Proof of the most general multiple-scalar field theory in Minkowski space-time free of Ostrogradski Ghost

Vishagan Sivanesan

School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK

(Dated: May 11, 2014)

We give a proof of the most general multiple scalar field theory in flat space-time free of Ostrogradski ghosts. We start from the assumption that the action is a functional of the scalar fields and their derivatives of order up to two (which in general can have field equations of derivative order up to four), and arrive at the most-general field theory that can be constructed such that the field equations are of derivative order up to two. Naturally, this theory includes all physically viable multi-scalar field theories in flat space-time, in the sense that they don’t suffer from Ostrogradski instability [10].

I. INTRODUCTION

Recently, there has been a lot of interest in scalar field theories with higher derivative structure, and with derivative self-interaction. These theories have been investigated, especially in the context of modifying gravity in the infra-red. Famous examples include, the effective field theory of the DGP model [1] on the brane, more precisely the decoupling limit of this theory, and the generalisation of this effective theory to higher order in fields, known as, the Galileon field theory [11]. These field theories with derivative self-interactions exhibit novel properties at both the classical and quantum mechanical level. The Vainshtein mechanism [7, 12, 13] which suppresses the scalar force near dense environments, local strong coupling [2], stable violation of null-energy condition [14][3][4][5] are a few such properties. In [9] the authors generalised the single Galileon field theory to find the most general scalar field theory in Minkowski space-time with field equations of derivative order up to two. Remarkably when this theory was covariantised and appropriate non-minimal couplings added, it reduced to the most general scalar-tensor theory in arbitrary background in 4-dimensions which was proved by Horndeski in 1974 [8]. Later this covariantisation procedure was carried out for the multiple scalar field theory (that is proved in this paper to be the most general one), and was conjectured to be the most general multi-scalar-tensor theory in arbitrary dimensions and arbitrary backgrounds [6]. This paper can be taken to be a companion to [6] where the conjectured starting point, the most general multiple scalar field theory, is proved.

II. STATEMENT OF THE PROOF

The most general multiple scalar field theory in flat space-time satisfying the conditions

i) Lagrangian contains up to second order derivatives of the fields

ii) Field equations contain up to second order derivatives of the fields

is given by,

\[ S_{\text{multi-scalar}} = \int_M d^Dx A(\bar{X}_{ij}, \pi_i) + \sum_{m=1}^{D-1} A^{k_1...k_m} (\bar{X}_{ij}, \pi_i) \partial^{a_1} \partial_k \pi_{k_1} \cdots \partial^{a_m} \partial_{k_m} \pi_{k_m} \]  (1)

where \( \bar{X}_{ij} = \frac{1}{2} \partial_i \pi_j \partial^a \pi_j \) and \( \frac{\partial A^{i_1...i_m}}{\partial X_{kl}} \) is symmetric in all of its indices \( i_1, \ldots, i_m, k, l \).

III. POLYNOMIALITY OF THE SECOND DERIVATIVES OF \( \pi_i \) IN THE LAGRANGIAN

We start with the general multi-field action in D dimensions of the form,

\[ S = \int d^Dx \mathcal{L}(\pi_i, \partial_\alpha \pi_j, \partial_\beta \partial_\epsilon \pi_k) \]  (2)

svishagan@gmail.com
Here $i, j, k$ are used to label different fields ($i \in \{1 \ldots N\}$). The Euler-Lagrange equations of this action is
\[
\frac{\partial L}{\partial \pi_i} - \partial_a \left( \frac{\partial L}{\partial \pi_{ia}} \right) + \partial_b \partial_b \left( \frac{\partial L}{\partial \pi_{iab}} \right) = 0
\] (3)
Note that we occasionally resort to the notation $\pi_{i...b} \equiv \partial_b \ldots \partial_a \pi_i$ for brevity. In general only the third term on the LHS will have fourth derivatives, explicitly this is,
\[
\frac{\partial L}{\partial \pi_{icd}} \pi_{abcd} + \frac{\partial L}{\partial \pi_{icd}} \pi_{iab} \pi_{abcd} = 0
\] (4)
imposing the constraint that this term vanishes we get,
\[
L_{icd} \pi_{abcd} = 0
\] (5)
where $(\ldots)$ stands for symmetrisation. Furthermore, the following symmetries are trivial.
\[
L_{icd} = L_{cic} = L_{dc} = L_{dc}
\] (7)
On account of these symmetries, (5) is satisfied if and only if the following cyclic identity holds
\[
L_{icd} + L_{icd} + L_{icd} + L_{icd} = 0
\] (8)
We define the tensor,
\[
L_{\ldots ab} \equiv \frac{\partial}{\partial \pi_{aa} \ldots \partial \pi_{ab}} L
\] (9)
This tensor naturally inherits the cyclic identity
\[
\frac{\partial}{\partial \pi_{aa} \ldots \partial \pi_{ab}} + \frac{\partial}{\partial \pi_{ab} \ldots \partial \pi_{bc}} + \frac{\partial}{\partial \pi_{bc} \ldots \partial \pi_{cd}} = 0
\] (10)
Consider the $(D+1)$th derivative of the Lagrangian with respect to $\pi_{ab}$ that is given by the tensor of rank $(0,2D+2),
\[
L_{\ldots ab} \equiv \frac{\partial}{\partial \pi_{aD+bD+1}} L
\] (11)
In D dimensions it is easy to see that the components of this tensor will have atleast 3 identical space-time indices. Using the symmetry any component of this tensor can be cast in the following two forms,
\[
F_1 = L_{\ldots ab},
F_2 = L_{\ldots ab}
\] (12)
Now using the cyclic identity on the identical indices $a$ of $F_1$ yields,
\[
F_1 = 0
\] (13)
and using the cyclic identity on the first three indices $a, b, a$ of $F_2$ and subsequently on the similar indices gives,
\[
F_2 = L_{\ldots ab}, aD = -\frac{1}{2} L_{\ldots ab} = 0
\] (14)
Thus we have established,

**Theorem 1.** The most general action that yields equations of motion of derivative order up to two, has the Lagrangian that depends polynomially on the second derivatives of the field and the polynomial order is bounded above by $D$ in $D$-dimensions ie, Order of the polynomial $\leq D$. 
Having established the polynomiality of the second derivatives in $\mathcal{L}$, we write down a generic term in the Lagrangian that is constrained by this fact.

$$
\mathcal{L}_{2p,q} = A(\bar{X}_{ij}, \pi_i)^{i_1 \ldots i_p j_1 \ldots j_q} \sum_{\sigma \in S_{p+q}} \sum_{p \in S_q} c(\sigma)d(p) \left[ \left( \partial_{a_1} \pi_{i_1} \partial^{\sigma(a_1)} \pi_{i_2} \right) \ldots \left( \partial_{a_p} \pi_{i_{2p-1}} \partial^{\sigma(a_p)} \pi_{i_{2p}} \right) \times \partial_{p(b_1)} \pi_{j_1} \ldots \partial_{p(b_q)} \pi_{j_q} \right]
$$

(15)

A note on notations is in order. Here $i, j, k \ldots$ denote internal indices and $a, b, c \ldots$ are space-time indices. $\bar{X}_{ij} = \frac{1}{2} \partial_{a_i} \partial^{a_j} \pi_j$ and $A(\bar{X}_{ij}, \pi_i)^{i_1 \ldots i_p j_1 \ldots j_q}$ can in general be a non-polynomial function. $\sigma \in S_{p+q}$ is an element of the symmetric permutation group acting on the positions of the space-time labels $a_i, b_j$ upstairs, similarly $p \in S_q$ is an element of the symmetric permutation group acting on the positions of the space-time labels $b_j$ downstairs. $C_\sigma, d_p$ are coefficients that depend on $\sigma, p$ resp (in general they can also be functions of $X_{ij}, \pi_k$). We observe the following facts.

i) Any term generated by the action of $\sigma$ within the positions of the indices $b_i$ can also be generated via the action of $p$ on $b_i$ downstairs since the indices are summed over.

ii) Any term generated by the action of $\sigma$ within the positions of the indices $a_i$ can also be generated by shuffling the order of internal indices $i_j$.

On account of these facts we can avoid over-counting by reducing the symmetry group of $\sigma$ to the quotient group defined by,

$$
S_{p+q} \rightarrow \frac{S_{p+q}}{S_p S_q}
$$

(16)

A. Symmetry of $A(\bar{X}_{ij}, \pi_k)^{i_1 \ldots i_2 \ldots}$

So far we have established that a generic term in the Lagrangian can be cast in the form,

$$
\mathcal{L}_{2p,q} \left[ A(\bar{X}_{ij}, \pi_i)^{i_1 \ldots i_p j_1 \ldots j_q} \right] \equiv A(\bar{X}_{ij}, \pi_i)^{i_1 \ldots i_p j_1 \ldots j_q} \times \sum_{\sigma \in S_{p+q}} \sum_{p \in S_q} c(\sigma)d(p) \left[ \left( \partial_{a_1} \pi_{i_1} \partial^{\sigma(a_1)} \pi_{i_2} \right) \ldots \left( \partial_{a_p} \pi_{i_{2p-1}} \partial^{\sigma(a_p)} \pi_{i_{2p}} \right) \times \partial_{p(b_1)} \pi_{j_1} \ldots \partial_{p(b_q)} \pi_{j_q} \right]
$$

(17)

We note that any cancellation of 3’rd and 4’th derivatives coming from the variational principle should occur within each of these terms as they are structurally different from each other. We ignore the variation of $A^{i_1 \ldots}$ for now, and focus on terms that are of derivatives order 3 (containing $(\pi_i \pi \pi)$). Let us isolate a generic term that would give rise to 3’rd order derivatives in $\mathcal{L}_{(2p,q)}$.

$$
\mathcal{L}_{2p,q} \supset B_{k,l} \equiv A^{\ldots k \ldots} \sum_{\sigma, p} c_\sigma d_p \left( \partial^{\sigma(a)} \pi_k \partial^{\sigma(a)} \pi_l \right) \ldots
$$

(18)

where we have suppressed the dependence of the arbitrary function $A^{i \ldots}(\bar{X}_{ij}, \pi_k)$ and other factors of the fields for brevity. For the $\delta \pi_k$ variation this will contain the term,

$$
\delta B_{k,l} = -A^{\ldots k \ldots} \sum_{\sigma, p} c_\sigma d_p \left( \partial \pi_k \partial^{\sigma(a)} \partial^{\sigma(a)} \pi_l \right) \ldots
$$

(19)

Note however that under the interchange of indices $k, l$ the corresponding variation yields a similar term i.e,

$$
\delta B_{l,k} = +A^{\ldots k \ldots} \sum_{\sigma, p} c_\sigma d_p \left( \partial \pi_k \partial^{\sigma(a)} \partial^{\sigma(a)} \pi_l \right) \ldots
$$

(20)

These would cancel if,

$$
A^{\ldots k \ldots} = A^{\ldots l \ldots}
$$

(21)
This implies that $A^{i_1 \cdots i_p} (X_{ij}, \pi_k)$ is completely symmetric in its indices.

As we have seen the symmetry of the $\sigma$ would just amount to the interchange of space-time labels $a_i \leftrightarrow b_j$. On account of the symmetry in $A^{i_1 \cdots i_p}$, it does not matter which labels are interchanged i.e., all such operations form an equivalence class labelled by how many pairs are interchanged. Now consider the case where two interchanges are made, we write a generic term of this type, and omit the internal indices and the function $A^{i_1 \cdots i_p}$ as they don’t play any role in this argument, due to the symmetry of $A^{i_1 \cdots i_p}$,

\[ C \equiv c_p c_a \pi_{a_1} \pi_{a_2}^{b_k} \pi_{p(b_h)}^{a_j} \pi_{p(b_i)}^{a_i} \cdots \]  

(22)

Here we have interchanged $a_i \leftrightarrow b_k$, $a_j \leftrightarrow b_l$. This term would give rise to 4’th derivatives in the eom given by,

\[ \delta_{ij} C \supset 2c_p c_a \pi_{a_1} \pi_{a_2}^{b_k} \left( \pi_{p(b_h)}^{a_j} \pi_{p(b_i)}^{a_i} \right) \delta \pi \cdots \]  

(23)

Here $\delta_{ij}$ denotes the variation and the subsequent integration by parts restricted to the factors $\pi_{p(b_h)}^{a_j} \pi_{p(b_i)}^{a_i}$. Note that $C$ is invariant under the interchange of either $a_i \leftrightarrow a_j$ or $p(b_i) \leftrightarrow p(b_l)$, this means that the 4’th derivative term in $\delta C$ would vanish only if all terms proportional to $C$ vanish. Thus we have established,

**Theorem 2.** The action of $\sigma$ is necessarily limited to just the interchange of at most 1 index $a_i \leftrightarrow b_j$, if the condition (ii) is to be satisfied. Thus the $L_{2p,q}$ is restricted to

\[ L_{0,q} \left[ A^{i_1 \cdots i_q} \right] = A^{i_1 \cdots i_q} \sum_{p \in S_n} c_p \partial^{b_k}_{p(b_h)} \pi_{(i_1)} \cdots \partial^{b_k}_{p(b_h)} \pi_{(i_q)} \]  

(24)

\[ L_{2,q} \left[ A^{i_1 \cdots i_{q+2}} \right] = A^{i_1 \cdots i_{q+2}} \sum_{p \in S_n} c_p \partial a \pi_{i_1} \partial^{b_1}_{p(b_h)} \pi_{(i_2)} \partial^{b_1}_{p(b_h)} \pi_{(i_3)} \cdots \partial^{b_1}_{p(b_h)} \pi_{(i_{q+2})} \]  

(25)

Having resolved the $\sigma$ permutations we focus on the permutation $p$ acting on the indices $b_i$ downstairs.

**Definition 3.** $T^{ij}$ is a transposition map that acts on individual instances of the permutations generated by $p$ and interchanges $b_i \leftrightarrow b_j$,

\[ T^{ij} : P^{ij}_k \rightarrow P^{ij}_k \]  

Note that $T^{ij} \circ T^{ij} = id$. Here $P^{ij}_k$ is an instance of the permutation where $b_i$ appears before $b_j$ and $k \in \{1, \ldots q! / 2\}$.

Two permutations $P^{ij}_k, P^{ij}_k'$ are identical if they are related by the relabelling of the indices $b_i$. Relabelling of $b_k \leftrightarrow b_l$ would change their order both upstairs and downstairs, which can be put back into the allowed form by interchanging the fields that carry the relabelled indices upstairs. Let us illustrate this by an example. Consider the following term in the series, where we have again omitted the symmetric function $A^{i_1 \cdots i_p}$ and the remaining factors of the field,

\[ \partial^{b_k}_{b_i} \pi_{i_1} \partial^{b_k}_{a_i} \pi_{i_{r+1}} \cdots \partial^{b_k}_{b_i} \pi_{i_{r+5}} \partial^{b_k}_{b_i} \pi_{i_{r+6}} \]  

(26)

relabelling $b_2 \leftrightarrow b_3$ yields the identical term,

\[ \partial^{b_2}_{b_i} \pi_{i_1} \partial^{b_3}_{a_i} \pi_{i_{r+1}} \cdots \partial^{b_2}_{b_i} \pi_{i_{r+5}} \partial^{b_3}_{b_i} \pi_{i_{r+6}} \]  

(27)

But permutation group $p$ would not alter the indices upstairs therefore we have to move the fields back such that the order upstairs is unchanged. We can interchange the positions of the terms $\partial^{b_2}_{b_i} \pi_{i_1}$, $\partial^{b_2}_{b_i} \pi_{i_{r+1}}$ invariably since the function $A^{i_1 \cdots i_p}$ is symmetric. We get the identical term

\[ \partial^{b_2}_{b_i} \pi_{i_{r+2}} \partial^{b_2}_{b_i} \pi_{i_{r+3}} \partial^{b_2}_{b_i} \pi_{i_{r+4}} \partial^{b_2}_{b_i} \pi_{i_{r+5}} \partial^{b_2}_{b_i} \pi_{i_{r+6}} \]  

(28)

comparing this to (26) we see that it differs by two transpositions acting on the original term interchanging $b_2 \leftrightarrow b_3$ and $b_4 \leftrightarrow b_5$, in other words it is operated by the map $T^{23} \circ T^{46}$. Thus we have established,

**Theorem 4.** Two identical terms generated by the permutation group $p$ necessarily differ by an even number of transpositions acting downstairs.
The following condition, $\partial A$,

In order for the 3'rd derivative terms coming from the variation of the function $A$, $A'$ to cancel. At the level of Lagrangian this implies that, interchanging any indices $\delta^{ij}$, $\delta^{kl}$ care of theorem 4. Varying the second derivatives and integrating by parts would yield for each of these terms

$$\delta_{ij} \left( c_p P^{ij}_k \right) \supset 2c_p \cdots \delta^{i_k} \delta^{j^s} \pi_i \pi_j \cdots$$

and

$$\delta_{ji} \left( c_p P^{ij}_k \right) \supset 2c_p' \cdots \delta^{i_k} \delta^{j^s} \pi_i \pi_j \cdots$$

Note that there is also a natural induced map between the coefficients, $T^{ij}_{\text{induced}} : c_p \rightarrow c'_p$. Now let us observe the following facts,

i) Every 4'th derivative term occurring in the field equation originates from the pairs $(P^{ij}_k, P^{ij}_k)$, $(P^{ij}_{kl}, P^{ij}_{kl})$ connected by transposition map and the map $M^{[ik][jl]}$ as shown in fig(1).

ii) We can set the coefficients $c_p$ of terms in Lagrangian differing by relabelling of the indices (i.e. same terms) to be equal to each other without loss of generality.

This implies that under the action of any transposition map $T^{ij}$ we must have $T^{ij}_{\text{induced}} : c_p \rightarrow -c_p$, if the 4'th derivative terms are to cancel. At the level of Lagrangian this implies that, interchanging any indices $b_i$ downstairs would flip the sign of the Lagrangian. Thus we have resolved the permutation coefficients $c_p$, that is to say, all the indices $b_i$ downstairs are anti-symmetrised. Now we can express $L_{2,q}, L_q$ as,

$$L_{2,q} \left[ A^{i_1i_2 \cdots i_{q+2}} \right] = A^{i_1i_2 \cdots i_{q+2}} \partial_{\alpha_i} \pi_{i_1} \partial^{i_i} \pi_{i_2} \partial^{i_2} \pi_{i_3} \cdots \partial^{i_{q+2}} \pi_{i_{q+2}}$$

$$L_q \left[ A^{i_1i_2 \cdots i_k} \right] = A^{i_1i_2 \cdots i_k} \partial_{\pi_{i_1}} \partial_{\pi_{i_2}} \cdots \partial_{\pi_{i_k}}$$

In order for the 3'rd derivative terms coming from the variation of the function $A$, $A'$ to cancel, we impose the following condition, $\frac{\partial A^{i_1 \cdots i_n}}{\partial X_{k_{i_j}}} \pi_{k_{i_j}}$ is symmetric in all of its indices $i_1, \ldots, i_n, k, l$.

Finally, generalising to multiple fields, a recursive relation shown in [9], we find,

$$L_{2,q} \left[ B^{i_1 \cdots i_n} \right] = -L_{0,q} \left[ B^{i_1 \cdots i_n} \right] - 2L_{0,q-1} \left[ \frac{\partial B^{i_1 \cdots i_n}}{\partial \pi_k} X_{k_{i_1}} \right] + (n-1)L_{2,q-1} \left[ \frac{\partial B^{i_1 \cdots i_n}}{\partial \pi_k} \right]$$
Where $B^{i_1 \ldots i_n}$ is defined by a line integral along an arbitrary curve in the $Y_{ij}$ space ending at point taking values in $\bar{X}_{ij}'$. Namely,

$$B^{i_1 \ldots i_n} = \int_{X_{ij}} G_{i_1 \ldots i_n k l} (Y_{mn}, \pi_r) dY_{kl}$$

(37)

On account of this recursive relationship we can see that $L_{2,q}$ is in fact a linear combination of $L_k$ with $k \in \{0 \ldots q\}$. Further, for $q = D$ we have,

$$L_{0,D} [A^{i_1 \ldots i_{D+2}} \bar{X}_{ij}] = \frac{D}{2} L_{2,D} [A^{i_1 \ldots i_{D+2}}]$$

(38)

Thus we can see from [30] that $L_{0,D}$ is a linear combination of $L_{0,k}$ with $k \in \{0 \ldots (D - 1)\}$. Hence, we have established our final result,

**Theorem 6.** The most general multiple scalar field theory in flat space-time satisfying conditions (i), (ii) is given by,

$$S_{\text{multi-scalar}} = \int_{\mathcal{M}} d^Dx A(\bar{X}_{ij}, \pi_i) + \sum_{m=1}^{D-1} A^{k_1 \ldots k_m} (\bar{X}_{ij}, \pi_i) \partial^{k_1} \partial_{a_1} \pi_{k_1} \ldots \partial^{k_m} \partial_{a_m} \pi_{k_m}$$

where $\bar{X}_{ij} = \frac{1}{2} \partial_{a_i} \pi_{j} \partial_{a_j} \pi_{i}$ and $\frac{\partial A^{i_1 \ldots i_m}}{\partial X_{kl}}$ is symmetric in all of its indices $i_1 \ldots i_m$, $k, l$.

[1] G. R. Dvali, G. Gabadadze and M. Porrati, Phys. Lett. B 485 (2000) 208 [hep-th/0005016].

[2] A. Nicolis and R. Rattazzi, JHEP 0406 (2004) 059 [hep-th/0404159].

[3] P. Creminelli, K. Hinterbichler, J. Khoury, A. Nicolis and E. Trincherini, JHEP 1302 (2013) 006 [arXiv:1209.3768 [hep-th]].

[4] K. Hinterbichler, A. Joyce, J. Khoury and G. E. J. Miller, arXiv:1212.3607 [hep-th].

[5] V. A. Rubakov, arXiv:1305.2614 [hep-th].

[6] A. Padilla and V. Sivanesan, JHEP 1304 (2013) 032 [arXiv:1210.4026 [gr-qc]].

[7] C. Deffayet and D. A. Steer, arXiv:1307.2150 [hep-th].

[8] G. W. Horndeski, Int. J. Theor. Phys. 10 (1974) 363-384.

[9] C. Deffayet, X. Gao, D. A. Steer and G. Zahariade, Phys. Rev. D 84 (2011) 064039 [arXiv:1103.3260 [hep-th]].

[10] M. Ostrogradsky, Memoires de lAcademie Imperiale des Science de Saint-Petersbourg, 4:385, 1850.

[11] A. Nicolis, R. Rattazzi and E. Trincherini, Phys. Rev. D 79 (2009) 064036 [arXiv:0811.2197 [hep-th]].

[12] A. De Felice, R. Kase and S. Tsujikawa, Phys. Rev. D 85 (2012) 044059 [arXiv:1111.5090 [gr-qc]]. R. Kimura, T. Kobayashi and K. Yamamoto, Phys. Rev. D 85 (2012) 024023 [arXiv:1111.6749 [astro-ph.CO]].

[13] A. I. Vainshtein, Phys. Lett. B 39 (1972) 393.

[14] A. Nicolis, R. Rattazzi and E. Trincherini, JHEP 1005 (2010) 095 [Erratum-ibid. 1111 (2011) 128] [arXiv:0912.4258 [hep-th]].

A. Masoumi and X. Xiao, arXiv:1201.3132 [hep-th]. S. -Y. Zhou, Phys. Rev. D 85 (2012) 104005 [arXiv:1202.5769 [hep-th]].