

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

On Higher-Derivative Operators in Supersymmetric Theories

Relatore:
Prof. Francisco Gil Pedro

Presentata da:
Lorenzo Paoloni

Anno Accademico 2020/2021

Abstract

The occurrence of higher-derivative terms both in classical and in quantum theories can be problematic. In case of impossibility of eliminating them by performing an integration by parts they give rise, in fact, to multiple problems, like extra unphysical degrees of freedom, negative unbounded energies and pathological negative norm states called ghosts. The possibility of curing them comes from a procedure developed by Jaén, Llosa and Molina in a classical Hamiltonian framework. It is applied when these terms are introduced as perturbative corrections to a healthy leading order theory multiplied by a perturbative expansion parameter. We first review this method from the classical point of view, then generalize it for Field Theories and finally apply it to SUSY. It's indeed in the latter context that one has always had difficulties in treating certain operators, those whose explicit field expansions contained unusual terms, like second (or higher) derivatives of a dynamical field or derivatives of an auxiliary one.

As we'll show, the approach developed is able to cure these unmanageable and unjustifiable extra degrees of freedom, paving the way for the construction of a new range of supersymmetric theories.

Contents

1	Introduction	2
2	The JLM Procedure	6
2.1	Classical version	6
2.1.1	Lagrangian systems of any order	7
2.1.2	Approximated Lagrangian systems	8
2.1.3	Application: the Double Harmonic Oscillator in a Gravitational Field	12
2.2	Classical Field Theory version	27
2.2.1	Application: Light-Heavy Scalar Fields Interaction	30
3	Higher-Derivatives in SUSY	38
3.1	Notation and Conventions	38
3.2	Higher-Derivatives Treatment in SUSY	40
3.3	Application: the DBI Action	57
4	Conclusions	63
A	Treatment of ϵ-polynomials	65
B	Details on Kähler potential and Superpotential dependence	67
C	Details on DBI Lagrangian expansion	70

Chapter 1

Introduction

The descriptions of physical systems often involves Lagrangians or equations of motion of a given order in time derivatives. In particular for each particle we usually have a Lagrangian containing just the coordinate and its first time-derivative¹ and/or second order differential equations of motion. Models with Lagrangian containing derivatives higher than one or, equivalently, with third or higher order equations of motion describe the so called **higher-derivative theories**.

These theories are often avoided since they all share very peculiar features. In particular, they do not describe – at least as they stand – any known physical system, either in a classical or in a quantum setup. Nevertheless, it's also true that higher-derivative terms in the Lagrangian or in equations of motion can arise spontaneously or be necessary because of physical considerations.

Such theories can be classified in three families ([1]):

- 1) Theories where higher derivatives are added to describe particular physical features, because it's the only way to describe them or because it's the cleanest or the smartest path. Doing this always generates different problems which require the imposition of constraints to be fixed. An example can be the Abraham-Lorentz model for radiating charged particles, where their EOM's takes the form²

¹This is true in classical physics, in Field Theory the Lagrangian density will be composed by the field and its first spacetime derivative.

²Such an EOM is clearly different in form from the usual Newton law

$$\ddot{x} = m^{-1}F(x, \dot{x}) , \tag{1.1}$$

where the second derivative is a function of only the zero-th and the first one. This change, as small as it may seem, would lead to big problems if left unconstrained.

$$\ddot{x} = m^{-1}F(x, \dot{x}) + \varepsilon \ddot{\ddot{x}} \quad (1.2)$$

(with $\varepsilon = 2e^2/(3mc^3)$). While this new modified equation describes correctly the energy loss of an accelerating charge (Larmor formula), it also gives rise to two common problems, extra degrees of freedom and runaway solutions. The former can be seen analysing (1.2), the motion is in fact specified when initial position, velocity *and* acceleration are given. The latter are found for example when one sets an initial null force ($F = 0$) on a non-moving body at the origin ($x_0 = 0, v_0 = 0$): in this situation any non-zero initial acceleration (no matter how small) would generate an exponentially growing solution. It's clear then how, in this class of theories, not every solution is physical.

- 2) Theories where higher derivatives arise spontaneously in the process of constructing a perturbative approximation to a more fundamental theory. From this assumption follows the need for the new solution to be a perturbative expansion of the old one as well, but since this is not assured, the use of constraints becomes necessary even for this class. An example can be found looking at non-local theories, where non-local expressions become series of local higher-derivative terms through a Taylor expansion

$$x(t + \varepsilon T) = \sum_{n=0}^{\infty} \frac{(\varepsilon T)^n}{n!} \frac{d^n}{dt^n} x(t) . \quad (1.3)$$

Here every higher derivative term should in principle correct the theory, approximating better and better the real solution at every order. Notice that this is not what happens: if left unconstrained higher-derivative corrections, although multiplied by a small coefficient (like the ε above could be), give rise to undesired and unphysical behaviours, as we'll see later. This is the class which we will mainly focus on.

- 3) Theories that are "*genuinely higher order*", where every solution of the higher order equations of motion have to be taken into account and considered physical. These theories present strange behaviours absent in second order models (like negative kinetic energies) but are mathematically consistent: there are no theoretical reasons they should not exist either. Anyway, no physical examples are known for this class.

As we anticipated, we will focus on the second type of models described here. In particular, we will add higher derivatives as a correction – through small coefficients – to a standard lower-derivative theory. The problem is

that this does not only perturb the original theory, like a lower-derivative correction would have done, but changes it drastically. For this reason every higher-derivative correction has to be treated properly in order not to cause a dramatic change in the fundamental behaviour of the system. In other words, they have to be treated from the very beginning as corrections to the leading order theory.

This brief introduction overlooks two important aspects: the fact that all the unconstrained theories with higher-derivative terms share same common features, and that the treatment that introduces the necessary constraints will be based on the assumption that higher derivatives are just perturbative corrections to a lower order theory.

Let's analyze the first point, the peculiar features shared by higher-derivative models. These are mainly two: the presence of extra degrees of freedom and the lack of a lower energy bound.

The arising of extra degrees of freedom might also be physically more accurate – although in general just signals that the previous lower-order theory was defective – but becomes very disturbing when higher and higher derivative terms arise as a progression of corrections (like in non-local theories): here, increasing the order of the approximation leads to an increase of the initial data necessary to completely specify the motion.

Notice that switching to quantum mechanics opens to something equally disturbing. Quantising systems with extra d.o.f.'s leads indeed to a deep change in the usual well-known uncertainty principle: now position and velocity commute, becoming then freely specifiable and so – from an experimental point of view – simultaneously measurable.

As regards the lack of a lower energy bound, it appears in every higher-derivative theory (whatever small the coefficients multiplying higher-derivative terms are) and it assumes different forms depending on the framework we are working with. In classical physics it is encoded in an indefinite unbounded energy, quantum mechanically it appears as an infinite tower spectrum of energy eigenvalues from $-\infty$ to $+\infty$, and in Quantum Field Theory it is related to the problem of *ghosts*, field excitations (that is particles) of negative energy. The problem about ghosts concerns their production: since their creation is energetically favourable (it produces an excess of energy), they would be bound to be generated in infinite number.

These features are very particular and, in general, problematic. They prompt us to find a method able to correctly treat higher-derivative terms in order to make such problems disappear.

The procedure in question is known as the method of *perturbative constraints*, and is based on the imposition of several constraints in order to keep only solutions that are perturbatively expandible in the small perturbative

parameter. Taking as example the approximation of a non-local theory into a truncated series of local higher-derivative terms, such constraints will have the role of allowing us to consider the expansion as a proper perturbative expansion. The presence of the small coefficients alone (appropriately elevated to the corresponding power) is indeed not sufficient to guarantee that terms belonging to higher approximation orders contribute proportionally less. Without constraints each higher-derivative term would play the same role in the physics of the model.

In this way, the perturbative solutions – selected by the constraints – will be the physical ones, whereas all the other – discarded by the constraints – will be seen as spurious pseudosolutions.

The most powerful feature of this method is its range of applicability. It in fact does not require the knowledge of the full theory we are approximating, but instead just needs the terms of the truncated series to be multiplied by a small coefficient, appropriately elevated to some power. Higher-derivative terms can thus be added by hand and assumed to be part of a truncated expansion.

It's important to stress that, as clearly stated in [1], this method is not just an ad hoc procedure, but rather completely natural and necessary when a higher-derivative theory is, or is assumed to be, the approximation of a more fundamental one.

The aim of this thesis is to systematically apply the method of perturbative constraints to supersymmetric theories, thereby clarifying and extending the discussion of higher-derivative theories including chiral superfields. In practice what we'll do is to add new higher-derivative operators to a 0-th order Kähler potential and to assume that they belong to some sort of perturbative expansion. Since such operators have to be accompanied by a coefficient which restores the correct mass dimension, this quantity can become the necessary *small coefficient* for the applicability of the procedure (for example taking the inverse of the cutoff scale). One of the missing points in previous treatment was indeed the avoidance of mass dimension problems: the introduction of higher-derivative operators must always be accompanied by dimensionful constants. Once considered, this immediately suggests a difference between the 0-th order and the other higher-derivative operators added by hand, and right here the method finds its foundations.

This work is structured as follows: first, in Chapter 2, we'll review (and generalize) the classical method to find the perturbative constraints for an higher-derivative theory and then, in Chapter 3, we'll focus on the application of this procedure to Global Supersymmetric Theories. Conclusions follow in Chapter 4.

Chapter 2

The JLM Procedure

We are going to review and (slightly) generalize the method developed by Jaén, Llosa and Molina, exposed in [2], that allows to analyze higher derivative Lagrangians without introducing new degrees of freedom or instabilities. They implement this procedure in a classical Hamiltonian framework, so first we review the original version and then we move on to a field theory update. We would point out that the method we are going to present, at a certain point differs from the [2]'s one: since we are mostly interested in the Lagrangian formulation (rather than the Hamiltonian) for Field Theory applications, we will only use Lagrangians in those contexts.

2.1 Classical version

We start by reviewing how to deal with Lagrangians that contains coordinates and their derivatives up to their n -th order¹, like

$$L = L(t; q, q^{(1)}, \dots, q^{(n)}) . \quad (2.1)$$

Here $q^{(s)}$, with $s \geq 1$, represent the s -th time derivative of q . Sometimes $q^{(0)}$ can be used to indicate q and \dot{q} , \ddot{q} to indicate $q^{(1)}$, $q^{(2)}$.

We will keep this notation for L (and for every quantity that depends on coordinates and their derivatives) even if the system is composed by multiple particles: in that case $q^{(s)} \equiv \{q_\alpha^{(s)}\}$ with $\alpha = 1, \dots, N$ and $s \geq 0$.

In this section to label different particles – or, more generally, components of the system – greek letters will always be used.

¹See [1, 2, 3]

2.1.1 Lagrangian systems of any order

As for standard Lagrangians the variational principle

$$\delta \int_{t_1, P}^{t_2, Q} L(t; q, q^{(1)}, \dots, q^{(n)}) dt = 0 \quad (2.2)$$

is assumed to hold. The fundamental difference is that here the points P and Q , belonging to the configuration space, are identified not only by their coordinates q , but by the n -tuple $\{q, q^{(1)}, \dots, q^{(n-1)}\}$.

The Euler-Lagrange equations which solve the variational problem are

$$\sum_{k=0}^n (-1)^k D^k \frac{\partial L}{\partial q_\alpha^{(k)}} = 0 \quad (2.3)$$

with

$$D \equiv \frac{d}{dt}. \quad (2.4)$$

For the Hamiltonian formulation (that we just briefly present) also the conjugate momenta are necessary.

There is, as always, a momentum for every configuration space variable $q_\alpha^{(j)}$ ($j = 0, 1, \dots, n-1$)

$$\Pi_{q_\alpha^{(j)}} \equiv \sum_{k=0}^{n-j-1} (-1)^k D^k \frac{\partial L}{\partial q_\alpha^{(k+j+1)}}, \quad (2.5)$$

and all together are used to construct the Hamiltonian in the usual way²

$$H = \sum_{\lambda} \sum_{k=0}^{n-1} \Pi_{q_\lambda^{(k)}} q_\lambda^{(k+1)} - L. \quad (2.6)$$

Here $L = L(t; q, q^{(1)}, \dots, q^{(n-1)}, \Pi_{q^{(n-1)}})$, indeed

$$q_\alpha^{(n)} = q_\alpha^{(n)}(t; q, q^{(1)}, \dots, q^{(n-1)}, \Pi_{q^{(n-1)}}) \quad (2.7)$$

whereas all the remaining $q^{(j)}$'s are independent generalized coordinates and so are not inverted.

The first order Hamilton-Jacobi equation of motion are

$$\frac{\partial H}{\partial \Pi_{q_\alpha^{(j)}}} = \dot{q}_\alpha^{(j)} \quad \frac{\partial H}{\partial q_\alpha^{(j)}} = -\dot{\Pi}_{q_\alpha^{(j)}} \quad j = 0, 1, \dots, n-1 \quad (2.8)$$

²If our Lagrangian is time-independent, the Hamiltonian is conserved and equal to the energy of the system.

where the dots have been used to denote a total time derivative ($\dot{q}_\alpha^{(j)} \equiv Dq_\alpha^{(j)} = q_\alpha^{(j+1)}$, $\dot{\Pi}_{q_\alpha^{(j)}} \equiv D\Pi_{q_\alpha^{(j)}}$) with the aim of recovering a more familiar form for these equations.

Let us stress the similarities with the usual two-derivative case: in going from a Lagrangian to a Hamiltonian description we move from a problem where the evolution of the system was ruled by a $2n$ -th order differential equation for n variables (configuration space) to one where it is governed by $2n$ first order differential equations for $2n$ variables (phase space).

2.1.2 Approximated Lagrangian systems

We are going to deal with Lagrangians that assume the form

$$L = \frac{1}{2} \sum_{\alpha=1}^N m_\alpha (q_\alpha^{(1)})^2 + \sum_{k=0}^n \epsilon^k \mathfrak{Y}_k(q, q^{(1)}, \dots, q^{(k)}) + O(\epsilon^{n+1}) \quad (2.9)$$

where m_α are constants³ and ϵ is the real parameter used to truncate at the desired perturbative order.

This Lagrangian, in fact, can be thought as an infinite-order one, approximated to ϵ^{n+1} . Specifically, an infinite-order Lagrangian will have the general form⁴

$$L = \sum_{k=0}^{\infty} \Gamma_k(q, q^{(1)}, \dots, q^{(k)}) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \Gamma_k(q, q^{(1)}, \dots, q^{(k)}) \equiv \lim_{n \rightarrow \infty} L_n, \quad (2.10)$$

but now we want to restrict our attention to a particular form of L_n : we only consider polynomials of degree n in the variable ϵ .

In this sense *approximated to order ϵ^{n+1}* means *equal modulo ϵ^{n+1}* , that is, equal if every term multiplied by ϵ^s , with $s \geq n + 1$, is neglected.

Then, every equation – or quantity – deriving from such a system will have to contain terms approximated to order ϵ^{n+1} , and so an $O(\epsilon^{n+1})$ will always appear on its right-hand side.

Formally, our quantities will belong to the quotient ring $\mathbb{R}[\epsilon]/(\epsilon^{n+1})$: the polynomials in ϵ (with real coefficients) modulo ϵ^{n+1} .

³These would be the masses in the usual $n = 0$ case.

⁴This sort of series are obtained for example when one has to expand a non-local Lagrangian, as the one describing the Wheeler-Feynman Electrodynamics (see [1, 2]), or when one has to integrate out a degree of freedom, as we'll see later.

This fact automatically excludes the ill-behaved solutions of equations of motion in the limit $\epsilon \rightarrow 0$ but since an algebraic definition has been stated, the future results will hold regardless of the size of ϵ .

Euler-Lagrange equations (2.3), applied to the Lagrangian (2.9), read as

$$-m_\alpha q_\alpha^{(2)} + \sum_{k=0}^n \epsilon^k \mathfrak{W}_{\alpha,k}(q, \dots, q^{(2k)}) = O(\epsilon^{n+1}) \quad \forall \alpha \quad (2.11)$$

with

$$\mathfrak{W}_{\alpha,k} \equiv \sum_{r=0}^k (-1)^r D^r \frac{\partial \mathfrak{W}_k}{\partial q_\alpha^{(r)}} . \quad (2.12)$$

From these equations we can easily derive the so called **primary constraints** of the system:

$$\epsilon \left[-m_\alpha q_\alpha^{(2)} + \sum_{k=0}^{n-1} \epsilon^k \mathfrak{W}_{\alpha,k}(q, \dots, q^{(2k)}) \right] = O(\epsilon^{n+1}) \quad \forall \alpha \quad (2.13)$$

where a primary constraint is defined as a relation between coordinates and their derivatives⁵ that holds without using the equations of motion.

For such a system will also exist **secondary constraints**, that hold when equations of motion are satisfied but need not hold if they are not satisfied, although the procedure to find them is a little bit more involved.

Before outlining this procedure and deriving the secondary constraints of this generic system we would point out that the definitions of primary and secondary constraints could be given in a clearer and more useful way: the primary constraints will be those "restrictions" to the system that emerges spontaneously, whereas the secondary will be the ones obtained by an iterative method that uses also the equations of motion (and then requires that these hold!).

So, starting from primary constraints, we multiply them by ϵ^{n-1} , getting

$$\epsilon^n q_\alpha^{(2)} = \frac{1}{m_\alpha} \mathfrak{W}_{\alpha,0}(q) \epsilon^n + O(\epsilon^{n+1}) , \quad (2.14)$$

to then begin differentiating with respect to time:

⁵In origin, Dirac defined a primary constraint as a relation between coordinates and *momenta* because he referred to the Hamiltonian formalism (as Jaén, Llosa and Molina do in [2]). Here we treat the problem under a Lagrangian point of view and then also the definition of primary constraint has to change.

- the first derivative will lead to

$$\epsilon^n q_\alpha^{(3)} = \frac{1}{m_\alpha} D \mathfrak{W}_{\alpha,0}(q) \epsilon^n + O(\epsilon^{n+1}) = \frac{1}{m_\alpha} \dot{\mathfrak{W}}_{\alpha,0}(q, q^{(1)}) \epsilon^n + O(\epsilon^{n+1}) ; \quad (2.15)$$

- the second derivative will lead to

$$\epsilon^n q_\alpha^{(4)} = \frac{1}{m_\alpha} D^2 \mathfrak{W}_{\alpha,0}(q) \epsilon^n + O(\epsilon^{n+1}) = \frac{1}{m_\alpha} \ddot{\mathfrak{W}}_{\alpha,0}(q, q^{(1)}, q^{(2)}) \epsilon^n + O(\epsilon^{n+1}) \quad (2.16)$$

and then will permit us to use (2.14) to substitute the $\epsilon^n q_\alpha^{(2)}$ dependence in $\ddot{\mathfrak{W}}_{\alpha,0} \epsilon^n$. The result will be

$$\begin{aligned} \epsilon^n q_\alpha^{(4)} &= \frac{1}{m_\alpha} \ddot{\mathfrak{W}}_{\alpha,0}(q, q^{(1)}, \frac{1}{m_\alpha} \mathfrak{W}_{\alpha,0}(q)) \epsilon^n + O(\epsilon^{n+1}) \\ &\equiv \frac{1}{m_\alpha} \mathfrak{Z}_{\alpha,4,0}(q, q^{(1)}) \epsilon^n + O(\epsilon^{n+1}) ; \end{aligned} \quad (2.17)$$

- the third derivative will lead to

$$\epsilon^n q_\alpha^{(5)} = \frac{1}{m_\alpha} D^3 \mathfrak{W}_{\alpha,0}(q) \epsilon^n + O(\epsilon^{n+1}) = \frac{1}{m_\alpha} \dot{\mathfrak{Z}}_{\alpha,4,0}(q, q^{(1)}, q^{(2)}) \epsilon^n + O(\epsilon^{n+1}) \quad (2.18)$$

and so the use of (2.14) will again be necessary to eliminate all the $\epsilon^n q_\alpha^{(2)}$ terms. The result will be

$$\epsilon^n q_\alpha^{(5)} \equiv \frac{1}{m_\alpha} \mathfrak{Z}_{\alpha,5,0}(q, q^{(1)}) \epsilon^n + O(\epsilon^{n+1}) \quad (2.19)$$

and so on.

Thus, in general, we obtain the relations

$$\epsilon^n q_\alpha^{(r+2)} \equiv \frac{1}{m_\alpha} \mathfrak{Z}_{\alpha,r+2,0}(q, q^{(1)}) \epsilon^n + O(\epsilon^{n+1}) , \quad (2.20)$$

for $r \geq 1$ ⁶, that can be plugged back – with $r = 2n - 4$ and $r = 2n - 5$ – into primary constraints (2.13) to eliminate the $\epsilon^n q^{(2n-2)}$ and $\epsilon^n q^{(2n-3)}$ terms that appears only in the last summation's addend, $\mathfrak{W}_{\alpha,n-1}$.

Now the highest derivatives are q^{2n-4} and q^{2n-5} , in $\mathfrak{W}_{\alpha,n-1}$ and $\mathfrak{W}_{\alpha,n-2}$, and to eliminate them we restart from primary constraints, this time multiplying by ϵ^{n-2} :

$$\epsilon^{n-1} q_\alpha^{(2)} = \frac{1}{m_\alpha} [\mathfrak{W}_{\alpha,0}(q) + \mathfrak{W}_{\alpha,1}(q, q^{(1)}, q^{(2)})] \epsilon^{n-1} + O(\epsilon^{n+1}) . \quad (2.21)$$

⁶Here $\dot{\mathfrak{W}}_{\alpha,0}(q, q^{(1)})$ has been implicitly redefined as $\mathfrak{Z}_{\alpha,3,0}(q, q^{(1)})$.

Again, (2.14) can be used to throw away the $q^{(2)}$ dependence on the r.h.s. of this equation, finding

$$\epsilon^{n-1} q_\alpha^{(2)} = \frac{1}{m_\alpha} \left[\mathfrak{W}_{\alpha,0}(q) + \tilde{\mathfrak{W}}_{\alpha,1}(q, q^{(1)}) \epsilon \right] \epsilon^{n-1} + O(\epsilon^{n+1}) . \quad (2.22)$$

From here we can start differentiating as before, substituting at every step $\epsilon^{n-1} q_\alpha^{(2)}$ with (2.22). The general result will be

$$\epsilon^{n-1} q_\alpha^{(r+2)} \equiv \frac{1}{m_\alpha} \left[\mathfrak{Z}_{\alpha,r+2,0}(q) + \mathfrak{Z}_{\alpha,r+2,1}(q, q^{(1)}) \epsilon \right] \epsilon^{n-1} + O(\epsilon^{n+1}) \quad (2.23)$$

with $r \geq 1$. Notice that this relation implies the (2.20) one, obtained by just multiplying by ϵ both sides.

We can continue iteratively restarting from the primary constraint and lowering progressively the power of ϵ , always eliminating the undesired q derivatives: at the end this algorithm leads to

$$\epsilon \left[q_\alpha^{(r)} - \frac{1}{m_\alpha} \sum_{k=0}^{n-1} \epsilon^k \mathfrak{Z}_{\alpha,r,k}(q, q^{(1)}) \right] = O(\epsilon^{n+1}) \quad (2.24)$$

for $r = 2, \dots, 2n - 1$. Taking the derivative of this equation when $r = 2n - 1$ (and of course substituting $q^{(2)}$ terms as it's been done so far), will give us

$$\epsilon \left[q_\alpha^{(2n)} - \frac{1}{m_\alpha} \sum_{k=0}^{n-1} \epsilon^k \mathfrak{Z}_{\alpha,2n,k}(q, q^{(1)}) \right] = O(\epsilon^{n+1}) , \quad (2.25)$$

thanks to which – with the help of all the other relations found along the way of this painstaking procedure – we can drastically simplify the equations of motion to obtain the new constraints

$$q_\alpha^{(2)} = \frac{1}{m_\alpha} \sum_{k=0}^n \epsilon^k \mathfrak{Z}_{\alpha,2,k}(q, q^{(1)}) + O(\epsilon^{n+1}) \quad (2.26)$$

where $\mathfrak{Z}_{\alpha,2,k}$ is just $\mathfrak{W}_{\alpha,k}$ with all the q 's derivatives higher than one eliminated.

Finally, differentiating and substituing (using this last equation found for $q_\alpha^{(2)}$) one the last time, we can write explicitly our full minimal set of secondary constraints:

$$q_\alpha^{(r)} = \frac{1}{m_\alpha} \sum_{k=0}^n \epsilon^k \mathfrak{Z}_{\alpha,r,k}(q, q^{(1)}) + O(\epsilon^{n+1}) \quad r = 2, \dots, 2n - 1 . \quad (2.27)$$

These constraints not only contain the primary ones, but they are also more restrictive than them and stable under time differentiation.

Even if this algorithm or its final formula could seem obscure or pretty difficult to apply, as we'll see the analysis of the effective constraints for a system can be very clearer, faster and – after all – much easier.

2.1.3 Application: the Double Harmonic Oscillator in a Gravitational Field

Let us analyze closely an example in classical physics, in order to compare a normal lower-derivative theory with a higher-derivative one.

Consider a system composed by two masses m and M such that $m \ll M$, and by two springs with the same harmonic constant k and the same rest length ℓ .

The system will be assembled as in Fig. 2.1: the coordinate x will label the mass m whereas the coordinate y will label the mass M .

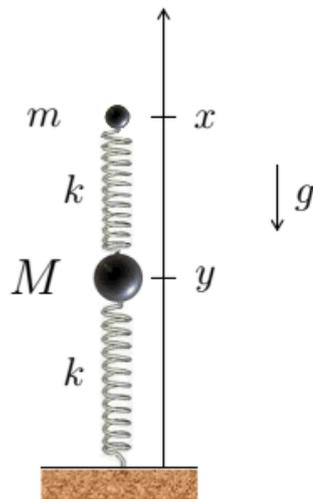


Figure 2.1: The system considered.

The Lagrangian of the system is

$$\begin{aligned}
L &= \frac{1}{2}m\dot{x}^2 + \frac{1}{2}M\dot{y}^2 - mgx - Mgy - \frac{1}{2}k(y - \ell)^2 - \frac{1}{2}k(x - y - \ell)^2 \\
&= \frac{1}{2}m\dot{x}^2 + \frac{1}{2}M\dot{y}^2 + (k\ell - mg)x - \frac{1}{2}kx^2 - Mgy - ky^2 + kxy + const .
\end{aligned} \tag{2.28}$$

Now we rescale this Lagrangian dividing by M and we define

$$\begin{aligned}
\Omega^2 &\equiv \frac{k}{M} \\
\epsilon^2 &\equiv \frac{m}{M} \\
\eta &\equiv \Omega^2\ell - \epsilon^2g
\end{aligned}$$

to get

$$L' = \frac{1}{2}\epsilon^2\dot{x}^2 + \frac{1}{2}\dot{y}^2 + \eta x - \frac{1}{2}\Omega^2x^2 - gy - \Omega^2y^2 + \Omega^2xy \tag{2.29}$$

where constants have been omitted.

Notice that, whereas $[L] = E$, $[L'] = L^2T^{-2} = E M^{-1}$. From now on we will remove the prime in every function, implying that quantities in exam now becomes *per unit mass* (Lagrangian per unit mass, Energy per unit mass, and so on).

Firstly we solve exactly the system as it stands, later we will make some assumptions that will drastically change the resulting physics.

So, we start from the equations of motion. For x we have

$$\eta - \Omega^2x + \Omega^2y - \epsilon^2\ddot{x} = 0 , \tag{2.30}$$

whereas for y

$$-g - 2\Omega^2y + \Omega^2x - \ddot{y} = 0 . \tag{2.31}$$

From this last equation we can isolate Ω^2x , differentiate twice, and plug back into x 's EOM. Notice that we are always free to do this: it is always allowed to substitute equations of motion inside other equations of motion (problems arises when we want to plug them into the Lagrangian, as we'll see later).

The result is the fourth order differential equation describing the motion of the heaviest mass:

$$\epsilon^2\Omega^{-2}y^{(4)} + (1 + 2\epsilon^2)\ddot{y} + \Omega^2y + \bar{g} = 0 , \tag{2.32}$$

where

$$\bar{g} = g - \eta .$$

The general solution of this differential equation is

$$y(t) = A_+ e^{i\lambda_+ t} + B_+ e^{-i\lambda_+ t} + A_- e^{i\lambda_- t} + B_- e^{-i\lambda_- t} - \bar{g} \Omega^{-2} \quad (2.33)$$

with

$$\lambda_{\pm}^2 = \frac{1 + 2\epsilon^2 \mp \sqrt{1 + 4\epsilon^4}}{2\epsilon^2} \Omega^2 . \quad (2.34)$$

For later convenience we also write down these two frequencies truncated at their first magnitude order

$$\lambda_+ = \Omega + O(\epsilon^2) \quad (2.35)$$

$$\lambda_- = \frac{1}{\epsilon} \Omega + O(1) . \quad (2.36)$$

Following the same steps we can also find the differential equation describing the motion of the lightest mass m

$$\epsilon^2 \Omega^{-2} x^{(4)} + (1 + 2\epsilon^2) \ddot{x} + \Omega^2 x + \bar{g} = 0 , \quad (2.37)$$

which has the same form of the one for y , with

$$\bar{\bar{g}} = \bar{g} - \eta = g - 2\eta$$

replacing \bar{g} .

The solution will be, thus, completely analogous to the y 's one

$$x(t) = \bar{A}_+ e^{i\lambda_+ t} + \bar{B}_+ e^{-i\lambda_+ t} + \bar{A}_- e^{i\lambda_- t} + \bar{B}_- e^{-i\lambda_- t} - \bar{\bar{g}} \Omega^{-2} \quad (2.38)$$

and using (2.30) (or (2.31)) one is able to relate the four free coefficients A_+ , B_+ , A_- , B_- in $y(t)$ with their barred versions in $x(t)$:

$$A_+ = (1 - \epsilon^2 \lambda_+^2 \Omega^{-2}) \bar{A}_+ \equiv Z_+^{-1} \bar{A}_+ \quad (2.39)$$

$$B_+ = (1 - \epsilon^2 \lambda_+^2 \Omega^{-2}) \bar{B}_+ = Z_+^{-1} \bar{B}_+ \quad (2.40)$$

$$A_- = (1 - \epsilon^2 \lambda_-^2 \Omega^{-2}) \bar{A}_- \equiv Z_-^{-1} \bar{A}_- \quad (2.41)$$

$$B_- = (1 - \epsilon^2 \lambda_-^2 \Omega^{-2}) \bar{B}_- = Z_-^{-1} \bar{B}_- . \quad (2.42)$$

To fix the four constants left we impose initial conditions on position and velocity for both x and y , as always.

Considering both at rest at $t = 0$ – that is, $\dot{x}(0) = 0$, $\dot{y}(0) = 0$, $x(0) \equiv x_0$, $y(0) \equiv y_0$ – we find

$$A_+ = B_+ = -\frac{1}{2(K_+ - K_-)} (K_- y_0 - x_0 + \Omega^{-2} (K_- \bar{g} - \bar{\bar{g}}))$$

$$A_- = B_- = \frac{1}{2(K_+ - K_-)} (K_+ y_0 - x_0 + \Omega^{-2} (K_+ \bar{g} - \bar{\bar{g}}))$$

and so the final time laws will assume the form

$$y(t) = 2A_+ \cos(\lambda_+ t) + 2A_- \cos(\lambda_- t) - \bar{g}\Omega^{-2} \quad (2.43)$$

$$x(t) = 2K_+ A_+ \cos(\lambda_+ t) + 2K_- A_- \cos(\lambda_- t) - \bar{\bar{g}}\Omega^{-2} . \quad (2.44)$$

In order to better understand the behaviour of the system, we replace λ_+ and λ_- with their first order expansion (2.35)-(2.36)⁷. We get

$$y(t) = 2A_+ \cos(\Omega t) + 2A_- \cos\left(\frac{1}{\epsilon}\Omega t\right) - \bar{g}\Omega^{-2} \quad (2.45)$$

$$x(t) = 2A_+ \cos(\Omega t) - 2\frac{A_-}{\epsilon^2} \cos\left(\frac{1}{\epsilon}\Omega t\right) - \bar{\bar{g}}\Omega^{-2} \quad (2.46)$$

with

$$A_+ \simeq \frac{1}{2} (y_0 + \bar{g}\Omega^{-2}) = O(1)$$

$$A_- \simeq \frac{1}{2}\epsilon^2 (y_0 - x_0 + (\bar{g} - \bar{\bar{g}})\Omega^{-2}) = O(\epsilon^2) .$$

In Fig. 2.2 we plot these approximate solutions⁸. We clearly see how – as we could expect – whereas the lightest mass is strongly influenced by the motion of M , the latter is just slightly perturbed around a common harmonic behaviour: it is lightly interested by the motion of m .

⁷In fact, λ_- needs its second order term in ϵ when plugged into K_- .

⁸The Fig. 2.1 lead us to believe that $x > y \forall t$, even if the starting Lagrangian never impose that. In this sense m and M have to be thought as non-impenetrable as well as the springs. On contrary, because of the form given to the gravity potential, the floor has to be impenetrable: the energy would be, otherwise, undefinite.

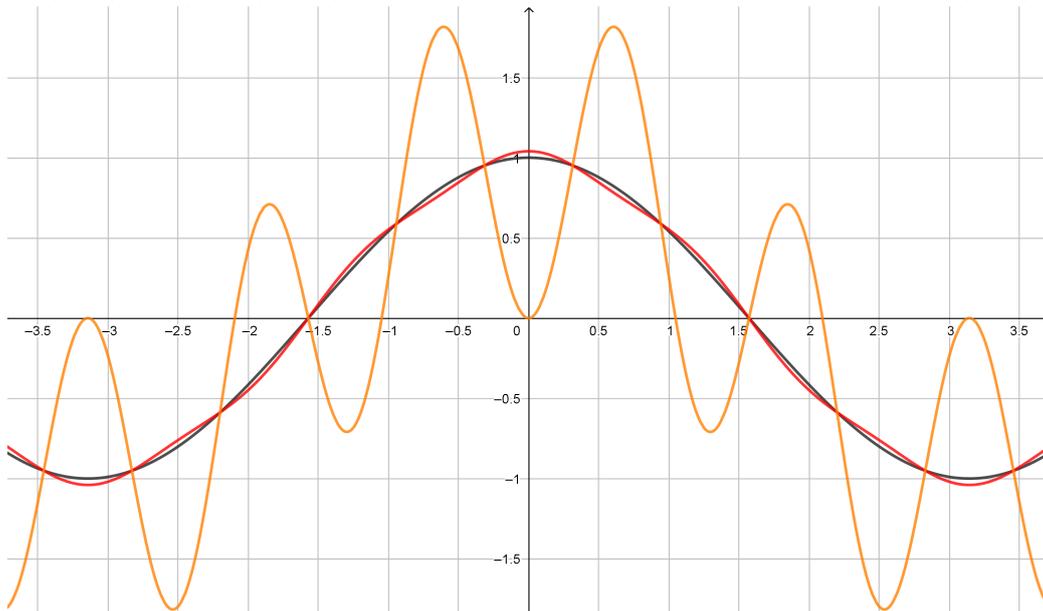


Figure 2.2: Plot of $f(t) = x(t) + \bar{g}\Omega^{-2}$ (orange) and $g(t) = y(t) + \bar{g}\Omega^{-2}$ (red) compared to $\cos(\Omega t)$ (black). The values assumed are: $\Omega = 1$ (Hz), $\epsilon = 0.2$ ($M = 5m$), $A_+ = \frac{1}{2}$ (50 cm), $A_- = \frac{1}{2}\epsilon^2$ (2 cm). Remember that $x(t), y(t) > 0$, since the floor must be impenetrable. With the values assumed here (and $g = 9.81$ m/s^2) this implies $\ell \geq 6.3$ m .

Now imagine that the lightest mass was inaccessible, in the sense that we are not able to measure its position (or velocity) throughout the time evolution of the system. We left with just y or, better, with its equations of motion (2.32). One can think that the situation is still the same, but the system is deeply changed. Let's see in details.

Since the EOM for y is unchanged after this simple assumption, the solution (2.33) is still valid. The big difference is that now to fix the four free constants A_+ , B_+ , A_- , B_- we need not only the initial position and the initial velocity of M , but also its initial acceleration and its initial jolt (\ddot{y}). This means that the body of mass M , whose position is labelled by y , has now four degrees of freedom: in a certain sense the two d.o.f.'s of x move to y changing drastically the physical features of the system. The inaccessibility to x leads to a classical system determined by its initial position, velocity, acceleration and jolt, something that we never meet in physics and something completely different – under a conceptual point of view – from the previous case.

In order to solve the system under this "*inaccessibility assumption*", we can impose for simplicity a non-jolting ($\dot{y}(0) \equiv j_0 = 0$), non-accelerating ($\ddot{y}(0) \equiv a_0 = 0$), non-moving ($\dot{y}(0) \equiv v_0 = 0$) body at $t = 0$. We find that

$$y(t) = 2C [\cos(\lambda_+ t) - r^2 \cos(\lambda_- t)] - \bar{g}\Omega^{-2} \quad (2.47)$$

with

$$r = \frac{\lambda_+}{\lambda_-} \quad (2.48)$$

and

$$C = A_+ = B_+ = -\frac{1}{r^2}A_- = -\frac{1}{r^2}B_- = \frac{1}{2(1-r^2)} (y_0 + \bar{g}\Omega^{-2}) . \quad (2.49)$$

In order to evaluate qualitatively the situation we consider again the approximation (2.35)-(2.36), for which

$$y(t) = 2C \left[\cos(\Omega t) - \epsilon^2 \cos\left(\frac{1}{\epsilon}\Omega t\right) \right] - \bar{g}\Omega^{-2} \quad (2.50)$$

with

$$C \simeq \frac{1}{2(1-\epsilon^2)} (y_0 + \bar{g}\Omega^{-2}) \quad (2.51)$$

– since

$$r \simeq \epsilon . \quad (2.52)$$

In Fig. 2.3 is plotted this latter solution together with the previous one. As expected, except for its initial position, the motion of M is exactly the same as before (its equation of motion are indeed unchanged).

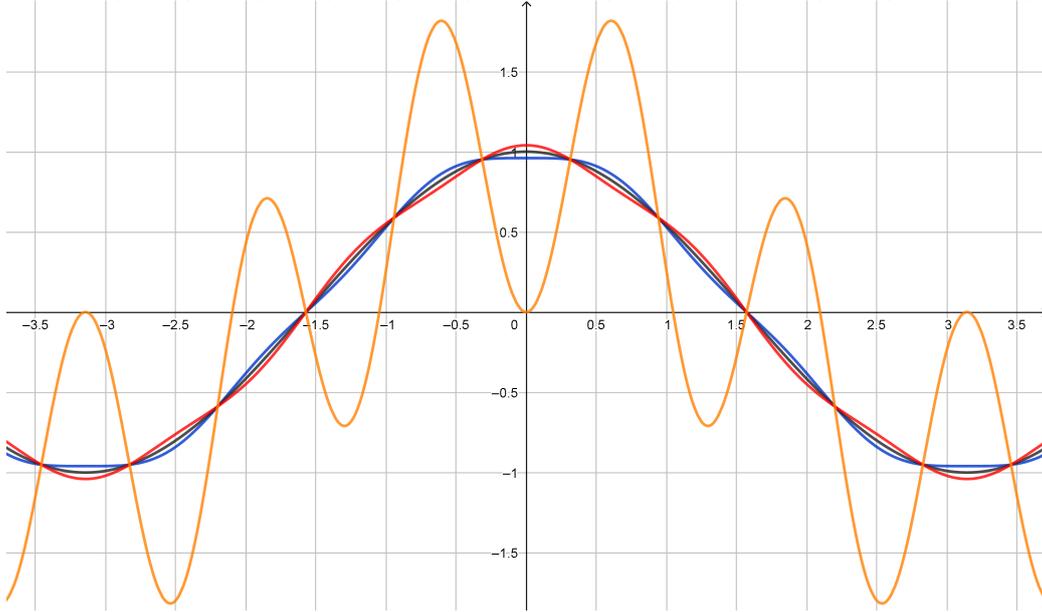


Figure 2.3: Plot of $h(t) = y(t) + \bar{g}\Omega^{-2}$ (blue) overlapped to the previous Fig. 2.2. The values assumed are: $\Omega = 1$ (Hz), $\epsilon = 0.2$, $C = \frac{1}{2}$ (50 cm).

We could stop here stating that, however unpleasant and singular it could be, this was a particular system which needed initial data on position, velocity, acceleration and jolt to just describe one body of mass M . But what does this statement really mean?

To answer this question we have to wonder about which Lagrangian – containing just y and its derivatives – would generate the equation of motion (2.33). The easiest Lagrangian we can think of is

$$L = \frac{1}{2}\dot{y}^2 - \bar{g}y - \frac{1}{2}\Omega^2 y^2 + \epsilon^2 \dot{y}^2 - \frac{1}{2}\epsilon^2 \Omega^{-2} \ddot{y}^2 . \quad (2.53)$$

From it, we can compute (with (2.5) and (2.6)) the conjugate momenta

$$\Pi_y = \frac{\partial L}{\partial \dot{y}} - D \frac{\partial L}{\partial \ddot{y}} = (1 + 2\epsilon^2)\dot{y} + \epsilon^2 \Omega^{-2} \ddot{y} \quad (2.54)$$

$$\Pi_{\dot{y}} = \frac{\partial L}{\partial \ddot{y}} = -\epsilon^2 \Omega^{-2} \ddot{y} \quad (2.55)$$

and, then, the Hamiltonian

$$H = \Pi_y \dot{y} - \frac{1}{2}\epsilon^{-2}\Omega^2 \Pi_{\dot{y}}^2 - \frac{1}{2}(1 + 2\epsilon^2)\dot{y}^2 + \bar{g}y + \frac{1}{2}\Omega^2 y^2 . \quad (2.56)$$

Several observations can be done now. First notice that we can cure the extra degrees of freedom in L or, better, in y 's equations of motion by taking the limit $\epsilon \rightarrow 0$: every higher-derivative term is indeed multiplied by this parameter. For this limit, $\eta \rightarrow \Omega^2 \ell$ and so the system becomes the one of Fig. 2.4. To prove it is sufficient to write down the Lagrangian for such a system

$$\begin{aligned} L_0 &= \frac{1}{2} \dot{y}^2 - gy - \frac{1}{2} \Omega^2 (y - \ell)^2 \\ &= \frac{1}{2} \dot{y}^2 - (g - \Omega^2 \ell) y - \frac{1}{2} \Omega^2 y^2 + \text{const} \\ &= \frac{1}{2} \dot{y}^2 - \bar{g} y - \frac{1}{2} \Omega^2 y^2 + \text{const} , \end{aligned}$$

where in the last line has been considered $\eta = \Omega^2 \ell$.

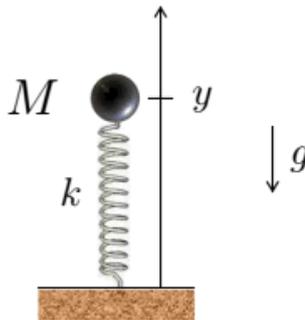


Figure 2.4: The effective system at zero-th order.

The three terms $-\epsilon^2 gy$, $\epsilon^2 \dot{y}^2$ and $-\frac{1}{2} \epsilon^2 \Omega^{-2} \ddot{y}^2$ could be then thought, incorrectly, as some sort of corrections to a common harmonic motion. This is false because not only the addition of such terms leads to extra degrees of freedom – as we just saw – but it also gives a big problem with energy. An anticipation can be found by inspecting H . Here in fact we are not allowed to take the limit $\epsilon \rightarrow 0$ anymore: the second term would blow up leading to $-\infty$ values. In the present case of vanishing v_0 , a_0 and j_0 the problem is disguised, indeed the energy of the system is

$$E = \bar{g} y_0 + \frac{1}{2} \Omega^2 y_0^2 , \quad (2.57)$$

the same of the system in Fig. 2.4 when the body of mass M has zero initial velocity. Since $y(t) \geq 0$ this energy is always positive for every allowed value of y_0 .

The higher-derivative case described by the Lagrangian (2.53) shows its energy problems when we consider an initial non-vanishing acceleration $a_0 \neq 0$. In this case we'll have

$$y(t) = 2C \left[\cos(\lambda_+ t) - \left(r^2 + \frac{a_0}{2C\lambda_-^2} \right) \cos(\lambda_- t) \right] - \bar{g}\Omega^{-2} \quad (2.58)$$

and, more importantly,

$$E = \bar{g}y_0 + \frac{1}{2}\Omega^2 y_0^2 - \frac{1}{2}\epsilon^2 \Omega^{-2} a_0^2 . \quad (2.59)$$

Indeed, as we can see in eqs. (2.54)-(2.56),⁹

$$E = H(y_0, v_0, \Pi_y(v_0, j_0), \Pi_{\dot{y}}(a_0)) \quad (2.60)$$

but while we still have $v_0 = 0$ and $j_0 = 0$, now $a_0 \neq 0$ and so $\Pi_{\dot{y}} \neq 0$ at $t = 0$. Specifically we get $\Pi_{\dot{y}}(\dot{y}(0)) = -\epsilon^2 \Omega^{-2} a_0$.

Eq. (2.59) clearly shows how the energy now is undefined and so how it can assume arbitrarily negative values for sufficiently high values of a_0 . In particular we start getting negative values for the energy when

$$a_0^2 > \epsilon^{-2} (2\bar{g} + \Omega^2 y_0) \Omega^2 y_0 \quad (2.61)$$

and in our case (see Fig. 2.3 for the values chosen) for a "rest length"¹⁰ $\ell = 10 \text{ m}$ – sufficient to assure $y(t) \geq 0 \forall t$ – we'd get such negative energies for every initial acceleration $a_0 \gtrsim 5 \text{ m/s}^2$. The problem here is not so much the negative values for E but rather the lack of a ground state.

This problem persists when the system is quantized. In order to see this we first define

$$z = y + \bar{g}\Omega^{-2} \quad (2.62)$$

so that the Lagrangian (2.53) assumes the form

$$L = \frac{1}{2}(1 + 2\epsilon^2)\dot{z}^2 - \frac{1}{2}\Omega^2 z^2 - \frac{1}{2}\epsilon^2 \Omega^{-2} \ddot{z}^2 . \quad (2.63)$$

The Hamiltonian for such a higher-derivative Lagrangian will be

$$H = \Pi_z \dot{z} - \frac{1}{2}\epsilon^{-2}\Omega^2 \Pi_{\dot{z}}^2 - \frac{1}{2}(1 + 2\epsilon^2)\dot{z}^2 + \frac{1}{2}\Omega^2 z^2 \quad (2.64)$$

⁹Thanks to the Noether's Theorem: the Hamiltonian does not depend explicitly on t and so we know it's a constant of motion.

¹⁰In this description the system cannot be considered anymore as composed by springs with bodies attached. A physical system described by the Lagrangian (2.53) as it stands is unknown.

and introducing the variables q_+ , p_+ , q_- and p_- as

$$\begin{cases} q_+ = -\frac{i}{\sqrt{2}\Omega} ((1 + \epsilon^2)\dot{z} - \Pi_z) \\ p_+ = \frac{\Omega}{\sqrt{2}\epsilon} (\sqrt{2}\epsilon z - \Pi_{\dot{z}}) \\ q_- = \frac{i}{\sqrt{2}\epsilon\Omega} (\epsilon^2\dot{z} - \Pi_z) \\ p_- = -\frac{\Omega}{\epsilon} \left(\frac{\epsilon}{\sqrt{2}}z - \Pi_{\dot{z}} \right) \end{cases} \quad (2.65)$$

– or, equivalently,

$$\begin{cases} z = \frac{\sqrt{2}}{\Omega} (\sqrt{2}p_+ + p_-) \\ \Pi_{\dot{z}} = \frac{\sqrt{2}\epsilon}{\Omega} (p_+ + \sqrt{2}p_-) \\ \dot{z} = \frac{i\sqrt{2}\Omega}{\epsilon} (\epsilon q_+ + q_-) \\ \Pi_z = i\sqrt{2}\epsilon\Omega \left(\epsilon q_+ + \frac{1+\epsilon^2}{\epsilon^2}q_- \right) \end{cases} \quad (2.66)$$

– it becomes

$$H = p_+^2 + \Omega^2 q_+^2 - \left(p_-^2 + \frac{\Omega^2}{\epsilon^2} q_-^2 \right), \quad (2.67)$$

difference of two harmonic oscillators' Hamiltonians.

From here the quantization is straightforward and leads to energy eigenvalues determined by two non-negative integers $n, m = 0, 1, 2, \dots$:

$$E = \left(n + \frac{1}{2} \right) \Omega - \left(m + \frac{1}{2} \right) \frac{\Omega}{\epsilon}. \quad (2.68)$$

This energy is again clearly indefinite: fixing ϵ , we have an infinite tower of energy levels from $-\infty$ to $+\infty$. Notice that now if we take $\epsilon \rightarrow 0$, then $E \rightarrow -\infty$ and we do not approach the purely simple harmonic oscillator ground state.

More in general the limit $\epsilon \rightarrow 0$ does not recover the system in Fig. 2.4 as we at first imagined studying the Lagrangian. That system has nothing to do with our case: once ϵ is non-vanishing the two systems describe very different situations with deeply different physics.

This lack of a ground state is unacceptable and signals – more than all the other problems (such as never-seen extra degrees of freedom or the impossibility of recovering the lower-derivative theory taking $\epsilon \rightarrow 0$) – an unhealthy treatment of this problem.

The correct approach is the one of *perturbative constraints*: eq. (2.32) does not describe the system by itself, it must be accompanied by the constraints provided by the JLM procedure. Under this approach the terms $\propto \epsilon^2$ are considered from the very beginning as perturbative correction (up to ϵ^2

order) of the 0-th order system of Fig. 2.4 and the constraints imposed are necessary to keep only the solutions that are perturbative corrections (up to ϵ^2 order) of the 0-th order solution.

For these reasons the Lagrangian (2.53) can be rewritten as

$$L = \frac{1}{2}\dot{y}^2 - (g - \Omega^2\ell)y - \frac{1}{2}\Omega^2y^2 - \epsilon^2gy + \epsilon^2\dot{y}^2 - \frac{1}{2}\epsilon^2\Omega^{-2}\ddot{y}^2 + O(\epsilon^3) . \quad (2.69)$$

This Lagrangian is now in the form (2.9) ($n = 2$) and the JLM method can thus be applied. We identify

- $m = 1$ ¹¹;
- $\mathfrak{W}_0 = -(g - \Omega^2\ell)y - \frac{1}{2}\Omega^2y^2$;
- $\mathfrak{W}_1 = 0$;
- $\mathfrak{W}_2 = -gy + \dot{y}^2 - \frac{1}{2}\Omega^{-2}\ddot{y}^2$.

Because of the very simple form of this theory and of the fact that we already know the EOM – to be fair they are our starting point, the Lagrangian has just been derived from them – there's no need for computing the \mathfrak{W}_k 's. The **primary constraints** are indeed obtained by the equation of motion

$$\epsilon^2\Omega^{-2}y^{(4)} + (1 + 2\epsilon^2)\ddot{y} + \Omega^2y + \bar{g} = O(\epsilon^3) \quad (2.70)$$

by multiplying by ϵ :

$$\epsilon [\ddot{y} + (g - \Omega^2\ell) + \Omega^2y] = O(\epsilon^3) . \quad (2.71)$$

Now all we need is a constraint for $\propto \epsilon^2\ddot{y}$ and a constraint for $\propto \epsilon^2y^{(4)}$. To find them we just multiply (2.71) by ϵ and differentiate this result twice substituting properly to get rid of the \ddot{y} dependence. The result is

$$\epsilon^2\ddot{y} = -(g - \Omega^2\ell + \Omega^2y)\epsilon^2 + O(\epsilon^3) \quad (2.72)$$

$$\epsilon^2y^{(4)} = \Omega^2(g - \Omega^2\ell + \Omega^2y)\epsilon^2 + O(\epsilon^3) . \quad (2.73)$$

Notice that these steps are exactly the first two steps of the JLM procedure illustrated above: (2.72) substitutes (2.14) and (2.73) substitutes (2.16).

The few relations found so far allow us to skip to the end of the JLM algorithm, where the equation of motion is used to find what generates the most

¹¹Here we deal with just one particle, then the α label is useless and, so, suppressed.

general set of secondary constraints: eq. (2.26). Such equation represents also the *healthy* EOM and in this case it is

$$\begin{aligned}\ddot{y} &= -g + (1 - \epsilon^2)\Omega^2\ell - \Omega^2(1 - \epsilon^2)y + O(\epsilon^3) \\ &= -\Omega^2(1 - \epsilon^2)(y + \bar{g}\Omega^{-2}) + O(\epsilon^3) .\end{aligned}\quad (2.74)$$

The solution to this equation is

$$y(t) = Ae^{i\lambda t} + Be^{-i\lambda t} - \bar{g}\Omega^{-2} \quad (2.75)$$

with

$$\lambda = \left(1 - \frac{1}{2}\epsilon^2\right)\Omega . \quad (2.76)$$

Considering as before $\dot{y}(0) = 0$ we get $A = B$ and then

$$y(t) = 2A \cos(\lambda t) - \bar{g}\Omega^{-2} , \quad (2.77)$$

an usual oscillatory behaviour. This solution can be visualized in Fig. 2.5. We realize that the perturbative correction is applied to the oscillation frequency: whereas the 0-th order system (Fig. 2.4) oscillates with the frequency Ω , this "corrected system" oscillates with the (lower) frequency $(1 - \frac{1}{2}\epsilon^2)\Omega$.

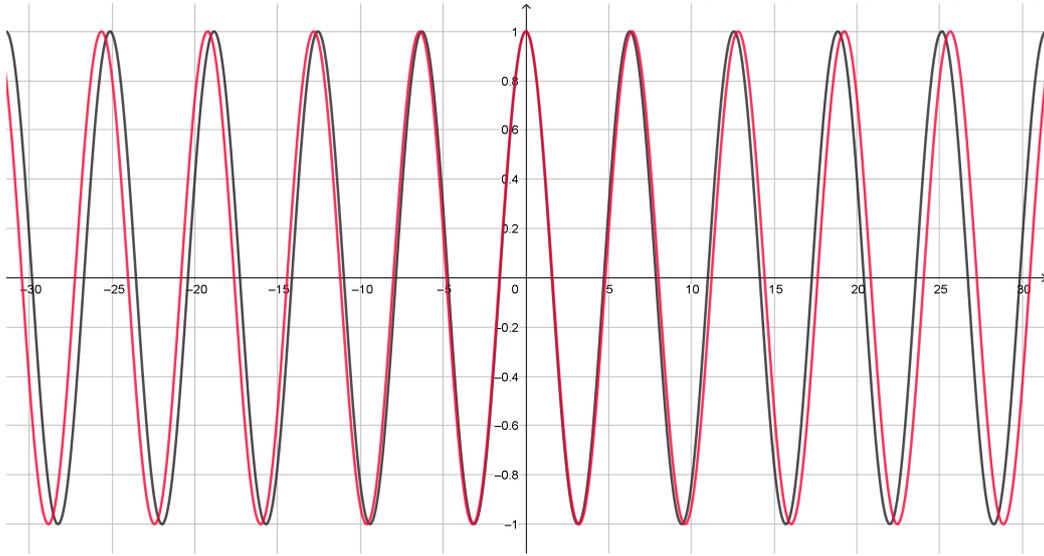


Figure 2.5: Plot of $k(t) = y(t) + \bar{g}\Omega^{-2}$ (pink) overlapped to $\cos(\Omega t)$ (black). The values assumed are: $\Omega = 1$ (Hz), $\epsilon = 0.2$, $A = \frac{1}{2}$ (50 cm).

Comparing (2.75) with (2.33) we immediately realize how now two initial conditions are sufficient to determine the motion: we now have only two free

constants, A and B , instead of four.

However, the imposition of the correct constraints not only eliminated the extra d.o.f.'s, it also cured the lack of a ground state. Indeed taking the constraints in their Hamiltonian form, that is¹²

$$\Pi_y = (1 + \epsilon^2)\dot{y} + O(\epsilon^3) \quad (2.78)$$

$$\Pi_{\dot{y}} = \epsilon^2 \Omega^{-2} (g - \Omega^2 \ell + \Omega^2 y) + O(\epsilon^3) , \quad (2.79)$$

and substituting them into the higher-derivative Hamiltonian (2.56), we obtain the **reduced Hamiltonian**

$$H_R = \frac{1}{2}(1 - 2\epsilon^2)\Pi_y^2 + (1 - \epsilon^2)\bar{g}y + \frac{1}{2}(1 - \epsilon^2)\Omega^2 y^2 - E_0 + O(\epsilon^3) \quad (2.80)$$

with

$$E_0 = \frac{1}{2}\epsilon^2 \Omega^{-2} (g - \Omega^2 \ell)^2 . \quad (2.81)$$

We can observe that now this has a lower bound (remember that $y > 0$), that is $-E_0$, and that it gives the correct Hamiltonian for the $\epsilon \rightarrow 0$ limit, the one we obtain from L_0 for the 0-th order system.

We could stop here: we cured the equations of motions (obtaining eq. (2.74)) and we got a well-behaved energy (obtainable from (2.80)). In particular, considering as always $\dot{y}(0) = 0$ the latter is

$$E_R = (1 - \epsilon^2)\bar{g}y_0 + \frac{1}{2}(1 - \epsilon^2)\Omega^2 y_0^2 - E_0 + O(\epsilon^3) . \quad (2.82)$$

Nevertheless, we could also ask ourselves about how to recover the EOM (2.74) from the Hamiltonian (2.80). The path is not as straightforward as it could seem. The correct equations of motion for a constrained Hamiltonian system are indeed

$$\frac{df}{dt} = \{f, H_R\}_{DB} + \frac{\partial f}{\partial t} \quad (2.83)$$

where $f = f(y, \Pi_y, t)$ is the quantity we want to time-evolve and $\{\cdot, \cdot\}_{DB}$ are the Dirac Brackets, a generalization of the Poisson Brackets that involve the constraints in their definition. Using the same notation of [2], if we write the constraints as

$$\chi = \Pi_{\dot{y}} - \epsilon^2 \Omega^{-2} (g - \Omega^2 \ell + \Omega^2 y) = O(\epsilon^3) \quad (2.84)$$

$$\omega = (1 + \epsilon^2) - \Pi_y = O(\epsilon^3) \quad (2.85)$$

¹²To obtain these relations we used the constraint (2.73) and the time-derivative of (2.72).

we can define the Dirac Brackets through

$$\{f, g\}_{DB} = \{f, g\}_{PB} + \{f, \chi\}_{PB} \{g, \omega\}_{PB} - \{g, \chi\}_{PB} \{f, \omega\}_{PB} , \quad (2.86)$$

where $\{\cdot, \cdot\}_{PB}$ are the Poisson Brackets. Notice that, in such a higher-derivative framework, these latter read as

$$\{f, g\}_{PB} = \sum_{k=0}^{n-1} \left(\frac{\partial f}{\partial y^{(k)}} \frac{\partial g}{\partial \Pi_{y^{(k)}}} - \frac{\partial f}{\partial \Pi_{y^{(k)}}} \frac{\partial g}{\partial y^{(k)}} \right) . \quad (2.87)$$

Finally we can compute

$$\frac{d\Pi_y}{dt} = \{\Pi_y, H_R\}_{DB} = -(1 + \epsilon^2) \frac{\partial H_R}{\partial y} \quad (2.88)$$

that is equivalent to (2.74) once is taken into account (2.78).

An easier way to check the consistency of the EOM with the Hamiltonian (2.80) consists in getting first a **reduced Lagrangian**, and then to use the usual Euler-Lagrange equations to obtain eq. (2.74).

This method is illustrated in [1] and is based on finding a new conjugate momentum to y so that usual Hamilton-Jacobi equations (or equivalently Euler-Lagrange equations) hold true. In general, in order to find this new momentum p is necessary to express the energy of our constrained system as a function of y and \dot{y} . What we do is then to plug eq. (2.78) into (2.80), getting

$$E_R(y, \dot{y}) = \frac{1}{2} \dot{y}^2 + (1 - \epsilon^2) \bar{g} y + \frac{1}{2} (1 - \epsilon^2) \Omega^2 y^2 - E_0 + O(\epsilon^3) , \quad (2.89)$$

to then apply the formula for $p(y, \dot{y})$

$$p(y, \dot{y}) = \int_0^{\dot{y}} \frac{dv}{v} \frac{\partial E_R(y, v)}{\partial \dot{y}} + p(y, 0) = \dot{y} . \quad (2.90)$$

In the last step $p(y, 0)$ has been set to zero. The freedom of choosing $p(y, 0)$ is linked to the freedom of adding total derivatives to L without changing the physics of the system.

We are finally ready to compute the reduced Lagrangian:

$$\begin{aligned} L_R &= p(y, \dot{y}) \dot{y} - E_R(y, \dot{y}) \\ &= \frac{1}{2} \dot{y}^2 - (1 - \epsilon^2) \bar{g} y - \frac{1}{2} (1 - \epsilon^2) \Omega^2 y^2 + O(\epsilon^3) \end{aligned} \quad (2.91)$$

from which the equation of motion (2.74) is easily obtainable.

It is interesting to see what would happen if we took another way: once found all the constraints we decide, instead of building the Hamiltonian, to plug them directly in our higher-derivative Lagrangian (2.69). What we'd obtain is

$$\tilde{L}_R = \frac{1}{2}(1 + 2\epsilon^2)\dot{y}^2 - (1 + \epsilon^2)\bar{g}y - \frac{1}{2}(1 + \epsilon^2)\Omega^2 y^2 + O(\epsilon^3) . \quad (2.92)$$

This Lagrangian gives the same equations of motion – once treated correctly the ϵ -polynomials (see Appendix A) – since it's basically the same Lagrangian as L_R rescaled

$$(1 + 2\epsilon^2)L_R = \tilde{L}_R + O(\epsilon^3) . \quad (2.93)$$

At first sight this \tilde{L}_R seems to generate a completely different Hamiltonian, however one has to keep in mind that the conjugate momentum is not the previous one:

$$\tilde{\Pi}_y = \frac{\partial \tilde{L}_R}{\partial \dot{y}} = (1 + 2\epsilon^2)\dot{y} \quad (2.94)$$

so that

$$\tilde{H}_R = \frac{1}{2}(1 - 2\epsilon^2)\tilde{\Pi}_y^2 + (1 + \epsilon^2)\bar{g}y + \frac{1}{2}(1 + \epsilon^2)\Omega^2 y^2 + O(\epsilon^3) . \quad (2.95)$$

To correctly compare this Hamiltonian with (2.80) we have to express \tilde{H}_R in terms of Π_y . To do this we use the relation (2.78), which tells us that

$$\tilde{\Pi}_y = (1 + \epsilon^2)\Pi_y \quad (2.96)$$

and, finally, that¹³

$$(1 + 2\epsilon^2)H_R = \tilde{H}_R + O(\epsilon^3) . \quad (2.97)$$

The Hamiltonian, and so the energy of the system, is just the previous one rescaled. This fact is not particularly disturbing since in classical physics E (or equivalently H in this case) is just an integral of motion, a fixed constant independent on time that, even if rescaled, does not affects the dynamics of the system: its most important feature is its constant nature. To prove it we can once again compute EOM's, this time using \tilde{H}_R . In this case the system is healthy from the very beginning, thus we have to apply the usual Poisson Brackets. We can proceed in two equivalent ways: we can compute

$$\frac{d\tilde{\Pi}_y}{dt} = \{\tilde{\Pi}_y, \tilde{H}_R\}_{PB} = -\frac{\partial \tilde{H}_R}{\partial y} \quad (2.98)$$

¹³Here H_R has been shifted to $H_R + E_0$.

and then use (2.78) and (2.96) to get the desired result; or directly calculate

$$\frac{d\Pi_y}{dt} = \{\Pi_y, \tilde{H}_R\}_{PB} = -\frac{\partial\Pi_y}{\partial\tilde{\Pi}_y} \frac{\partial\tilde{H}_R}{\partial y} = -(1 - \epsilon^2) \frac{\partial\tilde{H}_R}{\partial y} \quad (2.99)$$

that, again using (2.78), gives the correct equation of motion.

There's then an ambiguity in the choice between these two ways of acting: we could state that since these two paths lead to the same EOM and to the same energy (except for a rescaling constant), they are equivalent. This statement, **in classical physics**, could at least sound dangerous. The problem lies in the substitution of equations of motion inside the functional generating them. Indeed we are, in principle, not allowed to perform such a substitution – it in general modifies the very nature of the variational principle underlying the formalism – but while for the Hamiltonian treatment it is part of the correct procedure to treat constrained systems (one computes \mathcal{H}_R and then uses Poisson brackets to find the correct EOM's¹⁴), it seems just an illegal shortcut as regards the Lagrangian (a correct treatment here would require the introduction of constraints through Lagrangian multipliers). *Maybe* here the result is equivalent because we are eliminating, through substitution of EOM's, unphysical degrees of freedom as higher derivatives.

We will see that in Classical Field Theory this operation is safely permitted and, so, that working with only Lagrangians is possible.

2.2 Classical Field Theory version

The treatment in case of Field Theory has been generalized naturally thinking on the comparison between the usual Euler-Lagrange equations of motion:

$$\frac{\partial\mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial\mathcal{L}}{\partial\dot{q}} = 0 \quad \longleftrightarrow \quad \frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} = 0. \quad (2.100)$$

It is then straightforward to apply the same method – with the substitutions

$$\begin{aligned} q &\longrightarrow \varphi \\ D &\longrightarrow \partial_\mu \\ q^{(n)} &\longrightarrow \partial_\alpha\partial_\beta\dots\partial_\nu\varphi \equiv \partial^{(n)}\varphi \\ D^n &\longrightarrow \partial_\alpha\partial_\beta\dots\partial_\nu \end{aligned} \quad (2.101)$$

¹⁴See [2].

– to an higher-derivative classical Field Theory Lagrangian

$$\mathcal{L} = -\frac{1}{2} \sum_{a=1}^N (\partial\varphi_a)^2 + \sum_{k=0}^n \epsilon^k \mathfrak{W}_k(\varphi, \partial\varphi, \dots, \partial^{(k)}\varphi) + O(\epsilon^{n+1}), \quad (2.102)$$

with $(\partial\varphi)^2 \equiv \partial\varphi \cdot \partial\varphi \equiv \partial_\mu\varphi\partial^\mu\varphi$ and $a = 1, \dots, N$ labeling N different real fields.

To maintain the notation of [4], the flat metric used to contract space-time indices is taken to be the *mostly plus* one: $\eta_{\mu\nu} = \text{diag}\{-1, +1, +1, +1\}$ (origin of the minus sign in front of the kinetic term in \mathcal{L}).

The Euler-Lagrange equations will assume the form¹⁵

$$\sum_{k=0}^n (-1)^k \underbrace{\partial_\alpha \partial_\beta \dots \partial_\kappa}_{k \text{ derivatives}} \frac{\partial \mathcal{L}}{\partial (\partial_\alpha \partial_\beta \dots \partial_\kappa \varphi_a)} = 0 \quad (2.103)$$

that, applied to the Lagrangian (2.102), read as¹⁶

$$\partial^2 \varphi_a + \sum_{k=0}^n \epsilon^k \mathfrak{W}_{a,k}(\varphi, \dots, \partial^{(2k)}\varphi) = O(\epsilon^{n+1}) \quad \forall \alpha \quad (2.104)$$

with $\partial^2 \varphi \equiv \partial_\mu \partial^\mu \varphi$ and

$$\mathfrak{W}_{a,k} \equiv \sum_{r=0}^k (-1)^r \partial_\alpha \partial_\beta \dots \partial_\rho \frac{\partial \mathfrak{W}_k}{\partial (\partial_\alpha \partial_\beta \dots \partial_\rho \varphi_a)}. \quad (2.105)$$

From these equations we can easily derive the **primary constraints** of the system

$$\epsilon \left[\partial^2 \varphi_a + \sum_{k=0}^{n-1} \epsilon^k \mathfrak{W}_{a,k}(\varphi, \dots, \partial^{(2k)}\varphi) \right] = O(\epsilon^{n+1}) \quad \forall \alpha. \quad (2.106)$$

Notice that one big difference with the classical case is about the mass of the particle: here it is never encoded in the kinetic term, neither in the

¹⁵From now on we won't write "k derivatives", implying that every time there's a match between greek and latin letters the number of derivatives is considered the same. E.g. the derivative with the Lorentz (greek) index "kappa" will be the k-th derivative whereas a derivative with the index "nu" will be the n-th derivative. Exceptions will be "alpha" and "beta", nearly always considered the first and the second derivatives.

¹⁶One could argue that the first term of this equation, as well as the Lagrangian's first one, has opposite sign with respect to its classical version (2.11). Actually, expanding the contracted indices with the mostly-plus metric, one would realize that time derivatives of the field have the same sign.

easiest $n = 0$ (usual) case where it would lie in $\mathfrak{W}_{a,0} \supset -\frac{1}{2}m_a\varphi_a^2$.

To get the secondary constraints the algorithm will be exactly the same: one starts from primary constraints, multiplies by ϵ^{n-1} and takes derivatives substituting $\epsilon^n\partial^2\varphi_a$ when it appears; then one restarts from primary constraints, multiplies by ϵ^{n-2} and takes derivatives substituting $\epsilon^{n-1}\partial^2\varphi_a$; the iteration of these steps will lead to an expression for $\epsilon\partial^2\varphi_a$ that doesn't contain derivatives of the field higher than one.

Notice that in this case we have to pay close attention to the contractions of indices when differentiating. In particular, while we can decide whether to contract or not a new derivative in the midst of the algorithm to find secondary constraints, we have no choice on the very first move: the primary constraint. Indeed such a restriction will always concern $\partial^2\varphi_a$ and never the more general $\partial_\mu\partial_\nu\varphi_a$, since primary constraints – descending directly from equations of motion – have their form fixed.

Then it follows that, because of the variety of possible contractions, the JLM method in the case of Classical Field Theory will be an ad hoc procedure, in the sense that it won't be much interested in finding every kind of constraint but more in searching those of our interest, relating to the treated Lagrangian. In turns the Lagrangian will be adapted to be constrained: this means that we will use the freedom of integrating by parts to make our Lagrangian suitable or *more* suitable for the applicability of the algorithm.

For example a term like

$$\epsilon^2\partial_\mu\varphi\partial_\nu\varphi\partial^\mu\partial^\nu\varphi$$

in \mathcal{L} will be written as

$$\frac{1}{2}\epsilon^2\partial_\mu\varphi\partial^\mu((\partial\varphi)^2)$$

and then integrated by parts to get

$$-\frac{1}{2}\epsilon^2\partial^2\varphi(\partial\varphi)^2,$$

on which the constraint

$$\epsilon^2 \left[\partial^2\varphi_a + \sum_{k=0}^{n-2} \epsilon^k \mathfrak{Z}_{a,2,k}(\varphi, \partial\varphi) \right] = O(\epsilon^{n+1}), \quad (2.107)$$

Classical Field Theory version of (2.24), can be used.

Without delay we are already going to analyse a practical application in case of Classical Field Theory.

2.2.1 Application: Light-Heavy Scalar Fields Interaction

This example is modeled on the case considered in [5] and is a typical Effective Field Theory application. As the previous one it shows how higher-derivative terms arise naturally when a light and a heavy degree of freedom are coupled.

The Lagrangian considered is ¹⁷

$$\mathcal{L} = -\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}(\partial\chi)^2 - \frac{1}{2}m^2\phi^2 - \frac{1}{2}M^2\chi^2 - gM\phi^2\chi \quad (2.108)$$

with ϕ and χ real fields and $m \ll M$.

Notice that this Lagrangian is the same as (2.1) in [5] except for the \sqrt{g} factor (which has been removed) and the substitution

$$\lambda \longrightarrow gM, \quad (2.109)$$

performed to simplify the comparison with the JLM method.

There is something to specify about this substitution. One could think that, since $[\lambda] = 1$, this replacement was little more than a renaming where, from a dimensional coupling constant, we extracted its explicit mass dependence (M) leaving a generic dimensionless free coupling g . Actually, the substitution considered is implying something stronger: we are asking for a λ of the same order of M and for a $g = O(1)$. Indeed if we considered $\lambda = O(m)$ this renaming would lead to a $g = O(m/M)$ that, since the method requires $m \ll M$, would compete with our dimensional expansion parameter $1/M$, going to belong to the same infinitesimal order.

So, to avoid every problem, we ask for a large coupling between the two fields in the 3-vertex interaction.

As we are interested in physics on energy scales $E \ll M$, it is sufficient to study the system through a low-energy effective Lagrangian obtained integrating out the heaviest field χ , since its mass scale is above the energy scale of interest.

Thus we start by writing the equation of motion for χ

$$(\partial^2 - M^2)\chi = gM\phi^2 \quad (2.110)$$

¹⁷ $\hbar = 1, c = 1.$

that, inverted, becomes

$$\begin{aligned}
\chi &= gM(\partial^2 - M^2)^{-1}\phi^2 \\
&= -\frac{g}{M} \left(1 + \frac{\partial^2}{M^2} + \frac{(\partial^2)^2}{M^4} + \frac{(\partial^2)^3}{M^6} + \dots \right) \phi^2 \\
&= -\frac{g}{M} \left[\sum_{k=0}^{\infty} \left(\frac{\partial^2}{M^2} \right)^k \right] \phi^2 .
\end{aligned} \tag{2.111}$$

As anticipated, we choose as expansion parameter

$$\epsilon \equiv \frac{1}{M^2} \tag{2.112}$$

which in this case is dimensionful.

Once computed some useful terms

$$\begin{aligned}
\partial\phi^2 &= 2\phi\partial\phi \\
\partial^2\phi^2 &= 2(\partial\phi)^2 + 2\phi\partial^2\phi \\
\partial\partial^2\phi^2 &= 6\partial\phi\partial^2\phi + 2\phi\partial\partial^2\phi \\
\partial^2\partial^2\phi^2 &= 6(\partial^2\phi)^2 + 8\partial\phi \cdot \partial\partial^2\phi + 2\phi\partial^2\partial^2\phi
\end{aligned}$$

(where \cdot has been used to denote the contraction between two omitted Lorentz indices), we can deal with

$$\partial^2\chi = -\frac{g}{M} \left(\partial^2 + \frac{(\partial^2)^2}{M^2} + \frac{(\partial^2)^3}{M^4} + \frac{(\partial^2)^4}{M^6} + \dots \right) \phi^2 \tag{2.113}$$

and, finally, integrate out χ from our Lagrangian

$$\mathcal{L} = \mathcal{L}|_{\chi=0} + \frac{1}{2}\chi\partial^2\chi - \frac{1}{2}M^2\chi^2 - gM\phi^2\chi \tag{2.114}$$

– where an integration by parts has been performed.

The result, up to total derivatives and approximated to order ϵ^4 , will be

$$\begin{aligned}
\mathcal{L} &= \mathcal{L}|_{\chi=0} + \frac{1}{2}g^2\phi^4 \\
&\quad + \frac{1}{2}\epsilon g^2\phi^2\partial^2\phi^2 + \frac{1}{2}\epsilon^2 g^2(\partial^2\phi^2)^2 \\
&\quad + \frac{1}{2}\epsilon^3 g^2\partial^2\phi^2\partial^2\partial^2\phi^2 + O(\epsilon^4) \\
&= -\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \frac{1}{2}g^2\phi^4 \\
&\quad - 2\epsilon g^2\phi^2(\partial\phi)^2 + 2\epsilon^2 g^2 [(\partial\phi)^4 + \phi^2(\partial^2\phi)^2 + 2\phi(\partial\phi)^2\partial^2\phi] \\
&\quad - 2\epsilon^3 g^2 [\phi(\partial\phi)^2\partial^2\partial^2\phi - \phi^2(\partial\partial^2\phi)^2 \\
&\quad \quad + 4(\partial\phi)^2\partial\phi \cdot \partial\partial^2\phi - 4\phi\partial^2\phi\partial\phi \cdot \partial\partial^2\phi] + O(\epsilon^4) .
\end{aligned} \tag{2.115}$$

This Lagrangian is in the form (2.102)¹⁸, with $n = 3$. In particular

- $\mathfrak{Y}_0 = -\frac{1}{2}m^2\phi^2 + \frac{1}{2}g^2\phi^4$;
- $\mathfrak{Y}_1 = -2g^2\phi^2(\partial\phi)^2$;
- $\mathfrak{Y}_2 = 2g^2[(\partial\phi)^4 + \phi^2(\partial^2\phi)^2 + 2\phi(\partial\phi)^2\partial^2\phi]$;
- $\mathfrak{Y}_3 = -2g^2[\phi(\partial\phi)^2\partial - \phi^2\partial\partial^2\phi + 4(\partial\phi)^2\partial\phi - 4\phi\partial^2\phi\partial\phi] \cdot \partial\partial^2\phi$.

Notice that \mathfrak{Y}_3 contains a term $\propto \partial^2\partial^2\phi$ that shouldn't exist, since from (2.102) we deduce that \mathfrak{Y}_3 is supposed to depend only on ϕ , $\partial_\alpha\phi$, $\partial_\alpha\partial_\beta\phi$ and $\partial_\alpha\partial_\beta\partial_\gamma\phi$. It is possible to remove this dependence integrating by parts the Lagrangian to obtain just first, second and third derivatives of the field, but such an operation would lead to a problematic uncontracted second derivative $\partial_\mu\partial_\nu\phi$ that does not appear in any kind of constraint – as we already said – since the equations of motion restrict the form of the second derivative ones¹⁹.

Fortunately, keeping this ϵ^3 -four-derivative term does not affect the procedure, as we will see in a moment.

From these \mathfrak{Y}_k 's we can compute the \mathfrak{W}_k 's through (2.105)

- $\mathfrak{W}_0 = \frac{\partial\mathfrak{Y}_0}{\partial\phi} = -m^2\phi + 2g^2\phi^3$;
- $\mathfrak{W}_1 = \frac{\partial\mathfrak{Y}_1}{\partial\phi} - \partial_\alpha\frac{\partial\mathfrak{Y}_1}{\partial(\partial_\alpha\phi)} = -4g^2\phi(\partial\phi)^2 + 4g^2\phi^2\partial^2\phi$;
- $\mathfrak{W}_2 = \frac{\partial\mathfrak{Y}_2}{\partial\phi} - \partial_\alpha\frac{\partial\mathfrak{Y}_2}{\partial(\partial_\alpha\phi)} + \partial_\alpha\partial_\beta\frac{\partial\mathfrak{Y}_2}{\partial(\partial_\alpha\partial_\beta\phi)}$;
- $\mathfrak{W}_3 = \frac{\partial\mathfrak{Y}_3}{\partial\phi} - \partial_\alpha\frac{\partial\mathfrak{Y}_3}{\partial(\partial_\alpha\phi)} + \partial_\alpha\partial_\beta\frac{\partial\mathfrak{Y}_3}{\partial(\partial_\alpha\partial_\beta\phi)} - \partial_\alpha\partial_\beta\partial_\gamma\frac{\partial\mathfrak{Y}_3}{\partial(\partial_\alpha\partial_\beta\partial_\gamma\phi)}$.

Since in this case we just care about finding those relations necessary to cure the higher derivatives in \mathcal{L} and we are not interested in computing the most general secondary constraint, we did not write \mathfrak{W}_2 and \mathfrak{W}_3 explicitly.

Indeed, although the primary constraint

$$\epsilon [\partial^2\phi + \mathfrak{W}_0 + \epsilon\mathfrak{W}_1 + \epsilon^2\mathfrak{W}_2] = O(\epsilon^4) \quad (2.116)$$

¹⁸Here we deal with just one particle, then the a label is useless and, so, suppressed.

¹⁹We can move from a $\partial_\mu\partial_\nu\phi$ constraint to a $\partial^2\phi$ one but we cannot do the inverse: if we are handling a scalar equation we are not allowed to generalise to a tensorial equation (in this case multiplying by the inverse of the metric).

depends on the explicit form of \mathfrak{W}_2 , there are no terms $\propto \epsilon \partial^2 \phi$ in our Lagrangian (2.115) (its ϵ term is healthy, it does not contain higher derivatives), and then this constraint result useless.

On the contrary, we have – and need to constraint – higher-derivative terms like

$$\begin{aligned} &\propto \epsilon^2 \partial^2 \phi \\ &\propto \epsilon^3 \partial^2 \phi \\ &\propto \epsilon^3 \partial \partial^2 \phi \\ &\propto \epsilon^3 \partial^2 \partial^2 \phi . \end{aligned}$$

So, as always, we start multiplying the primary constraint by ϵ^2

$$\epsilon^3 \partial^2 \phi = (m^2 \phi - 2g^2 \phi^3) \epsilon^3 + O(\epsilon^4) \quad (2.117)$$

to then differentiate once and twice

$$\epsilon^3 \partial \partial^2 \phi = (m^2 \partial \phi - 6g^2 \phi^2 \partial \phi) \epsilon^3 + O(\epsilon^4) \quad (2.118)$$

$$\epsilon^3 \partial^2 \partial^2 \phi = (m^2 \partial^2 \phi - 12g^2 \phi (\partial \phi)^2 - 6g^2 \phi^2 \partial^2 \phi) \epsilon^3 + O(\epsilon^4) . \quad (2.119)$$

We eliminate the $\epsilon^3 \partial^2 \phi$ dependence in (2.119) with (2.117) and get

$$\epsilon^3 \partial^2 \partial^2 \phi = (m^4 \phi - 8g^2 m^2 \phi^3 - 12g^2 \phi (\partial \phi)^2 + 12g^4 \phi^5) \epsilon^3 + O(\epsilon^4) \quad (2.120)$$

that, together with (2.118) and (2.117), identifies $\mathfrak{Z}_{4,0}$, $\mathfrak{Z}_{3,0}$ and $\mathfrak{Z}_{2,0}$ respectively.

The only terms left in \mathcal{L} that have to be constrained are those $\propto \epsilon^2 \partial^2 \phi$, thus we have to restart from the primary constraint and multiply it by ϵ . The result is

$$\epsilon^2 \partial^2 \phi = [m^2 \phi - 2g^2 \phi^3 + (4g^2 \phi (\partial \phi)^2 - 4g^2 \phi^2 \partial^2 \phi) \epsilon] \epsilon^2 + O(\epsilon^4) \quad (2.121)$$

which, once used (2.117), become

$$\epsilon^2 \partial^2 \phi = [m^2 \phi - 2g^2 \phi^3 + (4g^2 \phi (\partial \phi)^2 - 4g^2 m^2 \phi^3 + 8g^4 \phi^5) \epsilon] \epsilon^2 + O(\epsilon^4) . \quad (2.122)$$

The four constraints (2.117), (2.118), (2.120) and (2.122) are all what we need to eliminate the unphysical degrees of freedom in \mathcal{L} : once plugged

them we obtain the **reduced effective Lagrangian**

$$\begin{aligned}
\mathcal{L}_R = & -\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \frac{1}{2}g^2\phi^4 - 2\epsilon g^2\phi^2(\partial\phi)^2 \\
& + 2\epsilon^2 g^2 [(\partial\phi)^4 + m^4\phi^4 + 2m^2\phi^2(\partial\phi)^2 \\
& \quad - 4g^2\phi^4(\partial\phi)^2 - 4g^2m^2\phi^6 + 4g^4\phi^8] \\
& - 2\epsilon^3 g^2 [4m^2(\partial\phi)^4 - 4m^4\phi^2(\partial\phi)^2 + 8g^2m^4\phi^6 \\
& \quad + 12g^2m^2\phi^4(\partial\phi)^2 - 44g^2\phi^2(\partial\phi)^4 - 32g^4m^2\phi^8 \\
& \quad - 72g^4\phi^6(\partial\phi)^2 + 32g^6\phi^{10}] + O(\epsilon^4) , \tag{2.123}
\end{aligned}$$

that only depends on ϕ and $\partial\phi$.

To introduce the next application to Supersymmetry we want to group terms in this Lagrangian in a more convenient form. We can indeed observe that here the JLM procedure led simply to:

- a non-standard kinetic term

$$-\frac{1}{2}f(\phi)(\partial\phi)^2 \tag{2.124}$$

with

$$\begin{aligned}
f(\phi) = & 1 + 4\epsilon g^2\phi^2 - 8\epsilon^2 (m^2g^2\phi^2 - 2g^4\phi^4) \\
& - 16\epsilon^3 (g^2m^4\phi^2 - 3g^4m^2\phi^4 + 18g^6\phi^6) \tag{2.125}
\end{aligned}$$

- a squared-kinetic term²⁰

$$h(\phi)(\partial\phi)^4 \tag{2.126}$$

with

$$h(\phi) = 2\epsilon^2 g^2 - 8\epsilon^3 (g^2m^2 - 11g^4\phi^2) \tag{2.127}$$

- a modified potential

$$\begin{aligned}
V_{new}(\phi) = & \overbrace{-\frac{1}{2}m^2\phi^2 + \frac{1}{2}g^2\phi^4}^{V_{old}(\phi)} + 2\epsilon^2 (g^2m^4\phi^4 - 4g^4m^2\phi^6 + 4g^6\phi^8) \\
& - 16\epsilon^3 (g^4m^4\phi^6 - 4g^6m^2\phi^8 + 4g^8\phi^{10}) . \tag{2.128}
\end{aligned}$$

Of course the explicit expression for these terms depends on the original 2-real-field model but this structure is similar to the one we will find in the SUSY case.

²⁰We have to remember that this kind of terms are usually generated by an effective field theory treatment, when a more general theory is approximated to a lowest order.

It is worth noticing that we are substituting repeatedly equations of motion into the Lagrangian in this Classical Field Theory framework: even if this method is not uncommon in Field Theory – it is often used to integrate out undesired fields – its application is not without subtleties. Jaén, Llosa and Molina themselves describe a method that is different, also condemning the substitution into the Lagrangian we performed so far.

One could argue that the fact that this habit is of common use does not justify it, and we agree: in general the EOM must not be used to convert the Lagrangian. By the way, in Field Theory, one could perform a field redefinition that result equivalent to substitute the equations of motion for the field in question.

Following the path of [6], we briefly review such a trick that allow us to close this chapter peacefully.

Let us consider the Lagrangian (2.102) for just one field:

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{2}(\partial\varphi)^2 + \sum_{k=0}^n \epsilon^k \mathfrak{Y}_k(\varphi, \partial\varphi, \dots, \partial^{(k)}\varphi) + O(\epsilon^{n+1}) \\
&= \underbrace{-\frac{1}{2}(\partial\varphi)^2 + \mathfrak{Y}_0(\varphi)}_{\mathcal{L}_0} + \underbrace{\epsilon \mathfrak{Y}_1(\varphi, \partial\varphi) + \epsilon^2 \mathfrak{Y}_2(\varphi, \partial\varphi, \partial^{(2)}\varphi) + \dots}_{\mathcal{L}_1} \\
&\quad + \underbrace{\epsilon^n \mathfrak{Y}_n(\varphi, \partial\varphi, \dots, \partial^{(n)}\varphi) + O(\epsilon^{n+1})}_{\mathcal{L}_2}. \tag{2.129}
\end{aligned}$$

We write²¹

$$\mathcal{L} = \mathcal{L}_{n-1} + \epsilon^n T_n \partial^2 \varphi + \epsilon^n \tilde{\mathfrak{Y}}_n + O(\epsilon^{n+1}) \tag{2.130}$$

with $\tilde{\mathfrak{Y}}_n \equiv \mathfrak{Y}_n - T_n \partial^2 \varphi$ and T_n arbitrary function of φ and its derivatives. It is now possible to perform the set of transformations on the field configuration space

$$\begin{aligned}
\varphi &\longmapsto \varphi - \epsilon^n T_n \\
\partial\varphi &\longmapsto \partial\varphi - \epsilon^n \partial T_n \\
&\vdots \\
\partial^{(n-1)}\varphi &\longmapsto \partial^{(n-1)}\varphi - \epsilon^n \partial^{(n-1)} T_n \tag{2.131}
\end{aligned}$$

²¹Notice that from a term like $f(\varphi, \partial\varphi, \dots) \partial^{(j)}\varphi$ can be always extracted a term $\propto \partial^2 \varphi$ integrating repeatedly by parts. The only case in which it is impossible is when $f = \text{const}$, but in that case we'd just have an eliminable total derivative.

for which the Lagrangian transforms as

$$\begin{aligned}
\mathcal{L} &\longmapsto \mathcal{L}_{n-1} + \epsilon^n T_n \partial^2 \varphi + \epsilon^n \tilde{\mathfrak{Y}}_n - \epsilon^n T_n \left(\frac{\partial \mathcal{L}_0}{\partial \varphi} - \partial_\alpha \frac{\partial \mathcal{L}_0}{\partial (\partial_\alpha \varphi)} \right) + O(\epsilon^{n+1}) \\
&= \mathcal{L}_{n-1} + \epsilon^n T_n \partial^2 \varphi + \epsilon^n \tilde{\mathfrak{Y}}_n - \epsilon^n T_n \left(\frac{\partial \mathfrak{Y}_0}{\partial \varphi} + \partial^2 \varphi \right) + O(\epsilon^{n+1}) \\
&= \mathcal{L}_{n-1} + \epsilon^n \tilde{\mathfrak{Y}}_n - \epsilon^n T_n \frac{\partial \mathfrak{Y}_0}{\partial \varphi} + O(\epsilon^{n+1}) .
\end{aligned} \tag{2.132}$$

We removed an undesired $\epsilon^n \partial^2 \varphi$ piece but a lot of more unwanted terms still remain in $\tilde{\mathfrak{Y}}_n$ and T_n . To get rid of them, we integrate by parts to make other $\epsilon^n \partial^2 \varphi$ terms emerge and again we redefine the field – in the same way – to remove them. This *first* step leads to

$$\mathcal{L} = \mathcal{L}_{n-1} + \epsilon^n \tilde{\mathfrak{Z}}_n(\varphi, \partial \varphi) : \tag{2.133}$$

we only have removed the higher derivatives from the last term.

Notice that, even if better motivated, this step did the same operation performed by the JLM algorithm when removing pieces $\propto \epsilon^n$. In fact, removing just the second derivative (moreover one by one, since a term like $(\partial^2 \varphi)^2$ must be treated twice with this method, including in T a $\partial^2 \varphi$ in the first move) and then integrating by parts to find other $\partial^2 \varphi$ terms, is equivalent to taking derivatives of the 0-th order EOM and substituting all at once.

Moreover, as we emphasised, this was the first – and easier – step.

Now we have to look at the $n-2$ order, so we take (2.133) and we expand it as

$$\mathcal{L} = \mathcal{L}_{n-2} + \epsilon^{n-1} \mathfrak{Y}_{n-1} + \epsilon^n \tilde{\mathfrak{Z}}_n . \tag{2.134}$$

As before, we write it as

$$\mathcal{L} = \mathcal{L}_{n-2} + \epsilon^{n-1} T_{n-1} \partial^2 \varphi + \epsilon^{n-1} \tilde{\mathfrak{Y}}_{n-1} + \epsilon^n \tilde{\mathfrak{Z}}_n \tag{2.135}$$

and perform the set of transformations

$$\partial^{(j)} \varphi \longmapsto \partial^{(j)} \varphi - \epsilon^{n-1} \partial^{(j)} T_{n-1} \quad j = 0, 1, \dots, n-2 \tag{2.136}$$

that transform the Lagrangian as

$$\begin{aligned}
\mathcal{L} &\longmapsto \mathcal{L}_{n-2} + \epsilon^{n-1} T_{n-1} \partial^2 \varphi + \epsilon^{n-1} \tilde{\mathfrak{Y}}_{n-1} + \epsilon^n \tilde{\mathfrak{Z}}_n \\
&\quad - \epsilon^{n-1} T_{n-1} \left(\frac{\partial \mathcal{L}_1}{\partial \varphi} - \partial_\alpha \frac{\partial \mathcal{L}_1}{\partial (\partial_\alpha \varphi)} \right) \\
&= \mathcal{L}_{n-2} + \epsilon^{n-1} T_{n-1} \partial^2 \varphi + \epsilon^{n-1} \tilde{\mathfrak{Y}}_{n-1} + \epsilon^n \tilde{\mathfrak{Z}}_n \\
&\quad - \epsilon^{n-1} T_{n-1} \left(\frac{\partial \mathfrak{Y}_0}{\partial \varphi} + \frac{\partial \mathfrak{Y}_1}{\partial \varphi} + \partial^2 \varphi - \epsilon \partial_\alpha \frac{\partial \mathfrak{Y}_1}{\partial (\partial_\alpha \varphi)} \right) \\
&= \mathcal{L}_{n-2} + \epsilon^{n-1} \tilde{\mathfrak{Y}}_{n-1} + \epsilon^n \tilde{\mathfrak{Z}}_n - \epsilon^{n-1} T_{n-1} \frac{\partial \mathfrak{Y}_0}{\partial \varphi} - \epsilon^{n-1} T_{n-1} \frac{\partial \mathfrak{Y}_1}{\partial \varphi} \\
&\quad + \epsilon^n T_{n-1} \partial_\alpha \frac{\partial \mathfrak{Y}_1}{\partial (\partial_\alpha \varphi)}. \tag{2.137}
\end{aligned}$$

Like in the first step, here we still potentially have a lot of undesired ϵ^{n-1} -higher-derivatives terms that will be cured iterating the integration by parts and the field transformation, but this is no longer our only problem: another ϵ^n term arose and one cannot be sure of its health, indeed terms

$$\propto \epsilon^n \partial^2 \varphi \subset \epsilon^n T_{n-1} \partial_\alpha \frac{\partial \mathfrak{Y}_1}{\partial (\partial_\alpha \varphi)}$$

in general.

Thus we already have to stop to re-treat the $\epsilon^n \partial^2 \varphi$ before taking care of ϵ^{n-1} pieces, extending even more the process that will last until the final higher derivative has been removed.

In addition, notice how in these first two steps the equations of motions that appear multiplied by the ϵ power and the T function are very simple: as the power of ϵ lowers throughout the procedure those become more and more complex. In general we'll have

$$\epsilon^{n-j} T_{n-j} \left(\sum_{k=0}^j (-1)^k \partial_\alpha \partial_\beta \cdots \partial_\kappa \frac{\partial \mathcal{L}_j}{\partial (\partial_\alpha \partial_\beta \cdots \partial_\kappa \varphi)} \right) \tag{2.138}$$

for $j \geq 1$.

The equivalence of these methods is crystal clear for the $n = 1$ case, which is the one of interest for the next application and the one exposed in [6] (to which one can refer for greater clarity).

Chapter 3

Higher-Derivatives in SUSY

This chapter is going to apply the JLM procedure to a (quite) general Higher-Derivative Supersymmetric Chiral Lagrangian with the aim of curing every extra degree of freedom. The latter will come both from higher-derivative terms of the scalar component and derivatives of the auxiliary field.

This chapter is structured as follows: firstly we'll present the notation and the conventions we are going to use in a preliminary section, then we'll dedicate a section to a class of higher-derivative operators in SUSY, and finally we'll see an application of the method aimed at recovering the DBI action in a supersymmetric theory.

3.1 Notation and Conventions

For both, notation and conventions, we will follow [4].

We label with latin beginning-alphabet letters the Lorentz indices. The metric is the mostly-plus one, as already anticipated,

$$\eta_{ab} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} = \eta^{ab} \quad (3.1)$$

and it is always the one used to raise, lower or contract two Lorentz indices.

Spinorial indices are instead labelled by greek beginning-alphabet letters, dotted or undotted for the conjugate or the fundamental representation.

To raise, lower or contract two spinorial indices the antisymmetric tensors

$$\epsilon^{\alpha\beta} = \epsilon^{\dot{\alpha}\dot{\beta}} = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix} = -\epsilon_{\alpha\beta} = -\epsilon_{\dot{\alpha}\dot{\beta}} \quad (3.2)$$

are used.

For the undotted and dotted indices there are the opposite contraction's conventions

$$\begin{array}{ccc}
 \alpha & & \dot{\alpha} \\
 & \searrow & \nearrow \\
 & \alpha & \dot{\alpha}
 \end{array} \quad (3.3)$$

and an inversion in this order implies the appearance of a minus sign.

Grassmann variables have to be introduced, as usual, in a SUSY context. Here we report some relations that involve these variables and the generalized Pauli matrices¹, which are often used in computations with superfields

$$\theta^\alpha \theta^\beta = -\frac{1}{2} \epsilon^{\alpha\beta} \theta\theta \quad , \quad \bar{\theta}^{\dot{\alpha}} \bar{\theta}^{\dot{\beta}} = \frac{1}{2} \epsilon^{\dot{\alpha}\dot{\beta}} \bar{\theta}\bar{\theta} \quad (3.4)$$

$$\theta \sigma^a \bar{\theta} \theta \sigma^b \bar{\theta} = -\frac{1}{2} \theta\theta \bar{\theta}\bar{\theta} \eta^{ab} \quad (3.5)$$

$$Tr\{\sigma^a \bar{\sigma}^b\} \equiv (\sigma^a)_{\alpha\dot{\alpha}} (\bar{\sigma}^b)^{\dot{\alpha}\alpha} = -2\eta^{ab} \quad . \quad (3.6)$$

Last but not least we write down the superspace derivatives²

$$D_A \equiv (\partial_a, D_\alpha, \bar{D}^{\dot{\alpha}}) \quad , \quad (3.7)$$

where

$$\begin{aligned}
 D_\alpha &= + \frac{\partial}{\partial \theta^\alpha} + i(\sigma^a \bar{\theta})_\alpha \frac{\partial}{\partial x^a} \\
 \bar{D}_{\dot{\alpha}} &= - \frac{\partial}{\partial \bar{\theta}^{\dot{\alpha}}} - i(\theta \sigma^a)_{\dot{\alpha}} \frac{\partial}{\partial x^a}
 \end{aligned} \quad (3.8)$$

are the SUSY covariant derivatives, and

$$\{D_\alpha, \bar{D}_{\dot{\alpha}}\} = -2i(\sigma^a)_{\alpha\dot{\alpha}} \partial_a \quad (3.9)$$

$$\{D_\alpha, D_\beta\} = \{\bar{D}_{\dot{\alpha}}, \bar{D}_{\dot{\beta}}\} = [D_\alpha, \partial_b] = [\bar{D}_{\dot{\alpha}}, \partial_b] = 0 \quad (3.10)$$

their graded algebra.

We stress that, even with this compact notation, D^2 is always meant to be appropriately contracted only in its spinorial indices – that is $D^2 \equiv D^\alpha D_\alpha$ – and of course the same is true for \bar{D}^2 for which, however, there is no ambiguity.

¹They are $(\sigma^a)_{\alpha\dot{\alpha}} = (\mathbf{1}, \vec{\sigma})$ and $(\bar{\sigma}^a)^{\dot{\alpha}\alpha} = (-\mathbf{1}, \vec{\sigma}) = \epsilon^{\alpha\beta} \epsilon^{\dot{\alpha}\dot{\beta}} (\sigma^a)_{\beta\dot{\beta}}$.

²This compact notation is still identical to the [4]'s one.

We will only use chiral and antichiral superfields – that is, superfields satisfying the relations

$$\bar{D}_{\dot{\alpha}}\Phi = 0 \quad , \quad D_{\alpha}\bar{\Phi} = 0 \quad (3.11)$$

– but we also want to treat just bosonic degrees of freedom, mainly to facilitate calculations but also because only these seem interesting from the inflationary point of view we will see in the application.

For this reason, as of common use, we simply neglect the fermionic field that naturally would appear in these superfields, practically setting it to zero from the very beginning.

As result, once expanded considering the conventions just written, chiral and antichiral superfields assume the simplified form

$$\Phi = A + i\theta\sigma^a\bar{\theta}\partial_a A + \frac{1}{4}\theta\theta\bar{\theta}\bar{\theta}\partial^2 A + \theta\theta F \quad (3.12)$$

$$\bar{\Phi} = A^* - i\theta\sigma^a\bar{\theta}\partial_a A^* + \frac{1}{4}\theta\theta\bar{\theta}\bar{\theta}\partial^2 A^* + \bar{\theta}\bar{\theta}F^* \quad (3.13)$$

with both $A(x)$ and $F(x)$ complex fields.

Having said that, we are ready to study the possible higher-derivative operators (for a certain mass dimension) in such a supersymmetric theory.

3.2 Higher-Derivatives Treatment in SUSY

We start from the most general Supersymmetric Chiral Lagrangian

$$\begin{aligned} \mathcal{L} = & \int d^4\theta K(\Phi, \bar{\Phi}; D_A\Phi, D_B\bar{\Phi}; D_AD_B\Phi, D_AD_B\bar{\Phi}; \dots) \\ & + \left[\int d^2\theta W(\Phi; \partial_a\Phi; \bar{D}^2\bar{\Phi}; \partial_a\partial_b\Phi; \dots) + h.c. \right] \end{aligned} \quad (3.14)$$

where the Kähler potential K is a real function that includes an arbitrary number of (properly contracted) superspace derivatives of the chiral superfield and of its antichiral partner, whereas the superpotential W is a holomorphic function that contains just chiral superfields, like Φ and its spacetime derivatives but also $\bar{D}^2\bar{\Phi}$ and its spacetime derivatives.

In order to obtain³ $[\mathcal{L}] = 4$, we must have $[K] = 2$ and $[W] = 3$ and then every operator with a different mass dimension must be multiplied by a

³ $\hbar = 1, c = 1.$

dimensionful constant. In the case of relevant operators such a constant will be multiplied to raise the mass-dimension, whereas in the case of irrelevant operators it will be divided and will assume the role of our perturbative expansion parameter ϵ .

To have a clearer discussion we are going to classify the components of K and W in terms of their mass dimension, often choosing particular cases to adapt to our next application. A more general treatment is left for a future work.

Before we choose a particular form for our Kähler potential and for our superpotential, we want to better study the way they depend on their (very general) arguments (see (3.14)). The path is similar to the one followed in [7], the only difference is that we do not discard spacetime derivatives of superfields imposing the condition $\partial_a \Phi = \partial_a \bar{\Phi} = 0$: this was aimed at rejecting every purely kinetic higher-derivative operator⁴. The details on the process of simplification can be found in Appendix B.

The Kähler potential turns out to depend on

$$\begin{aligned}
K &= K(\Phi, \bar{\Phi}; D_\alpha \Phi, \bar{D}_{\dot{\alpha}} \bar{\Phi}; D^2 \Phi, \bar{D}^2 \bar{\Phi}; \\
&\quad \partial_a \Phi, \partial_a \bar{\Phi}; D_\alpha \partial_a \Phi, \bar{D}_{\dot{\alpha}} \partial_a \bar{\Phi}; D^2 \partial_a \Phi, \bar{D}^2 \partial_a \bar{\Phi}; \dots \\
&\quad \partial^{(n)} \Phi, \partial^{(n)} \bar{\Phi}; D_\alpha \partial^{(n)} \Phi, \bar{D}_{\dot{\alpha}} \partial^{(n)} \bar{\Phi}; D^2 \partial^{(n)} \Phi, \bar{D}^2 \partial^{(n)} \bar{\Phi}; \dots) \\
&\equiv K(X(i), \bar{X}(j); Y(k), \bar{Y}(l); Z(p), \bar{Z}(q)) \quad \forall i, j, k, l, p, q \in \mathbb{N}_0 \quad (3.15)
\end{aligned}$$

with

- $X(i) \equiv \partial^{(i)} \Phi$, that increases the mass dimension of the operator it multiplies by a factor of $(i + 1)$;
- $Y(k) \equiv D_\alpha \partial^{(k)} \Phi$, that increases the mass dimension of the operator it multiplies by a factor of $(k + \frac{3}{2})$;
- $Z(p) \equiv D^2 \partial^{(p)} \Phi$, that increases the mass dimension of the operator it multiplies by a factor of $(p + 2)$.

and $\partial^{(0)} \Phi \equiv \Phi$.

Obviously in the last notation hidden Lorentz indices have to be properly contracted – as well as the spinorial ones – in the explicit expression of K .

⁴Such operators bring indeed unavoidable (and undesirable in [7]’s treatment) derivatives of the auxiliary field.

The superpotential's dependence instead results to be of a much easier form⁵:

$$W = W(X(i)) \quad \forall i \in \mathbb{N}_0 . \quad (3.16)$$

We point out that, even if it is completely equivalent, we prefer to use SUSY covariant derivatives rather than spacetime derivatives and then the latter will always be substituted using (3.9) repeatedly.

Before we proceed, let us make some assumptions on the form of the Kähler potential. We choose to not consider relevant operators in K but only marginal or irrelevant ones.

In particular, as marginal part we take the Kähler potential of the most general supersymmetric *renormalizable* Lagrangian involving just chiral superfields,

$$\Phi \bar{\Phi} , \quad (3.17)$$

while for the irrelevant part we decide to restrict the treatment to operators that could reproduce terms like $(\partial A)^2 (\partial A^*)^2$, as these are interesting for the study of higher-derivative scalar field theories of the form $P(X, \phi)$ (with $X \equiv -(\partial\phi)^2/2$ linked to our chiral superfield through $A \equiv \frac{1}{\sqrt{2}}(\phi + i\xi)$)⁶.

In the case of superpotential, our choice is of keeping a quite general form, just requiring that the dependence stops at Φ ($\partial^{(i)}\Phi$ with only $i = 0$).

With these assumptions we are able to apply the JLM method and to see its effects in a supersymmetric situation that will also find application in the next section.

Of course one can in principle apply the method to any kind of Lagrangian

$$\mathcal{L} = \int d^4\theta K(X(i), \bar{X}(j); Y(k), \bar{Y}(l); Z(p), \bar{Z}(q)) + \left[\int d^2\theta W(X(i)) + h.c. \right] \quad (3.18)$$

once explicit forms for K and W are assumed. We decided to treat this particular class of operators because, about this, two schools of thought were dominant: one (cf. [7, 8, 9, 10, 11, 12]) says that only the operators whose component expansion do not contain derivatives of the auxiliary field F or undesired higher-derivatives in A have to be considered in building supersymmetric $P(X, \phi)$ theories; whereas the other (cf. [13, 14]) states that one can consider every such operator provided that each F -derivative or too-high derivative of A is neglected in an EFT treatment.

Both paths have cons: the first leads, in practice, to consider just a particular

⁵Once considered a particular form for W . For details see Appendix B.

⁶See [8].

operator in the set, discarding the rest and significantly restricting the freedom we would otherwise have in constructing new theories⁷; while the second gives not a satisfactory explanation for rejecting the (clearly) unphysical and unwanted terms.

In a certain way our method hybridizes the two approaches since it starts by considering all the operators – as the second one does – to then obtain an expression with only physical terms – as the first one has – by consistently treating every undesired term in an EFT approach.

So, to recover a term like $(\partial A)^2(\partial A^*)^2$ we have to introduce two more fields and two more spacetime derivatives with respect to the ones we count in the component expansion⁸

$$\Phi\bar{\Phi} |_{\theta\theta\bar{\theta}\bar{\theta}} = -|\partial A|^2 + |F|^2 . \quad (3.19)$$

From the anticommutation relation (3.9) we deduce that we have to consider two chiral superfields Φ , their two antichiral correspondives $\bar{\Phi}$, and four SUSY covariant derivatives, two D 's and two \bar{D} 's⁹.

As stated in [8], there are only 8 operators that can be assembled with these elements, they are

$$\mathcal{O}_1 = |\Phi|^2 D^2 \Phi \bar{D}^2 \bar{\Phi} \quad (3.20)$$

$$\mathcal{O}_2 = \Phi D^2 \Phi (\bar{D}\bar{\Phi})^2 \quad (3.21)$$

$$\bar{\mathcal{O}}_2 = \bar{\Phi} \bar{D}^2 \bar{\Phi} (D\Phi)^2 \quad (3.22)$$

$$\mathcal{O}_3 = |\Phi|^2 D \bar{D} \bar{\Phi} \bar{D} D \Phi \quad (3.23)$$

$$\mathcal{O}_4 = \Phi^2 D \bar{D} \bar{\Phi} D \bar{D} \bar{\Phi} \quad (3.24)$$

$$\bar{\mathcal{O}}_4 = \bar{\Phi}^2 \bar{D} D \Phi \bar{D} D \Phi \quad (3.25)$$

$$\mathcal{O}_5 = \Phi D \Phi \bar{D} \bar{\Phi} D \bar{D} \bar{\Phi} \quad (3.26)$$

$$\bar{\mathcal{O}}_5 = \bar{\Phi} \bar{D} \bar{\Phi} D \Phi \bar{D} D \Phi \quad (3.27)$$

although not all are "independent": the last four can be related to the first four using integration by parts¹⁰.

⁷A proof of this fact is the unavoidable correction to the potential one obtains, reason that pushes to consider also the operators containing unphysical degrees of freedom in [13].

⁸We use the notation $|\partial A|^2 \equiv \partial_m A \partial^m A^*$.

⁹Indeed in $(\partial A)^2(\partial A^*)^2$ we have two A 's that must come from two Φ 's, two A^* 's that must come from two $\bar{\Phi}$'s, and two extra derivatives that can come from two D 's and two \bar{D} 's through the anticommutation relations (3.9).

¹⁰Using this linear dependence (that we are also going to analyse better) we will ignore \mathcal{O}_5 and $\bar{\mathcal{O}}_5$ from now on.

Another operator deserves mentioning, even if also this can be expressed as a linear combination of the previous ones. It is

$$\mathcal{O}_0 = D\Phi D\Phi \bar{D}\bar{\Phi} \bar{D}\bar{\Phi} \quad (3.28)$$

and is of great importance for its explicit expression. It is indeed quite special since it has no ∂F terms and, considering only bosonic fields, has just the $\theta\theta\bar{\theta}\bar{\theta}$ component.

These operators have to be added to the Kähler potential, that now is only composed by $\Phi\bar{\Phi}$, of mass dimension 2. They all have mass dimension 6, and then a dimensionful parameter of mass dimension -4 is required: let us call it M_s^{-4} . It is straightforward, for the field component expansion we are going to write, to choose $\epsilon = M_s^{-2}$, similarly to what we did in the previous Effective Field Theory application.

As we are adding these terms to K , just the $\theta\theta\bar{\theta}\bar{\theta}$ component of these operators survives in the Lagrangian.

They are (up to total derivatives)

$$\mathcal{O}_0|_{\theta\theta\bar{\theta}\bar{\theta}} = 16(\partial A)^2(\partial A^*)^2 - 32|F|^2|\partial A|^2 + 16|F|^4 \quad (3.29)$$

$$\begin{aligned} \mathcal{O}_1|_{\theta\theta\bar{\theta}\bar{\theta}} = & 16|F|^4 + 16|A|^2|\partial^2 A|^2 + 8(|A|^2 F \partial^2 F^* + |A|^2 F^* \partial^2 F) \\ & + 24(|F|^2 A \partial^2 A^* + |F|^2 A^* \partial^2 A) \\ & + 16(AF^* \partial A^* \cdot \partial F + A^* F \partial A \cdot \partial F^*) \end{aligned} \quad (3.30)$$

$$\begin{aligned} \mathcal{O}_2|_{\theta\theta\bar{\theta}\bar{\theta}} = & -16|F|^4 - 32|F|^2 A \partial^2 A^* - 32AF^* \partial A^* \cdot \partial F \\ & + 16A \partial^2 A (\partial A^*)^2 \end{aligned} \quad (3.31)$$

$$\begin{aligned} \bar{\mathcal{O}}_2|_{\theta\theta\bar{\theta}\bar{\theta}} = & -16|F|^4 - 32|F|^2 A^* \partial^2 A - 32A^* F \partial A \cdot \partial F^* \\ & + 16A^* \partial^2 A^* (\partial A)^2 \end{aligned} \quad (3.32)$$

$$\begin{aligned} \mathcal{O}_3|_{\theta\theta\bar{\theta}\bar{\theta}} = & 8(\partial A)^2(\partial A^*)^2 - 16|F|^2|\partial A|^2 + 8|A|^2|\partial^2 A|^2 \\ & + 4(|A|^2 F \partial^2 F^* + |A|^2 F^* \partial^2 F) \\ & - 4(|F|^2 A \partial^2 A^* + |F|^2 A^* \partial^2 A) \\ & - 8(AF^* \partial A^* \cdot \partial F + A^* F \partial A \cdot \partial F^*) \\ & + 8(A \partial^2 A (\partial A^*)^2 + A^* \partial^2 A^* (\partial A)^2) \end{aligned} \quad (3.33)$$

$$\begin{aligned} \mathcal{O}_4|_{\theta\theta\bar{\theta}\bar{\theta}} = & -16(\partial A)^2(\partial A^*)^2 + 32|F|^2|\partial A|^2 + 32|F|^2 A \partial^2 A^* \\ & + 32AF^* \partial A^* \cdot \partial F - 16A \partial^2 A (\partial A^*)^2 \end{aligned} \quad (3.34)$$

$$\begin{aligned} \bar{\mathcal{O}}_4|_{\theta\theta\bar{\theta}\bar{\theta}} = & -16(\partial A)^2(\partial A^*)^2 + 32|F|^2|\partial A|^2 + 32|F|^2 A^* \partial^2 A \\ & + 32A^* F \partial A \cdot \partial F^* - 16A^* \partial^2 A^* (\partial A)^2 \end{aligned} \quad (3.35)$$

where these particular components have been chosen for convenience, but integrating by parts also other terms can be obtained and, of course, used throughout the procedure.

To work better with these operators it is useful to define

$$A \equiv (\partial A)^2 (\partial A^*)^2 \quad (3.36)$$

$$B \equiv |F|^2 |\partial A|^2 \quad (3.37)$$

$$C \equiv |F|^4 \quad (3.38)$$

$$D \equiv |A|^2 |\partial^2 A|^2 \quad (3.39)$$

$$E \equiv |A|^2 F \partial^2 F^* + |A|^2 F^* \partial^2 F \quad (3.40)$$

$$F \equiv |F|^2 A \partial^2 A^* + |F|^2 A^* \partial^2 A \quad (3.41)$$

$$G \equiv A F^* \partial A^* \cdot \partial F + A^* F \partial A \cdot \partial F^* \quad (3.42)$$

$$H \equiv A \partial^2 A (\partial A^*)^2 + A^* \partial^2 A^* (\partial A)^2 \quad (3.43)$$

which can be used to write

$$\tilde{\mathcal{O}}_0 \equiv \frac{\mathcal{O}_0}{16} \Bigg|_{\theta\theta\bar{\theta}\bar{\theta}} = A - 2B + C \quad (3.44)$$

$$\tilde{\mathcal{O}}_1 \equiv \frac{\mathcal{O}_1}{8} \Bigg|_{\theta\theta\bar{\theta}\bar{\theta}} = 2C + 2D + E + 3F + 2G \quad (3.45)$$

$$\tilde{\mathcal{O}}_3 \equiv \frac{\mathcal{O}_3}{4} \Bigg|_{\theta\theta\bar{\theta}\bar{\theta}} = 2A - 4B + 2D + E - F - 2G + 2H \quad (3.46)$$

$$\tilde{\mathcal{O}}_2 \equiv \frac{\mathcal{O}_2 + \bar{\mathcal{O}}_2}{16} \Bigg|_{\theta\theta\bar{\theta}\bar{\theta}} = -2C - 2F - 2G + H \quad (3.47)$$

$$\tilde{\mathcal{O}}_4 \equiv \frac{\mathcal{O}_4 + \bar{\mathcal{O}}_4}{16} \Bigg|_{\theta\theta\bar{\theta}\bar{\theta}} = -2A + 4B + 2F + 2G - H \quad (3.48)$$

and thus to check easily the linear dependence with the Gauss algorithm. In order to use it, we consider each term (3.36)-(3.43) as an independent direction with the aim of writing the operators (3.44)-(3.48) as vectors:

$$\begin{array}{r} \\ \tilde{\mathcal{O}}_0 \\ \tilde{\mathcal{O}}_1 \\ \tilde{\mathcal{O}}_2 \\ \tilde{\mathcal{O}}_3 \\ \tilde{\mathcal{O}}_4 \end{array} = \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{cccccccc} A & B & C & D & E & F & G & H \\ \left(\begin{array}{cccccccc} 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 1 & 3 & 2 & 0 \\ 0 & 0 & -2 & 0 & 0 & -2 & -2 & 1 \\ 2 & -4 & 0 & 2 & 1 & -1 & -2 & 2 \\ -2 & 4 & 0 & 0 & 0 & 2 & 2 & -1 \end{array} \right) \end{array}$$

From here we can notice that some components appear just in fixed combinations. For this reason we define

$$\Lambda \equiv A - 2B \quad , \quad \Delta \equiv 2D + E \quad (3.49)$$

to simplify the preceding scheme to

$$\begin{array}{rcl}
& & \Lambda \quad C \quad \Delta \quad F \quad G \quad H \\
\tilde{\mathcal{O}}_0 & = & (\quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad) \\
\tilde{\mathcal{O}}_1 & = & (\quad 0 \quad 2 \quad 1 \quad 3 \quad 2 \quad 0 \quad) \\
\tilde{\mathcal{O}}_2 & = & (\quad 0 \quad -2 \quad 0 \quad -2 \quad -2 \quad 1 \quad) \\
\tilde{\mathcal{O}}_3 & = & (\quad 2 \quad 0 \quad 1 \quad -1 \quad -2 \quad 2 \quad) \\
\tilde{\mathcal{O}}_4 & = & (\quad -2 \quad 0 \quad 0 \quad 2 \quad 2 \quad -1 \quad)
\end{array}$$

Now we can apply the Gauss algorithm to these five "vectors"¹¹

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 3 & 2 & 0 \\ 0 & -2 & 0 & -2 & -2 & 1 \\ 2 & 0 & 1 & -1 & -2 & 2 \\ -2 & 0 & 0 & 2 & 2 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & 2 & -1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

finding that only three of them are linearly independent¹². In particular, the operators chosen – corresponding to the three non-null rows above – are

$$\tilde{\mathcal{O}}_0, -\tilde{\mathcal{O}}_2, \tilde{\mathcal{O}}_1 + \tilde{\mathcal{O}}_2.$$

The most general Lagrangian will thus assume the form

$$\begin{aligned}
\mathcal{L} &= |\Phi|^2 |_{\theta\theta\bar{\theta}\bar{\theta}} + (W(\Phi) |_{\theta\theta} + h.c.) + \epsilon^2 \left\{ \alpha \tilde{\mathcal{O}}_0 + \beta (-\tilde{\mathcal{O}}_2) + \gamma (\tilde{\mathcal{O}}_1 + \tilde{\mathcal{O}}_2) \right\} \\
&= -|\partial A|^2 + |F|^2 + \frac{\partial W}{\partial A} F + \frac{\partial \bar{W}}{\partial A^*} F^* \\
&\quad + \epsilon^2 \{ \alpha \Lambda + (\alpha + 2\beta) C + \gamma \Delta + (2\beta + \gamma) F + 2\beta G - (\beta - \gamma) H \} \\
&= -|\partial A|^2 + |F|^2 + \frac{\partial W}{\partial A} F + \frac{\partial \bar{W}}{\partial A^*} F^* + \epsilon^2 \{ \alpha (\partial A)^2 (\partial A^*)^2 - 2\alpha |F|^2 |\partial A|^2 \\
&\quad + (\alpha + 2\beta) |F|^4 + 2\gamma |A|^2 |\partial^2 A|^2 + \gamma (|A|^2 F \partial^2 F^* + |A|^2 F^* \partial^2 F) \\
&\quad + (2\beta + \gamma) (|F|^2 A \partial^2 A^* + |F|^2 A^* \partial^2 A) \\
&\quad + 2\beta (A F^* \partial A^* \cdot \partial F + A^* F \partial A \cdot \partial F^*) \\
&\quad - (\beta - \gamma) (A \partial^2 A (\partial A^*)^2 + A^* \partial^2 A^* (\partial A)^2) \} \tag{3.50}
\end{aligned}$$

with α, β, γ linear combination constants.

We almost have the desired Lagrangian we want to work with. The last term we have to consider to apply the JLM procedure comes from the assumption that such a Lagrangian contains the higher-order corrections to the

¹¹The moves are: (4)-2(1), (5)+2(1); (3)+(2), (4)+(2), (5)+(3); (4)-2(3); (2)-(3).

¹²E.g. $\tilde{\mathcal{O}}_0 = -\frac{1}{2}\tilde{\mathcal{O}}_1 - \tilde{\mathcal{O}}_2 + \frac{1}{2}\tilde{\mathcal{O}}_3$.

minimal chiral (3.19) one. It may seem trivial but it's not: we are demanding that (3.50) is not built by simply adding operators – as, instead, we just did – but rather that such a Lagrangian describes the same system with a better accuracy, considering perturbative corrections of order ϵ^2 . In this regard, (3.50) is viewed as an approximation to order ϵ^3 of a putative UV theory. This assumption leads in practice to add the necessary $O(\epsilon^3)$ in \mathcal{L} and allow us to treat this Lagrangian with the methods just reviewed.

Having said that, we have to pay close attention to F . In fact whereas in the 0-th order case it was clearly an auxiliary field, in this ϵ^3 approximation doubts about its nature may arise.

Terms $\propto \partial F$ or $\propto \partial^2 F$ make us indeed question the nature of F (and F^*), and wonder if it is still the auxiliary field.

A hint comes from thinking to the dynamical field, which has two physical degrees of freedom (four in this case since it is complex¹³) at the lowest order. Indeed if for A we have to constrain every derivative higher than one in \mathcal{L} , since they are all fictitious d.o.f.'s, for an auxiliary field as F , that has just one degree of freedom (two in this case since it is complex) at the 0-th order, we should constrain any kind of derivative, including a first derivative. Proof of this is the possibility of performing field redefinitions – as we showed at the end of Chapter 2 – to eliminate the undesired F -derivative terms. Again, these field redefinitions are equivalent to applying the JLM procedure but since they do not have exactly the same form of the ones previously seen we write them down explicitly.

Taking into account the general Lagrangian (3.50), we can notice that it is in the form

$$\mathcal{L} \supset \overbrace{|F|^2 + \frac{\partial W}{\partial A} F + \frac{\partial \bar{W}}{\partial A^*} F^*}_{\mathcal{L}_0} + \epsilon^2 \Upsilon^* \cdot \partial F + \epsilon^2 \Upsilon \cdot \partial F^* + \epsilon^2 T^* \cdot \partial^2 F + \epsilon^2 T \cdot \partial^2 F^* . \quad (3.51)$$

T and Υ are functions of the fields and the latter hides a Lorentz index: they explicitly read as

$$\begin{aligned} T &= \gamma |A|^2 F \\ T^* &= \gamma |A|^2 F^* \\ \Upsilon_a &= 2\beta A^* F \partial_a A \\ \Upsilon_a^* &= 2\beta A F^* \partial_a A^* . \end{aligned}$$

¹³The four d.o.f.'s for a complex dynamical field are the two real fields that compose it and by its first derivative.

With the fields redefinitions¹⁴

$$F \longrightarrow F + \epsilon^2 \partial \Upsilon \quad ; \quad F^* \longrightarrow F^* + \epsilon^2 \partial \Upsilon^* \quad (3.52)$$

\mathcal{L} transforms as

$$\begin{aligned} \mathcal{L} &\longrightarrow \mathcal{L} + \epsilon^2 \partial \Upsilon \frac{\partial \mathcal{L}_0}{\partial F} + \epsilon^2 \partial \Upsilon^* \frac{\partial \mathcal{L}_0}{\partial F^*} + O(\epsilon^3) \\ &= \mathcal{L} + \epsilon^2 \partial \Upsilon \left[F^* + \frac{\partial W}{\partial A} \right] + \epsilon^2 \partial \Upsilon^* \left[F + \frac{\partial \bar{W}}{\partial A^*} \right] + O(\epsilon^3) \\ &= \mathcal{L}_0 - \epsilon^2 \Upsilon^* \cdot \partial \frac{\partial \bar{W}}{\partial A^*} - \epsilon^2 \Upsilon \cdot \partial \frac{\partial W}{\partial A} + \epsilon^2 T^* \cdot \partial^2 F + \epsilon^2 T \cdot \partial^2 F^* + \dots \end{aligned} \quad (3.53)$$

where in the last line integration by parts has been used. We see that this field redefinition has been equivalent to substituing ∂F (extra degree of freedom) with F 's 0th-order EOM differentiated.

Before showing the equivalence between this method and the JLM procedure let's finish this treatment by performing the second field redefinition:

$$F \longrightarrow F - \epsilon^2 \partial^2 T \quad ; \quad F^* \longrightarrow F^* - \epsilon^2 \partial^2 T^* \quad (3.54)$$

This transformation leads to the new Lagrangian

$$\begin{aligned} \mathcal{L} &\longrightarrow \mathcal{L} - \epsilon^2 \partial^2 T \frac{\partial \mathcal{L}_0}{\partial F} - \epsilon^2 \partial^2 T^* \frac{\partial \mathcal{L}_0}{\partial F^*} \\ &= \mathcal{L} - \epsilon^2 \partial^2 T \left[F^* + \frac{\partial W}{\partial A} \right] - \epsilon^2 \partial^2 T^* \left[F + \frac{\partial \bar{W}}{\partial A^*} \right] \\ &= \mathcal{L}_0 - \epsilon^2 \Upsilon^* \cdot \partial \frac{\partial \bar{W}}{\partial A^*} - \epsilon^2 \Upsilon \cdot \partial \frac{\partial W}{\partial A} - \epsilon^2 T^* \partial^2 \frac{\partial \bar{W}}{\partial A^*} - \epsilon^2 T \cdot \partial^2 \frac{\partial W}{\partial A} + \dots \end{aligned} \quad (3.55)$$

again equivalent to substitute $\partial^2 F$ with F 's 0th-order EOM differentiated (twice).

The application of the JLM procedure to F and F^* requires the usual comparison with the Lagrangian in (2.102), so we identify:

- No 0-th order kinetic terms;
- $\mathfrak{Q}_0 = |F|^2 + \frac{\partial W}{\partial A} F + \frac{\partial \bar{W}}{\partial A^*} F^*$;

¹⁴ $\partial \Upsilon \equiv \partial^a \Upsilon_a$.

- $\mathfrak{Y}_1 = 0$;
- $\mathfrak{Y}_2 = -2\alpha|F|^2|\partial A|^2 + (\alpha + 2\beta)|F|^4 + \gamma(|A|^2F\partial^2F^* + |A|^2F^*\partial^2F) + (2\beta + \gamma)(|F|^2A\partial^2A^* + |F|^2A^*\partial^2A) + 2\beta(AF^*\partial A^* \cdot \partial F + A^*F\partial A \cdot \partial F^*)$.

We point out that, in principle, one has to treat F and F^* separately since they are different degrees of freedom. The symmetry between the two in (3.36)-(3.43) – and, more in general, in every Lagrangian term – lead us to treat them together with a sort of compact notation for the \mathfrak{Y}_k 's¹⁵. Fortunately, treating them in parallel does not cause any trouble.

So first we have to find the primary constraints for F and F^* . We then write their equations of motion

$$\frac{\partial \mathcal{L}}{\partial F} - \partial_a \frac{\partial \mathcal{L}}{\partial(\partial_a F)} + \partial_a \partial_b \frac{\partial \mathcal{L}}{\partial(\partial_a \partial_b F)} = 0 \quad (3.56)$$

$$\frac{\partial \mathcal{L}}{\partial F^*} - \partial_a \frac{\partial \mathcal{L}}{\partial(\partial_a F^*)} + \partial_a \partial_b \frac{\partial \mathcal{L}}{\partial(\partial_a \partial_b F^*)} = 0 \quad (3.57)$$

that read as¹⁶

$$\mathfrak{W}_{F,0} + \epsilon^2 \mathfrak{W}_{F,2} = O(\epsilon^3) \quad (3.58)$$

$$\mathfrak{W}_{F^*,0} + \epsilon^2 \mathfrak{W}_{F^*,2} = O(\epsilon^3) \quad (3.59)$$

or, explicitly,

$$\begin{aligned} F^* + \frac{\partial W}{\partial A} + \epsilon^2 \{ & (-2\alpha - 2\beta + 2\gamma)F^*|\partial A|^2 + 2(\alpha + 2\beta)F^*|F|^2 \\ & + 2\gamma|A|^2\partial^2F^* + 2\gamma AF^*\partial^2A^* + (2\beta + 2\gamma)A^*F^*\partial^2A \\ & + (2\beta + 2\gamma)A^*\partial A \cdot \partial F^* - (2\beta - 2\gamma)A\partial A^* \cdot \partial F^* \} = O(\epsilon^3) \end{aligned} \quad (3.60)$$

$$\begin{aligned} F + \frac{\partial \bar{W}}{\partial A^*} + \epsilon^2 \{ & (-2\alpha - 2\beta + 2\gamma)F|\partial A|^2 + 2(\alpha + 2\beta)F|F|^2 \\ & + 2\gamma|A|^2\partial^2F + 2\gamma A^*F\partial^2A + (2\beta + 2\gamma)AF\partial^2A^* \\ & + (2\beta + 2\gamma)A\partial A^* \cdot \partial F - (2\beta - 2\gamma)A^*\partial A \cdot \partial F \} = O(\epsilon^3) . \end{aligned} \quad (3.61)$$

¹⁵A further observation can be made. Keeping the auxiliary field complex is impossible to separate $\mathfrak{Y}_{F,0}$ and $\mathfrak{Y}_{F^*,0}$: they are linked together by the term $|F|^2$. To obtain exactly the form (2.102) we have to write F in terms of two real fields (as we will do for A) and not doing that should force us to take care of F and, once finished, of F^* (or viceversa). As said above, fortunately, treating them in parallel does not cause any trouble.

¹⁶Contrary to $\mathfrak{Y}_{F,0}$ and $\mathfrak{Y}_{F^*,0}$ that were impossible to write separately (see the previous note), $\mathfrak{W}_{F,0/2}$ and $\mathfrak{W}_{F^*,0/2}$ have sense and can be written, since the subscript specifies the field derivative that automatically decouple the term $|F|^2$.

At this point we should clarify a subtlety about the notation. We started using $\frac{\partial W}{\partial A}$ as a short notation for

$$\left. \frac{dW}{d\Phi} \right|_{\Phi=A}, \quad (3.62)$$

in fact it is equivalent to differentiate W with respect to A once substituted the latter every time Φ appears into the superpotential: $\partial W(A)/\partial A = dW(A)/dA$. The same considerations can be done for its hermitian conjugate: we'll have

$$\left. \frac{d\bar{W}}{d\bar{\Phi}} \right|_{\bar{\Phi}=A^*} = \left. \frac{dW(\bar{\Phi})}{d\bar{\Phi}} \right|_{\bar{\Phi}=A^*} \quad (3.63)$$

written as $\partial \bar{W}/\partial A^* \equiv \partial W(A^*)/\partial A^* = dW(A^*)/dA^*$. Having said that it is clear how¹⁷

$$\frac{\partial W(A^*)}{\partial A^*} = \left(\frac{\partial W(A)}{\partial A} \right)^* \quad (3.64)$$

simply because W depends only on Φ and thus its hermitian conjugate will be the same function but with $\bar{\Phi}$ replacing Φ . This still holds when one performs the substitutions $\Phi \rightarrow A$, $\bar{\Phi} \rightarrow A^*$.

The **primary constraints** for F and F^* are obtained by multiplying (3.60) and (3.61) by ϵ as always. They are

$$\epsilon \left[F^* + \left(\frac{\partial W}{\partial A} \right) \right] = O(\epsilon^3) \quad (3.65)$$

$$\epsilon \left[F + \left(\frac{\partial W}{\partial A} \right)^* \right] = O(\epsilon^3). \quad (3.66)$$

From these, we can find the full set of secondary constraints. They indeed are just the equations of motion (3.60) and (3.61) (and their derivatives) with terms depending on the undesired d.o.f.'s, ∂F and $\partial^2 F$, removed.

So first of all we need constraints for

$$\begin{aligned} &\epsilon^2 \partial_a F, \quad \epsilon^2 \partial_a F^* \\ &\epsilon^2 \partial^2 F, \quad \epsilon^2 \partial^2 F^*, \end{aligned}$$

thus

¹⁷From now on the explicit dependence of the superpotential will be omitted: it can be understood from the derivative.

(i) We multiply (3.65) and (3.66) by ϵ

$$\epsilon^2 \left[F^* + \left(\frac{\partial W}{\partial A} \right) \right] = O(\epsilon^3) \quad (3.67)$$

$$\epsilon^2 \left[F + \left(\frac{\partial W}{\partial A} \right)^* \right] = O(\epsilon^3) ; \quad (3.68)$$

(ii) We differentiate (3.67) and (3.68)

$$\epsilon^2 \left[\partial_a F^* + \left(\frac{\partial^2 W}{\partial A^2} \right) \partial_a A \right] = O(\epsilon^3) \quad (3.69)$$

$$\epsilon^2 \left[\partial_a F + \left(\frac{\partial^2 W}{\partial A^2} \right)^* \partial_a A^* \right] = O(\epsilon^3) ; \quad (3.70)$$

(iii) We differentiate (3.69) and (3.70)

$$\epsilon^2 \left[\partial^2 F^* + \left(\frac{\partial^3 W}{\partial A^3} \right) (\partial A)^2 + \left(\frac{\partial^2 W}{\partial A^2} \right) \partial^2 A \right] = O(\epsilon^3) \quad (3.71)$$

$$\epsilon^2 \left[\partial^2 F + \left(\frac{\partial^3 W}{\partial A^3} \right)^* (\partial A^*)^2 + \left(\frac{\partial^2 W}{\partial A^2} \right)^* \partial^2 A^* \right] = O(\epsilon^3) . \quad (3.72)$$

Plugging these last four equations in (3.60) and (3.61) gives us the two healthy equations of motion for F and F^*

$$\begin{aligned} F^* + \frac{\partial W}{\partial A} + \epsilon^2 \left\{ (-2\alpha - 2\beta + 2\gamma) F^* |\partial A|^2 + 2(\alpha + 2\beta) F^* |F|^2 \right. \\ \left. - 2\gamma |A|^2 \left[\frac{\partial^3 W}{\partial A^3} (\partial A)^2 + \frac{\partial^2 W}{\partial A^2} \partial^2 A \right] \right. \\ \left. + 2\gamma A F^* \partial^2 A^* + (2\beta + 2\gamma) A^* F^* \partial^2 A \right. \\ \left. - (2\beta + 2\gamma) A^* \frac{\partial^2 W}{\partial A^2} (\partial A)^2 \right. \\ \left. + (2\beta - 2\gamma) A \frac{\partial^2 W}{\partial A^2} |\partial A|^2 \right\} = O(\epsilon^3) \end{aligned} \quad (3.73)$$

$$\begin{aligned} F + \left(\frac{\partial W}{\partial A} \right)^* + \epsilon^2 \left\{ (-2\alpha - 2\beta + 2\gamma) F |\partial A|^2 + 2(\alpha + 2\beta) F |F|^2 \right. \\ \left. - 2\gamma |A|^2 \left[\left(\frac{\partial^3 W}{\partial A^3} \right)^* (\partial A^*)^2 + \left(\frac{\partial^2 W}{\partial A^2} \right)^* \partial^2 A^* \right] \right. \\ \left. + 2\gamma A^* F \partial^2 A + (2\beta + 2\gamma) A F \partial^2 A^* \right. \\ \left. - (2\beta + 2\gamma) A \left(\frac{\partial^2 W}{\partial A^2} \right)^* (\partial A^*)^2 \right. \\ \left. + (2\beta - 2\gamma) A^* \left(\frac{\partial^2 W}{\partial A^2} \right)^* |\partial A|^2 \right\} = O(\epsilon^3) . \end{aligned} \quad (3.74)$$

Now we can complete the task: we impose the secondary constraints (3.69)-(3.72) on our system plugging them in \mathcal{L} . The result will be exactly the Lagrangian (3.55), that in turn has as equations of motion (for F and F^*) (3.73) and (3.74).

The equivalence between the two methods is then crystal clear. The main advantage of the JLM procedure is that it is easier to apply systematically. The first procedure can be seen as a justification for the second one.

Notice that we left with a Lagrangian containing just F , F^* but not their derivatives: this means that now they are undoubtedly non-propagating auxiliary fields. We now wonder if we can integrate them out by using their EOM's as we were doing in the 0th-order case. We converted every term containing some sort of space-time derivative of the field F into a certain function of the field A and its derivatives, obtaining a Lagrangian density in the form

$$\mathcal{L} \supset f|F|^2 + gF + \bar{g}F^* + k|F|^4, \quad (3.75)$$

with $k = \alpha + 2\beta$ constant and f and g functions of A , A^* , ∂A , ∂A^* , $\partial^2 A$, $\partial^2 A^*$.

We can safely integrate out a field by substituting its equations of motion when the part of the action where it appears is in a Gaussian form. Here we see that this condition is assured when $k = 0$ ¹⁸ and that otherwise there is some ambiguity as to what the correct quantum theory should be (see [9]). Remaining in a context of Classical Field Theory we won't impose $k = 0$ and we will substitute the auxiliary field's equations of motion, as one always does. We will analyze the effect of this choice in our last application to the DBI Action in the next section.

Whereas the replacement of ∂F and $\partial^2 F$ (and their complex conjugates) has been straightforward, the substitution of F and F^* deserves a little observation. Regarding the part of the Lagrangian multiplied by ϵ^2 it is obvious that the effect of inserting (3.73) and (3.74) is the same of using the 0-th order EOM

$$F^* = - \left(\frac{\partial W}{\partial A} \right) \quad , \quad F = - \left(\frac{\partial W}{\partial A} \right)^* \quad , \quad (3.76)$$

¹⁸It can be seen more easily by performing multiple integrations by parts in order to get $f = 1$. Indeed every term multiplied by ϵ^2 in the explicit expression of f contains always some sort of A 's derivative that we can then move to $|F|^2$ to get terms which can be integrated in g or \bar{g} .

but when we have to plug them in the original part

$$|F|^2 + \left(\frac{\partial W}{\partial A}\right) F + \left(\frac{\partial W}{\partial A}\right)^* F^* \quad (3.77)$$

one could think that new ϵ^2 terms arise. Actually, a satisfying cancellation takes place. In order to see it clearly we momentarily define the term between brackets in (3.73) as X , so that the corresponding term in (3.74) will be X^* . The equations of motion for F and F^* then read as

$$F^* = - \left(\frac{\partial W}{\partial A}\right) - \epsilon^2 X + O(\epsilon^3) \quad (3.78)$$

$$F = - \left(\frac{\partial W}{\partial A}\right)^* - \epsilon^2 X^* + O(\epsilon^3) \quad (3.79)$$

and once plugged in (3.77) lead to

$$\begin{aligned} & \underbrace{\left| \frac{\partial W}{\partial A} \right|^2 + \epsilon^2 \left(\frac{\partial W}{\partial A}\right) X^* + \epsilon^2 \left(\frac{\partial W}{\partial A}\right)^* X + O(\epsilon^3)}_{\text{from } |F|^2} \\ & - \underbrace{\left| \frac{\partial W}{\partial A} \right|^2 - \epsilon^2 \left(\frac{\partial W}{\partial A}\right) X^* + O(\epsilon^3)}_{\text{from } (\partial W/\partial A)F} \\ & - \underbrace{\left| \frac{\partial W}{\partial A} \right|^2 - \epsilon^2 \left(\frac{\partial W}{\partial A}\right)^* X + O(\epsilon^3)}_{\text{from } (\partial W/\partial A)^* F^*} = - \left| \frac{\partial W}{\partial A} \right|^2 + O(\epsilon^3) \end{aligned}$$

that is the usual term remaining.

As for the minimal chiral model, we define

$$V(A, A^*) \equiv \left| \frac{\partial W}{\partial A} \right|^2 \quad (3.80)$$

which is the scalar potential of the 0-th order theory.

Even if we could finally be ready to move on and treat the complex scalar field A , one last observation is worthy of interest. Indeed comparing the application of the JLM method to a dynamical field with its version for an auxiliary field, we notice that the most general secondary constraints in the first case are always built to constrain the unphysical degrees of freedom: they are always associated to the second (or higher) derivative of the field by its equation of motion. An auxiliary field, instead, does not have such a

term in its EOM and thus can be eliminated exploiting the approximation to the desired order. In a certain sense the procedure is telling us that the auxiliary field, as any of its derivatives, is unphysical. In fact if we blindly apply it to F we could say that (3.67) and (3.68) are constraints as much as (3.69)-(3.72), with the difference that these last seem to constrain extra degrees of freedom – not expected in a theory described by just one chiral superfield Φ – whereas the former contain only F and F^* . Furthermore, still following the method to the letter, the healthy equations of motion (3.73) and (3.74) (with terms $\propto \epsilon^2 F$ and $\propto \epsilon^2 F^*$ eliminated by (3.67) and (3.68)) should be themselves constraints, in particular the generators of the most complete set of secondary constraints. Their substitution in \mathcal{L} should then result straightforward and not require further justifications.

So, in summary,

- 1) As in the usual "0-th order" case, we want to eliminate F ;
- 2) In order to do that, we can follow two main paths: one involves the removal of the extra degrees of freedom via field redefinitions together with a use of the healthy equations of motion, whereas the other sees the systematic application of the JLM procedure;
- 3) Both have the same core, that is substituting EOM's in \mathcal{L} taking advantage of the $O(\epsilon^3)$ (key of the method developed by Jaén, Llosa and Molina)¹⁹;
- 4) Both give the same result, equivalent to the usual procedure one always performs in order to integrate out a field, that is substituting repeatedly its equations of motion and their derivatives in the Lagrangian until it completely disappears because of the approximation order considered.

Now it is the turn of A . The Lagrangian after having integrated out F becomes

$$\begin{aligned}
\mathcal{L} = & - |\partial A|^2 - V + \epsilon^2 \{ \alpha (\partial A)^2 (\partial A^*)^2 - 2\alpha V |\partial A|^2 \\
& + (\alpha + 2\beta) V^2 + 2\gamma |A|^2 |\partial^2 A|^2 \\
& + \gamma |A|^2 (V_A \partial^2 A + V_{AA} (\partial A)^2 + V_{A^*} \partial^2 A^* + V_{A^*A^*} (\partial A^*)^2) \\
& + (2\beta + \gamma) V (A \partial^2 A^* + A^* \partial^2 A) \\
& + 2\beta (A V_{A^*} (\partial A^*)^2 + A^* V_A (\partial A)^2) \\
& - (\beta - \gamma) (A \partial^2 A (\partial A^*)^2 + A^* \partial^2 A^* (\partial A)^2) \} + O(\epsilon^3) \quad (3.81)
\end{aligned}$$

¹⁹The first path seems to be, nevertheless, better justified, not involving the direct substitution of EOM's into a Lagrangian.

where the notation

$$V_A \equiv \frac{\partial V}{\partial A} \quad (3.82)$$

$$V_{A^*} \equiv \frac{\partial V}{\partial A^*} = \left(\frac{\partial V}{\partial A} \right)^* \quad (3.83)$$

$$V_{AA} \equiv \frac{\partial^2 V}{\partial A^2} \quad (3.84)$$

$$V_{A^*A^*} \equiv \frac{\partial^2 V}{\partial A^{*2}} = \left(\frac{\partial^2 V}{\partial A^2} \right)^* \quad (3.85)$$

have been used once noticed that

$$W_{AA}W_{A^*} = V_A \quad (3.86)$$

$$W_AW_{A^*A^*} = V_{A^*} \quad (3.87)$$

$$W_{AAA}W_{A^*} = V_{AA} \quad (3.88)$$

$$W_AW_{A^*A^*A^*} = V_{A^*A^*} . \quad (3.89)$$

We can make the usual comparison with the Lagrangian in (2.102). We identify:

- The standard kinetic term for a complex scalar field, $-|\partial A|^2$ (completely equivalent to the one in (2.102) if one expands $A = \frac{1}{\sqrt{2}}(\phi + i\xi)$ with ϕ and ξ real fields);
- $\mathfrak{Q}_0 = -V(A, A^*)$;
- $\mathfrak{Q}_1 = 0$;
- $\mathfrak{Q}_2 = \mathfrak{Q}_2(A, A^*; \partial A, \partial A^*; \partial^2 A, \partial^2 A^*)$, including all the terms between brackets.

Following the same steps, we start from the equations of motion for A and A^{*20}

$$\partial^2 A^* - V_A + \epsilon^2 \{ \dots \} = O(\epsilon^3) \quad (3.90)$$

$$\partial^2 A - V_{A^*} + \epsilon^2 \{ \dots \} = O(\epsilon^3) \quad (3.91)$$

thanks to which one can derive the **primary constraints**

$$\epsilon [\partial^2 A^* - V_A] = O(\epsilon^3) \quad (3.92)$$

$$\epsilon [\partial^2 A - V_{A^*}] = O(\epsilon^3) . \quad (3.93)$$

²⁰The content of $\{ \dots \}$ comes from (3.81), in particular from the piece multiplied by ϵ^2 .

Now the only secondary constraint we need is the one necessary for the cancellation of the undesired $\epsilon^2 \partial^2 A$, $\epsilon^2 \partial^2 A^*$ terms in (3.81), obtained by just multiplying (3.92) and (3.93) by ϵ . Notice that the absence of \mathfrak{B}_1 enables us to apply a very simplified form of the JLM method that, in practice, consists in just substituting the 0-th order EOM for A and A^* into \mathcal{L} .

The final result is the **reduced Lagrangian**

$$\begin{aligned} \mathcal{L} = & -|\partial A|^2 - V + \epsilon^2 \{ \alpha (\partial A)^2 (\partial A^*)^2 - 2\alpha V |\partial A|^2 \\ & + (\gamma |A|^2 V_{AA} + (\beta + \gamma) A^* V_A) (\partial A)^2 \\ & + (\gamma |A|^2 V_{A^* A^*} + (\beta + \gamma) A V_{A^*}) (\partial A^*)^2 \\ & + (\alpha + 2\beta) V^2 + 4\gamma |A|^2 |V_A|^2 + (2\beta + \gamma) V (A V_A + A^* V_{A^*}) \} + O(\epsilon^3) \end{aligned} \quad (3.94)$$

that is a Lagrangian depending just on A , A^* and their first derivatives. It contains:

- A non-standard kinetic term

$$-f(A, A^*) |\partial A|^2 + \epsilon^2 \{ (\gamma |A|^2 V_{AA} + (\beta + \gamma) A^* V_A) (\partial A)^2 + c.c. \} \quad (3.95)$$

with

$$f(A, A^*) = 1 + 2\alpha \epsilon^2 V(A, A^*) \quad (3.96)$$

- The desired higher-derivative term

$$\alpha \epsilon^2 (\partial A)^2 (\partial A^*)^2 \quad (3.97)$$

- A modified potential

$$V(A, A^*) + \epsilon^2 \{ (\alpha + 2\beta) V^2 + 4\gamma |A|^2 |V_A|^2 + (2\beta + \gamma) V (A V_A + A^* V_{A^*}) \} \quad (3.98)$$

Each higher-derivative term has been eliminated through the constraints arose naturally from the system, leaving a Lagrangian with no unphysical degrees of freedom. We have recovered the squared-kinetic term demanded and varying the three free coefficients one can obtain different theories.

Of course not every kind of model is reachable tuning α , β and γ , in the next section for example we will have to impose some conditions for the scalar potential in order to obtain the desired theory.

3.3 Application: the DBI Action

In this section our aim is to recover the Dirac-Born-Infeld Action for a D7-brane in a 10-dimensional space. We focus on a compactification where the internal (compact) manifold is a toroidal orbifold. For such a situation we have ([12, 13, 15, 16, 17, 18, 19, 20])

$$S_{DBI} = \int d\mathcal{V}_{\mathcal{M}_{1,3}} M_s^4 \times \left[- \left(1 + \frac{1}{M_s^4} V(\phi, \xi) \right) \sqrt{\det \left(g_{mn} + \frac{1}{M_s^4} (\partial_m \phi \partial_n \phi + \partial_m \xi \partial_n \xi) \right) + 1} \right] \quad (3.99)$$

with ϕ and ξ the two real bosonic fields (brane's transverse coordinates) and M_s the string mass.

Defining

$$\varepsilon \equiv \frac{1}{M_s^4} \quad (3.100)$$

and expanding²¹, one finds

$$\mathcal{L}_{DBI} = - (1 + \varepsilon V) |\partial A|^2 - V(A, A^*) + \frac{1}{2} \varepsilon (\partial A)^2 (\partial A^*)^2 + O(\varepsilon^2) , \quad (3.101)$$

where the two real fields have been assembled in the complex scalar field

$$A \equiv \frac{1}{2} (\phi + i\xi) . \quad (3.102)$$

In what follows we will try to provide a supersymmetric description of the DBI action by adjusting the three free coefficients in (3.94). In other words we will try to identify the linear combination of higher-derivative operators that leads to the DBI action.

First of all we have to impose

$$\varepsilon \triangleq \epsilon^2 \quad (3.103)$$

so that we can immediately recognize the similarities between (3.101) and the first line of (3.94). Comparing we see that

$$\alpha = \frac{1}{2} . \quad (3.104)$$

²¹See Appendix C.

Furthermore we would like for every other term (in (3.94)) to disappear just by fixing the other two constants β and γ . Unfortunately, this does not happen and a correction to the scalar potential always arises, in general.

The only way to match the DBI action is to constrain the potential. Requiring every unwanted term of (3.101) to vanish, indeed, leaves us with the system of partial differential equations for the scalar potential

$$\begin{cases} (\frac{1}{2} + 2\beta) V^2 + 4\gamma|A|^2|V_A|^2 + (2\beta + \gamma) V (AV_A + A^*V_{A^*}) = 0 \\ \gamma|A|^2V_{AA} + (\beta + \gamma)A^*V_A = 0 \\ \gamma|A|^2V_{A^*A^*} + (\beta + \gamma)AV_{A^*} = 0 \end{cases} \quad (3.105)$$

where the first row follows from demanding a vanishing correction to the scalar potential whereas the other two removes the undesired corrections to the kinetic term. We must have a potential satisfying these three equations if we want to obtain the DBI Lagrangian at the desired approximation order.

For a more convenient treatment we now use (3.102) to work with real scalar fields. The system becomes

$$\begin{cases} (\frac{1}{2} + 2\beta) V^2 + \gamma(\phi^2 + \xi^2)(V_\phi^2 + V_\xi^2) + (2\beta + \gamma) V(\phi V_\phi + \xi V_\xi) = 0 \\ \gamma(\phi^2 + \xi^2)(V_{\phi\phi} - V_{\xi\xi}) + (\beta + \gamma)(\phi V_\phi - \xi V_\xi) = 0 \\ 2\gamma(\phi^2 + \xi^2)V_{\phi\xi} + (\beta + \gamma)(\xi V_\phi + \phi V_\xi) = 0 \end{cases} \quad (3.106)$$

with $V = V(\phi, \xi)$.

As in [20], we simplify further assuming that

$$\xi = 0 \quad , \quad V_\xi = 0 \quad (3.107)$$

– that is, suppressing a degree of freedom. The system reduces to

$$\begin{cases} (\frac{1}{2} + 2\beta) V^2 + \gamma\phi^2V_\phi^2 + (2\beta + \gamma) V\phi V_\phi = 0 \\ \gamma\phi^2V_{\phi\phi} + (\beta + \gamma)\phi V_\phi = 0 \end{cases} \quad (3.108)$$

with $V = V(\phi)$.

The second row can now be solved exactly. The result is the monomial²²

$$V = v_0\phi^{-\frac{\beta}{\gamma}} \quad (3.109)$$

that, once plugged in the first equation, gives the relation between β and γ

$$2\beta^2 - 2\gamma\beta - \gamma = 0 \quad . \quad (3.110)$$

²²To be fair an integration constant should be added to V here but, once plugged back in the other equation of the system, it turns out to be zero.

So we found that if the point (β, γ) belongs to the hyperbole (3.110), depicted in Fig. 3.1, and the potential has the form (3.109), then the Lagrangian (3.94) assumes the form (3.101) as desired.

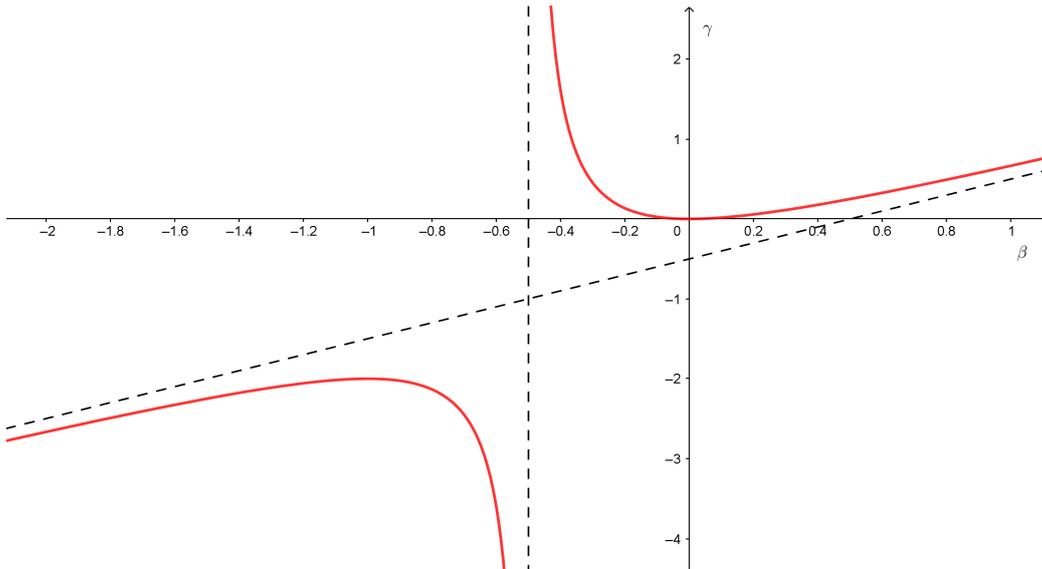


Figure 3.1: The hyperbole $2\beta^2 - 2\gamma\beta - \gamma = 0$.

A more interesting point of view can also be considered. Indeed, since a monomial potential for the inflaton field is often taken into account (cf. [5, 20]), we can change perspective and look at the parametric solution of the system (3.108): we can say that we are interested in a scalar potential of the form

$$V = v_0 \phi^n, \quad (3.111)$$

with $n \in \mathbb{R}^+$, and therefore that we want to find β and γ in terms of n .

Computing, it turns out that

$$\beta = -\frac{1}{2(n+1)} \quad (3.112)$$

$$\gamma = \frac{1}{2n(n+1)}. \quad (3.113)$$

We can observe that

- (i) $n = -\frac{\beta}{\gamma}$, as it must be from the comparison between (3.109) and (3.111);

- (ii) A potential of interest for inflation models – as the monomial one is – solves exactly the system (this is absolutely non-trivial);
- (iii) Since we have suppressed ξ , we now have that

$$A = A^* = \frac{1}{\sqrt{2}}\phi \quad \Rightarrow \quad \phi = 2^{\frac{1}{2}}A = 2^{\frac{1}{2}}A^*$$

and thus we are able to derive a superpotential that can reproduce the desired scalar potential. We first consider that

$$V = v_0\phi^n = \left(v_0^{\frac{1}{2}}\phi^{\frac{n}{2}}\right) \left(v_0^{\frac{1}{2}}\phi^{\frac{n}{2}}\right) = \left(2^{\frac{n}{4}}v_0^{\frac{1}{2}}A^{\frac{n}{2}}\right) \left(2^{\frac{n}{4}}v_0^{\frac{1}{2}}A^{*\frac{n}{2}}\right) \quad (3.114)$$

and then that

$$V = \frac{dW}{dA} \frac{dW}{dA^*} \quad (3.115)$$

to finally get

$$\frac{dW}{dA} = 2^{\frac{n}{4}}v_0^{\frac{1}{2}}A^{\frac{n}{2}}. \quad (3.116)$$

This simple differential equation has the solution

$$W(A) = \frac{2^{\frac{n}{4}+1}v_0^{\frac{1}{2}}}{n+2}A^{\frac{n}{2}+1} \quad (3.117)$$

which allow us to recover the superpotential

$$W(\Phi) = \frac{2^{\frac{n}{4}+1}v_0^{\frac{1}{2}}}{n+2}\Phi^{\frac{n}{2}+1} \quad (3.118)$$

that, once suppressed the second degree of freedom of A^{23} , gives in fact the right potential in ϕ .

To conclude, the combination of operators

$$\frac{1}{2}\tilde{\mathcal{O}}_0 + \frac{1}{2n(n+1)}\tilde{\mathcal{O}}_1 + \frac{1}{2n}\tilde{\mathcal{O}}_2 \quad (3.119)$$

in the Kähler potential and the superpotential (3.118) (with $\xi = 0$) is what we were looking for in order to write the DBI Lagrangian²⁴ in a Supersymmetric setup.

²³The scalar potential generated by such a superpotential would be

$$V(\phi, \xi) = v_0(\phi^2 + \xi^2)^{\frac{n}{2}}.$$

²⁴To be precise, it is the first order approximation of the DBI Lagrangian.

To find a better form for such Lagrangian we decide to use $\tilde{\mathcal{O}}_3$ instead of $\tilde{\mathcal{O}}_2$ in the linear combination, choosing the highest number of operators with the factor $|\Phi|^2$ (it is just an aesthetic choice).

Then, the explicit form for the supersymmetric Lagrangian that allow us to recover the DBI Lagrangian at its first-order expansion is given by

$$\begin{aligned}
K &= |\Phi|^2 \\
&+ \varepsilon \left[\frac{n-1}{32n} D\Phi D\Phi \bar{D}\bar{\Phi} \bar{D}\bar{\Phi} - \frac{n-1}{32n(n+1)} |\Phi|^2 D^2\Phi \bar{D}^2\bar{\Phi} + \frac{1}{16n} |\Phi|^2 D\bar{D}\bar{\Phi} \bar{D}D\Phi \right] \\
&+ O(\varepsilon^{3/2})
\end{aligned} \tag{3.120}$$

and

$$W = \frac{2^{\frac{n}{4}+1} v_0^{\frac{1}{2}}}{n+2} \Phi^{\frac{n}{2}+1}, \tag{3.121}$$

with $\xi = \psi = 0$.

Notice that, as announced at the very beginning of the previous section,

$$K = K(\Phi, \bar{\Phi}; D_\alpha\Phi D^\alpha\bar{\Phi}, \bar{D}_\alpha\bar{\Phi} \bar{D}^\alpha\Phi; D^2\Phi, \bar{D}^2\bar{\Phi}; \partial_a\Phi, \partial_a\bar{\Phi}) \tag{3.122}$$

and

$$W = W(\Phi), \tag{3.123}$$

where the $\partial\Phi, \partial\bar{\Phi}$ dependence in K can be seen by using (3.9) in \mathcal{O}_3 :

$$\mathcal{O}_3 = |\Phi|^2 D\bar{D}\bar{\Phi} \bar{D}D\Phi = -8|\Phi|^2 \partial\Phi \cdot \partial\bar{\Phi}. \tag{3.124}$$

We observe that the Kähler potential contains both \mathcal{O}_0 , used in [7, 8, 9, 10, 11, 12], and \mathcal{O}_3 , used in [13, 14], another proof that this method takes into account the two ways of approaching the problem.

As anticipated we see what happens if we want to exclude the term $\propto |F|^4$. Such a term, as we said, modifies the Gaussian form of the F 's path integral and therefore one would like to find an alternative description of the DBI Action that does not include this expression. A simple and non-invasive way to remove it is to set $\beta = -\frac{1}{2}\alpha = -\frac{1}{4}$. Doing this excludes such a term from the very beginning of the discussion without interfering with the procedure.

Imposing this new relation fixes n and thus all the three linear combination constants. The values will be

$$\alpha = \frac{1}{2}, \beta = -\frac{1}{4}, \gamma = \frac{1}{4} \tag{3.125}$$

$$n = 1 \tag{3.126}$$

The scalar potential can then only be

$$V = v_0 \phi \tag{3.127}$$

if we choose a linear combination of operators such that terms $\propto |F|^4$ disappear. The problem is thus still solvable with the monomial potential ansatz, even if we lost the freedom for the exponent n .

Chapter 4

Conclusions

In this thesis we have studied systems with higher-derivative terms both in mechanics and in field theory. We have reviewed how those terms may lead to new and problematic degrees of freedom often called Ostrogradski's ghosts. These degrees of freedom in a quantum theory lead to negative norm states and/or an absence of a lower energy state and therefore to instability.

We have illustrated how to treat these higher-derivative terms in effective (field) theories via the application of the method of perturbative constraints, providing two illustrative examples in mechanics and (scalar) field theory. These examples have showed clearly how the core of the JLM procedure has not been changed in the generalization from mechanics to field theory: in both cases it concerns the insertion of the leading order EOM into higher-derivative terms.

Finally we have seen how the procedure can be used also in SUSY as a systematic method to cure unphysical degrees of freedom generated by higher derivatives. In previous works several of supersymmetric operators have been discarded because of their undesired terms like higher-than-one derivatives of dynamical fields or derivatives of the auxiliary field: thanks to this way of approaching such systems these terms are no longer a problem. We point out that a similar treatment of auxiliary fields was performed by Weinberg in a context of gravity corrections in inflaton models (see [21]).

For a particular class of operators we showed how the Lagrangian had to be constrained and we obtained its generic expression with only desirable zero-th and first derivatives of the dynamical field:

$$\mathcal{L} = -\frac{1}{2}f(\phi)(\partial\phi)^2 + h(\phi)(\partial\phi)^4 - V_{new}(\phi) \quad (4.1)$$

(with the field ξ suppressed to facilitate the comparison with the application of the JLM method to Classical Field Theory).

In the expression above we have

$$f(\phi) = 1 + \epsilon^2 \{2\alpha V - (\beta + \gamma)\phi V_\phi - \gamma\phi^2 V_{\phi\phi}\}$$

$$h(\phi) = \frac{1}{4}\alpha\epsilon^2$$

$$V_{new}(\phi) = V(\phi) - \epsilon^2 \{(\alpha + 2\beta)V^2 + (2\beta + \gamma)\phi V V_\phi + \gamma\phi^2 V_\phi^2\}$$

where the free constants α , β and γ clearly show the general nature of the discussion.

Once achieved this remarkable result of curing a generic supersymmetric Lagrangian – that can be used as a pathfinder to handle every kind of SUSY Higher-Derivative Operator – we sought a way to recover the DBI Action in this flat superspace framework. What we found is that a tuning of the three free parameters is not enough and that we also have to constrain the potential. Surprisingly a monomial potential satisfies the requirements and its power fixes the constants. This means that if we plugged such a monomial potential – as we would have done – in \mathcal{L} , we would have obtained a Lagrangian traceable to the DBI one just by fixing α , β and γ .

It is important here to notice how exactly three parameters are necessary to obtain the desired result: the previous treatments of the problem couldn't reach the hoped outcome since the latter is reachable only considering the widest range of operators and not only one of them¹.

In future it would be interesting to apply this procedure to a more general Lagrangian, considering wider classes of operators. In particular being able to identify every operator up to a certain desired mass dimension and applying the method as it's been done in this paper would be the best. Other possible extensions could be of considering also spinorial fields when treating chiral multiplets (and not setting them to zero from the very beginning) or of taking into account vector multiplets, whose higher-derivative operators contain ∂D terms.

Another important goal would be the one of bringing this procedure from flat superspace (SUSY) to curved superspace (SUGRA), since a lot applications can be found in inflationary contexts and these obviously require the presence of gravity.

In conclusion, the *generalized JLM method* derived above can pave the way for the construction of new supersymmetric inflation theories and, hopefully, can try to fill the gap between Supergravity and String Theory via higher and higher order corrections.

¹In [7, 8, 9, 10, 11, 12] the only operator used was \mathcal{O}_0 whereas in [13, 14] it was \mathcal{O}_3 .

Appendix A

Treatment of ϵ -polynomials

We notice that from (2.92) descends the EOM

$$(1 + 2\epsilon^2) \ddot{y} = - (1 + \epsilon^2) (\bar{g} + \Omega^2 y) + O(\epsilon^3) \quad (\text{A.1})$$

that apparently disagrees with (2.74)

$$\ddot{y} = - (1 - \epsilon^2) (\bar{g} + \Omega^2 y) + O(\epsilon^3) . \quad (\text{A.2})$$

To make them coincide we have to think to the algebraic nature of the system, indeed we are dealing with the commutative ring of polynomials in ϵ with real coefficients¹, $\mathbb{R}[\epsilon]$, where the role of the zero element is played by $O(\epsilon^{n+1})$ (in this case $O(\epsilon^3)$).

Since we are handling a ring, the division is forbidden: once set the perturbative order the multiplicative inverse does not exist and so we cannot divide (A.1) by $1 + 2\epsilon^2$ to get rid of the undesired term in front of \ddot{y} . What we can do is finding a polynomial in ϵ that, multiplied to $1 + 2\epsilon^2$, gives $1 + O(\epsilon^3)$. It turns out that such polynomial is simply

$$1 - 2\epsilon^2 \quad (\text{A.3})$$

which also gives the correct $1 - \epsilon^2 + O(\epsilon^3)$ when multiplied to $1 + \epsilon^2$.

Another way of recovering (A.2) from (A.1) consists in digging out that constraint for $\epsilon^2 \ddot{y}$ we discovered throughout the procedure and plugging it in the l.h.s. of (A.1).

That constraint is (2.72). To facilitate the check we rewrite it here

$$\epsilon^2 \ddot{y} = - (g - \Omega^2 \ell + \Omega^2 y) \epsilon^2 + O(\epsilon^3) . \quad (\text{A.4})$$

¹In the present case of equations of motion, the real coefficients are those $\propto \bar{g} + \Omega^2 y$ or $\propto \ddot{y}$.

As expected, inserting this expression in (A.1) and moving properly to the r.h.s. gives the correct result.

This treatment must be applied everytime we are tempted to divide by an ϵ -polynomial.

Appendix B

Details on Kähler potential and Superpotential dependence

Kähler potential

We started by considering the general form

$$K = K(\Phi, \bar{\Phi}; D_A\Phi, D_B\bar{\Phi}; D_AD_B\Phi, D_AD_B\bar{\Phi}; \dots) \quad (\text{B.1})$$

for our Kähler potential. We are going to analyse every term in this dependence focusing only on $\Phi, \bar{\Phi}$ will be its "barred-mirror". D 's graded algebra and chiral fields' definition are constantly used.

The first term is the trivial $D_A\Phi$, it contains

$$\begin{aligned} \partial_a\Phi \\ D_\alpha\Phi \end{aligned}$$

Then we have $D_AD_B\Phi$, that contains

$$\begin{aligned} \partial_a\partial_b\Phi \\ \partial_a D_\beta\Phi = D_\beta\partial_a\Phi \\ D_\alpha D_\beta\Phi = \frac{1}{2}\epsilon_{\alpha\beta}D^2\Phi \end{aligned}$$

$$\boxed{\bar{D}_{\dot{\alpha}}D_\beta\Phi = -2i(\sigma^a)_{\beta\dot{\alpha}}\partial_a\Phi \subset D_A\Phi}$$

The boxed term does not bring a new dependence since it belongs to the previous case. From now on every boxed expression signals a repetition and, so, a superfluous term.

The next one is $D_A D_B D_C \Phi$, with

$$\begin{aligned}
& \partial_a \partial_b \partial_c \Phi \\
& \partial_a \partial_b D_\gamma \Phi = \partial_a D_\gamma \partial_b \Phi = D_\gamma \partial_a \partial_b \Phi = \dots \\
& \partial_a D_\beta D_\gamma \Phi = \frac{1}{2} \epsilon_{\beta\gamma} \partial_a D^2 \Phi = \frac{1}{2} \epsilon_{\beta\gamma} D^2 \partial_a \Phi \\
& \boxed{\partial_a \bar{D}_{\dot{\beta}} D_\gamma \Phi = -2i(\sigma^b)_{\gamma\dot{\beta}} \partial_a \partial_b \Phi \subset D_A D_B \Phi} \\
& \boxed{\bar{D}_{\dot{\alpha}} D_\beta D_\gamma \Phi = -2i(\sigma^a)_{\beta\dot{\alpha}} \partial_a D_\gamma \Phi + 2i(\sigma^a)_{\gamma\dot{\alpha}} \partial_a D_\beta \Phi \subset D_A D_B \Phi} \\
& \boxed{D_\alpha \bar{D}_{\dot{\beta}} D_\gamma \Phi = -2i(\sigma^b)_{\gamma\dot{\beta}} D_\alpha \partial_b \Phi \subset D_A D_B \Phi}
\end{aligned}$$

Even if we can already see the underlying scheme we perform the last computation for more clarity.

Considering $D_A D_B D_C D_D \Phi$, we obtain

$$\begin{aligned}
& \partial_a \partial_b \partial_c \partial_d \Phi \\
& \partial_a \partial_b \partial_c D_\delta \Phi = \partial_a \partial_b D_\delta \partial_c \Phi = \partial_a D_\delta \partial_b \partial_c \Phi = \dots \\
& \partial_a \partial_b D_\gamma D_\delta \Phi = \frac{1}{2} \epsilon_{\gamma\delta} \partial_a \partial_b D^2 \Phi = \frac{1}{2} \epsilon_{\gamma\delta} D^2 \partial_a \partial_b \Phi \\
& \boxed{\partial_a \partial_b \bar{D}_{\dot{\gamma}} D_\delta \Phi = -2i(\sigma^c)_{\delta\dot{\gamma}} \partial_a \partial_b \partial_c \Phi \subset D_A D_B D_C \Phi} \\
& \boxed{\partial_a \bar{D}_{\dot{\beta}} D_\gamma D_\delta \Phi = -2i(\sigma^b)_{\gamma\dot{\beta}} \partial_a \partial_b D_\delta \Phi + 2i(\sigma^b)_{\delta\dot{\beta}} \partial_a \partial_b D_\gamma \Phi \subset D_A D_B D_C \Phi} \\
& \boxed{\partial_a D_\beta \bar{D}_{\dot{\gamma}} D_\delta \Phi = -2i(\sigma^c)_{\delta\dot{\gamma}} \partial_a D_\beta \partial_c \Phi \subset D_A D_B D_C \Phi} \\
& \boxed{\bar{D}_{\dot{\alpha}} \bar{D}_{\dot{\beta}} D_\gamma D_\delta \Phi = 4(\sigma^a)_{\delta\dot{\alpha}} (\sigma^b)_{\gamma\dot{\beta}} \partial_a \partial_b \Phi - 4(\sigma^a)_{\gamma\dot{\alpha}} (\sigma^b)_{\delta\dot{\beta}} \partial_a \partial_b \Phi \subset D_A D_B \Phi} \\
& \boxed{\bar{D}_{\dot{\alpha}} D_\beta \bar{D}_{\dot{\gamma}} D_\delta \Phi = 4(\sigma^a)_{\beta\dot{\alpha}} (\sigma^b)_{\delta\dot{\gamma}} \partial_a \partial_b \Phi \subset D_A D_B \Phi}
\end{aligned}$$

It is clear how increasing by one the number of superspace derivatives acting on Φ leads exactly to 3 new terms, moreover always of the same form. In fact, we can group terms in the six groups (once considered also the $\bar{\Phi}$ case)

$\partial^{(i)} \Phi$	$D_\alpha \partial^{(k)} \Phi$	$D^2 \partial^{(p)} \Phi$	$\partial^{(j)} \bar{\Phi}$	$\bar{D}_{\dot{\alpha}} \partial^{(l)} \bar{\Phi}$	$\bar{D}^2 \partial^{(q)} \bar{\Phi}$
Φ	$D_\alpha \Phi$	$D^2 \Phi$	$\bar{\Phi}$	$\bar{D}_{\dot{\alpha}} \bar{\Phi}$	$\bar{D}^2 \bar{\Phi}$
$\partial_a \Phi$	$D_\alpha \partial_a \Phi$	$D^2 \partial_a \Phi$	$\partial_a \bar{\Phi}$	$\bar{D}_{\dot{\alpha}} \partial_a \bar{\Phi}$	$\bar{D}^2 \partial_a \bar{\Phi}$
$\partial_a \partial_b \Phi$	$D_\alpha \partial_a \partial_b \Phi$	$D^2 \partial_a \partial_b \Phi$	$\partial_a \partial_b \bar{\Phi}$	$\bar{D}_{\dot{\alpha}} \partial_a \partial_b \bar{\Phi}$	$\bar{D}^2 \partial_a \partial_b \bar{\Phi}$
$\partial_a \partial_b \partial_c \Phi$	$D_\alpha \partial_a \partial_b \partial_c \Phi$	$D^2 \partial_a \partial_b \partial_c \Phi$	$\partial_a \partial_b \partial_c \bar{\Phi}$	$\bar{D}_{\dot{\alpha}} \partial_a \partial_b \partial_c \bar{\Phi}$	$\bar{D}^2 \partial_a \partial_b \partial_c \bar{\Phi}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

where colors indicate the original membership to the class characterized by a certain number of superspace derivatives.

Thus the equation (3.15) has been proven.

Superpotential

As regards the Superpotential, we make an assumption similar to the one considered in [7], that is a W which is a power series of Φ , $\bar{D}^2\bar{\Phi}$ and their derivatives. An arbitrary term in this series is of the form $(i, k, l, q \in \mathbb{N}_0)$

$$(\partial^{(i)}\Phi)^k (\bar{D}^2\partial^{(q)}\bar{\Phi})^l = \bar{D}^2 \left((\partial^{(i)}\Phi)^k (\bar{D}^2\partial^{(q)}\bar{\Phi})^{l-1} \partial^{(q)}\bar{\Phi} \right) \quad (\text{B.2})$$

since derivatives of a chiral (antichiral) superfield are still chiral (antichiral) superfields.

Using the identity

$$\int d^4x d^2\theta \bar{D}^2 f(x, \theta, \bar{\theta}) = -4 \int d^4x d^4\theta f(x, \theta, \bar{\theta}) , \quad (\text{B.3})$$

with f and arbitrary superfield, we notice that the dependence of W on $\bar{D}^2\partial^{(q)}\bar{\Phi}$ can be entirely absorbed into K .

Thus we obtained that a Superpotential, sum of terms like (B.2), depends just on Φ and on its derivatives, proving the equation (3.16).

Appendix C

Details on DBI Lagrangian expansion

In order to find (3.101) we have to expand the determinant under the square root in (3.99). First we take the case for matrices to then generalize to tensors.

We start by considering a square $N \times N$ matrix A and computing (cf. [22])¹

$$\begin{aligned} \det(\mathbb{1} + A) &= \exp \left\{ \text{Tr} [\log(\mathbb{1} + A)] \right\} \\ &= \exp \left\{ \text{Tr} \left[\sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} A^p \right] \right\} \\ &= \exp \left\{ \sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} \text{Tr} [A^p] \right\} \\ &= \sum_{q=0}^{\infty} \frac{1}{q!} \left(\sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} \text{Tr} [A^p] \right)^q \\ &\equiv \sum_{q=0}^{\infty} \frac{1}{q!} X^q \end{aligned} \tag{C.1}$$

with

$$X = \sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} \text{Tr} [A^p] . \tag{C.2}$$

Since we want to have $\mathbb{1} + A$ close to the identity $\mathbb{1}$, we should find a consistent way to truncate these series above.

¹ $\mathbb{1} + A$ has to be nonsingular.

A great help comes from the properties of the trace of a matrix, indeed we know that if λ_i ($i = 1, \dots, N$) are the eigenvalues of A , then

$$\text{Tr} [A^p] = \sum_{i=1}^N \lambda_i^p \quad (\text{C.3})$$

and so, once defined

$$\varepsilon \equiv \max_{1 \leq i \leq N} |\lambda_i| \quad (\text{C.4})$$

and considered (realistically) a dimension $N = O(1)^2$, we'll have

$$\text{Tr} [A^p] = O(\varepsilon^p) . \quad (\text{C.5})$$

Now we notice that having $\mathbf{1} + A$ close to the identity $\mathbf{1}$ implies a *small* A , thus

$$\lambda_i \ll 1 \quad \forall \quad i \quad (\text{C.6})$$

and, more importantly,

$$\varepsilon \ll 1 . \quad (\text{C.7})$$

Thanks to this last relation we can consider ε as our expansion parameter and in terms of it we will apply the truncations.

So, approximating to ε^4 , we get

$$\begin{aligned} X &= \text{Tr} [A] - \frac{1}{2} \text{Tr} [A^2] + \frac{1}{3} \text{Tr} [A^3] + O(\varepsilon^4) \\ X^2 &= \text{Tr}^2 [A] - \text{Tr} [A] \text{Tr} [A^2] + O(\varepsilon^4) \\ X^3 &= \text{Tr}^3 [A] + O(\varepsilon^4) \end{aligned} \quad (\text{C.8})$$

that can be used in (C.1) to obtain

$$\begin{aligned} \det (\mathbf{1} + A) &= 1 + \text{Tr} [A] + \frac{1}{2} \left\{ \text{Tr}^2 [A] - \text{Tr} [A^2] \right\} \\ &\quad + \frac{1}{6} \left\{ \text{Tr}^3 [A] - 3 \text{Tr} [A] \text{Tr} [A^2] + 2 \text{Tr} [A^3] \right\} + O(\varepsilon^4) \end{aligned} \quad (\text{C.9})$$

We can finally switch to our case. The generalization from matrices to (0, 2)-tensors can be done by thinking on how we *multiply*³ two tensors of this

²In our case $N = 4$.

³The *multiplication* is not part of the permitted operations between two tensors. With this term we denote the *contraction* (2-3,4-5) of the tensorial *product* $\mathbf{A} \mathbf{g}^{-1} \mathbf{B}$, both allowed operations.

kind (i.e., in our situation, what is intended when we write A^2 or a generic power A^n) and on the properties of the identity matrix.

Two $(0, 2)$ -tensors \mathbf{A} and \mathbf{B} – of components A_{mn} and B_{mn} – are multiplied (see note) through the inverse of the metric, the $(2, 0)$ -tensor of components g^{mn} :

$$(\mathbf{A} \cdot \mathbf{B})_{mn} = A_{mk} g^{kl} B_{ln} . \quad (\text{C.10})$$

The pursuit of an identity $\mathbf{1}$ for these kind of tensors passes through the properties of the former, indeed from

- $\mathbf{1} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{1} = \mathbf{A}$,
- $\mathbf{1} = \mathbf{1}^{-1}$,

we derive the relations that this "identity tensor" must have

- $(\mathbf{1})_{mk} g^{kl} A_{ln} = A_{mk} g^{kl} (\mathbf{1})_{ln} = A_{mn}$;
- $(\mathbf{1})_{mk} g^{kl} (\mathbf{1})_{ln} = (\mathbf{1})_{mn}$,

that is

$$(\mathbf{1})_{mk} g^{kn} = \delta_m^n \quad (\text{C.11})$$

with δ_m^n Kronecker delta.

The latter equation identifies the metric tensor \mathbf{g} as the desired identity tensor $\mathbf{1}$

$$(\mathbf{1})_{mn} = g_{mn} \quad (\text{C.12})$$

and allow us to use (C.9) in the present case of ⁴

$$g_{mn} + \varepsilon(\partial_m \phi \partial_n \phi + \partial_m \xi \partial_n \xi) \quad (\text{C.13})$$

with

$$A_{mn} \equiv \varepsilon(\partial_m \phi \partial_n \phi + \partial_m \xi \partial_n \xi) . \quad (\text{C.14})$$

⁴One can take (C.1) as the definition of the determinant of a tensor near its associated metric, as often happens for the generalizations of known functions to unusual domains.

What we need to perform the computation can be summarized in

$$\begin{aligned}
\text{Tr} [A] &= g^{lk} A_{kl} \\
&= \varepsilon((\partial\phi)^2 + (\partial\xi)^2) \\
&= 2\varepsilon|\partial A|^2
\end{aligned} \tag{C.15}$$

$$\begin{aligned}
\text{Tr} [A^2] &= g^{lk} A_{kr} g^{rs} A_{sl} \\
&= \varepsilon^2((\partial\phi)^4 + (\partial\xi)^4 + 2(\partial\phi \cdot \partial\xi)^2) \\
&= 2\varepsilon^2(|\partial A|^4 + (\partial A)^2(\partial A^*)^2)
\end{aligned} \tag{C.16}$$

$$\begin{aligned}
\text{Tr} [A^3] &= g^{lk} A_{kr} g^{rs} A_{st} g^{th} A_{hl} \\
&= \varepsilon^3((\partial\phi)^6 + (\partial\xi)^6 + 3(\partial\phi \cdot \partial\xi)^2((\partial\phi)^2 + (\partial\xi)^2)) \\
&= 2\varepsilon^3|\partial A|^2(|\partial A|^4 + 3(\partial A)^2(\partial A^*)^2)
\end{aligned} \tag{C.17}$$

(where (3.102) has repeatedly been used).

Computing now (C.9) gives us a surprising result: the ε^3 -term disappear in the determinant expression

$$\det(\mathbf{1} + A) = 1 + 2\varepsilon|\partial A|^2 + \varepsilon^2(|\partial A|^4 - (\partial A)^2(\partial A^*)^2) + O(\varepsilon^4) \tag{C.18}$$

leaving a *hole* in the expansion.

As last step we have to take the square root. As it's been done so far, we truncate the expansion to the ε^3 -order

$$\begin{aligned}
\sqrt{1+x} &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + \dots \\
&= 1 + \varepsilon|\partial A|^2 - \frac{1}{2}\varepsilon^2(\partial A)^2(\partial A^*)^2 + \frac{1}{2}\varepsilon^3|\partial A|^2(\partial A)^2(\partial A^*)^2 + O(\varepsilon^4)
\end{aligned} \tag{C.19}$$

– where x has been defined as

$$x \equiv 2\varepsilon|\partial A|^2 + \varepsilon^2(|\partial A|^4 - (\partial A)^2(\partial A^*)^2) . \tag{C.20}$$

We can finally expand the DBI Lagrangian in (3.99)

$$\begin{aligned}
\mathcal{L}_{DBI} &= \varepsilon^{-1} \left[- (1 + \varepsilon V(\phi, \xi)) \sqrt{\det(g_{mn} + \varepsilon(\partial_m \phi \partial_n \phi + \partial_m \xi \partial_n \xi))} + 1 \right] \\
&= \varepsilon^{-1} \left[- (1 + \varepsilon V(A, A^*)) \times \right. \\
&\quad \left. \times \left(1 + \varepsilon|\partial A|^2 - \frac{1}{2}\varepsilon^2(\partial A)^2(\partial A^*)^2 + \frac{1}{2}\varepsilon^3|\partial A|^2(\partial A)^2(\partial A^*)^2 \right) + 1 \right] \\
&\quad + O(\varepsilon^3) \\
&= - (1 + \varepsilon V)|\partial A|^2 - V + \frac{1}{2}\varepsilon (1 + \varepsilon V) (\partial A)^2(\partial A^*)^2 \\
&\quad - \frac{1}{2}\varepsilon^2|\partial A|^2(\partial A)^2(\partial A^*)^2 + O(\varepsilon^3) .
\end{aligned} \tag{C.21}$$

This equation agrees with (3.101) if we consider its approximation to order ε^2 instead of ε^3 .

Thus we successfully showed the method to expand the DBI action, also pushing it further to derive the next unused ε^2 -order. About this latter we can notice how it brings a correction to the $(\partial\phi)^4$ term – moreover of the same form of the one brought from ε to the kinetic term – and introduces a new $(\partial\phi)^6$ term.

Bibliography

- [1] J. Z. Simon, “Higher derivative expansions and nonlocality with applications to gravity and the stability of flat space,” Master’s thesis, 1990.
- [2] X. Jaen, J. Llosa, and A. Molina, “A Reduction of order two for infinite order lagrangians,” *Phys. Rev. D*, vol. 34, 1986.
- [3] F. J. de Urries and J. Julve, “Ostrogradski formalism for higher derivative scalar field theories,” *J. Phys. A*, vol. 31, 1998.
- [4] J. Wess and J. Bagger, *Supersymmetry and supergravity*. Princeton, NJ, USA: Princeton University Press, 1992.
- [5] F. G. Pedro, “UV physics and the speed of sound during inflation,” *Eur. Phys. J. Plus*, vol. 134, no. 5, 2019.
- [6] C. Grosse-Knetter, “Effective Lagrangians with higher derivatives and equations of motion,” *Phys. Rev. D*, vol. 49, 1994.
- [7] D. Ciupke, “Scalar Potential from Higher Derivative $\mathcal{N} = 1$ Superspace,” 2016.
- [8] J. Khoury, J.-L. Lehners, and B. Ovrut, “Supersymmetric P(X, ϕ) and the Ghost Condensate,” *Phys. Rev. D*, vol. 83, 2011.
- [9] M. Koehn, J.-L. Lehners, and B. A. Ovrut, “Higher-Derivative Chiral Superfield Actions Coupled to N=1 Supergravity,” *Phys. Rev. D*, vol. 86, 2012.
- [10] M. Koehn, J.-L. Lehners, and B. A. Ovrut, “DBI Inflation in N=1 Supergravity,” *Phys. Rev. D*, vol. 86, 2012.
- [11] D. Ciupke, J. Louis, and A. Westphal, “Higher-Derivative Supergravity and Moduli Stabilization,” *JHEP*, vol. 10, 2015.

- [12] M. Cicoli, D. Ciupke, and S. Maraventano, “Superspace formulation of higher-derivative actions from d-branes,” 2017.
- [13] S. Bielleman, L. E. Ibañez, F. G. Pedro, I. Valenzuela, and C. Wieck, “The DBI Action, Higher-derivative Supergravity, and Flattening Inflaton Potentials,” *JHEP*, vol. 05, 2016.
- [14] S. Bielleman, L. E. Ibañez, F. G. Pedro, I. Valenzuela, and C. Wieck, “Higgs-otic Inflation and Moduli Stabilization,” *JHEP*, vol. 02, 2017.
- [15] D. Tong, “String Theory,” 2009.
- [16] B. Zwiebach, *A first course in string theory*. Cambridge University Press, 2006.
- [17] F. Quevedo, “Lectures on string/brane cosmology,” *Class. Quant. Grav.*, vol. 19, 2002.
- [18] C. V. Johnson, “D-brane primer,” in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 99): Strings, Branes, and Gravity*, 2000.
- [19] X. Yang, W.-L. Xu, and Y.-C. Huang, “New general DBI action, its solution to the paradox of the conversion of kinetic and potential energy in equal rights and their true applications to inflationary cosmology,” 2021.
- [20] F. G. Pedro and G. Gotti, “D-brane inflation in the non-relativistic and relativistic regimes,” 2018.
- [21] S. Weinberg, “Effective Field Theory for Inflation,” *Phys. Rev. D*, vol. 77, 2008.
- [22] I. C. Ipsen and D. J. Lee, “Determinant approximations,” 2011.
- [23] D. G. Cerdeño and C. Muñoz, “An introduction to supergravity,” *PoS*, vol. CORFU98, 1998.
- [24] F. Quevedo, S. Krippendorff, and O. Schlotterer, “Cambridge Lectures on Supersymmetry and Extra Dimensions,” 2010.
- [25] J. Louis, “Introduction to supersymmetry and supergravity,” 2014.
- [26] S. Weinberg, *The quantum theory of fields. Vol. 3: Supersymmetry*. Cambridge University Press, 2013.

- [27] D. Baumann and D. Green, “Supergravity for Effective Theories,” *JHEP*, vol. 03, 2012.
- [28] D. Bertolini, J. Thaler, and Z. Thomas, “Super-Tricks for Superspace,” in *Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales*, 2013.
- [29] J. M. Queiruga, “Supersymmetric galileons and auxiliary fields in 2+1 dimensions,” *Phys. Rev. D*, vol. 95, no. 12, 2017.
- [30] C. P. Burgess, M. W. Horbatsch, and S. P. Patil, “Inflating in a Trough: Single-Field Effective Theory from Multiple-Field Curved Valleys,” *JHEP*, vol. 01, 2013.
- [31] I. Antoniadis, E. Dudas, and D. M. Ghilencea, “Supersymmetric Models with Higher Dimensional Operators,” *JHEP*, vol. 03, 2008.
- [32] E. Dudas and D. M. Ghilencea, “Effective operators in SUSY, superfield constraints and searches for a UV completion,” *JHEP*, vol. 06, 2015.
- [33] C. Armendariz-Picon, T. Damour, and V. F. Mukhanov, “k - inflation,” *Phys. Lett. B*, vol. 458, 1999.
- [34] A. H. Guth, “The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems,” *Phys. Rev. D*, vol. 23, 1981.
- [35] A. D. Linde, *Particle physics and inflationary cosmology*, vol. 5. 1990.