to the pluri-Lagrangian structure as well. This way we obtained many previously unknown pluri-Lagrangian formulations for integrable hierarchies of PDEs, including the first instance of a continuous multi-component pluri-Lagrangian system.

The theory of pluri-Lagrangian systems is young and in need of further development. However, initial results show interesting connections with Emmy Noether's classical theory of variational symmetries and with the Hamiltonian formulation of integrable hierarchies. Developing these connections is ongoing work.

Pluri-Lagrangian systems are a potentially useful tool in the field of quantum integrable systems. Even though quantum mechanics is usually formulated as a Hamiltonian theory, when we move to quantum field theory the Lagrangian perspective is the preferred one. Initial steps towards a pluri-Lagrangian theory in a quantum context were made in [43].

Another big question where the pluri-Lagrangian point of view might provide some insight, is why high-dimensional integrable systems are rare. In one dimension (ODEs in the continuous case, or maps in the discrete case) countless integrable equations are known, and a pluri-Lagrangian structure is often obtained from the classical variational formulation. Indeed, a continuous pluri-Lagrangian 1-form

$$\mathcal{L} = \sum_j \mathcal{L}_j \,\mathrm{d}t_j$$

is typically built out of classical Lagrangians $\mathcal{L}_j$ for the individual equations in the hierarchy. For this reason we have not given much attention to 1-dimensional systems in the present work. For a pluri-Lagrangian 2-form,

$$\mathcal{L} = \sum_{i,j} \mathcal{L}_{ij} \, \mathrm{d}t_i \wedge \mathrm{d}t_j,$$

only the coefficients in the first row, the $\mathcal{L}_{1j}$, correspond to individual equations. The other coefficients do not have an interpretation as classical Lagrangians. These additional coefficients are a potential obstacle to the existence of pluri-Lagrangian structure. As we have seen, for many integrable hierarchies of $(1+1)$-dimensional PDEs this obstacle can be overcome by constructing all coefficients as a continuum limit.

If we increase the dimension, the number of coefficients without a classical interpretation grows. For 3-forms one would expect only the coefficients $\mathcal{L}_{12k}$ to correspond to a classical variational principle. The number of terms that do not correspond to an individual equation now grows cubically with the dimension of multi-time, as opposed to quadratically in the two-from case. In this sense, the potential obstructions to the existence of a pluri-Lagrangian structure increase with the dimension. This might explain why high-dimensional integrable systems are uncommon. No continuous pluri-Lagrangian 3-forms have been found so far. In the discrete case there is only the example of the AKP equation [49] and its Lagrangian is somewhat unsatisfactory as it is not naturally skew-symmetric. Unfortunately, so far it has also resisted our efforts to perform a continuum limit.

Before one can use the pluri-Lagrangian formalism to attempt to answer any of these big questions, the theory of pluri-Lagrangian systems itself needs to mature. I hope that (part II of) this thesis has provided a small but meaningful contribution in that direction.

Bibliography

[1] Abramowitz M. & Stegun I. A. *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables.* Courier Corporation, 1964.

[2] Adler V. E. *Bäcklund transformation for the Krichever-Novikov equation.* International Mathematics Research Notices, 1998 : 1–4, 1998.

[3] Adler V. E. & Suris Yu. B. *Q4: Integrable master equation related to an elliptic curve.* International Mathematics Research Notices, 2004 : 2523–2553, 2004.

[4] Adler V. E., Bobenko A. I. & Suris Yu. B. *Classification of integrable equations on quad-graphs. The consistency approach.* Communications in Mathematical Physics, 233 : 513–543, 2003.

[5] Adler V. E., Bobenko A. I. & Suris Yu. B. *Discrete nonlinear hyperbolic equations. Classification of integrable cases.* Functional Analysis and Its Applications, 43 : 3–17, 2009.

[6] Anderson I. M. *Introduction to the variational bicomplex.* In Gotay M., Marsden J. & Moncrief V., editors, *Mathematical Aspects of Classical Field Theory*, pages 51–73. AMS, 1992.

[7] Arnold V. I. *Mathematical Methods of Classical Mechanics.* Volume 60 of *Graduate Texts in Mathematics.* Springer, 1989.

[8] Arnold V. I. *On a theorem of liouville concerning integrable problems of dynamics.* In Givental A. B. & al, editors, *Collected Works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965*, pages 418–422. Springer, 2009.

[9] Batlle C. *Lecture notes on KdV hierarchies and pseudodifferential operators.* `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.30.9553`, 1992.

[10] Bobenko A. I. & Suris Yu. B. *On the Lagrangian structure of integrable quad-equations.* Letters in Mathematical Physics, 92 : 17–31, 2010.

[11] Bobenko A. I. & Suris Yu. B. *Discrete pluriharmonic functions as solutions of linear pluri-Lagrangian systems.* Communications in Mathematical Physics, 336 : 199–215, 2015.

[12] Boll R., Petrera M. & Suris Yu. B. *Multi-time Lagrangian 1-forms for families of Bäcklund transformations: Toda-type systems.* Journal of Physics A: Mathematical and Theoretical, 46 : 275204, 2013.

[13] Boll R., Petrera M. & Suris Yu. B. *What is integrability of discrete variational systems?* Proceedings of the Royal Society A, 470 : 20130550, 2014.

[14] Calvo M. P., Murua A. & Sanz-Serna J. M. *Modified equations for ODEs.* In Kloeden P. E., editor, *Chaotic Numerics: An International Workshop on the Approximation and Computation of Complicated Dynamical Behavior*, pages 63–74. AMS, 1994.

[15] Capel H. W., Nijhoff F. W. & Papageorgiou V. G. *Complete integrability of Lagrangian mappings and lattices of KdV type.* Physics Letters A, 155 : 377–387, 1991.

[16] Chartier P., Hairer E. & Vilmart G. *Numerical integrators based on modified differential equations.* Mathematics of computation, 76 : 1941–1953, 2007.

[17] Chin S. A. *Symplectic integrators from composite operator factorizations.* Physics Letters A, 226 : 344–348, 1997.

[18] Chin S. A. *Physics of symplectic integrators: Perihelion advances and symplectic corrector algorithms.* Physical Review E, 75 : 036701, 2007.

[19] Chin S. A. & Kidwell D. W. *Higher-order force gradient symplectic algorithms.* Physical Review E, 62 : 8746, 2000.

[20] Curtis L. J., Haar R. R. & Kummer M. *An expectation value formulation of the perturbed Kepler problem.* American Journal of Physics, 55 : 627–631, 1987.

[21] De La Torre G. & Murphey T. D. *On the benefits of surrogate Lagrangians in optimal control and planning algorithms.* In *Decision and Control, 55th Conference on*, pages 7384–7391. IEEE, 2016.

[22] De La Torre G. & Murphey T. D. *Surrogate Lagrangians for variational integrators: High order convergence with low order schemes.* arXiv:1709.03883, 2017.

[23] Dickey L. A. *Soliton Equations and Hamiltonian Systems.* World Scientific, 2nd edition, 2003.

[24] Dirac P. A. M. *The Lagrangian in quantum mechanics.* Physikalische Zeitschrift der Sowjetunion, 3 : 64–72, 1933.

[25] Faou E. *Geometric Numerical Integration and Schrödinger Equations.* European Mathematical Society, 2012.

[26] Forest E. & Ruth R. D. *Fourth order symplectic integration.* Physica D: Nonlinear Phenomena, 43 : 105–117, 1989.

[27] Fu W. & Nijhoff F. *Linear integral equations, infinite matrices and soliton hierarchies.* arXiv:1703.08137, 2017.

[28] Fu W. & Nijhoff F. *On reductions of the discrete Kadomtsev–Petviashvili-type equations.* Journal of Physics A: Mathematical and Theoretical, 50 : 505203, 2017.

[29] Gardner C. S. *Korteweg-de Vries equation and generalizations. IV. the Korteweg-de Vries equation as a Hamiltonian system.* Journal of Mathematical Physics, 12 : 1548–1551, 1971.

[30] Gelfand I. M. & Fomin S. V. *Calculus of Variations.* Prentice-Hall, 1963.

[31] Goldstein H. *Classical Mechanics.* Addison-Wesley Pub. Co., 1980.

[32] Hairer E. *Backward analysis of numerical integrators and symplectic methods.* Annals of Numerical Mathematics, 1 : 107–132, 1994.

[33] Hairer E. *Backward error analysis for multistep methods.* Numerische Mathematik, 84 : 199–232, 1999.

[34] Hairer E., Lubich C. & Wanner G. *Geometric numerical integration illustrated by the Störmer–Verlet method.* Acta Numerica, 12 : 399–450, 2003.

[35] Hairer E., Lubich C. & Wanner G. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations.* Springer, 2nd edition, 2006.

[36] Harris J. *Algebraic Geometry: A First Course.* Springer, 1992.

[37] Hietarinta J. & Zhang D.-J. *Soliton solutions for ABS lattice equations: II. Casoratians and bilinearization.* Journal of Physics A: Mathematical and Theoretical, 42 : 404006, 2009.

[38] Hietarinta J., Joshi N. & Nijhoff F. W. *Discrete Systems and Integrability.* Cambridge University Press, 2016.

[39] Hirota R. *Nonlinear partial difference equations I. A difference analogue of the Korteweg-de Vries equation.* Journal of the Physical Society of Japan, 43 : 1424–1433, 1977.

[40] Hirota R. *Nonlinear partial difference equations III. Discrete sine-Gordon equation.* Journal of the Physical Society of Japan, 43 : 2079–2086, 1977.

[41] Hydon P. E. & Mansfield E. L. *A variational complex for difference equations.* Foundations of Computational Mathematics, 4 : 187–217, 2004.

[42] Kasman A. *Glimpses of Soliton Theory: The Algebra and Geometry of Nonlinear PDEs.* Volume 54 of *Student Mathematical Library*. AMS, 2010.

[43] King S. D. & Nijhoff F. W. *Quantum variational principle and quantum multiform structure: the case of quadratic Lagrangians.* arXiv:1702.08709, 2017.

[44] Krichever I. M. & Novikov S. P. *Holomorphic bundles over algebraic curves and non-linear equations.* Russian Mathematical Surveys, 35 : 53, 1980.

[45] Leok M. & Zhang J. *Discrete Hamiltonian variational integrators.* IMA Journal of Numerical Analysis, 31 : 1497–1532, 2011.

[46] Lévy-Leblond J.-M. *Conservation laws for gauge-variant Lagrangians in classical mechanics.* American Journal of Physics, 39 : 502–506, 1971.

[47] Lobb S. & Nijhoff F. *Lagrangian multiforms and multidimensional consistency.* Journal of Physics A: Mathematical and Theoretical, 42 : 454013, 2009.

[48] Lobb S. B. & Nijhoff F. W. *Lagrangian multiform structure for the lattice Gel'fand–Dikii hierarchy.* Journal of Physics A: Mathematical and Theoretical, 43 : 072003, 2010.

[49] Lobb S. B., Nijhoff F. W. & Quispel G. R. W. *Lagrangian multiform structure for the lattice KP system.* Journal of Physics A: Mathematical and Theoretical, 42 : 472002, 2009.

[50] Magri F. *A simple model of the integrable Hamiltonian equation.* Journal of Mathematical Physics, 19 : 1156–1162, 1978.

[51] Mansfield E. L. & Hydon P. E. *Difference forms.* Foundations of Computational Mathematics, 8 : 427–467, 2008.

[52] Marsden J. E. & West M. *Discrete mechanics and variational integrators.* Acta Numerica, 10 : 357–514, 2001.

[53] Marsden J. E., Patrick G. W. & Shkoller S. *Multisymplectic geometry, variational integrators, and nonlinear pdes.* Communications in Mathematical Physics, 199 : 351–395, 1998.

[54] Miura R. M. *Korteweg-de Vries equation and generalizations. I. A remarkable explicit nonlinear transformation.* Journal of Mathematical Physics, 9 : 1202–1204, 1968.

[55] Miwa T. *On Hirota's difference equations.* Proceedings of the Japan Academy, Series A, Mathematical Sciences, 58 : 9–12, 1982.

[56] Moan P. C. *On modified equations for discretizations of ODEs.* Journal of Physics A: Mathematical and General, 39 : 5545, 2006.

[57] Morehead J. *Visualizing the extra symmetry of the Kepler problem.* American journal of physics, 73 : 234–239, 2005.

[58] Newell A. C. *Solitons in Mathematics and Physics.* SIAM, 1985.

[59] Newton P. K. *The N-vortex Problem: Analytical Techniques.* Springer, 2013.

[60] Nijhoff F. & Capel H. *The discrete Korteweg-de Vries equation.* Acta Applicandae Mathematica, 39 : 133–158, 1995.

[61] Nijhoff F., Hone A. & Joshi N. *On a schwarzian PDE associated with the KdV hierarchy.* Physics Letters A, 267 : 147–156, 2000.

[62] Nijhoff F., Atkinson J. & Hietarinta J. *Soliton solutions for ABS lattice equations: I. Cauchy matrix approach.* Journal of Physics A: Mathematical and Theoretical, 42 : 404005, 2009.

[63] Nijhoff F. W. *Lax pair for the Adler (lattice Krichever–Novikov) system.* Physics Letters A, 297 : 49–58, 2002.

[64] Nijhoff F. W., Quispel G. R. W. & Capel H. W. *Direct linearization of nonlinear difference-difference equations.* Physics Letters A, 97 : 125–128, 1983.

[65] Nijhoff F. W., Papageorgiou V. G., Capel H. W. & Quispel G. R. W. *The lattice Gel'fand-Dikii hierarchy.* Inverse Problems, 8 : 597–621, 1992.

[66] Noether E. *Invariante Variationsprobleme.* Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse, 1918 : 235–257, 1918.

[67] Oliver M. & Vasylkevych S. *A new construction of modified equations for variational integrators.* `http://math.jacobs-university.de/oliver/papers/newmodified.pdf`, 2012.

[68] Olver P. J. *Applications of Lie Groups to Differential Equations.* Volume 107 of *Graduate Texts in Mathematics.* Springer, 2000.

[69] Ovsienko V. & Tabachnikov S. *What is the Schwarzian derivative?* Notices of the AMS, 56 : 34–36, 2009.

[70] Palais R. *The symmetries of solitons*. Bulletin of the AMS, 34 : 339–403, 1997.

[71] Petrera M. & Suris Yu. B. *Variational symmetries and pluri-Lagrangian systems in classical mechanics*. Journal of Nonlinear Mathematical Physics, 24 (Sup. 1) : 121–145, 2017.

[72] Poincaré H. *Les Méthodes Nouvelles de la Mécanique Céleste*. Gauthier-Villars, 1892, 1893, 1899.

[73] Quispel G. R. W., Nijhoff F. W., Capel H. W. & Van der Linden J. *Linear integral equations and nonlinear difference-difference equations*. Physica A: Statistical Mechanics and its Applications, 125 : 344–380, 1984.

[74] Reich S. *Backward error analysis for numerical integrators*. SIAM Journal on Numerical Analysis, 36 : 1549–1570, 1999.

[75] Rogers H. H. *Symmetry transformations of the classical Kepler problem*. Journal of Mathematical Physics, 14 : 1125–1129, 1973.

[76] Rowley C. W. & Marsden J. E. *Variational integrators for degenerate Lagrangians, with application to point vortices*. In *Proceedings of the 41st IEEE Conference on Decision and Control*, pages 1521–1527. IEEE, 2002.

[77] Rudin W. *Function Theory in the Unit Ball of* $\mathbb{C}^n$. Volume 241 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1980.

[78] Sulem C. & Sulem P.-L. *The Nonlinear Schrödinger Equation: Self-Focusing and Wave Collapse*. Springer, 2007.

[79] Suris Yu. B. *The Problem of Integrable Discretization: Hamiltonian Approach*. Birkhäuser, 2003.

[80] Suris Yu. B. *Variational formulation of commuting Hamiltonian flows: Multi-time Lagrangian 1-forms*. Journal of Geometric Mechanics, 5 : 365–379, 2013.

[81] Suris Yu. B. *Variational symmetries and pluri-Lagrangian systems*. In *Dynamical Systems, Number Theory and Applications: A Festschrift in Honor of Armin Leutbecher's 80th Birthday*, pages 255–266. World Scientific, 2016.

[82] Suris Yu. B. & Vermeeren M. *On the Lagrangian structure of integrable hierarchies*. In Bobenko A. I., editor, *Advances in Discrete Differential Geometry*, pages 347–378. Springer, 2016.

[83] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 7.5.1)*. `http://www.sagemath.org`, 2017.

[84] Toda M. *Studies of a non-linear lattice*. Physics Reports, 18 : 1–123, 1975.

[85] Toda M. *Theory of Nonlinear Lattices*. Volume 20 of *Solid-State Sciences*. Springer, 1989.

[86] Tongas A. & Nijhoff F. *The Boussinesq integrable system: compatible lattice and continuum structures*. Glasgow Mathematical Journal, 47 : 205–219, 2005.

[87] Tyranowski T. M. & Desbrun M. *Variational partitioned Runge-Kutta methods for Lagrangians linear in velocities*. arXiv:1401.7904, 2014.

[88] Vermeeren M. *Numerical precession in variational integrators for the Kepler problem*. arXiv:1602.01049, 2016.
To appear in the volume *Discrete Mechanics, Geometric Integration, and Lie-Butcher Series* of Springer Proceedings in Mathematics & Statistics.

[89] Vermeeren M. *Continuum limits of pluri-Lagrangian systems*. arXiv:1706.06830, 2017.
Submitted to the Journal of Integrable Systems.

[90] Vermeeren M. *Modified equations for variational integrators*. Numerische Mathematik, 137 : 1001–1037, 2017.

[91] Vermeeren M. *Modified equations for variational integrators applied to Lagrangians linear in velocities*. arXiv:1709.09567, 2017.
Provisionally accepted in the Journal of Geometric Mechanics.

[92] Vermeeren M. *Modified equations and the Basel problem*. The Mathematical Intelligencer, 40(2) : 33–37, 2018.

[93] Walker A. J. *Similarity reductions and integrable lattice equations*. PhD thesis, University of Leeds. `http://etheses.whiterose.ac.uk/7190/`, 2001.

[94] Wiersma G. L. & Capel H. W. *Lattice equations, hierarchies and Hamiltonian structures*. Physica A: Statistical Mechanics and its Applications, 142 : 199–244, 1987.

[95] Will C. M. *Theory and Experiment in Gravitational Physics*. Cambridge University Press, 1981.

[96] Xenitidis P., Nijhoff F. & Lobb S. *On the Lagrangian formulation of multidimensionally consistent systems*. Proceedings of the Royal Society A, 467 : 3295–3317, 2011.

[97] Yoo-Kong S., Lobb S. & Nijhoff F. *Discrete-time Calogero-Moser system and Lagrangian 1-form structure.* Journal of Physics A: Mathematical and Theoretical, 44 : 365203, 2011.

[98] Zagier D. *The dilogarithm function.* In Cartier P., Moussa P., Julia B. & Vanhove P., editors, *Frontiers in number theory, physics, and geometry II*, pages 3–65. Springer, 2007.

[99] Zakharov V. E. & Faddeev L. D. *Korteweg-de Vries equation: A completely integrable Hamiltonian system.* Functional Analysis and Its Applications, 5 : 280–287, 1971.

[100] Zakharov V. E. & Mikhailov A. V. *Variational principle for equations integrable by the inverse problem method.* Functional Analysis and Its Applications, 14 : 43–44, 1980.