Darboux's Theorem, Lie series and the standardization of the Salerno and Ablowitz-Ladik models[⋆,⋆⋆]

Marco Calabrese[b], Simone Paleari[a,c], Tiziano Penati[a,c,*]

[a] *Department of Mathematics "F.Enriques", University of Milan, via Saldini 50, Milan, 20133, Italy*
[b] *Department of Mathematics and Statistics, University of Massachussets,, Amherst, 01003-9305, MA, USA*
[c] *GNFM (Gruppo Nazionale di Fisica Matematica) – Indam (Istituto Nazionale di Alta Matematica "F. Severi"), Roma, Italy*

Abstract

In the framework of nonlinear Hamiltonian lattices, we revisit the proof of Moser-Darboux's Theorem, in order to present a general scheme for its constructive applicability to Hamiltonian models with non-standard symplectic structures. We take as a guiding example the Salerno and Ablowitz-Ladik (AL) models: we justify the form of a well-known change of coordinates which is adapted to the Gauge symmetry, by showing that it comes out in a natural way within the general strategy outlined in the proof. Moreover, the full or truncated Lie-series technique in the extended phase-space is used to transform the Salerno model, at leading orders in the Darboux coordinates: thus the dNLS Hamiltonian turns out to be a normal form of the Salerno and AL models; as a byproduct we also get estimates of the dynamics of these models by means of dNLS one. We also stress that, once it is cast into the perturbative approach, the method allows to deal with the cases where the explicit trasformation is not known, or even worse it is not writable in terms of elementary functions.

Keywords: Darboux's Theorem, non linear chains, Lie-series technique, Ablowitz-Ladik and Salerno models, non standard symplectic form, discrete Nonlinear Schroedinger

1. Introduction

On a symplectic manifold (M, ω) of dimension $2n$, Darboux's Theorem of symplectic geometry [4] ensures the local existence of a set of coordinates, say (q_j, p_j), $j = 1, \ldots, n$, such that at any point $P \in M$ the symplectic 2-form ω reads $\omega_P = \sum_j dq_j \wedge dp_j$. The existence of such a standard set of coordinates is useful for the Hamiltonian formalism, and for Hamiltonian perturbation theory in particular (see [7]), since Hamiltonian equations, Hamiltonian vector fields X_H,

⋆The present version is the preprint one; for the accepted version on Physica D, please refere to the official page https://doi.org/10.1016/j.physd.2024.134183

*Corresponding author

Email addresses: Mcalabrese@umass.edu (Marco Calabrese), simone.paleari@unimi.it (Simone Paleari), tiziano.penati@unimi.it (Tiziano Penati)

symmetries and canonical transformations can be expressed in terms of Poisson brackets $\{\cdot,\cdot\}$ via the (constant) symplectic matrix J as follows

$$\dot{z}_j = (X_H(z))_j = \{z_j, H\} \qquad \{f, g\} = (\nabla f)^\top J \nabla g\,, \qquad J = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}, \tag{1}$$

where $\mathbb{I}$ represents the n-dimensional identity matrix and the gradient ∇ is intended with respect to the set of real coordinates $z = (q_j, p_j)$.

A classical example, in the field of Hamiltonian Lattices (see [21] for a recent review on the topic), is the discrete[1] Nonlinear Schrödinger model (dNLS)

[dNLS]
$$i\dot{\psi}_j = \varepsilon(\psi_{j+1} + \psi_{j-1}) + \gamma|\psi_j|^2\psi_j\,, \tag{2}$$

where $\psi_j \in \mathbb{C}$ and the lattice-index j may run in a finite $\mathcal{J} = \{1, \dots, N\}$ (with periodic or fixed boundary conditions) or infinite $\mathcal{J} = \mathbb{Z}$ sets (with ℓ^2-decay of the sequence $\{\psi_j\}_{j\in\mathcal{J}}$). It turns out ot be an Hamiltonian system with Hamiltonian given by

[$\mathbf{H_{dNLS}}$]
$$H(\psi) = \sum_{j\in\mathcal{J}} \left[\varepsilon(\psi_{j+1}\bar{\psi}_j + \bar{\psi}_{j+1}\psi_j) + \frac{\gamma}{2}|\psi_j|^4\right], \tag{3}$$

with the *standard Poisson structure* $\{\psi_j, H\} = -i\frac{\partial H}{\partial \bar{\psi}_j}$ in complex variables.

However, it might happen that the dynamics of a physical model is described by a vector field $X_H(z)$, with physical coordinates z, which is derived by the Hamiltonian $H(z)$ through a non-standard representation of the Poisson brackets $\{\cdot,\cdot\}$; in this case the 2-form ω is locally represented by a different (typically non constant) matrix $\Omega(z)$, such that $\omega_P(X, Y) = X^\top\Omega(z)Y$, being $\Omega(z)$ antisymmetric and non-degenerate. This is the case of the Ablowitz-Ladik (AL in the following) system

[AL]
$$i\dot{\psi}_j = \varepsilon(1 + \mu|\psi_j|^2)(\psi_{j+1} + \psi_{j-1})\,, \tag{4}$$

which is a celebrated integrable discretization of the NLS, and of the Salerno models

[Salerno]
$$i\dot{\psi}_j = \varepsilon(1 + \mu|\psi_j|^2)(\psi_{j+1} + \psi_{j-1}) + \gamma|\psi_j|^2\psi_j\,; \tag{5}$$

the two common parameters ε and μ can be taken as positive, while γ of any sign. It is well known that both models are Hamiltonian and share the same *nonstandard symplectic structure* given by the Poisson brackets

$$\dot{\psi}_j = (X_H)_j(\psi) = \{\psi_j, H\} \qquad \{\psi_j, H\} = -i(1 + \mu|\psi_j|^2)\frac{\partial H}{\partial \bar{\psi}_j}\,, \tag{6}$$

while the Hamilton function reads

$$H(\psi) = \sum_{j\in\mathcal{J}} \left[-\frac{\gamma}{\mu^2}\ln(1 + \mu|\psi_j|^2) + \varepsilon(\psi_{j+1}\bar{\psi}_j + \bar{\psi}_{j+1}\psi_j) + \frac{\gamma}{\mu}|\psi_j|^2\right]; \tag{7}$$

for $\gamma = 0$ we recover the AL model (4) while $\gamma \neq 0$ gives the Salerno model (5).

[1] It is indeed the most common discretization of the continuous Nonlinear Schroedinger Equation (NLS) model.

The two models also share a second conserved quantity

$$P = \frac{1}{\mu}\sum_{j\in\mathcal{J}} \ln(1+\mu|\psi_j|^2) \qquad \{H,P\} = 0\,, \tag{8}$$

which is related to the Gauge symmetry $e^{i\theta}$ of the equation; indeed, the flow of the Hamiltonian vector field $(X_P)_j = \{\psi_j, P\} = -i\psi_j$ is exactly given by the action of $e^{i\theta}$.

One of the reasons why it is preferable to standardize the symplectic structure (6) — which by the way coincides with the standard one when $\mu = 0$ — is the possibility of linking homotopically the AL to the dNLS at the Hamiltonian level, by setting $\gamma = 1-\mu$, with $\mu \in [0,1]$; indeed the Salerno model interpolates between the AL model (for $\gamma = 0$) and the dNLS (for $\gamma \neq 0$ and $\mu \to 0$), and a common symplectic structure might be preferable in order to explore the transition (AL-Salerno-dNLS) along the three models, especially in terms of dynamical features (see [12, 13] for a comparison between AL and dNLS in terms of vector fields and persistence of localized structures, see instead [18] for a comparison between AL and Salerno in terms of additional conserved quantities). On the other hand, standardization of (6) might be helpful to implement geometric numerical schemes which are more suitable for the integrability of AL (see for example [24]).

Motivated by the above mentioned issues concerning those Hamiltonian non linear lattices, the present manuscript aims at providing a new insight into the classical problem of applying Darboux's Theorem to specific models, when a given Hamiltonian system has to be explicitly transformed into Darboux's coordinates.

As a first result (see Theorem 2.1) we show that one of most used Darboux's transformation for AL and Salerno models (see again [24] and [11]), i.e.

$$\psi_j = \Psi_j\sigma\left(\frac{\mu}{2}|\Psi_j|^2\right) \qquad \text{with} \quad \sigma(s) = \sqrt{\frac{\exp(s)-1}{s}}\,, \tag{9}$$

can be derived directly from Moser's scheme of the proof (originally in [19], here taken from [17]). We implement such a transformation with the equivalent procedure of the Lie-series (see the classical works [5, 14, 10]): this method relies on the idea that, since the change of coordinate suggested by the proof is the flow of a vector field V at a given time $t = 1$, the transformed Hamiltonian systems can be obtained as a (totally convergent) series of iterated Lie derivatives $L_V H$ along V. In the scheme of Moser, such a vector field is time dependent, hence the Lie-series representation of (9) is performed in the extended $\mathbb{R}\times\mathbb{R}^{2n}$ phase-space

$$H(\Psi\sigma(\Psi)) = \exp\left(-L_{\tilde{V}}H\right)(\Psi) = \sum_{l\geq 0}(-1)^l L_{\tilde{V}}^l H(\Psi)\,,$$

where $L_{\tilde{V}}H$ is the Lie-derivative of H along the extended vector field $\tilde{V} = (1, V)$.

As a second result we quantitatively compare the models under investigation, both at the level of the Hamiltonians, in the spirit of normal forms, and at the level of solutions. Indeed, once put in the transformed coordinates Ψ, the AL and Salerno model can be compared to the dNLS dynamics by exploiting the Hamiltonian formalism, since they all share the same symplectic structure, at least in a small neighborhood of the origin. We thus provide estimates of the closeness between the AL and the dNLS model, as well as between the Salerno model (11), or a suitable cubic-quintic generalization (see Theorem 2.2 for a more precise formulation), and the dNLS: in all the cases small norm initial data have to be considered, since the Darboux transformation is only local. Similar results have been already obtained in [12]: however, at variance with these authors, we here make estimates at the level of the Hamiltonian formalism and with the use of Lie-series, rather than directly controlling the difference between the two vector fields.

A further result pertains the additional conserved quantity that all the models we are considering admit. With the above transformation (9), the AL, Salerno and dNLS models now share the same conserved quantity, since the quantity P, defined in (8), becomes the ℓ^2 norm (which is the additional conserved quantity of the dNLS model (3)). In terms of Lie-series, the integral given by the ℓ^2 norm has to coincide (modulo a prefactor) with $\exp(-L_{\tilde{V}}P)(\Psi) = \frac{1}{2}\|\Psi\|^2$. It is possible to prove (see Proposition 2.1) — and easily anticipate by means of a numerical evidence (using Mathematica) — a geometric convergence of the truncated Lie-series $\exp_K(-L_{\tilde{V}}P)(\Psi) = \sum_{l=0}^{K}(-1)^l L_{\tilde{V}}^l H(\Psi)$ to the ℓ^2 norm, for increasing values of K. Indirectly, this also numerically confirms the correct relationship between (9) and the (time dependent) vector field V explicitly computed going through Moser's scheme.

Clearly, in these specific applications, in particular for the second result, we might have avoided the use of Lie-series, being in fact equipped with the explicit form of Ψ. However, in the spirit of perturbation theory, we stress that if one is interested only in some leading order approximation (normal form) of the transformed Hamilton function $H(\Psi\sigma(\Psi))$, it is enough to simply construct a proper polynomial approximation $\tilde{V}^{(L)}$ of $\tilde{V}$, which can be then used to transform H through the (truncated) Lie-series $\exp_K(L_{\tilde{V}^{(L)}}H)$; it is indeed possible, with some standard analytical estimates of Lie-series, to derive a priori bounds of the error $|\exp(L_{\tilde{V}}H) - \exp_K(L_{\tilde{V}^{(L)}}H)|$. And in fact such a procedure can be exploited even in cases where the explicit form of the transformation is not known or does not exist in terms of elementary functions (see [16] for a related approach).

The scheme of the manuscript is the following. Section 2 will be devoted to present the results of this paper: we show the emergence of the aforementioned transformation from the abstract geometric scheme of the proof and provide normal form statements in the framework of the Lie-series. Section 3 includes the proofs of the results. Some additional comments and future perspectives are included in the Conclusions, Section 4. In Appendix A we recall some analytical results on Lie-series in the non-autonomous case. In Appendix B we review Moser's proof and we extract the constructive scheme.

2. Results

In order to find a set of Darboux coordinates, according to the scheme of the proof here reviewed, we prefer to work with a set of canonical and cartesian variables $\{q_j, p_j\} \in \mathbb{R}^{2N}$ defined by

$$\psi_j = \frac{1}{\sqrt{2}}(q_j + ip_j) \qquad\qquad \overline{\psi}_j = \frac{1}{\sqrt{2}}(q_j - ip_j) \; ;$$

from now on we assume $\mathcal{J}$ finite with cardinality N, but the results here presented work also in the infinite lattice. We start rewriting the Poisson brackets of two smooth functions F, G as

$$\{F, G\} = \sum_j \left[1 + \nu(q_j^2 + p_j^2)\right]\{\partial_{q_j}F\partial_{p_j}G - \partial_{p_j}F\partial_{q_j}G\} \, ,$$

where $\nu = \frac{1}{2}\mu$ and the Poisson brackets are related to the 2-form ω_0

$$\omega_0(q, p) = \sum_j \frac{1}{1 + \nu(q_j^2 + p_j^2)} dq_j \wedge dp_j \, .$$

We notice that at the origin ω_0 coincides with the standard[2] symplectic form ω_1

$$\omega_1 = \sum_j dq_j \wedge dp_j = \omega_0(0,0)\ . \tag{10}$$

The AL and Salerno Hamiltonian can be then rewritten in cartesian variables

$$H = H_0 + H_1 \qquad \begin{cases} H_0 = & \sum_{j\in\mathcal{J}} \left[-\frac{\gamma}{4\nu^2}\ln[1+\nu(q_j^2+p_j^2)] + \frac{\gamma}{4\nu}(q_j^2+p_j^2) \right] \\ H_1 = \varepsilon \sum_{j\in\mathcal{J}} (q_{j+1}q_j + p_{j+1}p_j) \end{cases}\ , \tag{11}$$

and the corresponding Hamilton equations are

$$\begin{cases} \dot{q} = & \varepsilon(p_{j+1}+p_{j-1})\Big(1+\nu(q_j^2+p_j^2)\Big) + \frac{1}{2}p_j\gamma(q_j^2+p_j^2) \\ \dot{p} = & -\varepsilon(q_{j+1}+q_{j-1})\Big(1+\nu(q_j^2+p_j^2)\Big) - \frac{1}{2}q_j\gamma(q_j^2+p_j^2)\ ; \end{cases}$$

in the same set of coordinates, the Hamiltonian of the dNLS model (2) reads

$$H = H_0 + H_1 \qquad \begin{cases} H_0 = \frac{\gamma}{8}\sum_{j\in\mathcal{J}} (q_j^2+p_j^2)^2 \\ H_1 = \varepsilon \sum_{j\in\mathcal{J}} (q_{j+1}q_j + p_{j+1}p_j) \end{cases}\ , \tag{12}$$

with Hamilton equations given by

$$\begin{cases} \dot{q} = & \varepsilon(p_{j+1}+p_{j-1}) + \frac{1}{2}p_j\gamma(q_j^2+p_j^2) \\ \dot{p} = & -\varepsilon(q_{j+1}+q_{j-1}) - \frac{1}{2}q_j\gamma(q_j^2+p_j^2)\ ; \end{cases}$$

Since the Darboux transformation acts in the same way on each two-dimensional subspace with coordinates (q_j, p_j), we can restrict to the 2-dimensional manifold $\mathbb{R}^2$ and consider ω_0 as

$$\omega_0 = \frac{1}{1+\nu(q^2+p^2)} dq \wedge dp\ ,$$

by omitting all the indexes; in this way, denoting again by J the restriction of (1) on $\mathbb{R}^2$, one has

$$\Omega(q,p) = \frac{1}{1+\nu(q^2+p^2)} J\ , \tag{13}$$

and the additional conserved quantity (8) takes the form (on the 2-dimensional subspace)

$$P(q,p) = \frac{1}{2\nu}\ln[1+\nu(q^2+p^2)]\ . \tag{14}$$

As anticipated in (9), the nonlinear change of coordinates $(q,p) = \varphi^{-1}(x,y)$ with φ^{-1} given by

$$\begin{cases} q = x\sigma(\nu\,\|(x,y)\|^2) \\ p = y\sigma(\nu\,\|(x,y)\|^2) \end{cases}\ , \tag{15}$$

is a Darboux transformation[3] and it transforms P into the ℓ^2 norm (modulo a prefactor $\frac{1}{2}$)

$$P \circ \varphi^{-1}(x,y) = \frac{1}{2}(x^2+y^2). \tag{16}$$

[2]As will be briefly explained in Appendix B, it is always possible to set $\omega_0(P) = \omega_1$ at a given point P, with a linear change of coordinates.

[3]Indeed it satisfies $J = (D_{x,y}\varphi)^\top \Omega(\varphi(x,y))(D_{x,y}\varphi)$.

2.1. First result: Darboux's change of coordinates

As a first main result we show the relationship between (15) and Moser's scheme of the proof, where the last can be summed up in the three main steps:

1. find a vector potential $a(q, p) = (a_1, a_2)$ for the closed 2-form $\omega_0 - \omega_1$, namely such that

$$\partial_p a_1 - \partial_q a_2 = \frac{\nu(q^2 + p^2)}{1 + \nu(q^2 + p^2)} \; ; \tag{17}$$

2. compute the vector field $V_t(q, p)$

$$V_t(q, p) = \Omega_t^{-\top}(q, p) a(q, p) \, , \tag{18}$$

   where, by taking $t \in [0, 1]$, the matrix $\Omega_t(q, p)$ interpolates between $\Omega(q, p)$ (non-standard symplectic structure given by (13)) and J (the standard one)

$$\Omega_t(q, p) = tJ + (1 - t)\Omega(q, p) \qquad\qquad \Omega_1 = J \qquad\qquad \Omega_0 = \Omega \, .$$

3. solve (if possible) the dynamical system

$$(\dot{q}, \dot{p}) = V_t(q, p) \, ,$$

   whose time-one-flow $\Phi^1(q, p)$ defines the Darboux transformation $(x, y) = \varphi(q, p) = \Phi^1(q, p)$.

In the following, we make use of the Lie-series formalism to represent $\Phi^1(q, p)$; in particular, being the vector field V_t time dependent, we introduce the extended vector field $\tilde{V} = (1, V_t)$ defined in the extended phase space $(\tau, q, p) \in \mathbb{R}^3$ and the corresponding Lie-derivative of a analytic function $f : \mathbb{R}^3 \to \mathbb{R}$

$$L_{\tilde{V}} f = \partial_\tau f + \langle V_t(q, p), \nabla f(q, p) \rangle \, . \tag{19}$$

The extended field $\tilde{V}$ defines a flow $\tilde{\Phi}^t : \mathbb{R}^3 \to \mathbb{R}^3$ which can be expressed in term of the Lie-series operator $\exp(tL_{\tilde{V}})$

$$\tilde{\Phi}^t(\tau, q, p) = \exp(tL_{\tilde{V}}\mathbb{I})(\tau, q, p) = \sum_{k \geq 0} \frac{t^k}{k!}(L^k_{\tilde{V}}\mathbb{I})(\tau, q, p) \, , \tag{20}$$

where $\mathbb{I} : \mathbb{R}^3 \to \mathbb{R}^3$ is the identity map and $L^k_{\tilde{V}}\mathbb{I}$ has to be interpreted as $(L^k_{\tilde{V}}\mathbb{I})_j = L^k_{\tilde{V}}\mathbb{I}_j$, being $\mathbb{I}_j$ the identity on the j_{th} component. It turns out that

$$\tilde{\Phi}^t(0, q, p) = (t, \Phi^t(q, p)) \qquad \Longrightarrow \qquad \tilde{\Phi}^1(0, q, p) = (1, x, y) = (1, \varphi(q, p)) \, ,$$

and consequently

$$\tilde{\Phi}^{-1}(1, x, y) = \sum_{k \geq 0} \frac{(-1)^k}{k!}(L^k_{\tilde{V}}\mathbb{I})(1, x, y) = (0, q, p) = (0, \varphi^{-1}(x, y)) \, .$$

Theorem 2.1. *Let $V_t(q,p)$ be the time dependent vector filed given by*

$$V_t(q,p) = \chi(t, \nu \|(q,p)\|^2) \begin{pmatrix} q \\ p \end{pmatrix} \tag{21}$$

with

$$\chi(t, \nu \|(q,p)\|^2) = \frac{1+\nu \|(q,p)\|^2}{2(1+t\nu \|(q,p)\|^2)} \left[\frac{\ln(1+\nu \|(q,p)\|^2)}{\nu \|(q,p)\|^2} - 1 \right], \tag{22}$$

and consider its extension $\tilde{V}(\tau, q, p) = (1, V_t(q,p)) : \mathbb{R}^3 \to \mathbb{R}^3$. Then, the time-one-flow $\tilde{\Phi}^1(0,q,p) = (1, \varphi(q,p))$ generated by $\tilde{V}$ is well defined in a sufficiently small ball $B_\rho(0) = \{\|(q,p)\| < \rho\}$ of the origin. Moreover, the Darboux change of coordinates (15) *corresponds to the inverse transformation $(q,p) = \varphi^{-1}(x,y)$, where φ is a near-to-the-identity and analytic map with the following asymptotic expansion*

$$\varphi(q,p) \sim \left[1 - \frac{1}{4}\nu \|(q,p)\|^2 \right] \begin{pmatrix} q \\ p \end{pmatrix} . \tag{23}$$

Remark 2.1. *In agreement with* (23)*, we observe that* (15) *has the asymptotic expansion $\varphi^{-1}(x,y) \sim (x,y) + \frac{1}{4}\nu \|(x,y)\|^2 (x,y)$.*

Some other remarks on Theorem (2.1) are in order:

1. the vector field V_t is equivariant under the same Gauge symmetry of the models (11). This has been made more evident by a proper choice of the vector $a(q,p)$, solution of (17). Indeed the vector potential $a(q,p)$ in (17) is defined modulo a gradient ∇f of a scalar function f;
2. in general, we cannot expect to write the flow of the dynamical system given by V_t; but we can prove that in a small neighbourhood of the origin the flow $\Phi^t(q,p)$ is radial and close to the identity (and contracting). Hence, we can impose a precise structure to the unknown transformation $\Phi^1(q,p)$ (and to its inverse), which leads to (15) as its unique analytic solution;
3. the vector field V_t is asymptotically cubic for (q,p) in a sufficiently small neighborhood of the origin. This can be understood by observing that V_t is constructed from the nonlinear deformation of ω_0 with respect to ω_1. Indeed the vector $a(q,p)$ is obtained integrating the quadratic deformation $\frac{\nu(q^2+p^2)}{1+\nu(q^2+p^2)}$ in (17). As a consequence, the flow $\Phi^t(q,p)$ is a nonlinear deformation of the identity map in $B_\rho(0)$.

2.2. Second result: dNLS-like normal forms

The approach of Lie-series allows to transform any Hamiltonian system, once given the generator vector field V_t. Indeed, the explicit expression of the flow $\tilde{\Phi}^1$ is not necessary and the Lie-series $\exp(\pm L_{\tilde{V}})$ operator suffices to write the Hamilton equations in the transformed variables (x,y). However, the knowledge of V_t depends on the possibility to provide the vector potential $a(q,p)$ through an explicit integration: in the spirit of perturbation theory, the leading order approximation of the transformed Hamiltonian $\exp(-L_{\tilde{V}}H) = H \circ \varphi^{-1}$ can be obtained by a suitable truncation of V_t, hence by a polynomial approximation of the vector potential $a(q,p)$, which is always an accessible task. In order to formulate the next statement, we introduce the following notation for the Taylor expansion of H in (11) and of V_t in (21):

$$H_0 = \sum_{l\geq 2} H_{0,2l} \qquad H_1 = H_{1,2}\,, \qquad V_t = \sum_{l\geq 1} V_{t,2l+1}\,.$$

where the second index s in $H_{j,s}$ represents the polynomial degree of $H_{j,s}$ in the variables (q, p), while the first index j is the degree with respect to the parameter ε. Furthermore, $\tilde{V}_{2l+1}$ will denote the extension of $V_{t,2l+1}$ and $\Phi^t_H(x_0, y_0)$ denotes the Hamiltonian flow associated to H at time t, with initial datum $(x_0, y_0) \in \mathbb{R}^{2n}$.

Theorem 2.2. *In a sufficiently small ball $B_\rho(0) \subset \mathbb{R}^{2n}$ of the origin, the Hamiltonian* (11) *can be transformed by the inverse Lie-series along $\tilde{V}$*

$$H \circ \varphi^{-1}(x, y) = \exp(-L_{\tilde{V}} H) = \sum_{k\geq 0} \frac{(-1)^k}{k!} (L^k_{\tilde{V}} H)(1, x, y)\,.$$

At leading order (in ρ and ε), the Salerno model (11) *can be approximated by the dNLS Hamiltonian* (12)

$$\begin{aligned}\exp(-L_{\tilde{V}} H) &= Z^{(0)} + \mathcal{R}^{(0)}\\ Z^{(0)} &= H_{0,4} + H_{1,2} = \sum_{j\in\mathcal{J}} \left[\frac{\gamma}{8}\Big(x_j^2 + y_j^2\Big)^2 + \varepsilon\Big(x_{j+1}x_j + y_{j+1}y_j\Big)\right],\end{aligned}$$

where the remainder $\mathcal{R}^{(0)}$ satisfies

$$\sup_{(x,y)\in B_\rho} |\mathcal{R}^{(0)}(x, y)| \leq C_0 \nu \rho^4 (\rho^2 + \varepsilon)\,, \qquad C_0 > 0\,,$$

and for times $|t| \leq (\rho^2 + \varepsilon)^{-1}$ one has

$$\left\|\Phi^t_H(x_0, y_0) - \Phi^t_{Z^{(0)}}(x_0', y_0')\right\| \leq c_0\Big(\left\|(x_0, y_0) - (x_0', y_0')\right\| + \rho^3\Big) \qquad c_0 > 0\,. \tag{24}$$

At next order, the Salerno model (11) *admits the following cubic-quintic normal form*

$$\begin{aligned}\exp(-L_{\tilde{V}} H) &= Z^{(1)} + \mathcal{R}^{(1)}\\ Z^{(1)} &= H_{0,4} + H_{1,2} - L_{\tilde{V}_3} H_{0,4} - L_{\tilde{V}_3} H_{1,2} =\\ &= \sum_{j\in\mathcal{J}} \left[\frac{\gamma}{8}\Big(x_j^2 + y_j^2\Big)^2 + \varepsilon\Big(x_{j+1}x_j + y_{j+1}y_j\Big)\right] +\\ &+ \sum_{j\in\mathcal{J}} \left[\frac{\gamma}{24}\nu\Big(x_j^2 + y_j^2\Big)^3 + \frac{1}{4}\varepsilon\nu\Big(x_j^2 + y_j^2\Big)\Big((x_{j+1} + x_{j-1})x_j + (y_{j+1} + y_{j-1})y_j\Big)\right],\end{aligned} \tag{25}$$

where the remainder $\mathcal{R}^{(1)}$ satisfies

$$\sup_{(x,y)\in B_\rho} |\mathcal{R}^{(1)}(x, y)| \leq C_1 \nu^2 \rho^6 (\rho^2 + \varepsilon)\,, \qquad C_1 > 0\,,$$

and for times $|t| \leq (\rho^2 + \varepsilon)^{-1}$ one has

$$\left\|\Phi^t_H(x_0, y_0) - \Phi^t_{Z^{(1)}}(x_0', y_0')\right\| \leq c_1\Big(\left\|(x_0, y_0) - (x_0', y_0')\right\| + \rho^5\Big) \qquad c_1 > 0\,. \tag{26}$$

The dNLS Hamiltonian $Z^{(0)}$ is a normal form also for the AL model, with a remainder $\mathcal{R}^{(0)}$ satisfying

$$\sup_{(x,y)\in B_\rho} |\mathcal{R}^{(0)}(x,y)| \leq C_0'\rho^4(1+\varepsilon)\ , \qquad C_0' > 0\ ;$$

for times $|t| \leq \varepsilon^{-1}$ one has

$$\left\|\Phi^t_H(x_0,y_0) - \Phi^t_{Z^{(0)}}(x_0',y_0')\right\| \leq c_0'\left(\left\|(x_0,y_0)-(x_0',y_0')\right\| + \frac{\rho^3}{\varepsilon}\right) \qquad\qquad c_0' > 0\ . \tag{27}$$

Some remarks on Theorem (2.2) are also in order:

1. the dNLS approximation of the AL dynamics given by formula (27), for small (in norm) initial data, is in agreement with the first (and main) statement in [12]; indeed, for $\varepsilon = O(1)$, our estimate (27) claims that two initial data which are ρ^3-close and of order ρ, stay ρ^3-close for times of order $|t| \leq O(1)$. In this case, our proof can be adapted in order to consider a fixed time scale $|t| \leq T$ with arbitrary $T > 0$; however, as in [12], the price to pay would be a constant $c_0'(T)$ which is increasing with T. At variance with [12], we do not need to impose any condition on $P(0)$, since the AL dynamics is compared to the dNLS one only after the Darboux transformation has been performed: hence the two models share the same conserved quantity, namely the norm. In fact, the smallness condition on $P(0)$ in [12] is asked in order to ensure that the AL flow keeps its norm bounded for all times: this is obtained for free with our approach;
2. estimates (24) and (26) provide two different approximations of the Salerno model ($\gamma \neq 0$) on the same time scale $O((\rho^2+\varepsilon)^{-1})$. The time scale suggests to link ε to ρ in the regime $\varepsilon \sim \rho^2 \ll 1$, so that the two statements are formulated only in terms of energy (or amplitude). This regime is the typical one for which the dNLS is the normal form of the Klein-Gordon chain (see for example [20, 22]);
3. it is clear by applying the transformation (15) to the model (11) (by truncating the Taylor expansion at the identity) that the dNLS (3) is the first approximation of (11), when $\gamma \neq 0$. However, we here want to derive different levels of approximation of the model by exploiting the Lie-series method and the expansion of the vector field V_t, without any need of knowing the exact shape either of φ or of V_t. Furthermore, the correctness of the expansion of $H\circ\varphi^{-1}$ can be directly verified thanks to the explicit knowledge of φ in (15);
4. as it is usual in the non autonomous case, we move to the extended phase space, so to apply the standard Lie-series operator (20) to transform the Hamiltonian in the new Darboux coordinates. Since by construction we already know the symplectic form in the new coordinates, it is enough to transform the Hamiltonian in order to get the Hamilton equations in the new set of variables;
5. as already stressed at the beginning of this subsection, we remark that, if we are interested in a leading order expansion of the transformed Hamiltonian $\exp(L_{\tilde{V}}H)$, it might be enough to truncate $\tilde{V}$ at a suitable polynomial (or perturbative, whatever is the small parameter in the expansion) order L; the required order L can be determined on the base of the error $|\exp(L_{\tilde{V}}H) - \exp(L_{\tilde{V}^{(L)}}H)|$, which can be apriori estimated by exploiting Proposition Appendix A.2 in Appendix A.

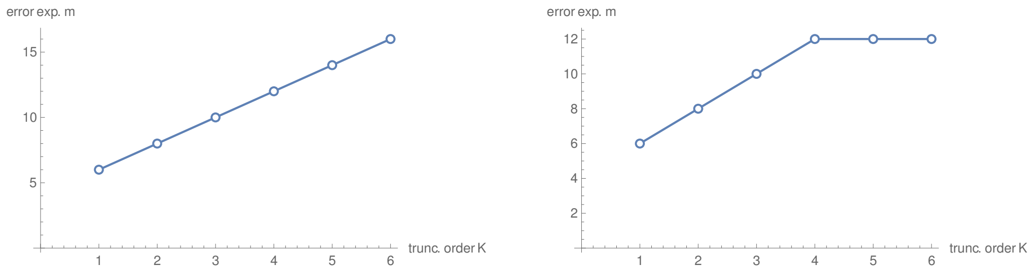

Figure 1: Order m of the truncated Lie-series error $O(\rho^m)$ as a function of the truncation order K. Left panel: only truncation of the Lie-series is performed, hence the decay of the remainder is of the order $O(\rho^{2K+4})$. Right panel: truncation of the vector field $\tilde{V}^{(2L+1)}$ is added (with $L = 4$), hence for $K \geq 4$ the error is kept constantly equal to $O(\rho^{12})$.

2.3. Third result: numerical evidence and the Lie-series of P.

In order to apply the Lie-series method to transform the conserved quantity P given in (14), and in general to transform any function $f(q, p)$ through $\varphi^{-1}(x, y)$, we need the inverse flow $\tilde{\Phi}^{-1}(1, x, y)$, which is indeed the flow of the opposite field $-\tilde{V}(\tau, x, y)$. As a consequence, the transformed quantity $P(\phi^{-1}(x, y)) = P(\tilde{\Phi}^{-1}(1, x, y))$ can be obtained from the Lie-series of P along the field $-\tilde{V}$ and has to coincide with (16)

$$P(\tilde{\Phi}^{-1}(1, x, y)) = \exp(L_{-\tilde{V}}P)(1, x, y) = \sum_{k\geq 0} \frac{(-1)^k}{k!} L^k_{\tilde{V}} P(1, x, y) = \frac{1}{2}\, \|(x, y)\|^2 \ .$$

This provides us with the possibility of numerically verifying the convergence of the Lie series and estimating the approximation errors, when the series is stopped at a certain order K or when the vector field is approximated with a suitable Taylor polynomial of order $2L + 1$. Let $V^{(2L+1)} = \sum_{l=1}^{L} V_{t,2l+1}$ be the Taylor polynomial at order $2L + 1$ of V_t, and denote by $\exp_K(tL_{\tilde{V}}) = \sum_{k=0}^{K} \frac{t^k}{k!} L^k_{\tilde{V}}$ the Lie-series truncated at order K; then the following result holds true:

Proposition 2.1. *In a sufficiently small ball $B_\rho(0) \subset \mathbb{R}^{2n}$ of the origin, the approximations of the Lie-series $\exp(L_{-\tilde{V}}P)(1, x, y)$ satisfy the following error estimates*

$$\begin{aligned} \sup_{(x,y)\in B_\rho} \left|\frac{1}{2}\, \|(x, y)\|^2 - \exp_K(L_{-\tilde{V}}P)(1, x, y)\right| &\leq C_1 \nu^{K+1} \rho^{2K+4} \\ \sup_{(x,y)\in B_\rho} \left|\frac{1}{2}\, \|(x, y)\|^2 - \exp_K(L_{-\tilde{V}^{(2L+1)}}P)(1, x, y)\right| &\leq C_2 \nu^{M+1} \rho^{2M+4} \ , \end{aligned} \tag{28}$$

with $M = \min\{L, K\}$ and suitable positive constants $C_{1,2}$.

Figure 1 shows the numerical computation[4] of the above errors for $K = 1, \ldots, 6$ and (in the right panel) $L = 4$. The exponent m of the leading polynomial term in the error is plotted against the truncation K. The left panel indicates a geometric convergence of $\exp(L^{(K)}_{\tilde{V}} P)(1, x, y)$ to (16) as K is increased, in agreement with the first estimate in (28); the right panel shows that for $K \geq 4$ the error is stabilized at values $O(\rho^{12})$, in agreement with the second estimate in (28), since $\min\{L, K\} = 4$ for such values of K.

[4]Performed with Mathematica Release 12

3. Proofs of Theorems

Given $x \in \mathbb{R}^n$, we define the polydisk $D_\rho = \bigotimes_{j=1,\ldots,n} \mathcal{B}_\rho(x_j) \subset \mathbb{C}^n$ as the product of n-copies of the complex disk $\mathcal{B}_\rho(x_j)$ with radius $\rho \le 1$ centered at the elements x_j. For $T > 1$ and $t \in \mathcal{G} = (-T, T)$, we also define the extended complex domain $\mathcal{G}_\delta = \bigcup_{t\in\mathcal{G}} \mathcal{B}_\delta(t) \subset \mathbb{C}$. Let $\tilde{X}$ be a time dependent and analytic vector field $\tilde{X}(t, x) : \mathcal{G}_\delta \times D_\rho \to \mathbb{R}^{n+1}$, and let f be a real-valued function $f(t, x) : \mathcal{G}_\delta \times D_\rho \to \mathbb{R}$ which is analytic in the same domain, then we introduce the following norms

$$\left|f\right|_{\delta,\rho} := \sup_{(\zeta,z)\in\mathcal{G}_\delta\times D_\rho} |f(\zeta, z)| \qquad \left|\tilde{X}\right|_{\delta,\rho} := \sup_{(\zeta,z)\in\mathcal{G}_\delta\times D_\rho} |X(\zeta, z)| ,$$

where $|\tilde{X}(\zeta, z)|$ is the norm of a complex vector in $\mathbb{C}^{n+1}$. Clearly, if f or $\tilde{X}$ depend only on the phase space variable z, we will consider only $\left|f\right|_\rho$ or $\left|\tilde{X}\right|_\rho$. Moreover, we use the symbol $\prec$ to compare to quantities, like the norms of two functions (or vector fields), modulo a numerical constant

$$A \prec B \qquad \Longleftrightarrow \qquad \exists C > 0 \quad \text{s.t.} \quad A \le CB . \tag{29}$$

3.1. Proof of Theorem 2.1

We first have to observe that (17) represents the integration of the closed 2-form $-\eta = \omega_0-\omega_1$, where $\omega_1 = dq \wedge dp$ and the unknown $a(q, p)$ is the potential vector of a 1-form α such that $d\alpha = -\eta$. Hence the problem reduces to find α_0 and α_1 such that

$$d\alpha_0 = \omega_0 \qquad\qquad d\alpha_1 = \omega_1 \qquad \Rightarrow \qquad \alpha = \alpha_0 - \alpha_1 .$$

We notice that the two 2-forms already coincide at the origin, $\omega_0(0, 0) = \omega_1$, hence no preliminary linear transformation is required. We look for a potential of ω_0. By passing to action-angle like coordinates

$$\begin{cases} q = \sqrt{A}\cos(\theta) \\ p = -\sqrt{A}\sin(\theta) \end{cases} \qquad\qquad \begin{cases} A = q^2 + p^2 \\ \theta = -\arctan\left(\frac{p}{q}\right) \end{cases}$$

one gets

$$dq \wedge dp = \frac{1}{2} d\theta \wedge dA .$$

We can rewrite ω_0 as

$$\omega_0 = \frac{1}{2(\nu A + 1)} d\theta \wedge dA$$

whose potential can be chosen as a $d\theta$-form $\alpha_0 = -\frac{1}{2\nu}\ln(1 + \nu A)d\theta$, or in cartesian coordinates (q, p)

$$\alpha_0 = \frac{\ln(1 + \nu\,\|(q, p)\|^2)}{2\nu\,\|(q, p)\|^2}\,(-pdq + qdp) .$$

The main point in the solution of (17) is to choose the potential α_1 of ω_1 as a $d\theta$-form as α_0, namely $\alpha_1 = -\frac{1}{2}Ad\theta$, or in cartesian coordinates $\alpha_1 = \frac{1}{2}(-pdq + qdp)$, so that

$$\alpha_0 - \alpha_1 = \frac{1}{2}\left[\frac{\ln(1 + \nu\,\|(q, p)\|^2)}{\nu\,\|(q, p)\|^2} - 1\right](-pdq + qdp) ,$$

which implies a solution $a(q, p)$ of (17) of the form

$$a(q, p) = \frac{1}{2}\left[\frac{\ln(1 + \nu\,\|(q, p)\|^2)}{\nu\,\|(q, p)\|^2} - 1\right]\begin{pmatrix}-p\\ q\end{pmatrix} .$$

For $\Omega_t^{-\top}$ is given by

$$\Omega_t^{-\top} = g(t, \nu\,\|(q, p)\|^2)J \qquad\qquad g = \frac{(1 + \nu\,\|(q, p)\|^2)}{1 + t\nu\,\|(q, p)\|^2} ,$$

the vector field $V_t = \Omega_t^{-\top} a(q, p)$ reads

$$V_t = \chi(t, q, p)\begin{pmatrix}q\\ p\end{pmatrix} , \qquad\qquad \chi = \frac{1 + \nu\,\|(q, p)\|^2}{2(1 + t\nu\,\|(q, p)\|^2)}\left[\frac{\ln(1 + \nu\,\|(q, p)\|^2)}{\nu\,\|(q, p)\|^2} - 1\right] ,$$

which is (22).

Lemma 3.1. *Given $T > 1$, the vector field V_t is contracting and analytic in $\mathcal{G}_\delta \times D_\rho$ for $\delta \le \frac{1}{\nu\rho^2} - T$ and $\rho \le \rho_* := \frac{1}{\sqrt{2T\nu}}$. Moreover:*

1. *it leaves the origin O fixed and for $\|(q, p)\|$ small enough it admits the time-independent asymptotic expansion*
$$V_t \sim V_{t,3} = -\frac{1}{4}\nu\,\|(q, p)\|^2\begin{pmatrix}q\\ p\end{pmatrix} ;$$
2. *it is symmetric under the action of the rotation group*
$$R(s) = \begin{pmatrix}\cos(s) & \sin(s)\\ -\sin(s) & \cos(s)\end{pmatrix} ;$$
3. *it satisfies the estimate*
$$\left|V_t\right|_{T,\rho} \prec \nu\rho^3 .$$

Proof. The field V_t is clearly decomposed in a coefficient which depends only on the norm $(q^2 + p^2)$ and a radial direction (q, p), hence it is invariant under the group action of $R(s)$. Indeed V_t commutes with $(p, -q)$, the generator of $R(s)$. Since $\ln(1 + \nu\,\|(q, p)\|^2) < \nu\,\|(q, p)\|^2$ for any $(q, p) \neq O$, the flow is contracting in the future and the origin is the only equilibrium of the dynamical system defined by V_t. The asymptotic expansion $V_t \sim V_{t,3}$ is immediately derived from the Taylor expansion with respect to the phase space variables (q, p). We rewrite $\chi = \chi_1\chi_2$ with

$$\chi_1 = \frac{1 + \nu\,\|(q, p)\|^2}{2(1 + t\nu\,\|(q, p)\|^2)} , \qquad\qquad \chi_2 = \frac{\ln(1 + \nu\,\|(q, p)\|^2)}{\nu\,\|(q, p)\|^2} - 1 .$$

The second factor χ_2 is analytic in D_ρ with $\rho < \frac{1}{\sqrt{\nu}}$; the first factor is analytic in polydisks $B_\delta \times D_\rho$ where δ, radius of the disks $B_\delta(\zeta)$ for $\zeta \in (-T, T)$, satisfies

$$\delta = \inf_{|\zeta|<T}\left|\frac{1}{\nu\,\|(q, p)\|^2} + \zeta\right| = \frac{1}{\nu\rho^2} - T ,$$

with $T \leq \frac{1}{\nu\rho^2} - T$ due to the condition $\rho \leq \rho^*$. Hence we are allowed to take $\delta = T$ in the estimate of V_t: the factor χ_1 can be uniformly bounded by a constant in the given domain, while the factor χ_2 and the radial direction provide the cubic dependence on ρ. □

The above Lemma implies that its flow Φ^1 is a (contracting in the future) nonlinear analytic deformation of the identity map, for any $|t| \leq 1$, provided ρ is small enough. Indeed, by its definition in terms of Lie-series[5], in Lemma (3.1) we can take $\delta = T = \frac{1}{2\nu\rho^2}$ (which means $\rho = \rho^*$) and $d_1 = d_2 = d$ so that

$$\Gamma < \frac{\nu\rho^2}{d} < \Gamma^*$$

holds true for sufficiently small ρ; the asymptotic expansion (23) follows immediately from the local behavior of V_t

$$\Phi^1(q,p) \sim (1 + V_t(1,q,p))\begin{pmatrix} q \\ p \end{pmatrix} \sim \left(1 - \frac{1}{4}\nu\,\|(q,p)\|^2\right)\begin{pmatrix} q \\ p \end{pmatrix} .$$

We finally have to show that $\Phi^1(q,p)$ is the inverse of (15). We can assume the time-one-flow $\varphi(q,p) = \Phi^1(q,p)$ to be a radial and close to the identity transformation having the form

$$\Phi^1(q,p) = \xi(\sqrt{\nu}\,\|(q,p)\|)\begin{pmatrix} q \\ p \end{pmatrix} \qquad\qquad \xi(0) = 1 ,$$

with ξ being analytic in the norm $\|(q,p)\|$; the same can be assumed also for the inverse transformation $\varphi^{-1}(x,y)$

$$\varphi^{-1}(x,y) = \sigma(\sqrt{\nu}\,\|(x,y)\|)\begin{pmatrix} x \\ y \end{pmatrix} \qquad\qquad \sigma(0) = 1 .$$

By imposing for $\varphi^{-1}(x,y)$ the condition of being a Darboux transformation $(\varphi^{-1})^*\omega_0 = \omega_1$, we get the following equation for σ

$$\sigma'\sigma\varrho + \sigma^2 = 1 + \varrho^2\sigma^2 \qquad \varrho = \sqrt{\nu}\,\|(x,y)\|$$

which becomes, by introducing the more suitable variable $h(\varrho) = \sigma^2(\varrho)$, a linear and non-homogeneous equation of the form

$$h' + \frac{2}{\varrho}(1 - \varrho^2)h - \frac{2}{\varrho} = 0 . \qquad (30)$$

The unique analytic solution of (30) is given by

$$h(\varrho) = \frac{1}{\varrho^2}\left(e^{\varrho^2} - 1\right),$$

which gives $\sigma(x,y)$ the expression in (15).

[5] Alternatively, one can construct a series of analytic approximating solutions which are uniformly convergent in B_ρ to the solution, then also the solution has to be analytic.

3.2. *Proof of Theorem 2.2*

In order to prove the normal form statement and the bound on $\mathcal{R}^{(1)}$, we make use of three different levels of approximation. As in the previous proof, we set $\delta = T = \frac{1}{2\nu\rho^2}$ and we assume ρ small enough to ensure $T > 1$. First expand the vector field $\tilde{V}$

$$\exp(L_{-\tilde{V}}H) = \exp(L_{-\tilde{V}_3}H) + \mathcal{R}_1 \ ,$$

where $\mathcal{R}_1$ can be bounded using Proposition (Appendix A.3) with $\Gamma_3 = \frac{1}{d\rho}\left|V - V_t\right|_{T,\rho} \prec \nu^2\rho^4$

$$\left|\mathcal{R}_1\right|_{(1-d)T,(1-d)\rho} \prec \Gamma_3\left|H\right|_{T,\rho} \prec \nu^3\rho^8 + \varepsilon\nu^2\rho^6 \ .$$

As a second step, we truncate the Lie series of $\tilde{V}_3$ at order $K = 1$

$$\exp(L_{-\tilde{V}_3}H) = \exp_1(L_{-\tilde{V}_3}H) + \mathcal{R}_2 \ ,$$

where

$$\exp_1(L_{-\tilde{V}_3}H) = H - L_{\tilde{V}_3}H \ ,$$

and $\mathcal{R}_2$ can be bounded using Proposition (Appendix A.2) with $\Gamma = \frac{1}{d}\left(2\nu\rho^2 + \left|\tilde{V}\right|_{T,\rho}\right) \prec \nu\rho^2$

$$\left|\mathcal{R}_2\right|_{(1-d)T,(1-d)\rho} \prec \Gamma^2\left|H\right|_{T,\rho} \prec \nu^3\rho^8 + \varepsilon\nu^2\rho^6 \ .$$

Last and easiest step consists in a Taylor expansion of H_0 in $H - L_{\tilde{V}_3}H$ which gives

$$H - L_{\tilde{V}_3}H = H_{0,4} + H_{1,2} - L_{V_3}H_{0,4} - L_{V_3}H_{1,2} + \mathcal{R}_3 \ ,$$

with

$$\left|\mathcal{R}_3\right|_{\rho} \prec \varepsilon\nu^2\rho^6 + \nu^3\rho^8 \ .$$

Since $d < 1$ is arbitrary, ρ is assumed to be small enough and the three contributions $\mathcal{R}_j$ to the remainder are of the same order, the estimate follows. The estimates for $\mathcal{R}^{(0)}$ follows by minor variations; in the AL case just remember that the cubic nonlinearity is one of the leading terms in the remainder.

In order to prove the different bounds on the closeness between the AL/Salerno model and the normal forms, one has to apply Cauchy estimates (A.1) to get estimates on the vector fields $X_{\mathcal{R}^{(0,1)}}$, starting from the bounds on the remainders

$$\left|X_{\mathcal{R}}\right|_{(1-d)\rho} \leq \frac{1}{d\rho}\left|\mathcal{R}\right|_{\rho} \ , \qquad\qquad \left|X_{\mathcal{R}}\right|_{(1-d)\rho} = \max_{j=1,\ldots,2n}\left|X_{\mathcal{R},j}\right|_{(1-d)\rho} \ . \tag{31}$$

Hence from (31) one can obtain

$$\left|X_{\mathcal{R}^{(1)}}\right|_{(1-d)\rho} \prec \nu^2\rho^5(\rho^2 + \varepsilon) \ ,$$

for the cubic-quintic dNLS-like normal form and

$$\left|X_{\mathcal{R}^{(0)}}\right|_{(1-d)\rho} \prec \nu^2\rho^3(\rho^2 + \varepsilon) \ , \qquad\qquad \left|X_{\mathcal{R}^{(0)}}\right|_{(1-d)\rho} \prec \nu^2\rho^3(1 + \varepsilon) \ ,$$

for the standard dNLS normal form, in the Salerno and AL cases respectively. Then all the estimates follow from Gronwall Lemma and from the conservation of the norm $\|x(t), y(x)\|^2$ along the different Hamiltonian flows (so that all the orbits belong to the initial ball $B_\rho(0)$ for infinite times). We can sketch the procedure as follows: we have to compare the dynamics of $H = Z^{(0)} + \mathcal{R}^{(0)}$ with the one of the normal form $Z^{(0)}$

$$\dot{z} = X_H(z) \qquad\qquad \dot{\zeta} = X_{Z^{(0)}}(\zeta) = X_H(\zeta) - X_{\mathcal{R}^{(0)}}(\zeta) ,$$

so we introduce the error $\delta(t) = z(t) - \zeta(t)$ which solves

$$\dot{\delta} = [X_H(z) - X_H(\zeta)] + X_{\mathcal{R}^{(0)}}(\zeta) = [X_H(\zeta + \delta) - X_H(\zeta)] + X_{\mathcal{R}^{(0)}}(\zeta) .$$

Hence we can derive the differential inequality

$$\left\|\dot{\delta}\right\| \leq \|X_H(\eta)\| \, \|\delta\| + \|X_{\mathcal{R}^{(0)}}(\zeta)\| \ , \qquad\qquad \eta = \zeta + s\delta , \qquad s \in [0, 1] ;$$

then Gronwall estimate can be applied, once we provide estimates for $\|X_H(\eta)\|$ and $\|X_{\mathcal{R}^{(0)}}(\zeta)\|$, for ζ, η belonging to a polydisk of small radius.

3.3. Proof of Proposition 2.1

To prove both the estimates, we have first to set (as in the previous proof) $\delta = T = \frac{1}{2\nu\rho^2}$ and assume ρ small enough, so that $T > 1$ and we can evaluate the Lie-series at $\tau = 1$, lying inside the domain $\mathcal{G}_\delta$. The first of (28) is a consequence of (A.5) with $\left|f\right|_{\delta,\rho} = \left|P\right|_{\delta,\rho} \sim \rho^2$ and

$$\Gamma \prec \frac{\nu\rho^2}{d} \qquad \Rightarrow \qquad (e\Gamma)^{K+1}\left|f\right|_{\delta,\rho} \prec \nu^{K+1}\rho^{2K+4} .$$

The second of (28) can be derived combining the previous estimate with (A.9), where Γ_3, defined in (A.7), in this case fulfills

$$\Gamma_3 \prec \frac{1}{\rho}\left|Y\right|_{\delta,\rho} = \frac{1}{\rho}\left|V_t - V_t^{(2L+1)}\right|_{\delta,\rho} \sim \nu^{L+1}\rho^{2L+2} ,$$

since the leading term of the remainder $V_t - V_t^{(2L+1)}$ is of order $2L + 3$.

4. Conclusions

In this manuscript we have focused on the constructive aspects of Moser's proof of Darboux's Theorem, with the aim of a deeper understanding of the standardization procedure, in particular for the AL and Salerno model.

In general, Darboux's change of coordinates is local, and linear coordinates can always be chosen so that it represents a small perturbation of the identity map; hence a successful strategy is to combine the polynomial approximation of the vector field V with the use of truncated Lie-series (or even Lie-transform), in order to compute leading order terms of the transformed Hamiltonian $H(\Psi)$, *without any need to derive a complete explicit expression for* Ψ *or for* V *itself.* This approximation of the transformed Hamiltonian might be enough for the subsequent investigation of its dynamical features by means of perturbation techniques (for example, by normal form methods). For what concerns the specific case of the Salerno model, such a normal form can be used as a starting point for perturbation schemes which require the use of the standard Poisson brackets: for example, one can start to explore existence and stability of localized solutions, such as multi-breathers, quasi-periodic breathers, low dimensional tori (see [3, 25, 26, 15, 6]).

Outside the field of nonlinear lattices, a classical example that it is worth mentioning is that of the Lotka-Volterra system and its higher dimensional generalizations (Lotka-Volterra systems). It is well known that, for $(x, y) \in \mathbb{R}^2$ the predator-prey system $\dot{x} = \alpha x - \beta xy$, $\dot{y} = -\gamma y + \delta xy$ (usually with all the parameters taken positive) admits the constant of motion $H(x, y) = \beta y - \alpha \ln y + \delta x - \gamma \ln x$. In the original variables the system is not Hamiltonian with respect to the standar Poisson structures, but it is so with the non standard[6] Poisson brackets $\{x, y\} = xy$. Equivalently it is possible to use the change of variables $\xi = \ln x$, $\eta = \ln y$ to obtain an Hamiltonian form with the standard symplectic structure. In [23] it is discussed the effectiveness of a particular numerical integration scheme by showing that it turns out to be symplectic with respect to the non standard Poisson structure. A suitable generalization of the above mentioned non standard structure is also used in the higher dimensional extensions of LV models: the possibility to view also those systems as Hamiltonian open the way for the investigation of their integrability (see [2] for a recent work in that direction).

Appendix A. Lie series in the extended phase space.

In this Section we present some analytical result about Lie-series in the non autonomous case (see for example the Appendix of [8] for similar estimates in the autonomous case and [9] in the non autonomous Hamiltonian context). In order to estimate the Lie series of a given function f

$$\exp(tL_{\tilde{X}} f)(\zeta, z) = \sum_{k\geq 0} \frac{1}{k!} t^k L_{\tilde{X}}^k f(\zeta, z)$$

already defined in (19), it is necessary to provide an upper bound to the Lie derivative $L_{\tilde{X}} f$ in the shrinked domain $\mathcal{G}_{(1-d_1)\delta} \times D_{(1-d_2)\rho}$, according to the usual Cauchy inequality valid for analytic functions (in several complex variables)

$$\Big|\frac{\partial f}{\partial \zeta}\Big|_{(1-d_1)\delta,(1-d_2)\rho} \leq \frac{1}{d_1\delta}\Big|f\Big|_{\delta,(1-d_2)\rho} , \qquad \Big|\frac{\partial f}{\partial z_j}\Big|_{(1-d_1)\delta,(1-d_2)\rho} \leq \frac{1}{d_2\rho}\Big|f\Big|_{(1-d_1)\delta,\rho} , \tag{A.1}$$

which is a consequence of the (multidimensional) Cauchy formula[7]

$$\frac{\partial^k f}{\partial u^k}(u) = \frac{k!}{2\pi i}\int_{T_{\delta,\rho}} \frac{f(v)}{(v-u)^{k+1}} dv , \qquad u = (\zeta, z) ,$$

where $k = (k_0, k_1, \ldots, k_n) \in \mathbb{Z}^{n+1}$ is a derivative multi index and $T_{a,b} = \partial B_\delta(\zeta) \times \partial B_\rho(z_1) \times \ldots \times B_\rho(z_n)$ is the $n+1$-dimensional torus given by the product of the boundaries of all the distinguished disks (distinguished boundary). Hence, if f is analytic in D_ρ, the same Cauchy formula allows to bound the k-derivative of f (with respect to z) in $D_{\rho-\delta}$

$$\Big|\frac{\partial^l f}{\partial z^l}\Big|_{\rho-\delta} \leq \frac{l_1!\ldots l_n!}{\delta^k}\Big|f\Big|_\rho , \qquad |l| = k .$$

[6] Sometimes called log-canonical.

[7] One can differently consider $\frac{\partial f}{\partial u_j}(u) = \frac{\partial f(u+e_j z)}{\partial z}(0)$ and use the one dimensional version of the Cauchy formula.

Consider f, analytic on $\mathcal{G}_\delta \times D_1$, as an analytic function on $D_\rho \subset D_1$, and Taylor expand f with respect to $z \in D_\rho$

$$f(\zeta, z) = T_k(\zeta, z) + R_{k+1}(\zeta, z)\ , \qquad R_{k+1}(\zeta, z) = \sum_{|l|=k+1} \frac{1}{l!} f^{(l)}(\zeta, cz) z^l\ , \qquad c \in (0,1)\ .$$

Then the following estimate provide the exponential decay of the reminder ot the truncated Taylor series

Lemma Appendix A.1. *Let $\rho < \frac{1}{2}$, then there exists $E = E(k, n)$ such that*

$$\Big|R_{k+1}\Big|_{\delta,\rho} \leq (2\rho)^{k+1} \Big|f\Big|_1 E\ .$$

Proof. It is enough to exploit the above bound of the derivative with respect to z and the combinatorial formula of the number of partial derivatives of given order $k+1$. Indeed from

$$|R_{k+1}(\zeta, z)| \leq \sum_{|l|=k+1} \frac{1}{l!} |f^{(l)}(\zeta, cz)| \rho^{k+1}\ ,$$

since $z \in D_\rho$ implies $cz \in D_{c\rho} \subset D_\rho$, we have

$$\Big|R_{k+1}\Big|_{\delta,\rho} \leq \rho^{k+1} \sum_{|l|=k+1} \frac{1}{l!} \Big|f^{(l)}\Big|_{\delta,\rho}\ ,$$

with

$$\Big|f^{(l)}\Big|_{\delta,\rho} \leq \frac{l!}{(1-\rho)^{k+1}} \Big|f\Big|_1 < 2^{k+1} l! \Big|f\Big|_1\ .$$

Hence

$$\Big|R_{k+1}\Big|_{\delta,\rho} \leq (2\rho^{k+1}) \Big|f\Big|_1 \left(\sum_{|l|=k+1} \right) = (2\rho^{k+1}) \Big|f\Big|_1 E\ ,$$

where $E = C^*_{n,k+1} = \binom{n+k}{k+1}$. ☐

Lemma Appendix A.2. *Let $\tilde{X} = (1, X)$ with X analytic in $\mathcal{G}_\delta \times D_\rho$ and f analytic in $\mathcal{G}_\delta \times D_\rho$, and let $0 < d_j < 1$, then*

$$\Big|L_{\tilde{X}} f\Big|_{(1-d_1)\delta, (1-d_2)\rho} \leq \Gamma \Big|f\Big|_{\delta,\rho} \qquad \Gamma = \frac{1}{d_1 \delta} + \frac{\Big|X\Big|_{\delta,\rho}}{d_2 \rho}\ . \tag{A.2}$$

Proof. Notice that the Lie differential operator $L_{\tilde{X}}$ acts on f as

$$L_{\tilde{X}} f = \partial_\zeta f + \langle X, \nabla_z f \rangle\ ,$$

hence from the previous Cauchy estimate and the usual bound on the scalar product we get (A.2). ☐

Remark Appendix A.1. *The basic estimate* (A.2) *and the ones which follow are equivalent to the ones due to Gröbner in [10], obtained originally with the classical methods of majorants due to Cauchy.*

Remark Appendix A.2. *If f and X depend only on $z \in D_\rho$, then formula* (A.2) *becomes the usual estimate*

$$\left|L_X f\right|_{(1-d)\rho} \le \frac{1}{d\rho}\left|X\right|_\rho \left|f\right|_\rho \ .$$

Next Lemma provides the main estimate when dealing with the convergence of the Lie series and related results

Lemma Appendix A.3. *Let $\tilde{X} = (1, X)$ and f as in Lemma Appendix A.2, and let $0 < d_j < 1$, then*

$$\left|L^k_{\tilde{X}} f\right|_{(1-d_1)\delta,(1-d_2)\rho} \le \frac{k!}{e}(\Gamma e)^k \left|f\right|_{\delta,\rho} \ . \tag{A.3}$$

Proof. The result follows by repeating formula (A.2) for k-times to each Lie derivative, starting from $L^k_{\tilde{X}} f = L_{\tilde{X}}\left(L^{k-1}_{\tilde{X}}\right) f$, by reducing each time the domain of the supremum norm by the factors $d'_j = \frac{1}{k} d_j$. Then, to conclude, one has to use the basic inequality $k^k \le k! e^{k-1}$. □

The above estimate (A.3) can be generalized in the following way

Lemma Appendix A.4. *Let $\tilde{X}_j = (1, X_j)$, $j = 1, \ldots, k$, be a sequence of vector fields with X_j analytic in $\mathcal{G}_\delta \times D_\rho$ and f analytic in $\mathcal{G}_\delta \times D_\rho$, and let $0 < d_j < 1$, then*

$$\left|L_{\tilde{X}_k} \ldots L_{\tilde{X}_1} f\right|_{(1-d_1)\delta,(1-d_2)\rho} \le \frac{k! e^k}{e}\left(\prod_{j=1}^{k} \Gamma_j\right)\left|f\right|_{\delta,\rho} \qquad \Gamma_j = \frac{1}{d_1\delta} + \frac{\left|X_j\right|_{\delta,\rho}}{d_2\rho} \ . \tag{A.4}$$

Next Lemma exploits the previous estimates to get total convergence of the Lie series on a smaller domain $\mathcal{G}'_\delta \times D'_\rho \subset \mathcal{G}_\delta \times D_\rho$:

Lemma Appendix A.5. *Let $\tilde{X}$ and f be as in Lemma Appendix A.2. Then for any $0 < d_j < 1$ the Lie series $\exp(tL_{\tilde{X}} f)$ is totally convergent in $\mathcal{G}_{(1-d_1)\delta} \times D_{(1-d_2)\rho}$ for times $|t| < \frac{1}{e\Gamma}$.*

Proof. By using (A.4) one obtains

$$\left|\exp(tL_{\tilde{X}} f)\right|_{(1-d_1)\delta,(1-d_2)\rho} \le \sum_{k\ge0} \frac{1}{k!}|t|^k \left|L^k_{\tilde{X}} f\right|_{(1-d_1)\delta,(1-d_2)\rho} \le \frac{\left|f\right|_{\delta,\rho}}{e} \sum_{k\ge0} (|t|\Gamma e)^k < \infty \ ,$$

for times $|t|\Gamma e < 1$. □

In the following we focus on the Lie series at times $t = \pm 1$; indeed we are mainly interested in the transformation $\varphi(z) := \{\varphi_j(z)\}_{j=1,\ldots,n}$, with $\varphi_j(z) = \exp(L_{\tilde{X}} \mathbb{I}_j)(0, z)$, and in its inverse $\varphi^{-1}(z)$. The following result holds true:

Proposition Appendix A.1. *Let $\tilde{X}$ be as in Lemma Appendix A.2. Then for any $j = 1, \ldots, n$ the Lie series $\exp(tL_{\tilde{X}} \mathbb{I}_j)(0, z)$ is totally convergent in $z \in D_{(1-d)\rho}$, for times $|t| < \frac{1}{\Gamma e}$ and $0 < d < 1$. Moreover, if*

$$\Gamma < \Gamma^* = \frac{e}{1 + e^2} < \frac{1}{e} \ ,$$

then the transformations $\varphi^{\pm1}(z)$ are well defined as Lie-series and, for $0 < d < \frac{1}{2}$, satisfy the inclusion

$$D_{(1-2d)\rho} \subset \varphi^{\pm1}(D_{(1-d)\rho}) \subset D_\rho \ .$$

Proof. For $z \in D_{(1-d_2)\rho}$ we have

$$|\exp(tL_{\tilde{X}}\mathbb{I}_j)(0,z)| \leq \left|\exp(tL_{\tilde{X}}\mathbb{I}_j)\right|_{(1-d_1)\delta,(1-d_2)\rho} \leq \frac{\left|\mathbb{I}_j\right|_{\rho}}{e}\sum_{k\geq 0}(|t|\Gamma e)^k < \infty\ .$$

Moreover, condition $\Gamma < \Gamma^*$ ensures $\frac{1}{e\Gamma} > 1$, so that $t = \pm 1$ makes sense in the Lie-series. The deformation of the domain can be bounded by taking the Lie series from $k \geq 1$

$$\left|\exp(L_{\tilde{X}}\mathbb{I}_j) - \mathbb{I}_j\right|_{(1-d_1)\delta,(1-d_2)\rho} \leq \sum_{k\geq 1}\frac{1}{k!}\left|L^k_{\tilde{X}}\mathbb{I}_j\right|_{(1-d_1)\delta,(1-d_2)\rho} = \sum_{k\geq 1}\frac{1}{k!}\left|L^{k-1}_{\tilde{X}}X_j\right|_{(1-d_1)\delta,(1-d_2)\rho}\ ;$$

then

$$\begin{aligned}\left|L^{k-1}_{\tilde{X}}X_j\right|_{(1-d_1)\delta,(1-d_2)\rho} &\leq \frac{(k-1)!}{e}(e\Gamma)^{k-1}\left|X_j\right|_{\delta,\rho} = (d_2\rho)\frac{(k-1)!}{e}(e\Gamma)^{k-1}\frac{\left|X_j\right|_{\delta,\rho}}{d_2\rho} \leq \\ &\leq (d_2\rho)\frac{(k-1)!}{e^2}(e\Gamma)^k\ .\end{aligned}$$

Hence by condition $\Gamma < \Gamma^*$ one has

$$\sum_{k\geq 1}\frac{1}{k!}\left|L^k_{\tilde{X}}\mathbb{I}_j\right|_{(1-d_1)\delta,(1-d_2)\rho} \leq \frac{d_2\rho}{e^2}\left(\frac{e\Gamma}{1-e\Gamma}\right) < d_2\rho\ .$$

The same estimate holds by replacing $(1-d_2)\rho$ with $(1-2d_2)\rho$ and ρ with $(1-d_2)\rho$ respectively in the inequalities involving the Lie derivatives. □

The proof clearly shows that the size of $\left|X\right|_{\delta,\rho}$ provides the leading deformation of $\varphi(z)$ with respect to the identity

$$\left|\varphi_j(z) - z_j\right|_{(1-d_2)\rho} \prec \left|X_j\right|_{\delta,\rho}\ .$$

Two important issues to be addressed in perturbation theory are to estimate the error when either the Lie series is truncated at some finite order K or the analytic vector field $\tilde{X}$ is approximated with its Taylor polynomial, truncated at order L. To provide standard estimates of this type is the goal of next results. We first introduce the following notation: given two integers $K, L \geq 1$ we denote by $\exp_K(L_{\tilde{X}}f) = \sum_{k=0}^K L^k_{\tilde{X}}f$, the truncation of the Lie series at order K, and by $\exp(L_{\tilde{X}^{(L)}}f) = \sum_{k\geq 0} L^k_{\tilde{X}^{(L)}}f$, the Lie series of the truncated vector field $\tilde{X}^{(L)}$ at order L. In the following, to simplify some estimates, we also make use of the symbol $\prec$ defined in (29).

Proposition Appendix A.2. *Let $\tilde{X}$ and f as in Lemma Appendix A.2,* $0 < d_j < 1$ *and* $K \geq 1$. *If* $\Gamma < \Gamma^*$ *then*

$$\left|\exp(L_{\tilde{X}}f) - \exp_K(L_{\tilde{X}}f)\right|_{(1-d_1)\delta,(1-d_2)\rho} \prec (e\Gamma)^{K+1}\left|f\right|_{\delta,\rho}\ . \tag{A.5}$$

Proof. The result is a consequence of the estimate of the reminder

$$\Big|\sum_{k\geq K+1} L^k_{\tilde{X}}f\Big|_{(1-d_1)\delta,(1-d_2)\rho} \leq \frac{1}{e}\left|f\right|_{\delta,\rho}\sum_{k\geq K+1}(e\Gamma)^k = \frac{1}{e}(e\Gamma)^{K+1}\left|f\right|_{\delta,\rho}\left(\sum_{k\geq 0}(e\Gamma)^k\right).$$

□

In order to study the error due to the truncation of a vector field, we start considering the Lie series of two vector fields $\tilde{X}_{1,2}$. Next Lemma allows to properly rewrite the difference of Lie derivatives:

Lemma Appendix A.6. *Let $\tilde{X}_1$ and $\tilde{X}_2$ two vector fields as in Lemma Appendix A.2, and define $\tilde{Y} = \tilde{X}_1 - \tilde{X}_2 = (0, X_1 - X_2)$. Then the following holds true*

$$\begin{aligned} & L_{\tilde{X}_1} - L_{\tilde{X}_2} = L_{\tilde{Y}} \ , \\ & L^k_{\tilde{X}_1} - L^k_{\tilde{X}_2} = L_{\tilde{Y}} L^{k-1}_{\tilde{X}_2} + \sum_{l=1}^{k-1} L^{k-l}_{\tilde{X}_1} L_{\tilde{Y}} L^{l-1}_{\tilde{X}_2} \qquad k \geq 2 \ , \end{aligned} \tag{A.6}$$

where obviously $L_{\tilde{Y}} = L_Y$.

Proof. By induction. Formula (A.6) works for $k = 1$, since by linearity we have $L_{\tilde{X}_1} - L_{\tilde{X}_2} = L_{\tilde{Y}}$, and for $k = 2$, since

$$L^2_{\tilde{X}_1} - L^2_{\tilde{X}_2} = L_{\tilde{X}_1} L_{\tilde{X}_2} + L_{\tilde{X}_1} L_{\tilde{Y}} - L^2_{\tilde{X}_2} = L_{\tilde{X}_1} L_{\tilde{Y}} + L_{\tilde{Y}} L_{\tilde{X}_2} \ .$$

We assume the above formula to hold for $k \geq 2$ and we show its validity for $k + 1$. Indeed

$$\begin{aligned} L^{k+1}_{\tilde{X}_1} - L^{k+1}_{\tilde{X}_2} &= L_{\tilde{X}_1} L^k_{\tilde{X}_1} - L^{k+1}_{\tilde{X}_2} = L_{\tilde{X}_1} \left[L^k_{\tilde{X}_2} + L_{\tilde{Y}} L^{k-1}_{\tilde{X}_2} + \sum_{l=1}^{k-1} L^{k-l}_{\tilde{X}_1} L_{\tilde{Y}} L^{l-1}_{\tilde{X}_2} \right] - L^{k+1}_{\tilde{X}_2} = \\ &= L_{\tilde{X}_1} L^k_{\tilde{X}_2} + L_{\tilde{X}_1} L_{\tilde{Y}} L^{k-1}_{\tilde{X}_2} + \sum_{l=1}^{k-1} L^{k+1-l}_{\tilde{X}_1} L_{\tilde{Y}} L^{l-1}_{\tilde{X}_2} - L^{k+1}_{\tilde{X}_2} = \\ &= L_{\tilde{Y}} L^k_{\tilde{X}_2} + \sum_{l=1}^{k} L^{k+1-l}_{\tilde{X}_1} L_{\tilde{Y}} L^{l-1}_{\tilde{X}_2} \ . \end{aligned}$$

□

Proposition Appendix A.3. *Let $\tilde{X}_{1,2}$ and $\tilde{Y}$ as in Proposition Appendix A.2 and let $\Gamma_{j=0,1,2,3}$ be defined by*

$$\Gamma_1 = \frac{1}{d_1 \delta} + \frac{\left| X_1 \right|_{\delta,\rho}}{d_2 \rho} \ , \qquad \Gamma_2 = \frac{1}{d_1 \delta} + \frac{\left| X_2 \right|_{\delta,\rho}}{d_2 \rho} \ , \qquad \Gamma_3 = \frac{\left| Y \right|_{\delta,\rho}}{d_2 \rho} \ , \qquad \Gamma_0 = \max\{\Gamma_1, \Gamma_2\} \ . \tag{A.7}$$

Then for any f analytic in $\mathcal{G}_\delta \times D_\rho$, if $e\Gamma_0 < 1$ we have

$$\left| \exp(L_{\tilde{X}_1} f) - \exp(L_{\tilde{X}_2} f) \right|_{(1-d_1)\delta,(1-d_2)\rho} < \Gamma_3 \left| f \right|_{\delta,\rho} \ . \tag{A.8}$$

Proof. We have to exploit Lemma Appendix A.6 to estimate the difference of the Lie derivatives for $k \geq 2$

$$\begin{aligned} \left| L^k_{\tilde{X}_1} f - L^k_{\tilde{X}_2} f \right|_{(1-d_1)\delta,(1-d_2)\rho} &\leq \left| L_{\tilde{Y}} L^{k-1}_{\tilde{X}_2} f \right|_{(1-d_1)\delta,(1-d_2)\rho} + \sum_{l=1}^{k-1} \left| L^{k-l}_{\tilde{X}_1} L_{\tilde{Y}} L^{l-1}_{\tilde{X}_2} f \right|_{(1-d_1)\delta,(1-d_2)\rho} \leq \\ &\leq k! e^{k-1} \left| f \right|_{\delta,\rho} \left[\Gamma_3 \Gamma_2^{k-1} + \sum_{l=1}^{k-1} \Gamma_1^{k-l} \Gamma_3 \Gamma_2^{l-1} \right] = \\ &= k! e^{k-1} \left| f \right|_{\delta,\rho} \Gamma_3 \left[\sum_{l=1}^{k} \Gamma_1^{k-l} \Gamma_2^{l-1} \right] \leq k! (k-1) \Gamma_3 (e\Gamma_0)^{k-1} \left| f \right|_{\delta,\rho} \ , \end{aligned}$$

while for $k = 1$ one easily has

$$\left| L_{\tilde{X}_1} f - L_{\tilde{X}_2} f \right|_{(1-d_1)\delta,(1-d_2)\rho} = \left| L_Y f \right|_{(1-d_1)\delta,(1-d_2)\rho} \leq \Gamma_3 \left| f \right|_{\delta,\rho} .$$

Hence we have

$$\left| \exp(L_{\tilde{X}_1} f) - \exp(L_{\tilde{X}_2} f) \right|_{(1-d_1)\delta,(1-d_2)\rho} \leq \Gamma_3 \left| f \right|_{\delta,\rho} \left(1 + \sum_{k\geq 1} k(e\Gamma_0)^k \right) \prec \Gamma_3 \left| f \right|_{\delta,\rho} .$$

□

The previous Proposition allows to estimate the error due to the truncation of a vector field. Indeed, if $X_2 = X^{(L)}$, namely the truncation at order L of a given field X, then $Y = X - X^{(L)}$ is the remainder. Due to Lemma Appendix A.1 it holds also $\left|Y\right|_{\delta,\rho} \leq \left|X\right|_{\delta,1} \rho^{L+1} E$, hence (A.8) takes the form

$$\left| \exp(L_{\tilde{X}} f) - \exp(L_{\tilde{X}^{(L)}} f) \right|_{(1-d_1)\delta,(1-d_2)\rho} \prec \frac{E}{d_2} \left|X\right|_{\delta,1} \left|f\right|_{\delta,\rho} \rho^L . \tag{A.9}$$

Appendix B. Outline of Moser's constructive scheme

The scheme of the proof here reported is taken from [17]. Since we are mostly interested in translating the idea of the proof in a applicable scheme, we omit most of the geometric details, while focusing on the objects one has to construct (in a given set of coordinates) and on their manipulation.

Theorem Appendix B.1. *Let (M, ω) a symplectic manifold and ω be a non-degenerate closed 2-form in a neighborhood U of $P \in M$. Then there exists V, a neighborhood of P, and a set of coordinates $(y_1, ..., y_n, x_1, ..., x_n)$ defined in V such that ω is represented in the standard form*

$$\omega = \sum_{i=1}^{n} dx_i \wedge dy_i .$$

Proof. Let ω_0 be a representation of the ω in a local chart (U_0, φ_0) around the arbitrary point $P \in M$, hence $\omega = \varphi_0^* \omega_0$; let $\{q_i\}_{i=1,...,2n}$ the coordinates related to this chart, hence

$$\omega_0 = \sum_{i<j} \omega_{0,ij}(q) dq_i \wedge dq_j .$$

We recall that the coefficients $\omega_{0,ij}$ uniquely define the antisymmetric matrix Ω which represents the action of ω_0 on tangent vectors X, Y, namely

$$\omega_0(X, Y) = X^\top \Omega(q) Y \qquad\qquad \Omega_{i<j}(q) = \omega_{0,ij}(q) .$$

We have to show that there exists a local chart (U_1, φ_1) centered in P, and a second set of coordinates $\{x_i, y_i\}_{i=1,...,n}$ such that $\varphi_1^*(\omega_1) = \omega$ where ω_1 is the standard symplectic form on $\mathbb{R}^{2n}$

$$\omega_1 = \sum_{i=1}^{n} dx_i \wedge dy_i = \sum_{i<j} \omega_{1,ij}(q) dz_i \wedge dz_j ;$$

here $z_i = x_i$ for $i = 1, \ldots, n$ or $z_i = y_i$ for $i = n+1, \ldots, 2n$ and $\omega_{1,ij} = J_{i<j}$. This is equivalent to saying that there exists a local change of coordinates $(x, y) = \varphi(q)$ such that

$$\varphi^* \omega_1(q) = \omega_1(\varphi(q)) = \omega_0(q) \; .$$

Hence, we can work directly in a neighborhood $U \subset \mathcal{R}^{2n}$ of the origin $O = \varphi_0(P) = \varphi_1(P)$. From Darboux's theorem on vector spaces (see [1]) there exists a linear and symplectic change of coordinates (simplectomorphism) such that $\omega_1|_O = \omega_0|_O$; this means that $\Omega(0) = J$ where J represents the standard symplectic form ω_1. We consider ω_0 written in the "unknown" coordinates $\{z_i\}_i$.

Let $\eta = \omega_1 - \omega_0$ be the deformation of ω_0 with respect to ω_1 (indeed they coincide ath the origin); η is closed and exact [8]. Let α be the 1-form such that $\eta = -d\alpha$, uniquely defined modulo adding df, with $f : U \to \mathbb{R}$. We can always assume[9] $\alpha_O = 0$ and write in coordinates

$$\alpha(x) = \sum_{i=1}^{2n} a_i(x) dx_i \qquad\qquad a_i(0) = 0 \; .$$

We introduce time dependent 2-form on U

$$\omega_t(x) = \omega_0 + t\eta = t\omega_1 + (1-t)\omega_0 \qquad \forall\, t \in \mathbb{R}$$

represented by the matrix

$$\Omega_t(x) = tJ + (1-t)\Omega(x) \qquad\qquad \Omega_t(0) = J \quad \forall t \in \mathbb{R} \; ,$$

which is invertible, being ω_t always closed and non-degenerate.

The idea of the proof is to construct a time-dependent vector field $V_t(q)$ on U whose time-one-flow, $\Psi^t(q)\Big|_{t=1}$, defines the change of coordinates $(x, y) = \varphi(q)$ which puts ω_0 (and hence ω) into its standard form ω_1, namely

$$\omega_1(\varphi(q)) = \sum_{i<j} \omega_{1,ij} d\varphi_i(q) \wedge d\varphi_j(q) = \omega_0(q) \; .$$

This is done by showing that, with a suitable definition of V_t, one has

$$\frac{d}{dt}\omega_t(\Psi^t(q)) = \frac{d}{dt}(\Psi^t)^* \omega_t = 0$$

so that $(\Psi^t)^*\omega_t$ is constant $\forall t$, and then $\omega_1(\Psi^1(q)) = \omega_0(\Psi^0(q)) = \omega_0(q)$, since Ψ^0 is the identity transformation of coordinates $z = (x, y) = q$. The vector field $V_t(q)$ is obtained as a not unique[10] solution to the so-called "Moser equation"

$$\omega_t(V_t, Y) = \alpha(Y) \qquad\qquad \forall Y$$

or more explicitly

$$\Omega_t(q) V_t(q) = a(q) \qquad \Rightarrow \qquad V_t(q) = \Omega_t^{-1}(q) a(q) \; .$$

[8] As we are working on U which is convex therefore in particular star-shaped, then Poincaré's Lemma holds true, i.e. η is exact.

[9] It is enough to redefine $\alpha = \alpha - \alpha_O$

[10] Because α, and hence the vector a, in defined modulo a differential df, as already noticed.

The dynamical system $\dot{q} = V_t(q)$ defines the flow $\Psi^t(q)$

$$\frac{d}{dt}\Psi^t(q) = V_t(\Psi^t(q)) \ .$$

With such a definition of V_t it is possible to explicitly show

$$\frac{d}{dt}\omega_t(\Psi^t(q)) = \frac{d}{dt}\sum_{i<j}\omega_{t,ij}(\Psi^t(q))d\Psi^t_i(q) \wedge d\Psi_j(q) \ ,$$

by performing the various derivatives and obtaining in coordinates that that

$$\frac{d}{dt}\omega_t(\Psi^t(q)) = (\Psi^t)^*\big[\eta(x) + d\alpha(x)\big] = 0 \ .$$

□

Acknowledgments T.P. and S.P. thank Vassilis Koukouloyannis for his hospitality in Samos in September 2022, which inspired useful and intensive discussions about the dynamics of the Ablowitz-Ladik model. T.P. has been supported by the MIUR-PRIN 20178CJA2B "New Frontiers of Celestial Mechanics: Theory and Applications".

References

[1] Abraham, Ralph and Marsden, Jerrold E, Foundations of mechanics. *American Mathematical Soc.*, n. 368, 2008.

[2] Christodoulidi H, Hone ANW and Kouloukas TE A new class of integrable Lotka-Volterra systems *Journal of Computational Dynamics*, 6 (2), 223–237, 2019.

[3] V. Danesi, M. Sansottera, S. Paleari and T. Penati Continuation of spatially localized periodic solutions in discrete NLS lattices via normal forms *Communications in Nonlinear Science and Numerical Simulations* 108 (4), 2022.

[4] G. Darboux, Sur le problème de Pfaff *Bulletin des sciences mathématiques et astronomiques 2 e série*, tome 6, n. 1, 1882.

[5] A. Deprit, Canonical transformations depending on a small parameter, *Cel.Mech.* 1, 1969.

[6] J. Geng, J. Viveros and Y. Yi Quasi-periodic breathers in Hamiltonian networks of long-range coupling *Physica D*, 237, 2008.

[7] A. Giorgilli, Notes on Hamiltonian Dynamical Systems *London Mathematical Society Student Texts*, Cambridge University Press, April 2022.

[8] A. Giorgilli, U. Locatelli and M. Sansottera Improved convergence estimates for the Schröder-Siegel problem *AMPA*, 194, 2015.

[9] A. Giorgilli and E. Zehnder, Exponential stability for time dependent potentials. *ZAMP* 43, 1992

[10] W. Gröbner, Serie di Lie e loro applicazioni (italian translation) Ed. Cremonese, Roma, 1973.

[11] E. Hairer, C. Lubich and G. Wanner, Geometric Numerical Integration *Springer Series in Computational Mathematics*, 31, 2006.

[12] D. Hennig and N. I. Karachalios and J. Cuevas-Maraver, The closeness of the Ablowitz-Ladik lattice to the Discrete Nonlinear Schrödinger equation *Journal of Differential Equations*, Vol. 316, 2022.

[13] Dirk Hennig and Nikos I. Karachalios and Jesus Cuevas-Maraver, The closeness of localized structures between the Ablowitz-Ladik lattice and discrete nonlinear Schrödinger equations: Generalized AL and DNLS systems *Journal of Math. Physics*, Vol. 63, 4, 2022.

[14] G. Hori, Theory of general perturbations with unspecified canonical variables, *Publ. Astron. Soc. Japan*, 18, 1966.

[15] M. Johanson and S. Aubry Existence and stability of quasiperiodic breathers in the discrete nonlinear Schrödinger equation *Nonlinearity*, 10,1997.

[16] A. Junginger, J. Main and G. Wunner, Construction of Darboux coordinates and Poincaré-Birkhoff normal forms in noncanonical Hamiltonian systems *Physica D: Nonlinear Phenomena*, Vol. 348, 2017.

[17] M.J. Lee, Introduction to Smooth Manifolds. *Graduate Texts in Mathematics*, Vol 218, 2nd ed., 2012,

[18] Mithun, Thudiyangal and Maluckov, Aleksandra and Mančić, Ana and Khare, Avinash and Kevrekidis, Panayotis G, How close are integrable and nonintegrable models: A parametric case study based on the Salerno model *Phys. Rev. E*, Vol. 107, 2, 2023.

[19] J. Moser, On the Volume Elements on a Manifold *Transactions of the American Mathematical Society*, Vol. 120, No. 2, 1965.

[20] S. Paleari and T. Penati, An extensive resonant normal form for an arbitrary large Klein-Gordon model. *Annali Matematica Pura ed Applicata*, 2014.

[21] S. Paleari and T. Penati, Hamiltonian Lattice Dynamics. *Editorial for the Special Issue "Hamiltonian Lattice Dynamics", Mathematics in Engineering*, 1 (4), 2019.

[22] D.E. Pelinovsky, T. Penati and S. Paleari, Approximation of small-amplitude weakly coupled oscillators with discrete nonlinear Schroedinger equations *Reviews in Mathematical Physics*, 28, n 7, 2016.

[23] Sanz-Serna JM, An unconventional symplectic integrator of W.Kahan *Applied Numerical Mathematics*, 16, 245–250, 1994.

[24] Yifa Tang, Jianwen Cao, Xiangtao Liu and Yuanchang Sun, Symplectic methods for the Ablowitz-Ladik discrete nonlinear Schrödinger equation *J. Phys. A: Math. Theor.*, n. 40, 2007.

[25] J. You Perturbations of Lower Dimensional Tori for Hamiltonian Systems *Journal of differential equations* 152, 1999.

[26] X. Yuan Construction of Quasi-Periodic Breathers via KAM Technique *Commun. Math. Phys.*, 226, 2002.